

**proceedings of the
Fifteenth International
MACHINE TOOL
DESIGN AND RESEARCH
Conference**

edited by

S. A. TOBIAS and F. KOENIGSBERGER

PROCEEDINGS OF THE
FIFTEENTH INTERNATIONAL
MACHINE TOOL DESIGN AND RESEARCH
CONFERENCE

PROCEEDINGS OF THE
FIFTEENTH INTERNATIONAL
MACHINE TOOL DESIGN AND RESEARCH
CONFERENCE

*held in Birmingham
18-20 September 1974*

Edited by

S. A. TOBIAS
*Chance Professor and Head of Department
Department of Mechanical Engineering
University of Birmingham*

and

F. KOENIGSBERGER
*Professor of Machine Tool Engineering
University of Manchester Institute of Science and Technology*

M

All rights reserved. No part of this
publication may be reproduced or transmitted,
in any form or by any means, without permission.

First published 1975 by
THE MACMILLAN PRESS LTD
London and Basingstoke
Associated companies in New York
Dublin Melbourne Johannesburg and Madras

SBN 333 15057 0

ISBN 978-1-349-01988-5 ISBN 978-1-349-01986-1 (eBook)
DOI 10.1007/978-1-349-01986-1

© The Macmillan Press Limited 1975

Softcover reprint of the hardcover 1st edition 1975

Published in the U.S.A.
by Halsted Press, a Division
of John Wiley & Sons, Inc.
New York

Library of Congress Catalog Card No: 63-19240
ISBN 978-0-470-87532-2

Set by William Clowes and Sons Ltd.,
London, Colchester and Beccles

CONTENTS

	<i>Page</i>
OPENING SESSION	
Opening address. J. W. ATWELL	1
GROUP TECHNOLOGY AND CELLULAR MANUFACTURE	
Group technology developments and modes of application. W. EVERSHEIM and M. MIESE	7
Total company appraisal for group technology. K. RATHMILL, P. BRUNN and R. LEONARD	19
Batch size selection for group technology. K. RATHMILL, P. BRUNN and R. LEONARD	29
The interaction of NC and GT. F. W. CRAVEN	37
Some organizational aspects of cellular manufacture based on computer simulation. D. ATHERSMITH and J. R. CROOKALL	45
CNC AND DNC	
BCLCNC—a computer numerical control system. N. AKGERMAN	57
Closed loop positioning control for a DNC-milling machine. A. M. LOSS and J. FRISCH	63
Adaptive control constraint and automatic cut distribution system for turning operations. E. GIESEKE	71
The improvement in accuracy of high precision machine tools by means of adaptive control. I. WASIUKIEWICZ	79
The machining-centre concept; an analysis of a tool-changing manufacturing system. J. R. CROOKALL and A. T. M. JAMIL	89
The NEL/MTIRA technological forecast for the computer control of machine tools. B. DAVIES, J. W. BRUCE and A. E. De BARR	97
CAD AND CAM	
Combination of CAD and CAM in machine tool manufacture. P. STÖCKMANN and G. WÖLLNER	107
The application of CAD techniques to machine tool component design, production and manufacture. A. JEBB, C. B. BESANT and R. C. EDNEY	115
Automatic design of machine tool components. T. G. FOSTER	121
A new generalized concept of computer aided production. J. FERENCZY and J. LEHOTZKY	129
Computer-aided design of axisymmetric hot forging dies. S. K. BISWAS and W. A. KNIGHT	135
Graph theory applied to computer aided plant layout. A. S. CARRIE	145
MACHINE TOOL ELEMENTS	
An analysis of a new type of high precision multiwedge clearanceless bearing for grinding wheel spindles. G. K. ARORA and T. S. R. MURTHY	155
Characteristics of electro-hydraulic feed drives under dynamic load conditions. D. AVRAMOVIC and P. FORT	165
Computer-aided design of standardized hydrostatic journal bearings. J. M. SANSINENEA and R. M. BUENO	175
METAL CUTTING	
Coolants and cutting tool temperatures. E. F. SMART and E. M. TRENT	187
Tool wear or tool design (a case study). R. A. ETHERIDGE and A. J. A. SCOTT	197
Ultra-high-speed machining: notes on metal cutting at speeds up to 7300 ft/s. G. ARNDT	203
On line measurement of bending and torque in milling. R. BEDINI and P. C. PINOTTI	209
Power hacksawing. P. J. THOMPSON and M. SARWAR	217
Optimization of drill geometry for orthopaedic surgery. G. H. FARNWORTH and J. A. BURTON	227
MACHINE TOOL DYNAMICS	
The dynamic characteristics of epoxy resin bonded machine tool structures. M. I. CHOWDHURY, M. M. SADEK and S. A. TOBIAS	237

	<i>Page</i>
On-line identification of machine tool receptances from cutting experiments. N. H. HANNA and A. W. KWIATKOWSKI	245
Influence of lathe tool wear on the vibrations sustained in cutting. P. MARTIN, B. MUTEL and J.-P. DRAPIER	251
Investigation of parametric instability in milling simulated by screw chasing operation. G. DROUBI and M. M. SADEK	259
Elastic wave transmission and reflection in long uniform bars due to the end impact of a rigid mass. W. JOHNSON	267
Self-excited vibration of circular saws cutting aluminium. D. S. DUGDALE	279
NOISE AND ITS REDUCTION	
Noise generation from an impact forming machine structure. A. E. M. OSMAN, W. A. KNIGHT and M. M. SADEK	289
Noise reduction in a HERF cropping machine. G. DROUBI, M. M. SADEK and S. A. TOBIAS	295
The effect of blade wear on noise levels of power hacksaws. G. J. McNULTY	305
GRINDING AND SURFACE TOPOGRAPHY	
Optimization of a grinding process and criteria for wheel life. G. TRMAL and H. KALISZER	311
The collection of data for the assessment of a grinding wheel dressing treatment. E. J. PATTINSON and J. LYON	317
Chattering phenomena as the criterion of redress life of grinding wheel: a study on the establishment of optimum operational condition in precision grinding of hardened steel. H. MAKINO	325
Finishing and running-in of plain bearings, with an economic appraisal. G. TRMAL, H. KALISZER and G. W. ROWE	333
In-process indication of surface roughness using a fibre-optics transducer. D. SPURGEON and R. A. C. SLATER	339
Performance analysis of the stylus technique of surface roughness assessment: a random field approach. J. B. AGULLO and J. PAGES-FITA	349
EDM AND ECM	
Conditioning a diamond-grit cup-wheel for electrochemical grinding. M. M. SFANTSIKOPOULOS and C. F. NOBLE	365
Electro-discharge machined surfaces. J. R. CROOKALL and B. C. KHOR	373
Some effects of debris concentration on erosion and electrode wear in electro-discharge machining. J. R. CROOKALL and P. W. LEE	385
Investigation of the modulation of interface waves in explosive welding. S. R. REID and D. J. LANGDALE	391
Implosive welding of composite metal-foil cylinders. H. EL-SOBKY and T. Z. BLAZYNSKI	399
The cold drawing of implosively welded bi-metallic tubes. S. TOWNLEY and T. Z. BLAZYNSKI	407
HOT FORMING	
Transverse deformation in section rolling and forging. G. W. ROWE, I. M. DESAI and H. S. SHIN	417
A study on plastic working of alloys in their mushy state. S. FUKUOKA and M. KIUCHI	423
Stress strain curves from the ring test. A. N. BRAMLEY and N. A. ABDUL	431
An elemental upper-bound technique for general use in forging analysis. R. P. McDERMOTT and A. N. BRAMLEY	437
Application of a modular approach to estimate load and energy in closed die forging. S. K. BISWAS and B. W. ROOKS	445
The elevated temperature extrusion of fluted sections in high-speed steels. C. E. N. STURGESS and T. A. DEAN	455
DIE LIFE IN HOT FORMING	
Cracking and fracture of hot-work die steels. A. THOMAS	467
Die life estimation in forging. S. STØREN, J. EBBESEN, J. SLUTÅS and I. SÆTRE	473

	<i>Page</i>
Wear in drop forging dies. T. M. SILVA and T. A. DEAN	479
The effect of die temperature on metal flow and die wear during high-speed hot forging. B. W. ROOKS	487
ROLLING	
Ball rolling: a literature survey and some experimental results. N. R. CHITKARA, W. JOHNSON and J. R. S. UTTLEY	497
Cold rolling of ring gears. J. B. HAWKYARD, F. R. NAVARATNE and W. JOHNSON	507
An experimental investigation into the S-mill rolling process. D. K. ROBERTSON and D. H. SANSOME	515
DRAWING, TUBE AND SHEET FORMING	
A preliminary investigation of the die-less drawing of titanium and some steels. J. M. ALEXANDER and T. W. TURNER	525
A fundamental study on dieless drawing. H. SEKIGUCHI, K. KOBATAKE and K. OSAKADA	539
Theoretical study of the drawing of bi-metal rod and tube. H. T. CHIA and D. H. SANSOME	545
An oscillatory deep-drawing analogue. M. J. R. YOUNG and D. H. SANSOME	551
Rigid-plastic analysis of bore expanding and flange drawing with anisotropic sheet metals by the matrix method. S. H. LEE and S. KOBAYASHI	561
Tube inversion: a process for obtaining gross expansions in the diameter of thin-wall tubes. S. T. S. AL-HASSANI	571
COLD FORMING, BLANKING AND CROPPING	
The strength of the material in cold formed parts. J. DATSKO	579
The development of a cold forming process on an automatic high-speed forming machine. T. SCAIFE, B. PARSONS and B. N. COLE	585
Forward cold extrusion of aluminium—estimation of the maximum extrusion force. L. E. FARMER and B. Z. PIROG	597
Rigid-plastic analysis of cold heading by the matrix method. S. N. SHAH and S. KOBAYASHI	603
Production of crankshafts by clamp shearing. T. NAKAGAWA and K. SHIOMORI	611
Analysis of stresses in bar cropping. V. G. WONG and M. K. DAS	617
COMPACTION AND SINTER FORGING	
The dynamic compaction of aluminium and iron powder. D. RAYBOULD	627
The hot extrusion of metal powder preforms. M. NEGM and R. DAVIES	637
Plasticity of sintered iron-powder compacts. S. D. EL WAKIL	645
The effect of presintering conditions on the cold forgeability of powder preforms. M. A. RIFAI	651
The production of extruded material from metal powders. T. SHEPPARD	659
Comparison between uniaxial and isostatic densified stainless-steel powder preforms forged in a closed die. H. ALTMANN	669
Powder compaction at very high pressures. P. N. TOMLINSON, R. L. HEWITT and R. D. VENTER	677
GENERAL ASPECTS OF FORMING	
Strong composites based on all-metal systems. D. V. WILSON, W. T. ROBERTS, J. J. MOORE, R. F. STOKES and J. FARMER	685
Characteristic features in the hole flanging and piercing of thin and thick circular plates using conical and ogival punches. W. JOHNSON, N. R. CHITKARA and P. A. BEX	695
The influence of hydrostatic pressure on the plastic deformation of metallic materials. W. LORREK and O. PAWELSKI	703
The toughness of M2 high-speed steel. G. BERRY and M. J. KADHIM AL-TORNACHI	713
Analysis of large, natural strain increments using the Mohr's circle. J. WOLAK	721
Smooth hole wall piercing in the manufacture of precision mechanisms. R. JOHNSTON and K. SWIFT	729
Author index	735
Subject index	737

OPENING ADDRESS

by

J. W. ATWELL*

I am greatly honoured to be invited to speak to you at the opening of this Matador Conference, not only because over many years these conferences have acquired an international reputation, which the list of delegates amply demonstrates, but also because I am encouraged to think that industry—certainly your industry—recognizes that the Research Requirements Boards have now become established as an important element in the system of government support of research and development.

In its first annual report, published recently, my own Board expressed concern that the work of the Board and perhaps even the purpose of the Board, was not well enough understood, especially in industry. My presence here today suggests that we may well have been doing less than justice to the machine tool industry and to those who are involved in its design and research programmes.

When I came to consider what I might say to you today and turned to the conference programme for inspiration, I was immediately impressed by the number and range of subjects which the organizers had been able to include, and I would guess that many other papers had to be declined simply because of time limitation. As I thought about all this work going on in universities and research centres in many countries, I was reminded yet again of the problem of communicating these ideas and new developments to engineers in manufacturing industry who should be making use of the information and I wondered what plans were being made to improve the communication of this knowledge—what has come to be known as 'technology transfer'.

I would like, therefore, to put some thoughts to you on the importance of technology transfer, but before doing so perhaps I should talk briefly about impressions gained in the Mechanical Engineering and Machine Tool Requirements Board, now that it has completed a survey of all the R. and D. projects in its field that had government support in the past; a great deal of which was concerned with manufacturing.

Somewhat to our surprise we found a disappointingly small number of projects directed towards

the manufacturing process and to the well-being of the production departments. You will, of course, appreciate that I am not referring to the components of the manufacturing system—the machine tool, fork lift truck, tool and cutter, jig and fixture—but to the combination of these components into an efficient and effective production unit.

For several years now, at international machine tool exhibitions, we have been shown a range of sophisticated approaches toward the automatic factory, and I use this rather old-fashioned term deliberately, since many engineers who are concerned with the rising costs of production and the increasing difficulty of finding skilled craftsmen—and I count myself one of these—have been confused by some of the terminology used by the specialists: group technology, cellular systems, CNC, DNC and many others. But these all have one thing in common. They are attempting to take groups of machine tools, together with other pieces of hardware in a workshop, and to devise a system of using them which will be more effective and more flexible than existing systems.

My Board feels that in a great many cases it is difficult to justify going to the complicated lengths that have been seen in some of these systems. Instead, we believe that better utilization of existing plant, aided by new ideas and new techniques, would for many manufacturing companies in this country bring substantial benefit and I suspect this would also be the case in other countries. Unfortunately, not every medium and small manufacturing company is able to take advantage of these new techniques, either because of management or equipment limitations, which in themselves are matters for concern.

The Board also found a wide variation in the quality of submissions made to it; some I am sorry to say with not very well defined objectives and no clear definition of the benefits to be expected. This was not altogether surprising, as my experience has been that, too often, research programmes are started without the objectives being defined sufficiently clearly and then, when completed, often with quite useful results, there is great disappointment that

* Chairman of Engineering Division, The Weir Group Ltd. and Chairman of Requirement Board for Mechanical Engineering and Machine Tools, Department of Industry

industry is not queueing up for the benefits. None of us has unlimited resources of men, equipment and money, and I have long held the view that if we are to utilize our resources efficiently we must match our technological developments to the needs of the market for our R. and D. output. This, of course, is equally as true for product development programmes in industry as it is for research projects in government research establishments.

It also became evident, from the Board's study of the situation, that unless corrective action is taken, the gap between the research worker—the scientist and technologist—and the point of application in industry will become wider. We shall simply keep adding to our bank of knowledge and neglect the work needing to be done to improve productivity in industry. Indeed, it has been said that we could almost call a halt to all R. and D. in many branches of engineering for several years without retarding the flow of new ideas to industry.

Let me just elaborate on this problem of applying the results of research and development programmes in industry. In many cases, the purpose of a research project is to lead to improvement in design of an engineering product, whether it be a machine tool or a pump for the oil industry. The problem, then, to be solved is the translation of the research output into design data; not without its difficulties, not always done effectively, but in most cases the designer is glad to have the new information and is ready to apply it.

On the other hand there are cases where research projects which are intended to offer improved manufacturing systems receive only limited response. Consequently, the research team expresses great dismay that its work is not being immediately utilized in industry and there is renewed criticism of industry's negative attitude to accepting new ideas. When this happens it is simply another example of the problem of technology transfer; of the difficulty of being able to perceive benefits from the application of some fresh thinking. It implies that in some sectors of industry there is failure to recognize when a problem exists and to see the possibility of a solution. And it underlines industry's need for engineers with the capability of bridging the gap between the research laboratory and the factory floor; with the capability of interpreting the results of research at the point of application.

The answer to these problems is not simple and requires constant attention from all of us—not only the Requirements Boards. No doubt a whole range of solutions will occur to you. For my part, I would like to suggest two courses of action: one which is comparatively short term, the other will take longer to become effective.

In the short term we should be encouraging projects which are directed towards manufacturing technology rather than machine tool technology, so that at least for a time the engineers who have produced the advanced machines available today can begin to direct their skills to the application of the advanced techniques in the factories. It could well be that the ability and performance of the modern machine tool is outstripping the competence of the user to employ these facilities to the best advantage. The research

workers who have achieved these improvements are surely the very men who should become involved in the application of their ideas.

The longer term solution will be more difficult to achieve and will need the understanding and support of many organizations. The objective should be to ensure that the user industries themselves employ manufacturing engineers who have the ability to understand what should and should not be done with the techniques that are now available, and who will be excited rather than frightened by the prospect of a small computer attached to each machine tool. This is a problem of education and training, and one which will not be solved without the collaboration and help of the universities and the professional institutions.

Recognition of this problem by the Science Research Council has led to the formation of a Manufacturing Technology Committee. My own Board—The Mechanical Engineering and Machine Tool Requirements Board—has also had submissions which have highlighted the difficulty. We saw how vital it would be to try to win over both industry and the universities to an understanding of the problem as we saw it, and after discussion with the Science Research Council, a joint committee has been formed under the chairmanship of Sir Richard Young. We all hope their deliberations will help us find an answer, but we are well aware that the most careful plans for training and education will fail unless the right kind of man can be attracted to this career.

It cannot be denied that we have some way to go before we can change the standing of production engineering to compete with, for example, law or medicine in the eyes of the school-leaver, but I would like to think that the dedication which brings all of you to this conference, year by year, can be directed to improving your prestige and standing in the eyes of the world at large. If our long-term strategy is to increase the number and improve the quality of production engineers, we must provide them with a new image and an incentive to achieve it. In this connection the proposals announced recently by the Institutions of Civil, Mechanical and Electrical Engineers for a unified engineering profession should, if accepted, make a substantial contribution towards establishing a new status for engineers in this country.

Long-term strategy is a vital part of the direction of R. and D. and all Research Requirements Boards have been considering their approach to it during the past month or so. There is, as you would expect, no approach common to the whole of British Industry, but all Boards recognize that it is essential they discover a formula for evaluating research ideas which will apply equally fairly to existing projects as it does to the injection of new ideas and new approaches. We realize that we must create an environment in which it is possible to make the best use of undoubted research skills and facilities, without necessarily perpetuating programmes which have passed their peak of inventiveness. I hope it will not be regarded as an unfair observation if I say that the Requirements Boards believe there is room for improvement in the attitude of some research workers to the exploitation of the results of their efforts. I think it will be conceded that in this competitive age the satisfaction of R. and

D. can no longer be in its completion, but only in its adoption. Even in nuclear physics it is not the completion of a programme of research which means success, but its adoption as a contribution to the solution of the unknown. Research programmes are simply not complete without a full consideration of the proposed means of transferring the results from the research worker to the end user. Some government research establishments and university departments could be open to this criticism and even these conferences, valuable as they are, internationally, as a forum for exchange between research workers, probably make only a limited contribution to communication between the research worker and the potential beneficiary.

Much has still to be learned about technology transfer, and although we can usually see when it has been effective, it remains a most unpredictable art. It

needs our most serious consideration, for I am sure that those of you who have produced papers of the standard we are seeing this week, must surely have a substantial contribution to make to this vital end-stage to their own researches.

Mr. Chairman, it was good of you to give me the opportunity of talking to such a distinguished gathering. There must be few opportunities of meeting such a representative cross-section of those engaged in machine tool design and research, the importance of which is recognized by the distinction given to machine tools in the title of my own Requirements Board—The Mechanical Engineering and *Machine Tool* Requirements Board. May I express the hope that you will have a successful and enjoyable conference and that you will spare some thought for this vital problem of communicating your results to those who can apply them effectively in industry.

GROUP TECHNOLOGY AND CELLULAR MANUFACTURE

GROUP TECHNOLOGY DEVELOPMENTS AND MODES OF APPLICATION

by

W. EVERSHEIM* and M. MIESE*

SUMMARY

Group technology is defined as an organization principle that can be applied to all departments of an enterprise concerned with production. This principle is based on the systematic grouping of identical or similar components by using classification systems. All components combined into a group are subject to common planning and machining, thus achieving a high rationalization effect particularly in the case of single and small batch production. This application of group technology is generally known as part family manufacturing.

Further rationalization can be achieved by applying the same principle to other production areas, mainly through standardizing all production data and through simplifying and automating the necessary planning activities using electronic data-processing systems.

1 INTRODUCTION

Developments in labour and sales markets are forcing enterprises increasingly to look for additional methods of rationalization within the whole manufacturing process. The objective is economic production in the face of a continuous increase in general costs.

This is particularly true of companies with small or medium batch production, for which such a high grade of mechanization and automation, such as can be encountered almost everywhere in mass production, is more or less illusory because of rather small job lot sizes.

Moreover, previous rationalization activities have been substantially concerned with the manufacturing process only. However, as manufacturing is only one area within the complex production process, its satisfactory performance alone does not automatically lead to an overall success of the enterprise. It is only by applying all available technological and organizational measures to every department concerned with production that rationalization successes similar to those obtained in mass production can also be achieved in small and medium batch production.

The production process is generally understood as being the sum of all the activities which are directly related to the manufacture of a product, from design through work planning and manufacturing to assembly. Where rationalization efforts are directed to individual departments alone, the effect will *a priori* remain limited. This is because both the exploitation of technological feasibilities in manufacture and assembly and the economic success of

production as a whole largely depend on the preparatory work carried out in the preceding stages of production, for example in work planning and designing. The enterprise can only achieve high overall profitability through the purposeful and mutual co-ordination of the activities of all departments.

The application of group technology can be a key to this overall rationalization success involving all departments of production.

2 DEVELOPMENT OF GROUP TECHNOLOGY

The concept of group technology was developed by Mitrofanov¹ on whose works all later publications in socialist countries are based. All efforts to implement group technology in industry were initially concentrated on single and small batch production.

With the aid of a classification system Mitrofanov groups all those machine components which are to undergo the same or similar operations on identical machine tools whose adjustments and equipment are also identical. For each group of components a standard work plan is set up based on the characteristics of a so-called 'complex component' which contains all feasible geometrical elements for all similar parts.

In the Federal Republic of Germany first attempts to develop group technology were made in the form of, a component plan by Lange and Rossberg². Comprehensive studies and research work were also carried out at the Laboratory for Machine Tools and Production Engineering at the Aachen Technical University. First results were published in 1958

* Institute for Machine Tools and Production Engineering, Technical University, Aachen

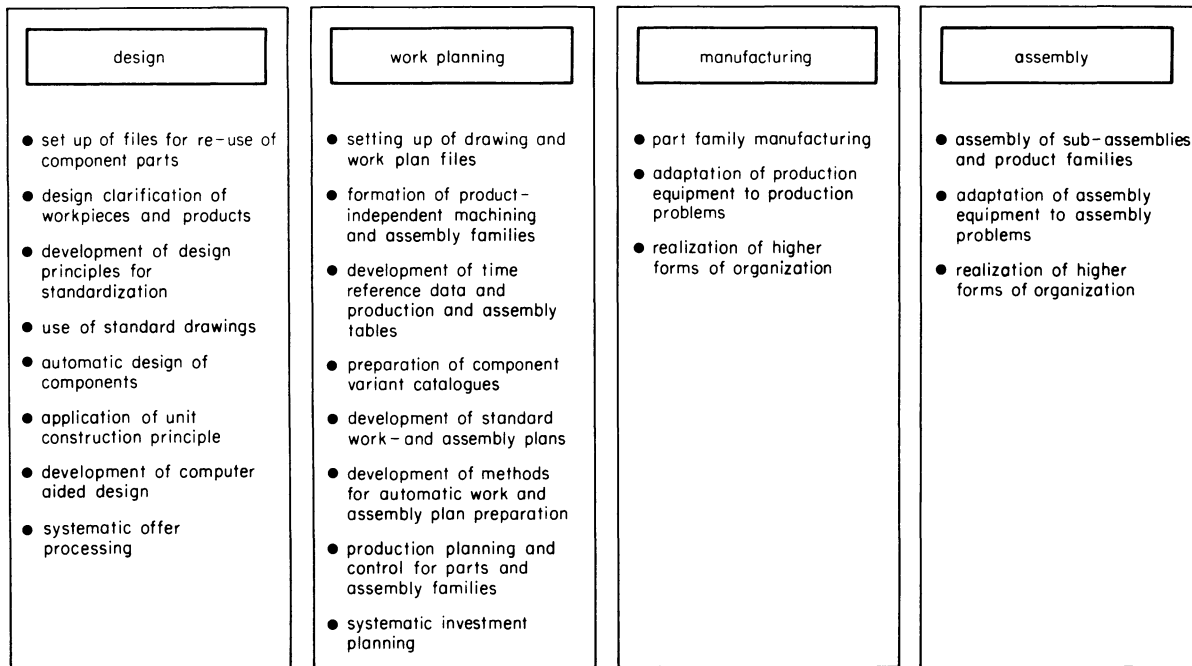


Figure 1. Applications of group technology.

entitled 'Adaption of Machine Tools to the Production Problem'.

In this connection the term 'part family manufacturing' was preferred to 'group technology', and was subsequently widely adopted in Germany. This term was chosen to express the following two principles: technological similarity is more important than the exact geometrical similarity of the components; various graduations of this relationship are possible, which nevertheless allow for common machining operations.

Thus a number of classification systems were developed as an aid to grouping part families, for example the systems by Mitrofanov, Lange-Rosberg, Brisch-Copic, ZAFO and the work piece describing a classification system by Opitz⁵.

In the course of the continuing development of group technology, the meaning of this term has changed even more. Originally group technology was synonymous only with part family manufacturing. Today it is defined as an organizational principle that can be implemented in all production-linked departments of an enterprise in each of which it can have manifold forms of application (figure 1).

The general objective in applying group technology is the systematic analysis of all properties of the product from the individual workpieces to the completed products, so as to achieve optimum working and machining methods for each specific element in question. These methods should involve a minimum of costs in order processing and at the same time permit the application of modern auxiliary devices, for example electronic data processing.

The results of part family manufacturing made it necessary to adapt the working methods in the design and work planning departments to the changed manufacturing conditions. In the course of the further development of group technology the research activities concentrated more and more on these departments. At the same time the evolution of

working methods based on group technology was not restricted to components, but was also extended to more complex units such as sub-assemblies and even complete products.

3 EXAMPLES OF THE APPLICATION OF GROUP TECHNOLOGY IN VARIOUS PRODUCTION DEPARTMENTS

All technological or organizational evolutions related to group technology are somehow based on the fact that within a production department there exist activities or data with a certain degree of similarity. They can thus be combined into groups for which general solutions or methods can be developed. Such a group is usually termed a 'family'. Figure 2 shows the kind of families which generally can be formed depending on the different production areas and the complexity of the objects. The most important auxiliary device for setting up such families are appropriate classification systems, which have been developed for a range of different groups of parts. Classification systems necessarily consist of the criteria needed for the formation of such families. The application of classification systems permits a survey over all objects existing in an enterprise. To realize the last-mentioned objective it is necessary that individual classification systems for each special kind of object must be integrated into a comprehensive coding system. For this purpose a brief definition, for example plants, machines, sub-assemblies, components and so on, is associated with the first digit of the code number. The second digit then permits a decimal break-down of these definitions by which all related classification systems can be incorporated into the overall coding system.

The application of such a systematic and homogeneous coding system is a most important prerequisite for an effective improvement of the communication system in order processing, and is also

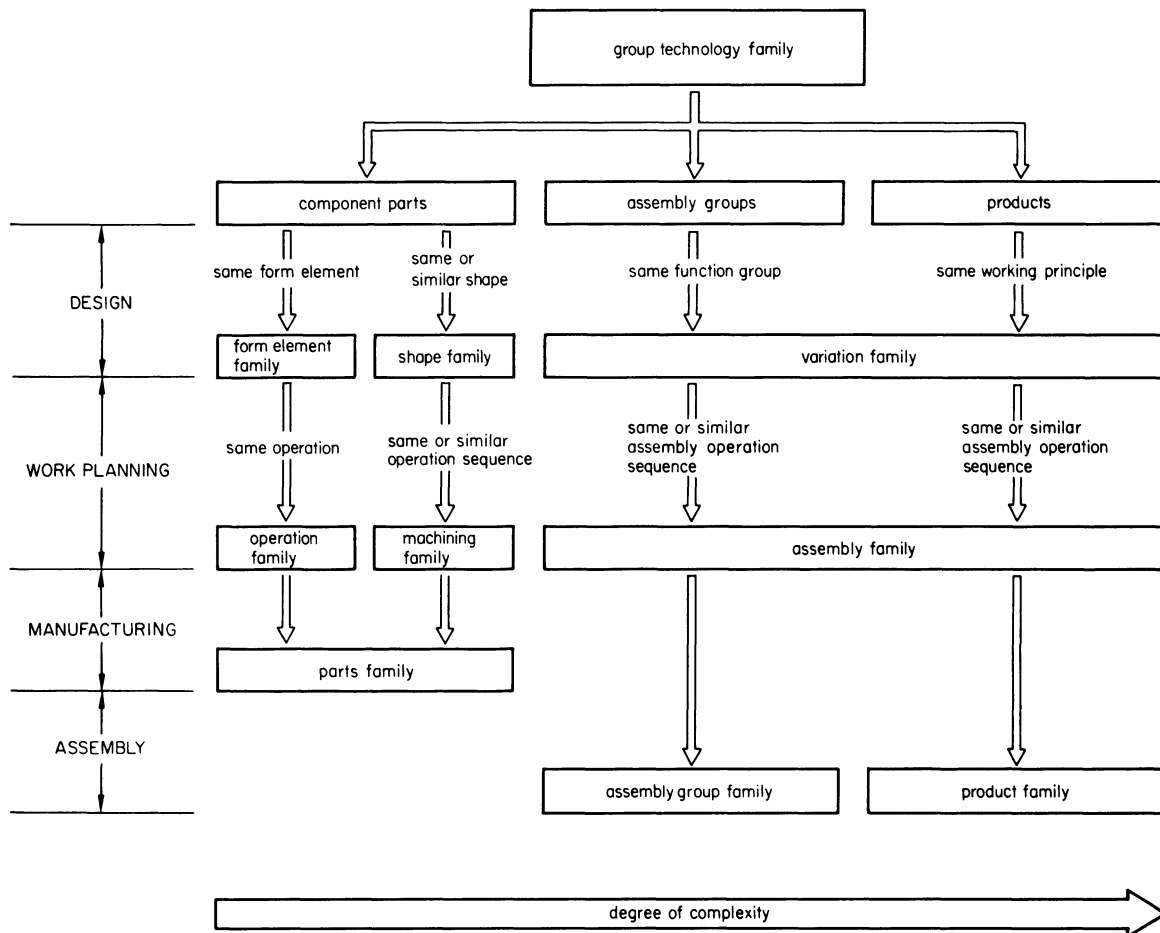


Figure 2. Group technology families.

necessary for the optimal layout of information documents such as drawings, parts lists, work plans and so on.

Classification should be performed by a special coding office inserted into the data-flow between design, work planning and manufacturing, thus ensuring that the classification of all objects is based on the same standard principles³⁻¹⁶.

3.1 Systematic tender planning

In enterprises with mainly individual order processing the short-term preparation of binding offers frequently involves a number of problems. In many cases special customer specifications require new designs and calculations causing relatively high expenditure. Companies are forced to keep the planning effort as small as possible, because time is precious and because the tender-to-order conversion rate in many cases is far below 10 per cent. The part played by new planning can be substantially reduced by a systematic and purposeful exploitation of all data already accumulated in the enterprise which can be used in the working out of new offers.

This calls for direct recourse to all existing problem solutions which can be easily achieved only by the application of appropriate classification systems¹⁷.

Figure 3 illustrates the approach to offer planning on the basis of such classification systems. As soon as all customer specifications are known and the procedural layout of the product is established, the

classification will assist in finding out whether similar products have been manufactured before which will more or less match the specified requirements.

To facilitate this procedure it is useful to set up film card files which contain all existing project documents in the form of drawings, parts lists and calculations sorted by classification numbers. If a drawing is available for the project sought after, all other necessary documents will be accessible through the reference identification numbers contained in the parts list. From these data, the required offer can readily be compiled without much time being lost in searching. In practice this hardly ever happens. For units of less complexity, however, such as sub-assemblies and so on, the changes of re-using existing data are much higher³.

A prerequisite for this approach is the systematic break-down of a project or design into sub-units, a procedure which should apply not only in the case of the project in hand, but also to all previous projects on the same common and systematic procedural basis. This system in combination with an appropriate classification, will facilitate recourse to previously designed sub-assemblies¹⁸.

3.2 Design department

The design department holds a key position with respect to the applicability of group technology. An investigation of the relative distribution of the throughput time in small batch production has shown

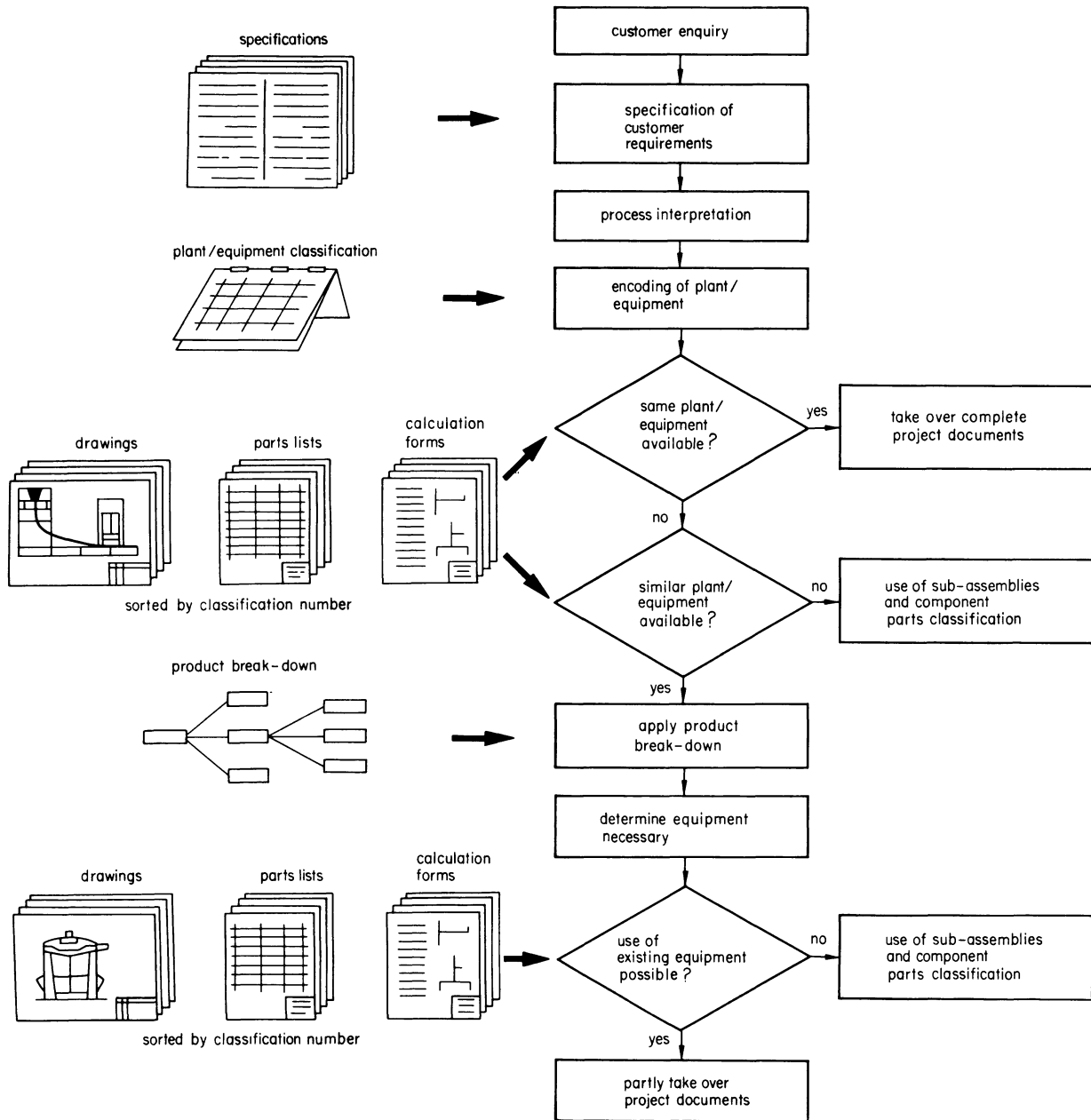


Figure 3. Use of classification systems for plants, equipment and sub-assemblies in planning.

that 55 per cent of this time is taken up by design, work planning and acquisition. In these departments particularly foundations of further order processing are laid. The information and documents leaving the design department will be the basis of all activities in the subsequent departments.

The working principle of a product, its material and its accuracy in combination with the selected manufacturing method largely predetermine the overall cost of the product. The application of group technology in the design department will set the pace for corresponding measures in work planning, manufacture and in the assembly department. Normally the designer is only bound to a few interface dimensions and is thus free to choose the optimal solution. In order to relieve the designer of routine work and to reduce the plurality of different designs for parts or products with identical functions, the designer should have an organizational aid through

which he can have access to solutions previously implemented in other products. The existing production programme can easily be analysed with the above mentioned classification systems^{19,20,21}.

The classification is the basis for any standardization in design and especially for the compilation of design principles. They also allow a survey of the frequency of components and sub-assemblies, and thus facilitate the decision as to which of these parts should be transferred into standard drawings or which parts should be incorporated into the works standard specifications. Any drawing organization based on such principles should serve the following purposes²² illustrated in figure 4:

- (1) direct re-use of parts or use of similar parts as basis for small drawing changes;
- (2) elaboration of process sheets for standardization of shape characteristics, form elements,

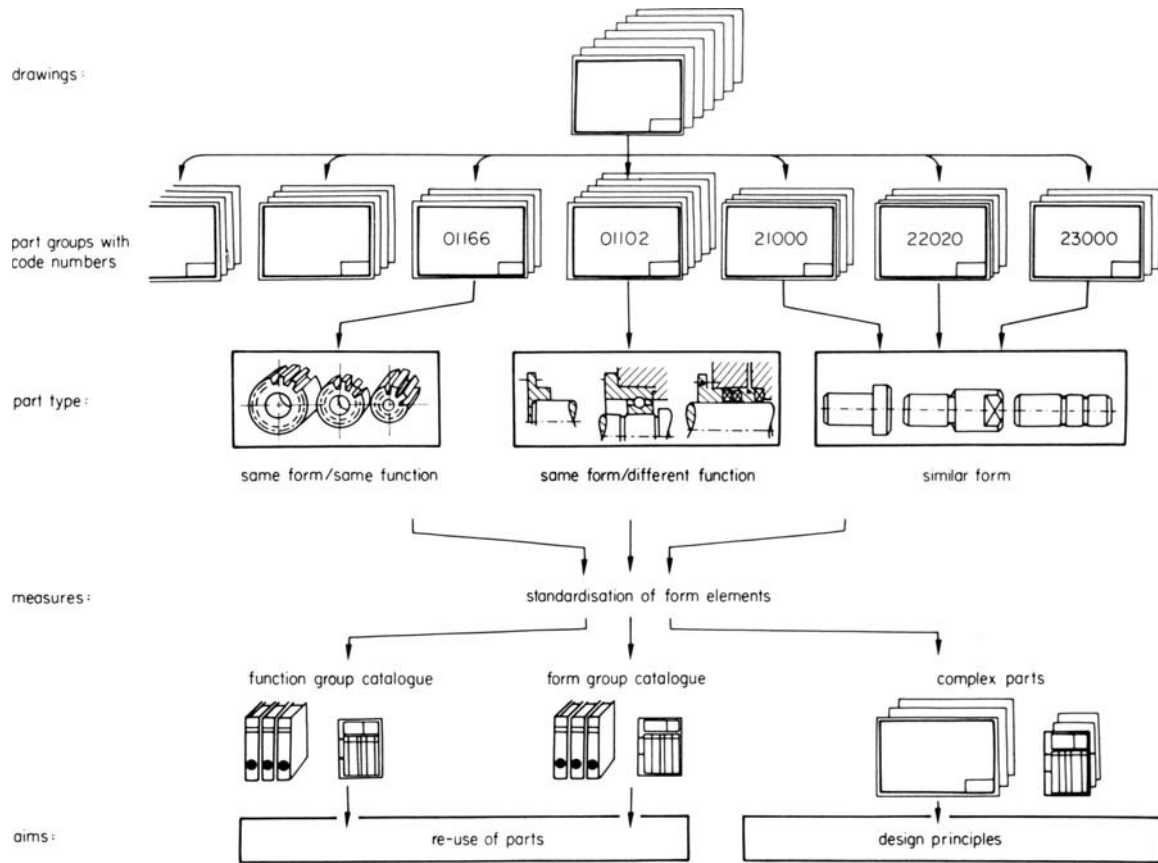


Figure 4. Development of design aids.

and design details with due consideration of the DIN standards;

- (3) extension of the works standard specification and of the development of design principles.

Any parts spectrum usually contains a substantial number of identical or at least similar components, so that these parts can be grouped according to geometrical and/or technical similarity. The resulting groups allow a comparison and survey of the plurality of variants which leads to the classification of the parts spectrum on the basis of standardization.

In the case of workpieces it is possible, for instance, to set up so-called complex parts² which can be defined as theoretical workpieces comprising all the variants of similar parts from a group of components with the same classification number. Drawings for such variants and complex parts can be greatly simplified by the development of master drawings which are not to scale. The relatively few variable dimensions for a new part can then be inserted. Such master drawings, however, are only suitable for parts which do not have too many variable dimensions but which on the other hand represent a large section of the overall parts spectrum. Hence, the expenditure for component drawings can be substantially reduced. The procedure for the set up of such master drawings²¹ is shown in figure 5.

The preparation of drawings can be further simplified by the use of electronic data processing. Master drawings are prepared on translucent foils and figure tables are printed by high-speed printers. The tables contain all the numerical data necessary and

correspond to the free places in the drawing. They are then overlaid and both are copied in one operation. The result is a production drawing containing all necessary data.

The automatic computation of workpiece drawings which are true to scale can be carried out by an electronic data processing programme system that

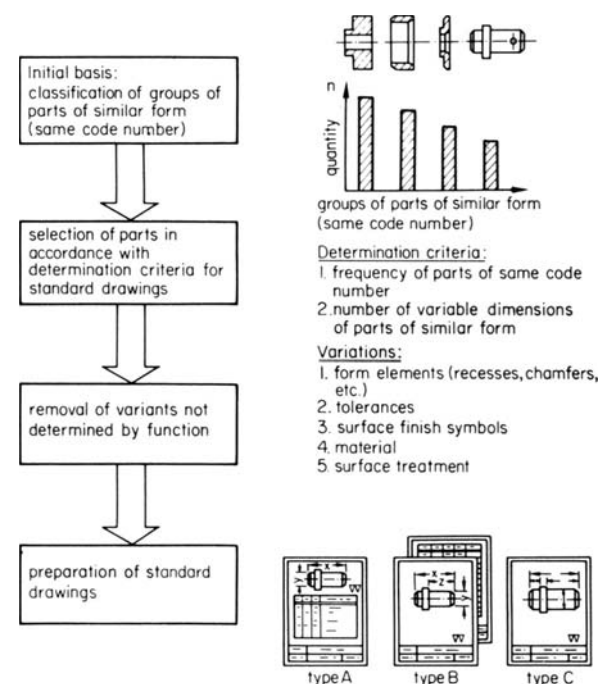


Figure 5. Set-up of standard drawings.

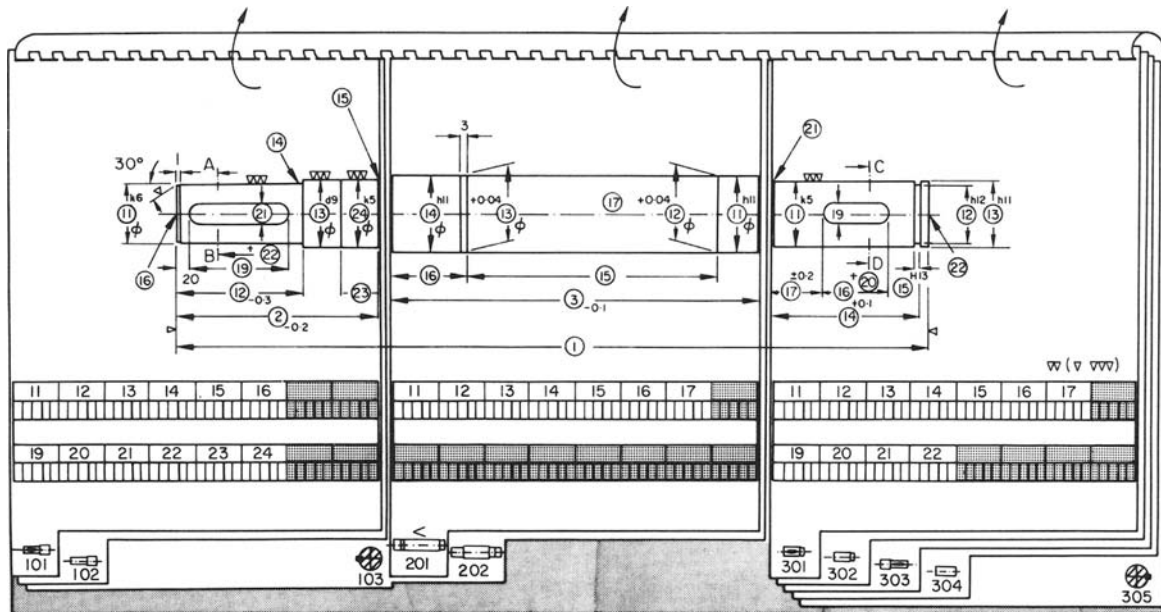


Figure 6. Data input for workpiece group SHAFT.

has been developed at the Institute for Machine Tools and Production Engineering at the Technical University of Aachen, West Germany. This programme system is based on the variant principle and has been successfully introduced in industry^{2,3}. A characteristic feature of the variant principle is the similarity of planning, that is planning on the basis of standardized documents for identical or similar components. In this system arbitrary elements of drawings—for example lines, curves and so on—are combined to make up a complete drawing. The instructions for the adaption of the drawing elements to the corresponding workpiece data and for their combination vary according to the nature of the parts. The rules of these variations for each group of parts are stored in a specific data file.

A universally applicable programme calculates all the necessary data required by the plotter for automatic drawing from the workpiece data in hand and from the specific drawing rules. The elements of the drawings were standardized using simple symbolic language.

The programming effort was greatly reduced and simplified by sub-dividing the programme system into a generally applicable routine and into special routines for each problem. Thus only the special data required for each individual group of parts need special programming. As a result, the programme system can easily be modified and extended as required.

These drawing rules based on group technology also resulted in a considerable reduction in the necessary input data. The designer makes use of a catalogue of variants and thus can readily survey all stored part groups. He is able to enter the appropriate input data even without much knowledge of programming.

Figure 6 shows a catalogue of variants for electric motor shafts. Each shaft is composed of three functional sections that add up to a complete shaft. For data input the required shaft sections are

selected. The variable workpiece data, identified by circled numbers, are entered into the fields for punching instructions shown in the bottom part of figure 6. Figure 7 shows a drawing of such a shaft produced by a plotter and controlled by the above mentioned programme system.

The advantage of computer-aided design will be even greater when complete products of higher complexity can be automatically processed. To do so it is necessary that for a given working principle of design there exist a lot of variants which depend on customer specifications. In group technology terms these variants also represent a group. The following example of a heat exchanger may illustrate the automatic variant design that has already been practised by an enterprise for several years.

The data required for the preparation of tender or for order processing are entered into an input-data form adapted to the subsequent punching operation. The computer then checks these data for logical and formal errors, and at the same time uses an optimizing strategy for the layout of the optimal solution. The output contains all desired data. Figure 8 shows the sectional view of a heat exchanger produced by a drum-type plotter¹⁹.

Compilation of parts lists can also be carried out using the same principles as in automatic drawing preparation. The parts lists of plants, machines or sub-assemblies contain information on the types and quantities of the pertaining component parts; in most cases it also indicates the mutual relationships of these parts and thus the structure of the product. Volume and limitation of the parts list depends mainly on the degree of product break-down. In this connection standard sets of parts lists should be useful for all production departments²⁴.

If the products of a group of variants largely consist of identical components, the obvious procedure is to subdivide the components and to compile them in so-called complex parts lists. A unit cannot be defined by such a list alone because it contains all

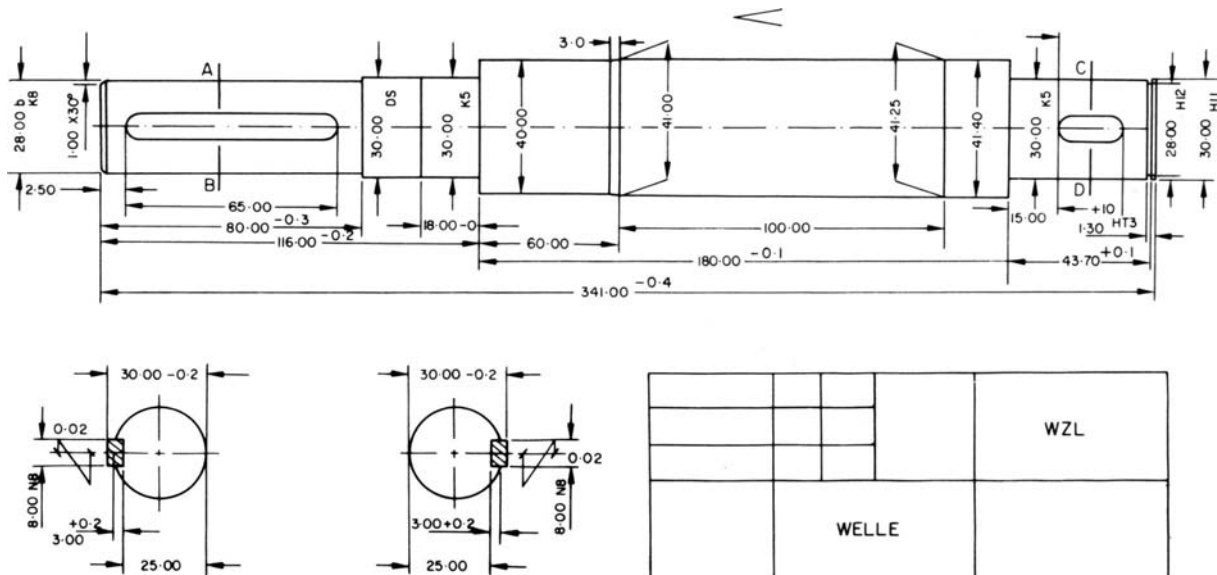


Figure 7. Automatically produced drawing variant of basic type SHAFT.

variable components and groups. To represent a complete variant group it is necessary to know those parts that are to be dropped from or retained in the complex list. This principle is well suited to electronic data processing and can easily be adapted to special requirements. The most important advantages of such structured parts lists are greater lucidity and a substantially reduced parts list volume which is of particular value in obtaining parts.

3.3 Work planning department

This department is growing in importance for order processing. Its tasks become more and more differentiated because of progress in technology and more and more voluminous because more and more tasks are allocated to it from the other production departments. For these reasons, rationalization measures in the work planning department turn out to be particularly efficient.

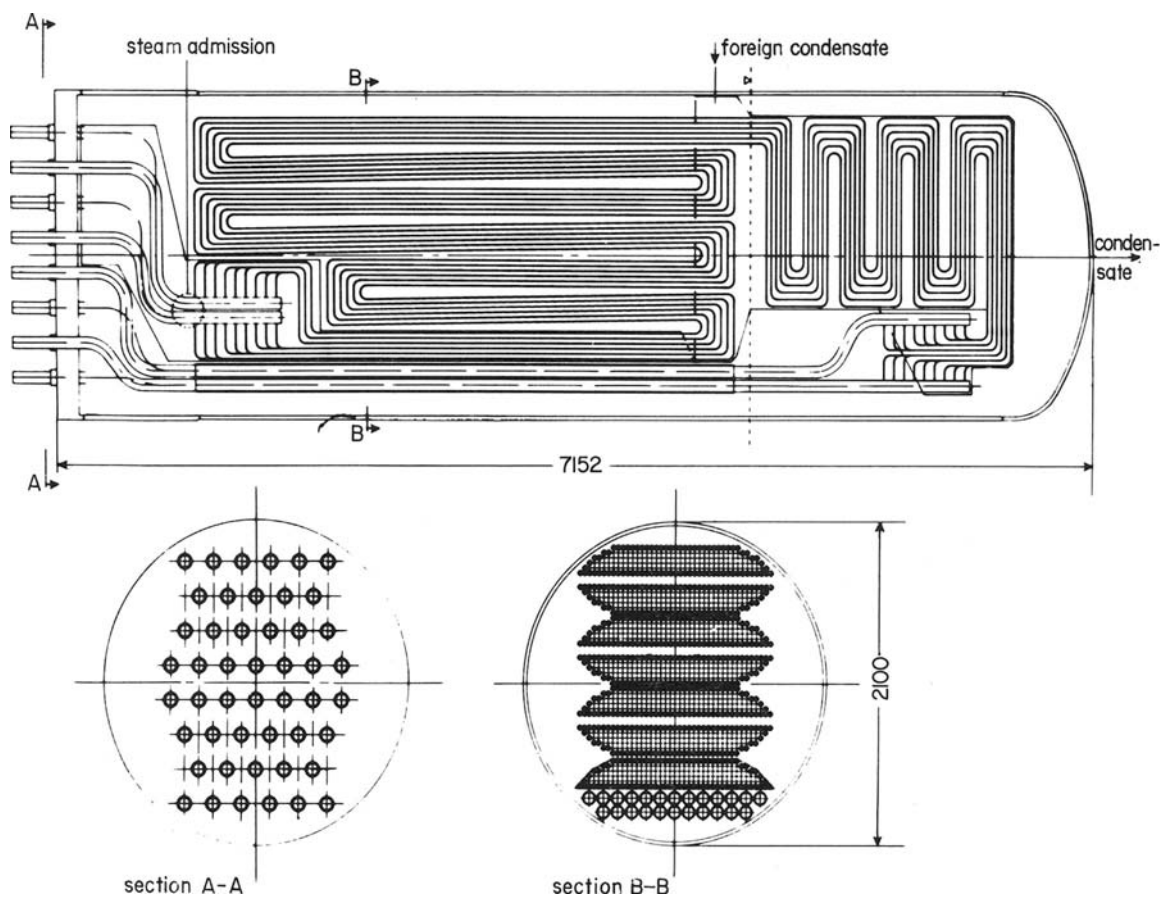


Figure 8. Computer aided design of a heat exchanger using the variant principle.

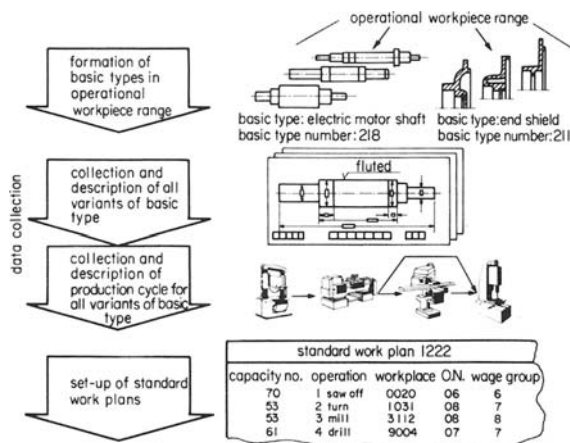


Figure 9. System of setting up standard work plans.

Group technology can also be applied to the work planning department, whose main tasks consist of production scheduling and production control. One of the most important problems here is the compilation of work plans for each component, that is, the conversion of design data into working instructions. This relatively time-consuming planning work can be rationalized in the long run only by electronic data-processing assisted automation. A feasible approach is the automatic set-up of work plans based on the principle of variants explained above. The Institute for Machine Tools and Production Engineering at the Aachen Technical University has developed a suitable procedure and a corresponding programme

system^{23,25}. The standard work plans for similar part groups are stored in data files that are processed by a generally valid routine programme.

Figure 9 illustrates the procedure for the setting-up of standard work plans. Starting from the basic types of similar part groups laid down in the catalogues of variants described on page 5, all necessary workpiece data and all associated work plan data for all variants of a basic type are compiled. By doing so any internal production know-how as well as all restrictions related to company production are documented. An analysis of these data quickly shows up planning errors and differences in production procedures; a further analysis reveals non-function deviations. These variants can be either standardized or eliminated from the parts spectrum.

In the next step, the quantitative influence of the variable parts features upon the standard time is determined. Statistical methods will permit us to establish the relation between part features and standard time objectively. The results thus obtained are then concentrated into standard work plans.

To set up a production schedule for a given workpiece (figure 10); the variable workpiece data are transferred from a corresponding catalogue of variants into the associated punching form. After the data input, the computer internally determines the file-stored standard work plan which comprises all computing functions, decision tables, editing and control instructions required by the processing programme. The variable workpiece data are checked for geometrical and technological limits and the dimen-

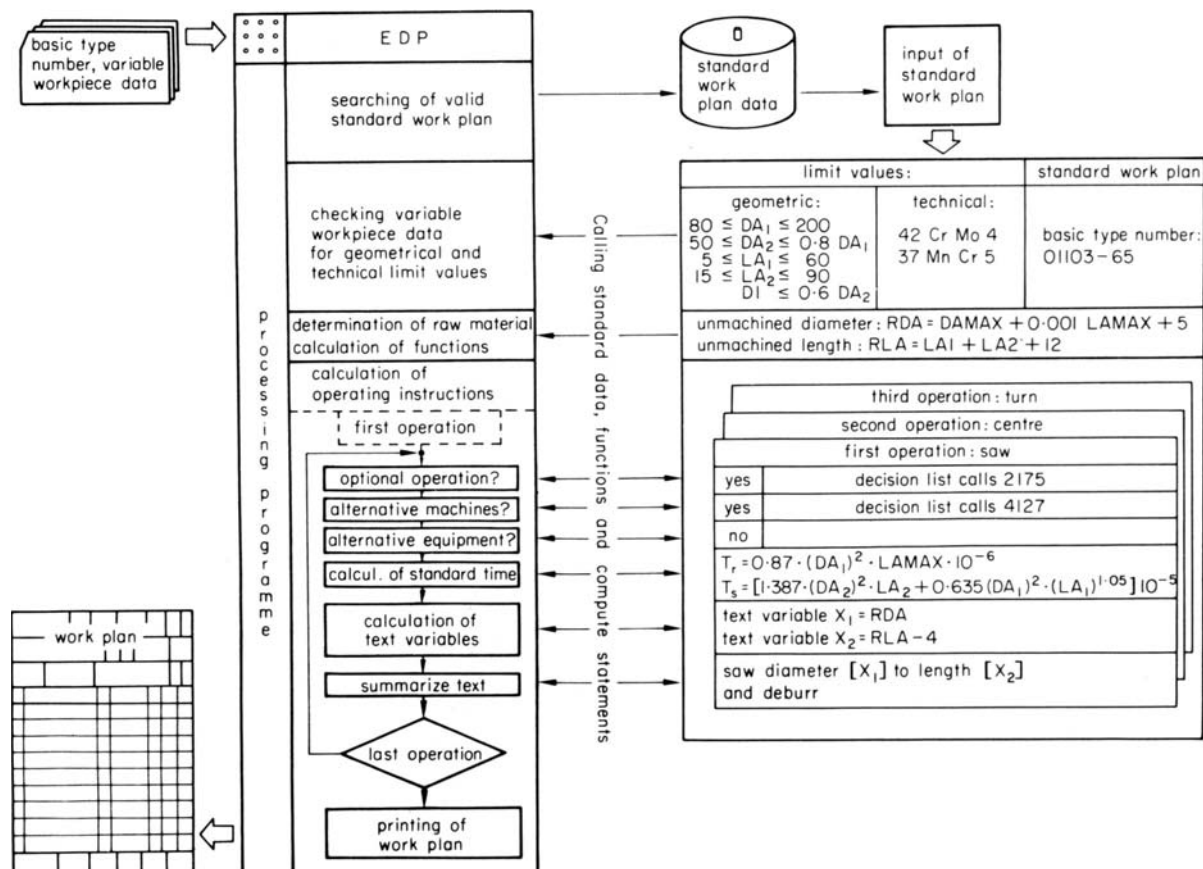


Figure 10. Automatic setting up of work plans using the variant principle.

sions of the raw material are determined. Then the operating instructions, production means and standard times are fixed for each manufacturing operation. When all manufacturing operations contained in the standard work plan have been processed, the complete work plan is printed.

This programme for the automatic setting up of work plans can be integrated with the programme for the automatic preparation of drawings into an overall system which is also based on the principle of variants. This integration not only increases the economy of the programme systems but at the same time greatly reduces the throughput time required to prepare the manufacturing documents.

In the planning of the production facilities, the application of group technology again allows rational and systematic procedures. The optimal exploitation of machinery is often a problem in companies with small or medium batch production. Job-lot sizes which are too low do not allow the use of expensive special machinery with its economic advantages. Frequently commercially available universal machines are not sufficiently exploited throughout their range of useful performance because sufficient information on the future parts spectrum to be machined is not available at the time of their acquisition.

Purposeful and economic exploitation of production devices will be possible through the application of part family manufacturing (see section 3.5), which means order processing independent of type of orders and products. This allows us to restructure the workshop so that non-interlinked machine lines (so-called cellular groups) are available for certain part families²⁸.

The machines are lined up in the general sequence of operations without exacting a prescribed transit or cycle time. This results in a number of advantages as, for instance, savings in set-up time, transport time and storage time. Production which does not depend on a particular type of products, however, also calls for a corresponding scheduling of production facilities. Therefore, increased attention must be paid to structural and investment planning because the acquisition of new machinery will roughly fix the production methods for at least ten to fifteen years thereafter, and thus determine the important factors which will affect future manufacturing costs.

A prerequisite for the planning of production facilities is a systematic survey of the parts spectrum to be machined and the resulting manufacturing problems²⁹. This survey can be obtained from analysis and sorting of the whole parts spectrum. This takes the form of catalogues of recurring parts or variants which are arrived at with the aid of the workpiece describing classification system (figure 11).

The parts spectrum is then subdivided into groups with similar operations ensuring the manufacture of these parts by a machine line. Examples of such parts families are short or long lathe workpieces.

The compilation of parts families is greatly facilitated by the use of a so-called field of code numbers. A field of code numbers includes the classification of all parts for which the same or a similar sequence of machining operations can be expected depending on certain criteria in the various

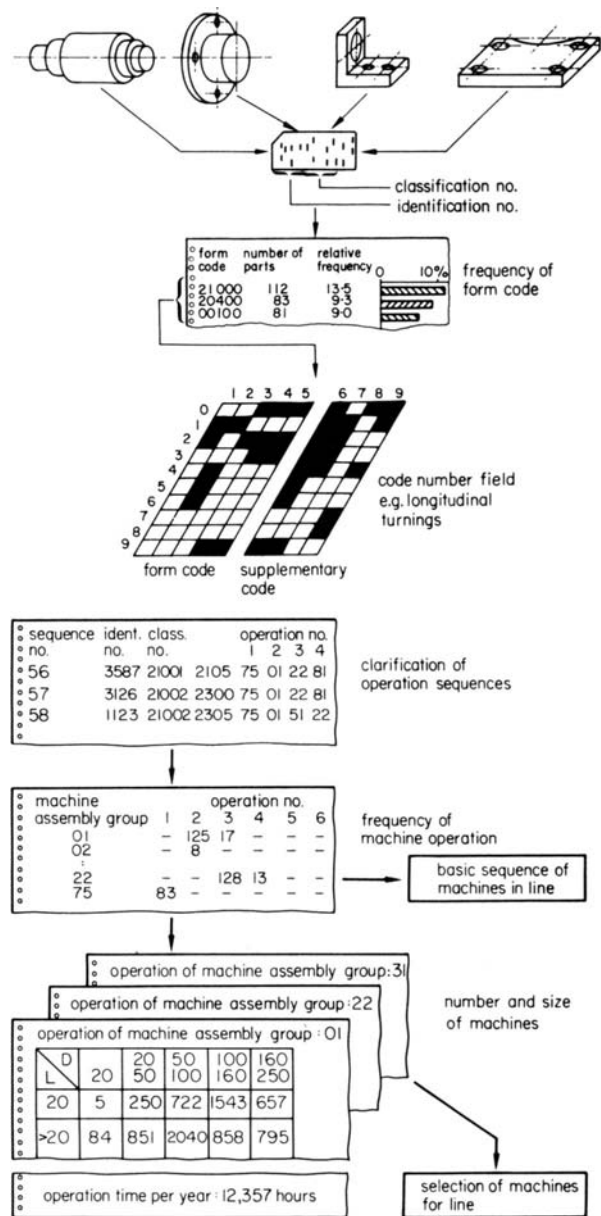


Figure 11. (a) Systematic procedure for setting up unlinked transfer lines (1). (b) systematic procedure for setting up unlinked transfer lines (2).

code number digits. The machining requirements for a part family are obtained by analysis of the associated work plan data. These include data on:

- (1) the type of machining,
- (2) the space of machining,
- (3) the accuracy of machining,
- (4) the machine loading time,
- (5) the equipment of machines.

From these data it is possible to determine the machine required for each operation. The basic sequence of the machines in a line can be found by determining those machines which are most frequently required for each sequence of operations. During manufacturing of the components not every machine in the basic sequence must be employed. It is possible to omit certain machines where necessary.

To determine the number of machines required for each operation the loading time for each operation is

	1	2	3	4	5	6	7	8	9	10	11	
	code number	main function	basic function	name of functional group	structural components	bearings and connections of structural comp.	maximum transmission ratio	maximum torque (kpm)	maximum dimension	number of components	weight	
0					without	attached	≤ 2	≤ 10	without mark	1	without mark	0
1					1 spur wheel	by closed linkage	$> 2 \leq 2.2$	$> 10 \leq 25$	$> 0.1 \leq 0.5$	2 or 3	≤ 10 kp	1
2	general unit assemb. group	transfer and modify power	torque change rotat.		bevel gears	by closing shape	$> 2.2 \leq 2.5$	$> 25 \leq 63$	$> 0.5 \leq 1$	$> 3 \leq 5$	$> 10 \leq 50$	2
3				shaft comb. not selectable	worm gears	0 and 1 or 2	$> 2.5 \leq 3$	$> 63 \leq 160$	$> 1 \leq 2$	$> 5 \leq 7$	$> 50 \leq 100$	3
4				shaft comb. selectable	pulleys	others	$> 3 \leq 3.6$	$> 160 \leq 400$	$> 2 \leq 3$	$> 7 \leq 10$	$> 100 \leq 500$	4
				other components		attached	$> 3.6 \leq 4.5$	$> 400 \leq 1000$	$> 3 \leq 4$	$> 10 \leq 20$	$> 500 \leq 1000$	5
				spur wheels bevel gears		by closed linkage	$> 4.5 \leq 5$	$> 1000 \leq 2500$	$> 4 \leq 5$	$> 20 \leq 40$	$> 1 \leq 5$ Mp	6
				worm wheels		by closing shape	$> 5 \leq 5.6$	$> 2500 \leq 6300$	$> 5 \leq 10$	$> 40 \leq 60$	$> 5 \leq 10$	7
				other combinations		5 and 6 or 7	$> 5.6 \leq 6.3$	$> 6300 \leq 16000$	$> 10 \leq 20$	$> 60 \leq 100$	$> 10 \leq 20$	8
						others	> 6.3	> 16000	> 20	> 100	> 20 Mp	9

Figure 12. Assembly classification as an aid for the selection of assembly families.

ascertained by adding up the times taken by each of the parts machined. The number of components to be produced within a given period of time is found either by analysing previous part spectrums or by considering a number of representative products and by extrapolation.

Using the above described methods, it is possible not only to determine new production capacities, but also to adapt the potential of machine facilities to changes in machining specifications. The last mentioned case is far more frequent in actual practice. The capacity demand is determined in the way we have described and is compared with the capacity actually available. The differences obtained in this comparison will point to special measures by which the machine potentials can be better adapted to the work in hand.

3.4 Manufacturing department

The implementation of group technology principles in the manufacturing department results in part family manufacture. Formation of part families is the logical consequence of the group technology measures discussed in the preceding sections. A part family is generally understood to be a group of components (figure 2, page 9) which are similar in respect of their manufacturing properties and which can thus be jointly processed in the manufacturing department. Parts are said to be similar in production engineering terms if the operations, their number, and their sequence are themselves similar and if they result in a similar flow of materials. The parameters affecting the operations are mainly machining paths, tools, clamping means, and cutting conditions. Moreover, the time-schedule should also be taken into account. The more similar the production factors of individual components, the fewer the problems which will be encountered in introducing part family manufacture. The essential advantages of part family production are:

- (1) savings in set-up time,
- (2) reduction of volume of documents,
- (3) facilitated progress control,
- (4) better calculation and work planning.

These advantages must be weighed against the cost of the planning of part families and of possible storage³⁰.

The types and volumes of part families may differ by degree. Similar components may be selected from sub-assemblies, complete products, or from the whole production programme. Moreover, part families can be based on one single operation or on a complete sequence of operations. Details for this procedure can be obtained from the references given in section 5. It can be generally stated that the individual items of part families should be selected from the whole production programme, that is the whole parts spectrum, and that they should also apply to the complete sequence of operations. Only in this way can a comprehensive exploitation of all advantages of group technology in every department be achieved³¹.

3.5 Assembly department

For a long time the assembly department was neglected in the rationalization of the production process. Companies with small and medium batch production in particular often manufacture their components using the most advanced methods while the subsequent assembly operations have hardly changed during industrial development. The scientific approach to the assembly process is still in its initial stage.

Application of group technology principles to the assembly process has also yielded new ways of rationalization. The same methods as in the manufacturing department may be employed. The only difference lies in the degree of complexity of the objects processed (figure 2, page 9).

The precondition for a standardization of assembly operations is the standardization of the products to be assembled. Therefore the first step must be a systematic break-down of the products by applying the unit construction principle. For the main-function groups, numbers of variants have to be developed and it must be possible for these variants to be assembled according to customer specifications. Standard assembly schedules should be set up for these individual groups in the same way as standard

work plans. The actual sequence of assembly operations can be obtained by combining them according to the unit construction principle³².

In this connection, so-called assembly families should be defined. By analogy with the parts family in the manufacturing department, the 'assembly family' is defined as a number of sub-assemblies or products which are similar in respect of the assembly technology required and which pass through the assembly department³³ jointly. In this way, the conditions for small or medium batch assembly will become comparable with those existing in mass production. The most important advantages are thus:

- (1) reduction of planning effort, in particular a simplified time and cost calculation;
- (2) facilitated time planning;
- (3) feasibility of employing more machines and fixtures.

Because of the more detailed division of the work load in the assembling process the employment of more semi-skilled, less expensive labour will be possible.

Selection criteria for the compilation of assembly families are included in a classification system for sub-assemblies, which was developed by the Institute of Machine Tools and Production Engineering at the Technical University of Aachen, West Germany (figure 12). It is part of the overall coding system presented in section 3 which is based on a total of eleven digits where the first four digits indicate the function and a verbal description³⁴. The heavy lines frame the so-called field of code numbers for the group of shafts shown in the lower left part of the figure as an example for an assembly family. For a joint planning of these units it is necessary to standardize the often rather different assembly plans in order to minimize the number of different assembly operations. This procedure finally leads to a so-called complex assembly plan. An example of such a plan is shown in Figure 13. This shows the above set of shafts, which are further characterized by the field of code numbers appearing in the upper left-hand part of the figure.

These standardizing measures frequently permit substitution of the stationary assembling, which is conventional in small and medium batch production, by assembly lines³³. To obtain a survey of the whole inventory of assembly groups, the frequency of each individual classification number has to be determined. The survey so obtained will permit a delimitation of those sub-assemblies which are similar in their design relationship. After a conclusive investigation of the necessary operation sequences, the assembly families are determined. A complex assembly plan is then provided for each assembly family. In combination with a thorough time analysis this plan will provide the basis of the layout of an alternative-flow assembly line. The precondition for the smooth operation of the assembly line is the elimination of those operations which cannot be planned in advance, such as adapting work and activities which contribute nothing to the progress of assembly work. The accuracy of NC machinery has largely eliminated the need for such activities.

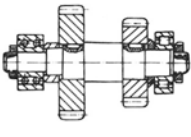
code number											assembly family no. 12												
position	0	1	2	3	4	5	6	7	8	9	10	11											
	group assembly plan											assembly group ident. no.											
operation no.	performance description	standard time t _s	fixture designation	special tool designation no.																			
					202401	224023	240801	241101	252201	260201	263301												
1	11	fitting leather key	5,0	measuring fixture	3821	X	X	X	X	X													
	12	pressing in leather key	1,0	pressing in fixture	3211	X	X	X	X	X													
	13	securing leather key with 2 allen screws	1,5	pneumatic screwdriver	2431	X			X	X													
2	21	fit gearwheel 1	1,5			X	X	X	X	X	X	X											
	22	fit spacer ring	1,0			X	X		X		X	X											
	23	fit gearwheel 2	1,5						X	X	X												
	24	secure lock ring	1,0			X		X	X	X	X	X											
	25	fit anti-friction bearing	1,2	mounting fixture		X	X	X	X	X													

Figure 13. Complex assembly plan.

4 CONCLUSIONS AND OUTLOOK

Beginning with a short survey of the development of group technology as an universal organizational principle for all production departments some examples of its application in the design, work planning, manufacture and assembly departments have been presented to show some of the effects of group technology upon rationalization in production. While previous applications mainly took place in individual departments of a company, the latest trend tries to provide a continuous flow of information for parts and functional groups throughout the whole process of production starting from the design department through to the preparation of punched tapes for NC machining. This trend is supported by the increasing introduction of numerical and graphic data processing. On the whole it can be said that the idea of group technology will gain even more importance in future.

5 REFERENCES

1. S. P. Mitrofanov. *Wissenschaftliche Grundlagen der Gruppentechnologie*, 1960.
2. Lange and Rossberg. *Wege zur wirtschaftlichen Fertigung im Arbeitsmaschinenbau*, 1960.
3. W. Michels. Anlagen- und Geräteklassifizierung—Ein Instrument zur Rationalisierung im Planungs- und Produktionsbereich. *Industrie-Anzeiger* (1971), 93, nos. 15 and 24.
4. K. Brankamp and W. von Bardeleben. Die Vorrichtungsklassifizierung—eine Rationalisierungsmethode im Betriebsmittelbereich. *Industrie-Anzeiger* (1970), 92, no. 11.
5. H. Opitz, K. Brankamp and H.-P. Wiendahl. Aufbau und Anwendung eines funktionsorientierten Baugruppenklassifizierungssystems. *Industrie-Anzeiger* (1970), 92, no. 31.
6. D. Zimmermann. *Eine allgemeine Formenordnung für Werkstücke—Gestaltung, Handhabung und Rationalisierungserfolg*. ZAFO, 1967.

7. H. Opitz. *Die richtige Sachnummer im Fertigungsbetrieb*, 1971.
8. K. Brankamp and E. Dirzus. Zielsetzungen und Aufbau eines einheitlichen Sachnummernsystems für einen Industriekonzern. *Industrie-Anzeiger* (1970), 92, no. 104.
9. J. Groebler. Untersuchung über den Aufbau eines Sachnummernsystems. *Industrie-Anzeiger* (1966), 88, no. 93.
10. W. Eversheim. Zur Einführung eines Klassifizierungssystems. *Industrie-Anzeiger* (1967), 89, no. 84.
11. H. Opitz. Einführung in die Probleme der Werkstückklassifikation. *Industrie-Anzeiger* (1967), 89, no. 41.
12. D. Grässler. Aufbau und Systematik des VDW-Werkstückklassifizierungssystems. *Industrie-Anzeiger* (1967), 89, no. 41.
13. H. W. Homann, H. Gühring and K. Brankamp. Ein Klassifizierungssystem für Stahlbaueinzelteile—Entwicklung und Beschreibung des Systems. *Industrie-Anzeiger* (1970), 92, nos. 6 and 15.
14. E. Dirzus, H. Figger, Ein Klassifizierungssystem für Stahlbaueinzelteile. *Industrie-Anzeiger* (1970), 92, no. 24.
15. K. Brankamp, H. Olbertz and R. Schütze. Aufbau und Anwendungsmöglichkeiten eines Klassifizierungssystems für Schweißteile. *Schweißen und Schneiden* (1970), 22, Heft 12.
16. H. Galland. Werkstücksystematik und Teilefamilienfertigung. *Industrie-Anzeiger* (1963), 85, no. 37.
17. J. Schleppegrell and K.-H. Kambartel. Systematische Angebotsplanung in Maschinenbauunternehmen. *Industrie-Anzeiger* (1970), 92, no. 79.
18. K. Brankamp and H.-P. Wiendahl. Systematische Erzeugnisgliederung als Voraussetzung für eine rationelle Auftragsabwicklung. *Industrie-Anzeiger* (1968), 91, no. 106.
19. Autorenkollektiv. Die Konstruktion—ein Schwerpunkt der Rationalisierung. *Industrie-Anzeiger* (1971), 93, no. 60.
20. H. Opitz. Produktplanung—Konstruktion—Arbeitsvorbereitung — Rationalisierungsschwerpunkte bei der Produktentstehung, 1971.
21. W. Eversheim and H.-P. Wiendahl. *Rationelle Auftragsabwicklung im Konstruktionsbereich*, 1971.
22. R. Simon. Einsatz des Klassifizierungssystems im Konstruktions- und Normungsbereich. *Industrie-Anzeiger* (1967), 89, no. 41.
23. G. Steinmetz. *Integrierte Konstruktion und Arbeitsplanung für Varianten*. Dissertation TH Aachen, 1973.
24. H. H. Gerlach and H.-P. Wiendahl. Das Speichern und Verarbeiten von Variantenstücklisten mittels elektronischer Datenverarbeitungsanlagen. *Werkstattstechnik* (1970), 60, Heft 8.
25. W. Olbrich. *Arbeitsplanerstellung unter Einsatz Elektronischer Datenverarbeitungsanlagen*. Dissertation TH Aachen, 1970.
26. Autorenkollektiv. Automatisierung der Arbeitsvorbereitung. *Industrie-Anzeiger* (1971), 93, no. 60.
27. B. Feldhege. Einsatz des Klassifizierungssystems in der Fertigungsplanung. *Industrie-Anzeiger* (1967), 89, No. 50.
28. J. Herrmann. Arbeitsvorbereitung bei Teilefamilienfertigung. *Industrie-Anzeiger* (1965), 87, no. 103.
29. H. Opitz, H.-P. Wiendahl and H. Grabowski. Investitionsplanung einer Maschinenfabrik unter Berücksichtigung der Werkstückanforderung. *Werkstattstechnik* (1971), 61, no. 1.
30. P. Stehle. Einsatz des Klassifizierungssystems im Fertigungsbereich. *Industrie-Anzeiger* (1967), 89, no. 66.
31. H. Opitz, K. Brankamp and W. Olbrich. *Untersuchungen über die Einsatzmöglichkeiten von Datenverarbeitungsanlagen für die Einführung der Teilefamilienfertigung*. Forschungsbericht des Landes Nordrhein-Westfalen Nr. 2093, Westdeutscher Verlag, Köln und Opladen, 1970.
32. K. Brankamp and J. Herrmann. Baukastensystematik—Grundlagen und Anwendung in Technik und Organisation. *Industrie-Anzeiger* (1969), 91, nos. 31 and 50.
33. J. Arlt and M. Miese. Rationalisierung der Montage (Voraussetzungen und Möglichkeiten). *Industrie-Anzeiger* (1971), 93, no. 49.
34. H.-P. Wiendahl. *Funktionsbetrachtungen technischer Gebilde—Ein Hilfsmittel zur Auftragsabwicklung in der Maschinenbauindustrie*. Dissertation TH Aachen, 1970.
35. R. Hahn, W. Kuhnert and K. Roschmann. Die Teileklassifizierung: Systematik und Anwendung im Rahmen der Betrieblichen Nummerung. *Handbuch der Rationalisierung. RKW Schriftenreihe* (1970), 21.

TOTAL COMPANY APPRAISAL FOR GROUP TECHNOLOGY

by

K. RATHMILL*, P. BRUNN* and R. LEONARD*

SUMMARY

The decision to change to Group operation carries profound commitments for a company. The total operation of the firm must be re-established on Group-Technology principles and this will require considerable amounts of time, effort and money. Hence it is imperative that an accurate estimation of the total potential benefits be established at the initial stage of company appraisal, such that the ensuing advantages may be compared with the cost of implementation.

This work demonstrates that the relative benefits of group technology may be evaluated in each department of an enterprise by breaking down the departmental activities into their basic constituents. Thus the total benefit to the company may be obtained by summing the individual contributions.

A complete 'suitability check list', successfully used for the appraisal of a medium-sized engineering firm is given. It is suggested that with suitable small modifications for specific company considerations the list forms a relatively speedy, yet thorough, method of appraisal.

INTRODUCTION

Group technology offers advantages in all facets of a company's operation. However, some enterprises, by virtue of the nature of their product, will benefit by the adoption of group technology to a greater extent than others. The time, effort and cost of changing to group operation can be considerable, and difficulties will be experienced in the management structure, Edwards¹. Thus it is strongly advocated that the company be given a thorough appraisal at the initial stage of consideration, such that the likely benefits may be assessed and compared with the cost of implementation.

Leonard and Koenigsberger² described the factors which an enterprise must consider before adopting group operation, these included considerations such as the typical batch size, the life of the product, and the skill required by the work force. Thornley³ listed the general achievements produced by group technology and indicated these advantages diagrammatically (figure 1). In a similar manner Crook⁴ specified areas of company operation where clear advantages would result from group technology (figure 2). Bennett and MacConnell⁵ stated the aims which motivate managements to apply group production and Knight⁶ suggested a cost comparison between functional and cellular layouts. In this work, an attempt is made to assess the total benefits, both financial and structural, which will result from group operation. Hence an answer can be ventured to the

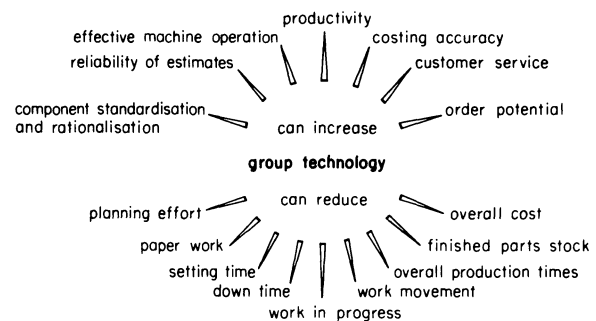


Figure 1. General achievements of group technology.

question: 'What will it cost and how much will we gain?'

Company suitability for group technology—the general case

The suitability for group technology will depend upon the type of company being considered, but as indicated by the works of Bennett and MacConnell⁵, Durie⁷, Ribbons and MacConnell⁸, and Marklew⁹, the field of application would appear to be very broad. The flexibility of the principles of group technology must surely be one of the technique's most dynamic characteristics. When a comparison is made between the jobbing shop, Ribbons¹⁰, Serck Audco, Ranson¹¹; Ferranti⁹; and Ferodo, Connolly et al.¹², the search for characteristics which indicate

* University of Manchester Institute of Science and Technology

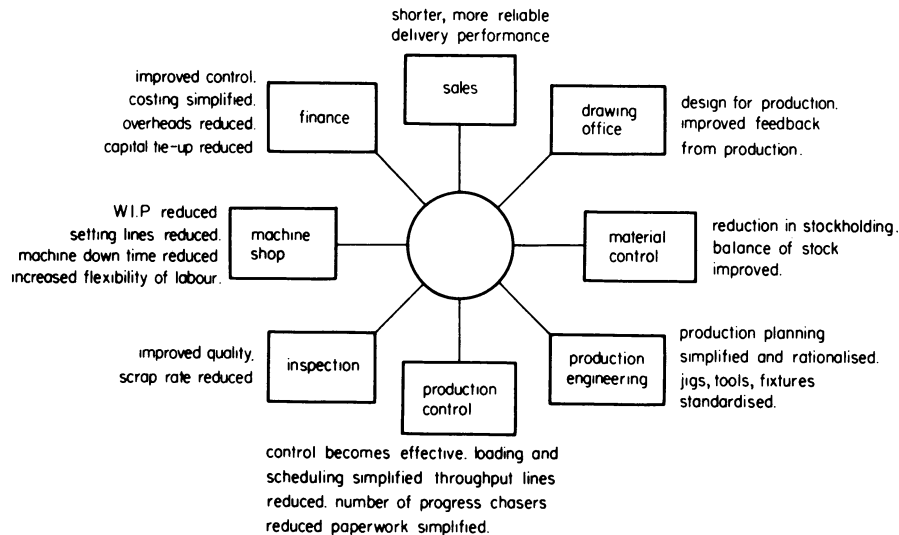


Figure 2. Assessing the benefits of group technology.

a common potential for group technology in other companies, is difficult.

Company suitability for group technology can be seen from two standpoints:

- (a) The compatibility of existing systems with respect to the introduction of group technology.
- (b) The compatibility of company parameters with respect to the economic use of group technology.

Factors affecting the cost of introducing group technology are of importance, although relatively short time in effect, they are the means by which the 'return on investment' can be assessed in order to justify the necessary changes. Factors involved in the long-term economies of group technology should be of overriding importance since, however formidable the challenge of the introductory problems, where a strong case for group technology has been made in the long-term, the responsibilities for management are clear.

Assessing the problems of introducing group technology

The accuracy of existing company data is of great importance, for the discipline of group technology relies upon accurate production, design and sales information¹. Where adequate information is not available a costly programme to generate data is required. The company may benefit by the use of a classification system. Although it is a rule which has exceptions, the use of classification systems is fundamental to the principles of group technology. To justify the costs involved in introducing a coding system, Middle et al.^{1,3}, the company should extract the maximum benefits. A company's product range must be assessed with respect to a given classification scheme before conclusions can be made. A relevant point concerning product life is made by Leonard and Koenigsberger: 'if the products made by the company have a relatively short life and new designs are frequent, the use and advantages of such standard parts can be quickly established. If, however, the product life is long and spare parts are required by

customers for some considerable time the change to standard components will occupy a number of years and the undoubted advantages of variety reduction will take time to emerge'². A value not easily quantifiable is the 'cost' resulting from changes in managerial outlook and responsibility. Ranson¹¹ indicates that resistance to change by management can be formidable and in some cases, decisive, thus changes within the management structure may be necessary to overcome such opposition.

Assessing the long-term benefits of group technology

The prime consideration for the introduction of group technology is that the type of manufacture should be batch production, that is, somewhere between one-off and high-volume production. A large number of small batches has been cited as being most suitable for group manufacture² and Lewis¹⁴ states that group technology presents a wide region of application between the extremes of batch sizes. Where trends in component-type demand can be determined, the use of group technology can be extended into the jobbing region¹⁰.

Customer delivery requirements are of importance when considering group technology. One of the major benefits is that of reduced and predictable production times resulting in an increased ability to meet demands for urgent deliveries. Hence where delivery problems are acknowledged under functional layout a potential benefit from group technology is indicated. Raw-material supply is of significance when contemplating group technology. Coupled with a reduced throughput time group technology is attributed with the advantage of less work in progress. Whilst this claim is borne out in practice¹¹ gaining the benefits of reduced work in progress puts a much greater dependence upon raw-material supply.

The physical weight of components is significant, for the concept of group production implies an efficient transfer of components such that queuing or stockpiling between processes is reduced. Although this may be overcome by handling equipment, it is clear that light components present fewer problems.

Where inspection cannot be reduced, the necessity to transfer components into inspection tables after

each operation is likely to cause excessive interruptions to work flow. Whilst the use of travelling inspectors with special inspection aids is recommended Marklew¹⁵ it is clear that components requiring little inspection or where inspection responsibility can be delegated to the operator, group production is more straightforward.

Low-skill-content work and flexibility of labour tend to make the operation of a group technology machining group efficient. In order to attain the best man/machine utilization labour should be flexible and able to move within the group as required. Flexibility requires the co-operation of the trades unions, also the skill requirements of the various jobs must not be excessive. For highly skilled operations such as spinning and accurate lathe work, it is doubtful whether the respective operators could achieve a satisfactory all-round performance. Where low skill operations are the norm a much greater distribution of proficiency within the group enables the maximum advantages to be achieved from the flexibility. Inexpensive plant is wholly compatible with group production and simplifies the problems of man/machine utilization. The costs considered when optimizing man/machine utilization are the machine-hour rate and the labour costs in terms of man-hour rate. Where plant is inexpensive the machine-hour rate will be considerably below the man-hour rate. Thus efficient group operation and manpower utilization can be attained at the expense of inferior machine-tool utilization.

The natural division of components into families is a feature which varies from firm to firm. The propensity for components to form families depends upon whether they require similar processes. Whilst much can be done to redesign and reorganize components to comply with the process requirements of others, certain components will defy integration and remain 'special'. The proportion of components which defy family formation indicates the extent to which group technology can be applied within the production system, and is therefore relevant when determining the savings offered by group technology.

Process divisibility greatly assists the design of group technology 'cells'. Examples of poor divisibility are heat treatment, shot blasting, electroplating, and paint spraying. Due to the centralized nature of these processes it is difficult to design effective group production systems, consequently the savings from group technology are reduced. A high dependence of the production system on indivisible processes, together with a lack of similarity of process requirements, greatly increases the difficulty of family formation and production-group design. A factor closely related to the above is the effects upon the production system of sub-contracted operations.

The appraisal procedure

Company appraisal for group technology can be subdivided into four, fairly well defined stages:

- (1) Critical examination of the company functions.
- (2) Defining the type or variety of group technology application most suitable.

- (3) Defining the most suitable approach to implementation.
- (4) Assessing the costs and savings of the proposed changes.

(1) Critical examination

The first step involves collecting information from each department of the organization. Much time can be devoted to areas which will contribute little to the conclusions of the appraisal. However, completeness is essential since the maximum benefits will ensue from the 'total' approach¹¹, this being based upon a complete examination of the company. A checklist provides a useful aid for the critical examination of a company, and enables the findings to be summarized. The 'suitability checklist' also provides a convenient means of comparing and weighing the advantages and disadvantages implied by group technology in terms of both quantitative and qualitative values.

The checklist given in Appendix 1 represents a series of questions which relate the various aspects of group technology to the functions of the company. Alongside each question two spaces are provided to record the value, either positive or negative, of that particular aspect of group technology in terms of the *status quo*. It must be recognized that some effects are quantified more easily than others. Thus values can be expressed in either quantitative or qualitative terms. The more easily quantifiable effects should be estimated as accurately as possible and the resultant value inserted into the appropriate column. For positive or negative quantifiable effects such as liberated investment in work in progress, and the cost of coding drawings, of a 'once off' or short-term nature, the total value should be stated in pounds sterling. If the effect is of a long-term or consistent nature the value is stated per annum (p.a.). For qualitative effects the treatment is similar, but the scale of the effect is indicated by means of an 'A' or 'B' classification. Where qualitative effects are relatively small or localized the classification 'B' is used, conversely if such effects are of high importance the classification 'A' is given.

The introduction of qualitative factors is detrimental to a rigorous approach. However, it is necessary to include qualitative effects because of their importance in assessing the total effects of group technology. Furthermore, a rigorous treatment of the effects of group technology, Jones¹⁶, is not within the scope of this work.

(2) Defining the 'type' of group technology application

The most suitable 'type' of group technology application for a company should be obtained from the critical examination, and the checklist should focus attention on major areas of potential.

The 'type' of group technology application referred to at this stage involves the following considerations:

- (a) The form of the machine groups indicated.
- (b) The extent of application of group technology methods in the manufacturing system.
- (c) The form of the changes in company systems.

- (d) The extent to which company systems, including management, will be affected.

(3) Defining the approach to group technology

From a knowledge of:

- (a) the state of existing systems within the company,
- (b) the 'type' of group technology application which is indicated as being suitable for the company,
- (c) other factors such as management policy, market conditions, labour relations, etc.

A suitable approach to implementing the changes necessary to achieve the objectives posed by (b) above is required.

Aspects to be considered in defining the approach are:

- (a) The method to be used in determining component families and machine groups.
- (b) Planning related activities such as value engineering, updating existing files, the generation of new data and educational programmes.
- (c) Changes in manufacturing and other systems and the rate of implementation.
- (d) The overall time-scale of the programme.

(4) Assessing the costs and savings of group technology

A most important part of the appraisal is that of assessing the relative value of gains and losses incurred by the implementation of group technology. A conservative approach to financial gains is considered to be prudent for cultivating the long-term support of management. Of equal importance in the short-term, is the need to obtain the confidence of management

in the integrity of the reasoning that underlies the recommendations of the appraisal report.

Some exponents of group technology believe that a 'hard sell' approach is advised when relating the subject of group technology to industrialists. However, higher management whilst, hopefully, ever ready for changes which will bring about increased profitability, are basically shrewd, cautious and unemotional in disposition, and therefore most likely to respond to a reasoned, cautious and unemotional argument. Results published by Serck Auco and others indicate that group technology can markedly improve performance in a wide range of industries and therefore does not require wildly exaggerated economic justifications.

Potential advantages

The advantages of group technology are maximized in the 'green-field' situation. This describes the case where a factory is as yet unbuilt and its management and production systems can be designed without limitations imposed by existing considerations. Although advantages of improved machine-tool investment will be achieved by a company changing from an existing organization to group technology, the time to achieve these benefits will be extended. In the 'green-field' situation, the maximum advantage, from a definite knowledge of machine tool requirements, can be speedily achieved. Thus the justification for group technology in the 'green-field' situation is much easier than for existing systems which require that problems and costs of change be taken into account. In certain situations partial 'green-field' conditions may prevail; e.g. the fact that a company is about to carry out a major expansion programme is of advantage in justifying the change to group technology.

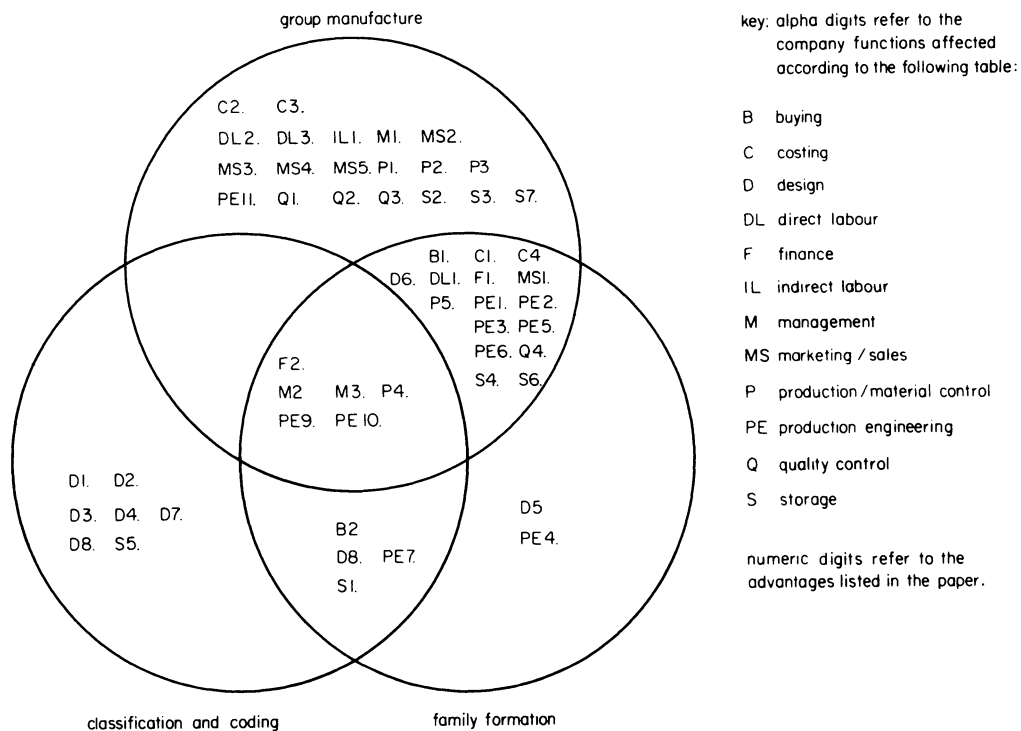


Figure 3. Corporate advantages related to the classification, component family, and group manufacturing aspects of group technology.

For a specific situation only a certain combination of potential advantages are likely to result from a given type of group technology application. A comprehensive list of potential advantages of group technology are as follows:

The code subscript should be used in conjunction with figure 3.

(1) Improved manufacturing capacity can reduce the level of sub-contracted work (B1).

(2) A reduced variety of bought out component and materials should improve the efficiency of the buying department (B2).

(3) Increased man, machine and space utilization should reduce overheads (C1).

(4) The group manufacturing system presents the opportunity to adopt a more effective incentive scheme (C2).

(5) Group manufacture enables a simplified yet accurate costing system to be implemented (C3).

(6) Group technology can reduce or eliminate the effort of a physical stocktake (C4).

(7) Classification and coding can reduce drawing retrieval effort (D1).

(8) Economies derived from variety reduction and standardization can be achieved through classification and coding (D2).

(9) The number of drawings in the filing system can be reduced as a result of variety reduction and standardization (D3).

(10) Duplication of design effort of components is avoided by a classification scheme (D4).

(11) The formation of component families can focus jig-and-tool-design effort (D5).

(12) The definition of production capacity in group technology enables draughtsmen to design for production (D6).

(13) Classification should prevent duplication of jig-and-tool-design effort (D7).

(14) The use of a coding system in conjunction with families of components presents an ideal foundation for value analysis and engineering (D8).

(15) Group technology can provide the means for higher earnings for direct labour (DL1).

(16) Group manufacturing systems offer greater job satisfaction to direct labour, thereby improving industrial relations (DL2).

(17) Better working conditions, offered by group manufacture, reduces labour turnover (DL3).

(18) Reduced investment in terms of floorspace, work in progress, and stocks improves liquidity (F1).

(19) Group technology offers improved profitability in an overall sense (F2).

(20) Group manufacture reduces the cost of indirect labour by reducing the need for inter-process transportation (IL1).

(21) Group manufacture benefits from the unification of responsibility of first-line management (M1).

(22) Group technology improves the definition of management responsibility, thus creating a more effective management structure (M2).

(23) Group technology initiates more effective company operation, based upon integrated systems (M3).

(24) Group technology can reduce manufacturing

costs and thus allow greater flexibility with profit margins (MS1).

(25) Group technology improves the competitive ability of a company by reducing the time between an order and delivery (MS2).

(26) Group technology can improve customer relations by using a defined production plan, known in advance, thus delivery dates can be accurately predicted (MS3).

(27) Accurate costs of 'specials' are more easily determined: this enables realistic prices to be quoted (MS4).

(28) A more accurate product cost enables decisions regarding product range rationalization to be made (MS5).

(29) Group technology presents a simplified and more effective method of controlling production (P1).

(30) Group technology defines production capacity more accurately and therefore aids disciplined production planning (P2).

(31) Group manufacturing systems plus improved production control eliminates the need for progress chasers (P3).

(32) Group technology systems reduce, in general, the complexity of paperwork systems (P4).

(33) The problem of 'shortages' can be drastically reduced by a better stock balance achieved by period batch control (P5).

(34) In certain circumstances group technology can improve man/machine utilization (PE1).

(35) Group manufacturing methods may reduce setting times (PE2).

(36) Floorspace utilization should be improved by group technology (PE3).

(37) Group technology provides a basis for determining future plant investment (PE4).

(38) Group manufacture increases the use of high-volume tooling methods, thus reducing machining time (PE5).

(39) Group technology facilitates the use of standard plannings (PE6).

(40) Component classification and family formation reduce the effort of planning similar components (PE7).

(41) A master planning file, in code-number sequence, can greatly assist production engineering projects (PE8).

(42) Due to variety reduction and the use of standard plannings, the number of plannings on file is reduced (PE9).

(43) Group technology generates accurate data for use in future production-engineering projects (PE10).

(44) Work handling can be reduced by the use of conveyors (PE11).

(45) Group technology improves quality control by unifying responsibility (Q1).

(46) Greater job satisfaction offered by group manufacture improves component quality (Q2).

(47) Group manufacture allows the delegation of inspection responsibility to direct labour (Q3).

(48) The specialization which occurs within the group leads to the development of more efficient inspection techniques (Q4).

(49) Group technology offers simplified materials storage due to variety reduction of raw materials (S1).

(50) Better control improves the stock balances, hence reduced investment in material stocks (S2).

(51) Group technology ensures the discipline of batches and stores procedures, hence control in the finished parts stores is improved (S3).

(52) Smaller batches made possible by group technology reduce the stocks of finished components thereby improving storage-space utilization (S4).

(53) Component and material classification produce a logical stores layout (S5).

(54) Reduced stocks of finished components imply less investment in stocks (S6).

(55) The elimination of inter-process storage represents a great reduction of investment in work in progress (S7).

Potential disadvantages

The number of disadvantages experienced by a company depends upon the type of group technology application introduced. Treating the cost of the changes as a disadvantage, a list of disadvantages of group technology follows:

(1) Where high and low-value equipment is used within the same group costing is not simple.

(2) The cost and effort entailed in classification achieve little variety reduction and standardization where product obsolescence is slow.

(3) The task of classification and coding of drawings is costly and time-consuming.

(4) Operating a classification system may require a coding officer.

(5) If the existing drawing system has much cross referencing and many detailed drawings are on one sheet, the costs of redrawing components on to individual sheets can be large.

(6) The costs of training men in a variety of skills to achieve flexibility is costly and time-consuming.

(7) For skilled work, misused flexibility may lead to a reduction in overall efficiency, the man may become a 'Jack of all trades but master of none'.

(8) The change to group manufacture may require lengthy negotiations with men and unions.

(9) The unification of responsibility achieved through group manufacture may produce a lack of specialized supervision.

(10) The implementation of group technology takes time, thus a lack of total commitment by top management may result in premature abandonment.

(11) Group technology represents considerable change throughout all aspects of an organization, companies not used to innovation may experience difficulties.

(12) During the transition stage much effort is required of middle management, where management is weak or inflexible, the implementation can be extended while changes in personnel occur.

(13) With group technology, machine group loading is critical for the optimum man/machine utilization, effort may be required for this function.

(14) Group manufacture may reduce machine-tool utilization due to facility duplication in different groups.

(15) Group manufacture reduces the flexibility of the manufacturing system; problems may occur for changes in component-type demand.

(16) An over-reduction of stock levels may give increased sensitivity to short-term supply problems.

(17) Investment in work-handling equipment may be necessary to reduce work in progress if components are heavy.

(18) The cost of reorganizing the plant layout may be large.

(19) Greater maintenance problems occur with group technology due to increased 'down time' losses caused by individual machine failure.

CONCLUSIONS

Companies differ in their products, size, and general operation. However, group technology has been shown to be suitable for the majority of firms engaged in batch production in industries of a most diverse nature. The decision to adopt group operation is readily made for the 'green-fields' situation and for the case of a factory extension. However, if the decision requires the total reorganization of an existing enterprise, considerable effort will be needed before the outcome is successful. Thus a thorough 'Suitability Appraisal' is necessary at the commencement of the exercise, such that the firm may assess the relative values of effort and reward.

APPENDIX 1

Suitability Check List (Data from small engineering firm)

	Remarks	Value	
		+	-
<i>The design function</i>			
To what extent could a classification and coding system improve drawing retrieval?		£415 p.a.	£
Could variety reduction and standardization be achieved by use of a classification and coding system?		(A)	A
Could the duplication of component design effort be avoided by use of a classification and coding system?		B	B
Would the use of a classification scheme + family formation of components assist jig and tool design?		A	A
		(B)	B

<i>The design function</i>	<i>Remarks</i>	<i>Value</i>	
		+	-
Would a group manufacturing system encourage designing for production?		A ⓑ	A B
Could classification and coding prevent duplication of jig and tooling effort?		A ⓑ	A B
Is the product life compatible with the short-term gains achieved by standardization?		A B	A ⓑ
What cost would be incurred by coding existing drawings?		£	£450
What are the annual running costs of the classification system?	Negligible	£	£
What would be the cost of redrawing detailed parts onto individual sheets?		£	£750
Could the size of the drawing file be decreased by the introduction of a classification scheme?		A B	A ⓑ
Would G.T. present opportunities for savings from exercises such as value analysis and metrication?		Ⓐ B	A B
<i>The production and works engineering functions</i>			
Could G.T. improve man/machine utilization?		A ⓑ	A B
Would a reduction in setting and machining times be achieved by G.T.?	By a small amount	£	£
To what extent might available floorspace be increased by G.T.?	By reduction in W.I.P.	£13,000	£
Could G.T. provide an improved basis for future plant justification?	Urgently required	Ⓐ B	A B
To what extent could the company benefit from the use of standard plannings as a result of G.T.?		A ⓑ	A B
Would a classification scheme reduce planning effort?		A ⓑ	A B
Could a reduction in the number of plannings be achieved by G.T.?	Present system incomplete	A B	A ⓑ
Will subcontracted intermediate operations interfere with machine-group formation and operation?	Very little	£	£
Is the indivisibility of processes within the company likely to interfere with machine-group formation and operation?	Shot-blast, etc. to form separate group	A B	A ⓑ
What costs would be incurred in bringing production data up to required standards?	data good but incomplete	£	£
Is capital value of equipment likely to cause difficulties in obtaining an economic machine utilization?	Extremely low-value equipment	£	£
To what extent does the company's component range suggest a potential for group manufacture?	Good potential indicated	Ⓐ B	A B
Are large or heavy components likely to require additional handling costs under group manufacture?	Small components	£	£
Are changes in product mix likely to result in changes in production requirements which could not be met by G.T.?	Insignificant flexibility requirements indicated	£	£
Could G.T. bring benefits from a decentralization of jigs and tooling?	None	£	£
Would conveyors assist in streamlining material flow under G.T.?	Yes	A ⓑ	A B
What costs would be involved in rearranging the factory layout for G.T.?	Based on the movement of most of the plant	£	£5000
To what extent could G.T. bring about a reduction in work handling indirect labour costs?		£1500 p.a.	£
Would G.T. involve higher maintenance costs?		£	£300 p.a.
<i>Quality control function</i>			
To what extent could G.T. improve quality, reduce scrap and reduce rectification costs?	£1000 p.a. if subcontracted welding made-in	£1000 p.a.	-

		<i>Value</i>	
<i>Quality control function</i>	<i>Remarks</i>	+	—
Are the levels of inspection compatible with G.T.?	Very compatible	—	—
To what extent are the types of inspection compatible with G.T.?	Very compatible	—	—
To what extent could the quality control effort be reduced by the unification of responsibility, the delegation of inspection responsibility to direct labour and the development of specialized inspection techniques?		A ⓐ	A B
<i>The buying function</i>			
Might a reduction of material stocks be supported by improved supplier relations?	Supply situation currently deteriorating	—	—
To what extent could G.T. lead to a reduction of subcontracted work?	indications of excellent machine-tool investment	ⓐ B	A B
To what extent could G.T. lead to a reduction in B/O components?	Small benefits indicated	—	—
<i>The costing function</i>			
Does G.T. present an opportunity for improving the costing system?		ⓐ B	A B
To what extent does G.T. offer an opportunity for improving present incentive-scheme arrangements?		A ⓐ	A B
<i>The production planning control and stock control functions</i>			
Could G.T. bring about a more effective production-control system?	+ £5000 p.a.	ⓐ B	A B
Would G.T. improve present production-planning procedures?		ⓐ B	A B
Could a reduction in progress-chasing staff be made with G.T.?		£1500 p.a.	£
Would G.T. lead to an increase or decrease in paperwork?		A ⓐ	A B
Are batch sizes capable of being profitably reduced under G.T.?	Reductions possible	—	—
To what extent could G.T. reduce the stock-taking effort?		A ⓐ	A B
Would G.T. lead to a reduction in raw material stocks?	Not determined	—	—
Could G.T. lead to a reduction in finished-parts stocks?	Interest not included	£26,000	£
To what extent could G.T. lead to a reduction in W.I.P.?	Interest not included	£30,000	£
Could G.T. improve discipline and control within the stores?		A ⓐ	A B
Might G.T. effect better layout of and better space utilization within the storage function?		A ⓐ	A B
Would G.T. require extra staff for handling the greater effort required in group loading?	Staff used from chasing	—	—
<i>Direct labour</i>			
Will G.T. provide the opportunity for improving labour relations from considerations such as working conditions, job satisfaction, and remuneration?	Has importance for future	A ⓐ	A B
Are problems likely to be met in terms of homogeneity of skill and social compatibility in the formation of G.T. cells?	Negligible	A B	A B
Does training to achieve flexibility cause problems?		A B	A ⓐ
To what extent could G.T. affect absenteeism?		A ⓐ	A B
To what extent is G.T. likely to cause difficult negotiations with trade unions?	Experience indicates few problems	A B	A B

	Remarks	Value	
		+	-
<i>The sales and marketing function</i>			
To what extent could G.T. improve the competitive ability of the company by reducing costs, reducing sales price, or increasing profit margins?		(A) B	A B
Does G.T. offer a means of improving delivery performance in terms of reduced delivery periods and punctuality?		(A) B	A B
To what extent could G.T. provide information of accurate costs of spares and special units which could lead to a rationalization of the product range?		A (B)	A B
<i>The management function</i>			
Could G.T. bring benefits from a unification of responsibility at first-line management?	Not at present	A B	A B
Could G.T. provide a basis for more effective company operation founded upon integrated systems?		(A) B	A B
Could G.T. provide a means of improving the definition of management responsibility and a more effective management structure?		A (B)	A B
Would, if a case were sufficiently proved, the implementation of G.T. have the full support of management, particularly at top level?	Yes	-	-
Is the company sufficiently acquainted with innovation to be able to cope with the change of established method and procedure?	Yes	-	-

Abridged conclusions of suitability appraisal

- (1) The firm is both profitable and well managed.
- (2) Changes are inevitable because of expansion, metrication and competition.
- (3) The company can increase output without extending the premises by liberating floor space occupied by work in progress if group technology is introduced.
- (4) The firm, its products, and operation are well suited to the disciplines of group technology.
- (5) Group technology should be adopted in 'total' but care is required at each stage of implementation, and the excellent-level of industrial relations must be preserved.

REFERENCES

1. G. A. B. Edwards. The management problems of introducing group technology, *Proc. 1st Int. Seminar on G.T., I.L.O., 1970*.
2. R. Leonard and F. Koenigsberger. Conditions for the introduction of G.T., *Proc. 13th Int. M.T.D.R. Conf., Birmingham, September, 1972*.
3. R. H. Thornley. Group Technology—A complete manufacturing system, Inaugural lecture, University of Aston in Birmingham, 14 October, 1971.
4. M. Crook. Group technology—A production technique backed by design, *Design Engineering*, pp. 94-5, October, 1971.
5. D. Bennett and W. MacConnell. Management motivation to apply group technology. *Proc. 14th Int. M.T.D.R. Conf. Manchester, September, 1973*.
6. W. A. Knight. Economic benefits of group technology. *Machinery and Production Engineering*, 23 June, 1971.
7. F. R. E. Durie. A survey of group technology and its potential for application in the U.K., Institute of Production Engineering, G. Bray Memorial Lecture, 1969.
8. F. Ribbons and W. MacConnell. *Case Studies in Group Technology*, Iss. 3, G. T. Centre, Blacknest, September, 1972.
9. J. J. Marklew. The Cell System—A variety of applications in the same factory. *Machinery and Production Engineering*, pp. 442-7, 24 March, 1971.
10. F. Ribbons. Application of G.T. in a conventional jobbing shop, *Case Studies in Group Technology*, Iss. 3, G.T. Centre, Blacknest, September, 1972.
11. G. M. Ranson. *Group Technology*, McGraw-Hill Book Company (UK) Limited, 1972.
12. R. Connolly, G. H. Middle and R. H. Thornley. Organising the manufacturing facilities in order to obtain a short and reliable manufacturing time, *Proc. 11th Int. M.T.D.R. Conf., Birmingham, September, 1970*.
13. G. H. Middle, R. Connolly and R. H. Thornley. Organisational problems and the relevant manufacturing system, *Int. J. Prod. Res.* (1971) 9, no. 2.
14. F. A. Lewis. Some factors affecting design of production systems. *Proc. 14th Int. M.T.D.R. Conf., Manchester, 1973*.
15. J. J. Marklew. An example of the cell system of manufacture at Ferranti Machinery and Production Engineering, August, 1970.
16. A. Jones. Manufacturing cost and batch quantity equations with reference to functional and group layout, M.Sc. Dissertation, Machine Tool Engineering Dept., October, 1973.

BATCH SIZE SELECTION FOR GROUP TECHNOLOGY

by

K. RATHMILL*, P. BRUNN* and R. LEONARD*

SUMMARY

Of all the potential advantages offered by group technology (G.T.), a reduction in throughput time and a corresponding fall in work in progress are the most frequently quoted. It is commonly held that further savings of investments in stocks and improvement in response from the manufacturing system can be achieved by a decrease in batch quantities, made economic by the setting-time reductions achieved by G.T.

In practice the opinions of major companies indicate that there is no single answer to the question of whether G.T. should lead to a reduction in batch sizes. Mr Ranson, Managing Director of Serck Audco Valves Limited, speaking of his own company states: 'Piece parts will only be manufactured for an assembly programme for products, and not held in store'¹. F. W. Craven, Director of Herbert Machine Tools, Coventry, making more general comment has said: 'Batch size reduction is certainly not essential to G.T. application, but can be a further added benefit in certain circumstances'².

Clearly any apparent difference of opinion on this matter should be based upon fundamental operational differences between the companies involved.

The present work is the result of an investigation, into various types of companies operating G.T., in which factors affecting the ability of a given company to reduce its batch sizes are indicated. Initially, from an analysis of cost parameters involved, a general examination is detailed relating various types of industry to an ability to achieve economically justifiable batch quantity reductions. The distinction between a reduction in optimum batch size and a reduction in cost-sensitivity to batch size changes is made. A cost sensitivity factor, S_f , is developed to express the ability of a company to change its batch sizes without financial penalty, and the operational benefits of a reduction in S_f is stressed. The theory is then examined against a background provided by some of the major G.T. implementations in Great Britain. This confirms that reductions in batch size are not in fact a natural sequel to implementing G.T., and proceeds to indicate factors outside the general theory, which in certain cases can greatly modify the arguments for or against a reduction in batch size.

NOTATION

C_I	Investment costs of stock in raw material and finished parts stores.	S_u	Setting up time (total for all operations) in hours.
C_p	Capital price of component (£).	T_c	Total cost/component while in factory.
C_s	Storage costs/component/year.	T_{co}	Minimum cost per component.
D_L	Direct labour costs per hour.	T_h	Delay time/production in works.
I_r	Company Interest rate.	T_p	Machining time (total for all operations), hours.
K_c	$C_p \cdot I_r + C_s$ (£/hour).	Δ	A proportion of T_{co} by which component costs are increased above the minimum cost for a given batch size, e.g. for Q_x , $\Delta = 0.05$.
L_t	Lead time = $T_p + T_h$.	ϕ	A function.
M_c	Plant depreciation and investment costs.		
O_h	General Overhead rate (£/hour).		
Q	Batch size.		
Q_m	Minimum stock level.		
Q_o	Optimum batch size (minimum cost/part)		
Q_x	Batch quantity below Q_o corresponding to a component cost 5 per cent above T_{co} .		
R	Annual rate of usage of components (parts/year).		
S_f	Cost-sensitivity factor (see appendix).		
S_R	Set-up ratio = S_u/T_p .		

INTRODUCTION

The subject of determining batch sizes must be to a certain extent open ended. The number of factors involved, both in the computation of the 'ideal' and the inevitable modification because of the practical constraints that exist within each manufacturing system, are such as to allow only general observations to be made in any proposed theory. An adequate

* University of Manchester Institute of Science and Technology

theory is required, however, since G.T. is rapidly being recognized as the standard method of batch manufacture and those considering the change to G.T. should be fully aware of all its implications. The answer to the question: 'how will G.T. affect the company's work in progress and lead time situation?' can only be fully appreciated if the likely effects of G.T. can be assessed with relation to the company's pattern of batch sizes.

Current methods in use for the determination of batch quantities

In a typical engineering company batch sizes are determined to a large extent by the stock controllers 'feel for the job'; this embraces a knowledge of the manufacturing problems of a given component, its value, the production time and weekly demand. In addition the company may have a general stock control philosophy which influences individual batch quantities. Some of the philosophies found within the industry are:

Fixed stock investment

A value of stock, usually related to a period of production, is defined and batch quantities are determined for a stock investment below this limit. Under this system the financial planning implications of capital tied up in stocks are of prime importance.

Fixed capacity

After an examination of the setting and production times of the most heavily loaded machines, the number of batches within a given period is maximized to achieve the output required. This system strives to minimize stock investment on the basis of a known output requirement and production capacity.

Cyclic ordering systems

Components are ordered according to a single or a number of fixed cycles. A single-cycle system or 'period batch control' necessitates orders for all components required during the next period to be ordered simultaneously. Thus the controlling system has a discipline and simplicity which is not attainable by simple inventory control. A multi-cycle ordering system is similar to the single-cycle system except that depending upon such factors as lead time, value, and demand, different components are given varying ordering frequencies according to a standard series such as 1, 2, 4, 8, 16. The advantages of the single-cycle ordering system are tempered by the differences in the economic batch-size requirements of washer-like components when contrasted with a complex plano-milled casting.

Limited ordering facility

The number of batches the control system can handle is determined and from this the sizes of batches are obtained. This approach attempts to reduce ordering and inventory control costs.

Fixed batch time

Where batch sizes are such that batch throughput times are excessive, the batch size is reduced in order to optimize the operator's job interest. Short batch

times result in an unacceptable increase in setting costs. Large batch times can produce frustration and loss of performance. This approach is based upon the human aspect and aids labour relations.

Economic batch size theory

The economic batch quantity (E.B.Q.) approach is widely used by British industry. E.B.Q. theory defines a mathematical model involving the product cost parameters. Expressions are derived to give batch quantities for the minimum component cost or the maximum return on investment, etc. The complexity of available analyses varies considerably from the simple form of equation quoted by Gilbert³ to the sophisticated analysis of Jones⁴.

Due to the arithmetic effort involved the simple formulae are most commonly used in industry. The E.B.Q. approach attempts to base the decision making associated with batch quantity determination on an analysis of the major economic factors involved. E.B.Q. theory has a wide appeal and is quoted widely in management texts, however, its inherent limitation is that it is severely restricted in the number of used parameters⁵. Burbidge⁶ states that the relative economies of investment in stock and alternative modes of investment must be considered, and that the most 'economic' form of material flow is by no means certain to result from a multi-cycle ordering system where each item is made to a different batch frequency. Simplifications and approximations which must be made in an E.B.Q. formula suitable for a specific application can be seen as a potential danger to the accuracy of such a formula in a different situation.

The E.B.Q. approach has limitations. However, it provides a method of assessing the economies of batch quantities selected by other philosophies, when it is used in a suitable form and when due regard is paid to its application.

Cost parameters related to batch quantity

An analysis of E.B.Q. formulae highlights a group of major cost parameters which form the basis of a function which indicates the optimum (for minimum cost) batch.

From the group of parameters it can be observed that some are inversely related to Q_o while others are directly related. Assuming that the annual rate of usage, R , remains constant the relationship may be summarized thus

i.e.

$$Q_o = \phi \left[\frac{O_h, S_u, D_L, M_c}{T_p, C_s, C_i, T_h} \right] \quad (1)$$

An increase in general overhead costs (O_h) tends to increase Q_o whereas an accompanying increase in storage costs (C_s) could counteract this. Therefore the situation can occur, see figure 1, where the introduction of G.T. effects a decrease in various cost parameters yet leaves Q_o unchanged. When considering component cost the introduction of G.T. can have a dramatic effect. The component cost equation (see Equation (2) Appendix), given by Hagan and Leonard⁷, indicates that a reduction in

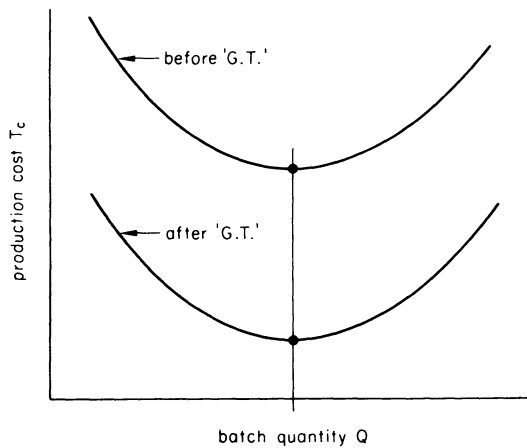


Figure 1. Reduced production costs with no change in optimum batch size.

any of the cost parameters results in a reduction of total component cost.

Certain costs may be increased by introducing G.T.⁸ but in general the cost parameters involved in the determination of Q_o and T_c are reduced. Hence a company experiencing a proportionately greater reduction in costs, T_p , C_s , C_l , and T_h , would be able to further reduce costs by increasing batch quantities. Similarly a greater reduction of O_h , S_u , D_L and M_c by G.T. would allow costs to be reduced by smaller batch sizes. In both these cases (figure 2) the cost is reduced, but the changes in optimum batch size are opposite in nature.

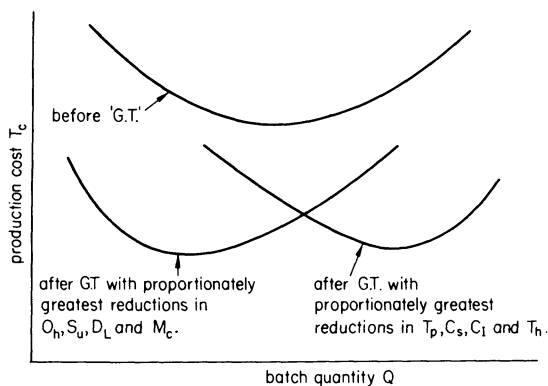


Figure 2. Increased and decreased optimum batch sizes as a result of introducing G.T.

Thus the effect of a change in batch sizes resulting from G.T. depends upon the proportional changes between the cost parameters involved and not the changes in individual values.

Batch size reduction as an objective

Because of obsolescence, lead time, storage, and work-load balancing, most companies prefer to operate on the minimum batch quantity consistent with economic production. If batch-size reduction is taken as a criterion the company that achieves a proportionately greater reduction in the cost parameters O_h , S_u , D_L and M_c will achieve the most successful G.T. implementation. Thus it is necessary to determine the relationships that exist between company features and the ability to benefit from a reduction in batch size.

General overhead costs (O_h)

General overhead embraces a multitude of indirect costs including management, research, development, design, quality control, indirect labour, maintenance, production engineering, buying, costing, production control, sales, marketing, storage, electrical power, heating, and rates.

Reductions in overhead costs are likely to result from an improved operating efficiency. Group technology has been described by Ranson¹ as 'a foundation for better total company operation' and by Thornley⁹ as 'a complete manufacturing system' and as such G.T. is widely accepted as being of greatest value when it is introduced in the most comprehensive manner involving change in all aspects of company operations. Such an implementation would be the result of a 'total approach' to G.T.

A distinction must be made between a company which has observed a family resemblance in a section of their component range and set up a corresponding machine group, and an enterprise that has carried out a complete analysis of the geometric, dimensional and production features of its entire component and assembly requirements, from which has evolved an integrated G.T. manufacturing system. Generally the company that can embrace G.T. in the most complete sense is the one which will be most able to reduce its batch sizes.

Set-up costs (S_u)

Initially G.T. was closely identified with a 'complex component' approach. The extremely high degree of component similarity implied by this approach leads to large savings in setting and machining costs due to improved production methods and tooling specialization. Although component families having a high degree of similarity are important in G.T. they are only possible where the scale of company operations provide the necessary demand for an economic utilization of machine groups. More generally the variation of component geometry and dimensions within the family is such as to require set-up costs consistent with those incurred prior to the introduction of G.T.

In general the form, size and operation of G.T. cells depends upon the intrinsic ability of the component range to form families. Conversely the machine group design depends upon the component family definition, hence different approaches towards component family definition have a major effect upon the eventual machine groups and particularly the extent to which set-up costs can be reduced. The aim should be to obtain a method of component family/machine group definition which integrates an analysis of the production requirements with consideration of component design features.

Thus the benefits resulting from a reduction in set-up and tool changing time depend upon many factors, two of which are:

- (a) the degree of component similarity exhibited. This depends upon the product and the family resemblance exhibited within the product range, and

- (b) The scale and consistency of demand for the components within the families. Where demand involves an inconsistent product mix, the degree of family resemblance is of importance when considering the degree of component specialization, or inflexibility, in the machine group design.

Hence a company manufacturing a relatively simple product, with a high degree of family resemblance within its component ranges and operating in a market in which demand is fairly stable, should benefit from production specialization under G.T. Hence by attaining a reduction in set-up time the company can achieve a reduction in the minimum cost batch quantity.

Direct labour (D_L)

Often by introducing G.T. higher wages are possible for direct labour¹ as a result of increased efficiency. The extent of improvements in direct labour largely depend upon the product. For example where components are simple, requiring little manufacturing sophistication, and where family resemblance allows savings from specialized tooling set-ups, reductions in direct labour costs are probable. Craven⁸ states that certain machine tools, due to the degree of freedom intrinsic to their mode of operation and control, are more able to take advantage of component family resemblance than others. N.C. machines fully utilized in G.T. cells reduce machining time and direct labour costs.

Another factor for consideration is the skill required in the manufacturing processes. Semi-skilled labour should lead to increases in direct labour utilization because of the relative ease with which flexibility of labour can be achieved.

Machine depreciation and investment costs (M_c)

New plant must be considered primarily on a basis of component cost. The increased machine-hour rate may mean that set-up costs are prohibitive for small batches unless setting times are reduced. With good utilization high-cost plant may be justified on a component cost basis. Figure 3 depicts a high cost sensitivity to batch sizes below Q_0 exhibited by a company employing high cost plant. Thus the batch

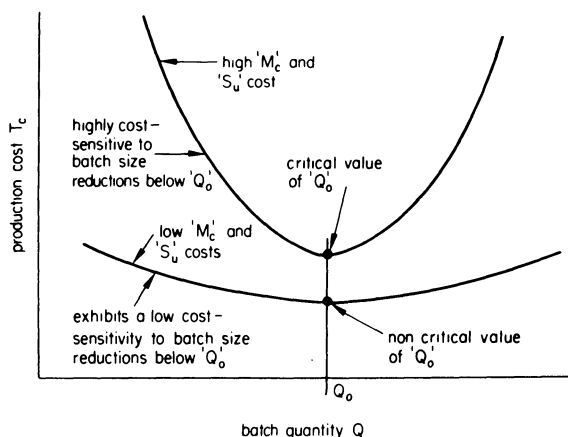


Figure 3. Cost-sensitivity variation for the same optimum batch size.

sizes will be similar to those prior to G.T. In contrast the company retaining low-cost plant after the implementation of G.T. may find that due to the horizontal nature of the batch/cost curves, figure 3, it is expedient to reduce batch sizes well below Q_0 in order to achieve the advantages of reductions in lead time and work in progress. Such considerations necessitate an examination of the form of different companies' batch quantity cost curves, particularly the degree to which Q_0 is seen to be critical from a cost standpoint.

Cost-sensitivity to variations in Q

Figure 3 demonstrates that companies with different values for parameters O_h , S_u , D_L , L_c , may have the same optimum batch sizes but very different batch quantity cost curves. For a company with high machine set-up and storage costs the batch size can be critical, and small variations about Q_0 lead to significant increases above the minimum cost. Thus the company is highly cost-sensitive to changes in batch quantity (Q). A company with a low set-up and storage cost is likely to exhibit a low cost-sensitivity to changes in batch size.

A dimensionless factor (S_f) is useful for quantifying cost sensitivity. The T_c - Q curve is asymmetrical about Q_0 , and since interest is generally shown in batch sizes below Q_0 , S_f (see Appendix) reflects the maximum reduction in batch quantity relative to the optimum (Q_0) which can be permitted before production costs (T_c) exceed 5 per cent of the minimum value. As the need to operate at Q_0 becomes more critical the value of S_f increases.

A reduction in cost-sensitivity to batch-size variation achieved by G.T. is valuable in its own right since the batch sizes used in practice will vary considerably due to practical constraints within the system. A reduction in S_f can form the basis for a reduction in batch costs, in addition it may justify a review of manufacturing policy with the aim of decreasing batch sizes to achieve considerable improvements in manufacturing control.

An examination of the major parameters involved (Equation (2)) suggests that for large batches; machining time (T_p), capital price of the component (C_p), company interest rate (I_r), and the storage costs (C_s) are most influential with regard to the total manufacturing cost. Therefore a company manufacturing low-value components, with high W.I.P. and storage costs, and whose component demand is such that group manufacture could achieve a reduction in total machining time, is likely to benefit from a reduction in cost-sensitivity to batch size increments above Q_0 . For small batches the most influential parameters are set-up time (S_u), general overhead (C_h), direct labour costs (D_L), and machine depreciation costs (M_c). Hence a company with systems problems, disproportionately high overhead costs, poor resource and labour utilization, and whose component demand indicates good family definition, is most likely to achieve a reduction in cost-sensitivity (S_f) to batch sizes below Q_0 . The company most able to benefit from a reduction in S_f and Q_0 is most likely to be the company which adopts a 'total approach' to G.T.

Industrial cases examined

Using company data in Equation 2, component costs can be calculated for various values of batch quantity. The results can be compared against the practical background of the company and the effect of G.T. on the batch sizes within the company considered. In most companies widely varying components exist having different batch-size requirements. Each company has its own market, product and manufacturing constraints. However, a survey of companies operating G.T. produced characteristic cases. Four such cases will now be detailed.

Example 1

The company manufactures a product of medium complexity, the average component values (C_p) are of the order of £35. The product range exhibits family resemblance, but size variations and design variants reduce component similarity. Demand is typically under 100 orders of any one model per year. A total approach to G.T. has been pursued. However, certain systems, particularly production control, require attention. The company has ceased manufacturing for stock since G.T. Setting times are of the order of two or three times the machining time for one component.

G.T. has had a major effect by greatly reducing lead time, typically from thirty to seventeen days for the complete manufacture of finished assemblies. A major factor in this reduction is the inclusion of assembly operations in the G.T. scheme.

The low volume production of any one component and the low family resemblance of components necessitated flexibility in machine-group design. Machining times have been reduced by 20 per cent in some cases, but setting times have remained unaltered.

Batch sizes have not been altered since G.T. and this policy is supported by the batch quantity cost results of figure 4. Since the major influence on Q_o in

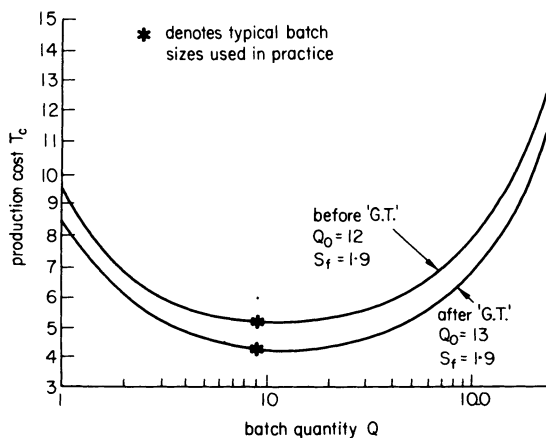


Figure 4. 'G.T.' reducing production costs with negligible change in Batch size constraints.

this case is the change in machining time (T_p), a slight increase in Q_o occurs. However, a large reduction in lead time and reduced machining costs produce a steady reduction in costs over the normal batch quantity range. Since the batch quantity cost curve retains its initial form, the cost-sensitivity (S_f) remains unaltered. When examined from the standpoint

of S_f the company is seen to exhibit ($S_f = 1.9$) a restrictively high sensitivity to batch-size changes. Thus batch sizes should not be significantly altered since they are within 5 per cent of the optimum.

Example 2

The company has a relatively simple product with few component parts. A high product and component family resemblance exist. A high demand situation for most models exists and the company may be defined as a large-batch manufacturing organization. Major component batch sizes of the order of 500 are usual. Component cost values are low, being typically of the order of £2.

A significant proportion of production has a low skill requirement and the company employs a large number of semiskilled personnel. Set-up ratios (S_R) tended to be in the region of 8 : 1 and the machine and plant investment costs are low.

A total approach to G.T. has produced large reductions in setting times, lead times and W.I.P. The major benefits emanate from (a) its ability to determine definitive component families, (b) its ability to justify and use flow-line machine groups, and (c) to improve labour utilization by increased mobility and flexibility.

Overhead costs have been reduced by improvements in company systems, material flow, and floor-space utilization. Inexpensive plant has concentrated improvements in resources utilization, on flexibility of direct and indirect labour, and improvements in ordering and production control.

To maximize the advantage of the reduced lead times obtained with G.T. a single-cycle ordering system operating on an extremely short production period exists. Manufacturing according to a short-term production programme has reduced batch sizes.

The company has reduced the cost parameters O_h , S_u and D_L , these being directly related to Q_o (Equation (1)). Figure 5 shows that the introduction

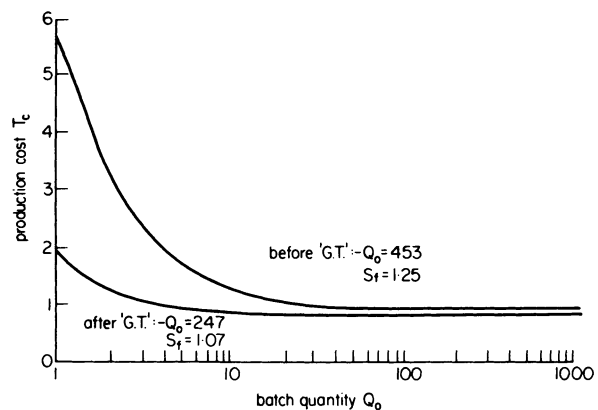


Figure 5. G.T. effecting a major reduction in cost-sensitivity S_f , and optimum batch size Q_o .

of G.T. has reduced the optimum batch size to some 55 per cent of its previous value. Of greater significance, however, is the fact that as a result of the reductions in O_h , S_u and D_L the company has achieved a very significant reduction in sensitivity factor S_f .

The present low value of S_f demonstrates an ability to produce batch sizes from 10 to 1000 without significant increases in cost. This is important because an overriding concern has been to create a highly responsive manufacturing system to enhance the competitive ability of the sales and marketing functions.

Example 3

The firm manufactures a very restricted range of components. Whilst considerable variations in both size and form exist within the component range, the volume and consistency of demand has allowed the formation of certain specialized G.T. cells. The average capital value (C_p) of the components is typically £4. Lead times, prior to G.T., were of the order of eight weeks, and set-up ratios (S_R) around 20 : 1. The high degree of component similarity has enabled setting time to be reduced by 40 per cent, and group manufacture has halved lead times. A 'total approach' to G.T. with production control at group level has reduced problems and costs generally. Although some aspects of overhead cost have been reduced by G.T. the overhead figure is similar to that existing before implementation because of increased investment in sophisticated plant and supporting personnel.

Batch sizes have not changed since G.T., and figure 6 shows that the company remains fairly close to the

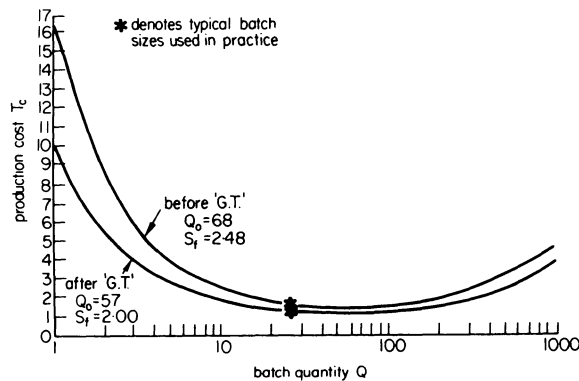


Figure 6. G.T. effecting a minor reduction in optimum batch size, Q_o , and cost-sensitivity, S_f .

optimum batch size, Q_o . The major effects from G.T. have been a reduced lead time (L_t) and set-up time (S_u), S_u influences the value of Q_o and the 'before' and 'after G.T.' curves in figure 6 indicate a small reduction in Q_o .

As seen in Example 2 (figure 5) the reduction in Q_o is not of importance, the reduction in S_f , however, is of more benefit to the company. As a result of G.T. the company has achieved a reduction in overall production costs and widened the range of batch sizes it can manufacture economically.

Example 4

This company is outside general engineering. Research, development, and the short run nature of the product impose a high overhead burden on production. Prior to G.T. the company suffered from extended lead times, and set-up ratios (S_R) were of

the order of 10 : 1. The product range is very loosely defined, with highly complex short-run quantity products being produced with few dependable repeat orders. The inconsistency of demand at product level has necessitated a flexible approach, and this has resulted in generalized machine groups with negligible savings in setting times. The major advantage from G.T. is a 50 per cent reduction in lead time (L_t). The components are generally of high cost and intricacy, requiring costly machining capacity including a high proportion of N.C. machines.

Batch sizes have not changed since the introduction of G.T. Figure 7 illustrates that the reduction

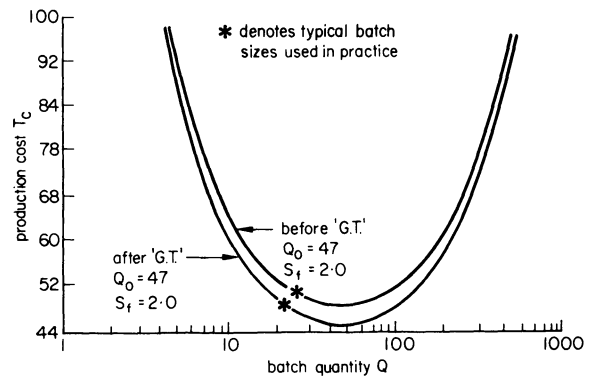


Figure 7. G.T. reducing production cost with no change in batch-size constraints.

in lead time has reduced the total production costs (T_c) by 7 per cent but the optimum batch quantity (Q_o) has not changed. The high and generally unaffected machine (M_c), overhead (O_h), and set-up (S_u) costs generate a high-cost sensitivity factor (S_f) which has not been changed by G.T. The company is therefore justified in retaining its batch sizes. The company tends to choose batch sizes at the lower limit of economic production; this preference for minimum batch quantities stems from sound economic considerations. Due to the high cost of the complete product assembly, reliable lead times are imperative since late components result in investment losses on the whole assembly. Also because sophistication and innovation are major constituents of the products' value, obsolescence and modifications make short manufacturing times additionally important. Reducing batch sizes well below the 'optimum' (Q_o) is justified in cases where delivery dates are critical or where the threat of obsolescence is particularly strong.

Other relevant considerations

In any simple treatment of a complex problem it is prudent to acknowledge the importance of factors not included in the analysis. The manufacturing environment in most batch-producing companies is dynamic, and the factors affecting optimum batch-size determination will vary with the changes in demand, component type, and capacity. However, factors which are difficult to quantify in a rigorous treatment of batch-size determination can be observed in a general sense and their modifying effect upon Q_o noted.

Capacity—utilization

The determination of Q_o by a simple analysis⁷ assumes that utilization of capacity is not affected by changes in batch size. Where a down-turn in trade results in excess capacity, personnel problems associated with a short-term excess of capacity could be eased by reduced batch sizes. Alternatively if component values and obsolescence risks are low, and where standardization is high, an excess of capacity may justify large batches of components being manufactured for stock.

Ordering system constraints

Where ordering costs are high a reduction in orders may be achieved by an overall increase of batch sizes.

Manufacturing system response

A company operating G.T. should benefit from short lead times. However, where delivery response is paramount, further batch size reductions may be justifiable.

Maximizing job interest

An important aspect of the manufacturing situation is job interest and satisfaction for the operator. Where batch sizes give rise to monotony, their reduction should be considered. Alternatively where total batch times are low they may be increased in order to make the work load more stable.

Although the personnel considerations of batch-size determination only apply to a section of industry, where demand is stable and set-up costs are low, the problems of attracting and maintaining a competent labour force can make such considerations worthwhile.

Obsolescence

In certain industries where total manufacturing times are long and the rate of technological development is high, long lead times and large batches can result in expensive products being obsolete before completion. Thus the penalties of obsolescence can offset the higher cost of small batches.

Total investment considerations

Where the product is an assembly of high capital value, or penalty clauses exist, short and reliable lead times are paramount. In such cases the importance of minimum-cost batches are lessened and batch sizes well below the optimum may be justified.

CONCLUSIONS

Group technology achieves a high level of output and system efficiency for batch production, and can reduce most of the cost parameters concerned with batch sizes. However, the effect of a change in batch sizes as a result of introducing G.T. depends upon the reductions in individual costs and the proportional changes between the cost parameters involved. Batch size reduction is not an automatic sequel to the introduction of G.T., and in certain circumstances batch sizes can increase.

It is often desirable to manufacture the smallest batches that are consistent with economic pro-

duction. Companies which by virtue of their product range, product type, demand situation, manufacturing and labour requirement characteristics are able, through G.T., to achieve savings in overheads, setting-up, and direct labour costs, will be able to reduce batch sizes.

Companies having the same optimum batch sizes can have markedly different cost-sensitivities to changes in batch size. This indicates an important difference between companies and their abilities to reduce batch sizes. A company which attains its greatest benefits from reductions in set-up time, general overheads, direct labour, and plant costs, will achieve a reduction in its cost sensitivity (S_f) to batch-size variations below Q_o .

Hence conditions conducive to reductions in optimum batch sizes are identical to those required for a reduction in S_f .

The 'total approach' to G.T. is seen to offer great potential for reductions in,

- (a) component cost,
- (b) optimum batch size,
- (c) cost sensitivity to batch size changes below Q_o .

In the majority of cases only marginal changes in Q_o may result; however, the change in S_f is likely to be more marked, and it is this reduction in cost sensitivity to batch-size variations that is of most value. In particular a general reduction in cost sensitivity is highly compatible with the operation of a period batch-control system since a lack of sensitivity breaks down the division between conflicting component requirements which may otherwise necessitate the use of stock control, or hybrid ordering systems.

ACKNOWLEDGMENTS

The authors would like to express their appreciation for the co-operation given by the following companies during the course of this work: Alfred Herbert Limited, Coventry; G.E.C. Elliot Fisher Valves Limited, Rochford, Kent; Ferranti Limited, Edinburgh; Serck-Audco Valves Limited, Shropshire; Ferodo Limited, Chapel en le Frith; Mather and Platt Limited, Manchester; Platt International Limited, Bolton; Hopkinsons Valves Limited, Huddersfield.

APPENDIX

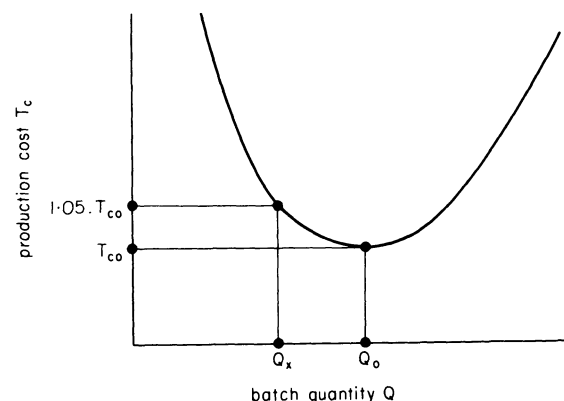


Figure 8.

From figure 8, the cost sensitivity factor $S_f = Q_o / (Q_o - Q_x)$.

It has been shown⁷ that the total manufacturing cost per part can be given as

$$T_c = \left(\frac{K_c Q_m}{R} + K_c T_h + T_p O_h \right) + Q K_c \left(\frac{1}{2R} + T_p \right) + \frac{S_u O_h}{Q} \quad (2)$$

from which the optimum batch size can be determined

$$Q_o = \sqrt{\left[\frac{2RS_u O_h}{K_c(1 + 2RT_p)} \right]} \quad (3)$$

It has also been shown that the batch size below Q_o corresponding to a 5 per cent increase over minimum cost T_{co} can be obtained from

$$Q_x = \left[Q_o + \Delta \left(\frac{2Q_o + Q_o^2 K}{2} \right) \right] - \frac{Q_o}{2} \sqrt{[(Q_o K \Delta^2 + 4\Delta)(Q_o K + 4) - 8\Delta]} \quad (4)$$

where

$$K = \frac{T_p O_h}{S_u O_h} + \frac{K_c T_h}{S_u O_h} + \frac{K_c Q_m}{2RS_u O_h}$$

S_f can therefore be calculated using Equations (3) and (4) or, can be determined graphically using Equation (2).

REFERENCES

1. G. M. Ranson. *Group Technology*, McGraw-Hill, 1972.
2. F. W. Craven. Discussion following 'Strategies for increasing the utilization and output of machine tools at the 14th Int. M.T.D.R. Conference, U.M.I.S.T., 1973.
3. R. Gilbert. Analogical Reasoning Applied to Product Rationalization and Manufacturing Systems. *Proc. Int. Seminar on Group Technology, I.L.O. Turin*. 1969.
4. A. P. Jones. Manufacturing Cost and Batch Quantity Equations, with reference to functional and group layout. M.Sc. Dissertation, Mechanical Engineering Department U.M.I.S.T., 1973.
5. S. Eilon. *Elements of Production Planning and Control*. Macmillan, New York. 1962.
6. J. L. Burbidge. *The principles of production control*. MacDonald and Evans, London. 1971.
7. P. C. Hagan and R. Leonard. Strategies for increasing the utilization and output of machine tools. *Proc. 14th Int. M.T.D.R. Conf., U.M.I.S.T., 1973*.
8. F. W. Craven. Some constraints, fallacies and solutions in Group Technology Applications. *Proc. 14th Int. M.T.D.R. Conf., U.M.I.S.T., 1973*.
9. R. H. Thornley. Group Technology—A complete manufacturing system, Inaugural Lecture, University of Aston in Birmingham, 14 October 1971.

THE INTERACTION OF NC AND GT

by

F. W. CRAVEN*

SUMMARY

This paper, having explained the background, leads on to a discussion on the general question of NC/GT. Because set-up problems receive considerable attention, particularly in the turning field, the effects of NC and GT on these problems are examined in some detail. Following a more general description of machine characteristics and capabilities, the interaction of machining centres and NC lathes with GT is analysed, leading to the conclusion that the use of NC in GT realizes additional benefits from both.

INTRODUCTION

The purpose of this paper is to stimulate discussion and argument about the different areas of interaction between NC and GT. As the subject is new, complex and a matter of continual development, it would be imprudent of the author to attempt to define specific rules.

The paper has been prepared with the help of a number of groups and individuals, all of whom have particular involvement in the fields of NC and GT. They include the HMT Systems Engineering Department, NC Product Managers and Specialists and the GT In-house and External Survey Teams.

We are fortunate in having in our own company a large, multi-product GT application; externally we continue to deal with many GT surveys in manufacturing organizations which frequently involve the use of NC machine tools of various types and makes. This practical, real-life background has involved us with a variety of applications of NC in cells which has been of value, not only to our GT Specialist Teams, but also to our NC Systems Engineering Department who are able to develop their philosophies against various scales of multi-product and single-product GT applications. Of the forty-seven machining cells planned for installation on three separate manufacturing sites of the writer's company, twenty are already commissioned.

1 GENERAL COMMENTS ON NC/GT

1.1 The interaction of NC and GT

There are many advantages claimed for NC machine tools; some of these are of greater importance in a GT environment. It is not the purpose of this paper to

discuss all the pros and cons of NC, but some of the advantages need to be identified. Numerical control of machine tools offers the ability to machine small batches of components economically by reducing set-up time and by enhancing the effectiveness of skilled manpower, a commodity which continues to become more scarce. NC machines are invariably more expensive than conventional machines of similar capacity, and opportunities should be sought for maximizing the return on these higher investments.

Arguments will be developed concerning not only the use of NC machine tools in a GT environment, but also the question of their mutual interdependence. It will be suggested that in some cases NC and GT will not only be mutually viable, but that together they can be self-improving and self-sustaining; the implied corollary is that if they are used independently, maximum benefits may not be realized. Clearly, we will not expect to find that NC and GT are wholly symbiotic. We have to try to establish the extent of their inter-relationship. As in the theory of sets, in order to determine the degree of parity between two functions it is necessary to define substantially more than the over-lapping common qualities. We have first to take a proper look at both GT and NC. At a later stage, with the ground clear, an attempt will be made to determine their inter-relationship in some detail. First, we must draw attention to one of the more important aspects of NC machine-tool operation—that of work in progress.

1.2 Work in progress

One of the advantages frequently claimed for NC machine tools is their ability to reduce work in progress. This statement is often true but its validity

* Herbert Machine Tools Limited, Edgwick Works, Coventry

is dependent upon the position of the NC machine in the machining cycle. In the past, the majority of NC machines in general engineering have been drilling machines, and as drilling operations are frequently towards the end of the component cycle, such machines have been truly effective in reducing work in progress. As the NC lathe is generally at the start of the machining cycle, the reverse can be true; NC lathes, properly used, are most effective generators of work in progress, and unless the subsequent operations are adequately dealt with, work in progress will rise. The NC lathe, with its many other advantages, makes an ideal marriage with cellular manufacture, and within such an environment the high and efficient output from this type of machine can be more quickly turned into saleable goods. A similar argument can be developed for the application of machining centres. The position of the machining centre in the machining cycle is often at a different point from that of the NC lathe; the NC lathe is usually the initial pacemaker whereas the machining centre is generally part way through the process flow. The effectiveness of machining centres depends, among other things, on the machine being properly fed and the subsequent operations being speedily dealt with. These requirements are readily met by placing the machining centre in a cell, with feeding and off-loading machines dedicated to the support of this principal investment.

We now proceed to deal with one of the most important matters in GT, that is the question of set-ups.

1.3 Set-ups

(a) *Natural families*

Many articles on GT suggest that cellular manufacture automatically leads to a reduction in set-up costs. It appears that the question of set-up costs in cells is frequently misunderstood. In many cases, the grouping of components into families will not in itself reduce overall setting time but may increase it. In functional layouts there are many opportunities for grouping similar set-ups which can lead to actual setting times in such layouts being less than in cellular layouts, due to easier matching of component requirements to machine facility. It is probably true to say that in the case of products which fall into natural families, there may be an automatic reduction in setting cost when moving from a functional to a cellular layout. Much of the literature deals with examples of this type. In the general engineering industry where there is a wide variety of component types, unless adequate production engineering work is carried out (and sometimes even if this is done) actual set-up times may increase, particularly on turning operations. It is unfortunate that the techniques and benefits from single-product natural family solutions are often extrapolated into multi-product situations, and fallacious conclusions are drawn. In general engineering applications of GT there can be a conflict between cell size and set-up optimization which is not so significant in single product applications.

(b) *Influence of machine type on nature of set-ups*

There is considerable difference in the areas of work holding, tooling and setting on turning machines, especially turret lathes, when compared with drilling machines and milling machines. Set-up variations are more easily condensed and higher levels of flexibility of set-up can be developed on drilling and milling machines. There are fundamental differences between turning operations and milling and drilling operations, particularly because of the effect of the three main elements of set-up, namely:

- (1) holding the workpiece,
- (2) tool location/size/capability,
- (3) relative movements in several planes of work-piece/tool.

Consideration of these features leads to the conclusion that set-ups can be more simply standardized with greater flexibility of application on milling and drilling operations rather than on turning with conventional lathes; this is certainly found to be normally true in actual practice.

(c) *Turning set-ups*

In most general engineering situations, product geometry and size variation within families can still display considerable variety unless cells are made so small as to be virtually unworkable. The restrictions which limit the lower level of cell size are primarily load balancing problems, the effects of absenteeism, unnecessary high flexibility of labour and unacceptably low utilization of second-operation machines.

Many statements on GT applications in the turning field refer to the composite component or some similarly termed device, with the implication that the use of such techniques automatically substantially reduces or eliminates alterations to set-ups. Whereas this may be true to some degree in certain cross-sliding lathe applications, it is certainly questionable on turret lathes and capstan lathes. It seems hardly necessary to state that different sizes and settings of roller boxes are needed to turn different finished diameters and that a whole range of end-working tools are required to produce a variety of different-sized holes. Turret lathes have to be set for turning to specific diameters and lengths, and unless the features and dimensions of consecutive components are identical and not just approximately similar, many alterations in tooling and stop setting have to take place.

In some cases the effect of component size is of importance, and with components of similar geometry it is naturally found that the ratio of set-up to operation time is generally considerably more adverse for components of smaller size.

The development of consolidated set-ups should be concentrated towards producing set-ups which are inexpensive to operate and which are either comprehensive, that is capable of producing a variety of components with no tool change (a limited solution), or flexible, that is able to produce a variety of components with minimum tool change. It is possible, in a cell environment, to engineer such set-ups, since the work is presented to the Production

Engineering Department in more manageable parcels. The opportunity is afforded to define preferred set-ups and preferred planning methods which, if rigorously applied, can substantially reduce on-going planning costs. Using standardized set-ups one can afford to have a higher provisioning cost to reduce set-up time, because overall there can be a worthwhile reduction in the number of such set-ups. However, the turning set-up problem in cellular manufacture can be difficult and if, for other reasons, batch sizes are reduced, the problem is exacerbated. This problem is completely resolved by the use of the NC lathe, provided batch sizes are appropriate. It will be dealt with in more detail later in the paper.

(d) Set-up sequencing

There is an opportunity for making considerable savings in set-up time by sequencing the flow of work of a machine so as to minimize the changes to the set-up. This is a complex problem, particularly in a multi-product environment, and it is difficult to develop a fully optimized solution because of the interaction and conflict of the various sequencing criteria. The problem is eased by the adoption of period batch control combined with cellular manufacture. Using this technique the component mix is the same every period, so that once the best sequence of components has been developed it will remain substantially unchanged unless there is a significant alteration to the product mix. With conventional stock control systems the component mix changes continually from week to week and the optimizing of set-up sequences is extremely difficult.

NC machines enable solutions to this type of problem to be obtained much more easily because generally the tooling requirements are similar for a wide range of components. The advantages gained are greater with NC lathes because at the worst only special-purpose tools for grooving or threading plus some end-working tools need to be changed for consecutive components within a GT family; setting is thus reduced to very low levels without the need for an elaborate sequencing system.

It is of great importance in a production-limited operation to achieve minimum total set-up costs because excess costs of this type will, in the extreme, have to be converted into additional capital investment. In conclusion, bearing in mind that batch sizes are often influenced by existing machine facilities, the efficiencies obtained from reduced set-up costs can give the opportunity for further reduction in work in progress by batch size reduction.

(e) Set-up summary

The work we have carried out in the GT set-up field shows quite clearly that the set-up problem is most difficult in the turning field when using conventional lathes. Good production engineering can, in many cases, alleviate this problem. On the other hand, if there is a total advantage throughout the cell in reducing batch sizes, the deterioration of the operation: set-up time ratio on turning machines may counteract this benefit.

The all-embracing solution to this problem is the NC lathe which will be dealt with in more detail in a later section.

2 MACHINING CENTRES

2.1 General comments

NC machining centres, by virtue of their capabilities, consolidate a number of set-ups into one. They eliminate queuing time (within the consolidated operations) and cellular layout of the supporting machines has a similar effect. We can regard machining centres as sub-cells with the capability of reducing lead time and work in progress. Such machining centres are high capital-cost machines, and their use can be optimized in cellular environments because they can be continually fed with properly prepared workpieces, and the subsequent operations can be completed speedily.

2.2 Machining centre types

Many machining centres, by virtue of their capabilities, are used solely for carrying out a wide variety of operations on a single complex component.

We should distinguish not only between single and multi-component applications but also between general-purpose machining centres and special-purpose machining centres with which we are not concerned. The latter type of machine is designed and dedicated towards a single type of complex component or a few types of complex components such as low/medium volume diesel engine blocks or crank cases.

The choice of controls available, both in capability and in the number of axes controlled, is wide and the selection of which control to use has a considerable effect on the set-up and programming requirements.

2.3 Machining centres/GT—tooling requirements

In many applications of machining centres the tool magazines are found to have inadequate capacity even for single components; furthermore even with larger sizes of magazines (which are expensive) many tools have to be changed in the magazine to suit the requirements of different components. With a machining centre in a cellular layout, having chosen the component family, it is simpler to carry out value engineering not only to reduce direct manufacturing cost but also to minimize tool variety. The limited variety of sizing tools required for a chosen family may lead to the tool magazine having adequate capacity not only for a single component but also for a whole family or sub-family. Ultimately this may lead to machining centres requiring smaller magazines in cellular layouts than similar machines in functional layouts.

An example of the benefits stemming from this approach is the top slide cell in the Edgwick (Coventry) Works of HMT. This cell has a DeVlieg NC borer and a Marwin Max-E-Mill machining centre as pacemakers. Prior to implementing this cell a value engineering analysis was carried out which showed that minor redesign of the components could bring considerable savings in tooling, and reduce setting and machining times. The Max-E-Mill has a forty-tool magazine and twin spindles which limits the number of different tools available to twenty per component. By adopting two preferred sizes of tapped holes, standardizing drain hole sizes, corner radii and so on, all top slides are completed with twenty or less tools.

The average reduction in the number of tools required is 65 per cent and the reduction in floor-to-floor times, solely due to less tool changing, is 10 per cent. There is a further saving in tooling costs. For example, prior to the analysis there were eleven different tee slots requiring twenty-six cutters; there is now one size of tee slot. This improvement stems directly from the determination of the family viewed against the capability of the manufacturing equipment.

2.4 Machining centres/GT—set-ups

The application of a machining centre to a family of components gives the opportunity for more attention to be paid to the improvement of set-ups so as to lower the total provisioning costs, to reduce the application time and to increase the flexibility. Such work carried out over a clearly defined family enables set-ups to be structured in a rational manner. Because the family is relatively restricted the outcome may usefully be a smaller number of basic set-ups which are flexible or comprehensive and cheaper to use.

2.5 Machining centres/GT—operator involvement

In a typical NC machining centre application where the tape may run for an hour or more, considerable time savings are possible with a manual data input (MDI) facility both during tape proving and for optimizing cutting conditions during the cycle. With the very low batch quantities often dealt with by this type of machine, the saving in tape-proving time can be substantial. On a machine fitted with contouring facilities, the benefits are even greater. MDI can be used to compensate for the use of undersized cutters during orbital milling, for example. A further use of MDI could be in making more powerful canned cycles. These are but a few examples where the operator, who is usually highly skilled, can contribute effectively towards the efficiency of the machine—something he cannot do without such a control system. We believe that NC systems ought to have the facility of involving the operator in the machining process, both for sociological reasons and to improve efficiency.

2.6 Machining centres/GT—summary

In many cases where the purchase of a machining centre is being considered, the maximum benefits will only be obtained by due attention being paid to the whole machining complex which supports and off-loads the machining centre, by determining the precise family range and by limiting variations of sizing tools. In other words, machining centres will generally benefit from being cell pacemakers. The role of the operator must also be considered carefully in order to realize the immense potential of the NC machining centre working in a GT environment.

3 NC LATHES

3.1 General comments

If this paper had been written a few years ago, the discussion on the interaction of NC lathes and GT would have been more simple because at that time the differences between various makes of NC lathes lay largely in their mechanical construction. For example, there was a wide variety of configurations

of tool changer, magazines, turrets and so on. The control systems were relatively simple, and in the early days many were only point to point. This meant that the differences in actual application of many of these earlier NC lathes were not very significant. Today, however, matters are quite different; NC lathes have developed in various ways and are much more capable machines. Nowadays there are greater variations in the application aspects of machines from different suppliers. Many of these options stem from the more recent developments made to the control systems, and indeed some of these improvements now bring greater relevance to the question of interaction between GT and NC turning.

Let us first identify a few of the key features of modern NC lathes:

- (a) High power permits heavy roughing cuts.
- (b) High spindle speeds allow planners to exploit modern tool materials.
- (c) Wide speed ranges enable optimum cutting speeds to be used throughout the machining cycle.

These are general advantages. In the next section we will consider more specific variations between different types of NC lathes.

3.2 NC lathe types and capabilities

Reference has been made earlier in this paper to the use of read-out and its particular advantage on cross-sliding machines. The natural development from the cross-slider fitted with read-out, as far as motions and geometrical capabilities are concerned, is the NC lathe. This is because the geometry and dimensions of the components are generated by tape command using a limited range of standard tools. It is a matter of continuing surprise that whereas most medium/high volume cam-operated machines or PSC machines have facilities for two slides cutting simultaneously, the majority of NC lathes are two-axis machines. The versatility of the basic features of NC lathes, including rigidity, speed range, feed rates, power and so on, lead to considerable flexibility in cutting conditions, and the opportunity of optimizing these conditions over a wide range of materials and diameters. It is curious that the historical and successful use of two slides cutting simultaneously on the machines referred to above appears not to have been generally recognized or applied by NC lathe manufacturers, particularly for machines at the smaller end of the range. Clearly, three-axis machines reduce cycle times because with typical components work is carried out in the bore and on the OD simultaneously. This is of greater importance when machining smaller-diameter components for two reasons. Firstly, many small components are either made from solid bar or do not have a cored or forged hole, so that the hole has to be produced from the solid. Secondly, the effect of the speed differential between OD and ID tools is generally less on smaller components. With larger components, the surface speed differential between tools in the bore and on the OD can be so marked as to make a two-axis machine a more feasible proposition.

Other features of machine design can also play a part in reducing non-cutting time. The distance and speed of tool withdrawal movements to allow indexing to take place, the indexing speed and general machine accessibility all affect lathe idle times. The reduction in tool requirement gained from the interaction of NC and GT has therefore a double effect on the non-cutting time by reducing both the tool setting and the tool changing time by turret indexing or other means.

3.3 NC lathes/GT—set-ups

Earlier in this paper it was stated that the main problems with set-up variations lay in the turning field. With conventional turret lathes, the use of identical set-ups for two components necessitates either exactly identical features and identical dimensions (or the omission of certain features), or alternatively, comprehensive set-ups with certain tools being left on the machine, some of which are omitted from the machining cycle. The opportunity to leave machines completely set is less frequent in cellular layouts than in functional layouts, particularly in general engineering work. Another approach is to develop set-ups which are flexible rather than comprehensive. This enables the adjustment of tools and stops to produce a range of diameters and lengths. Such solutions have more relevance in medium batch production, where the higher volumes improve the ratio of operation time to set-up time.

It is possible to develop sub-families with preferred and clearly identified categories of set-up, recognizing that even if this is done, in the majority of cases tools will still have to be changed, and this can lead to a strong argument for a greater use of pre-set tooling and quick change tooling.

The application of NC lathes virtually solves the set-up problem. The prime reason for this is the extremely low time taken to set the machine completely. A supplementary reason is that a wide variety of components only require a limited number of diameter and length turning tools. It is interesting to see the reduction that has taken place over the last few years in the number of NC lathes which previously had large tool magazine capacity.

Whereas the NC lathe can handle a wide variety of component types with limited tool change, the development of sub-families can further refine this opportunity. In addition to the great advantages of reduced set-up costs and the flexibility of the NC lathe in handling random dimensions of a variety of components, its capabilities are further maximized by making such machines cell pacemakers. As indicated earlier, there is little point in having a first-operation machine generating work in progress without back-up facilities, and there is a clear argument for the NC lathe to be the lead machine in a cell. This argument is of even greater importance with larger NC lathes, not only because material costs are generally higher with bigger components, but also because subsequent operations are generally more costly and queuing times in functional layouts are longer for these bigger components.

3.4 NC lathes/GT—programming and the operator's involvement

(a) Control

A common advantage claimed for NC machine tools is that of improved management control, but it is suggested that this statement is frequently fallacious. Control of the machining cycle is transferred from the operator not to a manager but to the part-programmer, who frequently is an ex-operator working in an office instead of operating a machine. This is not to imply that programmers are not responsible people (indeed the success of the NC application is largely dependent upon their capability), but that the control that is given is technical control rather than management control. The part-programmer, even with aids such as computer terminals, has a complex and difficult task to carry out; he is frequently subjected to criticism by shop floor management and others who cannot accept tape errors as reasonable, and such criticisms can lead to inherent caution being built into the program by the part-programmer. The programmer, who is generally removed from the NC lathe operation both by time and distance, has to allow for possible hazards such as surplus material on forgings, variations in casting size and so on. With the majority of conventional lathes, even PSC machines, the operator can take short cuts and improve the machining performance, particularly where problems already allowed for by the process engineer do not arise. As with most NC lathes, all machine instructions are generally on tape and not quickly capable of being altered, the machine operator does not have the facility to make adjustments of any consequence, with the result that the machine may be cutting at lower conditions than are ideally possible.

(b) The operator's involvement

This point has been recognized in the development of the HMT mini-computer system so that the operator can modify part-programs on the machine to optimize the cutting conditions.

There is little point in having rigid 'management control' if the result of the control is to give cutting conditions which are less than perfect. Surely it is better to release some of this control so that the operator can be less exasperated and can himself take command. This paper does not set out to deal with the sociological aspects of NC/GT but there is no doubt in the author's mind that the greater the involvement of the people actually cutting metal, the better.

A further development in this field which improves the operator's control is the opportunity to enter the tape, to correct any errors, to optimize performance and then to punch a new tape in the machine control cabinet under the command of the machine control system. For example on a three-axis machine, having optimized the OD turning conditions the operator can then balance the cycle by turning the end-working feeds.

(c) Programming requirements

Traditionally, most production engineers have been rightly concerned with the need to increase spindle utilization time. A great deal of their work has been

concentrated on the non-running time in order to make it smaller. Because the NC lathe generally doubles the spindle utilization time, it is of importance that the concentration on non-cutting time should be reduced and a great deal more attention should be paid to increasing the effectiveness and not just the quantity of spindle running time. With modern tooling, the range of speeds and feeds available and so on, metal removal rates on NC lathes can be substantially higher than in many current applications. Referring to an earlier comment in this section on the opportunities given by MDI, one can raise the question of whether the part-programmer should assume ideal conditions and provide a 'tight tape' which the operator can de-tune, or whether the programmer should provide a coarse tape which the operator can up-rate. This not only depends upon having a control system with MDI but also upon the ease with which certain modern control systems can be used for tape proving. There is little point in giving the operator the opportunity to adjust the tape command if he is going to spend a great deal of time determining errors on the first run-through. The question of speed of tape proving, and the facilities required for this, will assume increased importance, especially if the batch sizes are small.

(d) Future trends

Reference has already been made to the benefits to be gained from an improved coupling of the part-programmer's work and the operator's work. This section of the paper deals with at least one of the options for simplifying programming with the attendant advantage of reduced errors. Families of parts manufactured on NC lathes in a cell environment can have their part-programs prepared at less time and cost. The development of a library of data, against a range of sub-families, in itself simplifies this task. The next stage of such work is to write a master computer program for a prescribed family or sub-family, dealing with speeds, feeds, materials, tooling and so on. This can lead to the designer describing the component using a sub-family identification instead of drawing it. The part-programmer's input is simply the insertion of dimensional numbers, and not ordinates, into the master program for the sub-family. It has been estimated that in the Coventry Works of the writer's company, at least 60 per cent of rotational parts to be turned on NC lathes can be placed in sub-families and programmed by this method. There are other approaches to this question, all of which lead to reduced cost of NC lathe programming in a GT structure.

We should next consider the relationship between NC/GT and part-program on-machine storage. In the writer's view NC lathe systems need to have the facility for storage of a complete program for a single component of the most complex type. We should also give the machine operator comprehensive MDI facilities both in type and extent (for example multi-offsets). Ultimately the machine operator should have sufficient MDI opportunities to effectively countermand and replace every tape instruction. This is necessary, as explained earlier, in

order for the operator to fully optimize the operation sequence and the cutting conditions. To supplement this the control system must have adequate storage or memory capacity. This memory must also be sufficient to reduce the tape-reader's operation to reading the tape once per batch and not once per component. The final objective of single-program storage is to optimize cutting conditions progressively through the first batch and then to punch a new tape via the machine's control system/memory with an in-built tape punch. Not only does this give the management a better return on its investment but, equally important, the machine operator has the satisfaction of participating and contributing to improvements with his own skills.

These objectives should be recognized by modern management as essential requirements.

The question of multi-program, on-machine storage follows naturally. At first sight this may appear to be attractive. Actual examination of the benefits gained against premiums paid will generally reveal that such proposals are expensive, albeit elegant solutions to an illusory need. Such considerations should include striking the correct balance between the high cost of disc storage (and its continuing reliability in a shop floor environment), the lower cost and higher reliability of core storage (but limited capacity), and the low cost but relatively high access time of cassette stores. It is suggested that more important than the technical solution is the question of whether multi-program storage is of intrinsic value in view of the difficulty of controlling this facility even if the cost of its provision could be reduced.

3.5 NC lathe/GT—summary

It is the writer's conviction that NC lathes and GT will progressively come to support one another to such a stage that in the majority of circumstances they may ultimately become mutually dependent. The principal reasons for this belief are:

- (a) The NC lathe solves the set-up problem and gives the facility for reducing work in progress still further because smaller batches can be manufactured economically.
- (b) The work-in-progress generating capabilities of NC lathes generally need to be countered by the installation of such machines in cells.

CONCLUSIONS

A paper dealing with the relatively early period of a complex and rapidly developing subject can hardly be regarded as comprehensively definitive. The inter-relationship of NC and GT is so new that many of the aspects cannot yet be fully understood and many of the rules and guidelines are not yet established. Effects of scale, component variety and many other matters can lead to certain implications being of great importance in some situations and trivial in others.

Cellular manufacture leads to a closer control of manufacturing operations. In addition to the inherent reduction of work in progress which is automatically obtained, because of the shorter time for replenishing stock-cuts one can further reduce work in progress by making batch sizes smaller. With conventional turning machines, however, this may produce a more adverse and frequently unacceptable set-up:operation time ratio. With an NC lathe, because of the reduced set-up time, smaller batches are practicable.

The application of NC machining centres and NC lathes to families of components can lead to simplified part-programming with reduced variable input data thereby reducing programming costs and lowering the error rate. Both these types of machines will benefit from being placed in cells and their total success may often be dependent upon the NC machine being the pacemaker in a cell; NC and GT are self-supporting and will progressively be seen to give each other considerable mutual benefits.

SOME ORGANIZATIONAL ASPECTS OF CELLULAR MANUFACTURE BASED ON COMPUTER SIMULATION

by

D. ATHERSMITH* and J. R. CROOKALL*

SUMMARY

This paper examines numerical measures of effectiveness of cellular production systems which may be used, such as throughput time, work-in-progress and plant utilization. Using computer simulation, the effects of composition of cellular groupings are investigated. Factors in system design are considered, based upon balance and compromise between the measures of effectiveness. Also some aspects of dynamic states of cellular systems are examined, including concepts of elasticity and stability of operation, interference and choking.

INTRODUCTION

As the primary requirement of a production system is to produce, the first-order parameters are those centred upon this function and it is with these factors that this paper is concerned.

It is thus necessary to define capacity and performance characteristics, and even a theoretically ideal state, so that achievements due to cellular (or for that matter, any other) organization may be judged both in relative and in absolute terms. It is then rather dubious to expect that practical systems are capable of achieving more than definable physical limitations will permit.

Cellular manufacture is examined at the macro-system level, using computer simulation as the tool of investigation. The strongest argument in favour of the use of simulation is that of internal consistency. Results are neither subjective nor influenced by intangible or simply unobserved, external or spurious factors. Responsibility lies in the building in of an adequate representation of the practical situation for the purposes of the kind of conclusions sought. Definition of the problem is thus important, as is that of the measures of effectiveness to be employed.

Conceptually there is a sliding scale of cell design. At one extreme each cell contains a different set of machine types and as such is likely to manufacture a particular range of parts. At the other extreme, each cell contains the same machine types and therefore can potentially manufacture the same parts. Between these extremes a situation exists where certain cells have the same machine types, in part or in total, whilst other cells have different machine types. There are therefore fundamental differences in the nature

and purpose of cellular manufacturing systems. Whatever the purpose, most system planners share common aims such as increased capacity, a low level of work in progress and so on, and the ensuing consequences.

In an earlier computer simulation study, Baldwin and Crookall¹ sought to approach fairly closely the sort of detail and variety created in dealing with real components in the organization of a factory. After all, one basic principle in group technology is to reduce a multifarious and amorphous product mix to some semblance of manageable order and unification. The intention then was to examine improvements in machine resetting, transportation and work-in-progress in a cellular, in comparison with a functional, layout. The effect on resetting was very small, but the other two factors revealed more promising benefits. Furthermore, these findings were not in disagreement with some conclusions² resulting from the practical implementation of group systems in industry.

In the light of this it became evident that the use of highly detailed product characteristics are unnecessarily elaborate at the present stage, and that instead a more rigorous system model should enable the first-order characteristics, interactions, and dynamics to be studied. The distinction between approaches is therefore clear; we are here concerned with job/system performance analysis rather than with component analysis factors.

A major aspect of the work of which this paper reports a part, was therefore concerned with the development of a new system model for investigation purposes. The operation of this model will be briefly described.

* Imperial College of Science and Technology, London

OPERATION OF THE COMPUTER MODEL

The basic procedures and requirements of the computer simulation model can be summarized as:

- (1) part generation and input to the model;
- (2) matching of parts to cells;
- (3) maintaining an 'events' list, advancing time and progressing through the events;
- (4) repeated analysis of the state of the model and derivation of measures of effectiveness;
- (5) incorporation of constraints and decision rules.

Part generation involved generating by the Monte Carlo method up to six different operations to be carried out on six different types of machine. The sequence of operations must be preserved and back-tracking is not permitted. Operation times are also derived from a cumulative frequency distribution using the Monte Carlo method. Generated parts are entered as input to the system model at fixed inter-arrival periods (known as the JRATE) which can be changed for each new simulation.

Matching of parts to cells is on a priority basis, depending firstly upon the relative product requirements and cell provisions, and secondly on minimum jobs currently in progress. Inter-cell flow is not investigated, but would present no difficulty with the model.

The activity 'machine processing a part' is central to the design of the model, and the events are 'operation start time' and 'operation end time'. Figure 1 illustrates the functioning of the model in terms of the progression in time.

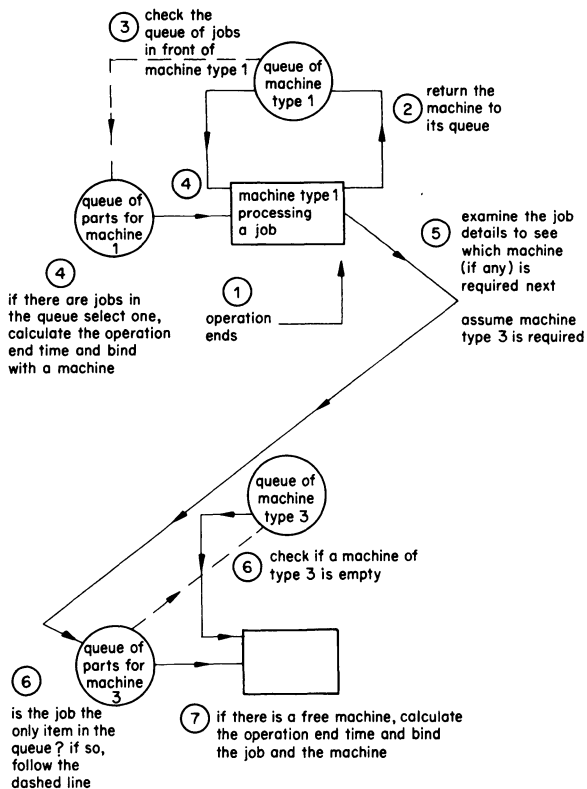


Figure 1. Representation of the simulated movement in time procedure.

- (1) When a control event occurs there are two separate entities, the machine and the job, which are both now demanding opportunity for further activity.
- (2) The machine is returned to the queue of machines of the same type (and in the context of the model, within the same cell). This occurs instantaneously with the operation ending.
- (3) The machine is now free and immediately looks to see if there is a queue of jobs waiting to be processed. If a queue of jobs exists, one is selected. The time is still the time of the control event.
- (4) It is now possible to calculate the time when the operation on the newly selected job will end. This event is a potential control event and is stored in a list of events, with identification to both the job and the machine.
- (5) The time is still the time of the control event as in (1). The job is now examined to establish whether or not it requires further operations, and if so on which machines. Thus the job is instantaneously moved into the queue of jobs which precede its next operation.
- (6) The next step is to check if there is a free machine of the type required for the next operation.
- (7) If there is a free machine the job and the machine are bound. The control event (the operation or activity end time) is obtained by adding the operation time to the time now.

The time now is still the time of the initial control event. However as all actions associated with the initial control event have been made, it is now possible to move on in time.

It is evident that, by examining both of the entities which are unbound by a control event, the system is geared to maintaining activity. Machines look back for jobs and jobs look forward for machines. The system operates whether queues are at zero or any integer value. The time of each operation ending control event is stored in the timeset.

Each time the event selected from the timeset is the time of the next or earliest arrival, a new job enters the system. The new arrival is directed to a queue of jobs ahead of a particular machine and thus undergoes the previously described procedure from (5) onwards.

The output from the model is in terms of the measures of effectiveness, namely

- (1) machine utilization—obtainable by machine type, by cell and by shop
- (2) work-in-progress level:
 - count of jobs by cells and by shop;
 - outstanding operation time by cell and shop
- (3) average time jobs remain in the cell and shop.

For the first two measures a value can be obtained for a period or interval of time, and also for the cumulative measure from zero time to the present time. The average job time is a cumulative average measure.

MEASURES OF EFFECTIVENESS OF PRODUCTION SYSTEMS

Measures of effectiveness follow from assumed criteria of success, and success itself can be subjective in reflecting the particular orientation from which a production system is viewed. Mechanistic measures of effectiveness have a familiar ring, for example:

- (1) work-in-progress level
- (2) machine utilization
- (3) labour utilization
- (4) job throughput time

of which the first three are system measures whereas the last is work orientated. Such measures may be cost-weighted, also. However even these seemingly factual and unambiguous factors possess more than one basis of interpretation, as will be shown.

That the optimization of production systems is more than a mechanistic problem may be obvious, but it is nevertheless based upon a desirable change in mechanistic terms. For example, the degree of management intervention and control, the predictability of throughput, social and quality measures are inextricably linked with the mechanics of system operation.

Work-in-progress

The components of work-in-progress may be the *number* of jobs in the production system, the material *value*, the labour or capital-added value at each operation, and even indirect costs such as storage. Overall, there is a relationship between work-in-progress level and capital employed in maintaining it.

But work-in-progress levels fluctuate over time (daily, weekly or monthly) and hence time enters into an effective measure of it. Making the assumption that value is added discretely when operations are completed, work-in-progress for a shop or cell is given by

$$L = \sum nM_n + \sum n_{i+1} (x_i + y_i + z_i) \quad (1)$$

where n is the number of jobs,

n_{i+1} is the number of jobs at the $(i+1)$ th operation,

M is the material value per job,

x_i is the direct labour and material-added value at each operation,

y_i is the capital-added value at each operation,

z_i is the indirect labour and material-added value at each operation.

This represents an *instantaneous* value. But on a balance-sheet analysis of effectiveness, an *average* value may be deemed to be more representative of achievement of aims. Because production systems are dynamic these measures can be expected sometimes to differ, the difference being representative of the chosen viewpoint.

The two terms above have distinct implications, hence $\sum nM_n$ depends upon the number of jobs in the system and the raw material content per job. Whilst important in relation to financial control of the system, material cost is not related to effectiveness and hence in this study the number of jobs only was monitored. It is noteworthy however, that in a situation where material costs are high or widely

different, expediting the material flow of high-cost items is worthwhile.

The second term in Equation (1) expresses elements of value which are time-dependent, and by assuming that they are invariant with the machine or operation performed, then time itself becomes a measure of the level of work-in-progress. The second term is thus a measure of the total completed operation time in the system at the chosen instant of observation.

A period measure is obtained by summing the measure (second term in equation (1)) over a desired period and dividing the summation by that period. By also making this a ratio of the number of machines in the system, a measure for comparison of different systems is obtained.

Two measures of work-in-progress are therefore used. A count of the number of jobs in the system is dependent upon the input and output rates of jobs in the system. If the input rate is comparatively too high the count builds up. However the level of completed operation time is reasonably constant here, but depends generally upon other system attributes.

Machine utilization

Similarly to work-in-progress, utilization can be expressed as an instantaneous or a periodic measure. For the present purposes machine utilization is a periodic measure of mean unutilized machine time per machine.

Job throughput time

This measure is an average value of the time required for a job to progress through the system.

RESULTS AND DISCUSSION

System behaviour with elapsed time

During a simulation run, jobs in the shop build up, and the effects of this in terms of the various measures may be observed for different inter-arrival rates, as is illustrated in figure 3. This is representative of the eleven-cell system ($N1 = 11$) of a twenty-eight machine shop with 'perfect matching' of parts to cells (figure 2). The virtually obvious result is that machine utilization (figure 3(a)) rises with a reduction in inter-arrival time (that is JRATE (min), which at present is a constant value), and utilization can reach upwards of 90 per cent. However the inter-arrival time of 8 min required to achieve this is below the critical value for the system, causing the system to fill up with jobs and the work-in-progress count to increase continuously (figure 3(b)). The importance of judging system performance by several measures is thus highlighted.

It is necessary to define sub-critical operating conditions as those in which measures can reach steady, equilibrium values, and it is most important thereafter to consider system performance in this regard. A critical value of inter-arrival time must therefore occur, below which the system is unstable and above which equilibrium conditions can be sought. Because there are many measures of system performance, there is a large number of possible solutions, and simple ideas of 'optima' must be

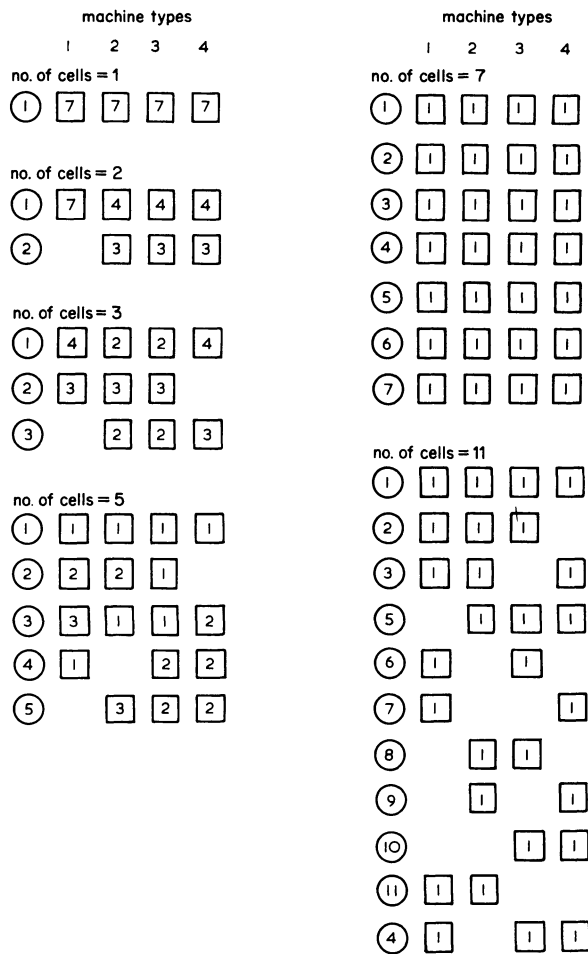


Figure 2. Cellular groupings for the twenty-eight machine shop.

regarded with due caution. It is salutary to note from figure 3 how sensitive are the measures to inter-arrival time (or input job rate) and hence the importance of this factor in system performance evaluation (cellular or otherwise!). Furthermore input rate affects the *response rate* to change. The input rate-output rate characteristics will be examined later in the context of different cellular constructions.

The work-in-progress level 'outstanding operation time' measure (figure 3(c)) does not show the rising trend of figure 3(b) for supercritical input rates. However the 'time jobs are in the shop' measure (figure 3(d)) indicates that a JRATE of 10 minutes is super-critical (in contrast to figure 3(b)) and hence appears to be a more sensitive indicator of system stability.

The effect of cellular grouping

Although plotted for a constant inter-arrival time (JRATE = 10 min), figure 4(a) suggests that system stability may be achieved at different inter-arrival times according to the cellular group formation, and this raises the problem of investigating stability and response of different systems under comparable conditions. This is confirmed in the work-in-progress 'count' and 'outstanding operation time' measures (figure 4(b) and (c), respectively), and that the ranking of the cellular groupings by these measures differs in certain respects. Certainly the seven-cell and

eleven-cell systems are super-critical under these conditions.

The count is dependent upon the input rate-output rate characteristics of the system (see figure 4(d)); thus once the system is operated beyond the 'elastic' region (that is, it is super-critical) the count builds up. It is noteworthy that the manner in which different systems become critical is of some significance. Thus whilst the eleven-cell system is the first to do so, its characteristics in the critical region are conducive to better stability in that the 'choking' phenomena (and even a marked reverse trend in the case of the seven-cell system) are absent.

Combined measures of effectiveness

At this juncture it is pertinent to consider some *combined* measures of system effectiveness and performance in relation to input rate as before, but also at different points in time of the dynamic build up of work within the system. Figure 5(a) shows the work-in-progress level in terms of the number of jobs in the shop per machine, and figure 5(b) in terms of outstanding operation time, both being plotted against cumulative machine utilization. The important feature is the relative effectiveness of cellular groups at different inter-arrival times. For example, at JRATE = 10 minutes the one-cell (functional shop) is most effective by both measures in figure 5(a), but in figure 5(b), whilst the one-cell system shows the best machine utilization, the eleven-cell system has a better outstanding operation time measure. A degree of inherent inconclusiveness is thus apparent due to differing stability states, in spite of the nominally equivalent basis of comparison. Hence the use of both a common inter-arrival rate *and* observation time in near-critical conditions leads to complex results which may be difficult to assess.

Under *stable* operating conditions (figure 6(a) and (b)) a definite pattern emerges, but interestingly the higher the inter-arrival time, the less is the difference in performance for different cell constructions. It is also evident that either work-in-progress measures and the utilization move smoothly together. Thus if it is construed that low work-in-progress and high utilization are good, it is immediately obvious that one factor is improved at the expense of the other. This is not to say, however, that a degree of under-utilization is not valuable in providing flexibility (elasticity) of operation, and likewise a certain level of work-in-progress assists in maintaining utilization at a reasonable level. It would hardly be a responsible strategy for a system planner to design for no flexibility under average conditions of working, but paradoxically the efficacy of correct system design in these terms decreases somewhat as utilization falls.

In practice, it is probable that the least certainty exists within the starting data, thus, for example, any plant capacity which is to be aimed for is at best a forecast and at worst a guess. It is relevant therefore to consider the effect of *variability* in input factors on system performance.

Variable operation times

Lack of knowledge of future operation time variations does not prevent examination of the *probable*

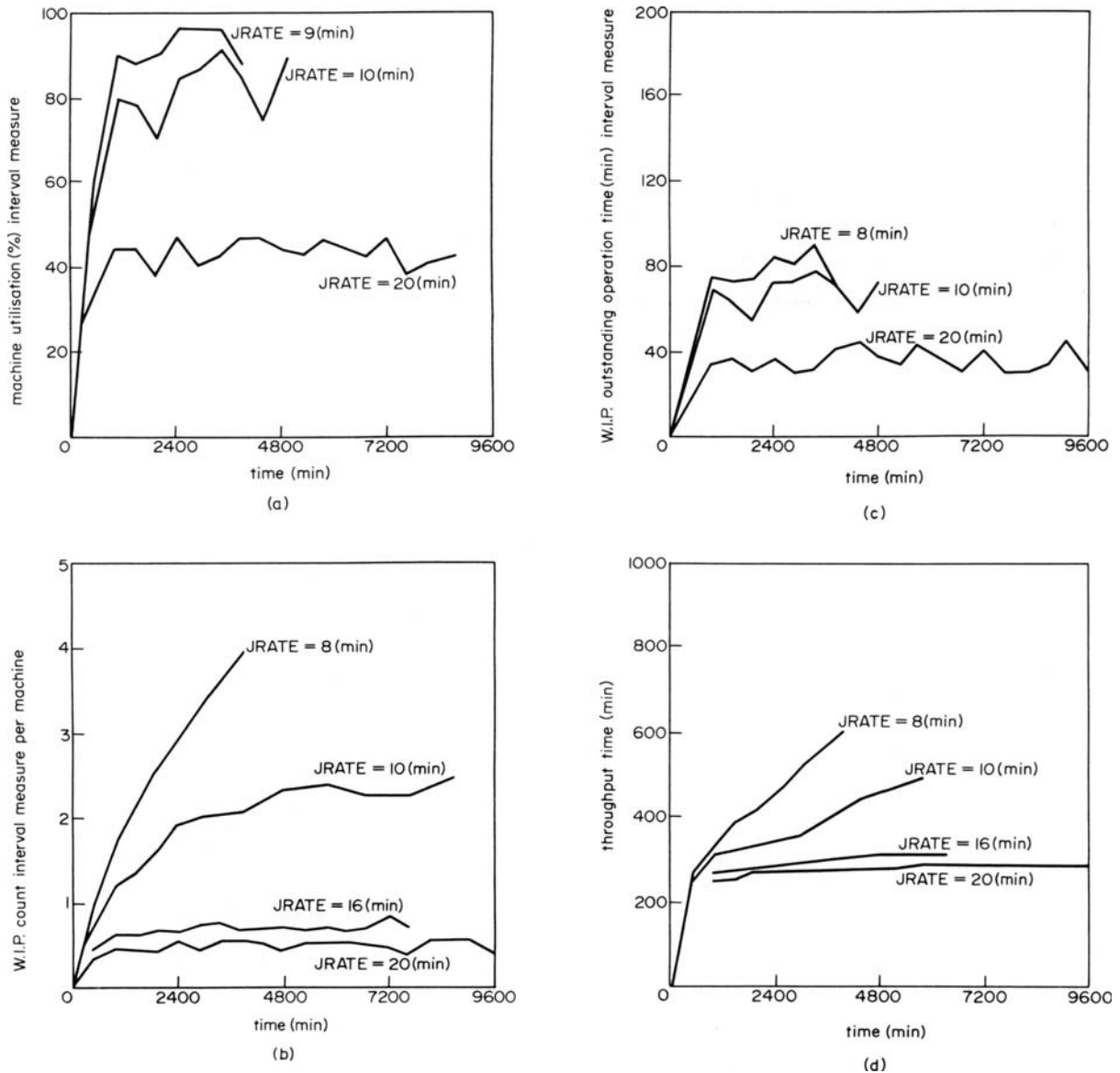


Figure 3. Measures of response of an eleven-cell, twenty-eight machine shop system with perfect matching. (a) Machine utilization, (b) work-in-progress count (interval measure), (c) work-in-progress outstanding operation time (interval measure), (d) time jobs are in shop.

effect of these compared with the regular values so far used. Operation times normally distributed with a mean of 100 minutes (equal to the regular value) and a standard deviation of 30 minutes have been used. The results for one-cell and eleven-cell systems are summarized in Table 1 for JRATE = 10 min.

For the eleven-cell shop the effect of random operation times can be seen to be greater than for the one-cell shop, for which relatively small increase of 6 per cent in work-in-progress count only was obtained.

TABLE 1 Variations in measures due to random operation times (standard deviation σ /mean = 0.3)

No. of cells	Operation time	Work-in-process count		Throughput time (minutes)	
		Regular ($\sigma = 0$)	Variable (σ /mean = 0.3)	Regular ($\sigma = 0$)	Variable (σ /mean = 0.3)
1		0.43	+6	250	0
11		0.51	+14	285	+8

It should be noted also that the *mean* values of measures for regular operation times ($\sigma = 0$) are also slightly higher for the eleven-cell shop; hence the percentage increases shown due to variable operation times are additional. Machine utilization was found to be virtually unaffected (normally less than 1 per cent) by the variable operation times.

System stability

The time that jobs are in the shop (that is throughput time) provides a good indication of system stability and the way it is reached. Figure 7 shows this plotted against elapsed time for JRATES of 9, 10 and 11 min respectively. At JRATE = 9 min none of the cellular constructions achieve stability, but at JRATE = 10 min stability is quickly reached for the one-cell and two-cell systems, and is being approached by the three-cell and eleven-cell systems. At JRATE = 11 min, all except the five-cell system achieve stability but their response rate differs, the one-cell showing the fastest growth to stability. Figure 7(d) shows the

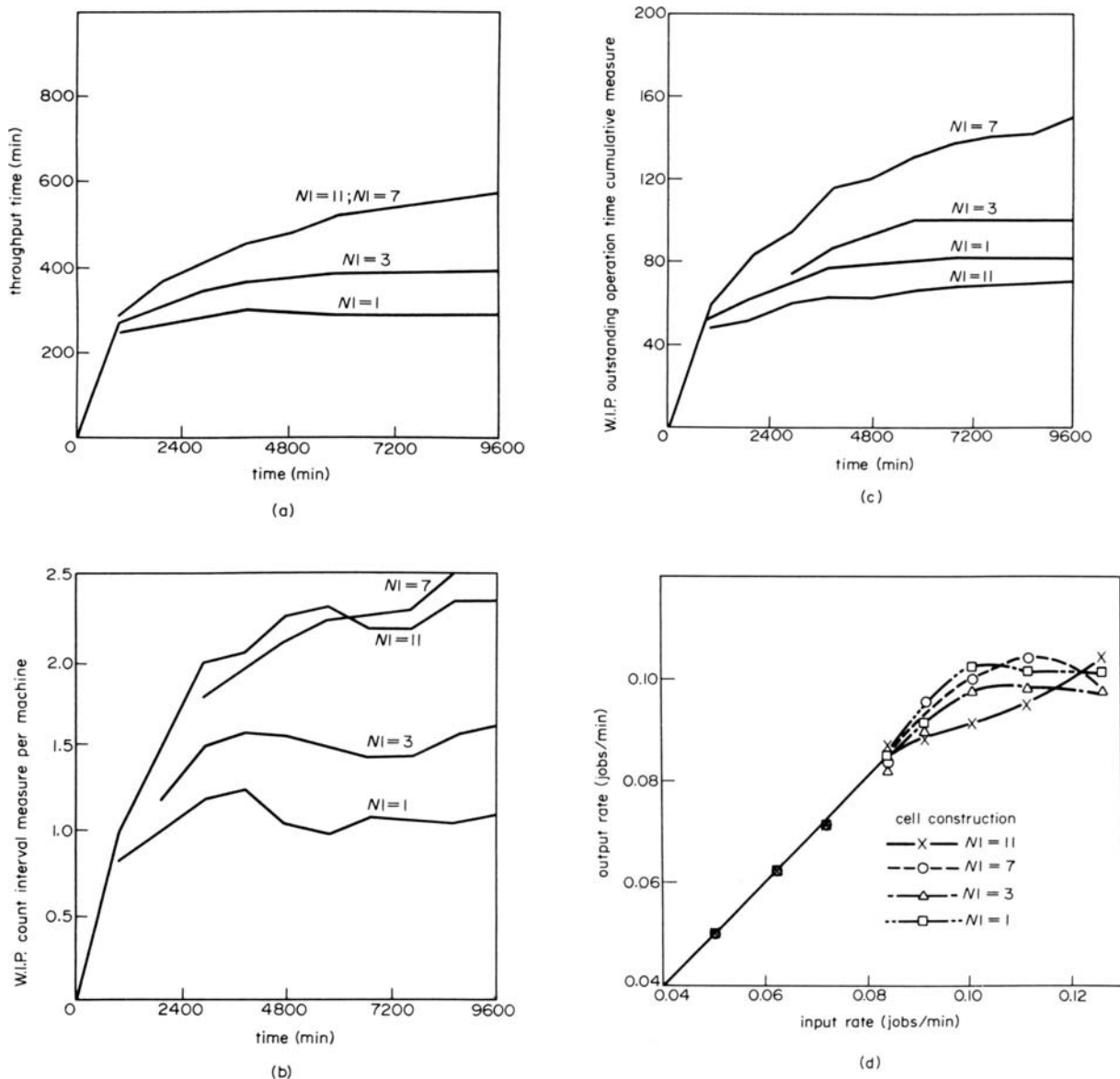


Figure 4. Effects of cellular groupings on production system behaviour, JRATE = 10 min. (a) Time jobs are in shop, (b) work-in-progress (W.I.P.) count (interval measure), (c) work-in-progress outstanding operation time (cumulative measure), (d) input rate/output rate characteristics.

values for the time that jobs are in the system under *stable* conditions against the inter-arrival rate; the ranking order under this measure is clear, and in fact is that of the number of constituent cells themselves, with the exception of the five-cell system which was unstable under all the conditions over this considered range. Greatest stability and least throughput time is achieved with the one-cell (functional) system.

Growth to stability in terms of the combined measures of work-in-progress count and machine utilization are shown from start-up to stable (or near-stable) operating conditions in figure 8 for a constant inter-arrival time of eleven min. The plotted points are at equal intervals of elapsed time. The system starts of course from the origin (zero work-in-progress and zero machine utilization). As jobs build up the initial growth is approximately linear; the greater the number of cells, the steeper is the slope in this region. There follows a non-linear section, with points becoming closer together as stable operating

conditions are approached. This non-linear behaviour is more pronounced the greater the number of cells in the system, and is due mainly to a higher rate of change in the work-in-progress, suggesting increased interference in relation to parts/cell matching.

This aspect of cellular manufacture involves the 'accessibility' of parts to cells. Whether a particular part is processed by a machine immediately on arrival to the system depends upon the cell construction and the inter-arrival time value relative to the operation time. For example, the eleven-cell shop, where each part is perfectly matched to a particular cell, may be contrasted with the functional shop. Due to part/cell capacity mismatching (which can often happen on a short-term basis even when the long-term situation is balanced), a situation can arise in which queues form ahead of certain cells within the eleven-cell shop, yet those parts would gain access immediately to the functional shop without forming queues. This is termed 'accessibility' of parts to cells, and it is

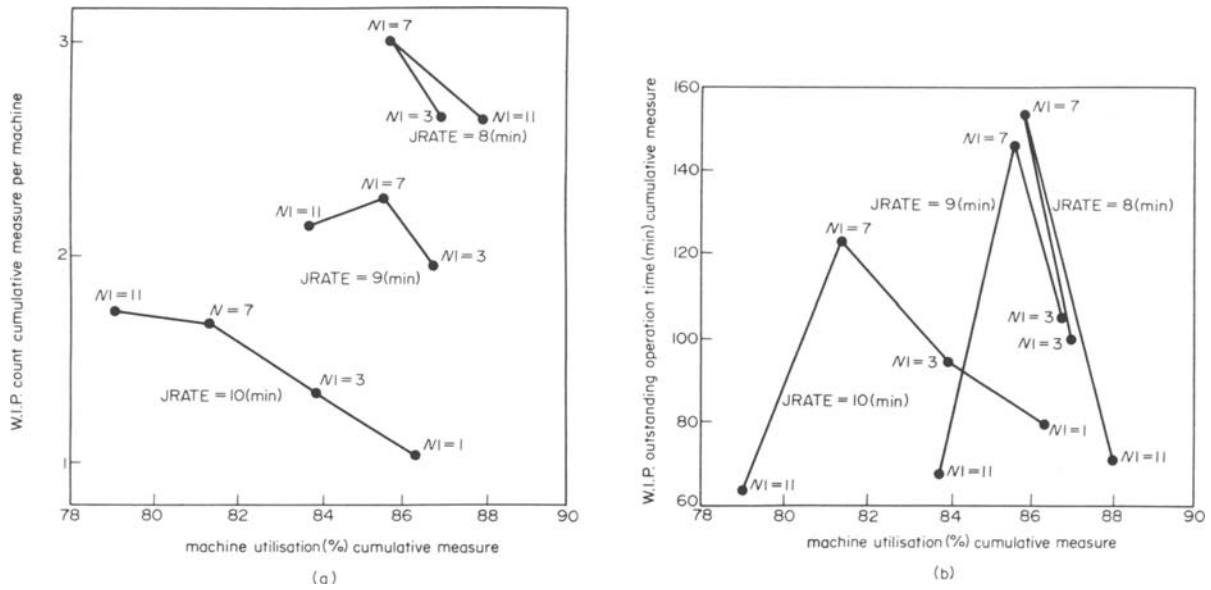


Figure 5. Combined effectiveness measures under supercritical operating conditions. (a) Work-in-progress (W.I.P.) count (cumulative measure), (b) work-in-progress outstanding operation time (cumulative measure).

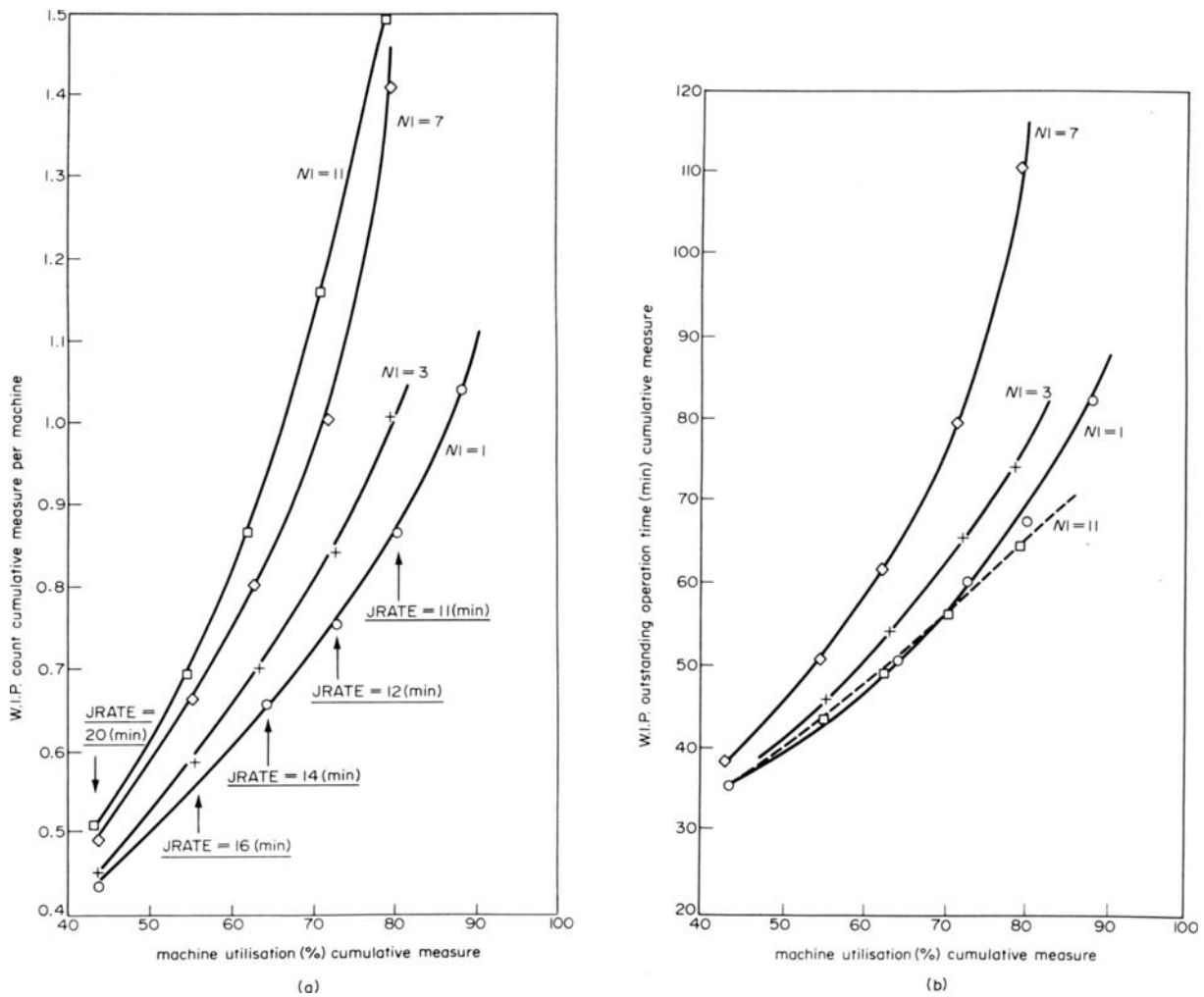


Figure 6. Combined effectiveness measures under subcritical operating conditions. (a) Work-in-progress (W.I.P.) count, (b) work-in-progress outstanding operation time.

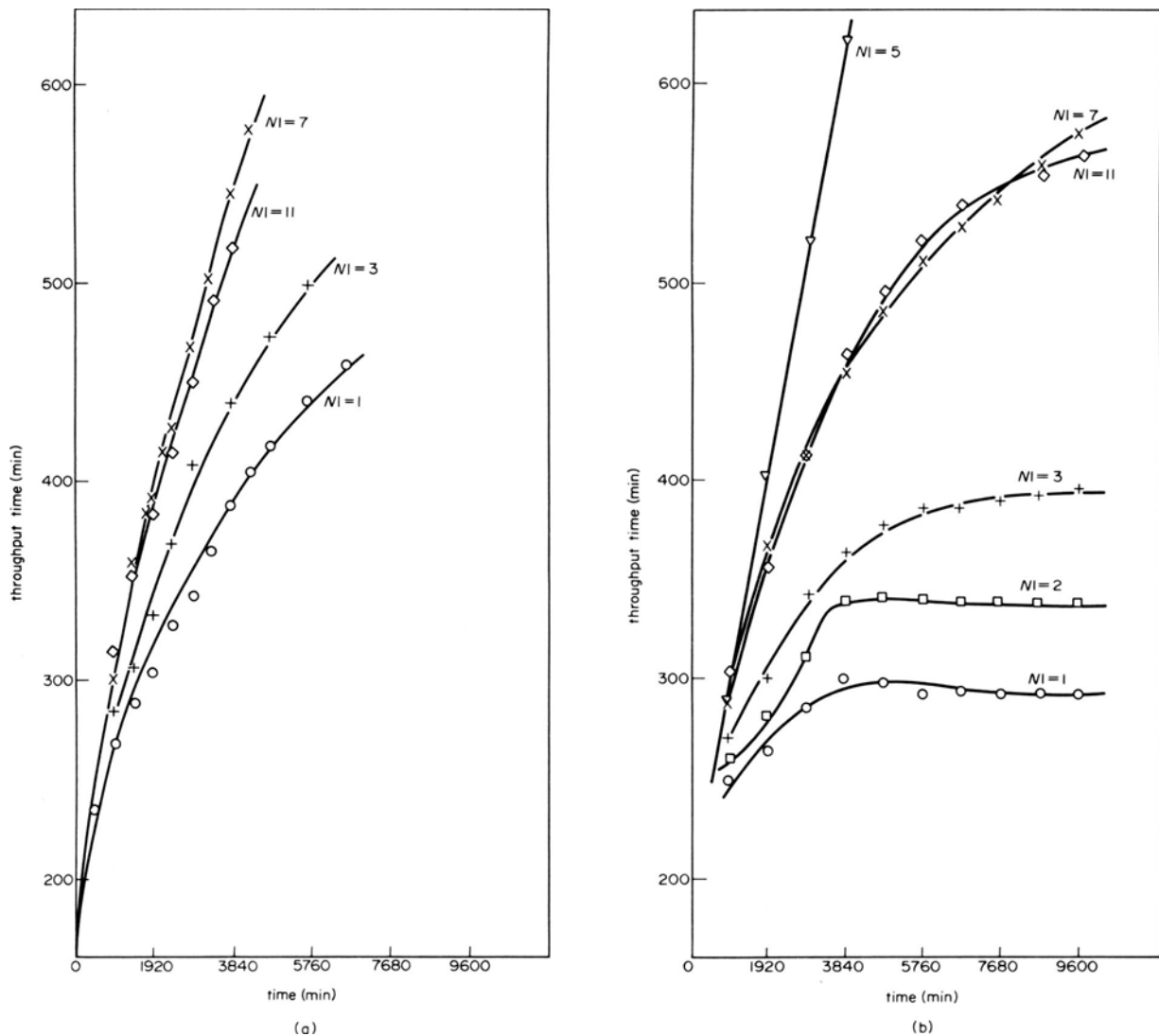


Figure 7. Throughput time characteristics illustrating system stability. (a) JRATE = 9 min, (b) JRATE = 10 min.

obviously a function of rigidity of grouping. No inter-cell movements have been permitted in the currently discussed simulations with the model, but evidently this would help to improve 'accessibility' of parts to cells. Nevertheless, taken too far this in itself negates the philosophy of cellular manufacture. A second aspect of accessibility to the system relates to the tactical cell selection mechanism. Whatever loading method is employed as a short-term tactic, a degree of limitation of access then results in relation to subsequent jobs.

For systems operating at the same job inter-arrival time and processing the same parts, access to the system is characterized by the work-in-progress count and the throughput time. Both of these tend to increase as the number of cells are increased, but there are cases where this generalization does not hold. For example, the cell selection rule for the jobs entering the seven-cell shop is the cause for the different ranking. A similar argument applies to the five-cell shop; the cell design for the five-cell shop is poor, and this is the major contributing factor in the relatively high work-in-progress results. Thus such differences in cell performance are due largely to the relative accessibility of parts to cellular systems under comparable job arrival patterns.

Related concepts of the perfect system

Consideration in this paper has been of a particular case, namely a twenty-eight machine shop, and it is thus pertinent to inquire how particular are the results in relation to other possible systems. Whilst it must be left to a further paper to deal with this aspect in detail, a general point arises in relation to this question.

A theoretically perfect system can be defined as one in which, whilst operating in a stable mode, interference and queuing is eliminated, the work-in-progress count per machine is unity and the machine utilization is 100 per cent. It is towards such an ideal that, by appropriate design, a production system operating under practical, non-ideal conditions is aimed. Furthermore the degree of success may be judged in relation to this concept which thus constitutes an absolute measure, as opposed to the use of local inter-comparisons, and the pitfalls inherent in not comparing like with like no longer apply.

Referring to figure 8 (page 10), if a point representing the above ideal condition is joined by a straight line to the origin, it will be found that the linear part of the one-cell (functional) construction is virtually coincident with it. Thus the point of divergence from this of the non-linear part represents

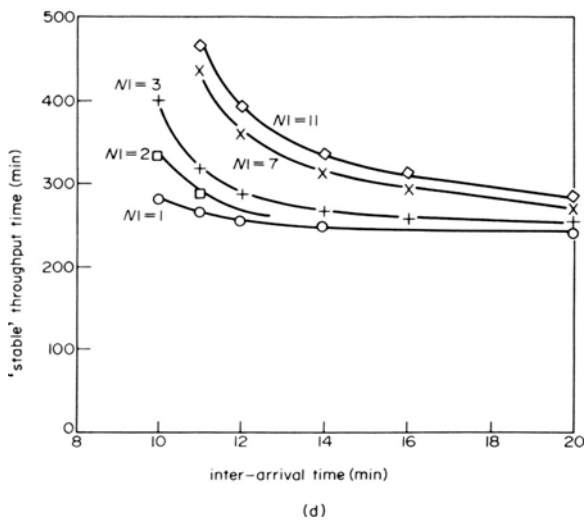
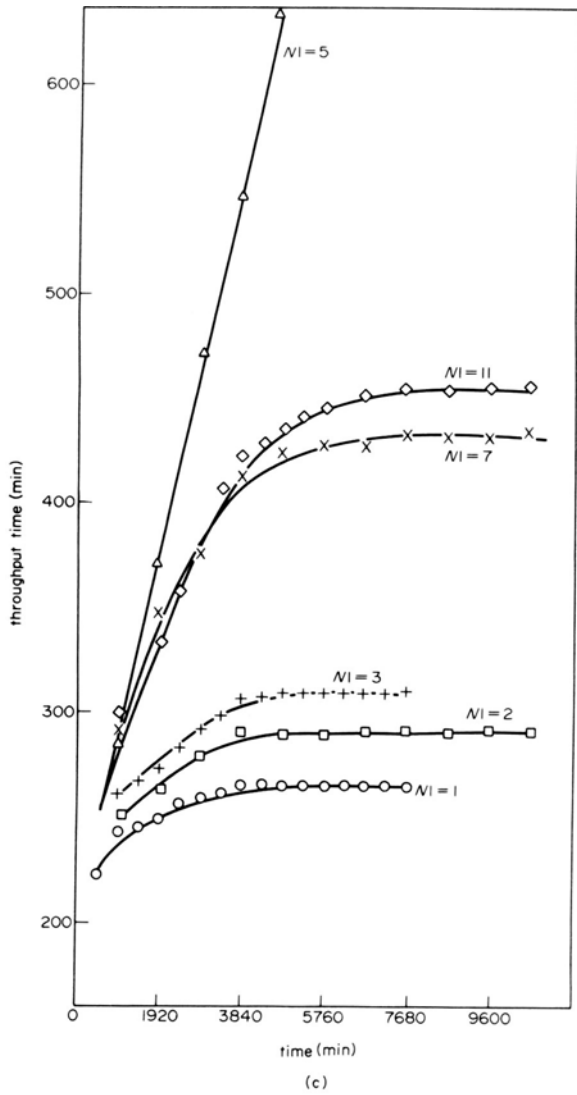


Figure 7. contd. (c) JRATE = 11 min, (d) throughput time for stable operation.

a deviation from the ideal behaviour, and similar conclusions apply to the other curves. Thus this system may be deemed ideal, paradoxically, provided it is under-loaded. As figure 8 depicts a sub-critical situation for all the cellular constructions used, and this has been seen to differ for super-critical conditions, certain inter-related general conclusions may be drawn.

The relative effectiveness of systems operated under *sub-critical* conditions depends upon the manner of sub-division of the system as it affects 'accessibility' of parts to cells: less sub-division is better than more. The operating point at which a system becomes critical depends upon the cell structure and the relationship between the input rate to total system size. In the *super-critical* region cell structure plays a more important role and the order of preference can be quite different. A cell structure which becomes critical early (such as the eleven-cell system here) may possess a more favourable operating characteristic in the super-critical range, particularly with regard to system stability and flexibility.

CONCLUSION

The work represents a basic simulation study of cellular manufacture in the sense that sub-division of the system, and the loading and allocation of parts, is purely numerical.

The various measures of effectiveness and conditions of operation investigated indicate that the ranking order of different cellular constructions may differ quite widely. In any particular situation therefore it is important to be convinced that the chosen criteria are the most relevant under the circumstances, and at least to be prepared to accept lower levels of performance in relation to others.

All the evidence points to the fact that solutions to cellular manufacture must be of a 'weighed compromise' nature, and if simple optima appear it is likely that a partial solution only has been obtained. Human judgment then is still indispensable, but can benefit considerably from aids to system planning such as have been described, and the development of which has been an essential aim during the course of this work.

According to the measures of effectiveness which have been used, two regions or modes of operation can be identified within which system characteristics are likely to be different. One is characterized by *stable* operation, that is in which equilibrium conditions can be reached (within a simulated time scale which is of practical relevance), and conversely the other exhibits *unstable* characteristics in which factors of an accumulative nature (particularly work-in-progress) deteriorate monotonically with elapsed time. Work-in-progress is in fact a convenient outward and visible sign of the mode of operation relative to the overall system capability.

A revealing quantifiable factor in this regard is the system input/output characteristics. In the stable mode this yields a proportional relationship, but in the critical region it exhibits a transition to 'choking' behaviour, and the trends which occur in this region

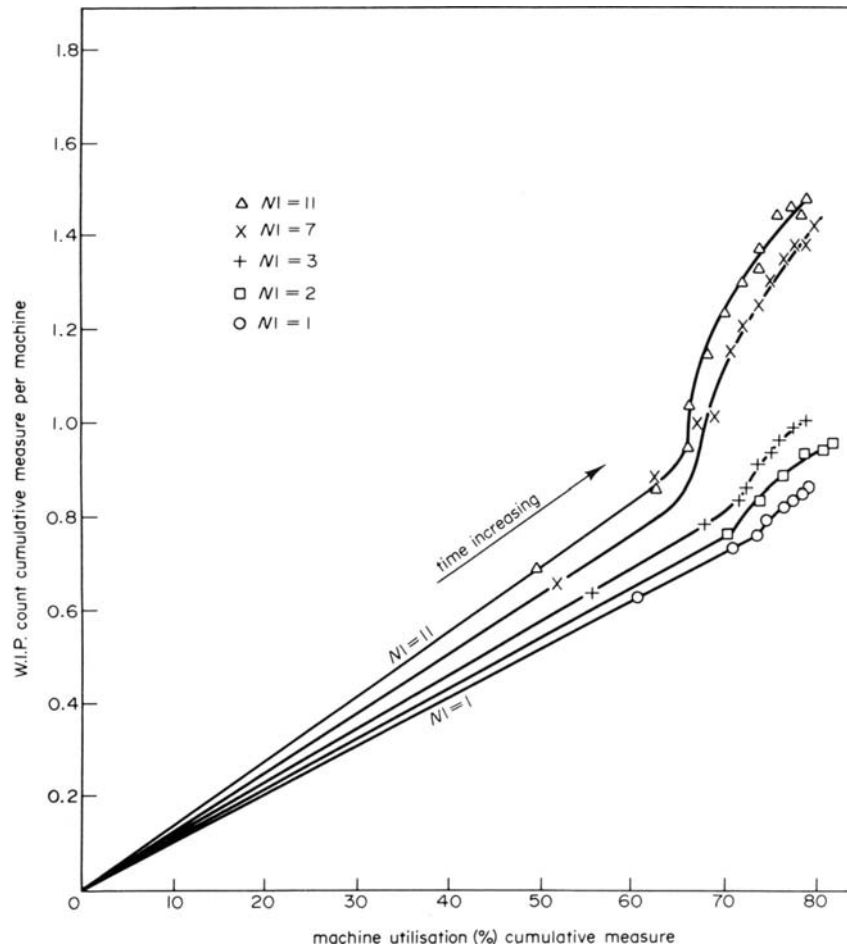


Figure 8. Characteristics of system growth to stability; JRATE = 11 min.

are plainly indicative of the manner in which the system becomes unstable. It is here that differences in system construction and organization become most evident, and that benefits from correct system design are therefore likely to be greatest. It is also probable that either range does not extend to all operation levels of a system in practice, and that 'worst case' as well as average conditions should be considered in relation to system design.

A significant factor which emerges is the 'accessibility' of parts to the cellular system as a whole and hence the rigidity of grouping in relation to the absolute size of the system. It is distinctly possible to classify and group too rigidly, and hence to have insufficient flexibility within the system to cope with inevitable random fluctuations in the input rate of individual jobs.

ACKNOWLEDGMENT

Thanks are due to the Science Research Council for an award to one of the authors (D. Athersmith) which enabled him to undertake this study as part of his advanced course in Production Technology at Imperial College.

REFERENCES

1. K. I. Baldwin and J. R. Crookall. An investigation into the application of grouping principles and cellular manufacture, using Monte Carlo simulation. *CIRP Journal on Manufacturing Systems*. (1972) 1, no. 3, pp. 193-208.
2. F. Craven. Some constraints, fallacies and solutions in G.T. applications. *Proc. 14th MTDRC Conf.*, Manchester (1973), paper no. 1013.

CNC AND DNC

BCLCNC—A COMPUTER NUMERICAL CONTROL SYSTEM

by

N. AKGERMAN*

SUMMARY

A Computer Numerical Control (CNC) system, which provides for cutter path programming for three axis contouring directly on the control minicomputer, is described. The system is built around a PDP-8/E computer and makes use of FOCAL an interpretive language both of which have been developed by Digital Equipment Corporation, Maynard, Massachusetts, USA. The computing capabilities of FOCAL were augmented with machine-tool control functions, switch sensing functions and a special routine for communicating with a large-scale computer over phone lines. All input/output and machine control is performed in the interrupt mode, thus, freeing the processor to perform computations most of the time. On line editing commands provide the operator with the flexibility needed to optimize the process.

NOTATION

F	Linear feed rate (in/min)
X, Y, Z	Absolute coordinates of cutter position (in $\times 10^{-3}$)
$\Delta X, \Delta Y, \Delta Z$	Incremental values of cutter motion (in $\times 10^{-3}$)
A	Angle in radians
U	A decimal value
N	A positive integer between 0 and 11 inclusive
L	Length in inches
K	A positive integer, sequence number.

INTRODUCTION

During the past few years, the use of minicomputers as machine tool controllers has gained more acceptance. Due to the continuing reductions in minicomputer prices, the economic crossover point between hard-wired Numerical Control (NC) and Computer Numerical Control (CNC) is at least at the level of three axis machining centre with contouring controls. Any CNC system only appears more complex than conventional NC. As the machine functions get more complex, so do the hard-wired systems, whereas, in CNC applications, only the control algorithms change, the hardware is therefore fairly stable. CNC is also a logical building block for Computer Aided Manufacturing (CAM) applications. NC data management, work station monitoring and reporting, com-

munication with a control computer are easily accomplished if the controller is a CNC.

The application of adaptive control loops for process optimization is also greatly facilitated in CNC systems; the major change would be in the control software which can be easily upgraded as better and more efficient adaptive control algorithms become available.

Another interesting advantage of the CNC systems is the possibility of having direct access to information processing equipment in the workshop itself, CNC systems introduce minicomputers to the shop floor. With proper configuration and software, they may be made available, at least part of the time, for part program processing and other computing needs. The minicomputer would be used by manufacturing personnel who were obliged before to wait for program processing in computing centres and who lost time for error corrections for each program. Even if the installation has a lower performance than big computing systems, an overall gain in time is given by the elimination of waiting and error correction times. The availability of the computer in the workshop is also an excellent incentive for manufacturing personnel who are growing used to these technologies and will be prepared for a more computerized system in the future, especially for the introduction of direct numerical control.

The work discussed in this paper was conducted at Battelle's Columbus Laboratories and is an improvement on work performed at the University of California at Berkeley by the author and his colleagues¹.

* Manufacturing Technology Department, Battelle's Columbus Laboratories, Columbus, Ohio, USA

EQUIPMENT

As illustrated in figure 1, the major components of the CNC system are

- (1) a minicomputer with associated peripherals,
- (2) manual control panel and step motor drivers,
- (3) a medium-sized milling machine fitted with ball-bearing lead screws coupled to stepping motors.

The minicomputer is a PDP-8/E with 12K words of memory and is equipped with a high-speed paper tape reader/punch, programmable real time clock, digital buffered I/O interface, A/D converter, teletype, and an asynchronous serial computer interface. A machine tool control interface which was designed and built at BCL, plugs into the PDP-8/E omnibus

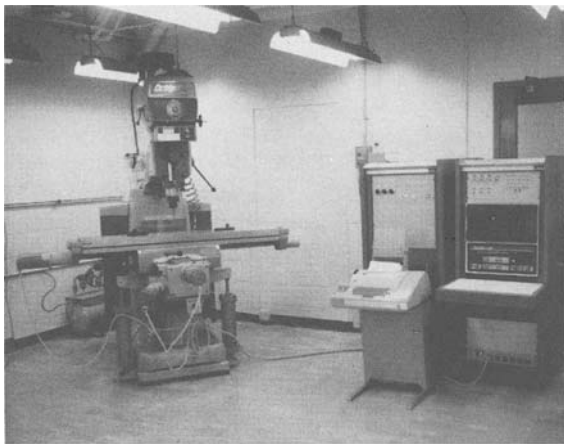


Figure 1. CNC machining system consisting of a minicomputer, high-speed paper tape reader/punch, teletype-writer, manual control console and step-motor drivers and the milling machine.

and supplies all the necessary signals to the bipolar-chopper drivers of the stepping motors which are coupled to the three lead screws. Both the bipolar chopper drivers and the stepping motors were manufactured by Sigma Instruments, Inc., Braintree, Massachusetts, USA. The machine tool is a Bridgeport Mark II with ball-bearing screws and preloaded ball-nuts on all three axes of motion. The table is balanced by air-cylinders, thus, eliminating the back-driving potential of its weight.

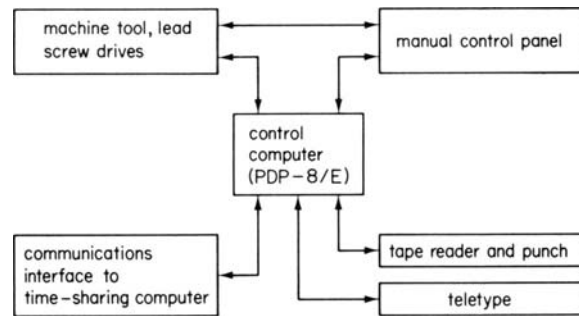


Figure 2. Outline of computer numerical control system.

Figure 2 outlines the overall relationship of the various hardware components. Except in the manual control mode, information is always processed through the minicomputer. In the manual control mode, a joystick is used to position the table, while during CNC mode, the minicomputer receives information from the tape reader, from the keyboard, from the manual control panel, or from a time-sharing computer and provides the necessary pulse train to the three stepping motors driving the table. It is also possible, using the teletype, the tape punch and the communications facilities, to create a part program on a host computer to process through an APT system and after postprocessing to punch the NC tape without leaving the machine site. Another mode of operation is to program cutter paths directly on the control computer using a fairly general purpose language called FOCAL-8², to which a set of machine tool control functions were added.

SOFTWARE SYSTEM

The machine tool control programs are written in PDP-8 assembly language, PAL III, and are overlaid to FOCAL-8 extending its capabilities. The system is called BCLCNC (Battelle's Columbus Laboratories Computer Numerical Control) to distinguish it from other similar systems also based on FOCAL.

Together with 8K FOCAL-8, BCLCNC requires a total of 12K memory space. FOCAL-8 resides in the first 8K and the last 4K is occupied by the machine tool control software. Since the system is memory resident at all times, it was possible to replace FOCAL-8 disk monitor command LIBRARY by a new command PERFORM to provide a logical means

TABLE 1 Available commands and their abbreviations²

A	ASK	user enters values defining each variable
C	COMMENTS or CONTINUE	null statement
D	DO	subroutine call
E	ERASE	program editing command
F	FOR	program loop control
G	GOTO	transfer program flow
I	IF	branch on arithmetic result
M	MODIFY	line editing command
P	PERFORM	perform machine tool related functions
Q	QUIT	return to command mode
R	RETURN	subroutine return
S	SET	evaluate expression and save result
T	TYPE	print-out the results in current format
W	WRITE	print-out program listing selectively
*		Read indirect program from the high speed reader

of access to the machine tool control routines, a list of the commands are given in Table 1. All the control routines can also be used in conjunction with the SET command, however, in most cases, this procedure is meaningless as only a few of the added functions return a value to FOCAL-8.

Almost all the standard functions of FOCAL-8 are retained in order to have a fairly complete computational capability. These mathematical functions are

FSIN (A) Sign of Angle A , where A is in radians
 FCOS (A) Cosine of angle A , where A is in radians
 FEXP (U) Exponential function, e^u
 FLOG (U) Natural Logarithm of U
 FATN (A) Arc tangent of A
 FSQT (U) Square root of U
 FABS (U) Absolute value of U
 FSGN (U) Sign of U
 FITR (U) Integer part of U

BCLCNC adds the following functions to the ones listed above.

BGN() as its name implies, is used to initialize the system. All buffer areas are cleared, device interrupts are energized and the system is put in CNC mode.

END() this routine cancels CNC mode, device interrupts are disabled and only FOCAL-8 is operational.

Usually, both BGN and END are executed in direct mode, however they are not used frequently. It is necessary to use BGN when BCLCNC is first loaded and after a call to END, in order to get the system back into CNC mode. The high speed reader input command of FOCAL-8, *, is only operational after CNC mode is disabled by a call to END. The following routines effect actual machine tool control.

CUT(X,Y,Z) moves the tool to the point with absolute coordinates X,Y,Z at feed rate F in/min. The motion is in a straight line from the previous position.

MOV($\Delta X,\Delta Y,\Delta Z$) moves the tool in a straight line at feed rate F in/min to its new position, which is defined by the incremental co-ordinates $\Delta X,\Delta Y,\Delta Z$, with respect to its previous position.

SFR(F) sets the feed rate at F in/min. This value will be in effect until it is changed by another call to SFR. The minimum possible feed rate is 0.25 in/min, which is also the minimum increment of table speed.

PNC(X,Y,Z) punches NC tape in EIA Standard RS-274 format and coded in accordance with EIA Standard RS-358, X,Y,Z are absolute coordinates as in CUT, however, the NC tape that is produced has incremental coordinate data which can be read back when using RNC.

LDR(L) punches L inches of leader/trailer code. If L is negative, a stop code, M00, precedes the leader trailer code.

PSQ(K) adds a sequence number to the next block of tape. This should be followed by PNC or LDR with $-L$ so that the block of data is in a permissible form.

IRS(X,Y,Z) initializes the reference system by defining the origin with respect to the existing tool position. In effect, this routine performs a co-

ordinate transformation on all points in all subsequent calls to CUT and PNC.

RNC(FO, XM, YM, ZM) reads EIA coded NC tape. FO is the feed rate over-ride parameter. If FO equals zero, the feed rate on the tape is used. If FO is greater than 0, then FO is the required feed rate. XM, YM, ZM are the mirror image parameters around the respective axis. In practice, this is one of the most commonly used routines, for many parts are programmed in APT (Automatically Programmed Tools), post-processed for BCLCNC and a command such as

P RNC (0,1,1,1)

is sufficient to initiate cutting. The process may be interrupted at the end of any block without loss of information by setting PDP-8/E switch register bit to 0. In the near future, this function, as well as mirror image options will be selected from a switch panel to be monitored by the digital buffered input interface. Since a teletype is not a suitable display device, sequence numbers, though decoded, are not displayed.

STN(N) is the routine used to set digital buffered output interface channel N .

CLN(N) is the routine which clears digital buffered output channel N .

The following routines return a value to FOCAL-8 and must therefore be used accordingly.

The FRDN(N) function reads digital buffered input channel N and returns a value of 1.0 if the channel is set, or 0.0 if it is not.

FADN(N) is an analogue to digital conversion routine. Converts channel N and returns a value between -1.0 and $+1.0$.

FGSR(N) (Get Switch Register bit N), this routine returns a value of 1.0 if the bit is set, otherwise, the returned value is zero. This enables the use of the SR as a sense switch panel.

COM() (Communications routine). This enables the user to log into a time-sharing system. Thus, it becomes possible to write and edit source programs and to compile and postprocess them directly from the machine site.

INTERNAL ORGANIZATION OF BCLCNC

As mentioned above, BCLCNC resides in the last 4K of the 12K memory. At present, it occupies approximately 3K of this memory stack. All computations in CNC functions are performed in integer arithmetic, usually in double precision. The range of numbers that can be expressed in twelve bits is from -2047 to $+2048$ which limits the maximum increment of motion to little over two inches. With double precision (using two words or a total of 24 bits) this range is extended to approximately 8388 inches. A double precision arithmetic package was coded by simulating the double precision instructions of the Extended Arithmetic Element (EAE) which is a hardware option with the PDP-8/E. Thus, it will only be necessary to reassemble the control routines if and when an EAE is added.

The CNC routines are structured around a set of memory locations that are used to store the co-ordinates of the present point, the coordinates of the

previous position and the increment of motion along the three axes. These locations are common among CUT, MOV, IRS, PNC, and RNC, and therefore, all of them reference the same co-ordinate system. For example, in order to return to the setpoint after a partial cut sequence, the following direct commands would be sufficient

*P IRS(0,0,0) defines the existing tool position as the origin.

*P RNC(0,1,1,1) starts reading NC tape and machining the part.

The operator can interrupt a cut sequence by setting the appropriate switch on the control panel.

*P CUT(0,0,0) Return to the origin or the setpoint. Of course, the setpoint need not be defined as the origin, it could be any point x,y,z then return would be to that point.

After preliminary processing by CUT, MOV or RNC, actual machine tool drive is effected by the same subroutine which also performs three-dimensional linear interpolation. The pulse train is stored in a circular buffer. At every clock interrupt a pulse is sent to the appropriate step motors by the interrupt service routine. Using circular buffers in conjunction with the interrupt system, frees the computer from waiting for the input/output devices. Instead, a device is serviced when it requests attention by causing an interrupt. Ample time is available for interpolation and other necessary calculations because only a minimum amount of time is spent servicing the peripherals. To enable this mode of operation, the FOCAL interrupt processor was extended to field 2 where all the CNC routines are stored. Each peripheral device is assigned a circular buffer space and has a couplet of device handler routines, one of which services the device in the event of an interrupt.

The feed rate is controlled by the interrupt frequency of a programmable real time clock. As mentioned above, at each program interrupt by the clock, a set of pulses is output to the stepping motors, driving the table at the specified feed rate.

MACHINE TOOL CONTROL INTERFACE

An interface was designed to output the necessary control signals to the machine tool. The following instructions were implemented

- (1) set Direction to Positive, four instructions for X,Y,Z,R axes
- (2) set Direction to Negative, four instructions for X,Y,Z,R axes
- (3) step an axis, four instructions for X,Y,Z,R
- (4) enable machine tool interrupt
- (5) disable machine tool interrupt
- (6) skip on machine tool emergency interrupt
- (7) skip if machine tool in computer control mode

All these functions can also be performed by a standard twelve channel input/output interface in which case minor modifications in the CNC routines would be necessitated. A total of four channels were implemented. The fourth channel, R , is reserved for rotary table motion.

EXAMPLE

An example of a short and simple BCLCNC program to cut an ellipse is given below

```

10.01 TYPE 'THIS PROGRAM CUTS AN
      ELLIPSE ON THE X-Y PLANE',!
10.05 ASK 'X-AXIS RADIUS?',XR
10.10 ASK 'Y-AXIS RADIUS?',YR
10.15 ASK 'DEPTH OF CUT?',DP
10.17 ASK 'FEED RATE?',RF
10.19 P SFR(RF)
10.20 SET IC=0
10.21 P IRS(0,0,0)
10.24 SET ZZ=100-DP
10.26 SET TH=0
10.30 SET DT=(360/60)*(3.1415927/180)
10.35 SET IC=IC+1
10.45 SET TH=TH+DT
10.50 SET XX=XR*FSIN(TH)
10.55 SET YY=YR*FCOS(TH)
10.60 P CUT(XX,YY,ZZ)
10.80 IF (IC-60)10.35,10.90,10.90
10.90 P CUT(0,0,0)
10.95 QUIT

```

The program starts by identifying its capability and asks for four pieces of input data. Line 10.19 sets the feed rate, in line 10.20 a counter is set to zero, then the reference system is initialized by defining the existing tool position as the origin. Line 10.24 assumes that the tool is 0-100 inch above the part surface. Line 10.30 calculates an incremental angle by dividing 360 degrees into sixty parts and converting the result to radians. Line 10.35 starts a loop, the counter is incremented, an angle, TH, is updated, and x,y co-ordinates of the next point on the ellipse is calculated. Line 10.10 performs the motion and line 10.80 checks if all the sixty cuts have been completed. The statement of line 10.90 takes the cutter back to the origin and control is transferred to the command mode by the QUIT instruction in line 10.95.

Cutter paths for analytically defined surfaces can be fairly easily programmed in BCLCNC. The technique is especially attractive for complex surfaces, such as a hyperbolic paraboloid, for which a very long NC tape would result if programmed by any of the major part programming languages. It is also possible

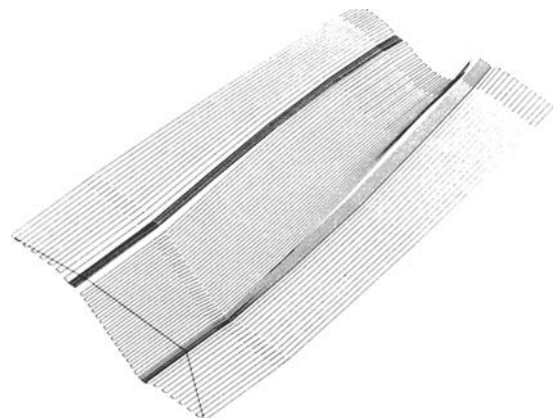


Figure 3. Isometric projection of cutter paths for the EDM electrode of the airfoil portion of a turbine blade die.

to program a set of subroutines in BCLCNC to cut the most common contours and to calculate tool offsets¹.

APT POSTPROCESSOR

In order to be able to program cutter paths in a more comprehensive language, such as APT, a post-processor was coded. Operating on the CLTAPE file (Cutter Location file), this program produces a NC tape coded according to applicable EIA standards. In addition, under user control, projections of the cutter

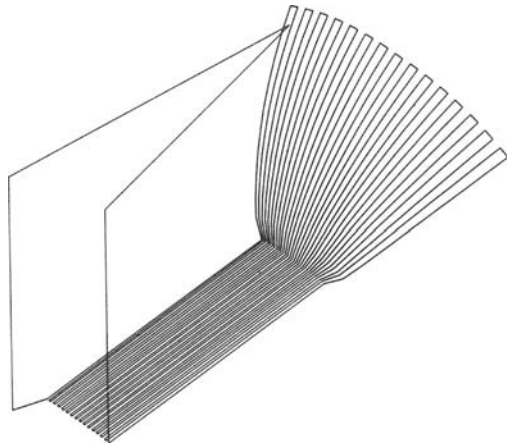


Figure 4. Isometric projection of cutter paths for a streamlined extrusion die for hexagon.

paths to any co-ordinate plane, plus an isometric projection, may be plotted. This three-dimensional display capability proved to be very useful for checking part programs before actual tape tryout. Figure 3 is a three-dimensional illustration of cutter

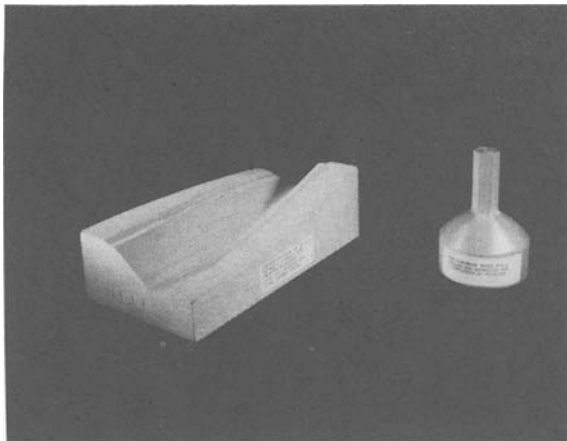


Figure 5. Model of the EDM electrode for the airfoil section of a turbine blade die on the left. Model of the EDM electrode for a streamlined extrusion die for a hexagon on the right. Cutter paths illustrated in figures 3 and 4 respectively.

paths for the EDM electrode to sink the airfoil portion of a turbine blade die. Figure 4 is an isometric projection of cutter paths for the EDM electrode of a streamlined extrusion die. As a circle was being blended into a hexagon, only $\frac{1}{6}$ th of the surface was programmed and the actual part was cut in six passes while it was positioned by a rotary index head. Both parts were machined on the CNC facility and a model of each is shown on figure 5.

CONCLUSION

A CNC system was developed that brings computing power to the shop floor with an interpretive language. The minicomputer adds flexibility to the controller, other peripheral equipment such as a cathode ray tube to display the cutter paths and a mass storage device for part program storage can be added with relative ease.

When cutter paths are programmed in BCLCNC, it is usually possible to write the program such that items having similar geometry can be cut using the same basic program. An NC tape corresponding to an existing BCLCNC program can be prepared by simply replacing calls to CUT with calls to PNC. In installations where manual programming is practiced, this feature would be useful in preparing tapes for other machines. Switches may be assigned functions such as terminate a cut sequence when a given position is reached, repeat a cut sequence, cut mirror images, etc. In this way, a machining operation can be automated in a more flexible manner and to a higher level than is possible with NC.

The increased versatility and the lower cost of minicomputers over conventional numerical controllers provides a more favourable solution to computer aided manufacturing.

ACKNOWLEDGMENTS

Acknowledgments are due to W. F. Scharenberg who supported this development activity by providing the finance, and by leading the project team so competently. The authors also thank Denis Rider who designed and built the computer interface.

REFERENCES

1. J. Frisch. A conversational mode for direct numerical machine tool control. *Proceedings of the 13th International M.T.D.R. Conference* (September 1972).
2. Programming languages. *PDP-8 Handbook Series*. Digital Equipment Corporation (1972).

CLOSED LOOP POSITIONING CONTROL FOR A DNC-MILLING MACHINE

by

A. M. LOSS* and J. FRISCH†

SUMMARY

The open loop positioning devices of a direct numerically controlled milling machine have been converted to provide feedback to the computer for a closed loop control system. The computer program to check and correct table movement and position after each step as well as the generalized point to point operating routine are presented.

NOMENCLATURE

Program command abbreviations

AC	Accumulator
BCD	Binary coded decimal
BCDBIN	Subroutine used by CONVRT for BCD to binary conversion
CONVRT	Service routine for BCD to binary conversion with respect to sign
DT	Delay time between time when computer issues a step command and when a change in BCD output occurs, 10^{-6} sec
EAE	Extended arithmetic element
I/O	Input/output
L	Link
MOVCKX	Table movement stepwise check and correction routine
NS	Prescribed number of steps to be taken \times stepwise \times 1000, inches
P_1	Table position before movement \times 1000, inches
P_2	Table position after movement \times 1000, inches
PT	Program execution time, 10^{-6} sec
R_{nc}	Result of negative check equation, inches
R_{pc}	Result of positive check equation, inches
R_{pnc}	Result of P-MOV check equation, inches
S	Stepsize \times 1000, inches
SAMPLX	Subroutine to place BCD data into computer memory
ST	Required minimum time between steps, 10^{-6} sec
TEDD	TRAV-A-DIAL Electronic Digital Display
UC/DNC/L	University of California Direct Numerical Control Language

INTRODUCTION

The direct numerical control (DNC) of machine tools

is now a well established procedure for on-line computerized management of various metal processing functions. Such an operating procedure invites the closed loop control of process variables and the machine motions were of particular interest with regard to a DNC-milling machine system. This direct numerically controlled machine tool is operated with a PDP-8e Digital Equipment Corporation computer and the programming language^{1,2,4} (UC/DNC/L) has been previously described in detail. The widely used method for determining the motion of machine components which are moved by stepping motors is the direct encoding of motorshaft rotation. Such a method, while most useful, does not give a direct value of the position of the table, cross slide, or knee of the milling machine. The displacements along the three axes is accomplished by stepping motors whose input pulse trains are controlled by two (UC/DNC/L) commands, P-MOV and P-CUT. The first command is used to move the table or other component, at the fastest possible rate during positioning only. The second command P-CUT, given feedrate and coordinate points, will perform a programmed cutting operation relative to a specified and initialized origin. Since these are open loop control schemes, there is no position feedback from the machine to the computer, and therefore the computer does not know whether the movement is the one requested by the control programs. However, the milling machine is equipped with TRAV-A-DIAL sensors manufactured by Southwestern Industries Incorporated as shown in figures 1 and 2. These sensors transmit digital electronic signals to a remote digital TEDD readout device. There the signals are used to drive numerical indicator tubes to display to the operator the incremental linear travel of the table. These displays are used only for initial table positioning. Within the readout device each displayed decimal digit is con-

* FMC Corporation, San Jose, California

† Department of Mechanical Engineering, University of California at Berkeley

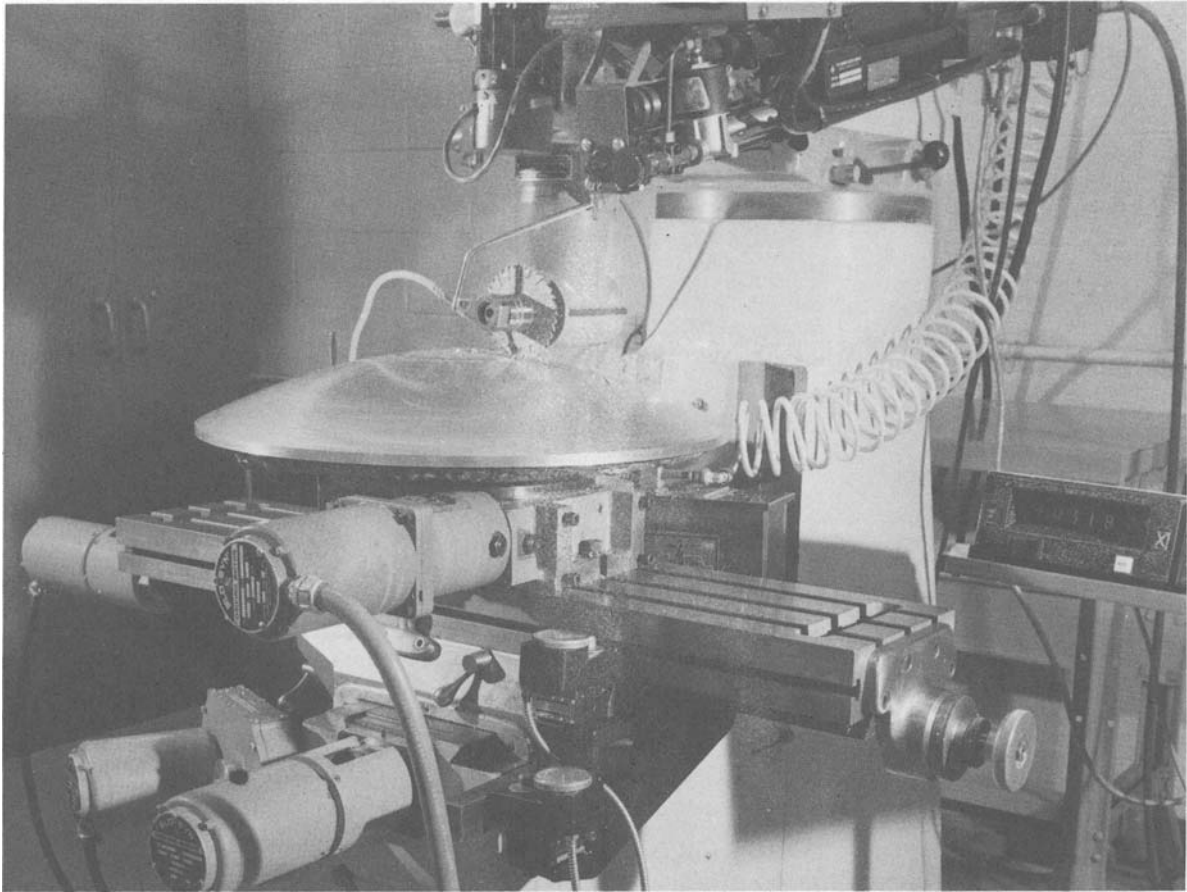


Figure 1. DNC-milling machine with sensors for the table and cross slide.

verted from a base ten numerical system to standard four-line (8-4-2-1) binary coding which can be transmitted over wires for use as desired.

Since this binary coded decimal (BCD) output was available, it was decided to use the TRAV-A-DIAL sensor and TEDD system for feedback of table position to the computer by converting to a closed-loop control scheme.

Overshooting at high feed rates due to inertia loads, backlash in the lead-screws, and lost motion due to the cutting load being greater than the stepping motor capacity, are a few sources of erroneous table movement which cannot be tolerated when dimensional accuracy is required. The proposed closed-loop system shown schematically for a single axis in figure 4 is required for accurate and precise metal removal. A software package in PDP-8e assembly language has been written for such a conversion and all programs are functioning as intended.

The general system for a single axis can be used for both the P-MOV and P-CUT operates in the following manner. First, the initial position coordinate is placed into the computer memory by gating the display system BCD output through a data storage buffer to the computer's accumulator (AC). This value of the axis coordinate consists of two double precision twelve-bit words. A subroutine converts the double precision BCD number to a double precision binary number which is then transferred to the appropriate storage location. The stepping motor then moves one step as commanded. Before the next step command is given, the second location is fed into the computer

memory and a check routine determines if the proper table movement has taken place. If not, an opposite directional step command is given to the stepping motor when overshooting occurs, and a same directional step command when understopping (a skipped step) occurs. Once the desired location is reached, control returns to either the command routine for another stepping command or to the UC/DNC/L control program if the table is in its desired final position.

The foregoing is the description of a closed-loop system which checks after each step command to determine if that step has been taken. This is a necessary requirement when metal is being removed to avoid carrying a stepping error, which occurs at the beginning, through the entire cut. As shown in figure 5, this could easily result in the undesired removal of large amounts of metal.

Because P-MOV is a point to point operation, the general closed-loop scheme had to be modified to allow the entire table movement to be completed before the check routine determines if the table is at its final location as prescribed by the operator's P-MOV instruction. A discussion of the instrumentation and conversion scheme is presented, however the algorithms and program details⁵ are omitted for brevity.

INSTRUMENTATION

The direct numerically controlled Bridgeport milling machine is equipped with sensors and digital readout



Figure 2. Close-up view of TEDD-system sensors on the DNC-milling machine.

units, which are shown in figures 1, 2 and 3 for the table and cross slide, that is the X axis and Y axis respectively. The readout is to 100 inches with a resolution of 0.0005 inches.

The X axis sensors friction wheel movement relative to the table and the Y axis sensors wheel movement relative to the knee. With this arrange-

ment, the X axis sensor remains stationary with respect to the table while the Y axis sensor actually moves with the table as it travels in the Y direction. Figure 2 shows the sensor units mounted on the machine.

The sensor unit is composed of the mechanical TRAV-A-DIAL and an internal rotary encoder

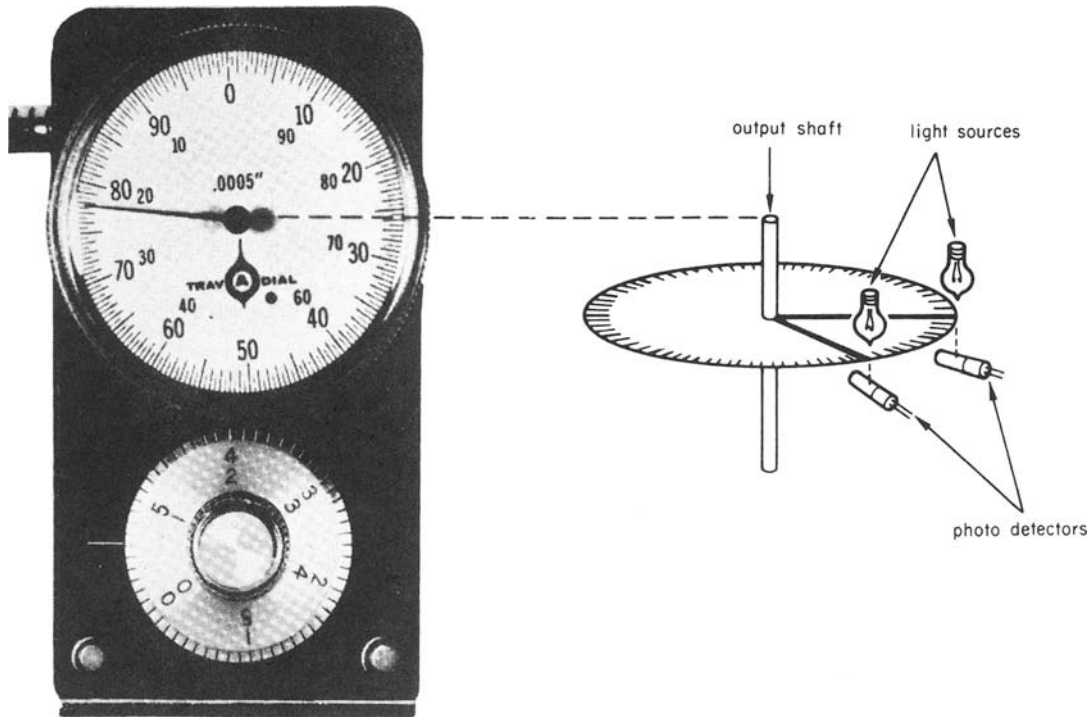


Figure 3. Sensor unit for the TRAV-A-DIAL electronic digital display system.

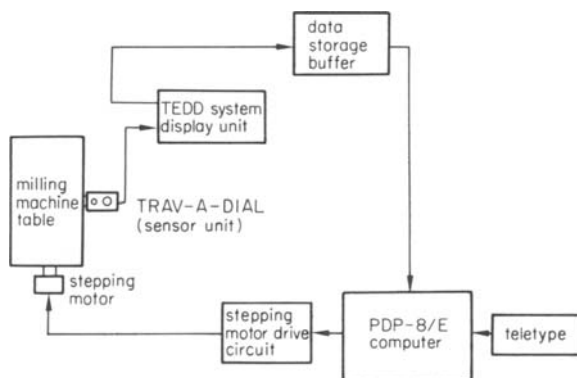


Figure 4. Schematic Diagram of experimental equipment for a single axis closed-loop control.

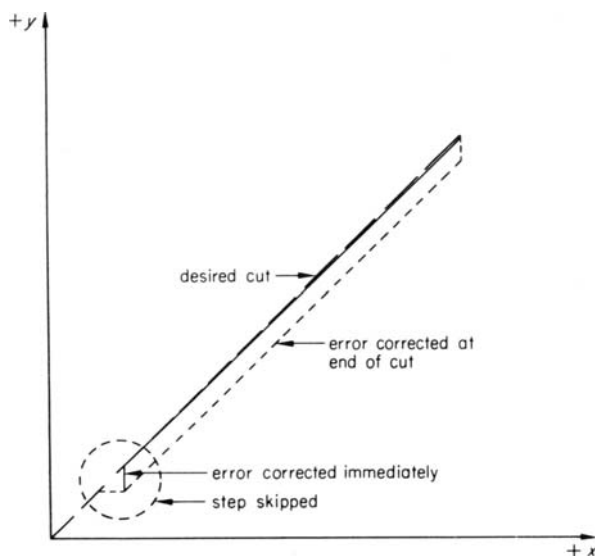


Figure 5. Error correction by two different closed-loop control schemes.

assembly as shown in figure 3. The output shaft on which the dial pointer is mounted rotates in response to circumferential travel of the friction-gauge wheel. The encoder is a photoelectric type and contains a rotating disc driven by the dial pointer shaft. The disc is made from a stabilized plastic film, and is photographically etched with a radial pattern of alternate transparent and opaque segments. The transparent segments allow light to pass from a light source on one side of the disc to a photodetector axially aligned on the other side. Since this is a bi-directional encoder, it has two such photodetector circuits, spaced 90 electrical degrees apart, which provide the means of determining the direction of the rotating disc.

The photodetector conducts current in proportion to the intensity of the light striking its surface. As the disc rotates, a changing voltage waveform is produced at the output of the photodetector circuit. The voltage is maximum when the light intensity is greatest (when the transparent section is between the source and the detector) and minimum when the opposite is the case.

The sinusoidal signals from each channel of the encoder are shaped and amplified in the dual channel preamplifier into square wave signals which drive the circuitry in the display unit via the interconnecting cable. Each signal change represents a discrete movement in the angular position of the disc which corresponds to linear travel of the table. The encoder provides 200 signals per revolution. There is a fixed 90-degree phase difference between the two channels, which is called a quadrature relationship of the signals³. The logic circuitry in the display unit determines the direction of rotation by detecting whether or not the pulse on Channel A is present while the pulse on Channel B occurs. If Channel A is present

while Channel B occurs, then the shaft is rotating clockwise. If absent, then rotation is counter-clockwise.

In the digital display units the dual channel input amplifier and pulse shaping network restores the signals in case any degradation occurs in transit from the sensor through the interconnecting cable and also rejects noise signals which may be picked up in transit.

The direction sensing and multiplying network performs the functions of sensing direction of encoder disc rotation and provides output signals to drive the bi-directional counters. These output signals which drive the counter are in the form of pulses on two lines, one for UP pulses and one for DOWN pulses. For each increment of travel in the positive direction, there occurs one UP pulse, and similarly for any negative travel DOWN pulses are produced. The direction control switch selects the direction of positive count. The bi-directional counter, related to the least significant digit in the display, operates in the following manner. Assuming that the initial setting is zero and that one increment of travel in the selected positive direction occurs, one UP pulse significant digit advances that digit to five. The next UP pulse advances the digit to zero, producing an UP pulse on its counter output line, which drives the second least significant digit to one . . . and so on.

The other bi-directional binary counters are cascaded and each counter stage has two input lines and two output lines. These counters operate as follows, assuming that a counter is initially in the zero stage. If one UP pulse is received the counter will advance to one, if another UP pulse is received the counter will advance to two, and so on up to a count of nine. As the tenth UP pulse is received, the counter goes to zero and a pulse is produced on its UP output line to drive the succeeding counter. If on the other hand one DOWN pulse is received, the counter will decrease by one.

In response to the operator's positioning of the direction select switch, the polarity control circuit sets the polarity display tube to show a positive (+) sign with sensor motion in the selected direction from a zero datum, and a minus (-) sign for the converse. A reset select switch sets all counters to zero. This enables the operator to select the zero datum at any desired location. The slew error sensing circuitry detects whether or not the allowable slew rate (10 in/sec) is exceeded, in which case an indicator light goes on to avoid miscounting.

Binary coded decimal (BCD) output is available in the internal circuitry of the display units. The bi-directional counters are mounted on a printed circuit card. There are seventy-two terminals. The topside ones labelled odd and the bottomside ones even. It is from these terminals that the BCD information is obtained. The BCD output for each axis consists of two twelve-bit words. The 'high' word contains the sign information and the whole number part of the co-ordinate value while the 'low' word contains the fractional (decimal) part. The least significant digit, 0-0005, is dropped because the stepsize of the motors is only 0-001 inches.

The data storage buffer consists of six parts since

the BCD input for each axis will be two twelve-bit words. Each part is subdivided into two twelve-bit components, a latch and a gate. With one input/output transfer (IOT) instruction, all encoders will be read. That is, the double precision BCD output from each axis is stored in the appropriate twelve-bit latch. Another IOT command, one of six, is then used to gate to the computer's accumulator (AC) the particular word needed by the command routine. This command also clears AC first and then transfers the twelve-bit data to the accumulator.

COMPUTER PROGRAMS

The P-MOV and P-CUT routines for direct numerical control of the machine tools are open-loop control schemes, and therefore no feedback from the machine to the computer is available. The commands to the stepping motors are timed by a clock interval to the computer so that enough time is available for the motors to complete one step before the next stepping pulse arrives. Therefore, erroneous table movement could take place and the computer would be unable to compensate for it. The conversion of the available encoder and display system for closed-loop control is accomplished by providing a software package that transfers the BCD output to the AC, converting it to a binary number, checking table movement. If required, take whatever correction action is necessary to properly position the table, or other moving component of the machine tool. The following description of the table (*X* axis) system applies equally well to the *Y* and *Z* axes. The complete program routines for the *X* axis, as well as the stepwise check and correct routines have been listed elsewhere⁵.

To make the control scheme workable, a means of feeding back the precise table position to the computer is necessary. In this case, electrical pulses are produced which correspond to table movement. Depending on direction of movement, these pulses are then added or subtracted in bi-directional counters. From these counters a six-digit BCD number is produced. The least significant digit, 0-0005, is dropped because the stepsize of the stepping motor is 0-001 inches. Since each digit consists of a four-bit binary code, there must be five four-bit words, as well as one one-bit word which means that twenty-one bits of information must be relayed to the computer. Because the accumulator is only twelve-bits and there are twenty-one bits of information required, the previously mentioned data storage buffer is used to gate the BCD output to the computer in the form of two twelve-bit words. The high word contains the sign information and the whole part of the coordinate value, and the low word contains the fractional (decimal) part.

The SAMPLX routine controls the transfer of the *X* axis BCD data to and from the data storage. Thus, it controls the gating of BCD data to the accumulator. Because of a lag between the time when the computer issues a step command and the time when the display system produces a change in the BCD output, a program delay had to be incorporated into the beginning of the subroutine. It is

accomplished by causing the computer to enter a loop in which it must transverse a set number of times before jumping out.

Once the delay is completed, an instruction code will cause the BCD output from the system display unit to be loaded into the appropriate latches. One instruction code gates the high BCD word to the accumulator where it is transferred to the high storage location. Another instruction code gates the low word to the accumulator where it is placed in the low word storage location. Control of the computer is then transferred to the subroutine, CONVRT which is used as a check for the BCD sign and is a general routine for all axes on the machine tool.

The conversion of BCD to binary is accomplished with the BCDBIN subroutine, written for the (EAE) extended arithmetic element of the computer. The routine converts a six-digit BCD number to its equivalent binary number in two computer words. Actually, the computer converts the six-digit BCD number to a binary number without regard for the decimal point. The binary value that is in the computer is the value of the coordinate times 1000. Instead of adding another conversion routine to place the true coordinate value into the computer, the stepsize for all mathematical operations is taken to be 1000 times its true value. In this case it is 1 instead of 0.001.

Figure 5 shows the results of corrective action during metal cutting and the consequences of a delay, relative to excessive metal removal. A stepwise table movement check and correction routine MOVCKX has been written to reduce possible errors during the process.

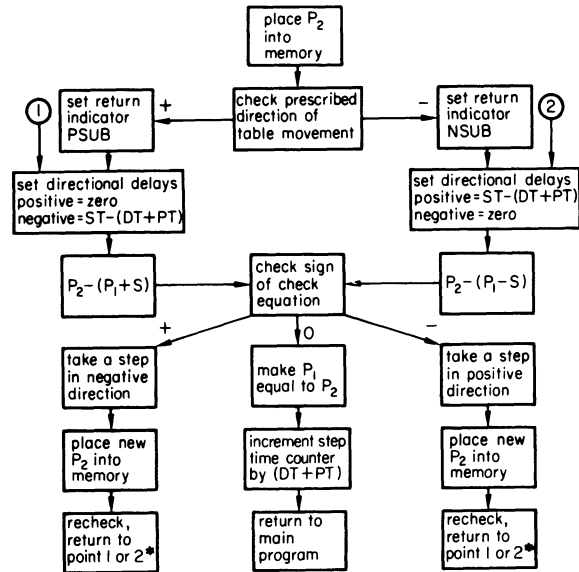
The subroutine MOVCKX is divided into two sections, namely positive and negative directional check logic, which are further subdivided into two parts, directional check equation, and corrective action logic. When the table movement is in the positive direction, the check equation is

$$R_{pc} = P_2 - (P_1 + S)$$

A positive value for R_{pc} indicates that overshooting has occurred, while a negative value means that a step was skipped. If R_{pc} equals zero, then proper table movement has occurred. For negative directional movement, the check equation is: $R_{nc} = P_2 - (P_1 - S)$. In this case a positive value for R_{nc} means that a step has been skipped while a negative value indicates that overshooting has occurred. A zero value implies that the table movement is correct.

It should be noted that when either R_{pc} or R_{nc} is positive, the prescribed corrective action is to take a step in the negative X direction. On the other hand, when either R_{pc} or R_{nc} is negative, the proper corrective action is to take a step in the positive X direction. If R_{pc} or R_{nc} is zero, then no corrective action is required. For this reason a correction action logic routine is used as a service routine for both directional check logic routines, thereby greatly reducing the amount of storage locations required for the MOVCKX subroutine.

The flow chart for the MOVCKX routine is shown in figure 6. After the initial table position co-ordinate



*control returns to point 1 or 2 depending upon which check equation was previously performed.

Figure 6. Flow chart for the stepwise check and correction program, MOVCKX.

is placed into the computer storage locations, and the first step pulse has been sent to the stepping motor, the MOVCKX subroutine begins to perform. First, the SAMPLX subroutine is used to place the new table position into the storage locations. Because the check logic is directional dependent, a check of the prescribed direction of table movement is made, and the proper logic scheme is selected. The check is done by determining if the content of location FLAGX is positive or negative. The lengthy description of this program has been given elsewhere⁵.

Figure 7 is a flow chart of the check and correction routine MOVCKX, modified for the P-MOV command to position the milling machine. To integrate a modified version of the stepwise check and correct routine into the P-MOV routine a complete understanding of the complex internal structure of that routine does not have to be known in order to accomplish such a conversion. Since P-MOV is a point to point service program, the check can be performed after the entire prescribed table motion has been completed. That is to say, a step by step check is not needed. To keep the explanation from getting too cumbersome, only the X axis modification scheme is presented, and the other two axes are converted in the same manner. There are only a few revisions that must be made to the subroutine MOVCKX in order to make it compatible with P-MOV.

It should be pointed out that the operator's P-MOV instruction places a signed integer number, corresponding to the number of steps to be taken, in the designated direction, into the computer memory. The P-MOV routine then converts this number to a double precision binary number. A negative number is represented by the two's complement of the positive. Because the number of steps to be taken is already a signed number, only one check equation is needed. The equation is $R_{pnc} = P_2 - (P + NS)$. Thus the high and low values are substituted. Since only

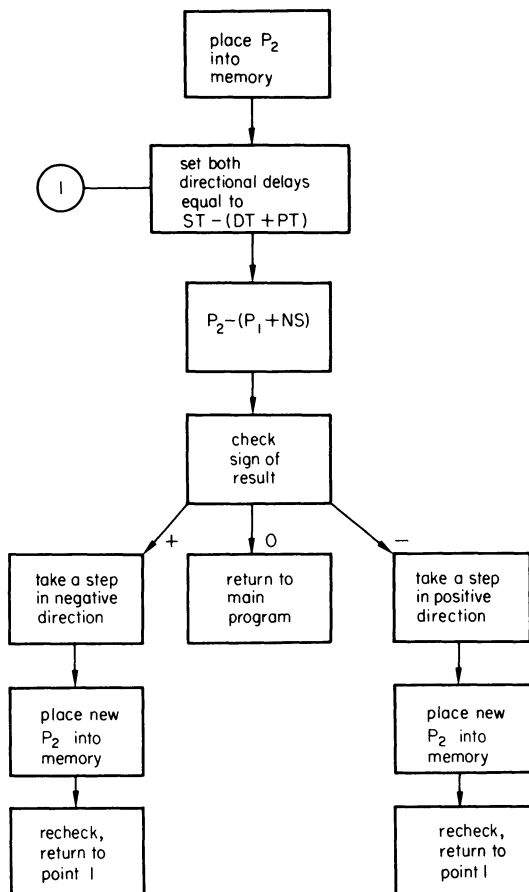


Figure 7. Flow chart for the check and correction routine of program P-MOV.

one check equation is necessary, the routine to determine the prescribed direction of table movement is eliminated. Also, there is no need to set the return indicators.

If a step is skipped, it is not known when it happens during the table movement. Therefore, both directional delays in the corrective action logic scheme are set equal to the minimum time between steps, minus the sum of the sample and program time.

The last modification to the subroutine MOVCKX is the change of the corrective action which is taken when the check equation equals zero. The only performed function is to switch the control of the computer back to the operator's UC/DNC/L control program.

When the modified MOVCKX is patched into P-MOV, two other minor additions to the P-MOV software must be made. First, an instruction is performed at the very beginning of P-MOV to accommodate the MOVCKX and SAMPLX routines in the (EAE) extended arithmetic element. Second, the initial table position is placed into the computer memory using the SAMPLX routine. The SAMPLX subroutine placed the double precision binary value of the co-ordinate into the contents of the high and low output of the BCDBIN routine. Therefore, the contents of these locations must be transferred to the ones for machine tool table location. Since both the CONVRT and the BCDBIN subroutines are general routines for the three axes, the amount of memory core required for those routines need to be provided only once.

DISCUSSION AND CONCLUSION

The closed-loop control software package has several distinct advantages which makes it a versatile routine. First, because it consists of many subroutines, it can be patched into either table control routine, P-CUT or P-MOV in as many separate parts as necessary. For this reason there is no need to provide one large area of the computer core memory for the movement check routine MOVCKX. It can be divided into as many sections as needed and placed into the memory core wherever there is a large enough field available. All programs are as compact as possible in order to keep the total core memory required to a minimum.

The second important advantage is that the check and correct logic is directional dependent and not origin dependent. There is no need to determine the quadrant in which the table is located. This allows the origin to be changed from one P-MOV or P-CUT instruction to the next. Both table control routines make the direction of the prescribed motion easily obtainable. Thus, the table movement check routine becomes even more compatible with them. This in turn simplifies the integration of the check subroutine with the table control routine.

Another advantage is that the check equations permit any step size that is a whole number multiple of 0.001 to be used. Thus, a change in step size does not require any major revision to the MOVCKX subroutine. Also, the double precision BCD output enables the check routine to be effective even if the table size is increased.

A closed-loop control system which provides only mill feedback cannot be used because there is the possibility that a distance greater than one step size could occur during overshooting. Thus, it was necessary to have the exact table position entered into the computer in the form of a six-digit BCD number. Therefore a double precision BCD to binary conversion subroutine is required and all mathematical expressions are carried out in double precision arithmetic. The software programs are therefore longer than required by a single precision routine. This added core memory requirement could be considered a disadvantage. However, the double precision routine does lend itself to a point to point operation where the error in table movement is more likely to be greater than a single step size. This is exactly the case when the stepwise movement routine is modified for P-MOV.

The MOVCKX subroutine is applicable to a control scheme that checks either after each step or at the end of the prescribed table movement. The only major adjustment is a change of the check equations. The corrective action logic for a single step remains the same even though in the later scheme more than a single step correction may be needed. The reason being that after each single step, correction control is returned to the beginning of the directional logic routine. Through recalculation of the check equation using the new table position, it is determined whether an additional corrective step needs to be taken. Since the stepping motor must wait at least 5 milliseconds between steps, and the computer program executor time is approximately 0.5 milli-

seconds plus the sample delay time, there is no need to change the corrective action logic to make more than one step at a time. This would only be necessary if the time between steps is less than the sum of the execution and delay time.

The assembly language table movement command instructions are preset as far as direction is concerned. The positive direction command always causes the table to be driven in a set direction, since polarity cannot be changed. On the other hand, the polarity switch on the system display unit can be used to change the direction of movement that corresponds to positive table displacement. This means that a positive step as prescribed by the table command instruction could be interpreted as a negative one by the encoder-display system, if the polarity switch is in the wrong position. Then the proposed closed loop software package would not operate properly. Thus, the polarity for all directional axis encoders must agree with that of the present milling machine coordinate system. This limitation must be imposed before the closed loop system can be adopted to

operate the direct numerically controlled machine tool with the described encoder-display instrumentation.

REFERENCES

1. J. Frisch. A conversational mode for direct numerical machine tool control. *Proceedings, 13th International Machine Tool Design and Research Conference*. MacMillan Press, Ltd. 161-66, 1973.
2. Introduction to programming. *PDP-8 Handbook Series* (1970) 1, Digital Equipment Corp.
3. *The TEDD-System*. Southwestern Industries Inc. 1971.
4. P. Mahr and J. Frisch. 'On the use of analog-digital conversion during DNC-metal cutting. *Proceedings North American Metal Working Research Conference* (May 1973) 3, 203-16.
5. A. M. Loss and J. Frisch, A closed-loop positioning control system for a direct numerical controlled milling machine. *University of California Engineering Research Report MD-74-1* (Feb. 1974).

ADAPTIVE CONTROL CONSTRAINT AND AUTOMATIC CUT DISTRIBUTION SYSTEM FOR TURNING OPERATIONS

by

E. GIESEKE*

SUMMARY

In this paper a control system will be presented which regulates cutting power according to a predetermined value, and which permits automatic distribution of the cutting volume in individual cuts, using a special cutting distribution plan. The NC program only contains technological data and the description of the finished part contour, so that the shape and size of the rough part need not be known. These can therefore vary as they occur, without unnecessary idle travel. The control system required, which was first presented to the public at IHA 73 in Hanover, consists of an analogue part for power control and a digital part for the automatic cut distribution. The functions of the digital part are handled to a large extent by the mini-computer, which is already available in the CNC system. The savings potential with respect to programming and cutting will be demonstrated in a practical example. The technological basis, the derivation and optimization of the control system, the theoretical considerations which led to this cut distribution system, together with the precise method of operation and applications of the system combined with machine programming systems will not be treated here, but will be described in detail in a progress report¹.

INTRODUCTION

Cutting-power control can be carried out either directly, that is with a power sensor and a power controller, or indirectly with a cutting speed and cutting force controller and the appropriate sensors. The advantage with the indirect method is that in addition to the performance regulation or supervision over the whole turning radius range, a defined cutting speed can be predetermined and a defined cutting force can be adhered to, so that apart from machine supervision a tool supervision is also possible. Whilst the cutting rate control can be obtained relatively easily, the cutting force control loop must meet certain particular requirements which are set by the cutting process on the control system. On the one hand the cutting force control loop must be capable of quick reaction in the event of excessive cutting force, thus fully exploiting the dynamics of the force measuring equipment and the feed drive. On the other hand it should be able to handle cutting force fluctuations which occur during each workpiece revolution caused by out-of-roundness and eccentric workpiece clamping, without feed vacillation. Additionally the control equipment must be capable

of changing over automatically to the most varied types of operation:

- covering of idle travel at full feed rate,
- cut-in on the workpiece,
- controlling the cutting force,
- interrupting the cut when the feed limit is reached,
- exit of the tool from the material.

There are very different methods of automatic cut distribution. In choosing a particular strategy it is important that it should be suited to machining problems, such as cut-in on the workpiece, build up of cutting chips, the most favourable direction of feed and the maximum possible tool usage. It should also involve the least possible time for control programming and require a low equipment outlay.

CUTTING FORCE CONTROL

In planning the cutting force controller the machining process was simulated as a cutting-force control system on the analogue computer and the controller was optimized for varying work conditions¹. It was shown that a P-I-controller whose control parameters are suited to the varying transfer behaviour of

* Laboratorium für Werkzeugmaschinen und Betriebslehre, TH Aachen, West Germany

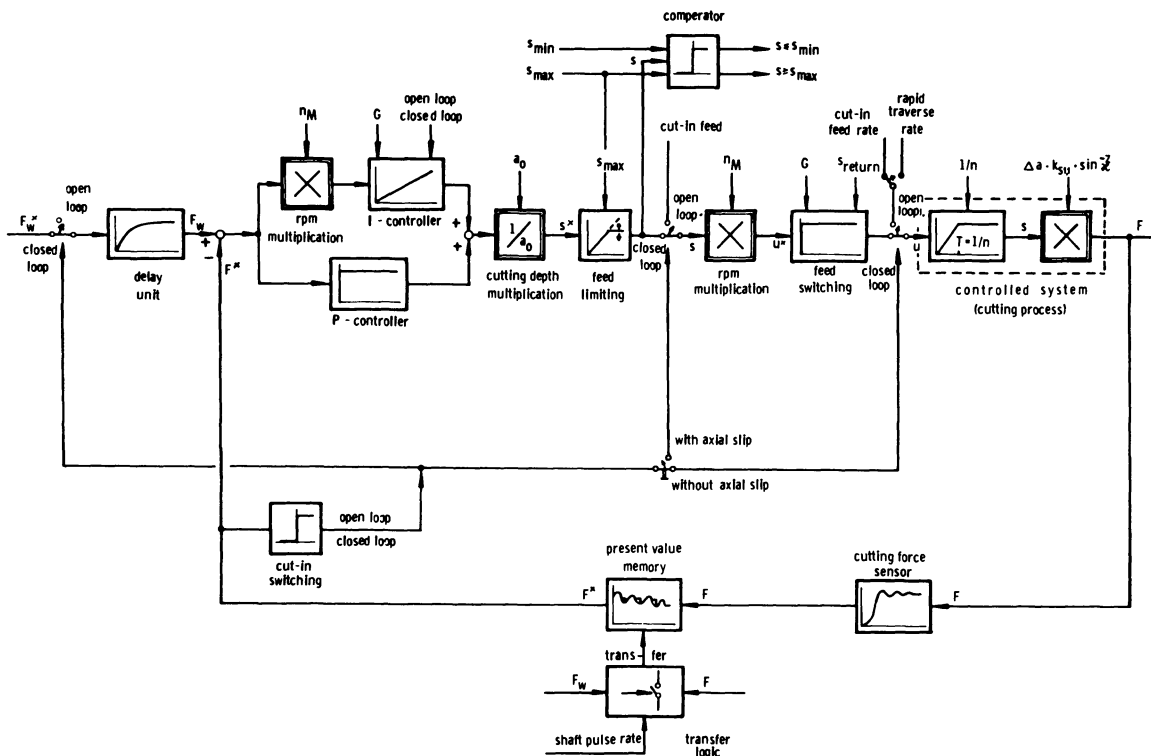


Figure 1. Cutting force control with control of out-of-roundness and cut-in.

the control system, primarily workpiece revolution per minute and cutting depth dependent, produced the best control results. As can also be shown theoretically, the control amplification V_R must be made proportional to the workpiece revolutions per minute and the integration constant K_I proportional to the square of the workpiece revolution per minute. Additionally, both values must fit the mean cutting depth a_0 . The P-I-controller employed must therefore be split into proportional part and integral part and then extended into an adaptive controller—in the sense used in control technology—by application of multiplying and dividing components.

As shown by the block diagram of the control installation in figure 1, it is not the shaft speed n which is multiplied up to the practically laid out controller, but the motor speed n_M because of the wide shaft-speed range. For this, the gear reduction between motor and shaft is accounted for by multiplying the gearing stage G .

The feed limiting component is located between the divider for dividing by the cutting depth a_0 and the multiplier for compensation of the revolutions-per-minute influence. The upper feed limit s_{max} plays an active part, that is if the comparator signal $s \geq s_{max}$ is not evaluated, as is the case in operation without automatic-cut distribution, then the process continues with the feed limited to the maximum value s_{max} . The lower feed limit, on the other hand, only has a passive role: that is, when the lower limit is reached a signal $s \leq s_{min}$ is produced by means of the comparator. The cut is then not continued with the minimum feed s_{min} —in the case of non-automatic cut distribution—but with the feed required to maintain the theoretical cutting force, even if this is less than the minimum feed s_{min} .

The component for multiplication of the shaft or motor speed is followed by a proportional component, with which the feed is raised to the value s_{return} in accordance with the setting angle \mathcal{H} if a cutting-depth reduction is to be made within the sphere of automatic cut distribution.

The proportional component for feed switching leads to the simplified block diagram of the control system. In the feedback branch of the control loop it meets the delay of the cutting force sensor and the symbolic illustration of a cutting force memory device, for which it is not the actual value which is switched through to the controller, but the cutting-force peak value per workpiece revolution, to avoid rev/min-dependent feed vacillation.

For handling idle travel the system is switched over to open loop control by means of cut-in switching which tests for zero value of the cutting force signal. Here three types of control must be differentiated:

- (1) Rapid traverse operation.
- (2) Cut-in at constant feed rate.
- (3) Cut-in at constant feed.

The switching signal 'open loop' simultaneously sets the cutting force theoretical value F_w and the output of the integral control section to zero, so that the controller is not overrun to saturation point.

Immediately after starting the cut-in, the system is switched over from open loop to closed loop operation by increasing the cutting force, whereby the feed s is also set to zero by the initial theoretical value $F_w = 0$. It then increases slowly again with theoretical value F_w which is increased by the delay unit, until the control condition has built up.

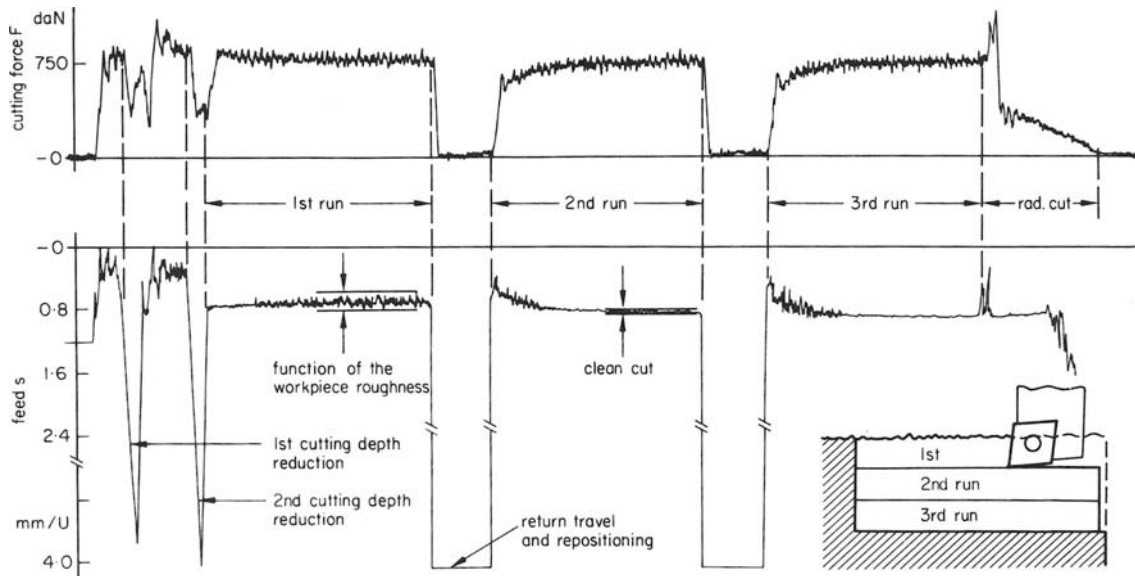


Figure 2. Curves for cutting force and speed in cut-in on the workpiece and in reducing cutting depth.

Figure 2 shows a cutting force sequence as a practical example and the corresponding speed of tool movement in feed, return and repositioning direction, in which several individual cuts determined by the automatic cut distribution can be followed. One can see that in spite of the workpiece cut-in at increased feed rate, no appreciable excess in cutting force occurs and a drop in cutting force is only noticeable during reduction of cutting depth; this is due to the limited return speed.

After the first tool run, the second and third runs are carried out after each cut-in sequence with almost constant feed. The operation in axis direction is followed by an operation in transverse direction to produce the angled contour. Here the cutting force can no longer be maintained at the preset value, because of the low cutting depth and the increased but still limited transverse feed.

AUTOMATIC CUT DISTRIBUTION

For the automatic cut distribution system a plan was adopted for which only the finished part contour and starting-point data are necessary, together with the data for the desired cutting depth. The rough piece is then sensed by the tool itself during the first tool run. This provides a contour in stages which permits handling by an integer number of cuts with full cutting depth and minimized idle travel. A brief explanation of the switching process for changing feed direction and the corresponding feed limits will be given, to describe this method.

As shown in figure 3, the cut-in on the workpiece generally commences at the starting point A on the height of the finished part contour, so that the first NC statement must be programmed from this point. The cut-in itself then takes place automatically, whereby the control equipment ensures a fast approach to the workpiece and speedy cutting force regulation during the cut-in.

If the feed reaches its minimum permitted value s_{min} because of an excessive dimension, the tool—as

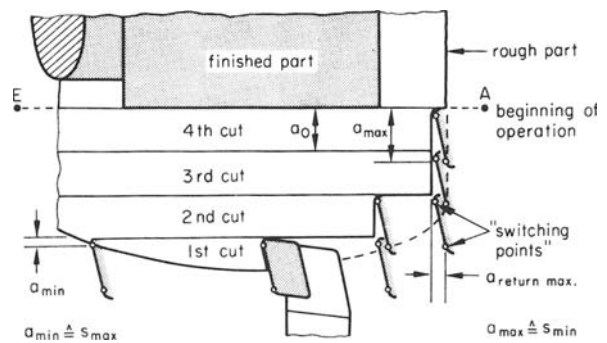


Figure 3. Plan for automatic cut distribution.

demonstrated in figure 3—is removed from the cut, uninfluenced by the remaining rough part contour along the defined distance corresponding to the programmed cutting depth a_0 .

If after completing this phase, the tool is presented with a cutting depth of less than $a_{return\ max}$ —with relation to the return direction—the cut is continued in the main direction of work, shifted by one cutting depth a_0 , until the need arises for another reduction in cutting depth. If, on the contrary, the cutting depth on withdrawal of the tool is equal to or greater than $a_{return\ max}$, the tool is withdrawn again or several times if necessary by a programmed cutting depth a_0 .

In light of the imaginary switching points on the tool cutter (figure 3), the tool path can be followed for a given rough part contour. The feed is suited to the cutting depth varying between a_{max} and a_{min} or $a_{return\ max}$ and $a_{return\ min}$ by the control equipment so that the programmed cutting force can be maintained in each and every case.

If in the main working direction the feed reaches its maximum permitted value s_{max} as a result of the cutting depth a_{min} , the feed is stopped and the cut ceases. The tool then returns immediately to the clearing point at which the cutting depth was last reduced, to carry out the second cut, but this time

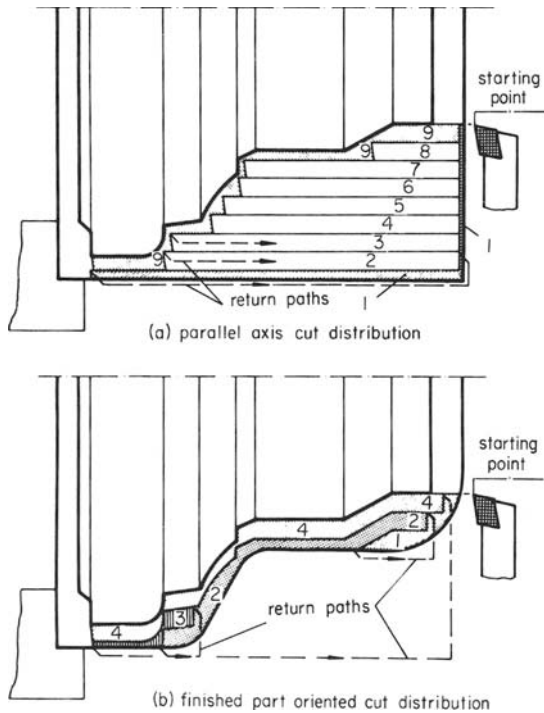


Figure 4. Principles of automatic cut distribution.

with the full, programmed cutting depth. At the end of this cut, the remainder of the first cut is added to it and then deducted with it. Cuts of a very small depth, with which the programmed cutting force or power can no longer be maintained because of the exceeded feed limit, can be avoided from the start.

Two different methods are available for the cutting distribution itself, as there is no one method which is equally suitable for all operating conditions¹. As shown in figure 4, the parallel-axis cut distribution is particularly applicable if a large cutting volume must be handled with many individual cuts, as is often the case with bar material. On the other hand, finished-part oriented cutting distribution is most effective when the rough parts have already

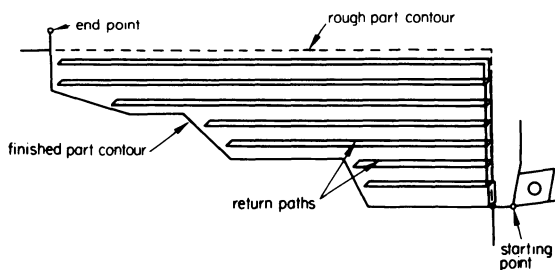


Figure 5. Tool paths determined automatically in handling a section of bar.

been preformed to the point where they can be handled with a few individual cuts.

The following will provide a practical example of parallel-axis cut distribution, carried out on an NC chuck lathe and drawn on an NC drawing board connected in parallel. Figure 5 reproduces the cutting sequence, the return and setting up paths, and the finished-part contour. The rough-part contour was drawn in later. The individual work phases can clearly be seen from the photographs in figure 6.

LAYOUT OF THE SYSTEM

The control and cut distribution system is laid out as a hybrid system, in which the control equipment is in the form of analogue hardware and the cut distribution plan is presented in digital software with a CNC system.

The analogue part for process control is directly connected to the machine, from which it receives information on the actual work situation by means of special measuring sensors. It provides the setting values, shaft speed and feed rate, which are transferred directly to the control unit for the shaft drive or the pulse generator of the CNC system.

Apart from the analogue control part, the limit controlling unit also has two digital counters whose task is to follow the tool position and inform the CNC system of a cut interruption. For this purpose the pulses transmitted to the feed drives in x and z direction are each counted in a binary counter. After a feed stoppage caused by the comparator, the tool position is passed on to the control computer in the form of a bit configuration.

The joint operation of the a.c. equipment with the hardware part of the CNC system and the connections to the CNC and cut distribution software are illustrated in figure 7, in which the hybrid character of the whole system is especially recognizable. In general, the following function groups stand out:

- (1) The CNC unit with pulse generator, parallel registers for transmission of path information, shift registers to provide the control pulses and digital delay to prevent path increment losses.
- (2) The control equipment with measured value input and set value output, by which the pulse generator frequency is controlled.
- (3) The comparator which interrupts the feed pulses immediately in advance of the stepping motor control—that is without being influenced by the digital delay—if the control range is exceeded.
- (4) Position determination which transmits the tool position with relation to the starting point of the sequence, to the cut distribution program after a cut interruption.

The comparator can be overridden by the cut distribution program, if a cut distribution is to be prevented. If the maximum feed is reached because of too low a cutting depth when the tool is located on the finished part contour, then the cut may not be interrupted, but can only be continued with the maximum feed. The signal ' $s \geq s_{max}$ ' is then transmitted to the cut distribution program and—only in the event that the tool is not on the finished part contour—passed on to the gate in the feed drive control lines. If on the other hand the minimum feed is reached, the signal ' $s \leq s_{min}$ ' is transmitted to the gate directly—provided that the cut interruption is authorized—and simultaneously passed on to the cut-distribution program. This only applies, however, to the first cut, with which a stepped contour is cut on the irregular rough-part contour. In making single cuts on the remaining material, the signal ' $s \leq s_{min}$ ' is ignored by closing the AND-gate.

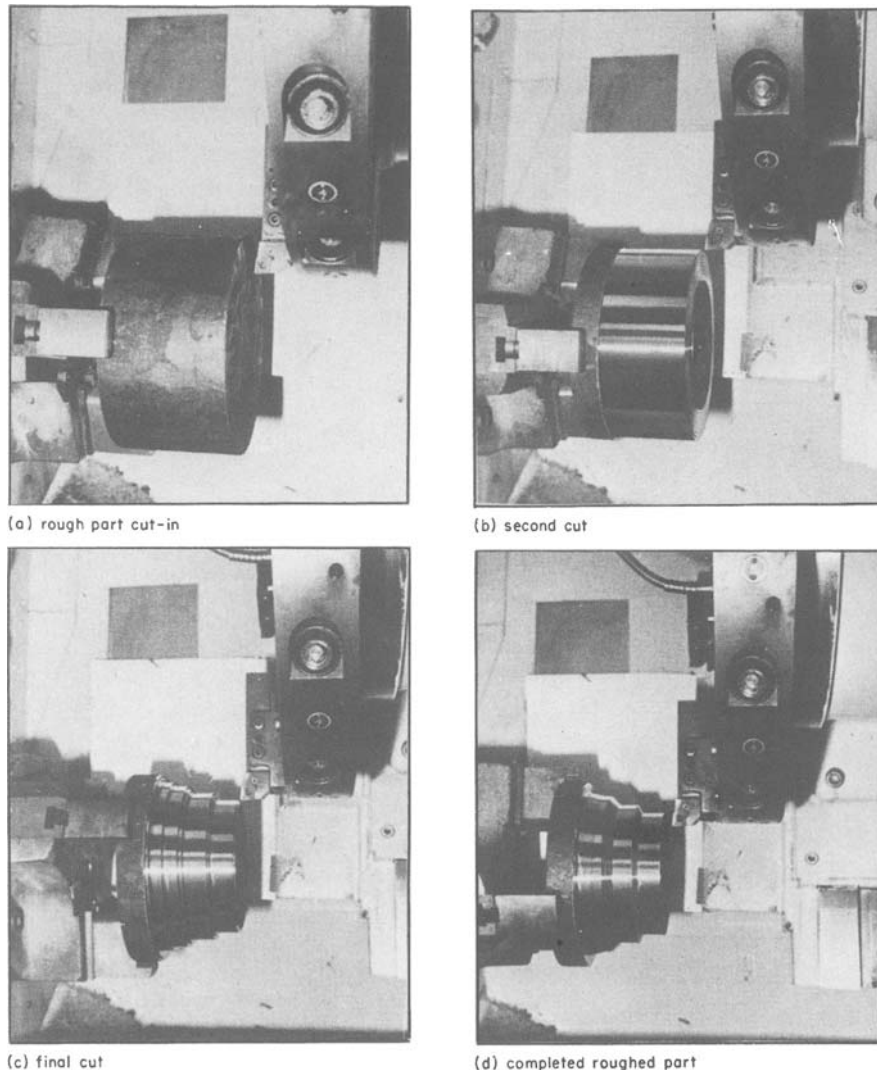


Figure 6. Work phases in handling a section of bar as in figure 5.

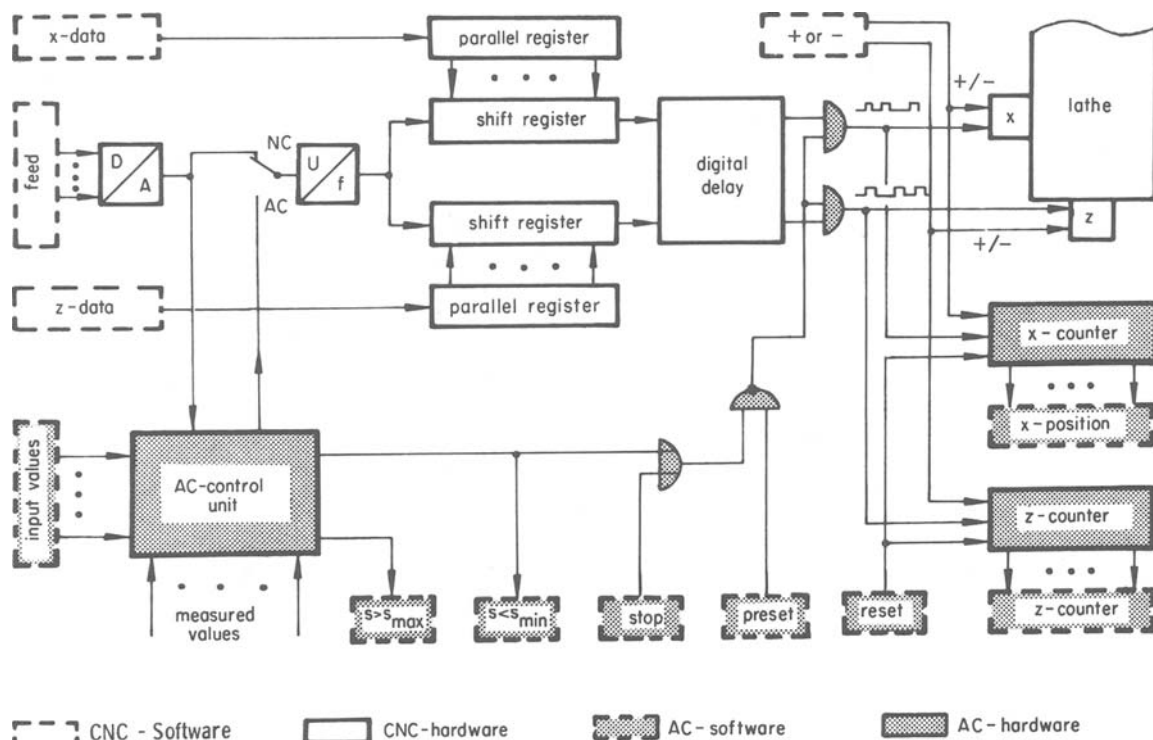


Figure 7. Joint operation of software and hardware of the cutting-force control system and of the automatic cut distribution with the CNC system.

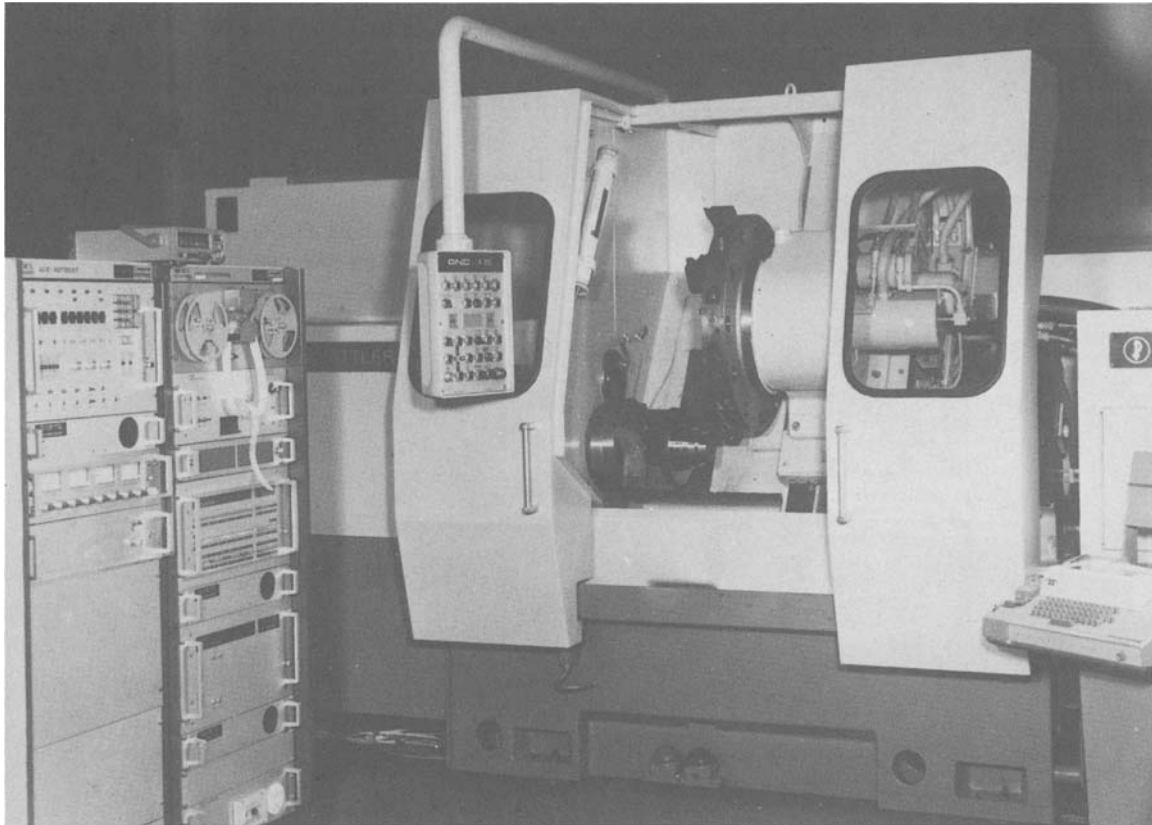


Figure 8. CNC system using power control and automatic cut distribution on an NC automatic chuck lathe.

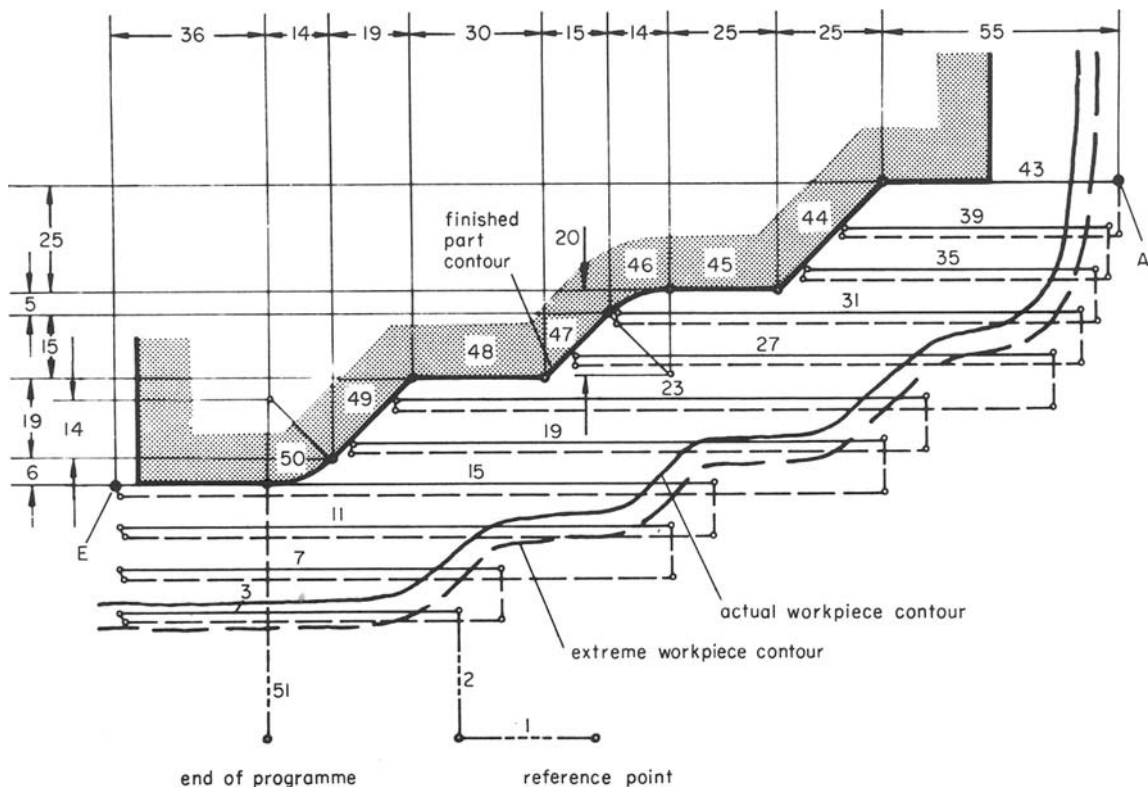


Figure 9. Example of programmed cut distribution.

Figure 8 shows the CNC unit next to the automatic chuck lathe on which the system described was successfully tested and demonstrated at IHA 73. In the unit on the left one can see the instruments to indicate machine load situation, the tool cutting force, the cutting speed and feed as well as manual input controls for influencing the theoretical values. Above these the input and instrument panel for the NC functions can be seen.

REALIZABLE SAVINGS

The following is a comparison between automatic cut distribution and fixed-program cut distribution:

firstly with respect to necessary programming outlay and secondly with regard to the saving in cutting and idle travel.

Figure 9 shows a simple turned part on which a defined area of work can be outlined by providing the final contour from point A to point E. If the cut distribution is programmed in advance, the sequence proceeds with the most extreme rough part contour occurring within a series of parts. The tool paths produced have been drawn in on figure 9 to clarify the individual cutting cycles and are partly provided with the appropriate NC statement number.

As shown in the corresponding programming

```

%
N01 G00 Z-27000
N02 G00 X-17000
S63 T0101 M16
M04 M08
    
```

```

%
N01 G00 Z-11500 U50 W50 D065 F41 S43 M04 M34 M42
N02 G00 X-27000 H44 L33 P48 Q42 T0101 M16 M08
    
```

```

N03 G01 Z-8000 F61
N04 G00 X-250 Z-250
N05 G00 Z-8800
N06 G00 X-1200
N07 G01 Z-9000 F61
N08 G00 X-250 Z-250
N09 G00 Z-12900
N10 G00 X-1200
N11 G01 Z-13000 F61
N12 G00 X-250 Z-250
N13 G00 Z-12800 S65
N14 G00 X-1200
N15 G01 Z-14000 F63
N16 G00 X-250 Z-250
N17 G00 Z-17900
N18 G00 X-1200
N19 G01 Z-12400 F63
N20 G00 X-200
N21 G00 Z-12400 S67
N22 G00 X-1200
N23 G01 Z-12500 F65
N24 G00 X-200
N25 G00 Z-15000
N26 G00 X-1200
N27 G01 Z-10600 F65
N28 G00 X-200
N29 G00 Z-11800 S69
N30 G00 X-1200
N31 G01 Z-10900 F67
N32 G00 X-200
N33 G00 Z-11200
N34 G00 X-1200
N35 G01 Z-6900 F67
N36 G00 X-200
N37 G00 Z-7200 S71
N38 G00 X-1200
N39 G01 Z-6200 F69
N40 G00 X-200
N41 G00 Z-6500 S73
N42 G00 X-1200
N43 G01 Z-5500 F71
N44 G01 X-2500 Z-2500 F69 S71
N45 G01 Z-2500 F67 S69
N46 G03 X-500 Z-1400 I-2000 F67
N47 G01 X-1500 Z-1500 F67
N48 G01 Z-3000 F65 S67
N49 G01 X-1900 Z-1900 F63 S65
N50 G02 X-600 Z-1400 I-1400 K-1400 F63
    
```

programme saving

```

N03 G11 Z-5500
N04 G11 X-2500 Z-2500
N05 G11 Z-2500
N06 G13 X-500 Z-1400 I-2000
N07 G11 X-1500 Z-1500
N08 G11 Z-3000
N09 G11 X-1900 Z-1900
N10 G12 X-600 Z-1400 I-1400 K-1400
N11 G11 Z-3600
N12 G05
    
```

```

N51 G00 X-20000
LF
    
```

```

N13 G00 X-20000
LF
    
```

Figure 10. Comparison between programmed cut distribution (left) and automatic cut distribution (right) for the example in figure 9.

example in the left-hand part of figure 10, the control program consists of initial statements for travelling from reference point to workpiece, the program section which contains the individual cutting cycles, the finished-part contour and the final statements for removing the tool from the workpiece. With the exception of the cutting cycles, these program sections must also be created for automatic cut distribution. As the comparison in figure 10 demonstrates, however, it is precisely the cutting cycle programming which is very lengthy, so that a saving on these NC statements leads to a considerable shortening of the control program. Additionally when the a.c. and cut distribution system for defining a finished part contour is applied as described, the remaining NC statements no longer need to contain the data on shaft speed and feed rate, as these data only have to be established by programming in the program initialization at the start of a work sequence.

Although the comparison of the control programs in figure 10 provides a very persuasive impression of the potential savings, no general figure in percentage can be given as the savings depend decisively on the figure ratio of the individual cutting cycles to the NC statements which describe the final contour. The percentage saving is therefore scattered over a wide range. Nevertheless for the typical example in figure 9, the number of NC statements was reduced from 51 to 13, so that in this case a program saving of 75 per cent is in evidence.

A quantitative comparison of the required cutting and idle travel can also only be reached by using a particular situation. Idle travel which is caused by dimensional extremes and very varied rough part shapes, generally comprises the greatest amount of idle travel occurring. It can be eliminated by automatic cut distribution but it cannot be represented numerically for a general statement.

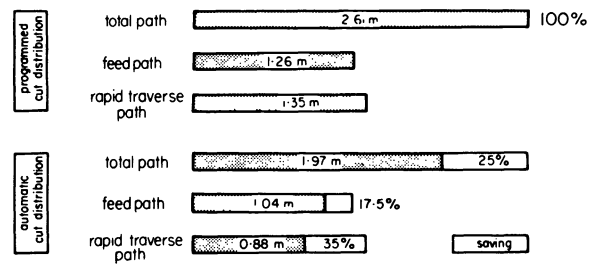


Figure 11. Saving in idle travel using automatic cut distribution for the example in figure 9.

The lesser amount of idle travel which can be saved by sensing the rough part contour as opposed to pre-programmed cut distribution already represents 25 per cent of the total travel in the example in figure 9. As this travel is partly in work feed and partly at rapid traverse, the division of idle travel into feed paths and paths covered at full operating speed is more useful for estimating the production time saved. This division shows for the case handled in figure 9, as can be seen in figure 11, that the first type of idle travel can be reduced by 17.5 per cent and the second type by 35 per cent. The total percentage of production time saved will lie between these two values. The idle travel covered at rapid traverse is less reflected in total production time however, because with the present state of development of feed drives, rapid traverse rates of up to 10 m/min can be reached. Savings in idle travel carried out in work feed, take on a particular significance if an assessment is made of realizable savings in production time.

REFERENCE

1. E. Gieseke. Adaptive Grenzregelung mit selbsttätiger Schnittaufteilung für die Drehbearbeitung, *Fortschrittbericht der VDI-Zeitschriften*, Reihe 2, Nr. 28, 1974.

THE IMPROVEMENT IN ACCURACY OF HIGH PRECISION MACHINE TOOLS BY MEANS OF ADAPTIVE CONTROL

by

I. WASIUKIEWICZ*

SUMMARY

High precision machine tools have almost reached the physical limits of their accuracy. The author analyses the nature of these limitations and discusses some possible solutions to overcome the existing accuracy boundaries. The particular importance of a convenient machine tool geometrical specification is also referred to in this paper.

INTRODUCTION

The primary aim of the machining process is to produce the workpiece within the imposed tolerances. The choice of the appropriate machining procedure is based, in general, on one of the following criteria: minimum production cost, maximum production rate or maximum profit. Whatever the modes of optimization in production, the main objective function remains the geometrical accuracy of the machined workpiece. It is therefore beneficial to search for new ways in which the working tolerance can be maintained when machining at high production rates.

For this reason, the modern jib boring machine was introduced into the general machine shop several years ago. As compared with the other production methods (jigs, fixtures, operator's skills) they are more versatile, more reliable and very often more economical. Particularly important is the reliability aspect, since some machined components are extremely expensive and in most cases corrective machining operations are not possible. The inherent machine accuracy allows very close machining tolerances to be achieved over a large working area within specified operating conditions. From the manufacturing point of view, they represent the highest standards in precision production techniques.

Nevertheless, they have already reached in many cases the physical limitations of their accuracy, which cannot be overcome solely by production and design techniques¹⁰.

Large machine structures are affected considerably in their dimensional stability by temperature and variable load distribution. The initial geometric accuracy of the machine is limited by the precision of the control equipment used during the different stages of the manufacturing process of the machine.

The increase in accuracy of machine tools follows (allowing for some delay) the development of dimensional metrology, which has been quite dramatic during the past 50 years (see figure 1). The boundaries of inherent accuracy must be overcome in order to keep pace with accuracy demands which increase continually. General aspects of growing accuracy demands are not the subject of this paper. It can only be said that the production technique and the technique in general influence each other.

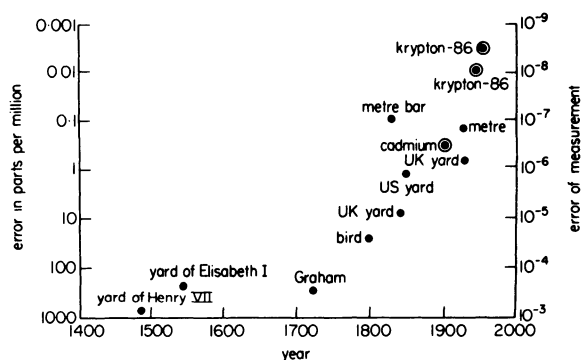


Figure 1. Accuracy of length measurement from 1400 to 1962 (after Professor J. Loxham).

The increased accuracy of boring machines was once necessary to manufacture the first steam engine, which marks the beginning of the tremendous technical progress of the industrial era⁹. The state of the art must advance to meet the state of demand. But in turn, the new state of the art can create a new state of demand and this eternal reciprocal action is a distinctive trait of mankind's development. It is also the natural desire of the production engineer to improve the ultimate accuracy of the manufacturing process and to achieve the complete elimination of

* DIXI II, Le Locle, Switzerland

scrap. Thus, the machine accuracy and its reliability will remain one of the major preoccupations of research and design engineers in this industry.

Most of the following considerations apply directly to the NC horizontal boring and milling machine, but they are to some extent also valid for other high precision machine tools.

THE ACTUAL STATE OF A METROLOGICAL MACHINE TOOL SPECIFICATION

The evaluation of the accuracy of NC machines

Visitors to the most recent international machine tool exhibition in Hanover will have noticed 'the sudden accuracy improvement' of the NC machine tools shown there. Makers' official specifications claimed positioning accuracies of 0.005–0.010 mm. There was great confusion at this exhibition, since no further particulars of the accuracy of the machines were available. This was not always just due to selling tactics. Some companies' representatives were unable to furnish such particulars because they were not in existence. A close look at these machines revealed that their measuring systems only were changed. Most of them were fitted this time with inductosyns.

The machining accuracy is the result of many factors² and the improvement of only one component of the system accuracy cannot influence to any degree the quality of machine tool performance.

The origins of the confusion are manifold. The NC systems are regarded as a simple sum of the machine and the NC control which is often furnished by specialized electronics firms. There are two different working groups in national and international standard organizations, one for machine tools in general, the other for NC machines. Finally, in factories we have different teams working in almost complete isolation from each other. The programmers, who ought to have a good knowledge of each NC machine, the existing tooling and organization imperatives, have seldom any knowledge of machine tool accuracy. This was once the duty of the machine acceptance commission.

The practice of machine tool testing is based on Schlessinger's work and conforms to the ISO R230 recommendation. Some national recommendations exist for NC machines.

It is interesting to examine the kind of metrological data included in the ISO standards. As a general rule, they specify the tolerance as a maximum admissible error for a given range of measuring parameter.

Thus, it is really a worst case approach, but unfortunately only for a limited set of measurements and very restrained testing conditions. The basic aim of any accuracy evaluation standard is to obtain a quality assurance of the finished component. The ISO standards do not fulfil this condition—particularly for high precision machines. Therefore, they can be considered only as a means for the classification of machine tools (see figure 2).

The problem is very acute for three-dimensional work, since existing methods are almost entirely restricted to one and two dimensions. After

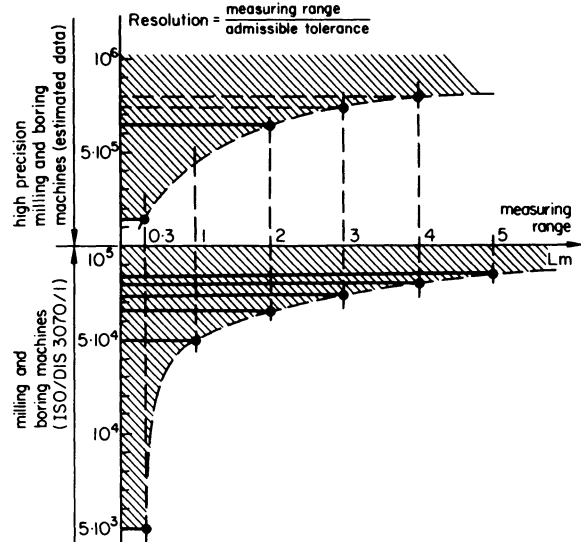


Figure 2. Machine tool dimensional resolution with respect to the straightness of the table motion in X -axis (X - Z plane) of horizontal boring machines.

Hemingray⁵, the medium precision machine with an alignment tolerance of 0.015 mm influenced by other sources of error can produce in the worst case a positioning error of ± 0.120 mm across the diagonal of a 500 mm cubic working zone.

It is evident that enlarging only the existing ISO tolerance set by the two- and three-dimensional accuracy characteristics cannot enrich our knowledge of the true machine performance. There is a need for the detailed machine tool geometrical specification to be produced in a convenient form. The importance of the work of Love and Scarr⁸ was to show that it is possible to obtain almost all the necessary information by means of the conventional ISO testing technique. Further work is necessary to establish a suitable data format.

Now, let us look at the actual state of the NC standards. The itemized data, as defined by draft ISO/DIS 2539, should include the additional data needed for preparing the work and selecting the right machine. It is thought to serve only as a guidance for NC users; nevertheless, the typical example shown there illustrates the kind of mechanical data of the machine to be provided for the overall specification of the system. It includes the maximum machine dimensions, spindle horse-power, maximum work-table load, programming increment, etc. Surprisingly, there is no mention of the machine tool accuracy. The draft states that the amount of data needed depends mainly on the type of work to be performed and on the kind of machines between which interchangeability is sought.

The true interchangeability of the two systems means full compatibility of their complete data sets. The machine tool accuracy specification has to be regarded as one of the most important sub-sets in the general specification of the system.

We have not only to compare the dimensions or the motor powers of the considered machines (initial choice) but also their precision (final choice of machine compatible with the type of work to be performed). Unfortunately, up to the present time,

there does not exist any complete accuracy specification in a convenient form. The existing recommendations VDI 3524 (1965) and NMTBA (1968) are confined to those elements of the machine tool accuracy set which have their origin in the control system. Any measurement is taken along one line and not under a fully specified machine environment state. It was found⁴ that both recommendations do not give an adequate picture of the positioning accuracy (progressive errors, lost motion, cyclic error, etc.) along the considered axis. Therefore, it is not possible to determine the likely workpiece accuracy on the basis of such an incomplete data sub-set.

Numerous attempts were made recently to extend the existing standards and to create new specifications. Special attention was given to the development of a practical testing system which would enable the resultant accuracy of the whole working zone of the machine to be established under different machine environment conditions.

The linear system proposed by Tlustý at the last International MTDR conference in Birmingham seems to offer real progress in the testing of the NC machines. It represents a good compromise between the inspection costs and the method accuracy as a result of the selection of the minimum basic set of measurements.

Nevertheless, all these systems are based on the worst case definition along one line or a set of parallel lines in the working zone of the machine. It is improbable that the workpieces operated by the machine tool user contain any critical dimension, as specified by this method. Thus, it would be very expensive to choose the machine tool for a given type of work by means of the worst case approach. It is a long tradition in the manufacture of precision scales that each scale is provided with an accuracy certificate. The relative position of every division line is measured with respect to the master scale and the error recorded in the form of a correction table. The same approach can be applied to the specification of the machine tool error matrix. The central role will be played by a small computer in conjunction with some auxiliary storage devices. It will perform the following functions:

- (1) Data collection. By means of the special program, a semi-automatic test of the machine

would be performed. Data from electronic gauges are sampled through the interface device, then reduced, translated and finally stored on the magnetic tape.

- (2) Data retrieval or the semi-automatic machine tool selection for a given workpiece with regard to the likely machining accuracy. Both features can be very useful in operation, both with standard and with NC machine tools. The stored machine tool error matrix has to be provided by the machine tool manufacturer. It can, nevertheless, be utilized by the machine user to take into account the wear of the guideways, etc.
- (3) Automatic input data correction for the CNC systems. At first, such an approach seems to be very expensive because of the extended data collection program. In practice, it would be sufficient to check carefully few machines of the same type and to establish the initial error matrix, than to test each machine for a limited basic set of the critical measurements, just to ensure that they are within a prescribed tolerance for this type of machine. Important deviations are to be included in the final machine tool error matrix.

THE CONCEPT OF MACHINE TOOL AS A SIGNAL STORAGE AND TRANSMISSION SYSTEM

The machine tool structure and kinematic chains can be considered as a special kind of main geometrical data, MGD, permanent memory. Machine tool testing (as described earlier) is a means of obtaining the complementary machine geometrical identification data, CGD, which can be recorded on a convenient data carrier. The NC enlarges the permanently stored main geometrical data as it contains some sub-programs, e.g. interpolation, but only special CNC systems can effectively integrate complementary geometrical identification data. This is a completely new aspect of manufacturing systems, which were generally considered as transmitting and translating systems only^{11,13}.

An active physical system can transmit, store, translate, create and exchange (with other systems)

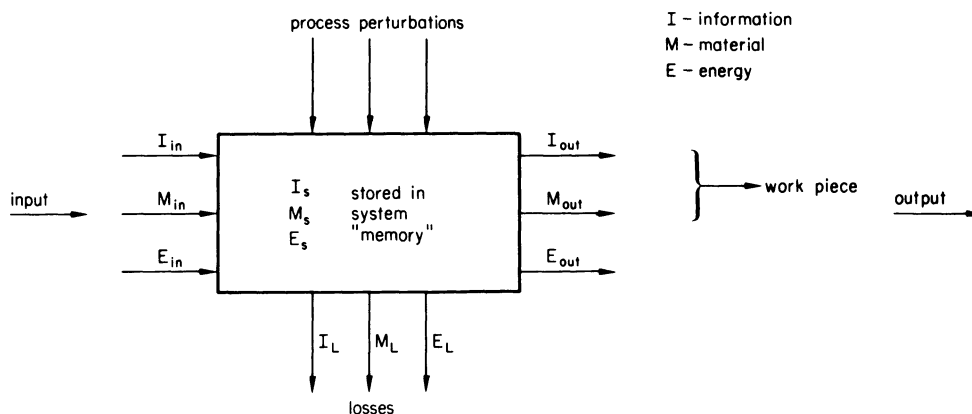


Figure 3. General concept of a manufacturing system.

data. The same applies to the energy and the material, as shown in figure 3.

In the manufacturing system one can distinguish two principal processes, namely, the forming process = data transformation and the cutting process = energy transformation which result in the storage of the input geometrical data in the workpiece. Both processes are inseparable and mutually influence each other¹⁶. Every physical energy transformation or transmission is connected with energy losses in the form of heat, which partially accumulates in the machine tool structure. The variation of the machine tool inner energy level causes geometrical distortion of the machine and thus changes the content of the main geometrical data in the permanent memory. The workpiece can thus be out of tolerance. The heat transmission and its accumulation in the machine structure is therefore the main limiting factor with regard to accuracy because the principal geometrical information is also stored in the same structure. This also applies to the machine distortion caused by a moving workpiece or by cutting forces. The wear of the guideways can also be interpreted as an irreversible, slowly progressing geometrical data loss (the data leakage in the machine basic information storage). To restore the inherent information content of the machine, it would be necessary to rescrape the guideways and to realign the machine to meet the initial accuracy requirements.

The information stored in the machine structure and the kinematic chains represent mainly continuous data. Thus, the machine absolute information content and the absolute entropy are infinite. During the machining process, the machine tool is emitting (generating) and transmitting signals—it is acting as an information channel. We wish to transmit a random variable XE over a channel. The message recovered by the receiver (workpiece) will be a continuous random variable XR . If the channels were noise-free, the received value XR will uniquely determine the transmitted value XE . But channel noise introduces a degree of uncertainty regarding the true value of XE .

The average information transmitted over the channel, called mutual information $I(XE:XR)$, is equal to the average information emitted by the source $H(XE)$ minus the average information loss over the channel $H(XE/XR)$.

$$I(XE:XR) = H(XE) - H(XE/XR)$$

The maximum amount of information that can be transmitted on average per value transmitted is called channel capacity CV .

$$CV = \max. I(XE:XR)$$

If the channel allows transmission of N values per second, then C , the channel capacity per second, is given by:

$$C = N \times CV \text{ bits/s}$$

It should be noted that C represents a finite value for all physically realizable channels. For a given channel, $I(XE:XR)$ is a function of the input probability density $p(x)$ alone. A channel means not only the

physical path for the propagation of messages, but also includes specifications for the kind of signals (continuous, binary, etc.) and the kind of receiver used (the receiver determines error probability). All these specifications are included in the channel matrix, which completely specifies a given channel. If we use binary symbols at the input (NC) the maximum value of $H(XE)$ is one bit (per binit) occurring when $p(0) = p(1) = 0.5$. Hence, for a binary channel $CV \leq 1$ bit per binit. There is always the possibility of converting a continuous message into a digital message by using an approximation technique known as quantization. A signal $XE(t)$ band limited to B Hz is completely specified by its samples taken at intervals $1/2B$ (or less) seconds apart⁷.

Thus, the problem of transmitting a continuous message reduces to that of transmitting a finite number of values per second. We reconstruct the data at the receiver (workpiece) to obtain the continuous signal. This is the basic idea of the NC. The reconstructed signal at the receiver will only be an approximation of the original signal because of quantization error and detection error.

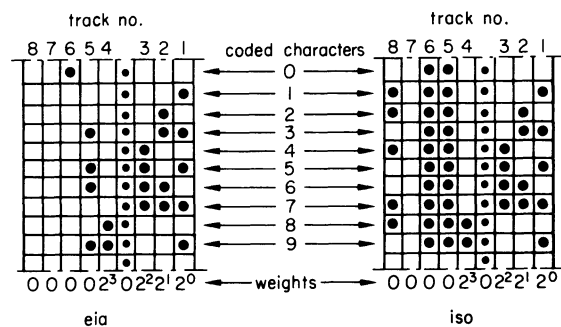


Figure 4. EIA code against ISO code.

The NC machine tool represents the hybrid, analogue-discrete communication system. Thus, the fidelity of the information transmission depends both upon the signal-to-noise ratio and the probability of error in making the decision at the receiver. For high precision NC machines, the dimension quantum is 0.001 mm. The detecting errors along one axis are at best 0.002–0.006 mm. The detecting errors in the machine working space can be as high as 0.030 mm.

The probability of the detection errors due to the data encoding is in the order of 10^{-4} to 10^{-7} . We shall therefore assume that the receiver (e.g. tape reader) will, at most, err in recognizing one of the digits in any character transmitted.

Two codes are commonly in use, namely, EIA-R.S. 244 and ISO/R 840 (see figure 4). Both apply to the 8-track punched tape. Thus, they can provide a set of $2^8 = 256$ different signals, which can represent 256 independent characters. Only some 50 characters are necessary for NC. The EIA coding system operates with the 6-bit words. Track five is an odd-parity check and track eight indicates the end of the block. The ISO makes use of 7 tracks, track eight serving for a parity check only. The (odd) parity check track enables single-error detecting in both systems because the minimum Hamming distance between any two coded characters on the punched tape is two.

SOME PERTURBATION MECHANISMS AFFECTING THE MACHINING ACCURACY

An elementary manufacturing system consists of the machine, the tool, the workpiece, the main manufacturing process and some disturbing processes^{1,3}. For the better understanding of a complex system structure, it is useful to consider separately each of the physical sub-systems which together determine

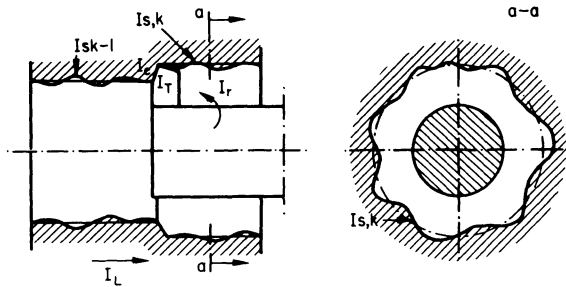


Figure 5. The cylindrical surface generation with the single-point tool.

the system performance. This approach can be illustrated by the boring operation (see figure 5). The hole position and its form can be considered as information stored in the workpiece I_s, k during the k th tool pass. It follows from the earlier discussion about the channel capacity that the received information I_s, k is equal to that emitted by source IP (program) minus the information loss over the channel I loss. I_s, k results from the information contributions of the following factors

- I_L, I_r = linear and rotation movements of the machine
- I_t = tool geometry
- I_c = cutting process
- I_m = workpiece and tool materials
- I_d = disturbances
- $I_s, k - 1$ = previous hole configuration

$$I_s, k = I_p - I_{loss} = f(I_L, I_r, I_t, I_c, I_m, I_d, I_s, k - 1).$$

It is very important to bear in mind that the inherent machine tool inaccuracy can represent a major error contribution in the machining of small components. The machine tool accuracy sets the absolute limits to the machining accuracy. As discussed earlier, the special CNC systems will facilitate the improvement of machine tool accuracy by means of the stored machine tool actual error matrix, but only for a limited class of geometrical errors. It is almost impossible, e.g. to improve the inherent spindle rotation accuracy. Tool geometry, cutting process stability, and materials, often play a decisive role in the high precision manufacturing processes, but they are not the subject of this paper. The relative importance of main process perturbations can be measured by their effects on the workpiece accuracy². They can affect: the dimensions D , the form F , the relative position of elements P and the surface quality S of the workpiece (see figure 6).

The thermal deformation of the machine affects its squareness, straightness, flatness and angular accuracy. It represents therefore a major challenge for

the design engineers. Professor Opitz once said: 'Efforts to control thermal effect errors should, first of all, begin with a suitable design. All heat sources other than those directly connected with the cutting process should be installed outside the work area.'³ The importance of the general machine concept is illustrated by figure 7. The electric motors and spindle bearings are the principal heat generators in the machine. Thus, the spindle and the headstock are mostly affected. Because of symmetric configuration, the horizontal spindle shows very small angular errors (0.002 mm/300 mm) in both horizontal and vertical planes. A parallel thermal drift error in the vertical plane can be eliminated by compensation or program correction. As the feed movement is carried out by the table, the form is practically not influenced by the thermal displacement of the spindle. If the feed movement is carried out by the boring spindle, both the form and the position of the hole are affected. A

factors influencing machining accuracy	D	F	P	S
geometrical inaccuracy of machine tool	●	●	●	●
thermal effects (internal sources of heat)	○	○	●	
weight effects	○	●	●	
foundation		○	○	●
tool	●	●	●	●
forced vibration	○	○		●
environment	●	●	●	○

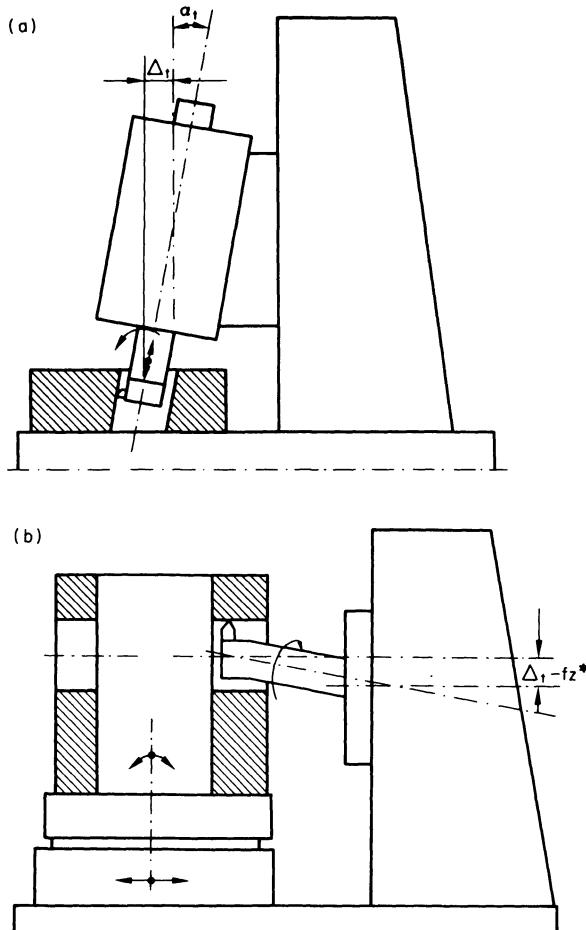
- basically influenced
- sometimes influenced

Figure 6. Some factors influencing the accuracy of the work piece.

special design of the horizontal machine main drive enables the motor to be fixed directly on the foundation outside the machine; thus, the motor heat and motor vibrations cannot affect the machine tool accuracy. It is almost impossible to do the same with the vertical jig boring machine.

A cooling system is necessary for both machine types to compensate the heat generated by the main and feed drives. But as Bryan puts it well³: 'The current practice of using constant-volume pumps and relief valves to regulate pressure is an engineering disgrace. The wasted pumping power is bad enough, but the cost of re-cooling the oil by refrigeration is unacceptable.'

Electrical motors result in a better efficiency of energy conversion, but they still need cooling systems. Fans very often blow the heat directly on to



*(see fig. 8)

Figure 7. Thermal deformations of the headstock and their effects on work piece accuracy: (a) vertical jig borer and (b) horizontal jig borer.

the machine and they also generate considerable noise.

Industry is now looking for pressure regulated pumps with low noise level or d.c. motors of special design which allow cooling to take place with a liquid, e.g. oil. Cooling will remain important as long as there are energy losses inside the machine. It will also be necessary to keep the machine as close as possible at the variable ambient temperature. Nevertheless, it is not possible to achieve absolute thermal stability of the machine because of random time and space energy distribution due to the variable machining programme. The adaptive control allows further reduction of the thermal residual errors.

A second important class of errors has its origins in the elastic deformation of the machine due to the variable time and space force distribution in the machine. With a suitable design the (own) weight and moving part clamping effects may be considerably reduced. However, the deformation of the machine structure due to the moving workpiece can still represent an important problem. The common practice of correcting the guideways form for half of the rated workpiece weight does not reduce the error band, but only changes its relative position². The clamping of the workpiece on the table may also be an important source of errors. The variable weight of

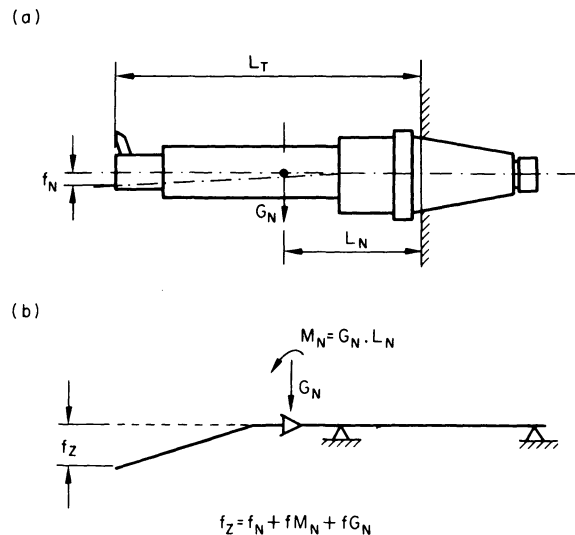


Figure 8. The deflection of the spindle rotation axis because of the tool weight. (a) Additional tool specification. (b) Calculation of the resulting error.

the boring bars causes the displacement of the tool rotation axis. Their effects can be eliminated by the simple program correction (see figure 8). Various compensation techniques are used to counterbalance the effects of the elastic machine deformation^{1,3}.

PERFORMANCE EVALUATION OF HIGH PRECISION MACHINES

It is common practice in the evaluation of manufacturing systems to firstly fix a set of system objectives and then to compare each attribute of the considered system with the corresponding objective^{1,3}. An approach presented by Warnecke and Scharf at the last international MTDR Conference in Manchester suggests the use of a single performance index for the comparison of alternative systems. The original set of system objectives consists of some main objectives, the rest being of secondary importance. Therefore, such a dramatic data reduction from the original objectives set into a single performance index incurs the danger of obscuring the basic system tasks.

The main system objectives can be: (1) maximum accuracy and reliability = the best production quality $Q \max$; (2) Maximum productivity $P \max$ and (3) Minimum production costs $C \min$. The modern high precision machines have to fulfil two requirements at once ($Q \max P \max$) or ($Q \max C \min$). Some machine users expect even ($Q \max P \max C \min$). The joint probability of two events $Q \max P \max$ occurring simultaneously is given by the following equation:

$$P(Q \max P \max) = P(Q \max) \times P(P \max/Q \max)$$

$P(P \max/Q \max)$ represents the conditional probability of $P \max$ given that the event $Q \max$ has occurred.

The problem of performance evaluation shifts therefore from the deterministic area into a probabilistic one. Peklenik and Kwiatkowski insist that a manufacturing system has to be treated predominantly as a statistical system^{6,11}. Unfortunately, the statistical dynamics of the manufacturing

process were not yet sufficiently investigated. However, for practical purposes it is still necessary to operate with the approximate performance functions which assure a given precision between the abstract and real system behaviour. The current practice in high precision machine evaluation consists of machining the customer test piece. The price of the machine, the accuracy of the finished work piece and the manufacturing time are generally considered as the main factors influencing the final choice of the machine.

The machine user is not interested in its abstract peak accuracy, as specified by the test code ISO R230. His real preoccupation is merely to get the assurance that all work pieces will be within the desired tolerances. It is evident that the work piece tolerances set the absolute limit to the machine production rate. Modern jig boring machines shift these limits to the truly physical limits which can be derived from the first law of uncertainty as follows:

$$\text{machining accuracy} \times \text{metal removal rate} = \text{constant} = \text{absolute figure of merit}$$

Therefore only machines which operate close to the absolute figure of merit can be considered as true high performance machine tools.

THE ADAPTIVE CONTROL OF MACHINE TOOLS

Because of its complex nature, it is sometimes difficult to distinguish whether a system is an adaptive control system or a variation thereof. Eveleigh defines an adaptive system as one which measures its performance relative to a given index of performance and modifies its parameters to approach an optimum set of values. Three basic functions performed by the

adaptive controller are: identification, decision and modification. The identification of the system dynamics is by far the most difficult of these functions. For practical applications, partial identification is necessary, because the dynamics of the manufacturing systems have not yet been fully explored. Still, there is a need for suitable sensors which can work reliably under industrial conditions, but have no side effects on the machine performance. The same applies to the development of convenient actuating devices. A further development of software is also necessary. Optimization of the manufacturing process is not possible without its reliable mathematical model as well as clearly defined optimization strategies. Hence, the importance of establishing the right indices of system performance. Spur¹³ and Peklenik¹¹ distinguish two basically different adaptive control systems for manufacturing system control: the technological (physical), and geometrical adaptive control system. Such a classification may create some confusion concerning the main objective of any manufacturing system, i.e. to produce the work piece within imposed tolerances. It is convenient to refer to two distinctive modes of system identification, namely:

- (1) The manufacturing system is identified with respect to the cutting process stability (chatter, maximum available machine power, tool life, etc.). This is the case in rough machining, where the manufacturing tolerances are very large and very seldom set the limits to the machining parameters.
- (2) The manufacturing system is identified with respect to its quality (accuracy) stability (machine accuracy, some form of chatter, tool wear, etc.). It is generally the case in finish machining, where the manufacturing tolerances

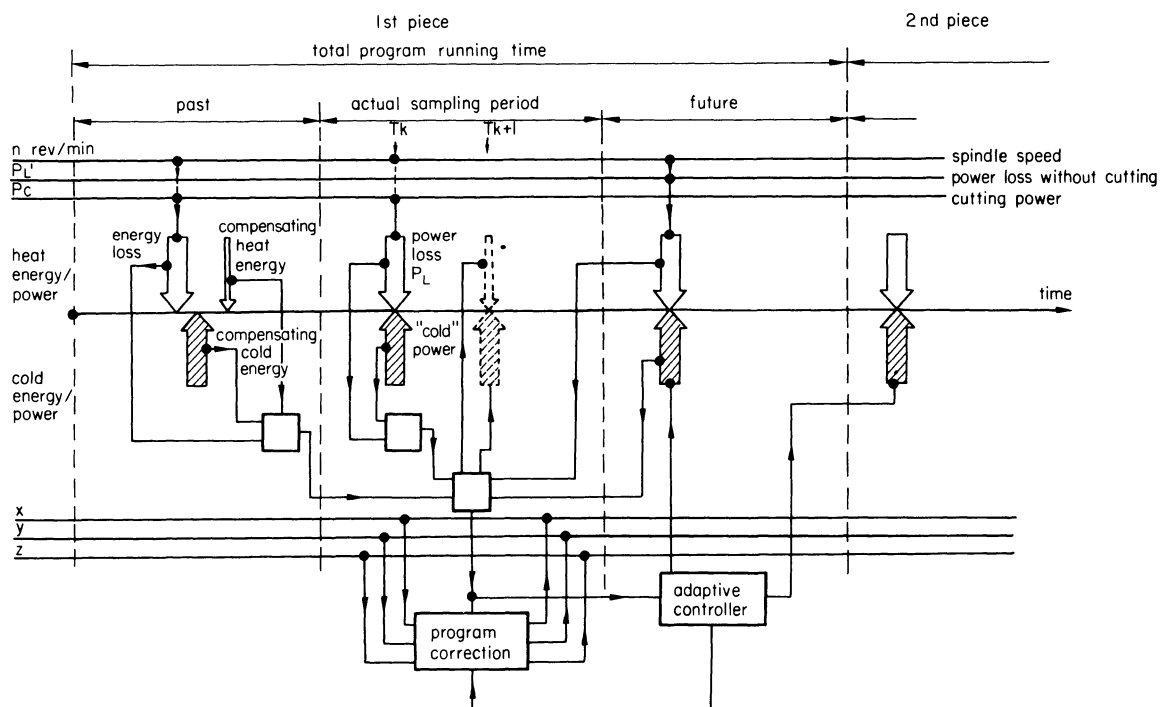


Figure 9. General concept of machine tool thermal stabilization and its thermal error compensation.

are very critical. The system accuracy level is variable and depends on the imposed work piece tolerances.

In both cases, the adaptive controller will act upon the machining parameters F, n , to assure the maximum productivity $P \max$ or the minimum production cost $C \min$. We are generally not interested in the absolute best quality, since $Q \max$ means only that all geometric parameters of the work piece are within the desired tolerances. To avoid confusion, it is more convenient to speak about the relative production quality Q_r . Thus, the complex objectives of the adaptive control can be expressed as follows

$$P(QR P \max) = \max \text{ or } P(QR C \min) = \max$$

Adaptive control allows for improvement in machine tool performance. It can also be proved that the true high performance of machine tools (as defined earlier) is not possible without adaptive control. Some excellent publications are available on adaptive control of machine tools^{1,3}; therefore this paper deals only with some particular aspects of the a.c., as applied to high precision boring machines.

These machines allow two modes of adaptive control operation. It is evident that the second mode has a particular importance for jig boring machines. Because of physical limitations, not all machine operations can be controlled by means of a.c. The on-line measurement of the bore diameter and the in-process boring diameter adjustment is almost im-

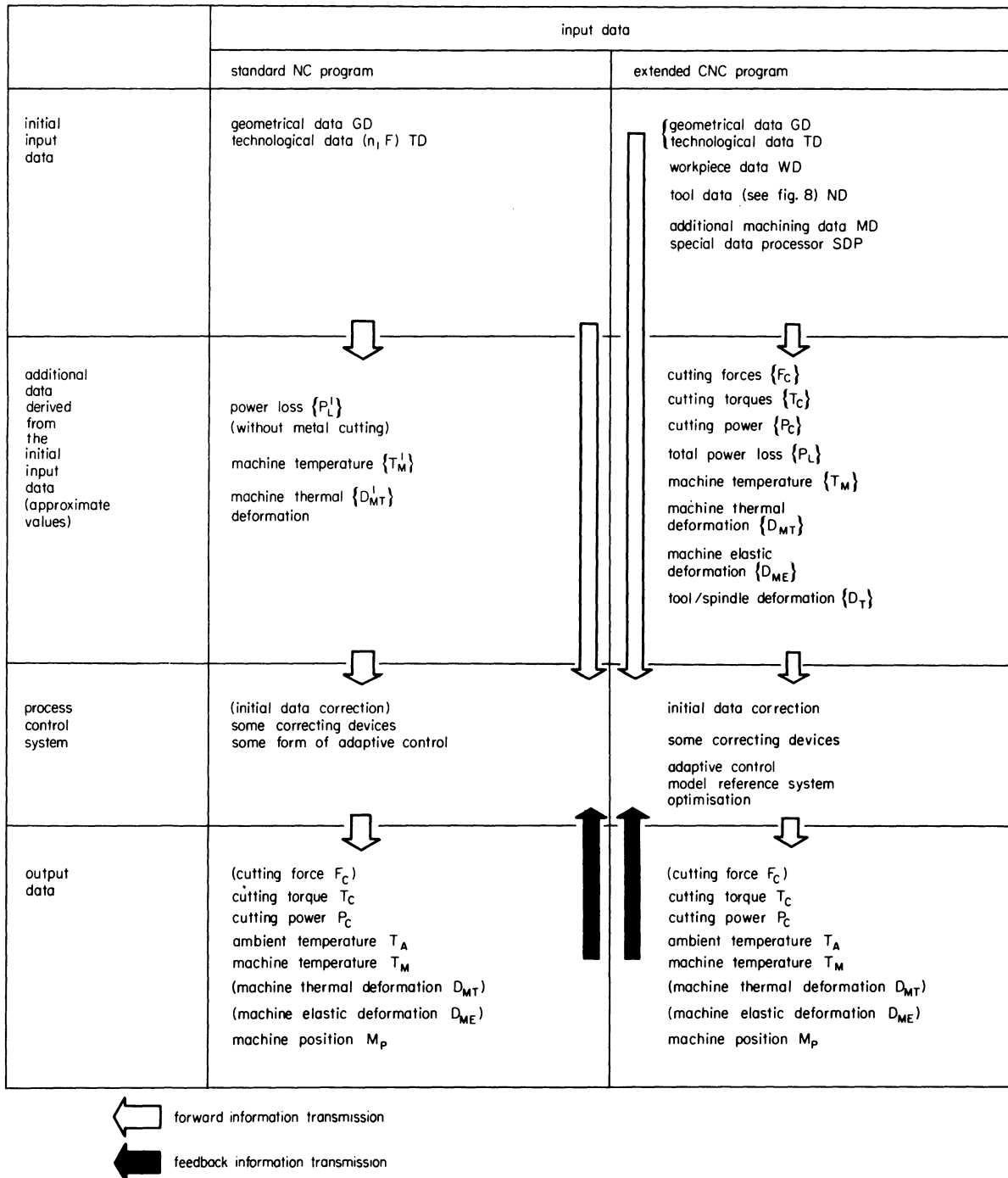


Figure 10. Some additional data necessary for the effective machine tool accuracy control.

possible. The few existing devices apply only to very limited classes of work pieces. The pre-setting of tool diameters for general precision operations and their manual re-adjustment for finish operations still represent the only efficient method in boring practice. For the same reason, the surface roughness cannot be assessed during the machining operation. The surface quality can be maintained within given bounds only by pre-programmed control.

Unfortunately, it is also impossible to measure directly the relative tool position with respect to the work piece by means of an independent reference system, as proposed by Professor Koenigsberger. Therefore, it is necessary to create a reliable machine tool co-ordinate system. The machine tool inherent error matrix + machine frame can be considered as the best physical realization of such an idealized reference system for high precision boring machines. The distortion of the machine frame, which affects the linearity of the initial system, can be partially counterbalanced by the convenient error matrix or program correction. One of the promising techniques consists of the model reference system, where it suffices to construct the model response on a computer.

Because of very complex manufacturing system dynamics, only a combination of a central adaptive controller and some correcting device allows the optimum system performance¹². This approach is particularly promising for the effective compensation of small thermal and elastic machine errors (see figure 9). Three kinds of adaptation are possible: instantaneous, from pass to pass, and from piece to piece. The necessary data for the process controller are provided by the process itself and by an extended part program (see figure 10).

CONCLUSIONS

Thermal and weight effect errors set an absolute limit for high precision machine tool accuracy. Adaptive control, in conjunction with a pre-programmed control and some correcting devices, can lead to a further improvement of the machine tool performance. Because of the flexibility of programming, a digital computer is more versatile than a specially designed controller. The magnitude of the admissible error

determines the degree of sophistication to be employed in the control system.

REFERENCES

1. J. W. Anderson. Machine tool testing—a practical approach. *Proc. of the 12th Internat. MTDR Conference*, 1971.
2. R. Bellmann and R. Schwaighofer. Stabilität und gesicherte Genauigkeit der Maschinen und Prozesse. *Fertigung* (1973) 3.
3. J. B. Bryan. International status of thermal error research. *Annals of CIRP*, **XVL**, 1968.
4. C. P. Hemingray, A. Cowley and M. Burdekin. Positioning accuracy of numerically controlled machine tools. *Proc. of the 12th Internat. MTDR Conference*, 1971.
5. C. P. Hemingray. Some aspects of the accuracy evaluation of machine tools. *Proc. of the 14th Internat. MTDR Conference*, 1973.
6. A. W. Kwiatkowski and H. M. A. L. Samaria. Progress in the application of analysis methods in the identification of machine tool structures. *Proc. of the 9th Internat. MTDR Conference*, 1968.
7. B. P. Lathi. *An introduction to Random Signals and Communication Theory*. Intertext Books, London, 1968.
8. W. J. Love and A. J. Scarr. The determination of the volumetric accuracy of multi-axis machines. *Proc. of the 14th Internat. MTDR Conference*, 1973.
9. J. Loxham. From science to technology. *Prod/IL/1966*.
10. A. Mottu. Procédés et limites des mesures dimensionnelles. *Ingénieurs et Techniciens*. 212, 1967.
11. J. Peklenik. Geometrical adaptive control of manufacturing systems. *Annals of CIRP*, **XVIII**, 1970.
12. R. Schultschik. Fehlerbetrachtung an einer Drehmaschine. Presented at *23rd CIRP Assembly*, Bled, Yugoslavia, 1973.
13. G. Spur. Optimierung des Fertigungssystems Werkzeugmaschine. Carl Hanser Verlag, Munich, 1972.
14. J. Tlustý and F. Koenigsberger. New concepts of machine tool accuracy. *Annals of the CIRP* **XIV**, 1971.
15. W. Sadowsky. Production jig boring. *Tooling and production*, March, 1968.
16. L. T. Wrotny. The principles of machine tool design. (In Polish). WNT, 1973.

THE MACHINING-CENTRE CONCEPT; AN ANALYSIS OF A TOOL-CHANGING MANUFACTURING SYSTEM

by

J. R. CROOKALL* and A. T. M. JAMIL*

SUMMARY

The machining centre concept represents a system of manufacture based upon the philosophy of performing complex cycles of operations at a single station. In carrying out cyclical combinations of tool changing, positioning and machining, it is evident that these basic operations interact not only with each other, but also with characteristics inherent in the products.

Computer simulation has been used to evaluate system performance, using a fairly fundamental approach. The work starts with an investigation of certain interacting elements such as tool changer/positioning systems, and proceeds to examine a basis for optimal programming, and system/component matching. Finally, comparison of single- and multi-station systems in terms of cost is exemplified and discussed.

INTRODUCTION

The machining centre concept represents a complete, automatic, data-controlled system of manufacture. It is designed to provide precise manipulation and positioning of the part, to bring into action a variety of different types of tools in a predetermined manner, and hence to perform complex cycles of generation processes. This is also the function of any system of manufacture based upon discrete conformation of shape. The obvious basic difference is that the machining centre is a single-station, multi-tool system, whereas the normal implication of the term 'manufacturing system' connotes a multi-station facility. Nevertheless it is evident that the form generation capabilities of both types of system are basically equivalent, and it is thus pertinent to attempt to establish criteria which will form a basis of assessment and comparison.

The conduct and control of manufacture in either case is subject to optimization, and must pay due regard to the specific tasks which are inherent within the product characteristics. This paper is concerned primarily with the machining-centre concept, representing one form of *arrangement* of facilities.

In performing cyclic combinations of tool changing, positioning and machining, it is also evident that these activities interact, and that the requirements of the product and the design of the system must be sensibly related. Taken in isolation, each element of the system is subject to many variations.

For example, the tool-changing facilities may range from a relatively simple turret, which usually limits both the complexity of sequence and the number of different operations which can be performed, to large-capacity random-selection tool-changing devices, which must be fairly complex in their mode of operation if an unacceptably long search time is to be avoided. Obviously it is important to be clear on the type of work which can be economically justified in the case of a more complex machine.

Several methods of approach are possible in the investigation of manufacturing systems. Indeed as far as numerical control itself is concerned, frequent recourse is made to the 'case study' approach. Whilst this is useful in its way, the exercise in hand is usually the *justification* for the proposed method, and it is in the authors' experience very seldom that the component or components which are used as a basis of the study have been ill chosen from this particular point of view. This usually means that the case has been favourably biased towards the adoption of numerical control.

Indeed it is difficult anyway to make a choice independent of machine characteristics. For example, certain arrays of holes which must be drilled are equally straightforward when performed upon machines having sequential, paraxial or simultaneous controllers, whereas for other patterns the former two categories of positioning system are at a disadvantage in terms of hole-to-hole positioning rate. In such a case a fairer comparison might be obtained by

* Imperial College of Science and Technology

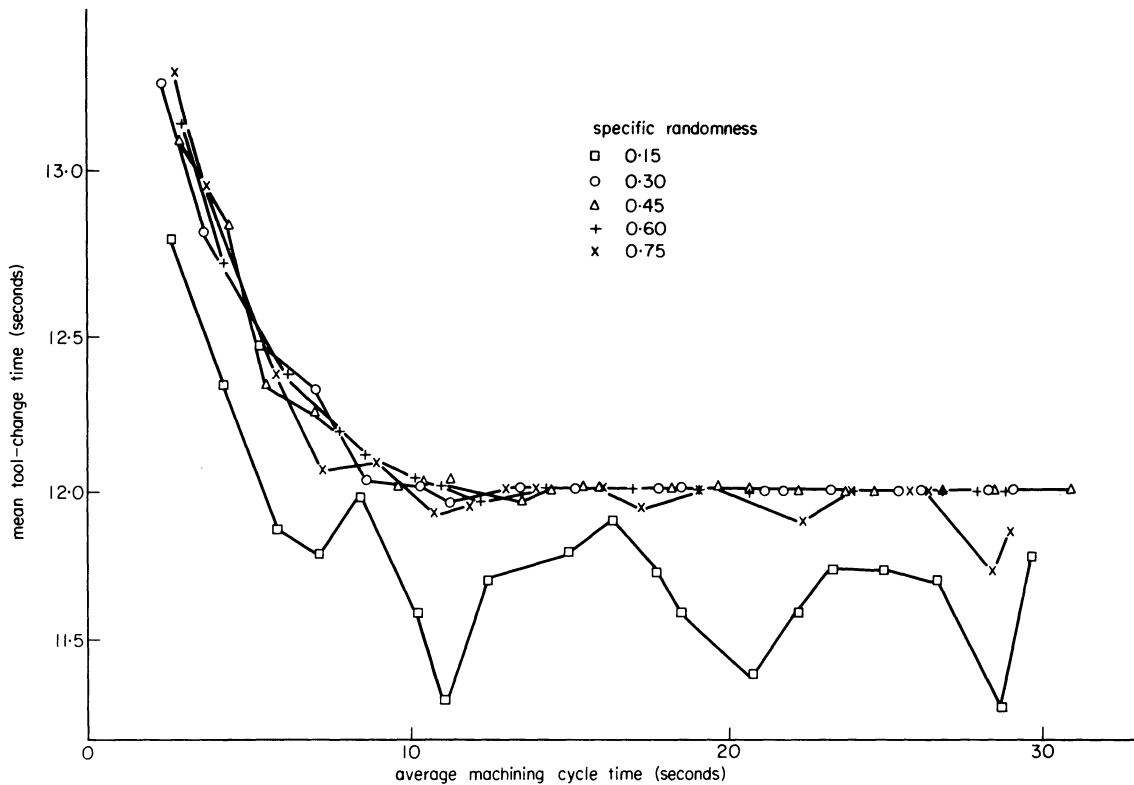


Figure 1. Simulated performance of a single-spindle variable-address tool changer. Unidirectional magazine rotation.

the greater utilization of an effective tool changer in the case of these two categories.

Even if an attempt is made to cover a wide variety of different components and hence to avoid the dangers of 'selection' of rather suitable types, a second difficulty arises. That is that there may be no relationship between the components which can be expressed in a numerical form, and which constitute some sort of scale of 'difficulty' or 'complexity'.

Recourse has therefore been made in this study to the 'generation' of components using the Monte Carlo method on a computer. The advantage is that successive components which are generated are now fundamentally 'similar' in that they are samples of the same population. Simulation can thus proceed until, notwithstanding individual variations from one component to another (which indeed represents the true situation) convergence upon realistically accurate parameters of machine performance can be obtained for each basic parameter of workpiece configuration. Furthermore investigation can be carried out with systematically varying parameters of workpiece characteristics, and hence a scale of workpiece 'complexity' can be envisaged along which individual points bear a direct and continuous relationship.

TOOL-CHANGING SYSTEMS

Many types of tool-changing systems are known in practice. It is not intended to enter into detailed inter-comparison of such systems here, as this will be reported elsewhere¹. Rather the intention is to illustrate some basic features, and to emphasize how certain elements of a tool-changing cycle interact with the functions of other parts of the system.

Suffice it to say that tool changers in general may be classified according to three main constituent elements, namely the spindle system, the tool-transfer mechanism, and the tool magazine. Variants on each of these elements are possible, including multiples thereof; for example more than one gripper system, multiple magazine arrays, and even double spindles, in which case the effective tool-changing time is related to the exchange of spindles rather than the changing of tools *per se*. Two types of tool changer are illustrated by way of example, namely the single-spindle variable-address system and the double-spindle fixed-address system. The performance of these has been evaluated by computer simulation, and the results are illustrated in figures 1 and 2. It is immediately evident that the effective tool-changing time is not a constant, but is affected by the average machining cycle time. This is because such systems are normally designed to achieve a degree of overlapping, and consequently for short average machining times the effective tool-change time is affected by search times, which for longer machining times are overlapped and therefore unimportant. For these tool changers the tool-change time approaches a constant value at or above an average machining time of about twenty seconds. It is also immediately evident that the tool-changing time for a double-spindle system is significantly lower than for the single-spindle type, for the reasons given above.

A further important factor concerns the extent to which tools are called up out of strict sequence of replacement in the magazine. Hence if tool replacement is strictly sequential, then a constant tool-change time could be expected. The measure of the non-sequential nature of tool selection related to the

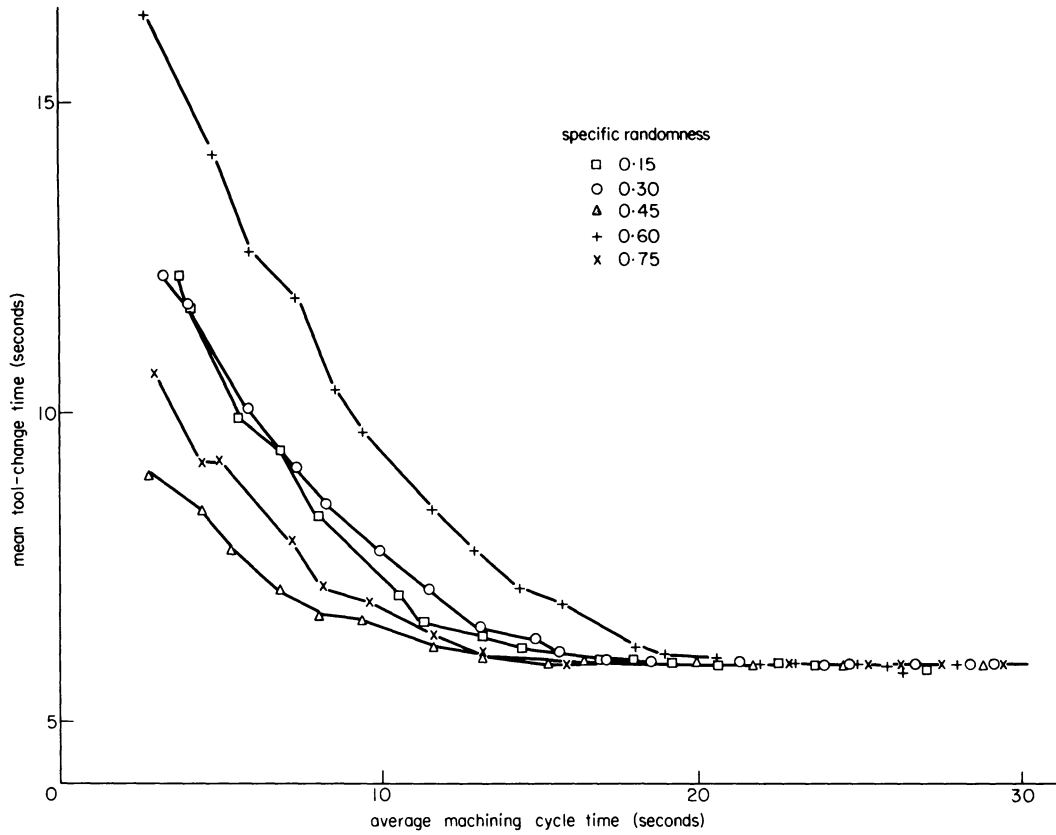


Figure 2. Simulated performance of a double-spindle fixed-address tool changer. Unidirectional magazine rotation.

original placement has been termed 'specific randomness of tool arrangement', and is expressed as a fraction of the full capacity of the magazine. Bearing in mind the shorter tool-change time of the double-spindle changer, it is evident that randomness of demand is more significant for this type than for the single-spindle system. The reason is the use of a fixed-address system in the example, implying that randomness of demand involves greater indexing activity in order to return tools to their original location. With a variable-address system, returning tools can be placed in the nearest available space in the magazine. A non-linear trend in the effect of

specific randomness may also be noted in the case of the double-spindle system illustrated, because magazine indexing is uni-directional, which is the one feature common to both changers.

A further point, which has been investigated, concerns the optimum arrangement of tools within the magazine to minimize tool-changing time for the product. Such an optimal arrangement will differ according to the type of tool changer used.

The action of a tool changer can be conveniently studied in terms of its activity network diagram and tool transportation chart. These are illustrated in figures 3 and 4, together with the tool/magazine

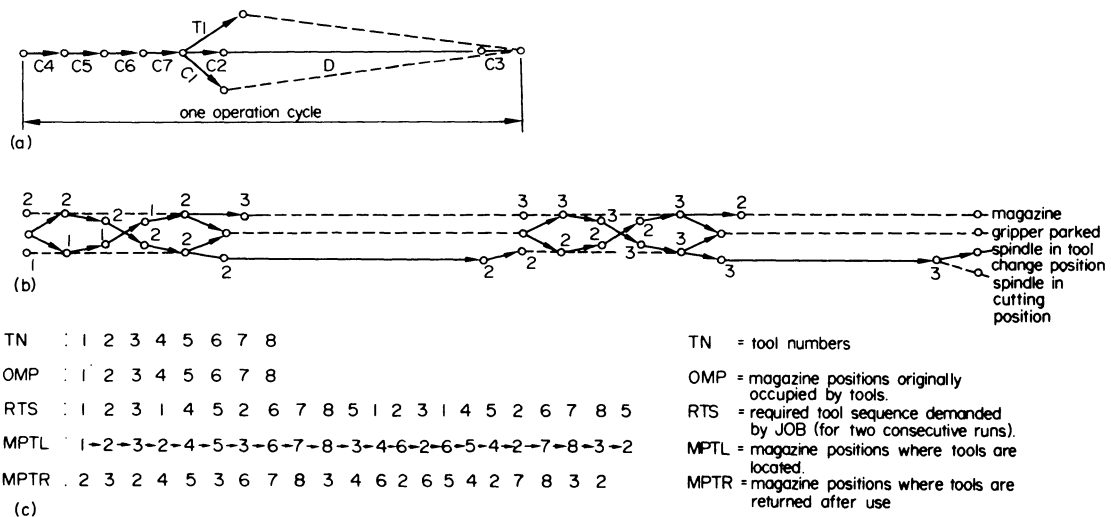


Figure 3. Single-spindle variable-address tool changer. (a) Network diagram. (b) Transportation chart. (c) Tool/magazine location chart for a hypothetical required tool sequence.

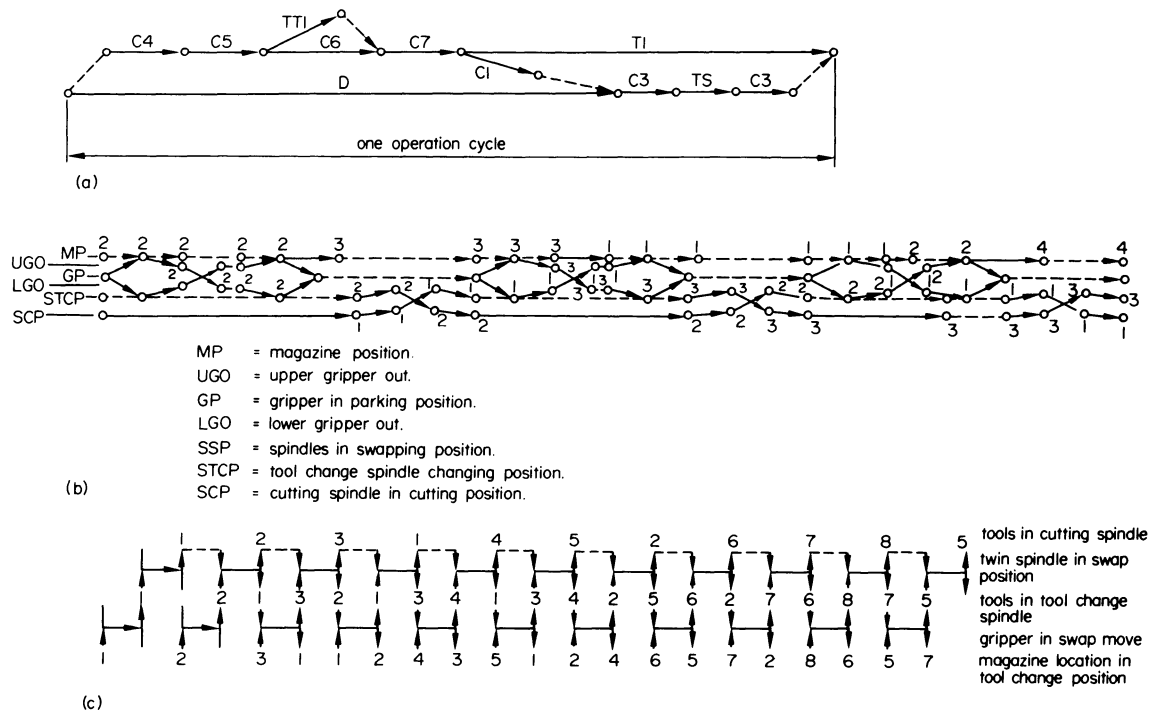


Figure 4. Double-spindle fixed-address tool changer. (a) Network diagram. (b) Transportation chart. (c) Tool/magazine location chart for a hypothetical required tool sequence.

location charts for a hypothetical required tool sequence, which illustrates the more complex indexing required for the fixed-address tool changer when random demands are made upon it. Comparison of the activity network and transportation charts illustrate how, by interposing more system elements between the magazine and the spindle, the action of these can proceed more independently, and hence provide greater scope for overlapping, thus reducing the effective tool-change time.

COMBINED N.C. POSITIONING AND TOOL-CHANGING SYSTEMS

Manufacturing systems, in the sense in which they are being discussed here, exhibit their essential differences, not primarily in the operations themselves, but in the arrangements, sequencing and manipulation of those operations. In terms of machining, the appropriate consideration therefore is 'non-cutting time', as it is in the minimization of this factor for complex sequences and cycles that the manufacturing system exhibits its efficiency.

However, in attempting numerically to define system performance, the characteristics of the task which it has to perform must also be decided, and it is evident that even for a single task a very large number of solutions exist. The question immediately arises as to which solution is to be adopted as the index of performance. In these circumstances, which are not uncommon, the logical approach is to seek the optimal solution. This has the twin advantages of uniqueness and repeatability. However, the statement is far easier to make than to carry out.

Consider, for example, a component requiring twenty separate operations; by conventional N.C.

standards this might be judged by no means a complex case. Seeking that route for which the total distance travelled by the positioning system is a minimum involves the selection from a total number of possible routes of $N!/2$. This represents a 'travelling salesman' situation, but it is immediately obvious that the application to numerical control yields a problem far too large for economical solution by this method.

In fact the problem does not end there, as the absolute performance of the positional and tool-changing systems are less important than their *relative* performance. Only in that circumstance in which the tool changer is so much faster than the average time taken by the positioning system in moving from one location to another, is a policy of allowing as many tool changes as may be necessary in covering the 'minimum path' route likely to be the most successful one. Conversely, if it is the tool changer which inevitably takes the longest time, then the appropriate policy is that of minimizing the number of tool changes, irrespective of the positional route which then follows. It is obvious therefore that most systems fall some way between these two extremes, and hence the optimization problem which must be solved is that of minimizing the combined positioning and tool-changing activity.

Furthermore it is evident that the characteristics of the product have entered the problem. Thus according to the distribution of operation locations which is assumed originally, the average distance between locations will vary. Hence a system specification whose relative speed of positioning and tool changing is well matched to one component, will be found to be poorly matched to another. The importance of matching of components to the system of manufacture, and illustration of the fact that system/

component matching cannot be perfectly achieved for a range of components, is illustrated thereby.

The basic parameters by which system flexibility and performance can be related to component characteristics and complexity in a rather fundamental way has already been implied. In the authors' submission, to approach the problem in any other way is to run the risk of obtaining as many answers as the number of solutions tried, or of placing unjustified reliance upon what is almost inevitably a chance result.

Derivation of optimal route

The combination of operations which are required in the manufacture of a hypothetical part are derived by the Monte Carlo method in terms of type (i.e. the tools) and their locations. The number of operations required is N , and the number of tool changes is M , but the latter is a variable according to the route chosen, and can take a value ranging from the minimum number of tool changes possible up to N .

To relate the positional and tool-changing elements of the system, a 'tool-change equivalent distance' is defined, and hence at each operation change within each simulated route, the positional distance is compared with the tool-change equivalent distance to provide the 'effective distance between operation locations'. This will be the longer of the combined operations, and in fact represents the 'critical path' in the activity network diagram, mentioned earlier. The sum to N of the effective distances yields the 'total effective distance', and the optimiza-

tion problem is thus the determination of the minimum value from all possible routes.

In order to avoid the almost unmanageably large problem which this represents, a simplified procedure is adopted. This involves beginning each simulated route in turn at a different starting point, and adopting the criterion of 'next nearest location' in deriving the remainder of that route. It is known that such a strategy, although not necessarily yielding the global optimum, nevertheless will provide a best solution (i.e. shortest total effective distance) which is most frequently quite near it.

Results of simulation

Figure 5 shows the simulated results of some solutions for a 'job' consisting of 40 operations and involving 20 different tools. Each plotted point represents one possible solution or route, and evidently the total number of possibilities is very greatly in excess of those which are shown. The best solutions are represented by the extreme left-hand points, which have been joined to form a 'job line', as any point on this line represents a solution yielding the minimum average positional distance according to the number of tool changes which that solution requires. Other solutions are of little interest because they are suboptimal.

The job line is so-called because it characterizes the simulated task; the line itself encompasses all possible variants in combined tool-changer/positional systems. It illustrates the fact that with a

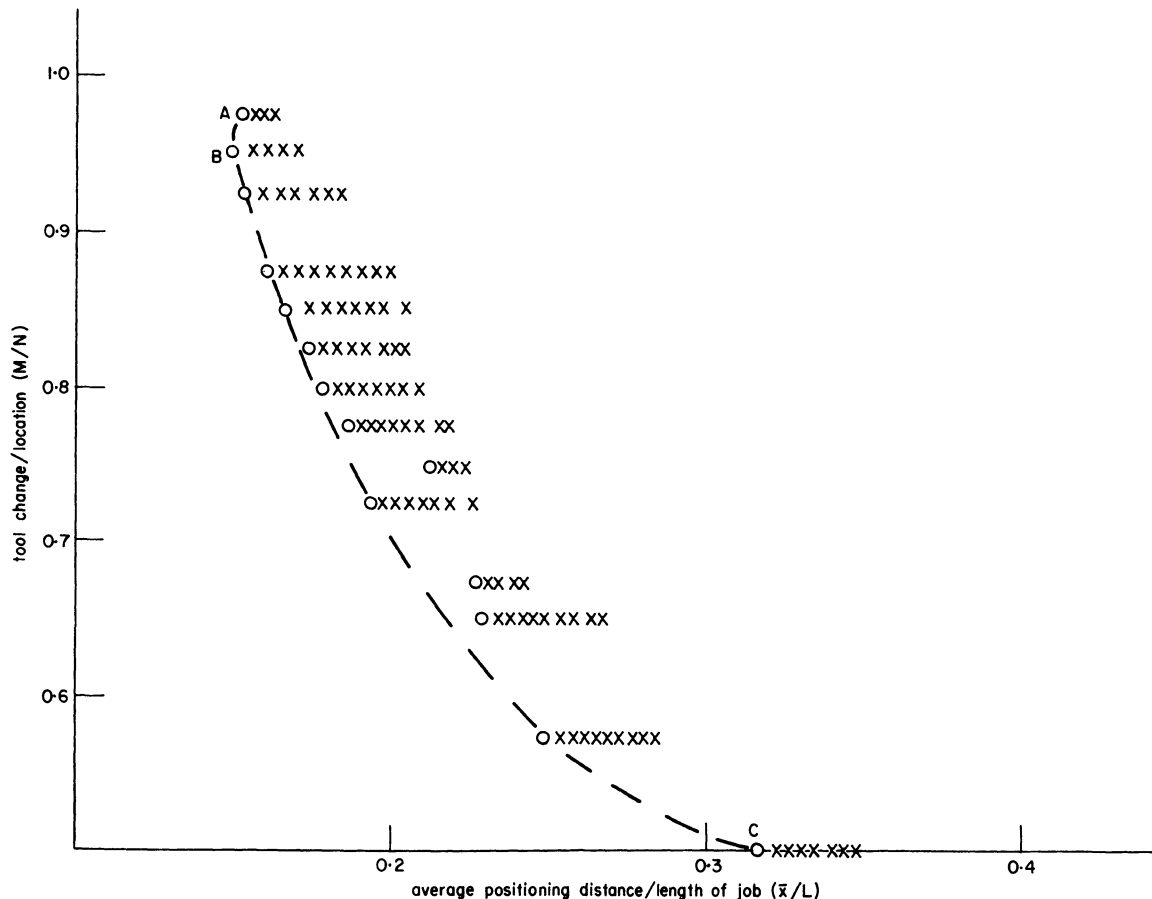


Figure 5. Simulated job-line for 40 operations and 20 tools. $L/B = 1.25$.

relatively fast tool-changing system the optimal policy would be to program the route towards point *A*, thus minimizing the average positioning distance, and conversely for a slow tool-changing system (or relatively fast positional system) the correct solution would be found towards *C* where the number of tool changes is at a minimum. The shortest positional route occurs at point *B*, and it will be noted that this is generally at some intermediate point between *A* and *C*.

Figure 5 is general in several respects. It is representative of the complete range of systems, as it is plotted for a tool-change equivalent distance from zero to infinity. The use of the ratio M/N permits the representation of a wide range of jobs having different values of N , and this applies also to the job 'size' (indicated by L and B) within which the N operations are distributed. The job line therefore represents a job family, the characteristics of which are inherent in the population from which its characteristics are generated. This means that the results of repeat simulations, though generating different 'jobs', should nevertheless converge with decreasing sample variance upon a result which is directly related to the assumed distributions. By varying the distributions in a consistent manner, a fundamentally valid indication of system performance can be correlated with a consistent and reliable scale of job parameters.

The relevance of this can perhaps be seen more clearly if it is considered that numerically controlled systems (or any other for that matter) should be designed or selected, not for one or two immediate tasks, but for a work span of perhaps ten years, during which particular characteristics of individual components lose all significance. The system can be envisaged therefore as operating upon a continuous succession of tasks which are statistically distributed. The origins of the populations from which they are derived lie in the process of component design.

Job-family lines and job-family groups

As the job line is representative of the population, individual jobs, representing samples of that population, must be expected to show variations. If the variance of the population is reduced, so will that of the samples, and hence the jobs generated are more precisely defined in terms of the populations for which they are generated. Such a philosophy forms the basis of dividing a population of jobs into family groups. Alternatively the total population is seen as comprising the envelope of distributions, each of which represents a different job-family group. In terms of the relationships represented by figure 5, each family group would be identified by a different job line. Evidently the performance of the system and the strategy for optimally programming it, will be different for each family group. The corollary is that the system specification can be more precisely matched to the job family group as the latter becomes more narrowly defined. It will be equally evident that optimal matching to a population is almost never achieved, except very occasionally by chance.

Further work (not reported here) concerns the extraction from simulation results of relationships

between job characteristics and optimal system programming, using regression analysis and hypothesis testing techniques.

SYSTEM PERFORMANCE AND MANUFACTURING COST

Consideration of manufacturing cost involves factors additional to those already discussed. An accepted feature of the single-station system, represented by the machining centre, consists in exploiting to the full the ability to carry out complex tasks, and frequently to complete the manufacture of a component at a single station. By comparison, the multi-station facility is normally arranged to provide less comprehensive facilities at each station, the manufacturing task being appropriately divided. The advantages gained are those of parallel working, whereas the disadvantages are those of multiple setups, multiple loading/unloading operations, and transportation. Such operations may be performed manually in part or in whole; nevertheless it is the extent to which such operations are saved which is relevant to the overall performance of the single-station system.

Indeed this situation arises in quite a general context in relation to manufacturing systems. Whilst in the minds of some, control of the ideal manufacturing system is either totally automatic, or at least the role of the human operator is not that of pace-maker. A more general, and possibly valid view, is that such systems are better considered as co-existent, complementary and carefully interlocked arrangements of men and machines. Indeed some careful analyses have indicated that totally automatic systems have a doubtful chance of practical success other than in rather limited areas of application. The question posed therefore is that of numerically evaluating such a mix of manual and automatically executed functions. In a search for a solution, recourse has been made here to the use of Standard Time Data, which provides an additional advantage in ensuring a better level of comparative assessment.

Returning to the main problem, it is evident that it is in the non-cutting time elements of manufacture that the system in question demonstrates its effectiveness in a multi-role task. Evaluation of non-cutting time is made using an equation of the following type:

$$t_{ed} = \frac{\sum_{j=1}^k t(m, d_j) \cdot p(m) \cdot p(m, d_j)}{\sum_{j=1}^k p(m) \cdot p(m, d_j)} \quad (1)$$

where t_{ed} = estimated duration of non-cutting activity on machine m .

$p(m)$ = probability of type m machine (e.g. milling) including different requirements to be used for a typical job

$p(m, d_j)$ = probability of strategy d_j being used in combination with m .

Some activities are known to be functions of job variables; for example, loading/unloading is a function of weight, and hence effort is required in its

manipulation. Other data used comprises established statistics of workpiece and machine characteristics. The following factors have been evaluated with respect to the type of system and task required:

- (1) Setting up time.
- (2) Loading/unloading time.
- (3) Tool-change time.
- (4) Co-ordinate positioning time.
- (5) Inter-machine transportation time.
- (6) Miscellaneous handling time.
- (7) On-machine inspection time.
- (8) Machine waiting time.

An explicit function for manufacturing time per job was derived in terms of the above activities, and was used as the basis for a simulation study.

Results of simulation

Because of the number and complexity of interacting factors, a single, consistent representation of total manufacturing time per job cannot be demonstrated. However, certain relatively straightforward relationships emerged, which are sufficiently general to be of some practical use, and will be useful as an illustration. The simulation procedure employed as variables the number of different types of machines, the number of operations, and the number of different tools required; these were generated by the Monte Carlo method from uniform distributions. Within certain constraints, jobs requiring up to thirty operations were simulated. Where necessary, simulations were repeated to improve the accuracy of the result.

It was found that the number of operations in a job yielded a fairly clear relationship with non-cutting time provided the latter excluded setup and machine-waiting time elements. This is illustrated in figure 6

for conventional machining. A standard deviation of 1.06 minutes was obtained, which is seen to be nearly constant. This illustrates a useful fact, in that the number of operations can be taken, in these terms, as an approximate measure of 'work complexity', and that the relative accuracy increases with the complexity of the work.

Relationship with manufacturing cost

Time-dependent costs may be directly associated with time elements which have been established in a manner described above. Other costs are quantity dependent, i.e. they are related to the batch size. There are also factors which differentiate systems, when considered in terms of cost, which have no counterpart so far as manufacturing time is concerned, examples being tooling and fixtures, planning, and scrap.

Furthermore the cost of a given manufacturing method or system is of less intrinsic value for investigational purposes than is an evaluation of *comparative* costs. Also, having made comparative studies, it is instructive to observe those aspects which constitute the primary source of difference, as it is frequently found that when a system is held, intuitively or otherwise, to be superior, then it is believed that such superiority extends throughout most or all of the factors involved. Obviously this is often not the case, and it is therefore important to be clear upon just what factors any superiority exists, and the bounds within which the advantages are retained, and conversely outside which the system is no longer effective in these terms.

One such comparison is illustrated in figure 7. The parameter $\Delta C/C$ represents a fraction of the total cost for conventional machining which is different for the

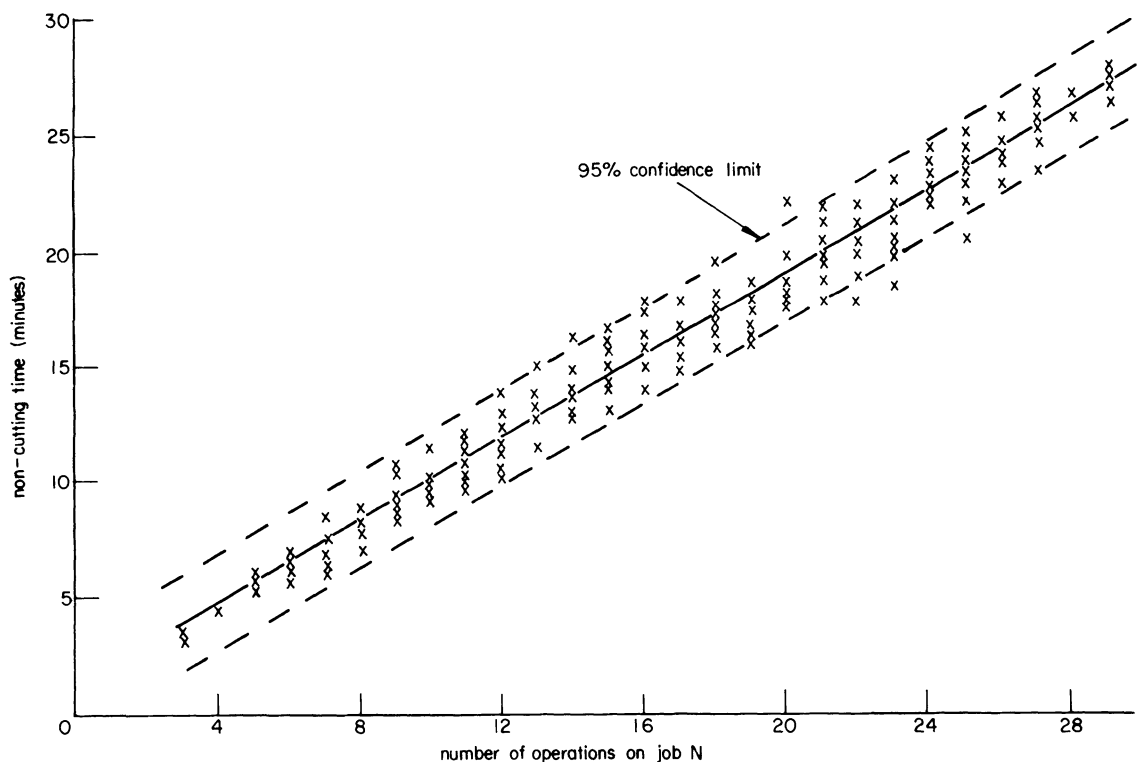


Figure 6. Non-cutting time as a function of work complexity N in conventional machining. Non-cutting time excludes setup and machine-waiting time.

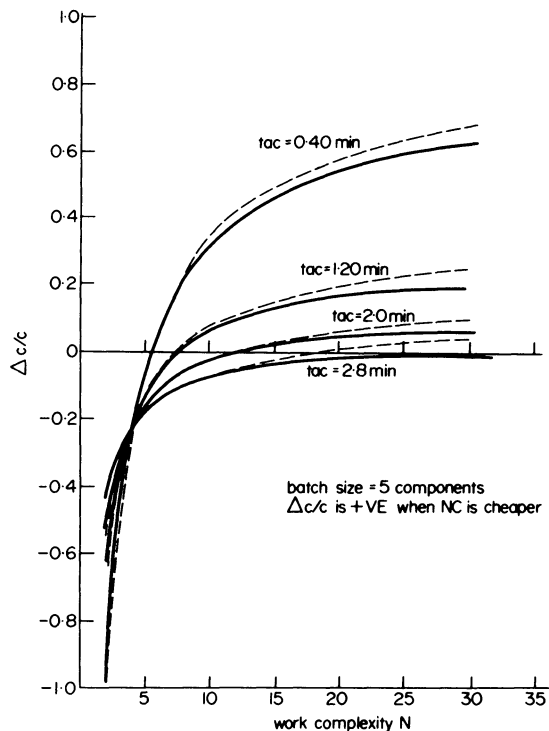


Figure 7. Simulated comparative cost ratio as a function of work complexity N .

two systems, and is positive when the numerically controlled system is cheaper. It is immediately apparent that work of low complexity (few operations) stands a high risk of being uneconomically produced on a numerically controlled system, and conversely as the complexity of the task rises, so does the efficacy of using numerical control. However, several curves are illustrated, representing different values of the average machining time for which simulation of the systems was carried out, and it is noteworthy that this has an important influence on the comparative economics of numerical control. Thus it is when the machining times are relatively short that the comparative benefit of numerical control is marked, as the major task is the manipulating and sequencing of operations, for which the numerically controlled system has a marked facility. Conversely, when the main task turns towards straightforward machining (for which no advantage for numerically controlled machines has been assumed), the relatively high cost of providing numerical control of machine axes is hardly worthwhile.

The effect of batch size on comparative cost ratio is shown in figure 8, from which it is seen that the economic balance is fairly quickly established over about the first ten parts, after which the situation does not change markedly. From a comparison of figures 7 and 8 it may be concluded that the type of demand made upon the system by the complexity and relative duration of the operations required significantly affects the comparative system economics.

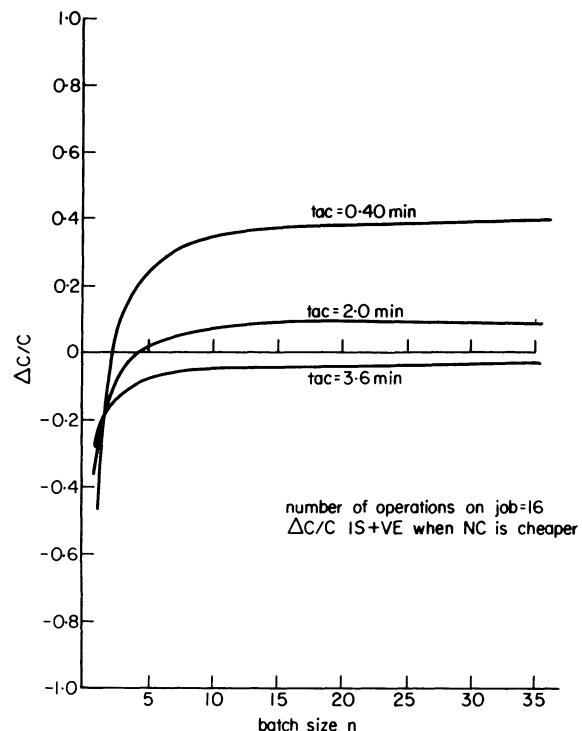


Figure 8. Simulated comparative cost ratio as a function of batch size n .

For the type of system which has been investigated here, i.e. numerically and conventionally controlled machining systems, the economies of 'scale' are less significant than might have been expected. Certainly this cannot be deemed a general conclusion extending to other production methods, but is relevant to the conduct and control of machining operations.

CONCLUDING REMARKS

The indications are that care and attention must be paid to establishing in a fundamental way the relationships between manufacturing system performance and product characteristics. It is necessary to ensure thereafter the continuance of the type of work for which positive advantages have been shown. The matching of the system elements and their performance closely to the production tasks involved, rather than simply in ensuring automatic operation at every level, is the only certain way of maximizing cost effectiveness in the design of manufacturing systems.

REFERENCE

1. J. R. Crookall and A. T. M. Jamil. An analysis of tool-changing systems for numerical control, *Proc. 3rd Prod. Development and Manuf. Tech. Conf., Univ. of Strathclyde* (September, 1973) Paper no. 12, p. 26.

NEL/MTIRA TECHNOLOGICAL FORECAST FOR THE COMPUTER CONTROL OF MACHINE TOOLS

by

B. DAVIES*, J. W. BRUCE* and A. E. De BARR†

SUMMARY

The object of the joint technological forecast conducted by NEL and MTIRA was to obtain an estimate of the likely technological changes in the field of the computer control of machine tools and related processes for the period 1975-1980. The technique used was the Delphi method developed by Rand which involves a structured interaction among a group of experts. The group involved in this forecast comprised some 50 experts drawn from Great Britain, Germany, Italy, Sweden, Norway, Netherlands, USA, Canada, Japan and Belgium.

The computer control of machine tools was divided into Computer Numerical Control (CNC), two classes of Direct Numerical Control (DNC) and hierarchical systems which might comprise systems of linked computers. Related processes have also been included and an attempt made to relate these to the above classes of application. Some primary considerations in production control and planning with reference to costing and concepts such as marketing requirements receive special consideration.

INTRODUCTION

During 1971 NEL and MTIRA decided that they should jointly sponsor a Long Range Technological Forecast with the objective of obtaining a reliable estimate from a panel of international experts of industrial needs and thence market opportunities for Computer Numerical Control (CNC), Direct Numerical Control (DNC) of machine tools and hierarchical system containing CNC and DNC¹.

The forecasting technique used was the now well-established Delphi method^{2,3,4} pioneered by the Rand Corporation. This is a form of structured interaction among a group of experts which has the merit of being more systematic than ad hoc methods which do not converge the questioning.

The period chosen 1975-1980 is intended to produce a short-to-medium range forecast. This has more relevance than and avoids the increasing uncertainties in longer-term extrapolations.

For the purpose of the forecast the following definitions were used:

- CNC—computer numerical control—essentially one computer, one machine tool
- DNC—1st level—direct numerical control—behind-tape-reader interfaces—BTR
- DNC—2nd level—direct numerical control—deeper interfaces between machine tool control and computer

Hierarchical systems—control by a number of computers operating at different levels of responsibility.

GENERAL EFFECTS OF CNC, DNC AND HIERARCHICAL SYSTEMS

The panel of experts were asked to appraise various general aspects of CNC, DNC and hierarchical systems and to assess their relationships. These aspects are listed as Items 1 to 28 in Table 1 but it is not claimed that this is a comprehensive breakdown.

Letters A, B, C and D refer to CNC, DNC (1st level), DNC (2nd Level) and hierarchical systems respectively. Although the respondents were in approximate agreement on most questions there were occasions where a lack of agreement had to be indicated by entering a letter in more than one category for a given question. Questions 2.1, 2.8 and 2.14 are examples and show an evenly divided opinion on these questions.

In addition the panel was asked to choose the three most significant items for each of the systems of computer control. Consideration of the panel's assessment of the importance of each item is given below.

* National Engineering Laboratory, East Kilbride, Glasgow

† Machine Tool Industry Research Association, Macclesfield

TABLE 1 Effects of CNC, DNC and the hierarchical system

No.	Item	Essential	Major importance	Relatively important	Relatively unimportant	Minor importance	Of no importance	Undesirable
2.1	Significant reductions in capital cost of computer-controlled machine tools over present levels of cost for numerically controlled machine tools			ABCD		D		
2.2	Reduction in running costs over present levels	A	C	BD				
2.3	Reductions of cost and complexity of program preparation		D		ABC			
2.4	Improvements in accuracy over present systems					ABCD		
2.5	Increased reliability over present manufacturing systems			ABCD				
2.6	Storage of programs in computer and removal of tape readers from the shop floor	B		CD	A			
2.7	Minimum maintenance requirements			BCD	A			
2.8	Life (say 5-10 years) of computer model before design obsolescence			ABCD				
2.9	Retrofit capability			ABD				C
2.10	Simple, standardized interfacing between computer and machine tool	A	BCD					
2.11	Single computer programming language				ABCD			
2.12	Ability of computer to reproduce skill of operator			B	AC	D		
2.13	'Technology' in computer programs	D		BC	A			
2.14	Provision of area-clearance cycles in programs		CD	ABC	A			
2.15	Machinability data banks	D		BC	A			
2.16	Group technology	D	B		A	C		
2.17	Adaptive control			AB	CD			
2.18	Automated inspection		D	BC	A			
2.19	Significant degree of automated work-handling		D	ABC				
2.20	Effective control of swarf	D		A		BC		
2.21	Long-life cutting tools			ABCD				
2.22	High stiffness of machine-tool structures			ABCD				
2.23	Modular constructions for machine tools	D	B	A		C		
2.24	Cheap and reliable drives			ABCD				
2.25	All electric systems					ABCD		
2.26	Simple provision for management data feedback	D	BC	A				
2.27	Appropriate training facilities for staff	ABCD						
2.28	General education in computer technology	D	C	AB				

A CNC
 B 1st level DNC
 C 2nd level DNC
 D Hierarchical

2.1 The capital cost of a machine tool is seen as being significantly reduced where computer technology is applied. However, there is uncertainty as to whether this applies to hierarchical systems. These opinions have been reaffirmed by respondents.

2.2 Reduction in running costs is clearly to be a part of future specifications of computer based systems as replacement technologies. In CNC this is regarded as essential and of major importance in second-level DNC but less important for first-level DNC and hierarchical systems.

2.3 It is interesting that costs and complexity of program preparation are regarded as important only where hierarchical systems are specified, probably because of the introduction of external library files for components and process data for adaptive control and operations planning.

2.4 The present levels of accuracy seem to be adequate for future systems.

2.5 Reliability of current NC controllers is not satisfactory and an improvement is relatively important for all classes of system.

2.6 The current trend toward eliminating intermittent control-tape reading in CNC systems is not supported but program storage is essential; this indicates increased automation.

2.7 Maintenance is clearly related to reliability, 2.5. CNC is not a radical change from conventional hard-wired NC for which maintenance procedures are well established, thus probably accounting for CNC not having special maintenance requirements.

2.8 There is agreement that computer design obsolescence is important.

2.9 Retrofit is declared to be irrelevant to 2nd level DNC systems which comes as a surprise in view of the growing interest in cellular manufacturing in which conventionally controlled NC machines may be grouped. In contradiction retrofit is regarded as having significance for hierarchical systems, CNC and 1st level DNC.

2.10 Simple standardized interfacing is clearly important in all cases.

Items 2.1 to 2.10 inclusive show the three most important considerations for CNC are increased reliability, minimum maintenance, and simple standardized interfacing between machine tool and computer. CNC should also be cheaper than the equivalent existing hard-wired system with lower running costs.

For 1st level DNC the prime consideration is again increased reliability with reduction in running costs as a secondary factor.

Second-level DNC systems must have low running costs, and require a minimum of maintenance.

Capital costs for hierarchical systems are not seen, surprisingly, as important relatively, whereas reduction of costs and complexity in part program preparation are. Improved relative reliability is required. Program storage in computer store, thereby again eliminating tape readers, is regarded as significant.

2.11 A single standardized programming language does not seem to be important, probably because APT and its derivatives have no serious contenders on the part programming side. The lower level,

assembler, languages also show little signs of moving towards any standard and remain completely system dependent.

2.12-2.17 These may be regarded as having common bases in reproducing the learning process of manually based systems, and also in terms of integrating a number of machine tools in a remotely controlled cell or group. As the sophistication of the system increases so does the relative importance of each of the items. 2.12 and 2.17 taken together play down the usefulness of adaptive control in all cases, which is rather surprising. Group Technology is clearly irrelevant to CNC in isolation.

2.18 Automated inspection is a function again related to a group of machines in its importance (as is item 2.19, automated work handling, which is also seen as useful to single-machine working).

2.19, 2.20, 2.22, 2.23, 2.24 and 2.25 These are all really machine design and specification subjects. The results again demonstrate the effect of integration of machine tools and the extension of automation. The poor showing of 'all-electric drives' reflects current technology. Item 2.22 may reflect the present state of confusion and disagreement over the subject of machine tool dynamics into which it seems agreed considerable research is still needed before agreement can be arrived at on dynamic testing standards and test methods.

2.26 The high rank of feedback data-reporting systems for management is thought to identify a clear industrial need in this context.

2.27 and 2.28 Education and training is seen to be of dominant importance in all cases.

MARKETING

Using the following designations for the following markets,

B—Great Britain

E—Europe

A—USA

the panel was asked to assess the importance of various alternatives upon the future sales of British computer-controlled machine tools. Their response is shown in Table 2 and discussed below.

3.1 It will be seen that joint arrangements between UK machine tool makers and overseas manufacturers are regarded as relatively important presumably because of overseas familiarity with their own home computer-based control systems.

3.2 Machine-tool makers will need to provide services covering machines and computer-based controllers, clearly a limiting factor to the growth, and possibly to the subject itself, because of skilled labour shortages.

3.3 The result seems to suggest that the availability of internationally acceptable computers appears to be less significant for the UK than Europe and the USA, though this is not supported in reality.

3.4 and 3.5 Whereas established home-based sales organizations are seen as being important, distribution appears to be less subject to local influences.

3.6 In all cases warranty or guarantee is seen as being essential.

TABLE 2 Marketing forecasts

No.	Alternatives	Essential	Major importance	Relatively important	Relatively unimportant	Minor importance	Of no importance	Undesirable
3.1	Joint arrangements between British machine-tool manufacturers and non-British manufacturers of computer control systems			BEA				
3.2	Availability of combined support service from machine-tool manufacturer covering both computers and machine tools	BEA						
3.3	Availability of internationally acceptable computers	AE	AE		B			
3.4	Employment of sales organizations native to, or well-established in, country of sale		BEA					
3.5	Appropriately positioned distribution centres			E	AB			
3.6	Warranty or guarantee	AEB						
3.7	Suitability for use with local computer services			BEA				

A USA; B UK; E Europe;

3.7 Suitability for use with local computers has some importance.

UK EXPORTS TO THE USA

The export of British computer-controlled machine tools to the USA would not be significantly restricted by licensing arrangements in the USA specifying the use of US computers with UK machine tools; there was 75 per cent agreement by the panel on this point.

ECONOMIC VIABILITY

The types of computer control of machine tools as defined in Table 3 were predicted by the panel as being economically viable as follows.

Table 3(a) shows a very strong opinion that CNC will be established by 1975. This is reinforced by the

increasing level of commercial activity in the field and as seen at major exhibitions and by increasing sales.

An almost equally large body of opinion supports 1st level DNC but a significant fraction of respondents disagreed.

Second-level DNC is seen by 70 per cent of opinion as becoming established by 1980 but again a significant opinion rejects it.

Whilst no one rejected hierarchical systems a small body of opinion regarded them as immediately economically feasible and the majority as being so from 1980.

Table 3(b) is also a prediction of when the items listed will become economically viable. The first three items referring to on-line part programming and factory data collection respectively are available and were thought to be economic. The last item is currently achievable and has been demonstrated, but

TABLE 3(a) Economic viability

Arrangements	1975	1980	Beyond 1980	Never
Computer numerical control (essentially one computer – one machine)	94%		6%	
Direct numerical control (bypass tape reader interfaces)	83%	4%		13%
Direct numerical control (deep interfaces with machine-tool control circuitry in computer)	17%	67%	4%	12%
Hierarchical systems (control by a number of computers operating at different levels of responsibility)	16%	60%	24%	

TABLE 3(b) Economic viability

Item	Year
Direct transmission of part-program information from computer store when required	1973/74
Modification on-line of part-program during tape-proving	1973/74
Collection of information to indicate the current state of the production process	1973/74
Introduction of computer programming to control the execution of a manufacturing programme	Beyond 1980

TABLE 4(a) Future demand for computer-controlled machine tools

Year	1973	1975	1980	1985	1990	Beyond 1990
Total machining capacity available from computer-controlled machine tools	100	230	590	1400	1600	
Machining capacity from computer-controlled machine tools	Large firm of over 1000 employees	75	160	325	500	1000
	Medium-size firm of between 250 and 1000 employees	15	40	100	170	330
	Small firm with less than 250 employees	4	12	40	70	170

NOTE: Cumulative totals for three subsets do not equal total machining capacities because of rounding errors and averaging.

TABLE 4(b) Future demand for computer-controlled machine tools

Arrangements	1973	1975	1980
Computer numerical control (essentially one computer—one machine)	50	130	260
Direct numerical control (bypass tape reader interfaces)	22	50	100
Direct numerical control (deep interfaces with machine tool circuitry in computer)	7	16	50
Hierarchical systems (control by a number of computers operating at different levels of responsibility)	4	13	60
Total machining capacity	100	200	450

the answer which refers to general applicability is seen as being beyond 1980.

FUTURE DEMAND FOR COMPUTER-CONTROLLED MACHINE TOOLS

Table 4(a) summarizes the response to the question what machining capacities will be needed in later years in relation to the total capacity in 1973 taken as a base of 100? The question was framed to obtain growth curves for computer-controlled machine tools for large, medium and small firms over the period 1973 to 1990.

Respondents were asked to allocate machine capacity between the four systems: the results shown in Table 4(b) are for the years 1973, 1975 and 1980, again 1973 has an index number of 100.

The rate of increase per 5-year period comes out at about 2.3 times, or 40 per cent per annum; this relates to 1973 manufacture and indicates a rather small increase in total population. The predicted rate of increase is about the same for each system.

The consensus view is that these predictions are optimistic.

CORPORATE ACTIVITIES

Integrated computer-control of machine tools sited over extended geographical areas was not seen by the panel to be evident by 1980. By 1990 however the trend of opinion reverses heavily in favour.

The prospect of sharing (very) large computer-controlled systems by different organizations by 1980 is completely rejected and the support for this happening by 1990 is small.

INDUSTRY DOMINATION

In reply to the question of which three industries are most likely to use computer-controlled machine tools when fully developed, the clear selection was

Aerospace, Automotive and General—the last included specific individual responses of shipbuilding, electrical engineering and printing.

PROCESS RANGE

Table 5 summarizes the panel's prediction of importance of computer-control to each of the processes itemized.

Computer control will have a major impact on drilling, milling, jig-boring, turning and die-sinking.

The relatively important group of processes comprises grinding, welding, hot and cold rolling and injection moulding.

Less important, being classed as relatively unimportant, are hot and cold forming, power metallurgy, pressure die-casting, precision casting, extrusion and wood-working.

The least important application seems to be gear-cutting, although error-correction could easily be applied between hob-spindle and drive to produce more accurate hobbing. Respondents have reaffirmed their views on these aspects.

RELATED ACTIVITIES

Table 6 extends the forecast into related activities, and again their relative importance is questioned.

The future importance and effectiveness of computer techniques was clearly stated for all the areas of application.

The most favoured area of application was specified as production control. The 66 per cent agreement is surprising in view of the smallness of activity to date.

Overall reductions in inventory and in production costs are seen as being consequent upon increased computer usage. A recent forecast⁵ suggests that by 1980 more than 25 per cent of all machine loading in

TABLE 5 Effect of computer control in various processes

Process	Essential	Major importance	Relatively important	Relatively unimportant	Minor importance	Of no importance	Unnecessary
Drilling		*					
Milling		*					
Jig-boring		*					
Grinding			*				
Gear cutting					*		
Turning		*					
Welding			*				
Die-sinking		*					
Cold forming				*			
Hot forming				*			
Hot rolling			*				
Cold rolling			*				
Powder metallurgy				*			
Pressure die-casting				*			
Precision casting				*			
Hot forging							
Injection moulding			*				
Extrusion				*			
Wood-working				*			

the USA will be done by computers, further that a new breed of production manager, familiar with computer methodology, will be commonplace. Production control by computer is also identified in the forecast as a significant growth area.

Almost equally favoured is production planning, again with 100 per cent in the 'essential/major importance' categories. The average weight in the 'essential/major importance' categories has been calculated as 80 per cent.

COSTS

The panel's opinion on the influence of computer techniques on costs was as follows.

Costs	Percentage Opinion
Capital increased	84
Factory Space decreased	76
Overheads increased	96
Direct Labour decreased	92
Inventory decreased	96
Maintenance increased	88
Production decreased	92

In paragraph 2.1 the costs of machine tools themselves are seen as decreasing.

COMMENT

The results of a Delphi-type forecast are based on expert knowledge and therefore represent the best available opinion over a given period of time. They

have no statistical significance and the appearance of precision and accuracy in Tables 3(a), 3(b), 4(a), 4(b) and 6 should not be misconstrued; they are at best estimates with wide tolerances. However Delphi forecasting has become so common that these cautionary remarks are hardly necessary.

It is useful to compare the NEL/MTIRA results with those of CIRP^{6,7} and an earlier Delphi forecast by NEL¹. There is a common measure of agreement concerning integrated manufacturing systems in which various related functions are automated. Significant practical advance is seen as being in evidence by 1980 whereas the fully automated factory still remains so far ahead in prediction as being outside the scope of forecasting—1990 and beyond.

The NEL/MTIRA survey forecasts that by 1980 various functions such as tool-life optimization—dependent on incorporation of machinability and technology data banks and on adaptive control—tool inventory control selection, operations planning and design will be integrated into second level and hierarchical systems. There is a puzzling and contradictory view (Table 1, item 2.17) that adaptive control is important for CNC and first level DNC but less so for second level DNC and hierarchical systems. This is not properly understood but the views here have since been reaffirmed by respondents

CIRP also saw 1980 as the date by which extensive development of manufacturing software and

TABLE 6 Related activities

Area	Essential	Major importance	Relatively important	Relatively unimportant	Minor importance	Of no importance	Unnecessary
Output from machine tools	17%	71%	12%				
Design office activities	17%	53%	30%				
Finance and costing	30%	57%	9%	4%			
Production planning	48%	52%					
Production control	66%	34%					
Sales and marketing		56%	44%				
Corporate planning	17%	58%	21%				4%

optimization of tool-life productivity and surface quality will have occurred. By this time in-process inspection and on-line optimization will also be widely used

This will entail on-line process identification and control so as to adapt the manufacturing process to manufacturing demands. Less clear is the date when more than 50 per cent (say) of new machine tools for use in engineering manufacture will be NC/CNC. All three surveys appear to put this at around 1990—a period so far ahead as to be outside useful forecast. The growth rate does not appear to be fast nor does the hard-wired versus soft-wired aspect appear to influence the growth rate or the date.

ACKNOWLEDGMENTS

The authors wish to express their thanks to the international group of respondents who gave their time to this forecasting exercise.

This paper is published by permission of the Directors of National Engineering Laboratory and

Machine Tool Industry Research Association. Crown Copyright Reserved.

REFERENCES

1. G. Sweeney and J. W. Bruce. Prediction of the future of computer-controlled machine tools. *Proc. Conf. Computer Control of Machine Tools*. East Kilbride, Glasgow: Birniehill Institute NEL, 1971.
2. OECD *Technological Forecasting in Perspective*, 1967.
3. University of Strathclyde. *Eur. Conf. on Technological Forecasting*, 1968.
4. J. W. Bruce. *A long-range technological forecast for the industrial application of small computers*. East Kilbride, Glasgow: Birniehill Institute NEL (including a comprehensive list of references), 1971.
5. Anon. University of Michigan Forecast Technology Advance *NC Scene* (Dec. 1973).
6. M. E. Merchant. Delphi-type forecast of the future production engineering. *CIRP Annals* (1971) **20**, 213-225.
7. M. E. Merchant. The future of batch manufacture. *Phil. Trans.* (1973) **275** (1250), 309-424.

CAD AND CAM

COMBINATION OF CAD AND CAM IN MACHINE TOOL MANUFACTURE

by

P. STÖCKMANN* and G. WÖLLNER*

SUMMARY

The machine tool is only one link in the various stages of a production process, and the numerical control of machine tools was only a single example of the application of electronic data processing to a production process. It was followed by the computer-aided production of punched tape as the next link in the chain of data transfer. In the course of the extending use of computers in production, further control functions such as AC, and particularly inspection and checking functions (for example, dimensions, surface finish, quantities) have been taken over by computers. Included in the field of CAD are design calculations for machine parts and the automatic production of drawings. The intended aim of CAD and CAM is an integrated data flow from an order through design and production planning to production and inspection. The production of control cams in a particular machine tool works before and after the employment of a computer, and matched production is described here as an example of such a combination of CAD and CAM.

The appearance of the first examples of NC machine tools twenty years ago ranked as a huge development. This combination of electronic data processing equipment with machine tools constituted the ideal solution of complicated machining problems, particularly in the aero-space industry. Problems in production planning, and also in the inspection of machined workpieces, soon showed that the application of computers to machine tools in isolation could not endure for long. It was recognized that NC machine tools form only one link in a chain in which electronic data transfer integrates all stages of the production process¹. This data flow is shown in figure 1 as a process emanating from the market in the shape of an order, and proceeding through the design planning and production stages. After inspection the result returns to the market.

Ideally, this cycle should be controlled by an uninterrupted data flow. The high degree of development of electronic data processing in NC machine tools makes it logical to employ electronic data processing methods and equipment for the data flow. The application of electronic data processing to the preparation of punched tapes in the production planning stage was already developed under the CAM concept soon after the introduction of NC machines. A multitude of systems for the most diverse problems were developed.

The best known is APT and its variants. These developments, which originate mainly from the USA, concentrated on the solution of complicated geometrical problems, while parallel work in Europe was

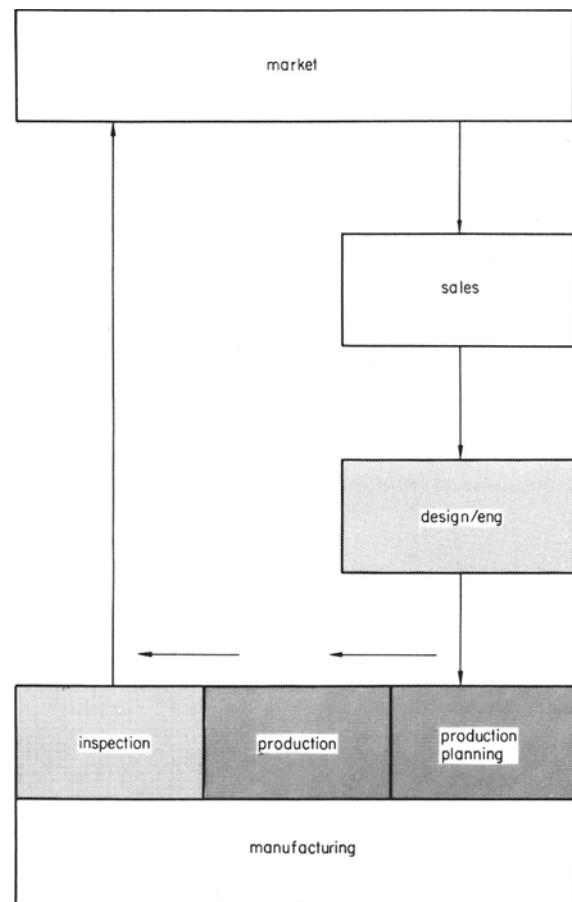


Figure 1. Information flow chart.

* Pittler Maschinenfabrik AG, Langen, W. Germany

directed increasingly towards the combined solution of geometrical and technological problems in the automated production of punched tapes, in order to satisfy the European aims of applying NC to the automation of small batch production². Thus AUTO-PIT, EXAPT and NEL programmes are extensively employed nowadays.

In parallel with this extending application of electronic data processing to more and more production plant, electronic data processing is also being applied to design. In the German Federal Republic an extensive research programme, supported by the German Ministry for Science and Technology, is in progress at the moment. The scope of the problems, of which several salient points are being attacked, is shown in figure 2. The final aim of all these efforts on CAD and CAM is an uninterrupted electronic data flow.

We have to manufacture a large number—ca. 2000 a year—of cams in the production of our cam-controlled multi-spindle automatics (figure 3). At the beginning of

total system	working system	auxiliary system
compatibility of programming languages structure of data total operating system user languages	<ol style="list-style-type: none"> <u>determination of application and working principle</u> search for solution principles <u>calculation and design</u> calculation of single parts and systems design of variations logic of design of single parts and sub-assemblies coupling of single parts and assemblies by dialogue with computer determining calculations of partial systems tools and fixtures design <u>preparation of production data</u> analysis of manufacturing conditions establishing of parts lists and work sheets by computer development of graphical display methods computer-assisted drafting 	<ol style="list-style-type: none"> <u>information</u> compilation of methods compilation of existing programs information systems independent from corporation information systems restricted to corporation <u>organisation</u> getting up a user organisation checking problems of users against existing programs user training

Figure 2. Main problems in CAD and CAM.

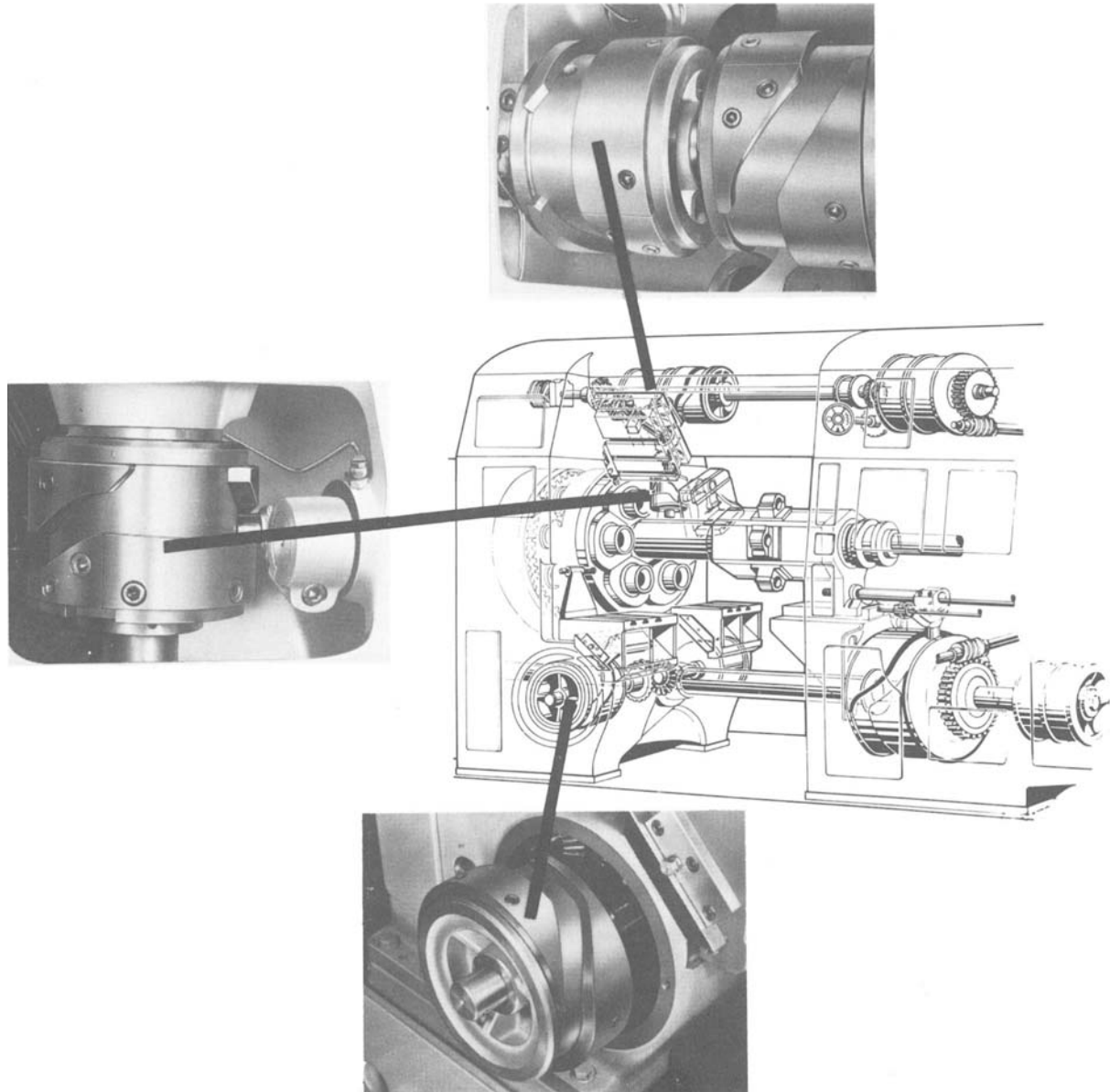


Figure 3. Cams in multi-spindle automatics.

1974, this production process was modernized and the work in the design office was also simplified appreciably by CAD. The problem is to manufacture cylinder cams from rings. The development of these cams is shown in figure 4. The light parts of the cam on the

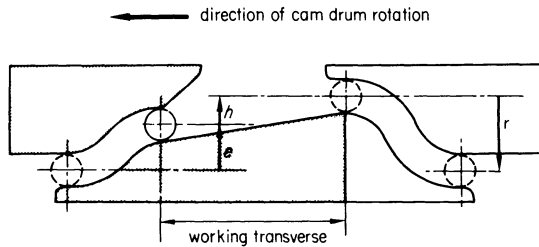


Figure 4. Cam layout.

right-hand side always remain unchanged, while the dark parts of the cam on the left-hand side have to be adapted to the particular requirements. Thus, h corresponds to the movement of the slide traversed at the machining feed rate, e is the approach of the slide at the rapid traverse rate and r is the return travel of the slide.

The conventional production method used so far will be described in order to facilitate understanding of the integrated computer-aided production. In figure 6 the individual departments involved are shown on the left, the locations of the work are indicated in the centre, and activities and media associated with data transfer processing are shown on the right. In the machine or tool design office, the track of the cam used to be drawn with compasses and ruler. The contour consisted of circular arc and straight line elements. The information on the drawing thus produced was used in the production planning department for drawing and cutting out

- (1) the marking out stencil,
- (2) the milling template,
- (3) the flame hardening template,

for the individual cam parts, as well as for producing

- (4) the drawing for the photo-electric tracer controlled flame cutting machine.

The milling and flame hardening templates were stuck on sheet metal plates in the tool room and cut out manually by means of a band saw. In production, the all-over machined cam ring was first marked out in accordance with the paper stencil and then flame cut in accordance with the flame cutting drawing. As the next operation, the contour was finish-milled on a copy milling machine and the sections of the contour were then hardened on a tracer-controlled flame hardening machine. Finally, the contoured surface was finished by a polishing process.

Cams produced by this method had contours composed of circular arcs and straight lines. The method of producing the templates resulted in a geometrical accuracy of ± 0.25 mm. In addition, a circular arc contour results in appreciably higher acceleration than other contours (polynomial or sinusoidal). The increased demands made on machine tools have in this instance been responsible for raising the rapid traverse speeds of slides by 45 per cent to

35 m/min. This in turn has increased the inertia forces, and hence wear and machine vibration, so much that it became necessary to produce

- (1) more accurate cams, and
- (2) cams of different contours.

As a first step in the development of new production methods, second degree polynomial curves (parabolas) were used instead of circular arcs for those parts of the cam which produce the rapid traverse, in order to reduce the acceleration peaks.

Cams with parabolic contours produce the least inertia forces (figure 5). As cams with such contours

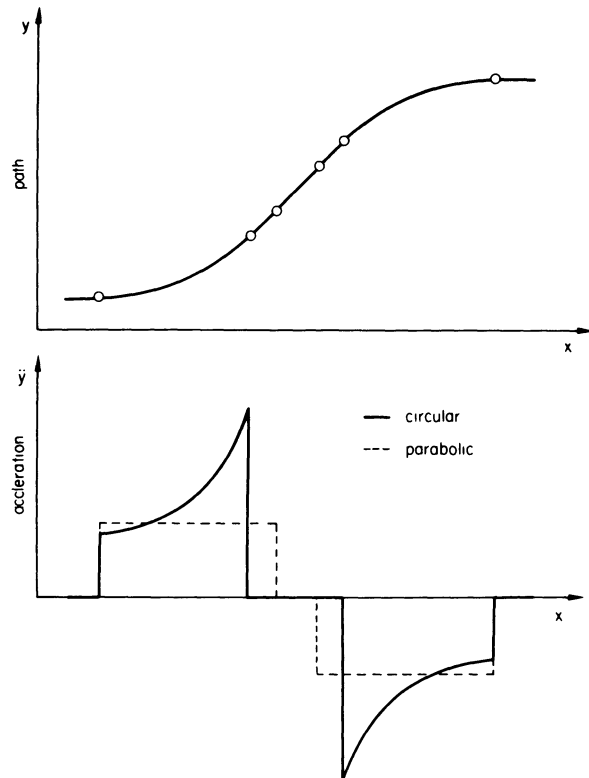


Figure 5. Paths and acceleration for various cam shapes.

could no longer be produced manually, the MERKUR computer programme (*Maschinelle Ermittlung von Kurven*) was developed for computing the mathematical equations of the track centre-line. As a second stage in the development, the data produced by the MERKUR programme was processed by a post-processor, so that the milling and hardening templates could be produced automatically by an NC nibbling machine. As a third stage in the development, the production of the drawing was simplified by producing the punched tapes required for the operation of a plotter by a further computer run with AUTO PLOT. The plotter produces the design and flame cutting drawings automatically (figure 7).

In order to achieve an accuracy of the cam track of ± 0.025 mm, the copy milling machine had to be replaced by a three-axis NC milling machine as a fourth step. This is the present state of development of cam manufacture. As a further development (figure 8), and in order to standardize the flow of data, it is intended to convert the flame cutting and hardening machines to NC.

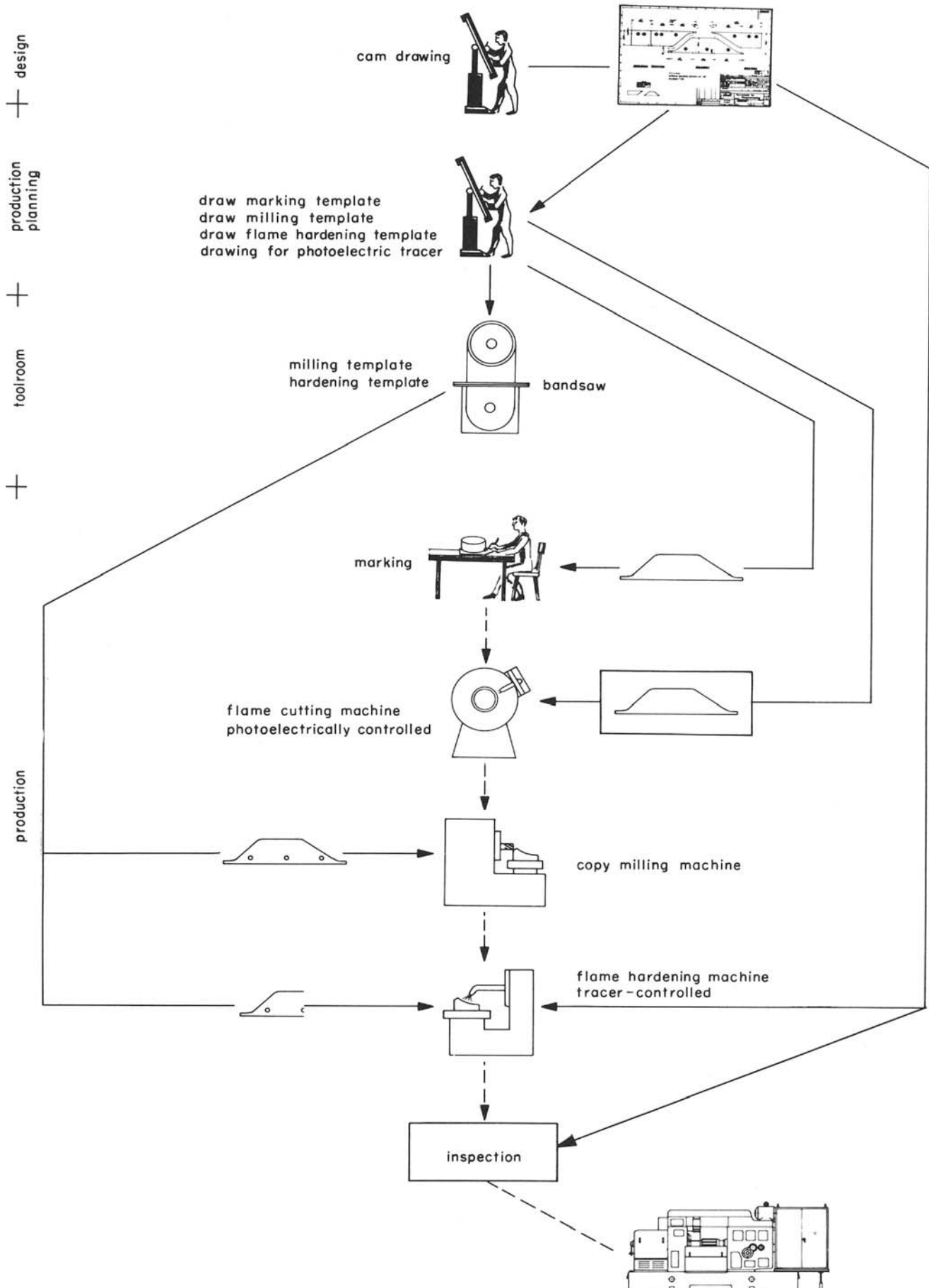


Figure 6. Conventional manufacture of cams for multi-spindle automatics.

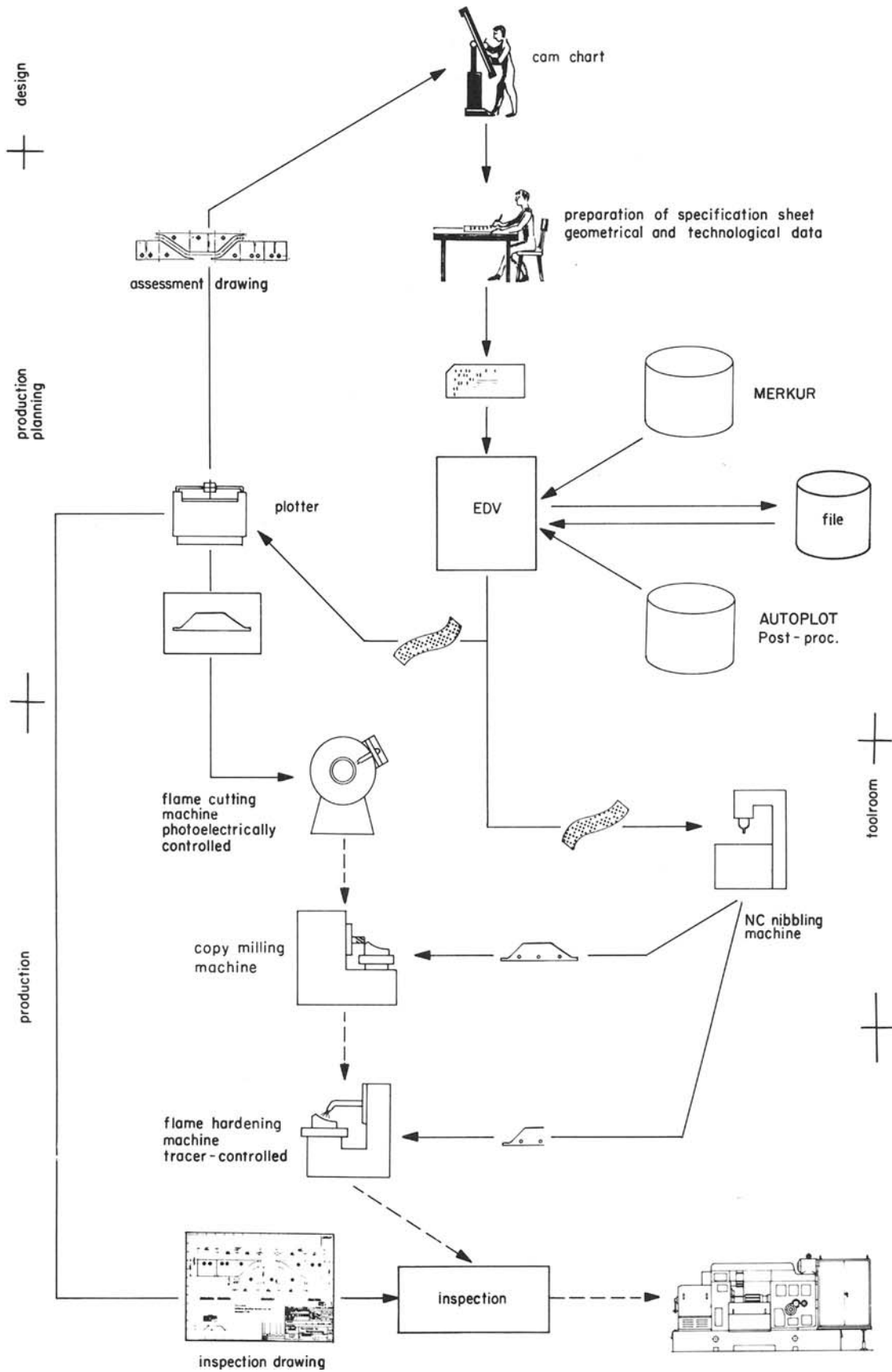


Figure 7. Actual manufacture of cams for multi-spindle automatics.

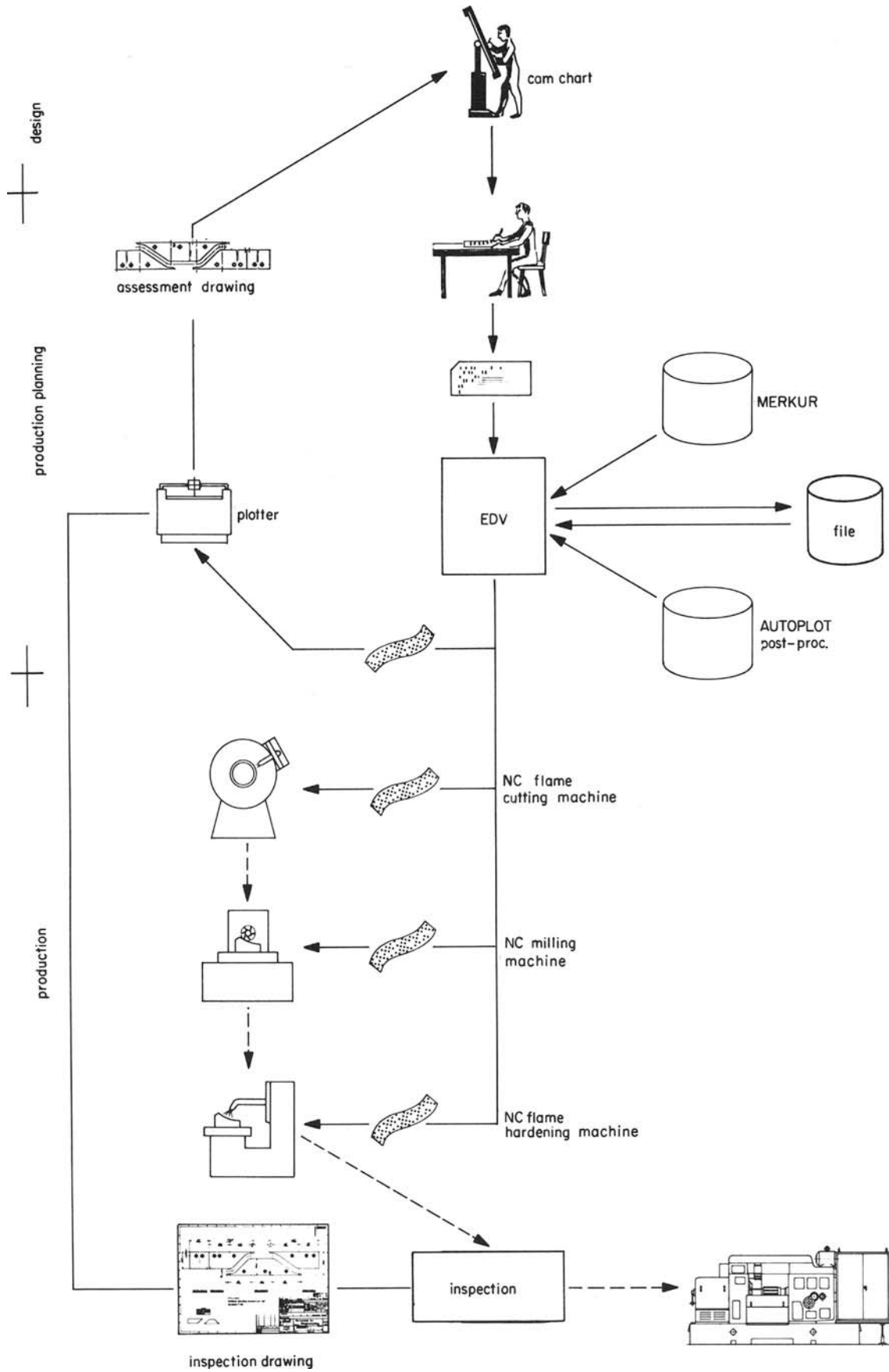


Figure 8. Future manufacture of cams for multi-spindle automatics.

The initial data is provided by the design department in the shape of a cam chart compiled from the slide travels required. This is punched on cards in the production planning department and then processed with the aid of MERKUR and AUTO PLOT. The drawing produced by the plotter can, if required, be submitted first to the design department for assessment and any necessary alterations. In addition, it serves as a reference for production and inspection purposes. The punched tapes for the flame cutting, milling and hardening machines are then produced by the various post-processors. The accuracy of the cams produced in this manner is ten times as great as that resulting from conventional production methods.

CONCLUSIONS

Numerous manual operations are made obsolete by the integrated electronic data flow. The time which elapses between receiving the customer's order and

delivery is also reduced. The labour-intensive manufacture of production aids is eliminated. The quality of the cams produced is appreciably better and more uniform, as it is unaffected by human factors.

This example from practical experience shows that CAD and CAM can be applied in many ways. The combination of both processes obtains the greatest advantages from an integrated data flow and relieves personnel extensively from routine work.

REFERENCES

1. P. Boese, K. Roschmann and P. Stöckmann. Der praktische Einsatz numerisch gesteuerter Werkzeugmaschinen im Betrieb. *Werkstattstechnik* 54 (1964) 10, pp. 483-490.
2. P. Stöckmann. Consequential application of the electronic data process for controlling and programming lathes. In *Advances in Machine Tool Design and Research*. Pergamon Press, London, 1965.

THE APPLICATION OF CAD TECHNIQUES TO MACHINE TOOL COMPONENT DESIGN, PRODUCTION AND MANUFACTURE

by

A. JEBB*, C. B. BESANT* and R. C. EDNEY*

SUMMARY

The falling costs of mini-computers in recent years has resulted in the introduction of low cost fully interactive CAD work stations in industrial drawing-office environments as well as in the pure design area, since much of the tedium in producing good quality drawings and documentation can be removed from the draughtsman.

Work at Imperial College resulted in the production of the CADMAC low-cost interactive computer-aided design system which incorporates a mini-computer, a visual display unit, plotting and digitizing tables. The use of a disc-operating system and overlay technique has led to the development of several application packages ranging from architecture to stress analysis.

The work described concerns the production of a software system for the CADMAC system which would allow mechanical component designers to produce drawings in much shorter times than by existing manual methods and at a lower cost. The software allows the user to create a library of previous and new designs which form part of a larger data base. The library can be used in a design retrieval mode which allows designers to examine quickly all relevant previously designed components, and should result in a reduction of the number of commonly used items. The data base also contains information on the tools, machines and methods used to manufacture a particular component. The components are classified according to manufacturing as well as functional design aspects.

The advantages of NC and DNC machine tools are well known but it is important to summarize one area where significant time and cost is involved. This concerns the production of NC machine tapes which involves:

- (1) the preparation of an engineering drawing and the necessary data for the machine tool operation;
- (2) the preparation of a process sheet from the drawing incorporating all the information that must be transferred to tape;
- (3) the conversion of the process sheet into a process tape for the computer;
- (4) the conversion of the process tape by the computer into a binary code for the control tape.

In addition to the direct use of the CADMAC system in the design of components for production by machine tools and the production of drawings, the authors have investigated the direct production of NC control tapes, circumventing intermediary processes. The elimination of these processes both speed up the overall process and bring the design and production engineers into much closer contact.

Throughout the work described the emphasis has been placed on:

- (1) low cost;
- (2) elimination of the need for the user to have computer programming or operational experience;
- (3) the retrieval and use of previous designs and production experience.

INTRODUCTION

The sophistication of the APT programming language and similar languages is often not necessary for the programming of numerical control equipment to produce a wide range of components. In addition, the use of these programs necessitates a knowledge of digital computer programming. This explains the current interest in developing simple and easy-to-use languages for producing NC tapes¹.

Using these methods, it is still necessary to prepare a process sheet for each component and then to

produce a process tape containing geometric and technical information. Finally, the NC control tape is prepared by a program which is run on a large time-sharing computer system.

A new method of creating an NC control tape is to perform the required component design on a fully interactive CAD system. The design is then in digital format and the NC tape can be produced by suitable processing of the digital information.

To achieve this a fully interactive post-processor is being developed for the CADMAC^{2,3,4} CAD system. No knowledge of programming is required to use the

* Department of Mechanical Engineering, Imperial College of Science and Technology, London

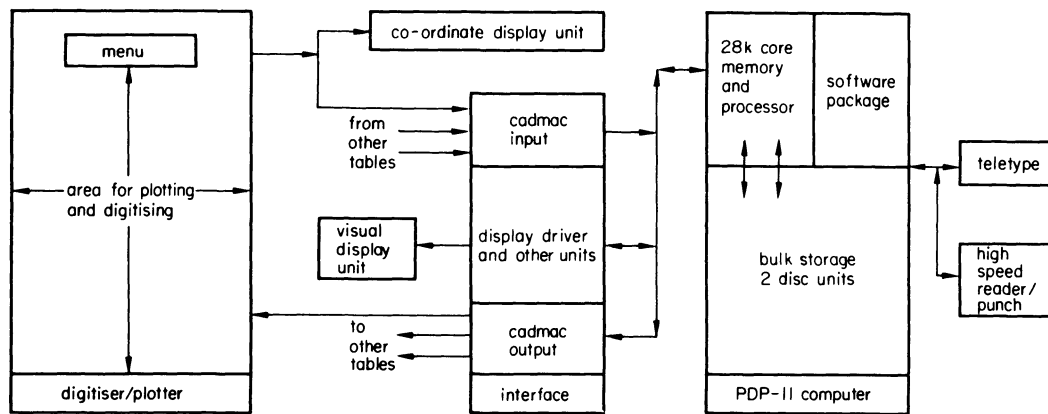


Figure 1. The CADMAC CAD system.

system, the NC tape being produced by a production engineer interacting with the system. At present only point-to-point operation is possible. It is envisaged that batch runs on a large computer will still be needed for the 'number crunching' associated with continuous path operations. However, the CADMAC system will be used to simplify considerably the data preparation stage.

CADMAC-11 GRAPHICS

Figure 1 shows the CAD system used for this work. It consists of a PDP-11/40 with 28 K of core, two 1.2 million word discs, a storage or refreshed-type visual display unit and a digitizing-plotting table. An engineering drawing can be digitized from the table and stored as points in I, X, Y co-ordinate form, one line in the stored data for every line in the drawing. In practice assemblies, components, dimension lines, scrap views, and so on may all be stored in different data 'levels' so that they can be easily separated and distinguished.

Component designs can be quickly built up in the computer using the MENU commands to input and position lines, circles and other symbols. Figure 2 shows a simple drawing produced in this way. An interesting feature is the system library which permits high-speed design information retrieval from a data base. Standard components such as spacers can be retrieved from the data base and tried in an assembly.

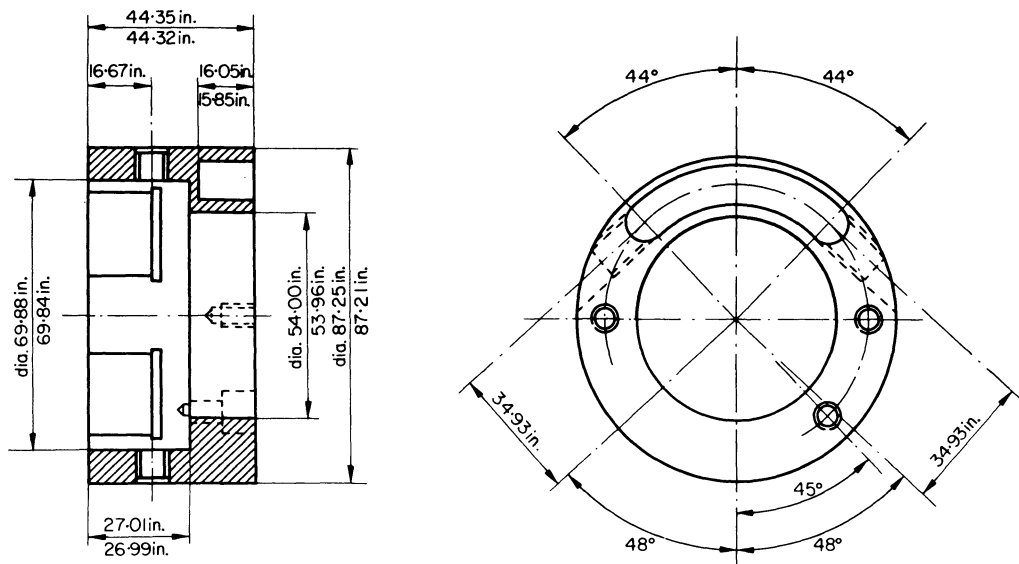


Figure 2. Type of drawing which can be easily digitized into the system.

The component can then be included in the design or rejected, or another component tried at the discretion of the designer. It is possible to try all the relevant existing designs in the data base, and to prove that a new design is necessary should this be the case.

INFORMATION RETRIEVAL

The system incorporates two discs, one to contain the programs and the other to be used for the data base. In the design retrieval mode components are classified in a library with a modified Opitz-type code⁵ which is a classification by design use. To retrieve a specific component type the design engineer enters the relevant reference code number via the teletype. This initiates a search of the system data base. A rapid data base search and fast information retrieval is achieved by the use of a disc and random access files.

The data base system serves a dual function since production information is also stored. The latter provides answers to questions similar to the following:

- (1) How was the tooling done for a particular job?
- (2) How was a similar feature cut on a previous job?
- (3) How often has use been made of this particular tool?
- (4) What advantage can be made of previous experience in tooling for the present job?

In order to answer these questions it is necessary to classify the components according to manufacturing aspects as well as design use. When a new component is to be tooled the data base can be searched and any relevant previous experience retrieved.

The value of this type of information retrieval depends entirely on the coding system employed. If a component is wrongly coded, and so put in the wrong group, it will be lost in the system.

It is necessary to decide what use you wish to make of the data and then choose a suitable coding system. For design purposes, we need to classify components according to their use or function. However, for production purposes, we need to classify the components according to how they are made. This poses a more difficult problem and a suitable coding system is being looked into.

A tooling file exists which stores information on tools according to their drawing numbers, that is in no particular order, and contains offset data, outline drawings and other necessary information on each tool. This file is referenced during the NC tape preparation process, the required tools being extracted from the file and displayed on the visual display unit.

DIRECT PRODUCTION OF NC TAPES

Once a design has been created within the computer and stored in I, X, Y co-ordinates several difficult problems have to be tackled to computerize fully the NC tape production process. The principal problems are:

- (1) All drawings are in two dimensions whereas components are in three dimensions. A sophisticated program is required to compile a three-dimensional representation of a component in I, X, Y, Z form from two digitized elevations in I, X, Y form.
- (2) Distinctions have to be made as to which surfaces require machining since tighter tolerances may be required for some dimensions. Additionally, components are often made in several operations, frequently on different machines.
- (3) It is impossible to digitize a drawing accurately enough for the X, Y co-ordinates to be used directly on the NC machine. Consequently these approximate X, Y co-ordinates need to be modified before the NC tape can be used.

A solution to these problems has been found by using an interactive graphical program which simulates the movements of the machine. Figure 3 gives a flow chart for programming a piece in the conventional manner and figure 4 a flow chart for the same job using the simulator.

When an elevation is digitized into the computer the designer starts by inputting reference points and lines, such as centre lines, whose co-ordinates in the data remain fixed. The basic outline is built up and by using MENU commands such as 'DIMENSION RADIUS', 'DIMENSION LINE', the designer can type in the accurate dimensions required. The pro-

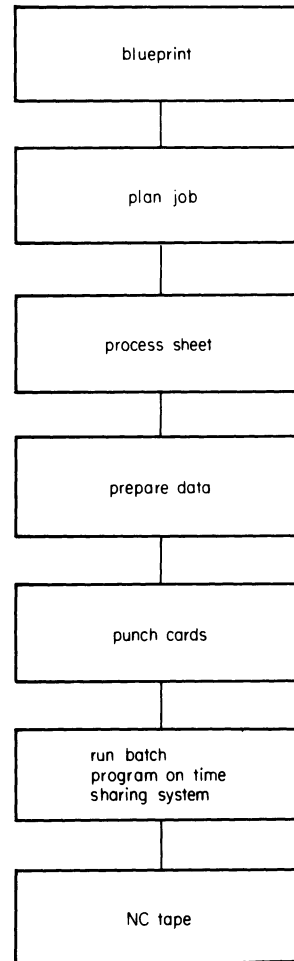


Figure 3. Flow chart for programming component by conventional means.

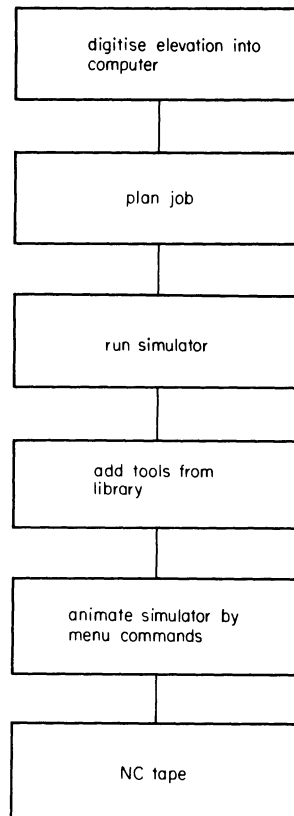


Figure 4. Flow chart for programming a component using the simulator.

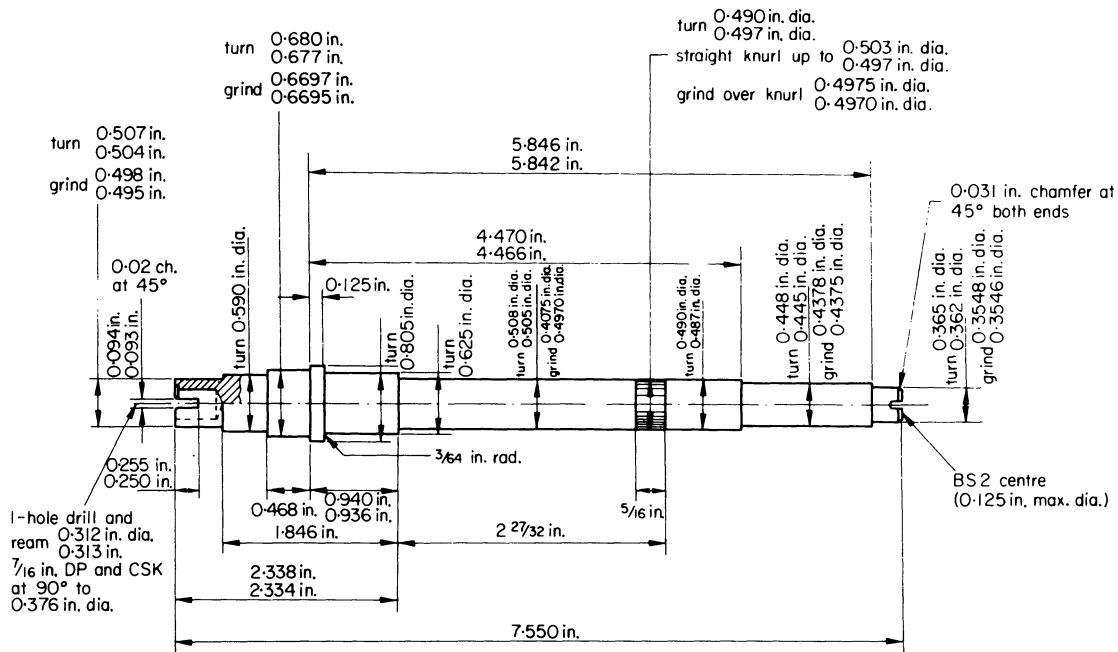


Figure 5. Component which has been programmed using the simulator.

gram then automatically adjusts the position of the digitized outline so that the distance between the outline and reference points corresponds to the typed value. Tolerances are specified by typing the maximum dimension preceded by a plus sign and the minimum dimension preceded by a minus sign. It is at this stage that problem (3) is overcome, with no conscious effort on the part of the user.

Initial attention has been confined to producing tapes for an NC lathe. Since all surfaces cut on a lathe are solids of revolutions, only one elevation is needed to define them; hence problem (1) is avoided rather than solved. When producing a tape, a production engineer first chooses which machine he is going to use and then the tools, and finally he defines the process. Thus problem (2) can be solved by the engineer interacting with the program.

When the machine has been chosen a diagrammatic representation of the slides and chuck/bar feed is displayed on the visual display unit. The engineer then selects the necessary tools and inputs any necessary qualifying dimensions. A file containing the tool number, machine station and setting dimensions for each tool used on the job is created at this stage. This is printed out for reference purposes for the setter-operator, and is also used by the system for feeds and cuts. When a particular machine has been selected, datum points and the fully retracted positions of the slides are retrieved from the program library.

By using such MENU commands as 'INDEX X TURRET', 'X TURRET TO CURSOR', 'COOLANT ON', 'COOLANT OFF', 'RAPID FEED' and so on the engineer can animate the simulator and readily see the results of the programming so far.

The simulator can also be 'reversed' and used to plot out the path of the cutting tools as defined by a given tape; this it is possible to check out a tape quickly and accurately, mistakes in the tape showing up on the screen.

Figure 5 shows a component programmed by the simulator. The total time from drawing to component

by conventional means was 5 hours. Using the simulator the same tape was produced in 2 hours.

CONCLUSIONS

Computer graphics using a dedicated mini-computer has much to offer the design and production engineer working in the field of mechanical component design. In the design function much of the dull repetitive draughting can be carried out by the computer, leaving the designer free to concentrate on the more creative aspects of his work. One important advantage of CAD is that component designs can be retrieved quickly from a large data base and variety reduction can be achieved.

By using a graphical simulator NC tapes can be produced more conveniently than by conventional means, since data preparation is greatly simplified. The time taken to program the component shown in figure 5 has been reduced from 5 to 2 hours, which included 1 hour to plan the job. Much of this reduction in time is due to the fact that all the necessary information is filed in the computer and the engineer need never leave the machine or experience any delays until the job is complete.

ACKNOWLEDGMENTS

The authors would like to thank all those in the CAD unit at Imperial College for their help, and Messrs Alfred Herbert Ltd for their support and encouragement.

REFERENCES

1. W. H. P. Leslie. *Numerical control programming languages*. North-Holland Publishing Co.
2. C. B. Besant, A. Jebb, A. Hamlyn, B. H. Koh and R. E. Grindley. *Cadmac-11*, a fully inter-

- active computer aided design system. *Journal of Computer Aided Design* (1972) 4, (5).
3. C. B. Besant. Computer aided design. *Chartered Mechanical Engineer* (February 1973).
 4. C. B. Besant, A. Hamlyn, A. Jebb and P. McClintock. The use of CADMAC systems in general draughting. *Journal of Computer Aided Design*.
 5. F. W. Craven. Private Communication. Alfred Herbert Ltd, 1974.

AUTOMATIC DESIGN OF MACHINE TOOL COMPONENTS

by

TRUMAN G. FOSTER*

SUMMARY

Techniques of computer-aided design for the automatic design of machine tool components are discussed. Two particular components are highlighted:

- (1) the automatic design of compact spur gear sets
- (2) the automatic design of hydrodynamic oil journal bearings

The main features illustrated are:

- (1) automatic iterative solutions using the IBM S/360 continuous system modelling program (CSMP)
- (2) emphasis on a method to include various non-linear functions using the CSMP AFGEN techniques
- (3) emphasis on inclusion of both the special logic of CSMP and the regular FORTRAN logic to accomplish automatic iteration
- (4) emphasis on equation order technique.

Two example problems are included which illustrate the details of the automatic design techniques advocated.

INTRODUCTION

Iterative design using the computer is emphasized in this paper. A system of equations is solved for fixed inputs, and the design is then iterated by changing some of these inputs. In automatic iteration or automatic design some of the inputs are changed automatically by the addition of computer logic. Besides the iteration situation there are a number of other situations where the use of the computer can be advantageous. If calculation by 'hand' is tedious, then the design problem could be a candidate for solution by the computer.

Often design problems can be treated by what Offner¹ has called 'Brute Force' optimization. In this approach, which is related to the iteration approach, an important input parameter is varied in a discrete manner; each new value producing a new solution and each solution representing a modified system. Thus a search is made for an admissible solution which satisfies some important performance criterion or criteria.

The purpose of this paper is to exhibit a few examples where some automatic design techniques have been applied to machine tool components. Chosen for exemplification are the automatic design of a set of spur gears where compactness is the design objective, and the automatic design of an 'oil-ring' hydrodynamic journal bearing where low oil-film

temperature is the criterion. In both of these examples a system of algebraic equations is automatically solved and iterated until the design criterion is met. The system of equations utilized in each of the computer programs discussed are taken, for the most part, from standard references. Thus herein no new theory is developed, however, new relationships easily could be incorporated into these computer programs.

UTILIZATION OF IBM S/360 CSMP TO SET UP AUTOMATIC DESIGN PROGRAMS

Some designers prefer to write a complete computer program including the required formatting for INPUT/OUTPUT. This should be encouraged. However, the advent of programs, wherein the INPUT/OUTPUT formatting is almost automatically taken care of by the program, has facilitated computer-aided design. IBM S/360 CSMP⁴ is such a program; the INPUT/OUTPUT formatting has been simplified, but the program retains considerable generality and both the special logic statements of the CSMP language and regular FORTRAN logic statements can be incorporated into the same program. The utilization of this program facilitates automatic design and encourages the development of general programs. For example, if a basic design objective, such as a low film tempera-

* Department of Mechanical Engineering, The Ohio State University, Columbus, Ohio, USA

ture criterion for a journal bearing, has been established then a program can be written which automatically adjusts such dependent variables as the oil film temperature, oil viscosity, and film thickness (relative eccentricity). Furthermore, it is a general solution. Not only can one specific problem, but also all such problems in a certain class be handled in a most routine fashion with accuracy and great speed.

In another example, an automatic program was written to solve for the design parameters for a set of straight-tooth spur gears based on a compactness criterion. The iteration is based on automatic adjustment of the so-called *diametral pitch*. Also, the estimated error in gear tooth manufacture is adjusted automatically by the logic of the program.

By using computer-aided design techniques, machine tool designers can model their design problems in a very sophisticated manner. Various types of non-linear functions can be handled by using the CSMP AFGEN statements. The output of the program can be automatic, and thus parametric studies can be made which are aimed at the discovery of the optimum solution or at least an admissible solution. Usually the actual computing time requires only seconds, and the 'turn-around time' for most programs is only minutes and at the most a few hours; the latter caused by queuing delays. Thus in a short period of time the essential features of the design performance of a machine tool component can be determined, and the effect of modifications investigated.

THREE TECHNIQUES FOR AUTOMATIC DESIGN

The AFGEN curve technique

IBM developed the CSMP arbitrary function generator, AFGEN, to represent non-linear functions. Example problem 2 exemplifies the situation where the phenomenon is represented by a system of n algebraic equations in n unknown variables, where $n - 4$ of the equations are linear and four are non-linear. The four non-linear equations are actually non-linear functions, and can be represented by curves; actually by a series of pairs of numbers (x_i, y_i). Thus four graphs are entered into the automatic design program as AFGEN CURVE 1, AFGEN CURVE 2, AFGEN CURVE 3, and AFGEN CURVE 4. Trial values of one or more dependent variables are selected to start an automatic iteration process. The CSMP RERUN statement was used to accomplish the iteration, and regular FORTRAN logic was added to test the convergence and terminate the iteration. See example 2 Automatic Bearing Design for an example where ten relationships in ten unknowns was solved in four iterations. In this example, the two dependent variables chosen to start the iteration were oil film temperature, (TF) and the relative eccentricity (ECC). The primary iteration variable was TF, and the logic involved in the iteration process tested TF for convergence, and terminated the automatic iteration process when the error in TF was less than 0.05 degrees fahrenheit.

The equation order technique

For the design situation represented by a system of n non-linear equations and functions in n unknowns,

the computer program statements are arranged in a proper order so that they are fed properly into each other. A method called NOSORT can be followed, or the CSMP-SORT procedure could be used. The latter method depends upon the CSMP program to SORT the statements and place them in their proper order. However, from hereon emphasis is placed on using the NOSORT approach and placing all statements in their proper order. Furthermore, if FORTRAN logic is to be employed it is imperative that the NOSORT approach be utilized.

Of course, the basic objective is to write the computer program so that the equations of the system are solved simultaneously. To this end, all that is required is to list the equations in the reverse order in which they were first written, and to ensure that any constants or variables required in each statement are listed earlier in the program. This simple rule has some important ramifications. Combining equations to reduce the number of dependent variables is not in general recommended. Not just because it is unnecessary, but more importantly, if equations are not combined, it is possible to ask the program to print out or plot out various dependent variables, sometimes just for a few iterations. Outputting the dependent variables is in itself important, but also it can be helpful in a *debugging* process because it permits focusing the designer's attention on each step of the analysis. Furthermore, standard diagnostics within the CSMP program will direct the programmer not only to the statement containing the error, but also to the position within the statement where the error has occurred.

Techniques of 'brute-force' optimization

In general, design problems require that certain variables be fixed. For example, in the journal bearing design program, two variables which are fixed are the 'type-of-oil' and the L/D ratio. It is a relatively simple matter to change the L/D (length over diameter) ratio by individually changing L and D in the input to the computer program. However, a new AFGEN CURVE 2 would have to be supplied since it is a function of the L/D ratio. Changing the type-of-oil requires entering a new AFGEN CURVE for the viscosity versus temperature relationship. Also new values for CP (specific heat) and GAMMA (specific weight) would be required. These types of changes and others can be considered as a crude search technique to find an optimum solution. As previously mentioned, this approach might be termed 'brute-force' optimization. In setting up a computer program, a very general approach usually can be used, and a new problem can be solved immediately by simply changing a card or two.

EXAMPLE PROBLEMS

Example 1—Compact spur gears

Using standard gear equations from Spotts² and a compactness criterion, a system of equations was devised to calculate the diametral pitch. This system was solved using the CSMP IMPLICIT statement which is a call for a subroutine which solves an implicit equation by an iterative trial-and-error

method. Also, iteration on the total gear tooth error (EG), sum of two errors for the two gears, was introduced by utilization of the CSMP RERUN statement. Thus two dependent variables DP and EG were chosen for the iterative automatic design. Arbitrarily chosen initial values, even unreasonable values, can be chosen for these variables and yet the computer program will calculate automatically the correct values.

To illustrate the use of the computer program the solution to the following problem will be discussed.

Problem statement

Design a set of spur gears to comply with the conditions given in Table 1. The design is to be based on a compactness criterion; that is, the number of teeth on the pinion is to be the minimum value recommended for the 20° involute full-depth system. Also the contact ratio and the bending strength capacity are to be checked.

TABLE 1 Example 1—automatic design, compact spur gear set

	Design conditions
1 Power	HP = 250.
2 System	Prang = 20°, involute full-depth
3 Material	Induction hardened steel
4 Gear ratio	GR = 2.
5 Pinion speed	N1 = 1725, rpm
6 Quality	AGMA No. 8
7 Tooth width	BMAX = 4. * PI/DP
8 Minimum number of teeth on pinion	NT1 = 17.

The INPUT to the computer program is given in Appendix B. Figure 1 illustrates the principal geometric parameters for the compact gear set. Values of both the fixed parameters and the design parameters for this example are exhibited in Table 2. The condensed set of equations which form the core of the computer program are given in Appendix A. The

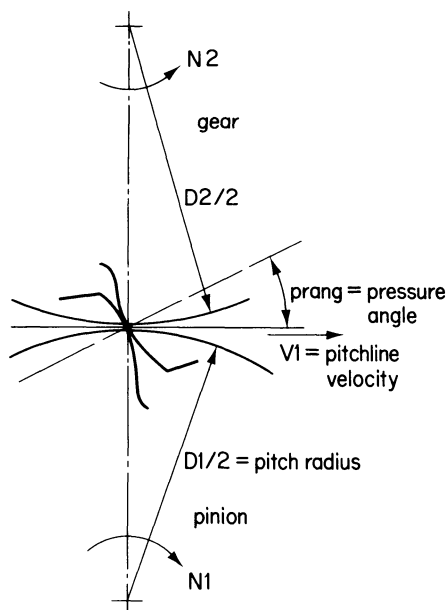


Figure 1. Geometry of the straight-tooth spur gear set.

TABLE 2 Notation for compact spur gear set

Principal fixed parameters	
NFS = 1.0	factor of safety
MA = 1.0	external impact factor
K = 417.	wear factor for induction hardened steel, Spotts ² , Table 10-2, p 392
N2 = N1/GR	rpm of gear
NT1 = 17.	minimum number of teeth for a 20° involute system to avoid fundamental interference
NT2 = GR*NT1	number of teeth on gear
ERROR = 0.001	error permitted in automatic calculation of diametral pitch, DP
Q = 2.*NT2/(NT1 + NT2)	geometry factor in calculation of wear load capacity, FW
S = 35 000.	fatigue strength of gear material for beam strength capacity, FB
SY = 0.096	Lewis form factor
C1 = PI * COS (PRANG)	
Principal design parameters	
BP = C1/DP	base pitch (in)
A1 = 1./DP	addendum of pinion-gear 1 (in)
A2 = A1	addendum of gear-gear 2 (in)
Z = length of path of contact (in)	
CR = Z/BP	contact ratio
CD = R1 + R2	center distance (in)
DP = NT1/D1 = NT2/D2	
EG = EG1 + EG2	error for gear 1 plus the error for gear 2, based on an assumed AGMA No. 8 quality (in)
V1 = pitch line velocity	
BMAX = 4.0*PI/DP	assumed estimate of the maximum permissible tooth width (in)
FT = 33 000*HP/V1	tangential load (lb)
DFD = estimate of dynamic load increment (amount to be added to FT to give the dynamic load) (lb)	
FD = NFS*MA*(FT + DFD)	estimated total dynamic load including a factor of safety (lbs)

rationale behind the choice of certain initial values of the two iteration variables is as follows:

- (1) The basis of the choice of $DPA = DP = 1.0$ used to start the iteration was simply that in many design situations the DP of the gear teeth would be larger than one. However, this choice really is arbitrary, and in some other situations a different choice might be made. However, the choice of zero is to be avoided to prevent division by zero.
- (2) The initial value for the total gear error was

$$EG = 0.0005 + 0.0005 = 0.0010$$

This value implies 'great precision', more than needed, and the program will return higher and more realistic values.

If the pair of gears are to be compact, that is, occupy minimum volume, then generally wear strength will govern. Thus the analysis was set up by equating wear strength to the total dynamic load. The key equation (FODP is a dummy name for DP) obtained from this balance is

$$FODP = NT1 * K * Q / FD$$

and if the indicated quantities on the right hand side are substituted (in equation form) an expression cubic in DP results. This expression is not difficult to

handle analytically, but the many substitutions are tedious and the computer can be used to make them. Furthermore, and of greater importance, once the computer program has been set up it constitutes a general solution for all sizes of spur gear sets based on the compactness criterion. It is only necessary to modify the input constants to obtain a solution to a new problem. The key equation, together with the equations that are substituted into it, were set up in a trial-and-error loop by means of the CSMP implicit statement. This is a call for a subroutine which will return a value of DP within a certain ERROR. The value of this ERROR is arbitrarily specified by the designer.

The total gear error (EG) is required in the first iteration. It is specifically required in the equation for DFD, the estimated dynamic load increment which, when added to the static load (FT), yields the estimated total dynamic load applied to a gear tooth. As mentioned above, the first iteration requires a value of EG, hence an initial value of EG has to be furnished. Referring to Appendix B, it can be observed that the TERMINAL section with its RERUN statement and FORTRAN logic was added to calculate a new value of EG based on the current value of DP. Since $EG = EG1 + EG2$ is required, EG1 and EG2 were set up as appropriate functions which approximately fit the data for AGMA quality No. 8 (Reference Table 10-3, p. 396, Spotts²). Iteration by the use of the RERUN statement then automatically permits convergence of both DP and EG, since they are interdependent.

Some statements which return the contact ratio are appended to the program. Also the bending strength capacity (FB) is calculated and compared with the total dynamic load.

The above problem statement and the design conditions given in Table 1, when introduced into the computer program in Appendix B, yield the following values of the important design parameters including

$$\begin{aligned} DP &= 4.433 \\ EG &= 0.000614 \text{ (in)} \\ FW &= 6043 \text{ (lb)} \\ FD &= 6043 \text{ (lb)} \\ FB &= 6748.4 \text{ (lb)} \\ CR &= 1.597 \\ IN &= 4. \end{aligned}$$

Here IN is the number of interactions required to converge within an accuracy of 0.0001 inches on the gear error. The computer time required was 3.23 seconds, and the program cost was 3.36 dollars.

Example 2—Hydrodynamic oil journal bearing

The performance of a hydrodynamic 'oil' journal bearing can be described by a system of ten relationships having ten unknown variables (Fuller³, 192-245). Of these ten relationships, seven are linear algebraic equations and four are non-linear relationships. The latter are represented by CSMP AFGEN statements using CSMP AFGEN CURVES. CURVE 1 returns the viscosity as a function of temperature. CURVE 2 returns ETA (the side leakage factor) as a function of the relative eccentricity ($ECC = \epsilon$) for a

fixed value of the L/D ratio. CURVE 3 returns a new value of ECC, (ϵ) as a function of A , the load parameter. A is defined as shown in the following expression

$$A = 132./ETA*(100.*M)**2*PAVG/Z/N$$

where M, PAVG, Z and N are defined in Table 4. Finally, AFGEN CURVE 4 returns a value of the coefficient KFJ as a function of ECC (ϵ), where KFJ is a friction factor for torque on a journal bearing. However, before any more computer techniques are discussed the example problem will be defined.

Problem statement

Devise a computer program to automatically design a hydrodynamic 'oil' journal bearing for the conditions shown in Table 3. Consider using SAE No. 10 oil, where the method of oil delivery to the bearing is an oil ring. Attempt to design the bearing without the application of forced flow cooling; that is, assume the bearing to be self-contained with some air movement such that the heat transfer coefficient is approximately $K = 7.0$, Btu/HR/FT²/°F. Can the equilibrium film temperature be held to less than 200° F?

TABLE 3 Example 2—automatic design, oil hydrodynamic journal bearing

Design conditions	
1 Radial load,	W = 1800 lb per side
2 Shaft diameter,	D = 3 in
3 Shaft speed,	N = 1200 rpm
4 Ambient temperature,	TA = 75° Fahr.
5 Length of diameter ratio,	L/D = 1.0, per side
6 Area of housing,	A = 1.0, ft ²
7 Ratio journal radius R, to radial clearance	R/C = 1000.

The condensed system of equations are shown in Appendix C. The INPUT computer program is exhibited in Appendix D. The OUTPUT from this program yields values of the important design parameters including

$$\begin{aligned} TF &= 195.92^\circ\text{F, film temperature,} \\ HPF &= 0.148, \text{ frictional horsepower loss per side} \\ TW &= 128.91^\circ\text{F} \\ ERROR &= TOO - TF = 0.0016, \text{ convergence error} \\ HO &= 0.00046 \text{ (in)} \\ Z &= 5.62 \text{ (cp)} \\ MU &= 8.15E-07 \text{ (reyns)} \\ PAVG &= 200. \\ IN &= 4. \text{ (number of iterations)} \end{aligned}$$

Here IN is the number of iterations required for convergence with an accuracy of less than 0.01° F on the oil-film temperature. The computer time required was 4.07 seconds and the program cost was 3.36 dollars.

Figure 2 illustrates the oil-ring lubricated journal bearing. Values of both the fixed parameters and design parameters for this example are exhibited in Table 4. Refer to Appendix D for the INPUT required for the computer program.

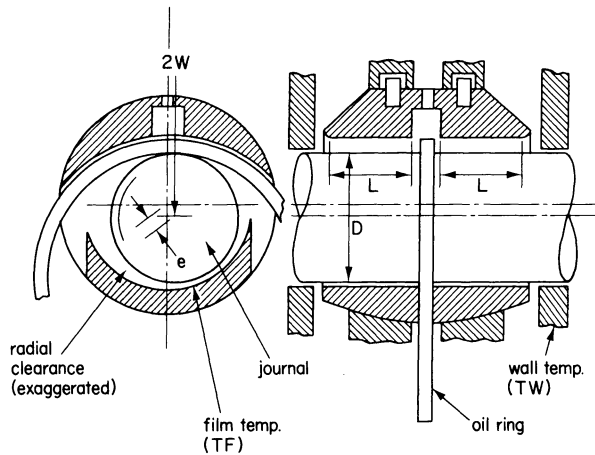


Figure 2. Principal geometric parameters for the 'oil-ring' lubricated hydrodynamic journal bearing.

TABLE 4 Notation for hydrodynamic journal bearing

Principal fixed parameters	
L = 3.	journal length per side (in)
D = 3.	journal diameter (in)
K = 7.	(Btu. hr ⁻¹ ft ⁻² °F ⁻¹)
AH = 1.0	area of the housing (ft ²)
TA = 75.	ambient temperature, (°F)
N = 1200.	journal speed (rpm)
C = 0.0015	radial clearance (in)
R = 1.5	journal radius (in)
M = C/R = 0.001	
W = 1800. lb	
PAVG = W/L/D	average load (lb in ⁻²)
M = C/R	
Principal design parameters	
MU = oil viscosity (reyns)	
Z = oil viscosity (cp)	
ETA = side leakage coefficient	
A = load factor	
ECC = relative eccentricity	
KFJ = friction factor for torque in journal	
MJ = friction torque on journal	
HPF = frictional horsepower loss	
TWA = wall temperature minus ambient temperature	
TF = oil film temperature	
HO = minimum oil film thickness	
IN = number of iterations required	

The iteration was started by selecting a mean value for ECC ($\epsilon = 0.5$) and a somewhat high value for TF = TFO = 200°F. Convergence was rapid and the cost was low.

DISCUSSION

In both examples a system of relationships was devised to realistically model a machine tool component. Careful consideration was given to which variables were independent, which were dependent, and what quantities were to be held fixed. Usually one or more dependent variables are chosen for an iteration process. This choice depended upon what quantity was chosen as the design objective.

Both examples illustrate the application of the techniques discussed in the section on CSMP tech-

niques. Iteration on one or more variables occurred in both. The equation order technique was used in both and the AFGEN representation of non-linear relationships was utilized in the second example. The examples give some insight into how to set up a system of equations to model a machine tool component using automatic computer techniques. The types of programs advocated are general programs in that a new problem can be solved immediately giving a method (somewhat crude method) of searching for an optimum solution. This has been referred to a 'Brute-Force' optimization.

As previously discussed, the order in which the equations are arranged in the computer program is a vital question. A NOSORT order was recommended and the examples illustrate how a *correct* order was established so that the computer program would automatically and simultaneously solve the system.

CONCLUSIONS

- (1) The examples illustrate the utilization of the type of logic statements required for automatic design. Both logic statements from regular FORTRAN and the unique statements within the CSMP language are employed.
- (2) The examples illustrate that automatic design programs can be developed. They show that convergence to a correct value of the dependent variable can be attained. In addition they are general solutions in the sense that a new problem can be solved immediately, provided it is of a certain class.
- (3) The examples illustrate that not only are rapid solutions possible but they are relatively economical.

The design of other machine tool components such as shafting, belt drives, and chain drives can be similarly developed. However, the design of indeterminate frames is best handled by one of the structural design codes. Some of the possible choices would be STRUDL I, STRUDL II, and SAP IV. Also the latter can be considered not only for the design of statically indeterminate frames, but also for the determination of the vibration modes and natural frequencies.

Though not illustrated here, the CSMP program also is very useful for modelling lumped-parameter systems where one non-linear differential equation is written for each degree-of-freedom. This approach is particularly useful for modelling machine tool drives so that both transient behaviour and steady-state behaviour can be determined and studied.

ACKNOWLEDGMENTS

The author would like to thank the Instruction and Research Computer Center at The Ohio State University for the use of the IBM/370/165 facility and the associated IBM/S360/CSMP program.

APPENDIX A

List of key equations for the compact spur gear program

The condensed list of equations which constitute the primary system of equations used to calculate the diametral pitch (DP) is as follows. The first statement calls the CSMP IMPLICIT subroutine.

$$\begin{aligned} DP &= \text{IMPL}(ZO, \text{ERROR}, \text{FODP}) \\ D1 &= NT1/DP \\ D2 &= NT2/DP \end{aligned}$$

$$\begin{aligned} R1 &= D1/2. \\ R2 &= D2/2. \\ V1 &= \pi * D1 * N1 / 12 \\ BMAX &= 4. * \pi / DP \\ FT &= 33000. * HP / V1 \\ DFD &= 1.46 * EG * N1 * NT1 * BMAX * \dots \\ &\quad R1 * R2 / (\text{SQRT}(R1^{**2} + R2^{**2})) \\ FD &= NFS * MA * (FT + DFD) \\ FODP &= NT1 * BMAX * K * Q / FD \end{aligned}$$

APPENDIX B

Program input for compact spur gear

```

INCON EG=0.0010,DPA=1.0
INITIAL
CONSTANT AA=.0003,BB=0.0015,CC=0.0006,E1=0.6,E2=0.2
CONSTANT NFS=1.0,MA=1.0,HP=250.,PI=3.1416,K=417.
CONSTANT GR=2.0,N1=1725.,NT1=17.,ERROR=0.001
CONSTANT S=35000.,SY=0.096,C1=2.9521,PRANG=20.
      N2=N1/GR
      NT2=GR*NT1
      Q=2.*NT2/(NT1+NT2)
      PHI=PI/9.
DYNAMIC
NOSORT
      ZO=DPA
      DP=IMPL(ZO,ERROR,FODP)
      D1=NT1/DP
      D2=NT2/DP
      R1=D1/2.
      R2=D2/2.
      V1=PI*D1*N1/12.
      BMAX=4.0*PI/DP
      FT=33000.*HP/V1
      DFD=1.46*EG*N1*NT1*BMAX*R1*R2/(SQRT(R1**2+R2**2))
      FD=NFS*MA*(FT+DFD)
      FODP=NT1*BMAX*K*Q/FD
TERMINAL
      CKK=EG0-EG
      IF(ABS(CKK).LT.0.0001) GO TO 40
      EG0=EG
      EG1=AA+BB*PI/(DP**E1)+CC*D1**E2
      EG2=AA+BB*PI/(DP**E1)+CC*D2**E2
      EG=EG1+EG2
      CALL RERUN
      GO TO 50
40 CONTINUE
      BP=C1/DP
      A1=1./DP
      A2=A1
      OA1=SQRT((R2+A2)**2-(R2**2)*(COS(PHI)**2))-R2*SIN(PHI)
      OB1=SQRT((R1+A1)**2-(R1**2)*(COS(PHI)**2))-R1*SIN(PHI)
      Z=OA1+OB1
      CR=Z/BP
      CK3=CR-1.4
      FW=D1*BMAX*K*Q
      FB=S*BMAX*PI*SY/DP
      CK1=FW-FD
      CK2=FB-FD
      CD=R1+R2
50 CONTINUE
      DPA=DP
TIMER PRDEL=0.1,FINTIM=0.2
PRINT DP,EG,EG1,EG2,CKK
END
PRINT DP,FW,FD,CK1,CK2,CD,D1,EG
END
PRINT DP,CR,PRANG,Z,BP,CK3,EG1,EG2
END

```

APPENDIX C

List of key equations for journal bearing program

The condensed list of equations which constitute the key system used to obtain the equilibrium oil film temperature are as follows

$$\begin{aligned} M &= C/R \\ \text{PAVG} &= W/L/D \\ A &= 132./\text{ETA}*(1000.*M)**2*\text{PAVG}/Z/N \\ \text{MJ} &= \text{MU}*L*R**2*N/M*KFJ \end{aligned}$$

$$\text{HPF} = \text{MJ}*2.*\text{PI}*N/12./33000.$$

$$\text{TWA} = \text{HPF}*2545./K/\text{AH}$$

$$\text{TW} = \text{TWA} + \text{TA}$$

$$\text{TF} = 1.15*\text{TWA} + 10. + \text{TW}$$

The last equation is a slope-intercept form of straight line equation arranged to fit the data in Fuller³, figure 150, p. 245 for the design conditions specified in the problem statement.

APPENDIX D

Program input for hydrodynamic oil journal bearing

```

INCON IN=1.0, TFO=200., ECC=0.5, TOO=0.0
INITIAL
CONSTANT L=3., D=3., K=7., AH=1.0, TA=75., PI=3.1416
CONSTANT N=1200., C=0.0015, R=1.5, W=1800.
      M=C/R
      PAVG=W/L/D
DYNAMIC
NOSORT
      INN=IN
      TF=TFO
AFGEN CURVE1=100., 6., 110., 4.5, 120., 3.46, 130., 2.7, 140., 2.15, ...
      150., 1.75, 160., 1.45, 170., 1.23, 180., 1.04, 190., 0.88, 200., 0.77, ...
      210., 0.67, 220., 0.585, 230., 0.52, 240., 0.466, 250., 0.421, ...
      260., 0.379, 270., 0.345, 280., 0.313, 290., 0.288, 300., 0.27
      MU=AFGEN(CURVE1, TF)*1.E-06
      Z=MU/1.45E-07
AFGEN CURVE2 =.2., .42, .4., .455, .6., .51, .8., .605, .9., .72, .92, .745, ...
      .94, .79, .95, .82, .96, .843, .97, .878
      ETA=AFGEN (CURVE2, ECC)
      A=132./ETA*(1000.*M)**2*PAVG/Z/N
AFGEN CURVE3 =0., 0., 1., .14, 2., .265, 3., .39, 4., .49, 5., .58, 6., .645, ...
      6.5, .67, 7.0, .691, 7.5, .705, 8.0, .721, 8.5, .74, 9.0, .75, ...
      10., .785, 11., .8
      ECC =AFGEN (CURVE3, A)
AFGEN CURVE4 =0., .658, .1, .666, .2, .691, .3, .734, .4, .798, .5, .886, ...
      .6, 1.010, .7, 1.193, .8, 1.495, .85, 1.785, .9, 2.162, .92, 2.477, ...
      .94, 2.872, .96, 3.532, .98, 4.915
      KFJ=AFGEN (CURVE4, ECC)
TERMINAL
      ERROR=TOO-TF
      IF (ERROR.GT.+0.05) GO TO 60
      IF (ERROR.LT. -0.05) GO TO 80
      GO TO 40
60 TFO=TF-ERROR/50.
      GO TO 90
80 TFO=TF+ERROR/50.
90 CONTINUE
      TOO=TF
      MJ=MU*L*R**2*N/M*KFJ
      HPF=MJ*2.*PI*N/12./33000.
      TWA=HPF*2545./K/AH
      TW=TWA+TA
      TF=1.15*TWA+10.+TW
      IN=INN+1.0
      IF (IN.GT. 20.) GO TO 40
      CALL RERUN
40 CONTINUE
      HO=C*(1.-ECC)
TIMER PRDEL=1.0, FINTIM=2.0
PRINT ERROR, TOO, TF, Z, ETA, ECC, KFJ, IN
END
PRINT MU, ETA, A, MJ, HPF, TW, TF, ERROR
END
PRINT HO, TWA, PAVG, Z
END

```

REFERENCES

1. D. H. Offner. Computer programs in a machine design course. *Engineering Education* (Jan, 1971) 375.
2. M. F. Spotts. *Design of Machine Elements*, 4th Ed. 1971.
3. D. D. Fuller. *Theory and Practice of Lubrication for Engineers*, 1st Ed, pp 192-245, 1956.
4. IBM SYSTEM/360. Continuous system modelling program. *User's Manual GH20-0367-4*, 1972.

A NEW GENERALIZED CONCEPT OF COMPUTER AIDED PRODUCTION

by

J. FERENCZY* and J. LEHOTZKY*

SUMMARY

The concept, described in this paper, is an approach to building up a CAD and a joint CAM system for use in machine production. It is a result of practical work, and any theoretical content is only a by-product. The system described is a compromise in processing time, memory requirements, investment and other factors. Hardware as well as software problems are considered.

The authors experiences are based originally upon a narrow, but highly vertical, tool design and technology CAD-CAM program. The new system is generalized for use in machine design and production. It is not restricted to special kinds of the nearly infinitely wide field of CAD and CAM, due to the flexibility of the system. Otherwise this system is really not a program, but an interactive program construction kit for machine designers and production engineers.

INTRODUCTION

It is well known that a long series of problem oriented machine design programs have been prepared for different purposes; i.e. gear box, pipeline, ship hull, car body, airplane wing, fuselage, cam, press tool design, etc. Another approach of machine design is more general, but represents only a phase of the work flow; simply graphic oriented like GINO and many others. Programs like this are tools (sometimes very powerful ones) which are perfectly independent of any application. A third type of approach is a group of systems, containing arithmetic parts, and, at the same time, possessing graphic properties.

The author's first experiences were realized by writing a program for industrial application. The topic being a press tool design. The program has a full verticality from the basis data to NC machining, but has a limited drawing capability^{2-5,9}

The idea for a new type of general machine design program, originated from the demands of machine designers. It is well known, that a decision like this involves a big risk, and so the risk is increased by extending it from CAD to CAM. In addition we had to accept that we could not count upon a giant computer for our purposes, but the real source of difficulties was not the hardware.

Mr Welbourn's warning that *the whole concept of machine design must be revised before building up a new system*^{1,3} was kept in mind. Thus we had to

forget all about existing systems before beginning this revision. It was only after this work had been accomplished that a sketch of the system was attempted.

THE NEW CONCEPT

Analysis of the flow of machine design

As a first step, we studied what the average machine designer really does during his work. Generally he drinks tea, smokes, has a breakfast, lunch and rings up his wife. Afterwards a great percentage of his work consists of seeking for data and standards, or looking up catalogues and drawings. In other words his first step is acquiring and grouping data. This is a time-consuming part of his work which could be replaced by storage packages, file organization and retrieval programs.

Another typical part of this activity involves *calculations*, most of which are generally used—long series of routine calculations. These can be computerized in the form of series of *independent subroutines*, and so a well selected *subroutines library* and a *combined retrieval system* may prove useful to the designer.

The real *creative part* of the design cannot be automated because intuition is a quality reserved for man, but it can be helped greatly by the computer. It must be noted, that our statistics about the designer's activities correspond to the data defined and

* Hungarian Tungstam Works

published by the NEL—concerning CAD—and provided by the statistics of some English factories (an appendix of the Feilden report).

At first sight it is clear, that the whole CAD work is a mosaic of deterministic routine works, indeterministic decisions, and creative parts. The realistic solution is a system, or a language, characterized by a high interactivity. Some points of the programs can be equally well handled in interactive or deterministic ways. The interactive version tends to be shorter and simpler because the deterministic type of program must contain all possible versions, decisions and essentially longer diagnostics, than a program written in interactive style.

Considerable effort has been made to determine the occurrence list of routine works. The range depends upon the nature of the work, but the outlines of the list do not change essentially. The headings of the occurrence list are

Primary calculations

- calculation of surface/area
- calculation of length of contours
- calculation of volume
- calculation of mass
- calculation of weight
- calculation of inertia to an axe
- calculation of inertia to a point
- calculation of centre of gravity
- calculation of driving moment
- etc.

Long series of *secondary calculations* are based upon the primary list, as

- stresses
- deformations
- reduced moment of inertia for combined systems of more elements
- swinging times
- resonance
- velocity equations and their higher differentials
- centrifugal forces
- kinematic curves
- handling vectors, having different physical meanings (forces, velocities, stresses, deformations, etc.)

The primary and the secondary routines cannot be linked in advance. The system is only flexible if all of the routines can be called in and joined optionally. The designer, who operates the system, has a free hand to change the routines, and to call and recall them as he wants.

The task of the design program

The basis of the program is the routine library. Its parts are independent modules which can be linked, optionally, by the operator to a temporary chain or network, fitted to the actual task. This is made interactively by the operator, who must be an engineer. He is helped, by a retrieval program, to find all the routines he needs. The second step is that one describes the momentarily needed connections. To avoid complicated organization, the retrieval program is not linked automatically to the library; it only gives call names of routines. A simple

alphanumeric display is implemented by the graphic display. Both are independent.

The retrieval program works by a simple tree structure. After some training, the operator can use it very efficiently. Every routine name is supplemented by a short instruction or information, regarding the use and ability of the program.

The surface of the alphanumeric display is divided into two parts. The upper part gives the name of the routine, instructions, and other information. Information, pointing to the next branches of the retrieval tree, are implemented, too (figure 1). Two or more versions of each routine are stored; each fulfilling nearly the same task, but containing different procedures. For instance one could adopt a simple procedure for broad work, giving only approximate results, or other procedures, giving higher precision, but having longer running times. Some routines make

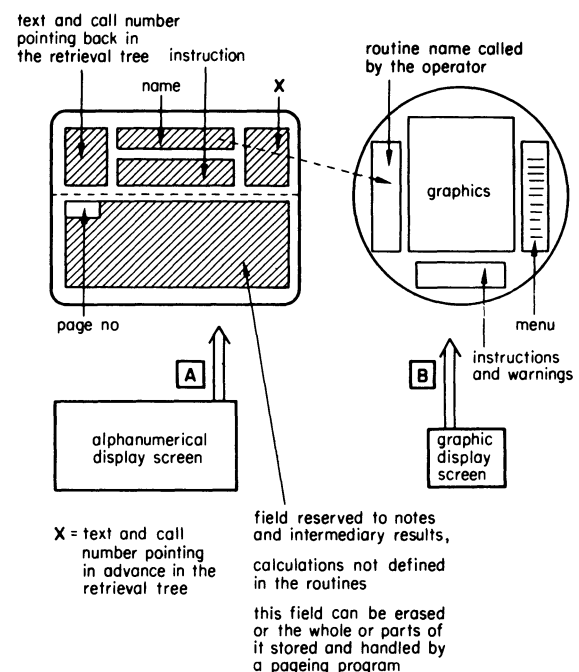


Figure 1. A and B display screens.

no calculations at all. These contain technical data, data of standards, etc. The lower part of the alphanumeric display field is free for use by the operator, for different purposes. If he wishes, he may store notes or intermediary results, and recall them for use later. He may use the little computer of the alphanumeric CRT to prepare calculations not written in the routine store, or keep data selected. If the half surface of the screen is not enough, he may use paging techniques.

So, due to the addition of a, not too expensive, alphanumeric hardware, and an independent storage system, the need for a very complicated organizing program becomes superfluous. Otherwise a great part of the operative memory would be used for organization between the instructions, retrievals and the routines.

The graphic display has the task of visualizing graphics, menus of light pen and cursor manipulations, and additional alphanumeric information. The

routines of calculations are kept in isolation from the retrieval programs, in a store connected to the graphic display computer. Many of the calculation routines use graphic figures and graphic manipulations. Numeric values used by the routines may be changed by light pen or cursor. Sometimes figures displayed and numeric values are strictly connected. Display basic software has been developed by a group of the Hungarian Academy of Sciences. Description language was a problem at the beginning. The use of APT or 2CL was a possibility, but finally a more modest and minimal interactive language was implemented.

Every routine, using graphics, is a specific problem. Some have graphic input, some graphic output; all of the variations may be described briefly in a little matrix (figure 2). Most types of program need close connections between graphics and calculations. Master programs fulfilling this task are always ready to be called in, as well as graphic manipulations. All these are only showing a broad view of the system. A bit more will be explained later by a functional diagram.

NI	GI	NO	GO
NI	∧	GI	NO
NI	--	NO	--
--	GI	NO	--
--	GI	--	GO
NI	--	--	GO
NI	∨	GI	NI
			∨
			GO

NI numeric input ∧ logic AND
 NO numeric output ∨ logic OR
 GI graphic input
 GO graphic output

Figure 2. Matrix of graphic and other variations.

To avoid the need for giant hardware, the capabilities of the computers we extended by discs and magnetic tapes. The system will be slower, but the times necessary for the operator's decisions tend to be longer than running times and gaps calling in peripheral memories and other changes of information.

There is no time to describe *hardware* in detail in this paper, only to point out some characteristic ways used, and to explain the system in broad lines. CAD and CAM are connected in this system. Normally it would be too big a task for relatively small hardware, but this is not the case because the method of refrigerating all the program parts is used, while only some small fragments are living in the processor. Having gained results from CAD programs, all previous data used in CAD are erased, with the exception of those useful for preparing machining instructions. Afterwards all the processing power is used to make punched tapes for CAM. Later we shall have, as a further help, a connection to a big computer in time-sharing.

Architecture of the program

It has been explained that unlimited computer capacity was not available. Thus, not only the flow of

design, but also the possible economic methods utilizing minimal computer capacity to gain maximal output had to be revised. It is not stated that this high level has realized, only that an approach was attempted with more or less success. However the idea is that the situation shows some analogy to a trivial example. Some of the well known programs use a Cadillac when a handtruck would have been more suitable.

Figure 3 shows a general distribution of the tasks of the alphanumeric program. The storage and display hardware is independent of the program. R i, . . . , R iv contain only names, addresses, pointers and instructions of the routines. The programme fragments, represented by R i, . . . , R iv, are stored in discs and tapes, linked to a display unit processor.

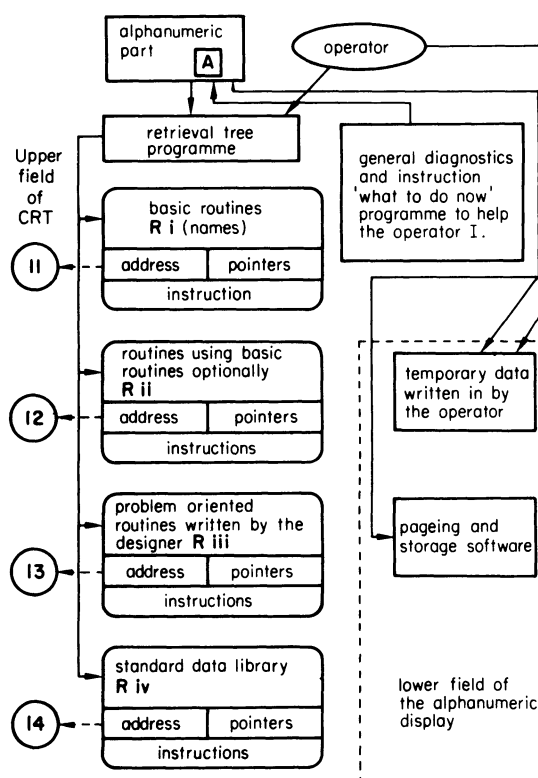


Figure 3. Flow diagram for alphanumeric part A.

The connection is only off line between the alphanumeric (A) and graphic (B) parts, and is drawn in the diagram by a dotted line.

Figure 4 shows the diagram of program B containing the routine programs and the graphic ones. The retrieval program is in A. All of the programs, R i, . . . , R iv, stored here have their diagnostics. General diagnostics are only to show eventual mistakes of call, start and editing. Dotted lines represent off line connections between A and B. Connections to graphics are partly off line, and are signed by double dotted lines pointing to graphics.

Other parts of program B are shown in figure 5. The lines drawn here are only to show functional connections. The middle point of the system is the operator. His role is pointed out in every diagram. Off line connections are signed by numbers, between A

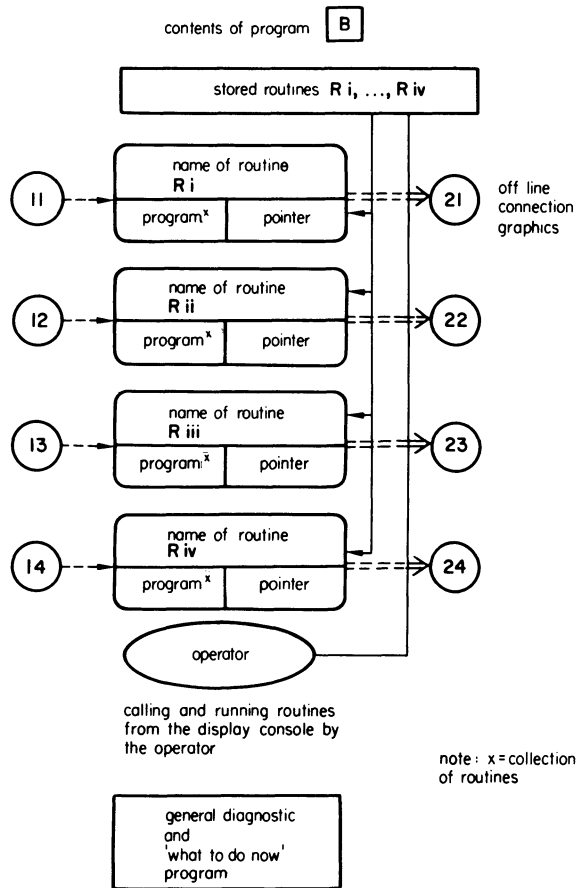


Figure 4. Flow diagram for the part B.

and B parts of programs. All of the program parts have a so-called 'what to do and error' program. This is a diagnostic and an instruction program. Every time a little fragment is finished, it supervises eventual errors and determines, for the operator, what he can do, or what he must do, or what the possibilities are for him to do. This sometimes gives a list, as in figure 6. These are naturally only a few of the many possibilities.

The high rate of interactivity, and the intensity of the contribution of the operator in cutting times, makes the programing shorter, but the price paid is that the operator must be an able engineer who can manage the system with a thorough knowledge of design and technology, and who is able to make decisions quickly and correctly. Big mistakes are filtered out by diagnostics, but it is impossible to think of all possible mistakes in advance.

Switching from CAD to CAM, all of the material of CAD will be erased from operative memory and sent back to the tapes. Only data necessary to CAM will be kept living in the processor and made ready to process punched tapes for the machining.

Example

The first impression is that the functions of the system described are so varied that a giant computer seems inadequate to fulfil this complicated task. The approach employed involved using time distribution (the double system and interactivity).

Now there is an example of how these three functions can solve a design task together. Interactivity is a bit more extended here, as usually, not

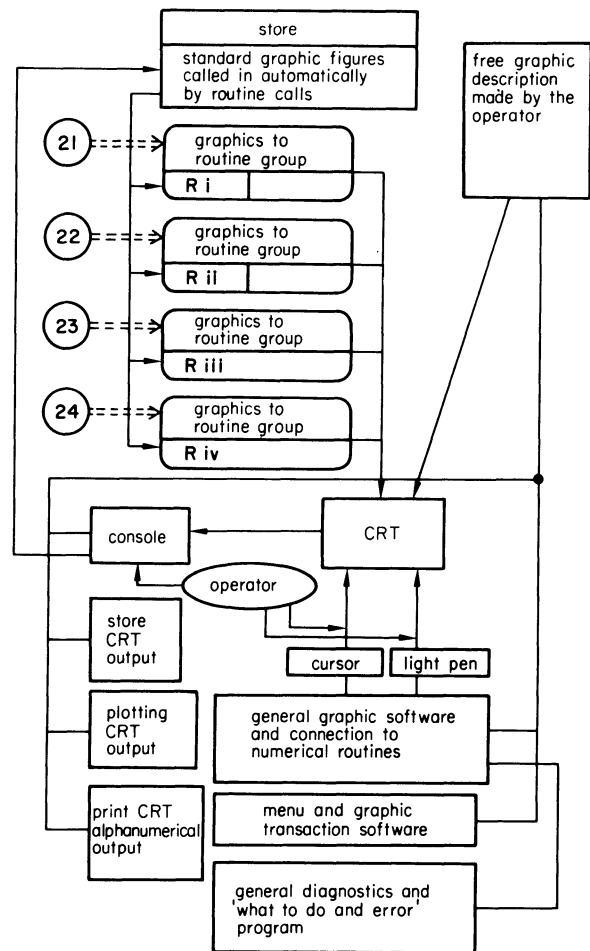


Figure 5. Flow diagram for the graphic part.

only for making decisions, but to replace sophisticated organization programs. As can be seen from figures 3 to 5, some connections between different parts of the program are off line. On these points organization is eased by the operator stepping into the system (see points 11 to 14, and 21 to 24 on figures 3, 4 and 5. Figure 7 is a simplified time-table, showing activities of the program and the role of the operator.

FURTHER DEVELOPMENT

Programs like this are never finished because the methods of design and machining have no deadlines, as development never stops. This is a big problem with a big batch program. By using a program like this, routines can always be changed, replaced, and out of dates erased and new ones added.

- will you call graphics? give sign CG
- will you call back another routine? (routine recall) RR
- will you have light pen, cursor menu? LM
- will you have plotted the contents of CRT screen? CM
- will you fix text on printer? PL
- PR

Figure 6. List of 'what to do, and error' program.

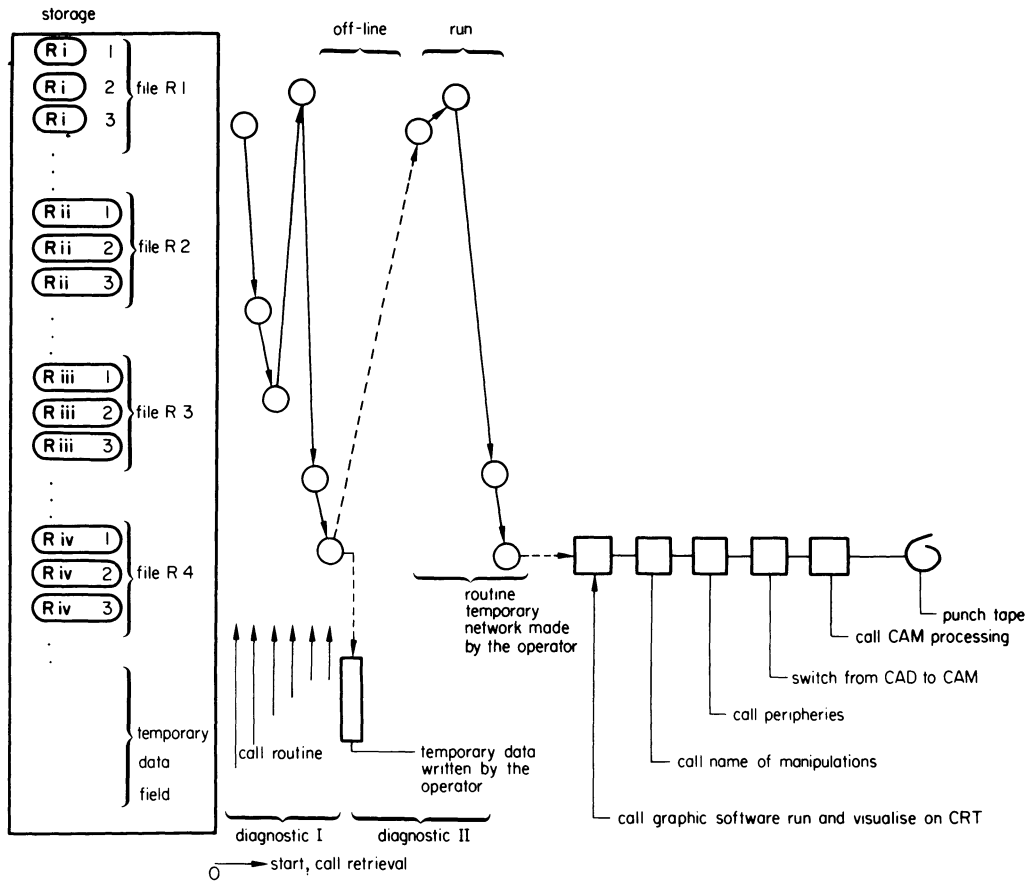


Figure 7. Simplified time-table.

Maintenance is a necessary, and continuous task, but this does not mean the rewriting of the program. Diagnostics, basic system, storage, file organization, retrieval system and others are not problem-oriented; consequently they do not change, except when refreshing of the system is necessary.

CONCLUSIONS

In this paper, the practical work has been emphasized, and general points have been mentioned only implicitly. Based on the experimental work, we may try to define characteristic and critical points of the engineers work. Generally, the methods used can be classified into three main groups

- (1) exact (predefined) methods
 - (2) intuitive phases
 - (3) multiple phases
- } (cannot be defined in advance)

The last two are characteristically interactive. Multiple phases are of a mixed nature, containing intuitive and deterministic building bricks. After some years of experimenting and obtaining results, it is believed that the new way to solve the computerizing of machine engineering will not be to write giant programs, but will be to build up flexible systems, in which routines can be freely and flexibly connected interactively in an optional network. The interconnections will not be constant, but changeable

from time-to-time, as the actual temporary task dictates. This means a *general program building method*, or we can call it a *language building for a machine production language*.

The results obtained are not yet considered as a ripe, finished work; just the beginning of something which is of some use for a designer. From the point of view of general development, it is only a rough sketch, delivered in the hope of stimulating constructive criticism.

REFERENCES

1. J. Ferenczy. Dialogue type computer program for the design of stamping dies and production technology. *Automation and Measuring* (1969) 1, 44-6.
2. J. Ferenczy. Interactive design system of press tool. *Conference paper of Comppontrol International Conference*, 1970.
3. J. Ferenczy. A CAD program for blanking die design. *Conference paper of the Uxbridge University* 1970.
4. J. Ferenczy and M. Kovács. Computer aided design. *Publication of the Automation Research Institute of the Hungarian Academy of Sciences*, 1969.
5. J. Ferenczy. An examination of alternative methods of strip layout design for press tools using a computer aided design centre. *CDAC Cambridge* (1972) 218.
6. Interactive design of press tools and technology. *Dissertation* (1969) p. 190.

7. J. Hatvany. New methods of NC adaptation *Kolloquium of Automation* (1966) 3.
8. Hermann. Graphical system '71. *CAD Seminar*, 1973.
9. Lacoste-Rothenberg. Communication in computer aided design and computer aided manufacturing. *Prolamat 1973*, 1973.
10. J. Lehotzky. Beweggründe and Möglichkeiten zum einsatz CAD und CAM. *Wissenschaftliche Tagung* 1973.
11. A. Leyer. (permanent publication). Maschinen-
konstruktion. *Lehrstuhl für Maschinenelemente, Techn. Hochschule Stuttgart*.
12. D. T. Ross and J. E. Rodriguez. Theoretical foundations for the CAD. *Proc. Spring Joint Comp. Conf.* p. 305-22, 1963.
13. D. R. Welbourn. The use of computers with graphical input/output. *The Chartered Mechanical Engineer* (1966) 11, 487-89.
14. H. Zemanek. Philosophie und Programmierung. *Elektrotechnik und Maschinenbau* (1967) 10, 413-21.

COMPUTER-AIDED DESIGN OF AXISYMMETRIC HOT FORGING DIES

by

S. K. BISWAS* and W. A. KNIGHT*

SUMMARY

A system of programmes, for the design of axisymmetric hot forgings and dies, is described. The input to the system is a geometric description of the required as-machined part, from which the appropriate forging is designed on the basis of stored design rules and material files. Subsequently, the billet size, forging load, flash geometry and forging tolerances to BS 4114, are automatically determined.

The system utilizes standard die sets, with interchangeable inserts, the profile of which is added to the forging impression, and an assembly drawing of the die set is plotted out. Other outputs are drawings of the forging, detail drawings of the variable die components, and a co-ordinate listing to facilitate manual NC programming for manufacture of the die parts. A further programme can be utilized to design appropriate preforms for components if sufficient quantities are required to make this economically justifiable.

INTRODUCTION

Three main stages must be completed before a forging is put into production: firstly, the appropriate forging must be designed from the machined part required; secondly, the method and sequence of operations, to produce the forging, must be decided upon (including estimation of the amount of material required); thirdly, the appropriate dies, including those for all preforming operations, must be designed. As a result of the highly specialized technology involved, these three stages are carried out with heavy reliance on the experience of the personnel involved, with the aid of largely empirically established design rules which may be far from formalized and interpreted individually. Coupled with the fact that forgings are produced in a large variety of shapes and sizes, this approach hinders the introduction of more rationalized design and manufacturing procedures which could lead ultimately to a greater use of computer-aided methods in the forging industry.

Experience in other areas of manufacture has shown that the introduction of systematic procedures into organizations with a diverse product range, such as forging companies, is facilitated by application of the principles of group technology¹⁻³. Classification of the product range, into families of forgings related by similar design and manufacturing requirements, forms the basis for rationalized procedures, incorporating, where appropriate, existing experience in a more formalized manner. Design rules for forgings can be specified on a class basis, and

associated with each class can be a number of alternative standard processing sequences or work plans, with variations mainly in the number and type of preforming operations used. Selection of the appropriate work plan, for a particular class of forging, would be based upon economic considerations of the quantities required. Each work plan could then have associated data, for estimating purposes, collated on the basis of current manufacturing practice, and associated die design procedures for all dies, including computer aided design systems.

Division of the forging range into suitable families can lead also to economies in die manufacturing costs, through the use of standardized dies and die components for restricted ranges of forgings¹⁻⁵. Such dies utilize interchangeable inserts and adaptors, enabling the main die assembly to be readily modified to produce any component within the associated family of forgings. This has two main advantages: firstly, the individual die costs are reduced since the design and manufacturing effort is applied only to the inserts for the main die assembly, as opposed to individual die sets for each component as is normally the case; secondly, the interchangeable inserts may enable machine changeover times to be greatly reduced^{2,3}. Both of these enable die and manufacturing costs to be reduced, and may make forging more economically viable for smaller quantities of parts.

In this paper a computer-aided system for the design of axisymmetric forgings and dies is described. It

* Department of Mechanical Engineering, University of Birmingham

would form one rationalized procedure for some classes of forging in an integrated system. Standard die sets, with appropriate inserts, are utilized, and the complete procedure, from the machined part required to the die insert design, is covered. The die sets employed are of a type used on the Petro-Forge machines⁶, but the programme could be readily adapted to the design of dies of similar construction for other machines. Some isolated computer programmes for certain stages in the design process, such as weight estimation and the determination of flash geometries⁷, have been developed previously. Where applicable, similar programmes have been incorporated in the system to be described. In addition, programmes for the design of preform shapes for axisymmetric forgings are outlined, together with the design principles adopted.

BASIC DESIGN PROCEDURES ADOPTED

The first stage in the design procedure is the design of the forging from the machined part required (figure 1(a)). This involves:

- (1) Selection of the die parting lines,
- (2) The allocation of machining allowances to various surfaces,
- (3) The selection of draft angles,
- (4) The addition of edge and fillet radii,
- (5) The allocation of tolerances.

The success of the forging operation depends to a considerable extent on the correct selection of these design features. A number of publications^{8,9} give general guide lines on the selection of forging features, but the only detailed specification is the German Standard DIN 7523¹⁰ for steel forgings, and this is used for the design rules incorporated in the computer system. The tolerancing system given in BS 4114¹¹ (and DIN 7523) is widely accepted in the forging industry, and is suitable for incorporation in a computer system.

Once the forging has been designed, this must be converted into the shape of the finished die cavity and any required preform shapes designed. The cavity of the finishing die (figure 1(b)) consists of the final forging shape, with a suitable flash land and gutter around the parting line. The purpose of the flash land is to restrict the lateral flow of material to facilitate die filling. Appropriate dimensions of the flash land are dependent on the shape of the forging and the material, and chosen so as not to result in excessive loading of the die. A number of semi-empirical methods for determining the flash geometry are available^{7,12}. The most widely applicable would seem to be that due to Teterin et al.¹³ which has been incorporated into a computer programme previously by Altan et al⁷. The flash gutter must be sufficiently wide to accommodate the excess material produced in the form of flash.

The geometry of the flash gap does not determine the amount of excess material which must be added to the forging weight in order to produce the forging. Figure 2 shows a characteristic load-deformation curve for a closed die forging. The material starts to flow into the flash after deformation D_1 . Although it

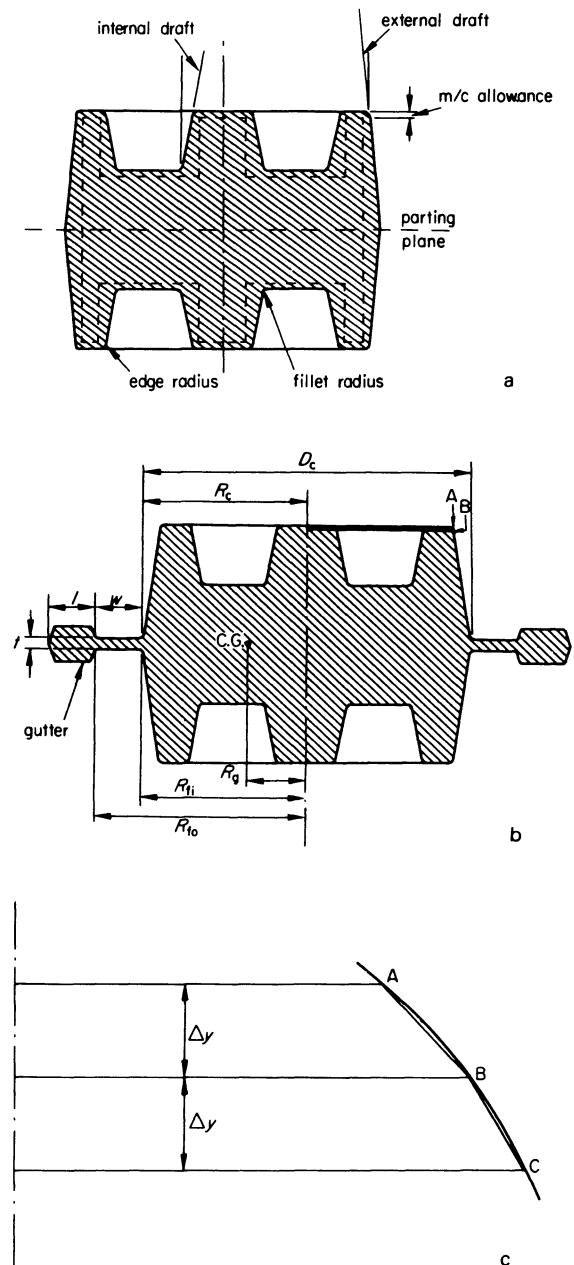


Figure 1. (a) Typical axisymmetric forging, (b) Forging cavity, including flash gutter, (c) Enlarged view of the radius segment.

is necessary for an amount of material to flow out through the flash between deformations D_1 and D_2 , only a certain proportion of it needs to be within the flash land at a given moment to provide the instantaneous restriction to outward lateral flow. Thus a volume of material, excess of that required for the component, and that in the final flash land, must be forged for adequate die filling.

Estimation of this excess material is important because material costs make up approximately 50 per cent of the total forging costs on average^{14,15}, and, of this, some 25 per cent is scrap in the form of flash (i.e. 14 per cent of the total cost). In the long term, the estimation of excess material requirements is best approached using statistical data collated as a class basis. However, in the programme to be described,

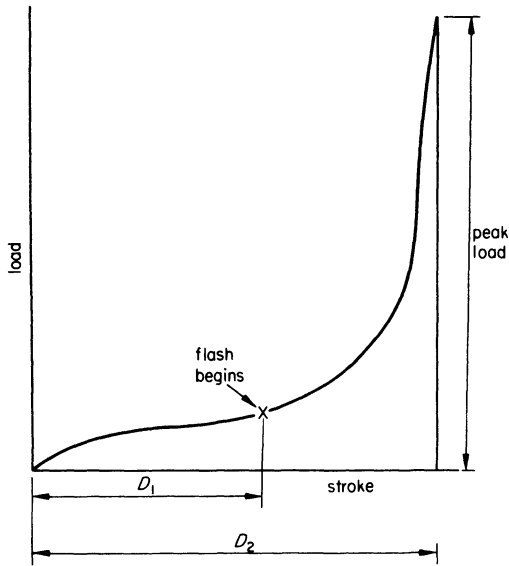


Figure 2. Forging load/deformation curve.

the method, suggested in the NAFDS recommendations¹⁶ on the costing of forgings, is used. Based on American practice, an equivalent width and thickness of flash is specified for various forging weight ranges. Thus knowing the length of the die line, the weight of flash can be determined. This, added to the forging weight, together with an allowance for scale losses, gives the total weight of the billet required. Programmes for the determination of the weight of axisymmetric forgings have been developed previously⁷, and a similar approach is adopted in the current design programme.

Once the shape of the finished die cavity is determined, this must be added to suitable die block or insert. Considerable savings, in manufacturing and die design costs, can be realized through rationalization and standardization of die blocks and die components¹⁻³. The current programme is applicable to the design of die sets for the Petro-Forge machines which utilize standard die inserts and die blocks. Figure 3 shows an assembly drawing of one of the standard die sets. The die block, clamping rings and inserts are all of standard dimensions, and

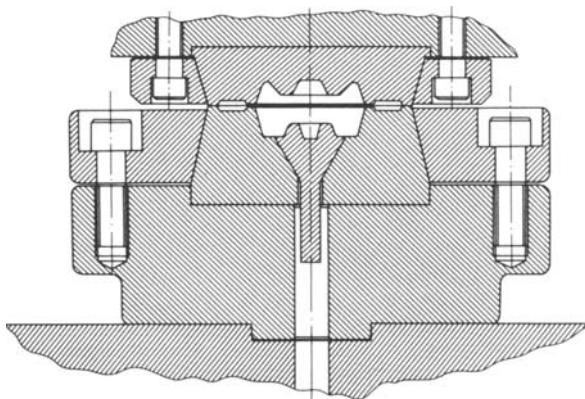


Figure 3. Standard die set with interchangeable inserts.

the only variables are the profiles of the forging cavities and the ejector pin. The particular die set shown is for components of up to 85 mm diameter and 50 mm total height. Further die sets could be available for taller components (85 mm diameter and 85 mm total height), or for smaller parts (less than 50 mm diameter).

The use of standard die sets is advantageous to the development of computer aided design programmes because the profiles of the standard inserts can be stored, and the design drawings obtained simply by adding the forging cavity shape to these.

COMPUTER AIDED DESIGN PROGRAMME FOR AXISYMMETRIC FORGING DIES

An outline block diagram for the system of programmes is shown in figure 4. The design procedure starts with the input of a geometric description of the

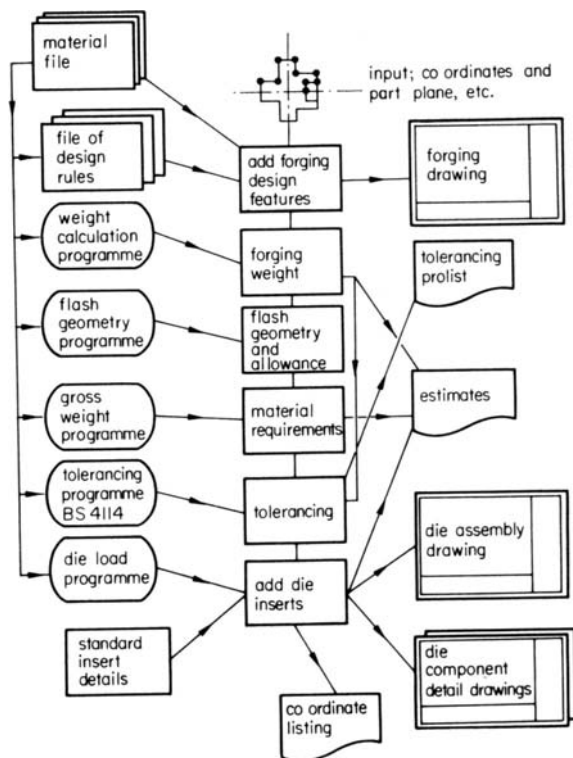


Figure 4. Outline flow diagram of computer programme.

required machined part, in the form of the co-ordinates of the change points of the cross section profile and the selected parting plane of the die. Further inputs are the material of the forging, and the forging conditions (speed and temperature), which enable data in a material file to be accessed at various stages in the execution of the programme.

The machined part profile is converted into an appropriate forging design by the addition of machining allowances, draft angles, and edge and fillet radii, using the design rules laid out in DIN 7523¹⁰. A drawing of the forging, superimposed on the original machine component, is plotted out automatically. Logic is included to omit small steps

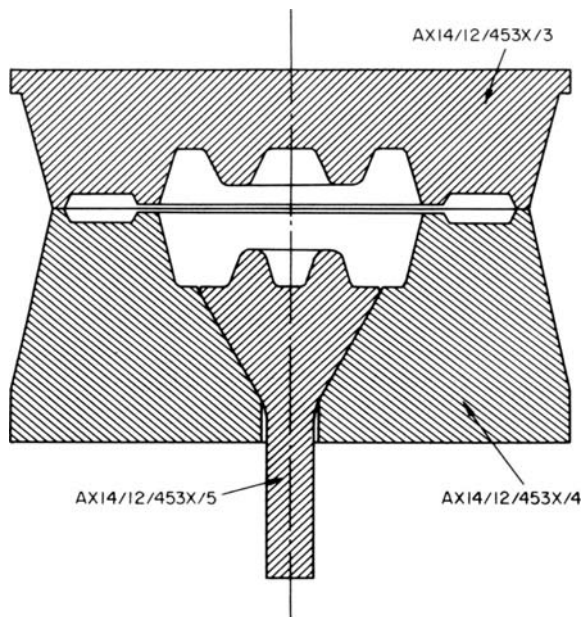


Figure 5. Typical die assembly drawing prepared on the incremental plotter.

and recesses in the machined component when determining the forging profile.

From this stage, the programmes employed calculate the weight of the forging, the flash geometry, the billet size required and the peak die load. The calculated weight and material of the forging serve as input to a programme for allocating tolerances to BS 4114¹¹, with the main values added to the forging drawing, or printed out, in tabular form, with other values.

The profiles of the appropriate inserts, for the standard die set, are stored in the computer, and an assembly drawing of the variable parts of the die set (top and bottom inserts and ejector pin) are plotted out by adding the profile of the forging to the insert profiles (figure 5). The programme selects the diameter of the ejector pin to coincide with a suitable point in the bottom profile of the forging. If required, the pin diameter can be specified in the input information, and the logic for its selection overridden. A detail drawing of each die insert, and the ejector pin, are plotted out subsequently (figure 6). The dimensioning is kept to a few easily checked items because of the added complexity of a fully dimensioned set of drawings. The drawings are augmented by a further output in the form of a co-ordinate listing of the points around each forging cavity. The co-ordinate data is in the form of X/Y

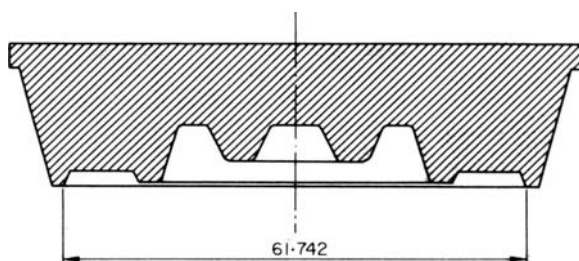


Figure 6. Detail drawing of top die insert.

co-ordinates for points, and the value and centre for any radii in the cavity cross section. This is the form required for the manual programming of NC machines to produce the die components, and would enable the manual programming to be quickly carried out. The same co-ordinate data can be used for inspection purposes during the die manufacturing process.

The system of programmes is written in Fortran IV, and runs on an IBM 1130 computer with 8k main store and a single disc store of half a million words capacity. The computer is linked to an 11 inch incremental drum plotter on which all the detail drawings are produced. The limitations of the working store means that the programme far exceeds the capacity, and is therefore written as a series of main sub programmes which are stored on the disc and suitably linked, to enable control to be transferred from one to the other. This also enables space to be saved by using only one peripheral device in each main programme, thus reducing the system software required to service peripherals in each sub-programme. Each main programme is built up of a series of sub-routines (about 40 in all) for specific tasks. These are also stored on the disc and called into the working store as required.

The programme takes some 30 minutes to run, but this must be weighed against the fact that for one set of input data, taking perhaps half an hour to prepare,

CARD INPUT RECORD

```
NEW DISK RUN
FORGING DESIGN TEST
AXI4/12/453X
20/12/73
```

```
START
POINT      0.00   55.00
MACHINE 1
POINT      5.00   55.00
POINT     10.00   45.00
POINT     20.00   45.00
POINT     25.00   55.00
POINT     30.00   55.00
POINT     35.00   40.00
POINT     30.00   20.00
POINT     18.00   20.00
POINT     15.00   30.00
POINT      6.00   30.00
POINT      3.00   20.00
POINT      0.00   20.00
PARTLINE   40.00
EN1B      13.00   50.00 1200.00
END
```

DATA RECORD

```
MATERIAL          EN1B
PARTING LINE AT   40.00 MM
FORGING SPEED     13.00 M/S
BILLET DIAMETER   50.00 MM
BILLET TEMPERATURE 1200.00 C
NORMAL TOLERANCES
```

Figure 7. Input data for programme.

a forging design, with tolerances and material estimation, drawings of the die components and suitable co-ordinate listings for NC programming or inspection are all obtained. By normal methods these would be carried out by a number of different personnel, and take perhaps several days to prepare.

PROGRAMME DETAILS

Input data

The input to the system is in geometric description of the machine part required (figure 7). The data is fed in on punched cards, but this could, at a later stage, be done from a drawing digitizing table. The component profile is defined in terms of points and arcs in a clockwise direction from the intersection of the centre line and the upper surface. This is achieved using a simple input language and an origin for co-ordinates which can be anywhere along the central axis of the component. The first main sub-programme in the system interprets the input language into the appropriate co-ordinate data for processing. In addition, it must be specified whether the surface is to be machined or unmachined. If it is to be machined, a machining allowance will be added at a subsequent stage. This is effected by the instruction *MACHINE*. Once set, this instruction continues to be effective for subsequent parts until cancelled by the alternative *UNMACH* instruction.

The final card which must be included contains details of the material and forging conditions. Two other input instructions are optional. Firstly the tolerancing grade (normal or close) can be specified. If this instruction is omitted, normal tolerances will be assumed. Secondly the ejector pin diameter can be specified, and the programme for selecting this quantity overridden.

Material file

Various sub-routines draw upon data, in a material file, which is accessed by the material name in the input data. The material file contains details of the difficulty factor of the material used for tolerance estimation, and coefficients for empirical material property functions which are used to calculate the peak forging load for the appropriate forging temperature and speed. Currently, the file contains data for a large proportion of the EN range of steels, but other materials could be readily included, when required; the current range being sufficient for demonstration purposes.

Forging weight

The weight of the forging is determined by multiplying the volume by the density of the material. The volume of an axisymmetric body is given by

$$V_c = 2\pi R_g(F/2) \quad (1)$$

The half (axial) cross sectional area ($F/2$) is calculated here by considering the half section to be enveloped by a polygon of n sides, the n th side being the component axis. The area of such a polygon is

$$F/2 = A_\phi = \frac{1}{2} \sum_1^n (x_{i+1}y_i - x_iy_{i+1}) \quad (2)$$

The centre of gravity of such a polygon, with respect to the component axis (Y axis) is

$$R_g = \frac{1}{4} \sum_1^n (x_{i+1}^2y_i - x_i^2y_{i+1}) \quad (3)$$

As seen in figure 1(b), the corners of a forging component are rounded. To facilitate the use of Equations (2) and (3), axial cross sectional profiles of the rounded corners are divided into very small arcs. An expanded view of one such arc is shown in Figure 1(c). With $\Delta y < 0.005$ in, arc AB can be assumed to be line AB. In this manner all the arcs can be transformed to make the sides of the considered polygon. In the programme, the radii are divided into 10° arcs which are required also for draughting purposes.

Flash land design

As the complexity and size of the die cavity are the main variables which determine the material flow for given forging conditions, they must be taken into consideration in designing the flash geometry. A comprehensive approach to the quantitative assessment of the complexity of the die cavity for forging, is that due to Teterin et al.¹³, and is expressed in terms of a *complexity* or shape difficulty factor, α . Referring to figure 1(b), the complexity factor for an axisymmetric forging, without any projection, is given by¹³

$$\alpha = \frac{P^2 H_c D_c}{F4(H_c + D_c)^2} \quad (4)$$

where P is the perimeter of the axial cross-section surface, and F is the surface area of the axial cross-section surface, obtained from Equation (2).

To accommodate projections (e.g. shafts, rims etc.) the authors¹³ proposed a lateral shape factor β which is used to express the degree of forging intricacy in lateral direction. Thus referring to figure 1(b)

$$\beta = 2 \frac{R_g}{R_c} \quad (5)$$

A shape difficulty factor S incorporating both α and β is defined as

$$S = \alpha\beta \quad (6)$$

No wide ranging method, whereby the optimum flash geometry can be analytically determined, is known at the present time. Thus, in the present paper, the following empirical expressions, derived from extensive experimental investigations¹³, are used to design the flash land.

(a) The flash thickness t is given by

$$t = 0.017 D_c + \frac{1}{(D_c + 5)^{1/2}}$$

where t and D_c are in mm

(b) The flash land width w is given by

$$\frac{w}{t} = 0.02 + 0.0038 s \left(\frac{D_0}{t} \right) + \frac{4.93}{(Q/2.2)^{0.2}}$$

where D_0 is the diameter of the material rough stock, w and t are expressed in mm, and Q is the forging weight in lbs. The complexity factor S , is obtained from Equation (6).

Excess volume and gross weight

The gross weight of the forging is obtained by adding a proportion of excess material, which is lost in the form of flash and scale, to the actual forging weight. The principle of establishing an average flash width and thickness for various weight ranges recommended

TABLE 1 Approximate flash thickness and width for hot trimmed forgings

Net weight (lb)	Thickness (in)	Width (in)	Weight of flash per inch (lb)
Up to 1	$\frac{1}{8}$	$\frac{3}{4}$	0.0266
1 to 5	$\frac{1}{8}$	1	0.0354
5 to 10	$\frac{5}{32}$	$1\frac{1}{4}$	0.0553
10 to 15	$\frac{3}{16}$	$1\frac{3}{8}$	0.0730
15 to 25	$\frac{7}{32}$	$1\frac{1}{2}$	0.0941
25 to 50	$\frac{1}{4}$	$1\frac{3}{4}$	0.1250
50 to 100	$\frac{5}{16}$	2	0.1790
100 to 200	$\frac{3}{8}$	$2\frac{1}{2}$	0.2670

by NAFDS¹⁶ is adopted (Table 1). Once this is assumed, the weight of the flash annulus for axisymmetric forgings is readily determined. A further contribution for scale loss is necessary, and 7.5 per cent of the net forging weight is recommended¹⁵ for forging up to 10 lbs in weight.

From the gross weight, the dimensions of the billet required to produce the forging is obtained. The original diameter of the billet D_0 is a designer selected input to the programme, and should be specified as the nearest standard bar size for ease of location in the die. With the diameter specified, the billet length is readily determined, and gives the appropriate setting for a cropping machine or the like to produce the billets.

Forging tolerances

The tolerances on the forging are allocated on the basis of BS 4114 for steel drop and upset forging¹¹. Two grades of tolerances (normal and close) are included. Tables of tolerances for each category are available, and include those for dimensions, mismatch, radii etc. The appropriate figures are determined from the following

- Weight of the forging
- Shape of the die line
- Type of steel used
- A shape complexity factor of the forging

For the forgings covered by this programme, the die parting line is straight. The steels used are allocated a difficulty factor (1 or 2) based upon their chemical composition, the appropriate factor being stored in the material file.

The shape complexity factor differs from that specified by Teterin et al.¹³, and is calculated from the ratio of the forging weight to the enveloping cylinder weight, with four ranges (< 0.16 ; 0.16 to 0.32 ; 0.32 to 0.63 ; and 0.63 to 1.0) leading to separate tolerances in the standard tables. The lower the complexity factor, the more difficult the component is to forge, and consequently wider tolerances are allowed.

In the programme, the difficulty factor is obtained from the material file, and the complexity factor from a separate sub-routine. These factors, in combination with the forging weight, lead to the appropriate values in a data file which contains the relevant items from the standard. The basic tolerancing information is printed on the forging drawing, and more detailed listings are printed out to accompany the drawing.

Determination of die load

For the purposes of die design and machine selection it is essential to estimate the peak load that the die is expected to experience. The peak load, as seen in figure 2, is experienced at the end of forging. Usually at this stage the cavity is completely filled, and the excess material extrudes into the flash by forming a shear zone^{17,18} in the cavity. As an approximation this shear zone can be treated¹⁹ as a disc of thickness equal to that of the flash.

An application of the slab method, which assumes the fixed shear stress ($\tau = f\bar{\sigma}$) theory¹⁸, is used here to estimate the load experienced by the assumed disc shaped shear zone. The equation to determine peak load is given by Equation (A2) in Appendix 1.

The flow stress $\bar{\sigma}$ is determined by using Equation (A1) in Appendix 1. This Equation has been used^{20,21} to predict the flow stress of metal, hot worked above $T_m/2$, where T_m is the melting point of metal.

Co-ordinate listing output

For axisymmetric forgings, the die inserts and ejector pin will be turned on a lathe, and the availability of numerical data suggests a numerically controlled machine. Although it may be preferable ultimately to

COMPONENT HALF-OUTLINE CO ORDINATES

ORIGIN AT INTERSECTION OF CENTRE LINE AND PARTING LINE

X CO-ORDINATES PARALLEL TO PARTING LINE, POSITIVE RIGHT.

Y CO-ORDINATES PARALLEL TO CENTRE LINE, POSITIVE UPWARDS.
SLOPE IS FOR APPROACHING LINE (START POINT FOR ARCS).

TOP DIE CAVITY

START POINT		END POINT		CENTRE		RADIUS	SLOPE
X	Y	X	Y	X	Y	MM	DEG
0.00	16.49						*****
5.03	16.49	6.82	15.39	5.03	14.49	1.99	89.99
10.72	7.60	12.51	6.49	12.51	8.49	2.00	26.56
16.25	6.49	19.82	8.71	16.25	10.49	3.99	89.99
23.17	15.39	24.96	16.49	24.96	14.49	2.00	-26.56
29.75	16.49	31.66	15.10	29.75	14.50	1.99	89.99
36.15	1.07						17.75
36.34	0.98						63.85
41.60	0.98						90.00
43.44	4.16						-29.99
59.90	4.16						90.00
61.74	0.00						23.76

Figure 8. Output listing of co-ordinate data of top die cavity.

produce the NC tapes directly by interfacing the current design programme onto an existing NC processor, it was felt at this stage that the most useful output could be a co-ordinate listing for parts around the profile of the die and ejector pin, to aid manual programming. This same data could serve for inspection purposes subsequent manufacture, in the absence of detailed dimensioning of the component drawings.

The output listing (figure 8) gives the datum point, and then the co-ordinates of the half outline of the die profile. For a radius, the start and end points are given, together with the centre and radius of the arc; this being the format used in the manual programming of a suitable NC lathe.

PREFORM DESIGN

The bulk of the work, to determine the preform shapes for complex forging components, has been carried out in Germany and the Soviet Union as reported by Spies²², Chamouard²³ in France, and Akgerman et al.²⁴ in the USA, dealt with some practical problems of preform design for certain forging cross sections.

The main considerations which dictate the designer's decision regarding whether a preform should be used in forging a part are:

- (1) the expected life of the final die impression for a given batch size. When using one or more preforms the excess cost due to the manufacturing of a preform die, increased overhead and work handling etc., is balanced, for a certain batch size, by the savings due to the extended die life of the final form and the reduced amount of excess material required. A functional relationship of the above factors, which yield the optimum number of preforms for a given batch size, is thus needed for each forging or class of forgings. The relationship can only be derived from past experience
- (2) if the power of a forming machine is insufficient to form a final shape from the billet, the whole operation needs to be performed in more than one stage
- (3) even with sufficient machine capacity, for a fairly complex forging, the metal flow, in the absence of a preform, might be such as to leave some corners, projections, ribs etc., of the die cavity, unfilled. In such cases a carefully designed preform which leads metal into the more inaccessible parts of the cavity can be essential.

Spies²³, as a result of experimental studies, suggests the following rules for designing preforms.

- (a) The cross section of the preform must be as great as that of the final form. There should be minimum flow of metal into the flash at the final stage as it causes considerable wear. Therefore during the formation of cross section, all, if possible, and at least the greater part, of this excess metal should be displaced into the flash.

- (b) The cross-section of the preform and the flash should be higher in the direction parallel to the movement of the tool, and narrower transversely to it. This arises out of the observation²² that, from the point of view of wear, it is more desirable to have an upsetting type, and not an extrusion type, of flow at the final stage of forging.

- (c) The radii of all concave curves in the preform should be greater than those of the final form.

The present approach to computerized preform design is based on the criteria set out above. As the axisymmetric components of H cross-sections need¹¹ careful preform designs to eliminate cracks and folds in the final form, the present method is illustrated by its application on one such component shown in figure 9.

Referring to figure 9(a), in order to facilitate the preform design of the component, it is necessary to replace the step CX–YZ by a continuous line CB such that the final operation is one of upsetting and not extrusion. The position and length of line CB is determined by using the condition that volumes f_x and f_y are equal.

The transformed component shown in figure 9(b) acts as the basis for preform design. The number of preforms required is determined by the ratio²² $(h - b)/a$. When $(h - b) < 3a$, the curve C'B''B' is of the exponential form

$$x = \frac{md}{4} \ln \frac{y}{h}$$

where m is a constant.

Referring to figure 9(b), $d/2$ is given by²³

$$d/2 = \frac{R}{1 + |\frac{1}{2} \ln b/h|}$$

The circular arc DC' is such as to be tangent to the curve C'B''B' at C'. The radius R is obtained by using the condition that the volume f_2 is equal to $(f_1 + f_3)$. The blend-in radius of the preform at the flash is influenced by the adjacent cavity depth, and it is suggested²⁴ that it takes the form

$$r_{ff} = r_f + C$$

where C is given by Table 2²⁴.

TABLE 2

Depth of cavity (in)	Value of C (in)
< 0.4	0.08
0.4–1	0.12
1–2	0.16
> 2	0.2

The value of m which determines the slope of the curve C'B''B' is conditional on the value of r_p . A certain critical value (say 0.2 in) of r_p , although satisfying other conditions, is envisaged to be restrictive to metal flow. Thus when $r_p < 0.2$ in, m is

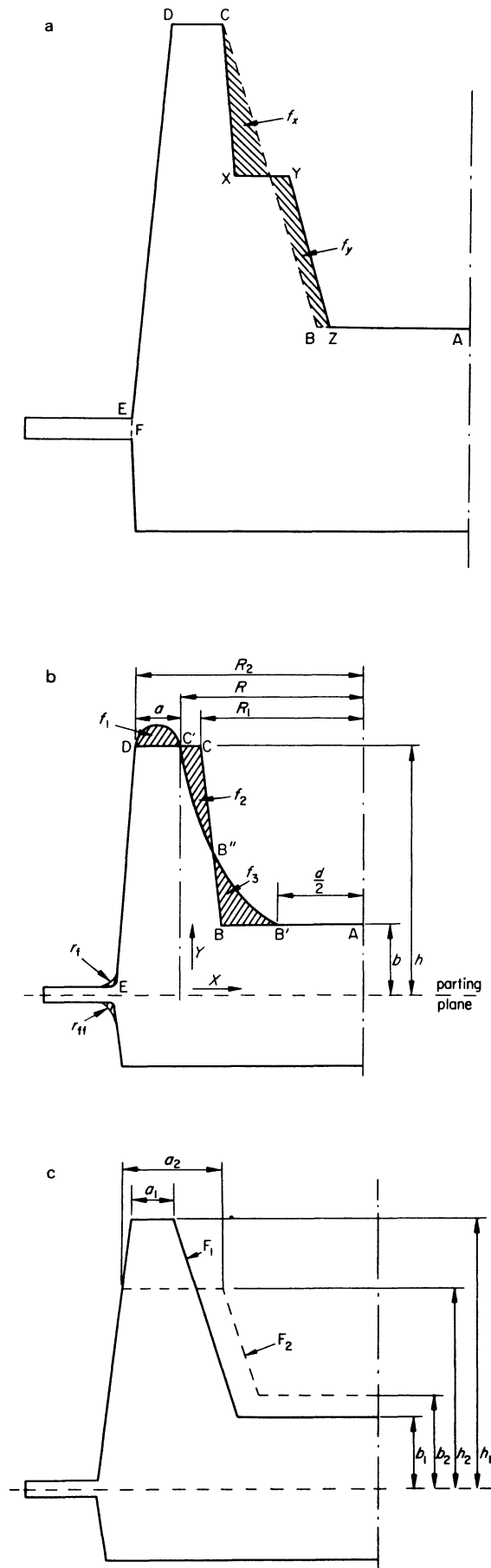


Figure 9. (a) Transformation of a stepped component to suit preform design, (b) Preform design, (c) Basis for more than one preform design.

reduced. This gives a flatter curve $CB''B'$ and greater r_p . When $(h - b) > 3a$, although the procedure for designing the final preform is as set out above, that for designing the preceding preforms is an adaptation of the method suggested by Akgerman et al²⁴.

Referring to figure 9(c), the cross sectional area of the preform base F_2 is the same as that of F_1 , the base from which the final preform is designed. The conditions $b_2 = x_b b_1$, $a_2 = x_a a_1$, area of $F_2 =$ area of F_1 , are satisfied. The designer chooses the value of constants x_a and x_b from experience, such that $(h_2 - b_2)/a_2$ is less than a critical value (say 2). The computer, having determined the exact shape of F_2 , determines the profile in the manner set out in the previous section.

CONCLUSIONS

A system of programmes, for the computer aided design of axisymmetric hot forgings, has been developed which could form part of an integrated system for all forgings, based on the principles of group technology. The whole design process, from machined part required to specification of die profiles and die components, is carried out. The system utilizes standard dies with interchangeable inserts for restricted ranges of forgings, and detail drawings of the die components, together with suitable co-ordinate data to facilitate NC programming for their manufacture, are automatically produced.

For suitably large quantities of forgings or difficult shapes, preforming operations may be justifiable. Programmes have been written to design appropriate preform shapes based upon empirically established design rules.

ACKNOWLEDGMENT

The work, on which this paper is based, is financed by a research grant from the Science Research Council. A major part of the programming for the main design programme was carried out as a project for the Graduate School of Machine Tool Technology, under the supervision of the authors. The contribution of the following students is gratefully acknowledged: P. J. Andrews, A. Dani, R. Horrocks, K. D. McDonald, R. D. Oakes.

APPENDIX

(a) Flow Stress. The flow stress $\bar{\sigma}$ is given by

$$\dot{\epsilon} = A_1 (\sin h A_2 \bar{\sigma})^{A_3} \exp\left(\frac{-A_4}{RT}\right) \quad (A1)$$

where A_1, A_2, A_3 and A_4 are strain rate, temperature and strain insensitive material constants: $\dot{\epsilon}$ and T are the mean strain rate and temperature (in °K) of the deforming component. R is the Boltzman Constant.

(b) Peak Load. Referring to figure 1(b), and assuming dry friction conditions at the flash, the peak load P is given by

$$P = 2\pi \left[\frac{A}{3} (R_{fo}^3) + \frac{B}{2} (R_{fo}^2) \right] \quad (A2)$$

where

$$A = \frac{-2\bar{\sigma}}{\sqrt{3t}}$$

and

$$B = -AR_{fo} + \bar{\sigma}$$

REFERENCES

1. C. C. Gallagher and W. A. Knight. *Group Technology*. Butterworth, 1973.
2. W. A. Knight. Group technology in forging and stamping. *Metallurgia and Metal Forming* (1971) 38, 245.
3. W. A. Knight. Part family methods for bulk metal forming. *Proc. 2nd Int. Prod. Res. Conf.* Copenhagen, (1973).
4. B. Stejskal. *Aplikace Skupinove Technologie pri Kovani*, Bratislava, 1966.
5. P. V. Kamnev. *Gruppovoi Metod Proizvodstvo Pokovok*, Lenizdat, 1961.
6. L. T. Chan, S. A. Tobias and F. Bakhtar. Design and performance characteristics of Petro-Forge mk I and mk II machines. *Proc. I. Mech. Engrs.* (1969) 184, 1097.
7. T. Altan and H. J. Henning. Closed die forging of round shapes—flash design and material savings. *Metallurgia and Metal Forming* (1972) 39, 83.
8. J. A. Schey. Principles of forging design. *Metal Treatment* (1966) 33, 263.
9. A. M. Sabroff, F. H. Bougler and H. J. Henning. *Forging Materials and Practice*. Reinhold Book Corp (1966).
10. DIN 7523. *Gestaltung von Gesenkschmiedestücken* (1972) 1–3.
11. British Standards Institution. *Tolerances for steel drop and press forgings and steel upset forgings made on horizontal forging machines* BS 4114, (1967).
12. K. Vierregge. *Die gestaltung der Gratspaltsan schmiedegesent*. *Industrie-Anzeiger* (1970) 76, 65, 1561.
13. G. P. Teterin et al. Shape difficulty criteria for forgings (Russian). *Kuznecno-Stampovocnoe Proizvodstvo* (1966) 7, 6.
14. Ministry of Technology. *A study of the U.K. Die and Mould Manufacturing Industry*, 1970.
15. F. Bakhtar. Economics of high energy rate forging. *Engineer* (1969) 227, 302.
16. National Association of Drop Forgers and Stampers. *Recommendations for Uniform Cost Estimating*, 1953.
17. E. P. Unskov. *An Engineering Theory of Plasticity*. Butterworths, 1961.
18. T. Altan, H. J. Henning and R. J. Florentino. The use of analytical methods to predict load and stresses in closed die forging. *Final Report of 'A Study of Mechanics of Closed Die Forging'*. Battelle Memorial Institute, 1969.
19. S. Kobayashi, V. McDonald and E. G. Thomsen. Some aspects of closed die forging. *Int. J. Mech. Sci.* (1960) 1, 282.
20. S. K. Samanta. The dynamic compression of steels at elevated temperatures. *Proc. 11th MTDR Conf.* 1970.
21. S. K. Biswas and B. W. Rooks. Application of a computer simulation technique to estimate load and energy in forging. *Proc. 13th MTDR. Conf.* 1972.
22. K. Von Spies. Intermediate forms in drop forging and their production by roll forming. *West Deutscher Verlag, Cologne* (Translated by Ministry of Technology, 1969), 1959.
23. A. Chamouard. *Estampage et Forge*. Dunod (Paris), 1964.
24. N. Akgerman, J. R. Becker and T. Altan. Preform design in closed die forging. *Metallurgia and Metal Forming* (1973) 40, 135.

GRAPH THEORY APPLIED TO COMPUTER AIDED PLANT LAYOUT

by

ALLAN S. CARRIE*

SUMMARY

Many of the flow diagrams used by industrial engineers for many years constitute graphs in the mathematical sense, and yet the realization that there is a considerable volume of mathematics dealing with graph theory which might be usefully applied has come only recently. This paper explores this possibility.

The paper starts by examining how present computer programs interpret the plant layout problem. Basic graph theory concepts are introduced and the paper then shows how the plant layout problem has been formulated in graph theory terms. The paper distinguishes between plant layout block diagrams and detailed machine shop layout. A graph theory algorithm is presented, together with its results in examples of both types of layout. Conclusions are drawn about the value of the technique and refinements are suggested.

INTRODUCTION

For many years industrial engineers have used charts and diagrams to present information in a visual, graphical manner. With the advent of the computer, methods have gradually been developed for quantifying the plant layout problem. These methods have introduced new concepts which are far removed from the visual and intuitive techniques used in practice. In many ways this has introduced more systematic procedures and more objective measures of evaluation than were previously available, but, like many computer-based techniques, introduce a communication barrier between theory and practice, between the men who can obtain a solution and the men who must make it work.

Many of the traditional graphical aids constitute graphs in the mathematical sense. The realization that there exists a considerable body of mathematics dealing with graph theory which might be fruitfully applied has come only recently. It is the purpose of this paper to explore this possibility.

Production involves the movement of work between work centres, departments or machines. The plant layout problem is to arrange the relative positions of the work centres so as to minimize the cost of material handling. This cost is usually quantified by calculating the sum, for all journeys, of the product of the volume of material making a journey and the length of the journey. This is referred to as the volume-distance product, and the optimal layout will minimize it.

The basic data is a travel chart showing the volume of material moving from each machine to each other

machine. The travel chart is a square matrix and, if directions of flow can be ignored, the two halves can be added together to provide a relationship matrix, or chart, of the form of figure 4, page 5.

It is extremely difficult to handle the problem of arranging the relative positions of work centres, and their shapes and areas, on a digital computer. Some means must be found of interpreting them in order to do so. If the work centres are functional departments, then the shape of the department is often unimportant so long as the required area is provided. This is the implicit assumption in the device used by most computer layout programs. They consider the factory floor to be made up of a large number of small rectangular units of area. The problem then becomes one of assigning each unit area to a department, while ensuring that each department gets the correct number of units and that those units form a continuous block of floor space. This method renders the problem highly amenable to computer treatment, but requires much computer space and time. The programs employing the technique adopt one of two approaches.

One approach, for example CRAFT¹, starts with an initial layout and re-arranges it to obtain an improved solution. The quality of its results depends on the initial layout supplied and on the rules used in re-arranging the departmental areas. The other approach, for example CORELAP² and Whitehead³, starts with an empty floor. It assigns the department with the heaviest traffic to the appropriate number of area units around the centre of the floor, and then gradually locates the other departments to the surrounding area units. The quality of solution is a

* Department of Production Engineering, University of Strathclyde

function of the sequence in which departments are located, and this is done in descending order of the volume of their material handling with the departments already allocated.

For practical problems the number of relationships between departments is large, and the number of possible solutions is very large, so that both these heuristic approaches examine a relatively small number of possibilities.

These methods have very little control over the shape of a work centre. When work centres are machine tools, as opposed to homogeneous departments, their usefulness is seriously limited. The graph theory techniques described in this paper are attractive in that they impose no restriction on the interpretation of work centres.

GRAPH THEORY AND PLANT LAYOUT

Basic concepts of graph theory

Figure 1 shows a simple graph which will be used to illustrate the basic concepts. Terms used throughout are first shown in italics.

A *graph* consists of a set of *vertices*, for example 1, 2, 3 in figure 1, and a set of *edges*, for example a, b, c. Edges may have some value associated with them, such as length, and may be directed or undirected. This paper concerns undirected edges. A graph may be drawn visually, or it may be defined in matrix form. The *incidence matrix* of a graph contains one row for each vertex and one column for each edge. The values in each column identify the vertices connected by the edge.

A *chain* or *path* is a sub-set of edges connecting two vertices, for example edges a, b and c form a chain between vertices 1 and 4. A *circuit* is a continuous chain, for example edges a, b, f, e, d. A *tree* is a chain with branches and may have more than two ends, but no circuit, for example edges a, b, c and f. A *spanning tree* is a tree connecting all vertices of the graph, for example edges a, b, c, f, e, h. A *minimal spanning tree* is the spanning tree for which the sum of the edge lengths is minimum. A spanning tree contains one less edge than the number of vertices on the graph. A graph may be *planar* or *non-planar*. A planar graph is one which can be drawn on a plane with no edges intersecting. The graph of figure 1 is planar because although edges e and g intersect in figure 1, the graph could have been drawn with one of these edges inside circuit a, b, f, h, i, d and the

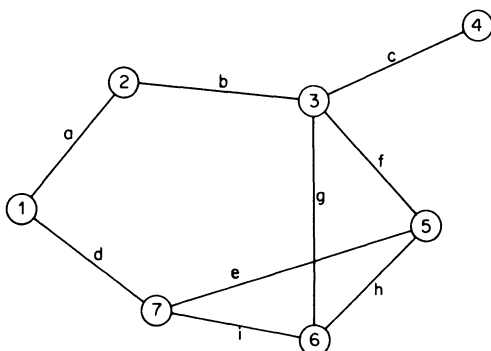


Figure 1. A simple graph.

other outside it. A spanning tree of any graph is planar.

Areas bounded by simple circuits are called *faces*, so that, with edge e redrawn outside the circuit, there are three faces bounded by abgid, fgh and eih

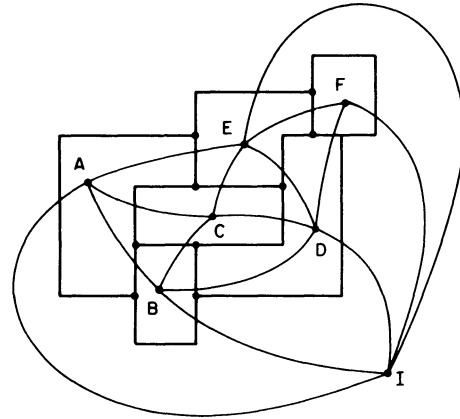


Figure 2. A plant layout block diagram is a planar graph. Its dual graph is also planar (after Seppanen and Moore⁴).

respectively. The area outside the graph, in this case the rest of the page, is termed the *infinite face*. Every graph has a *dual* graph. If a point is placed on each face of the graph, including the infinite face, then the dual graph has these points as vertices and edges connecting those vertices whose faces in the original graph are bounded by a common edge. If the original graph is planar, so is its dual. This is illustrated in figure 2.

Plant layout block diagram as a planar graph

Seppanen and Moore⁴ have observed that a plant layout block diagram may be thought of as a planar graph, and have formulated the plant layout problem in graph theory terms.

In a plant layout diagram, such as figure 2, each department is assigned an area, bounded by lines which may represent a physical boundary, such as a wall, or an organizational boundary, such as a yellow line painted on a concrete floor. Each departmental area can be considered a face and each boundary an edge of a graph whose vertices are the points of intersection of boundary lines. The dual graph of the plant layout diagram shows the lines of material flow between adjacent departments. Since the block diagram is a planar graph, so also is the graph of flows between adjacent departments. These flows are specified in the triangular relationship matrix, and it follows that if the complete relationship graph is planar, then all the relationships can be satisfied by a layout plan in which all the lines of flow are between adjacent departments. If the relationship diagram is non-planar, then no layout plan can satisfy them all.

Having deduced this result, Seppanen and Moore discuss a method of testing whether or not the relationship diagram is planar, and conclude, 'If this procedure establishes that the relationship graph is planar, the layout problem is solved. Developing the

dual graph results in a planar graph which depicts the block plan layout desired'. This is something of an over-simplification, as will be discussed later in this paper. Seppanen and Moore then go on to a discussion of a means of deleting the minimum number of relationships to provide a planar relationship graph and theoretical boundary conditions on the maximum number of relationships which can be satisfied by adjacent departments.

It is recognized that plants can operate with material movements that are longer than between adjacent departments, and hence with non-planar relationship diagrams. However, the objective of the plant layout problem as stated above is to minimize the distance-volume product, and intuitively this can best be done by placing adjacent to one another those departments between which a large volume is moved. If that planar graph which accounts for the maximum possible amount of material flow can be extracted from the relationship diagram, then the associated plant layout diagram should be optimal or very nearly so. This introduces a different criterion for optimality from that implied by Seppanen and Moore, and provides a means of extending earlier work by the author⁵, as will now be discussed.

Machine shop layout as a planar graph

The device of using unit areas is of limited value when the work centres are specific machine tools with a distinct shape and which cannot be meaningfully split up. An alternative procedure is required which does not impose any such restrictive interpretation. The author⁵ has used a principle of numerical taxonomy to identify the principal relationships. Since machines between which a large amount of material flows should be placed close together, the 'taxonomic distance' between machines can be inversely related to the material flow. A graph can then be formed whose vertices are the machines and whose edges are the paths of flow, the edge lengths being inversely related to the volume of flow. A minimal spanning tree of this graph will identify the principal paths of flow, and could be used as a schematic layout diagram. An example, from Carrie⁵, is shown in figure 8, page 7. This will be discussed in a later section of this paper.

One disadvantage of the tree as a layout diagram is that it does not give any information about the paths not included in the tree. Consequently it does not define the sequence in which the edges radiating from any vertex should be drawn (for example there are eight edges radiating from machine number 36, Dawson, just left of centre in figure 8), and therefore the diagram is not unique. It was previously stated that a minimal spanning tree of a graph is planar. It follows from this that if additional edges are added to the minimal spanning tree in descending order of material handling importance, until no more can be added without the resulting graph becoming non-planar, then the resulting graph will not only contain a very high proportion of the material handling but will also provide an explicit definition of the relative positions of the vertices and edges. It might possibly provide the unique optimum plant layout diagram.

DEVELOPMENT OF ALGORITHM

Planarity testing

The basic principle of planarity testing may be illustrated with reference to figure 1, page 2. First identify a circuit in the graph. Select abfhid. The remaining parts of the graph are referred to as bridges. In figure 1 there are three bridges, edge e, edge g and vertex 4 with edge c. Some of these bridges may be incompatible with each other, that is, if drawn within the circuit their edges would intersect. In figure 1 e and g intersect. But as was observed earlier, either of them could be drawn outside the circuit, thereby resolving the incompatibility.

Theorem 1. A graph is planar if all its bridges can be assigned either to the inside of the circuit or to the outside of the circuit such that no two incompatible bridges are on the same side, and provided these bridges themselves are planar graphs.

The condition for compatibility of bridges is simply stated:

Theorem 2. Two bridges are compatible if all vertices of attachment to the circuit of one bridge lie within, or coincide with, two adjacent vertices of attachment of the other bridge. In figure 1, vertices of attachment of bridge edge g are vertices 3 and 6. Edge g has one vertex of attachment within those of edge e, and the other outside; therefore the e and g are incompatible and must be assigned to opposite sides of the circuit for the graph to be planar.

The theorem for planarity required that the bridges themselves are planar graphs. There are three possible types of bridge:

- (1) a direct connection to the circuit, for example edges e and g in figure 1.
- (2) a vertex not in the circuit together with its edges of attachment, for example vertex 4 and edge c.
- (3) a group of interconnected vertices with their edges of attachment. (There are none in figure 1.)

Bridges of types 1 and 2 are obviously planar, but bridges of type 3 may not be. Fisher and Wing⁶ propose an iterative method of planarity testing. Initially the whole graph is considered. A circuit is selected and the bridges are identified, tested for incompatibility and assigned to inside or outside of the circuit. Then new sub-graphs are formed involving each type 3 bridge in turn and the process is repeated.

The Fisher and Wing algorithm utilizes the incidence matrix representation of a graph. It arranges the rows (vertices) and columns (edges) of the matrix in such a way that the unique visual drawing of the planar graph can be obtained quickly manually.

Interpretation for plant layout

Some modifications to the Fisher and Wing procedure are required, both in order to relate it to plant layout and to achieve the desired objectives.

Although a circuit can usually be formed from the relationship diagram, it may involve some minor work

flows, or indeed might incorporate a flow which is highly undesirable in the practical sense. However, in practice there are usually operations or machines which are distinctly first operations, for example saw, and others which are definitely final operations, for example inspect, and consequently it is more meaningful to replace the circuit by a chain. The two areas for assignment of bridges become above or below the chain, instead of inside or outside the circuit. The chain may be selected from the travel chart or from the minimal spanning tree. The substitution of a chain for a circuit makes no difference to the basic planarity theorem. This chain will be referred to as the principal chain.

The principal objective of the Fisher and Wing algorithm, in common with most other algorithms for

planarity testing, is to determine whether or not a distinct graph is planar. In practical terms the graph might represent an electronic circuit and one needs to know whether it can be effected on a single-layer printed circuit board. For plant layout the objective is to build up the maximal planar graph from the minimal spanning tree. This requires that Fisher and Wing's iterative method be used virtually in reverse order. Start with a sub-graph, the minimal spanning tree or principal chain, which is known to be planar, and gradually add edges, testing for planarity at each stage. Whenever a type 2 bridge evolves into a type 3 bridge, the type 3 bridge must be examined more closely. A subsidiary chain through the bridge, with ends on the principal chain, is identified. Sub-bridges within the type 3 bridge are identified and tested for

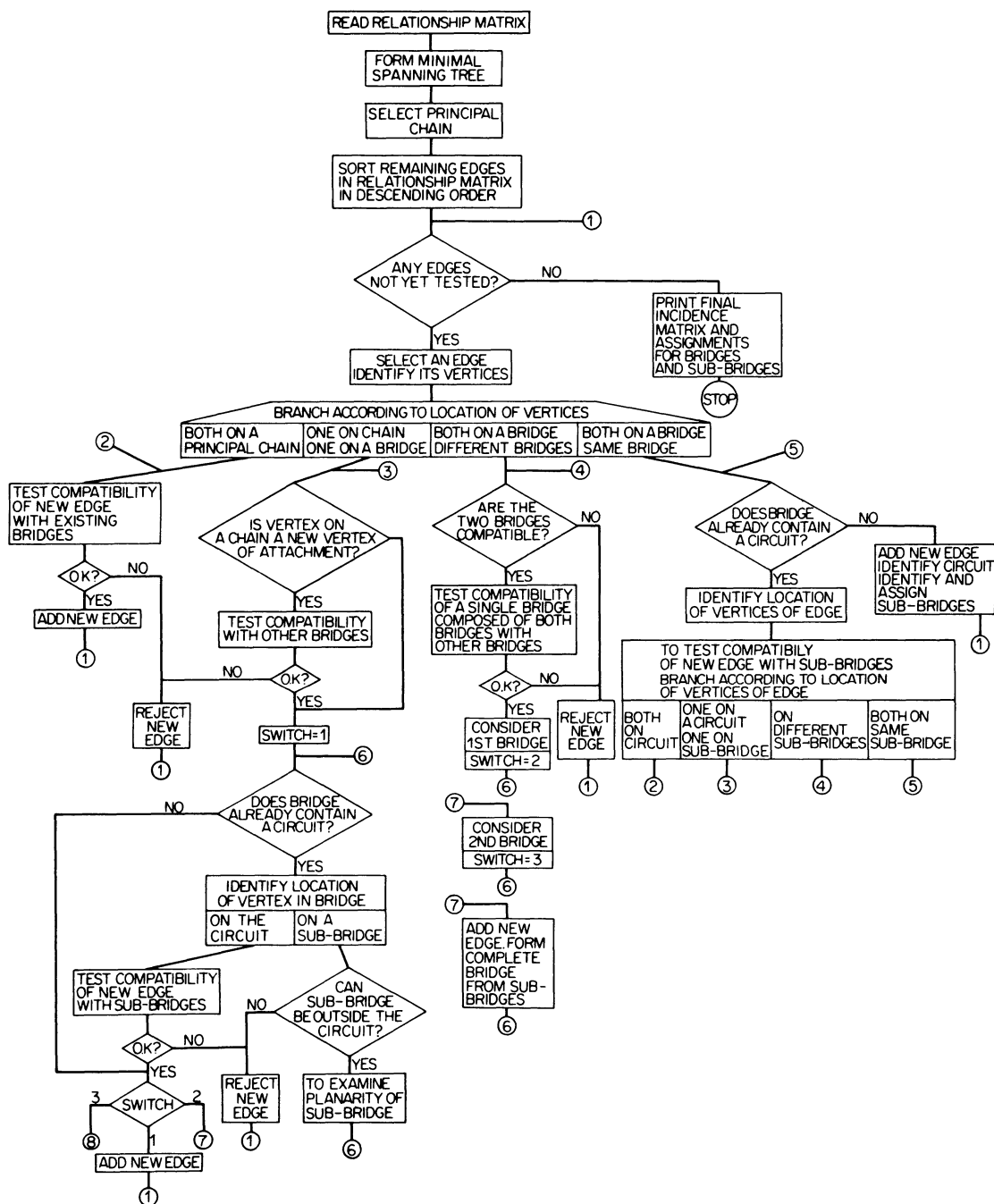


Figure 3. Flow diagram for plant layout planar graph algorithm.

compatibility. This procedure may be nested to successive levels. An outline flow diagram for the procedure is given in figure 3.

examined for the two aspects of the plant layout problem.

RESULTS

Proportion of material handling included in maximal planar graph

Carrie⁵ quoted the proportion of material handling included in the minimal spanning tree in five travel chart cases. These results are shown in Table 1, which also gives data on the distribution of entry values, and the results of the maximal planar graph in the same five cases.

The average proportion of zero entries was 76 per cent with more than half of the non-zero entries containing trivially small values. This accounts for the high proportion of material handling effort included in the minimal spanning tree, especially of the handling due to important entries. The success of the minimal spanning tree in this respect will be seen to depend on the large proportion of material handling accounted for by a small number of entries.

The maximal planar graph includes over 90 per cent of material handling effort in all cases, and almost 100 per cent in two of them. This appears a highly satisfactory result, and shows that a planar graph can be extracted from the relationship matrix with very little loss. The question of using the maximal planar graph as a layout diagram will now be

Preparation of plant layout block diagrams

To illustrate the possibilities of graph theory in the preparation of block diagrams for the layout of homogeneous activities, consider the simple engineer-

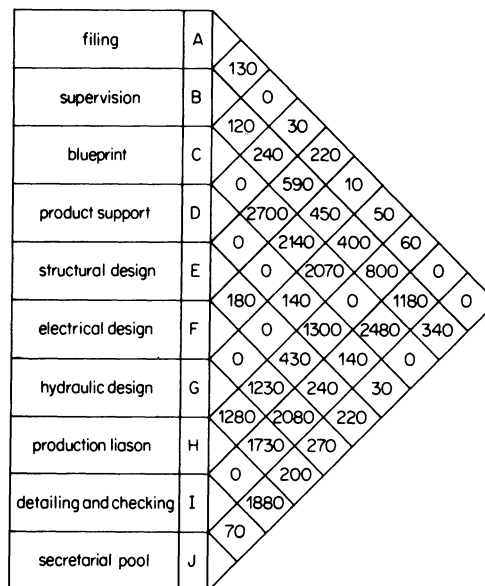


Figure 4. Relationship matrix for block diagram example (based on Buffa⁷).

TABLE 1 Proportion of material handling included in minimal spanning tree, and maximal planar graph

		travel chart case numbers					
		1	2	3	4	5	
<i>Data on the cases</i>							
number of machines/activities		59	27	20	17	10	
per cent of entries which were:	highly important*	} 1	2	3	5	16	
	important		3	7	3	9	
	trivial		1	3	5	7	7
	very trivial		6	3	16	7	17
	zero		92	89	69	78	51
per cent of total material handling accounted for by travel chart entries which were:	highly important	34	51	36	74	78	
	important	15	29	39	8	14	
	trivial	18	14	13	13	5	
	very trivial	33	6	12	5	3	
<i>Minimal spanning tree</i>							
cumulative percentage of material handling effort included in minimal spanning tree due to travel chart entries which were at least:	highly important	100	100	100	94	67	
	important	98	83	86	88	64	
	trivial	86	74	77	78	57	
	very trivial (all non-zero)	65	70	69	75	56	
<i>Maximal planar graph</i>							
cumulative percentage of material handling effort included in maximal planar graph due to travel chart entries which were at least:	highly important	100	100	100	100	97	
	important	100	100	100	98	96	
	trivial	99	98	100	96	93	
	very trivial (all non-zero)	98	98	92	95	91	

* Travel chart entries which accounted for 2 per cent or more of total material handling were considered highly important; those of 1 per cent or more were considered important. Those below 1 per cent were considered trivial, while those of less than 1/2 per cent were considered very trivial.

ing office layout problem involving ten activities, given by Buffa⁷. This is case 5 in Table 1.

The relationship matrix in figure 4 gives the number of face-to-face contacts between activities weighted by the wage rate of the personnel in each activity. The minimal spanning tree derived from it is given in figure 5. Table 1, column 5, shows that the

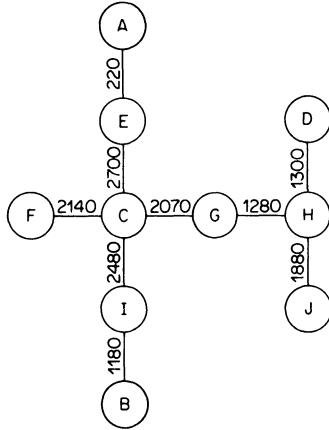


Figure 5. Minimal spanning tree for block diagram example.

minimal spanning tree includes 56 per cent of all material handling. The comparatively low proportion of handling due to very important and important entries included in the minimal spanning tree is due to the exclusion of certain very important relationships, such as E-I, F-I, G-I and F-H, which would form circuits in the tree. As observed earlier, this is related to the high proportion (25 per cent) of important entries. It is to be expected that the

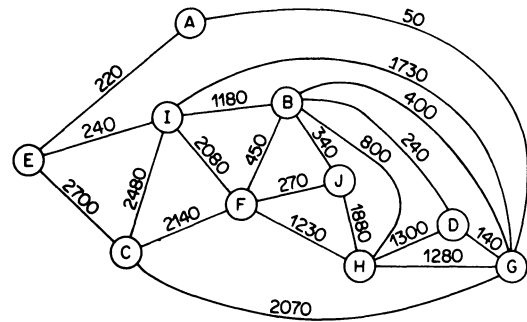
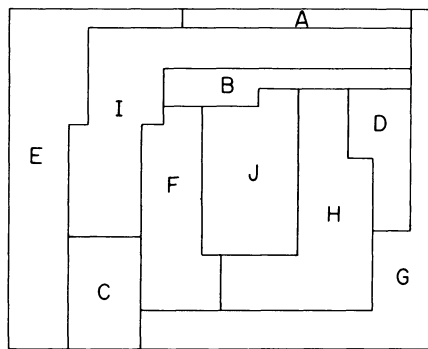


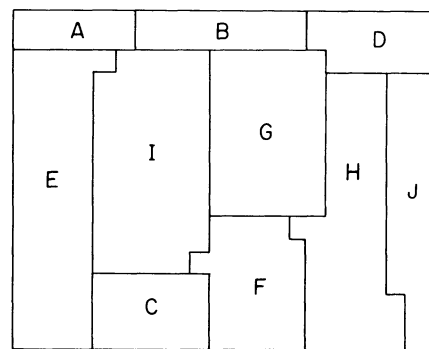
Figure 6. Maximal planar graph for block diagram example.

maximal planar graph would provide much better results and, in fact, includes 91 per cent of the handling effort. The maximal planar graph is shown in figure 6.

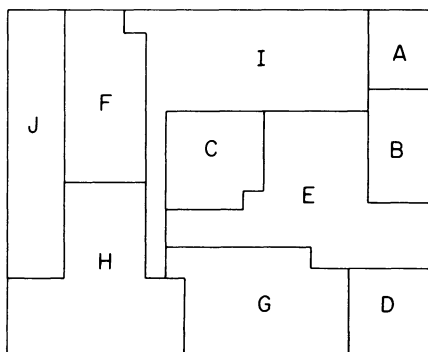
Seppanen and Moore's proposition concerning adjacency and satisfied relationships implies that, for example, activity G should have a common boundary with activities A, B, C, D, H and I, and that B should be adjacent to D, F, G, H, I and J. A block plan satisfying these adjacency requirements can be drawn by hand in a few minutes from the graph, and is shown in figure 7(a). It will be observed that satisfying all relationships of the maximal planar graph imposes some contortion of the activities, especially G. Buffa gives the final layout obtained by the CRAFT program for this problem, depicted in figure 7(b). Comparison of these two results indicates that the graph theory layout is superior to CRAFT's, with a cost of 89 per cent of CRAFT's. This result is due in part to CRAFT's method of evaluating solutions, which assumes that all movements are



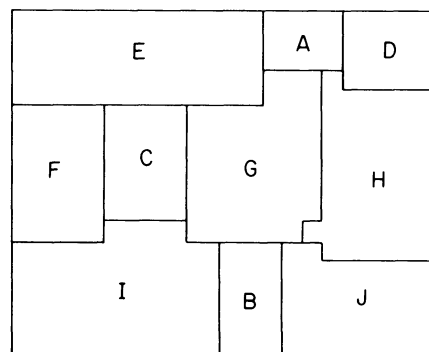
(a)



(c)



(b)



(d)

Figure 7. (a) Plant layout block diagram from maximal planar graph. (b) Plant layout block diagram by CRAFT (based on Buffa⁷). (c) Modified plant layout diagram from maximal planar graph. (d) Plant layout diagram from minimal spanning tree.

between the centres of gravity of activities and therefore does not penalize activities of G's shape. One can deduce that CRAFT would obtain zero cost if the departments were arranged as concentric rectangles.

The shape and narrowness of the activities renders a solution of figure 7(a) impractical. However, if one examines the maximal planar graph, figure 6, the basic cause is immediately apparent. The planar graph algorithm forced activities B, D, H and J to lie within the face bounded by I, F, C and G, and consequently G was forced out to the right to make way for them despite its very considerable contact with C and I. An improved solution is obtained by interposing G between I, F and C and B, J and H, figure 7(c). This reduces the cost to 87 per cent of CRAFT's solution and, allowing for the 'bit size' effect, all activities enjoy rectangular shape.

This discussion would not be complete without considering a block diagram prepared from the minimal spanning tree. Figure 7(d) shows such a layout. The cost of this layout is 87 per cent of CRAFT's, coinciding with the best result from the full planar graph. This is largely fortuitous, in this case, and arises from A and B being folded towards D and J respectively to fit inside the overall rectangle. This improves the position of I relative to E and F, with which it has much business. Nevertheless it shows that simple techniques can give good results.

Preparation of machine shop layouts

The case where the vertices of the graph are machine tools whose shape and area are not amenable to the unit area device, will be illustrated with reference to an earthmoving equipment manufacture study. This is case 1 in table 1, and was used by Carrie⁵ to illustrate the minimal spanning tree technique. The schematic layout diagram was obtained from the minimal spanning tree by providing at each vertex a circle of the required area and by arranging the tree within an overall rectangular area, as shown in figure 8. The size of each circle is proportional to the area required by a machine of that type, or by more than one machine where the work load required more. Each circle contains an abbreviation of the machine's name and a code number. The paths of flow included in the minimal spanning tree are also shown. The significance of the layout in terms of work flow were described⁵ as follows:

INSP2 near the centre of the figure is the department where completed components and sub-assemblies are inspected or where partially machined components are inspected prior to heat treatment. Situated close to this department, is the assembly department which is also the area where components are stored, if necessary, prior to assembly. The completed unit moves up to the right through testing, washing and packing to the despatch area. All the the heat treatment processes are together at the central right-hand side. These are close to INSP2, through which partly machined components requiring heat treatment must pass. After a component has been treated it must be inspected at INSP3 before returning for further machining. This further machining is

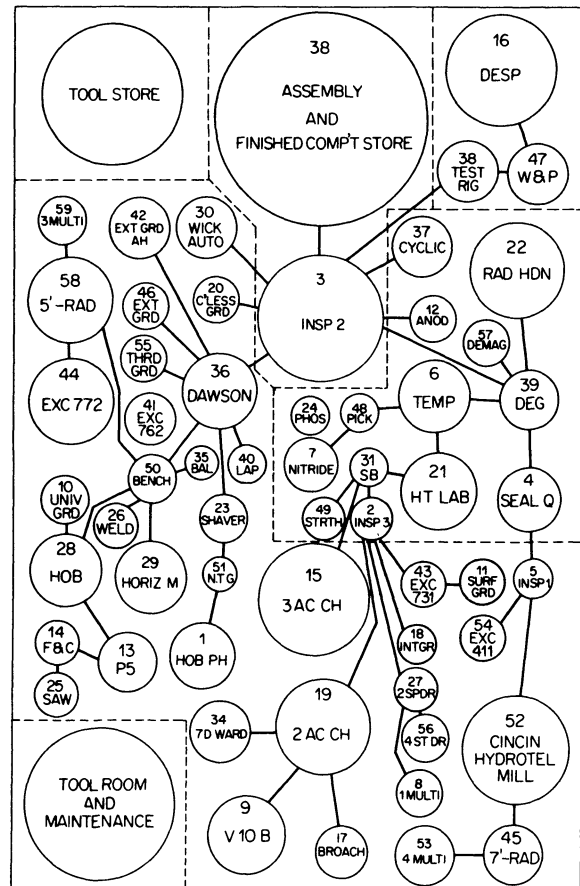


Figure 8. Machine shop layout from minimal spanning tree (after Carrie⁵).

usually drilling or grinding so that some machines of these types are located close to INSP3 at the lower right. On the other hand since grinding is usually a final operation most of the grinding machines are situated at the upper left, near the DAWSON automatic washing machine, which itself is adjacent to final inspection. Manual deburring at a BENCH, also usually a final operation, is also situated near the washing machine. Forgings and castings are checked, before processing, at INSP1, to the right of the diagram. They then move into the heat treatment section or to the machine shop for milling and drilling. Several splined shafts are components of the transmission unit and a cell for these is at the lower left, consisting of SAW, face and centering machine, (F&C), copy lathe (P5) and HOB. There are also several gears and the nucleus of a cell for these is in the centre of the diagram consisting of a hob, (HOBPH), deburring machine (NTG) and a gear shaving machine (SHAVER). The Wickman automatic lathe WICKAUTO is situated near INSP2 since many of the components produced by it are completed in a single operation. The lathes not mentioned already are in a functional group at the bottom centre of the diagram.

This layout ensures that 65 per cent of all handling is between adjacent machines, probably comparing well with industrial practice. However, as observed earlier, the minimal spanning tree gives no information about material flows not included in it,

and therefore no indication of the relative position of the edges incident on each vertex.

The maximal planar graph will give this information, and is shown in figure 9. In this case the

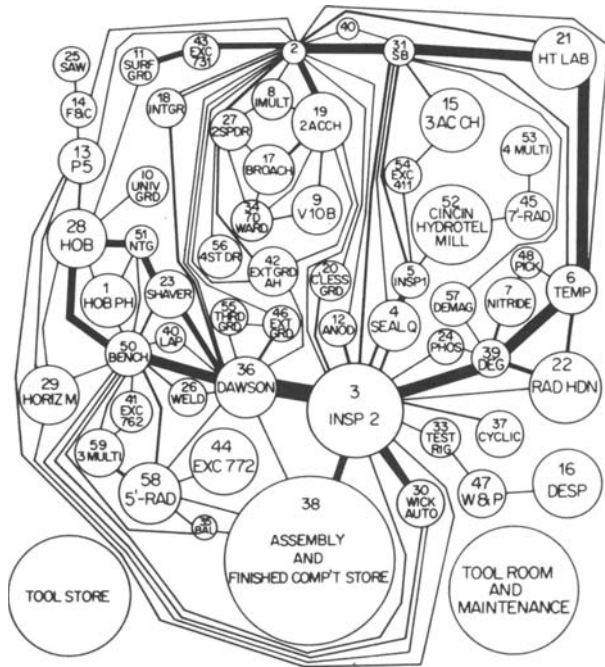


Figure 9. Machine shop layout from maximal planar graph.

connections in the graph largely determine the overall shape of the layout. The thickness of the lines indicates the relative magnitude of the flow. The graph accounts for 98 per cent of all the material handling. The main path, from 28 through 50 and 36 round to 2, 43 and 11, alone accounts for 34 per cent of all handling, and when minor flows parallel to it, for example 2 to 21, are included, this figure is raised to 39 per cent.

Apart from identifying the specific size and shape of the machines, very little adjustment would be required to produce a practical layout from this schematic one. However, two aspects of the graph theory are inappropriate, or perhaps over-rigorous, and require some adjustments.

In the previous example it was observed that the algorithm forced G outwards to make way for B, D, H and J. In this example the same effect can be observed in that 21, 31 and 2 have been moved outwards to the top to make way for two groups of machines, 15, 53, 45, 52, 54 at top right and 19, 9, 34, 17, 9, 27, 8, 56 and 42 at top centre. The overall material handling cost, the volume-distance product, could be improved by positioning 21, 31 and 2 below these machines and re-orienting them to suit. This would also bring all the heat treatment and related processes together, as would be needed in practice.

The other effect is the showing of flow paths around the periphery of the diagram, for example 2 to 6 and those below 38 at the bottom of the

diagram. In practice these flows could either be added to the main flow or be drawn in as straight lines. As a result 30 near the bottom right would be placed to the left of 3.

Nevertheless, the manual content of the exercise—drawing the diagram and deducing these adjustments—can be completed quickly. With the high proportion of material handling included in the maximal planar graph and the well-determined relative positions, the approach is promising.

CONCLUSION

This paper has given a brief statement of the principles and results of graph theory applied to plant layout. In particular, the layout problem was interpreted as a planar graph. It was pointed out that the planarity condition was unduly rigorous, but it was useful in that it provided a basic procedure.

The results obtained have been good in terms of the proportion of material handling taken into account, and in that a good plant layout diagram can be very quickly produced from the results. The technique therefore fulfils the basic essentials of any method. Two criticisms of the results were made, that heavy traffic activities were forced outwards to make way for light traffic ones, and that minor flows were not combined with parallel major ones.

The approach is promising. A computer program, which will eliminate or minimize these two limitations, is being developed.

The algorithm described in this paper has been found excessively cumbersome and an alternative algorithm has been developed based on Tarjan⁸.

REFERENCES

1. E. S. Buffa, G. C. Armour and T. E. Vollman. Allocating facilities with CRAFT, *Harvard Business Review* (March–April 1964), pp. 136–158.
2. R. C. Lee and J. M. Moore. CORELAP—computerised relationship layout planning, *J. Ind. Engng* (March 1967) 18, no. 3.
3. P. C. Portlock and B. Whitehead. A program for practical layout planning, *Building Science* (1970) 6, pp. 213–230.
4. J. Seppanen and J. M. Moore. Facilities Planning with graph theory, *Management Science* (Dec. 1970) 17, no. 4, pp. B242–B253.
5. A. S. Carrie. Numerical taxonomy applied to group technology and plant layout, *Int. J. Prod. Res.* (1973) 11, no. 4, pp. 399–416.
6. G. J. Fisher and O. Wing. Computer Recognition and extraction of planar graphs from the incidence matrix, *IEE Transactions on Circuit Theory* (June 1966), CT-13, no. 2, pp. 154–163.
7. E. S. Buffa. *Modern Production Management*, 3rd Edition. Wiley, 1969, pp. 353–356, also Instructor's Manual to this edition, p. 79.
8. R. E. Tarjan. An efficient planarity algorithm. Report No. STAN-CS-244-71, Stanford University, Nov. 1971.

MACHINE TOOL ELEMENTS

AN ANALYSIS OF A NEW TYPE OF HIGH PRECISION MULTIWEDGE CLEARANCELESS BEARING FOR GRINDING WHEEL SPINDLES

by

G. K. ARORA* and T. S. R. MURTHY*

SUMMARY

This paper presents a theoretical analysis of a new type of multiwedge clearanceless bearing for grinding wheel spindles. The analysis establishes the importance of three parameters, namely initial radial clearance, sleeve stiffness and spindle preload, on the performance of clearanceless bearing. The special feature of this bearing to self-adjust its working clearance according to the speed of the spindle is clearly borne out by this analysis.

NOTATION

<i>E</i>	modulus of elasticity of sleeve material	<i>t</i>	section containing minimum oil film thickness or from the support point
<i>I</i>	moment of inertia	α	sleeve thickness
<i>K</i>	stiffness of sleeve	δ	form factor defining oil film shape
K_{HD}	oil film stiffness	δ_0	initial radial clearance between undeformed sleeve and spindle
<i>L</i>	length of sleeve	ϵ	minimum oil film thickness at no load-spindle running
<i>P</i>	total preload on sleeve	η	e_c/δ
P_0	hydrodynamic force generated at supports 1, 2 and 3 under no load-spindle running		oil viscosity
P_1	hydrodynamic force generated at support 1 under external load <i>W</i> -spindle running; this is approximately equal to hydrodynamic force generated at support 2		
P_3	hydrodynamic force generated at support 3 under external load <i>W</i> -spindle running		
P_D	load required to deform the sleeve to three-lobe configuration so that the radial clearance δ is taken up		
P_S	preload on spindle		
<i>U</i>	surface speed of spindle		
<i>W</i>	external load		
<i>d</i>	diameter of spindle		
<i>e</i>	displacement of spindle centre on application of external load <i>W</i>		
e_c	displacement of spindle centre from reference centre on application of an external load <i>W</i>		
h_1	oil film thickness at support 1 under external load <i>W</i> -spindle running		
h_3	oil film thickness at support 3 under external load <i>W</i> -spindle running		
h_{max}	maximum oil film thickness		
h_{min}	minimum oil film thickness or oil film thickness at support point		
$h(\theta)$	thickness of oil film at an angle θ from the		

1 INTRODUCTION

In precision grinding operations, geometrical and form accuracies as well as surface finish of the components are, to a large extent, dependent on the performance of the spindle bearings. The principal requirements of such a bearing are high rigidity, smooth and accurate rotation of spindle, small temperature rise and minimum wear. Moreover, the recent efforts to achieve higher efficiency in grinding indicate that for maximum benefit grinding wheel speed should be changed according to work-piece material; this will require grinding wheels to run at peripheral speeds from 20 m/s to 60 m/s and, therefore, will be an additional requirement on the wheel spindle bearings.

A high-precision multiwedge bearing developed at CMTI¹ shown in figure 1(a) and its modified version² shown in figure 1(b) have given excellent results. A high degree of performance was achieved in respect to all the requirements mentioned above.

This paper presents a theoretical analysis of such a bearing and establishes the parameters which control its performance.

* Department of Research and Testing, Central Machine Tool Institute, Bangalore, India

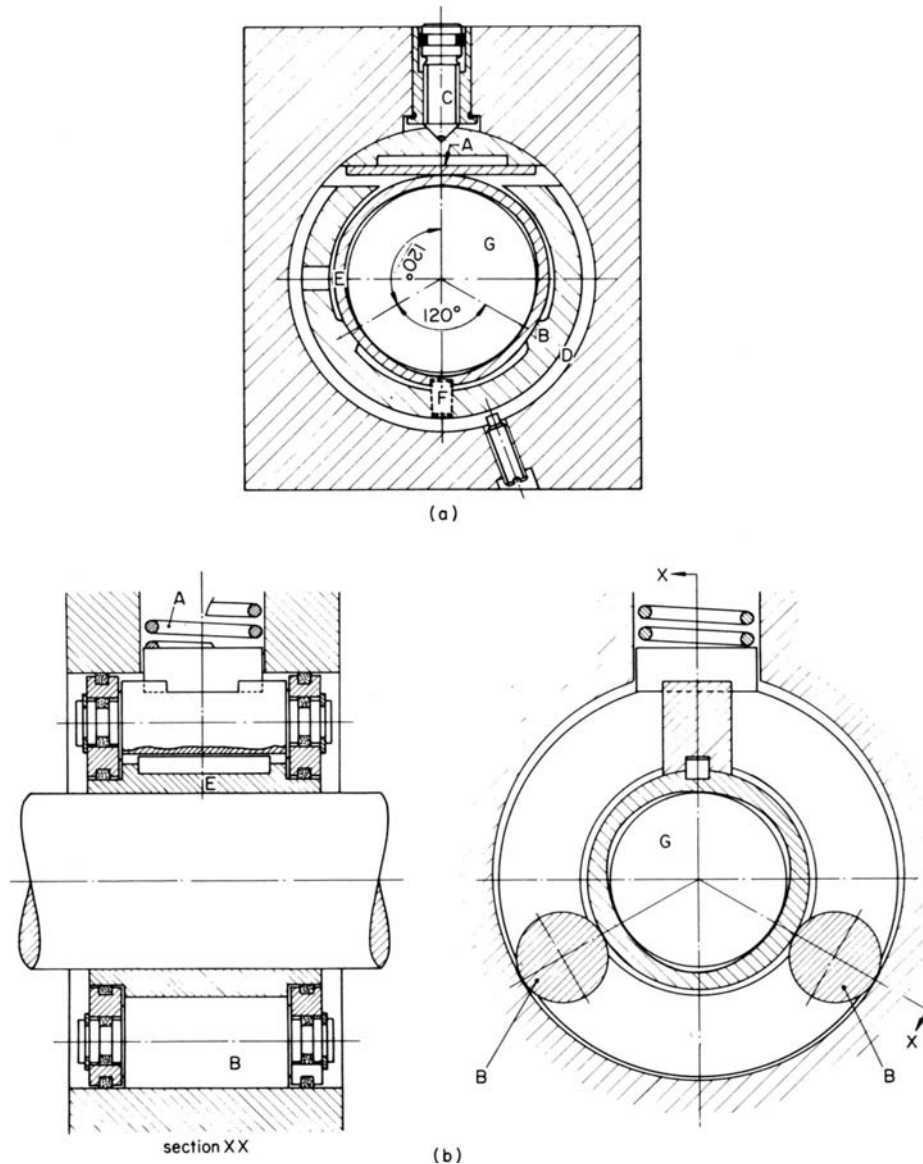


Figure 1. (a) Clearanceless journal bearing (after CMTI). A: loading spring; B: support; C: loading screw; D: oil groove, E: flexible metallic sleeve; F: retaining pin; G: spindle. (b) Hostar hydrodynamic journal bearing (after Linda and Rosberg²) A: loading spring; B: roller supports; E: flexible metallic sleeve; G: spindle.

2 FUNCTIONING AND MATHEMATICAL ANALYSIS

The main feature in the design of this bearing is a flexible metallic sleeve which can be easily deformed to a three lobe configuration. The three lobes which are equi-spaced at 120° provide the hydrodynamic oil wedges for supporting the spindle and the external load, and also provide the working clearance when the spindle is running.

2.1 Functioning of clearanceless bearing

For the sake of clarity, functioning of this bearing will be explained under three conditions:

- (1) no load- spindle stationary;
- (2) no load- spindle running;
- (3) external load- spindle running.

Condition 1

The spring (A) is compressed by the screw (C) (figure 1) to deform the sleeve (E) to a three-lobe shape so as

to take up all the radial clearance (δ) between the undeformed sleeve and the spindle (figures 1 and 2(a)). To achieve bearing operation with optimum working clearance, it is usually necessary to compress the spring further by a predetermined amount to create the preload P_S on the spindle. The total preload P on the sleeve is given by

$$P = P_D + P_S$$

Condition 2

When the spindle starts running, oil is drawn into the wedges; this partially restores the sleeve to its original circular form (figure 2(b)). If the external load at this stage is assumed to be zero, oil films of the same minimum thickness (δ_0) will be generated at all the three supports. Since the oil film thickness at all three supports is δ_0 and the supports are equally spaced, the spindle centre moves by $2\delta_0$ and the sleeve shifts by $3\delta_0$ against the static preloading force P .

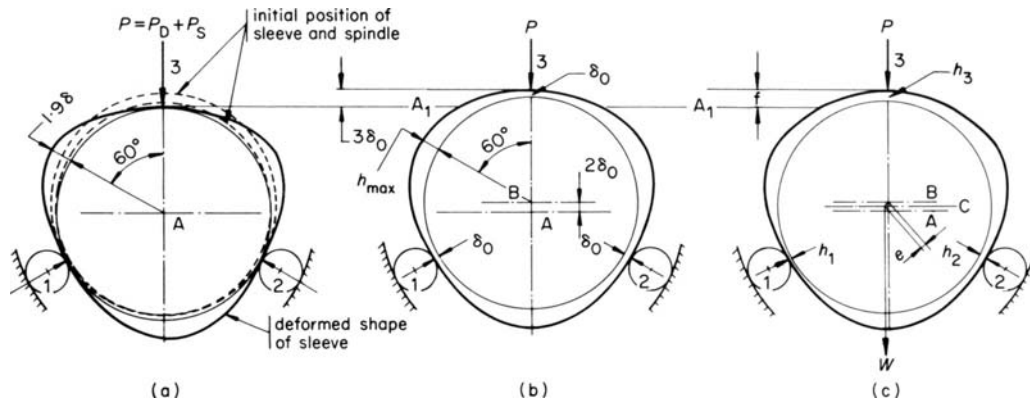


Figure 2. Spindle and sleeve positions: (a) sleeve deformed-spindle stationary; (b) spindle running-external load zero; (c) spindle running-external load W .

Condition 3

The application of an external load W to the spindle moves it to a new position of equilibrium. The external load here includes the weights of the spindle, the pulley and the grinding wheel, and is modified by the driving and the grinding forces while grinding operation is in progress. According to this new position of equilibrium the oil wedge shapes also change and new oil film thicknesses h_1 , h_2 and h_3 are established at supports 1, 2 and 3 respectively (figure 2(c)).

2.2 Mathematical analysis

An exact mathematical analysis of this bearing is not possible due to the complexity of its operation, but an analysis, under some reasonable assumptions, which provides a satisfactory solution is possible. The assumptions made in the present analysis are as follows:

(1) The oil wedges in the bearing are of exponential shape given by the relationship

$$h(\theta) = h_{\min} e^{-\alpha\theta} \quad (1)$$

This shape has been assumed because the mathematical relationship, defining it, is easily integrable and thus facilitates the mathematical analysis. The shape of oil wedge has virtually no influence on load carrying capacity³; it is the thickness of the oil wedge at its inlet and at its outlet which determines the load carrying capacity. This was further confirmed by comparing the hydrodynamic load generated by the exponential wedge shapes with that generated by the exact wedge shapes present in our case. The results were found to be close enough to justify this assumption.

(2) The line of action of the resultant hydrodynamic force generated by each oil wedge passes through its support point. According to Sommerfeld's theory⁴, the pressure in the clearance space is a continuous function so that, by reasons of symmetry, the pressure reaches zero at points of maximum and minimum oil film thickness, being negative during the diverging portion. Reynolds⁵, on the other hand, predicts that the pressure reaches zero where its first differential also becomes zero. In both instances the pressure distribution is supposed to begin at the

position of maximum oil film thickness. Experimental results⁶ show that pressure is never zero at minimum oil film thickness (this is consistent with Reynolds) and at higher external loads the resultant hydrodynamic force passes very close to the point of minimum oil film thickness, in our case the support points; for loads at which the bearing operates in practice, the line of action of the resultant hydrodynamic force does not show a shift of more than 10 to 15° from the position of minimum oil film thickness. These results suggest that the assumption made here should in fact provide a more realistic analysis.

(3) The hydrodynamic load will be calculated by using the Reynolds equation for a short bearing. In case of clearanceless bearings, the L/d ratio normally varies from 0.5 to 0.8 and is small enough to yield good results by approximating it to a short bearing.

2.3 Equations of equilibrium

Condition 2

When the spindle rotates, hydrodynamic pressure builds up in the wedge shapes. If the weight of the spindle is assumed to be zero, the spindle will take up an equilibrium position such that the minimum film thickness (δ_0) at all the three lobes is equal. Therefore,

$$h_{\min} = \delta_0$$

and according to Appendix A and figure 2(b),

$$h_{\max} = 1.9\delta - 0.9\delta_0 \quad (2)$$

The hydrodynamic load (P_0) generated at support 3 will act against the preload (P) such that

$$P = P_0 + K(\delta - \delta_0) \quad (3)$$

According to Appendix B the hydrodynamic load (P_0) is given by

$$P_0 = \frac{K_0}{\delta_0^2} \quad (4)$$

where

$$K_0 = \frac{U\eta L^3}{2} \cdot \frac{\alpha}{(1 + 4\alpha^2)} \left[2\alpha + \frac{(\sqrt{3} - 2\alpha)}{2} e^{-(2\alpha/3)} \right]$$

and α can be calculated from the relationship

$$h_{\max} = h_{\min} e^{-\alpha\pi/3}$$

Equation (3) on substitution from (4) can be expressed as

$$P = \frac{K_0}{\delta_0^2} + K(\delta - \delta_0)$$

Hence

$$P - K\delta = \frac{K_0}{\delta_0^2} - K\delta_0$$

but $P_D = K\delta$ and therefore

$$P_S = \frac{K_0}{\delta_0^2} - K\delta_0 \quad (5)$$

Condition 3

On the application of an external load, the spindle is shifted to a new equilibrium position (figure 3).

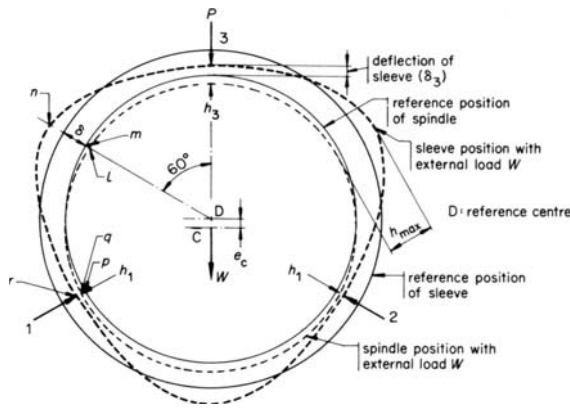


Figure 3. Spindle and sleeve positions for condition 3 (spindle running-external load W).

According to this new equilibrium position, the film thickness at an angle of 60° from support point 3 is given by

$$h_{\max} = lm + mn = \delta\epsilon \cos 60^\circ + 1.9\delta - 0.9(h_3 - \delta\epsilon) \quad (6)$$

Also,

$$h_3 = h_1 + 1.5\delta\epsilon \quad (7)$$

The necessary conditions for equilibrium of the spindle and of the sleeve are

$$P_1 = P_3 + W \quad (8)$$

and

$$P = P_3 + K[\delta(1 + \epsilon) - h_3]$$

which on simplification reduces to

$$P_S = P_3 + K(\delta\epsilon - h_3) \quad (9)$$

Equations (5), (8) and (9) are not solvable by deterministic techniques. However, an iteration technique using a computer and graphic representation can provide a solution.

2.4 Oil film stiffness of the bearing

From figure 3 it can be seen that on the application of an external load W the oil film thickness changes

from δ_0 to h_1 . However, the change in position of spindle centre from no-load to loaded condition is $2(\delta_0 - h_1)$. Therefore the oil film stiffness K_{HD} is given by

$$K_{HD} = W/[2(\delta_0 - h_1)]$$

3 EXAMPLE

This example illustrates the procedure for calculating oil film thickness and oil film stiffness.

Sleeve dimensions:

- length, $L = 30$ mm
- inside radius, $r = 26.5$ mm
- wall thickness, $t = 3$ mm
- initial radial clearance, δ , between the undeformed sleeve and the spindle = $15.0 \mu\text{m}$

- Material of sleeve:* bronze, for which $E = 10\,500$ kgf/mm²
 viscosity of oil, $\eta = 6^\circ\text{E}$
 spindle speed, $N = 2150$ rev/min

- (1) The computer output gives values of P_1 and P_3 for arbitrarily assumed values of h_1 and ϵ . Table 1 shows values for ϵ , h_1 , h_3 , h_{\max} , P_1 , P_3 and P_S . Figure 4 is a graph of P_1 against P_3 .

TABLE 1 Representative computer output

ϵ	h_1 (μm)	h_3 (μm)	h_{\max} (μm)	P_1 (kgf)	P_3 (kgf)	P_S (kgf)
0.00	1	1	27.60	19504.0	19504.0	19502.0
	2	2	26.70	4795.0	4795.0	4789.9
	3	3	25.80	2089.7	2089.7	2082.1
	4	4	24.90	1147.7	1147.7	1137.6
	5	5	24.00	713.16	713.16	700.57
	6	6	23.10	477.36	477.36	462.25
	7	7	22.20	334.85	334.85	317.23
	8	8	21.30	241.72	241.72	221.57
	9	9	20.40	177.00	177.00	154.34
	10	10	19.50	129.66	129.66	104.48
	11	11	18.60	93.41	93.41	65.71
	12	12	17.70	64.43	64.43	34.22
	13	13	16.80	40.26	40.26	7.52
	14	14	15.90	19.20	19.20	-16.05
	15	15	15.00	00.00	00.00	-37.77
0.02	1	1.45	27.61	19504.0	9212.3	9209.4
	2	2.45	26.71	4795.1	3170.8	3165.4
	3	3.45	25.81	2089.8	1566.0	1558.1
	4	4.45	24.91	1147.8	917.51	907.1
	5	5.45	24.01	713.22	592.68	579.7
	6	6.45	23.11	477.42	406.81	391.3
	7	7.45	22.21	334.92	290.20	272.2
	8	8.45	21.31	241.79	211.79	191.3
	9	9.45	20.41	177.07	156.05	133.0
	10	10.45	19.51	129.74	114.50	88.9
	11	11.45	18.61	93.50	82.14	55.07
	12	12.45	17.71	64.53	55.87	25.28
	13	13.45	16.81	40.37	33.65	00.54
	14	14.45	15.91	19.32	14.03	-21.58
	15	15.45	15.01	14.65	-4.05	-42.20

- (2) Let us assume that the external load on the bearing is equal to 80 kgf. Draw a line at 45° to the x-axis starting from the point $x_{-80}(w = 80$ kgf). This is the load line for 80 kgf external load on the bearing (figure 4). This line satisfies equation (8).
- (3) Find the points of intersections 1, 2, ..., 11 with P_1 against P_3 curves for various ϵ values. Table 2 shows the values for ϵ and P_1 .

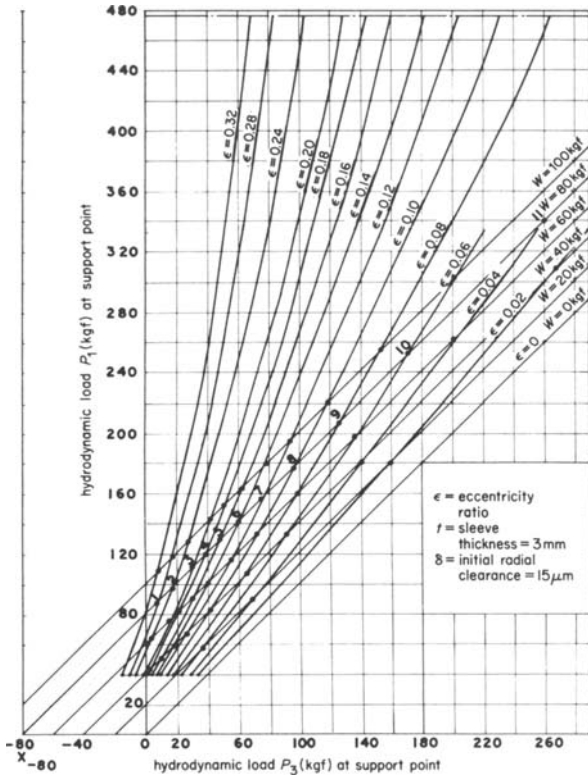


Figure 4. P_1 against P_3 curves.

TABLE 2 Operating points of the bearing

ϵ	P_1 (kgf)	(h_1) (μm)	h_3 (μm)	P_S (kgf)
0.24	97	10.9	16.3	-8
0.20	111	10.45	14.95	-1
0.18	120	10.25	14.30	8
0.16	129	10.00	13.6	19
0.14	141	9.75	12.9	32
0.12	156	9.4	12.1	51
0.10	177	9.0	11.25	71
0.08	207	8.5	10.3	104
0.06	251	7.9	9.25	148
0.04	340	6.95	7.85	238

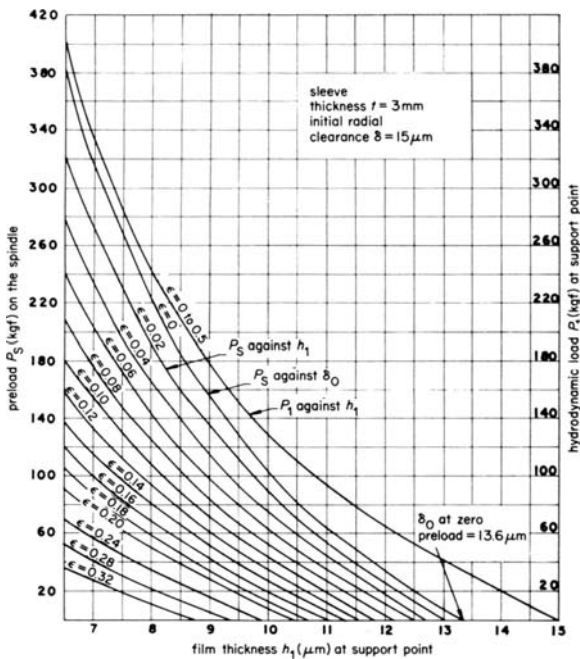


Figure 5. P_1 against h_1 , P_S against δ_0 and P_S against h_1 curves.

- (4) Draw the curves for P_1 against h_1 for different values of ϵ (computer output table 1) as shown in figure 5.
- (5) Find h_1 values from figure 5 corresponding to P_1 and ϵ values in table 2. These h_1 values are shown in table 2.
- (6) Determine the corresponding h_3 values by the relation

$$h_3 = h_1 + 1.5 \delta \epsilon$$

These values are shown in table 2.

- (7) Draw P_S against h_1 curves for different values of ϵ from table 1 as shown in figure 5.
- (8) Find P_S values from figure 5 corresponding to ϵ and h_1 values in table 2. These P_S are shown in table 2.
- (9) Draw P_S against ϵ curves and h_1 against ϵ curves (table 2) as shown in figure 6(a).

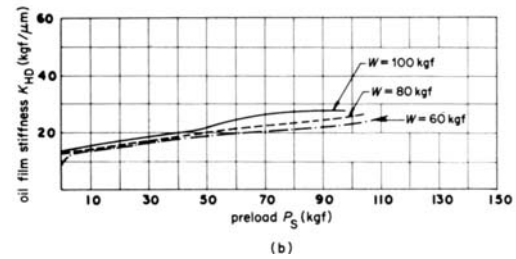
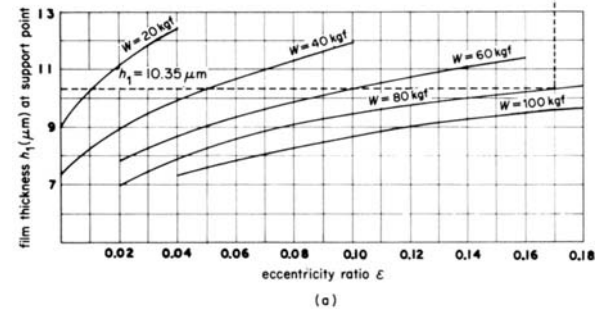
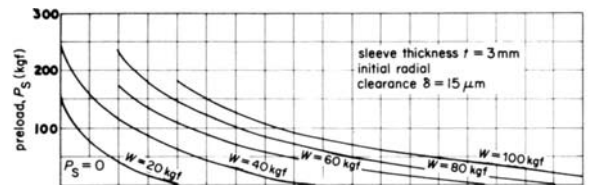


Figure 6. (a) P_S against ϵ and h_1 against ϵ curves. (b) Effect of P_S and W on K_{HD} .

- (10) Let us take a case where preload on the spindle is zero. From figure 6(a) the corresponding value for h_1 is $10.35 \mu\text{m}$ and from figure 5 $\delta_0 = 13.6 \mu\text{m}$.

TABLE 3 Oil film stiffness calculations

P_S (kgf)	δ_0 (μm)	h_1 (μm)	$\delta_0 - h_1$ (μm)	$2(\delta_0 - h_1)$ (μm)	K_{HD} ($= W/[2(\delta_0 - h_1)]$) (kgf/ μm)
0	13.6	10.35	3.25	6.5	12.3
8	13.0	10.25	2.75	5.5	14.55
19	12.55	10.0	2.55	5.1	15.7
32	12.1	9.75	2.35	4.7	17.1
51	11.45	9.4	2.05	4.1	20.0
71	10.8	9.0	1.8	3.6	22.2
104	10.0	8.5	1.5	3.0	26.7
148	9.1	7.9	1.2	2.4	33.4
238	7.8	6.95	0.85	1.7	47.0

(11) Therefore

$$K_{HD} = \frac{W}{2(\delta_0 - h_1)} = \frac{80}{2(13.6 - 10.35)} = 12.3 \text{ kgf}/\mu\text{m}$$

Table 3 lists the values of K_{HD} for different values of P_S ; these are shown graphically in figure 6(b).

4 DISCUSSIONS ON THE RESULTS

Since it is not within the scope of a paper of this nature to present all the data accumulated by computer simulation and graphic constructions, only representative data condensed to a suitable graphic form is given here.

4.1 Effect of initial radial clearance (δ)

Figure 7 shows that an increase in initial radial clearance δ causes a decrease in oil film stiffness

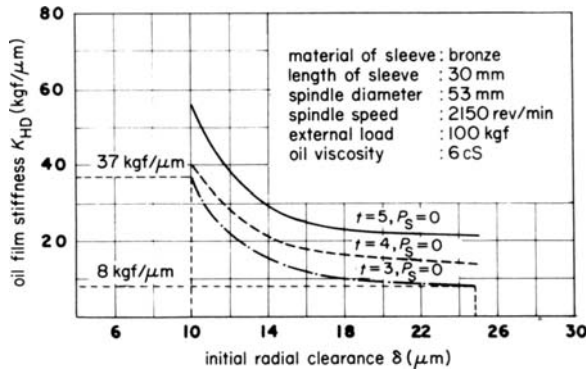
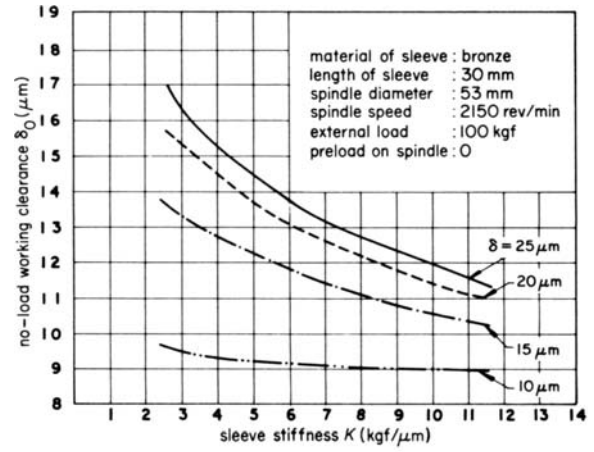


Figure 7. Effect of δ on K_{HD} .

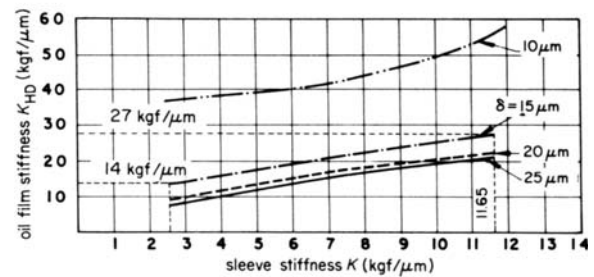
K_{HD} ; for example, for $t = 3$ mm and $P_S = 0$, an increase in δ from 10 to 25 μm ($2\frac{1}{2}$ times) causes a decrease of approximately $4\frac{1}{2}$ times (from 37 $\text{kgf}/\mu\text{m}$ to 8 $\text{kgf}/\mu\text{m}$) in K_{HD} . However, it can also be seen that these curves are steeper at smaller values of δ , in fact after a value of δ around 16 μm the curves tend to be flat and a decrease in the value of K_{HD} is not very significant.

From these results, one can draw the conclusion that for achieving higher oil film stiffness smaller values of initial radial clearance δ should be chosen. But δ has a limiting value which depends on the following two criteria:

- (1) heat generation; a clearanceless bearing with smaller initial radial clearance operates at smaller working clearance (figure 8(a)) and therefore generates more heat.
- (2) minimum permissible oil film thickness based on geometrical and mechanical considerations; the absolute minimum value should in no case be less than twice the shaft surface finish (CLA) added to half the deflection of the spindle over the width of the bearing.



(a)



(b)

Figure 8. (a) Effect of K on δ_0 at no-load running condition. (b) Effect of K on K_{HD} .

4.2 Effect of sleeve stiffness, K , or its thickness, t

Figure 8(b) shows that the effect of increasing K is to cause an increase in K_{HD} , but this increase in K_{HD} is not as pronounced as that obtained by changing δ ; for example for $P_S = 0$ and $\delta = 15$ μm , a change in the K value from 2.5 to 11.65 $\text{kgf}/\mu\text{m}$ (approximately $4\frac{1}{2}$ times) shows an increase of approximately 2 times (from 14 $\text{kgf}/\mu\text{m}$ to 27 $\text{kgf}/\mu\text{m}$) only in K_{HD} , whereas for $t = 3$ mm and $P_S = 0$ a decrease in δ from 25 μm to 10 μm ($2\frac{1}{2}$ times) causes an increase of approximately $4\frac{1}{2}$ times (from 8 $\text{kgf}/\mu\text{m}$ to 37 $\text{kgf}/\mu\text{m}$) in K_{HD} (figure 7). This suggests that for achieving higher oil film stiffness, the stiffness of the sleeve should be increased. But one should not fail to notice from figure 8(a) that at higher K values the bearing operates at smaller working clearance, and consequently at higher temperature.

4.3 Effect of preload P_S

The effect of P_S on K_{HD} is very similar to that of K . An increase in preload on the spindle increases the stiffness K_{HD} (figure 6(b)), but on the other hand it affects the life of the bearing. It is therefore advisable to keep the preload on the spindle at low values only and if possible it should be avoided altogether.

Moreover, one should be extra careful in case of machine tools where frequent starting and stopping is a necessity. In such cases a high preload may wreck the bearing in a short time and should definitely be avoided. The starting and stopping in a machine tool is normally associated with boundary lubrication in its bearings; under such conditions a lot of metal-to-metal friction is presented in the bearing. A higher

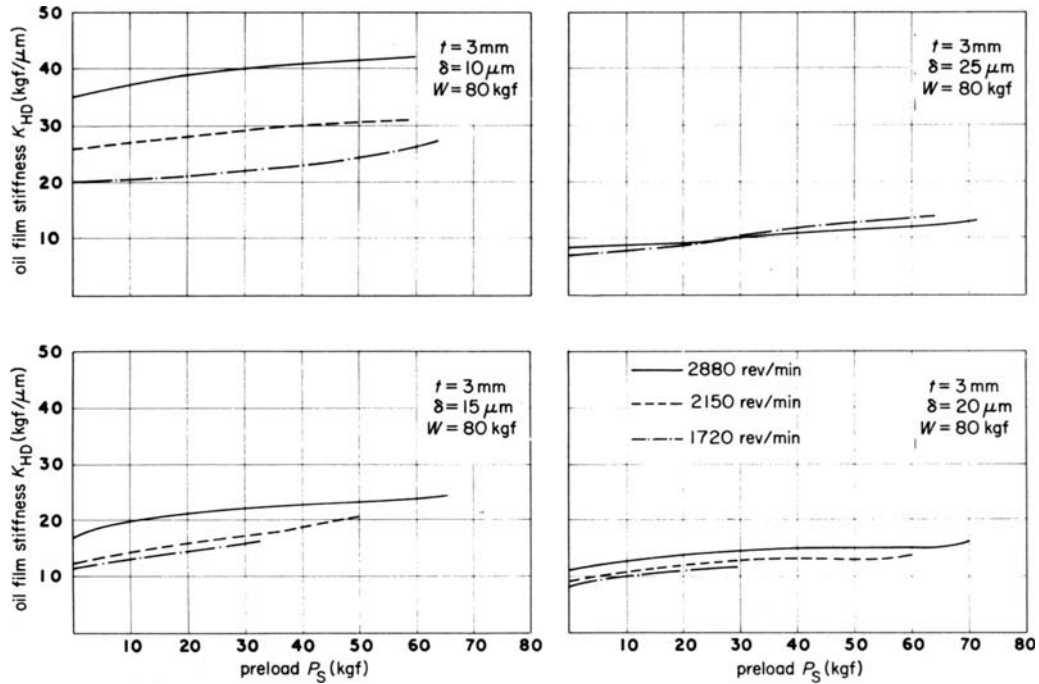


Figure 9. Effect of spindle speed on K_{HD} .

preload simply helps in augmenting this friction and therefore results in increased wear of the bearing.

A very good solution, to increase oil film stiffness of this bearing without affecting the life of the bearing, can be achieved by employing a high preload on the spindle that can be, hydraulically or by some other means, relieved for a very short time (1 to 2 s) at the starting and stopping of the spindle.

4.4 Effect of speed and viscosity

Effect of speed and viscosity on oil film stiffness is quite similar. Here the effect of speed only on K_{HD} (figure 9) is discussed.

The effect on K_{HD} of an increase in spindle speed is very much dependent on the value of initial radial clearance δ . For higher values of δ , a change in speed has practically no effect on K_{HD} . It is only at low values of δ ($\delta \leq 15\ \mu\text{m}$) that an increase in speed shows a noticeable increase in K_{HD} . In fact, by

choosing proper values of t and δ it may be possible to design a bearing whose stiffness is affected very little by speed variations.

With an increase in the speed of the spindle, the working clearance δ_0 at no load and minimum film thickness h_1 under external load show an increase in their values (figures 10 and 11); this gives an added advantage in that an increase in heat generation due to higher speed will be partially nullified by automatic increase in δ_0 and h_1 . This result is consistent with the practical observations made by Linda and Roseberg². Advantage may be taken of this feature of a clearanceless bearing in spindle units where the spindle is required to run at a number of speeds; in the existing designs clearance settings are normally changed by manual adjustments on complicated and difficult to manufacture arrangements.

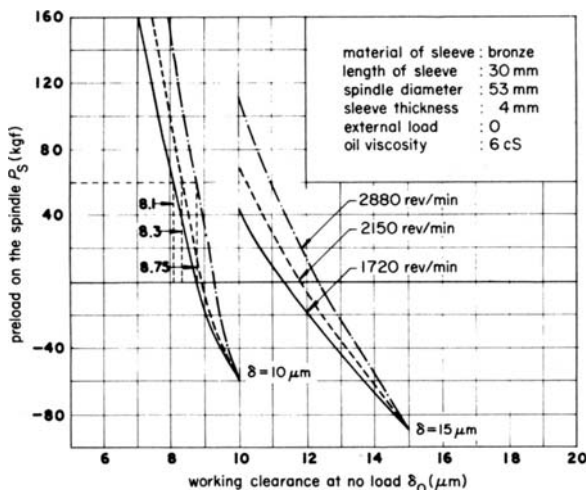


Figure 10. Effect of spindle speed on δ_0 at no-load running condition.

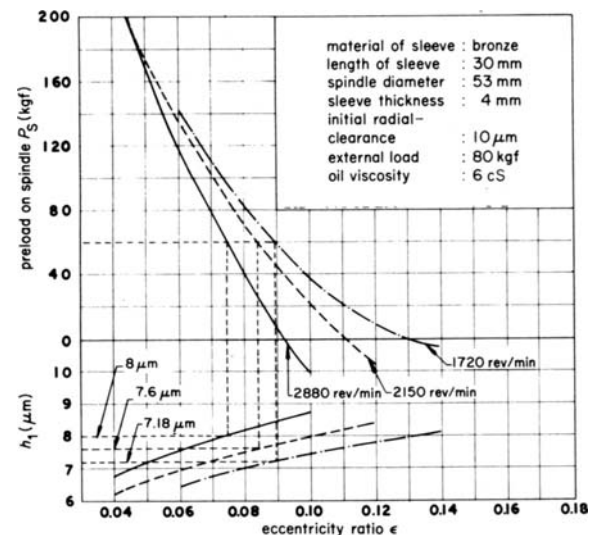


Figure 11. Effect of spindle speed on h_1 under condition 3, spindle running-external load 80 kgf.

5. CONCLUSIONS

The theoretical analysis presented here establishes the importance of three parameters, namely, initial radial clearance, sleeve stiffness and spindle preload, on the performance of clearanceless bearing. A smaller initial radial clearance, higher stiffness of sleeve and higher preload on the spindle are helpful in improving the oil film stiffness of the bearing; however, their values are limited by practical considerations like temperature rise in the bearing and the bearing life. Therefore a compromise should be made while designing this bearing according to the overall performance requirements demanded by the design. A special feature of this bearing to self-adjust its working clearance according to the spindle speed is clearly borne out by the analysis.

ACKNOWLEDGMENTS

This work was carried out in the Department of Research and Testing, Central Machine Tool Institute, Bangalore, India. The authors wish to thank Mr R. K. Gejji, Director, Central Machine Tool Institute, for granting facilities for this work.

APPENDIX A DEFORMATION SHAPE OF A RING UNDER THREE EQUAL LOADS ACTING AT THREE EQUISPACED RADIAL POSITIONS ON ITS PERIPHERY

According to Prescott⁷, the deflection $\delta(\theta)$ at an angle θ from the load line in the case of a ring loaded

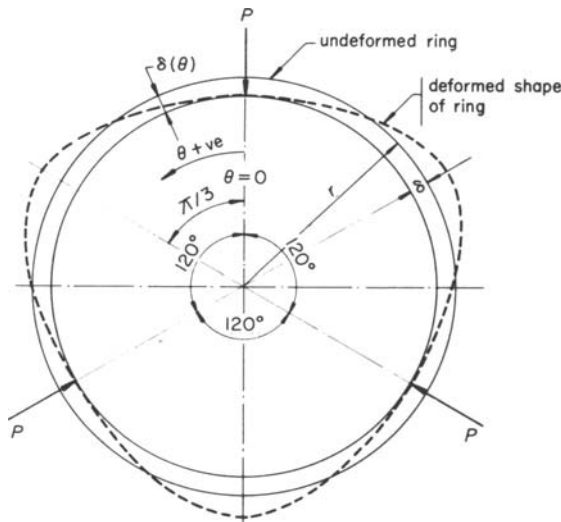


Figure A1. Deformed shape of a ring under three equispaced forces of the same magnitude.

by three equal loads (P) acting at three equispaced radial positions as shown in figure A 1 is given by

$$\delta(\theta) = \frac{Pr^3}{4EI} [1.9098 - 1.1547(\pi/3 \mp \theta) \sin(\pi/3 \mp \theta) - 1.852833 \cos(\pi/3 \mp \theta)] \quad (A1)$$

for $-\pi/3 \leq \theta \leq \pi/3$, where the $-$ sign is for positive

values of θ and the $+$ sign is for negative values of θ .

$$\delta(\theta) = \frac{Pr^3}{4EI} C(\theta)$$

where

$$C(\theta) = \left[1.9098 - 1.1547 \left(\frac{\pi}{3} \mp \theta \right) \sin \left(\frac{\pi}{3} \mp \theta \right) - 1.852833 \cos \left(\frac{\pi}{3} \mp \theta \right) \right]^3 \quad (A2)$$

The deflection under the load line is given by

$$\delta(0) = \frac{Pr^3}{4EI} C(0) \quad (A3)$$

where $\theta = 0$ in the expression for $C(\theta)$

Let there be a change in P to nP causing a change in deflection by $\Delta\delta(0)$. Therefore the total deflection under the new load nP is

$$\delta(0) + \Delta\delta(0) = \frac{nPr^3}{4EI} C(0) \quad (A4)$$

therefore

$$\Delta\delta(0) = \frac{(n-1)Pr^3}{4EI} C(0) \quad (A5)$$

Also,

$$\delta(\theta) + \Delta\delta(\theta) = \frac{nPr^3}{4EI} C(\theta)$$

$$\Delta\delta(\theta) = \frac{(n-1)Pr^3}{4EI} C(\theta) \quad (A6)$$

Therefore

$$\frac{\Delta\delta(\theta)}{\Delta\delta(0)} = \frac{C(\theta)}{C(0)} = K(\theta) \quad (A7)$$

For $\theta = 60^\circ$ (in case of clearanceless bearing)

$$C(60) = (1.9098 - 1.1547 \times 0 \times 0 - 1.8523) = 0.0576 \quad (A8)$$

and

$$C(0) = \left(1.9098 - 1.1547 \times \frac{\pi}{3} \times 0.866 - 1.8523 \times 0.5 \right) = -0.06225 \quad (A9)$$

Therefore

$$K(60) = \frac{C(60)}{C(0)} = -0.9 \quad (A10)$$

The ratio of change in deflection at $\theta = 60^\circ$ and at $\theta = 0^\circ$ is given by

$$\frac{\Delta\delta(60)}{\Delta\delta(0)} = \frac{C(60)}{C(0)} = K(60) = -0.9 \quad (A11)$$

APPENDIX B EQUATION OF HYDRODYNAMIC LOAD – CONVERGING WEDGE

The Reynolds' equation for very short bearings³ is given by

$$p = 3U\eta \frac{dh/dx}{h^3} \left(y^2 - \frac{L^2}{4} \right)$$

where p is the hydrodynamic pressure at the point (x, y)

U is the relative velocity between the two sliding surfaces in the x -direction

η is the viscosity of oil

h is the oil film thickness at a section distant x from the origin,

L is the length of the bearing.

Converting into polar co-ordinates,

$$x = r\theta \quad dx = r d\theta$$

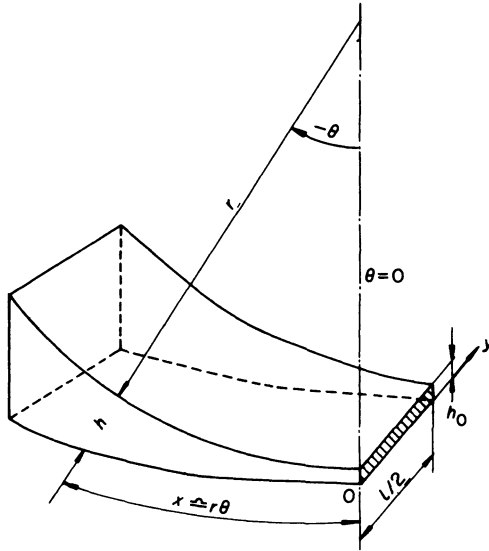


Figure B1. Exponential wedge shape.

Assuming the oil wedge shape (figure B1) to follow an exponential relationship of the form

$$h = h_0 e^{-\alpha\theta}$$

$$\frac{dh}{d\theta} = -h_0\alpha e^{-\alpha\theta}$$

Therefore,

$$p = -3U\eta \frac{(h_0/r)\alpha e^{-\alpha\theta}}{h_0^3 e^{-3\alpha\theta}} \left(y^2 - \frac{L^2}{4} \right)$$

$$= \frac{3U\eta\alpha}{h_0^2 r} e^{2\alpha\theta} \left(\frac{L^2}{4} - y^2 \right)$$

The component of the resultant hydrodynamic load in the x -direction is given by

$$W_x = \int_{-\theta}^0 \int_{-L/2}^{L/2} pr d\theta dy \cos \theta$$

$$= \iint \frac{3U\eta\alpha}{h_0^2} e^{2\alpha\theta} \cos \theta d\theta \left(\frac{L^2}{4} - y^2 \right) dy$$

Solving these integrals separately,

$$\int_{-L/2}^{L/2} \left(\frac{L^2}{4} - y^2 \right) dy = 2 \int_0^{L/2} \left(\frac{L^2}{4} - y^2 \right) dy$$

$$= 2 \left[\frac{L^2 y}{4} - \frac{y^3}{3} \right]_0^{L/2}$$

$$= \frac{L^3}{6}$$

$$\int_{-\theta}^0 e^{2\alpha\theta} \cos \theta d\theta$$

$$= \left[\frac{e^{2\alpha\theta}}{(1 + 4\alpha^2)} (\sin \theta + 2\alpha \cos \theta) \right]_{-\theta}^0$$

$$= \left\{ \frac{1}{(1 + 4\alpha^2)} [2\alpha + e^{-2\alpha\theta} (\sin \theta - 2\alpha \cos \theta)] \right\}$$

Therefore,

$$W_x = \frac{3U\eta\alpha}{h_0^2} \cdot \frac{L^3}{6} \cdot \frac{1}{(1 + 4\alpha^2)}$$

$$\times [2\alpha + e^{-2\alpha\theta} (\sin \theta - 2\alpha \cos \theta)]$$

$$= \frac{U\eta L^3}{2h_0^2} \cdot \frac{\alpha}{(1 + 4\alpha^2)}$$

$$\times [2\alpha + e^{-2\alpha\theta} (\sin \theta - 2\alpha \cos \theta)]$$

In our case $\theta = \pi/3$. Therefore

$$W_x = \frac{U\eta L^3}{2h_0^2} \cdot \frac{\alpha}{(1 + 4\alpha^2)}$$

$$\times \left[2\alpha + e^{-(2\pi\alpha/3)} \left(\sin \frac{\pi}{3} - 2\alpha \cos \frac{\pi}{3} \right) \right]$$

$$= \frac{U\eta L^3}{2h_0^2} \cdot \frac{\alpha}{(1 + 4\alpha^2)} \left[2\alpha + e^{-(2\pi\alpha/3)} \left(\frac{\sqrt{3} - 2\alpha}{2} \right) \right]$$

$$= C_0 \frac{U\eta L^3}{h_0^2}$$

where

$$C_0 = \frac{\alpha}{2(1 + 4\alpha^2)} \left[2\alpha + e^{-(2\pi\alpha/3)} \left(\frac{\sqrt{3} - 2\alpha}{2} \right) \right]$$

In a particular case if U, η and L are constant,

$$W_x = \frac{K_0}{h_0^2}$$

where

$$K_0 = \frac{U\eta L^3}{2} \cdot \frac{\alpha}{(1 + 4\alpha^2)} \left[2\alpha + e^{-(2\pi\alpha/3)} \left(\frac{\sqrt{3} - 2\alpha}{2} \right) \right]$$

REFERENCES

1. S. Z. Abdin and M. K. Bhat. *Testing of clearanceless bearings*, CMTI Test Report no. T73, 1968.

2. J. Linda and F. Roseberg. Hydrodynamic bearings for grinding spindles. *Machinery and Production Eng.* (1972), 16 Feb., p. 226.
3. A. Cameron. *Basic Lubrication Theory*. Longman, 1971.
4. A. Sommerfeld. *Zur Hydrodynamischen Theorie der Schmiermittelreibung*. Oswald Klassiker no. 218, 1904.
5. O. Reynolds. *Über die Theorie der Schmierung*, Oswald Klassiker no. 218, 1904.
6. T. E. Carl. An experimental investigation of a cylindrical journal bearing under constant and sinusoidal loading. *Proc. Inst. of Mech. Engrs.* Vol. 178 Pt. 3N (1963-4).
7. J. Prescott. *Applied Elasticity*, Dover Publications Inc. New York.

DISCUSSION

In the authors' opinion, out of the three types of bearing—hydrodynamic, rolling and hydrostatic—the first one seems to be most suited for precision grinding machines. It is the superior inherent ability of the hydrodynamic bearing to damp out the vibrations and the economic considerations that make it more acceptable for such an application. The disadvantage that a high degree of skill and craftsmanship is required to manufacture hydrodynamic

bearings has been eliminated in the bearing discussed in this paper. In the suggested design, where a large initial clearance (δ) is maintained and a heavy preload (which can be removed for a short period at starting and stopping of the spindle) is applied to the spindle, normally achievable production tolerance on the sleeve bore will be quite adequate. For best results, it is advisable to orientate the bearing so that the direction of load lies in between the two supports.

CHARACTERISTICS OF ELECTRO-HYDRAULIC FEED DRIVES UNDER DYNAMIC LOAD CONDITIONS

by

D. AVRAMOVIC* and P. FORT*

SUMMARY

Differences were found in the behaviour of two concepts of feed drives when fitted to and tested on the same machine. A comparative study has therefore been made, and the necessary adaptive steps for the improvement of the drives will be discussed with the purpose of comparing the dynamic performances of both feed drives under dynamic disturbing torque loads (interrupted cutting forces), and of achieving the best possible behaviour of the feed drive under given conditions.

For more comprehensive comparison of both concepts the systems have been simplified in the first part of this work.

Finally both concepts were applied to an existing lathe and the differences in dynamic performance were shown by computer simulation.

NOTATION

To assure continuity the symbols used in previous papers have been retained as follows:

$\theta_A(s)$	angle of motor shaft
$\dot{\alpha}_A(s)$	motor shaft velocity
$\dot{\alpha}_E(s)$	rotary servovalve velocity
$\epsilon(s)$	valve displacement
B	compressibility of oil
J	inertia
i	electric current
K_T	tacho-generator gain
K_E	amplifier gain
K_X	servovalve gain
K_Q	valve flow gain
K_C	total flow pressure coefficient
$M_L(s)$	distributing torque load on the motor
p_S	hydraulic supply pressure
p_L	load pressure
R	factor of viscous friction
$S = i\omega$	laplace operator
U	voltage
V'_M	volumetric displacement of motor
V_T	total volume of fluid
ω_H	hydraulic undamped natural frequency
D_H	hydraulic damping ratio

INTRODUCTION

Demand for greater accuracy of NC and copying machines has led to the search for feed systems that will be free from speed fluctuations even under varying load, thus causing better surface finish and longer tool life. In some cases the feed drive operates under heavy dynamic cutting forces.

In order to assure a satisfactory performance of the feed drive the output quantity must be measured and a loop must be formed through feedback.

For this study, velocity-controlled feed drives were investigated. The basic concepts of control systems are the following:

- concept 1—classic velocity control loop,
- concept 2—follow-up control loop.

The purpose of this paper is to show the difference of principle between the two control systems and to discuss their respective advantages and drawbacks. Accordingly the non-linear aspects in the first part have not been taken into consideration.

The load given by a torque on the motor shaft results from dynamic cutting force. This torque, assumed for the simulation, was taken from actual measurements on the machine used for the test and can be approximated to a function of a frequency and of amplitude.

The results obtained are in close accordance with practical experience.

DESCRIPTION OF THE FEED DRIVES

Concept 1—Electrohydraulic velocity control loop

Figure 1(a) shows the drive system which operates as follows: An electrohydraulic servovalve controls a hydraulic motor, the shaft rotation of which is transposed by a screw into the feed motion of the slide.

The tacho-generator senses the velocity and feeds the measured voltage back to the amplifier, where it is compared with the control input. The error is transformed into a displacement of the servovalve.

* Institut für Werkzeugmaschinenbau und Fertigungstechnik Eidg. Technische Hochschule, Zürich

Figure 1(b) shows the axial piston hydromotor coupled to an electrohydraulic servovalve. Figure 1(c) shows a schematic block diagram of the velocity control loop.

follows: a rotary built-in servovalve controls a hydromotor of the same size as in concept 1. The rotary motion of the servovalve is generated by a velocity-controlled d.c. motor.

The controlling orifices of the valve sensor are disposed in such a way that the flow distribution varies according to the motion of the rotary valve in

Concept 2—Electrohydraulic follow-up control loop
 Figure 2(a) shows the drive system which operates as

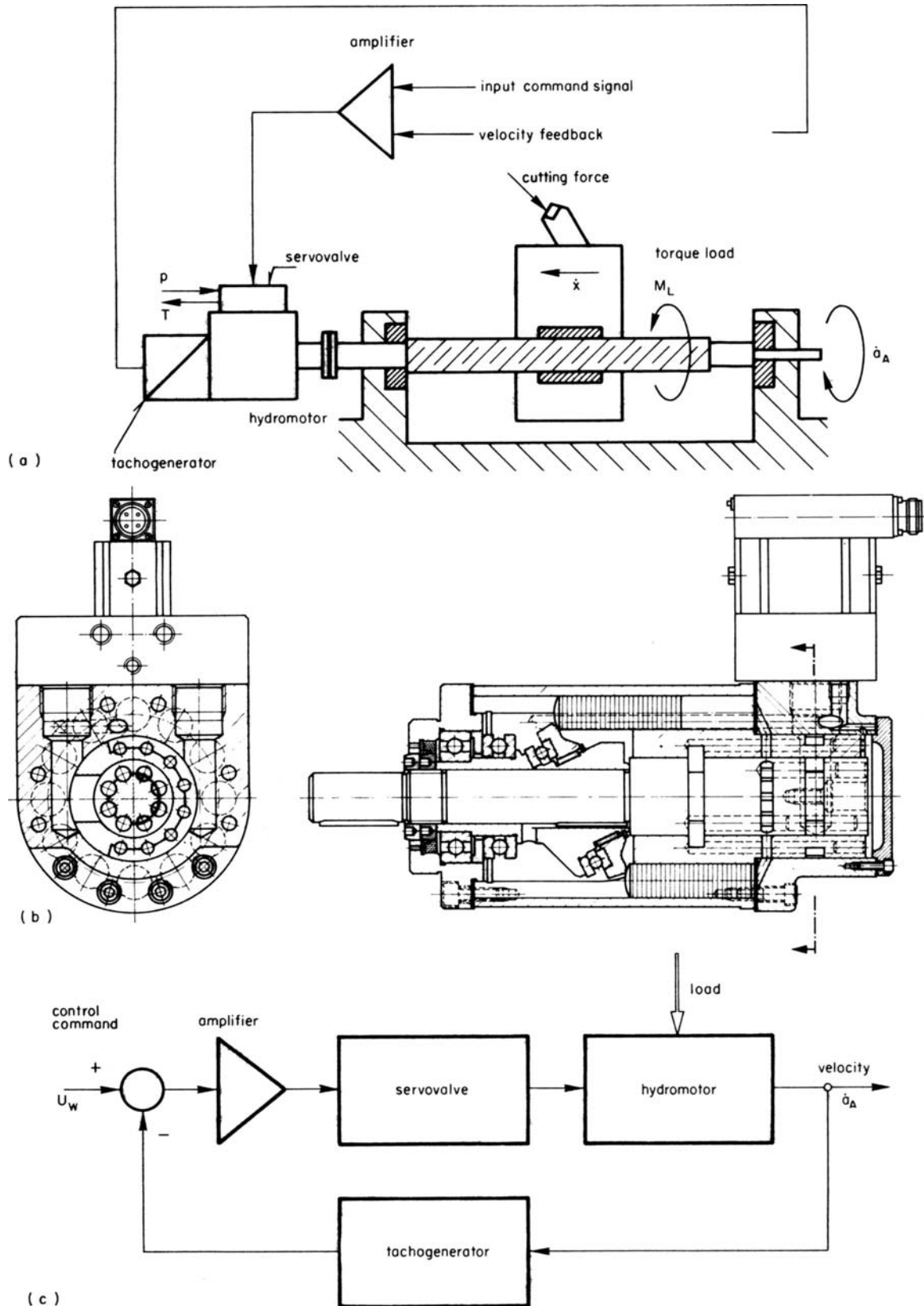


Figure 1. Servo drive concept 1. (a) Feed-drive diagram. (b) Axial piston-type hydromotor and electrohydraulic servovalve. (c) Block diagram of the velocity control loop.

order to revert to the original stable position in relation to the sensor.

The hydraulic torque amplifier³ operates as follow-up control loop (figure 2(b)).

The block diagram (figure 2(c)) shows two control loops of this system: the velocity control loop of the electrical part and the position feedback of the hydromotor.

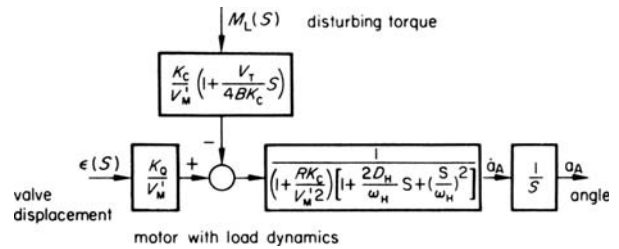


Figure 3. Open loop block diagram.

LINEAR MODELS OF CONTROL SYSTEMS

Linear models⁵ of control systems have to be set up in order to transmit the transfer function from the disturbing torque to the velocity error. The linearity

range near the operating point has been confirmed by non-linear computer simulation.

Assuming that both concepts possess the same characteristic parameters (displacement, servovalve

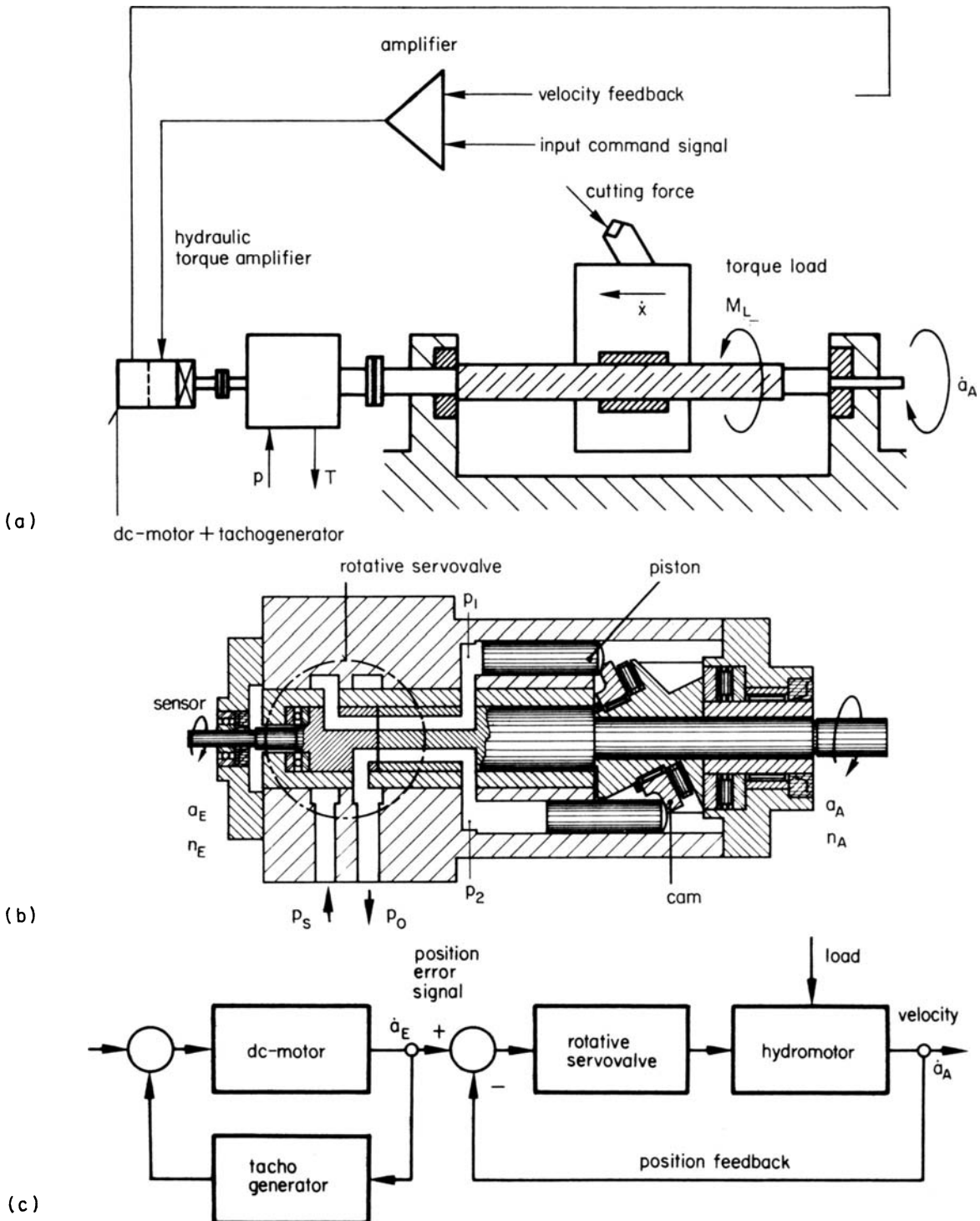


Figure 2. Servo drive concept 2. (a) Feed drive diagram. (b) Hydraulic torque amplifier. (c) Block diagram of the followup loop.

characteristics, leakage, etc.) the general response equation for the open loop³ defines the relation between the valve displacement ϵ , the disturbing torque M_L and the motor shaft position α_A

$$\alpha_A(s) = \frac{\frac{K_Q}{V'_M} \epsilon(s) - \frac{K_c}{V'_M} \left(1 + \frac{V_T}{4BK_c} \cdot S\right) M_L(s)}{\left(1 + \frac{K_c R}{V'_M{}^2}\right) \left[1 + \frac{2D_H}{\omega_H} S + \left(\frac{S}{\omega_H}\right)^2\right] \cdot S} \quad (1)$$

where

$$\omega_H = \sqrt{\left(\frac{4BV'_M{}^2}{JV_T}\right)} \cdot \sqrt{\left(1 + \frac{RK_c}{V'_M{}^2}\right)}$$

$$D_H = \frac{\frac{K_c}{V'_M} \sqrt{\left(\frac{BJ}{V_T}\right)} + \frac{R}{4V'_M} \sqrt{\left(\frac{V_T}{BJ}\right)}}{\sqrt{\left(1 + \frac{RK_c}{V'_M{}^2}\right)}}$$

Equation (1) illustrated in figure 3 is common to both concepts. The basic difference resides in the type of feedback.

Block diagram related to disturbing torque

The break frequencies of the servovalve and of the electrical equipment are usually higher than the natural hydraulic frequency. Therefore only the transfer function of the hydraulic part of the loop can be taken into consideration. It is assumed that Equation (1) is applicable to the torque load on the motor and that the block diagram of the velocity control loop becomes the one shown in figure 4.

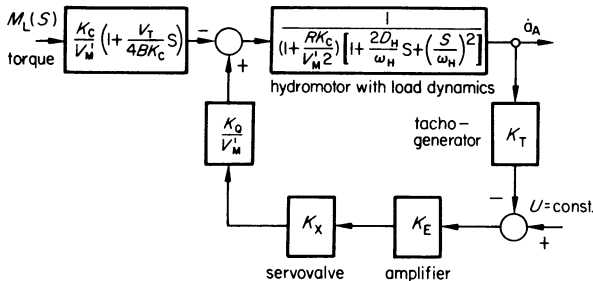


Figure 4. Block diagram of the control system concept 1.

Let us now determine the transfer from load M_L to velocity error $\Delta \dot{\alpha}_A$ (closed loop)

$$M_I(s) = \frac{\Delta \dot{\alpha}_A(s)}{\Delta M_L(s)} = \frac{-\frac{K_c}{V'_M} \left(1 + \frac{V_T}{4BK_c} \cdot S\right) \left[\frac{1}{1 + (RK_c/V'_M{}^2)}\right]}{K_{VI} + \left[1 + \frac{2D_H}{\omega_H} S + \left(\frac{S}{\omega_H}\right)^2\right]} \quad (2)$$

where

$$K_{VI} = K_T K_E K_X \cdot \frac{K_Q}{V'_M} \left[\frac{1}{1 + (RK_c/V'_M{}^2)}\right]$$

Setting the values of

$$F_H = \frac{1}{\left[1 + (2D_H/\omega_H)S + (S/\omega_H)^2\right]}$$

$$F_M = \frac{K_c}{V'_M} \left(1 + \frac{V_T}{4BK_c} \cdot S\right) \left[\frac{1}{1 + (RK_c/V'_M{}^2)}\right]$$

Equation (2) can be simplified to give

$$M_I(s) = \frac{-F_M}{K_{VI} + 1/F_H} \quad (3)$$

Concept 2—block diagram related to disturbing torque

The transfer from load torque M_L to velocity error $\Delta \dot{\alpha}_A$ can be computed from the block diagram (figure 5) of the following control loop.

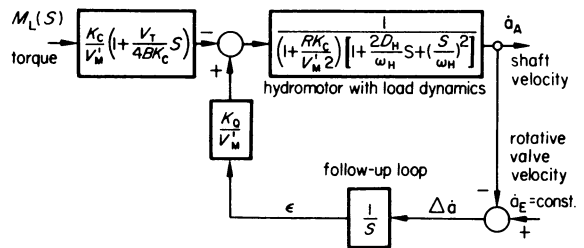


Figure 5. Block diagram of the control system concept 2.

The input command is given by the constant rotation α_E of the sensor. This sensor is so designed that the load torque cannot have any influence on the input command. Therefore the dynamic properties of the d.c. motor may be disregarded here.

The transfer function of the systems is valid in the linearity range near the operating point and it can be used as a common model in system design

$$M_{II}(s) = \frac{\Delta \dot{\alpha}_A(s)}{\Delta M_L(s)} = \frac{-\frac{K_c}{V'_M} \left(1 + \frac{V_T}{4BK_c} \cdot S\right) \left(\frac{1}{1 + (RK_c/V'_M{}^2)}\right) \cdot S}{K_{VII} + \left(1 + \frac{2D_H}{\omega_H} S + \left(\frac{S}{\omega_H}\right)^2\right)} \quad (4)$$

where the velocity gain is expressed as

$$K_{VII} = \frac{K_Q}{V'_M} \cdot \frac{1}{1 + (RK_c/V'_M{}^2)}$$

The final form of the transfer function (4) can be written as

$$M_{II}(s) = \frac{-F_M S}{K_{VII} + S/F_H} \quad (5)$$

Comparison between disturbing frequency responses
Equations (3) and (5) could be examined with a view to comparing the dynamic performance of both the simplified concepts.

The two transfer functions are different in the

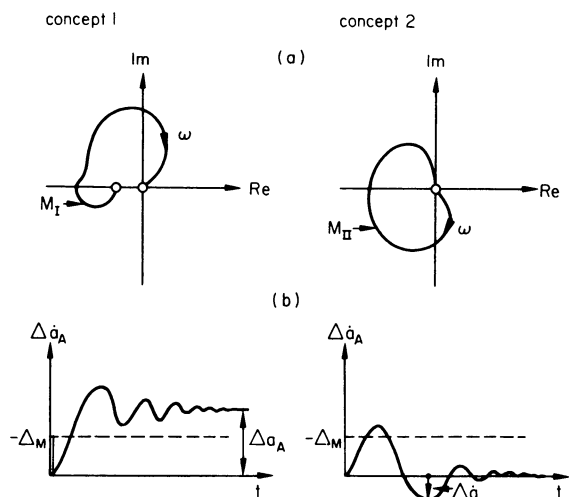


Figure 6. (a) Polar Nyquist diagram of transfer function $M(s)$. (b) Diagram of transient time response to a unit step.

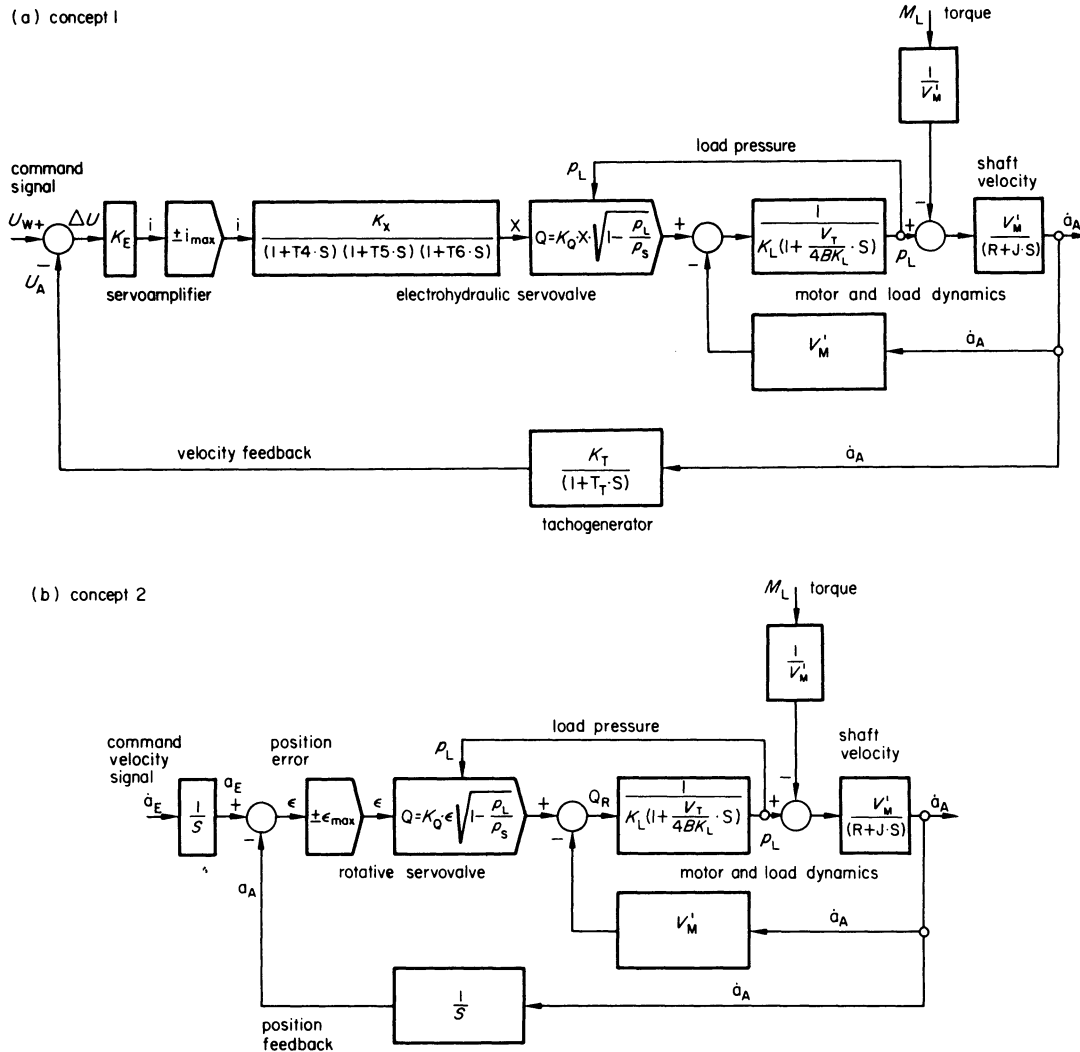


Figure 7. Block diagrams of the feed drives with actual dynamics and non-linear servovalve characteristics used for digital simulation.

degree of the denominator and the numerator. After algebraic simplification the result is

$$M_I(s) = \frac{-K_1(1 + T_1S)}{1 + A_1S + A_2S^2} \quad (6)$$

$$M_{II}(s) = \frac{-SK_2(1 + T_1S)}{1 + B_1S + B_2S^2 + B_3S^3} \quad (7)$$

Both transfer functions can be represented as a Nyquist diagram (figure 6a).

In the steady state concept 1 has a sustained deviation. This velocity error is a function of the

gain whereas the damping time depends on the position of the roots.

The step responses of the two systems are shown in figure 6(b).

TEST ON A LATHE

Both feed drives were fitted to a lathe and tested under variable load conditions. (Interrupted cutting force.)

The total inertia torque was $J: 4 \cdot 10^{-5} \text{ kgm}^2$ and the lead screw pitch $1 \cdot 10^{-2} \text{ m}$.

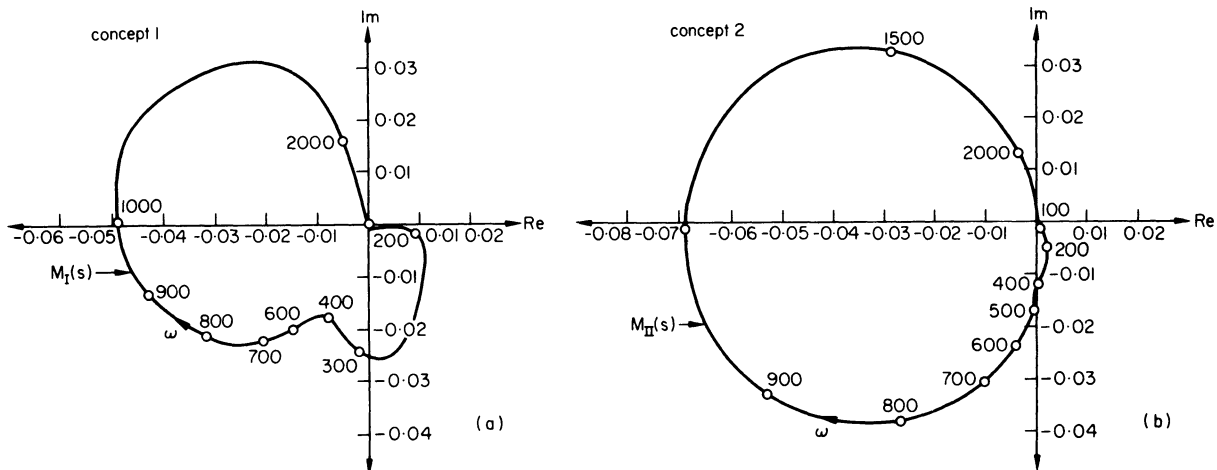
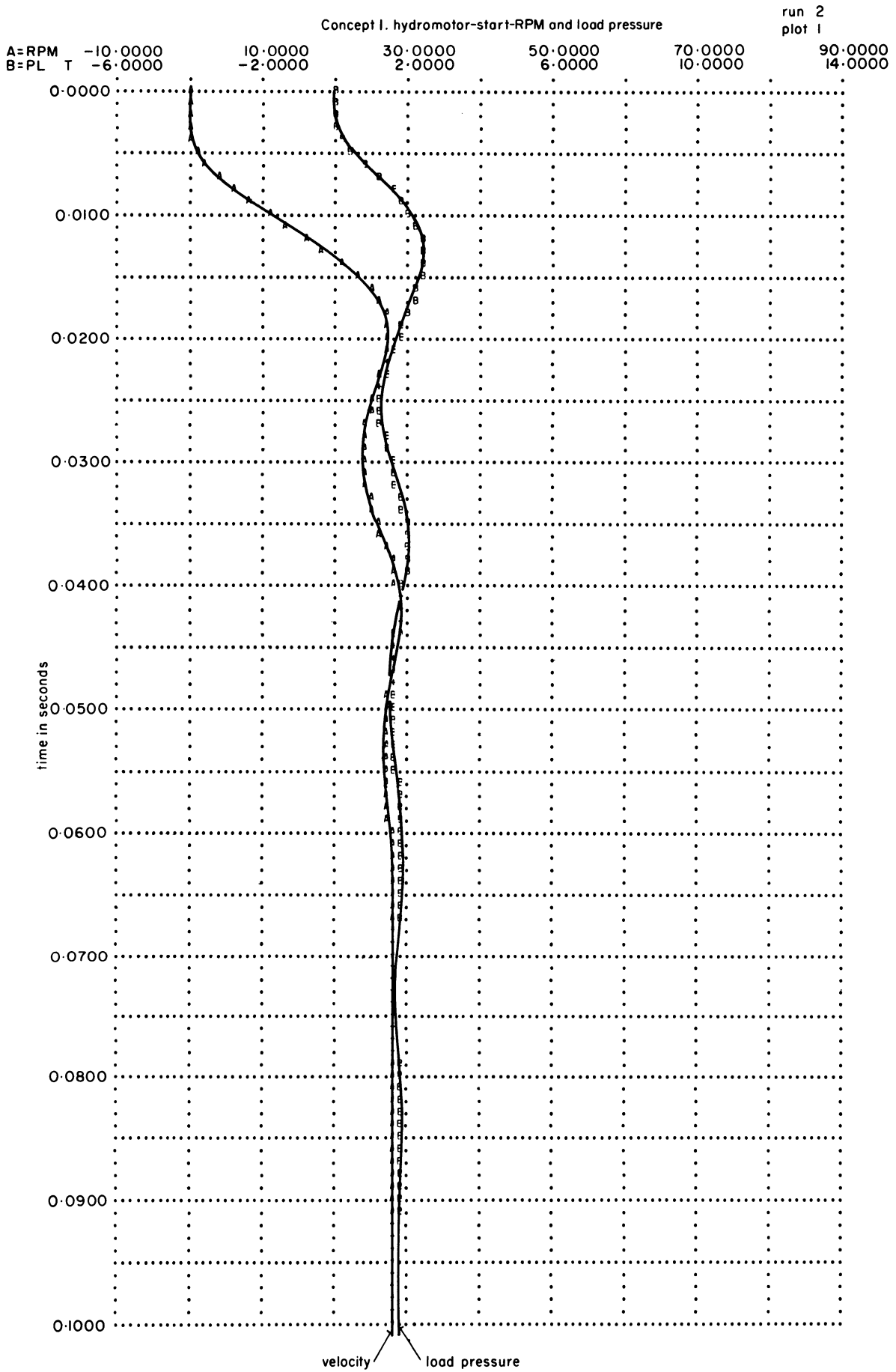
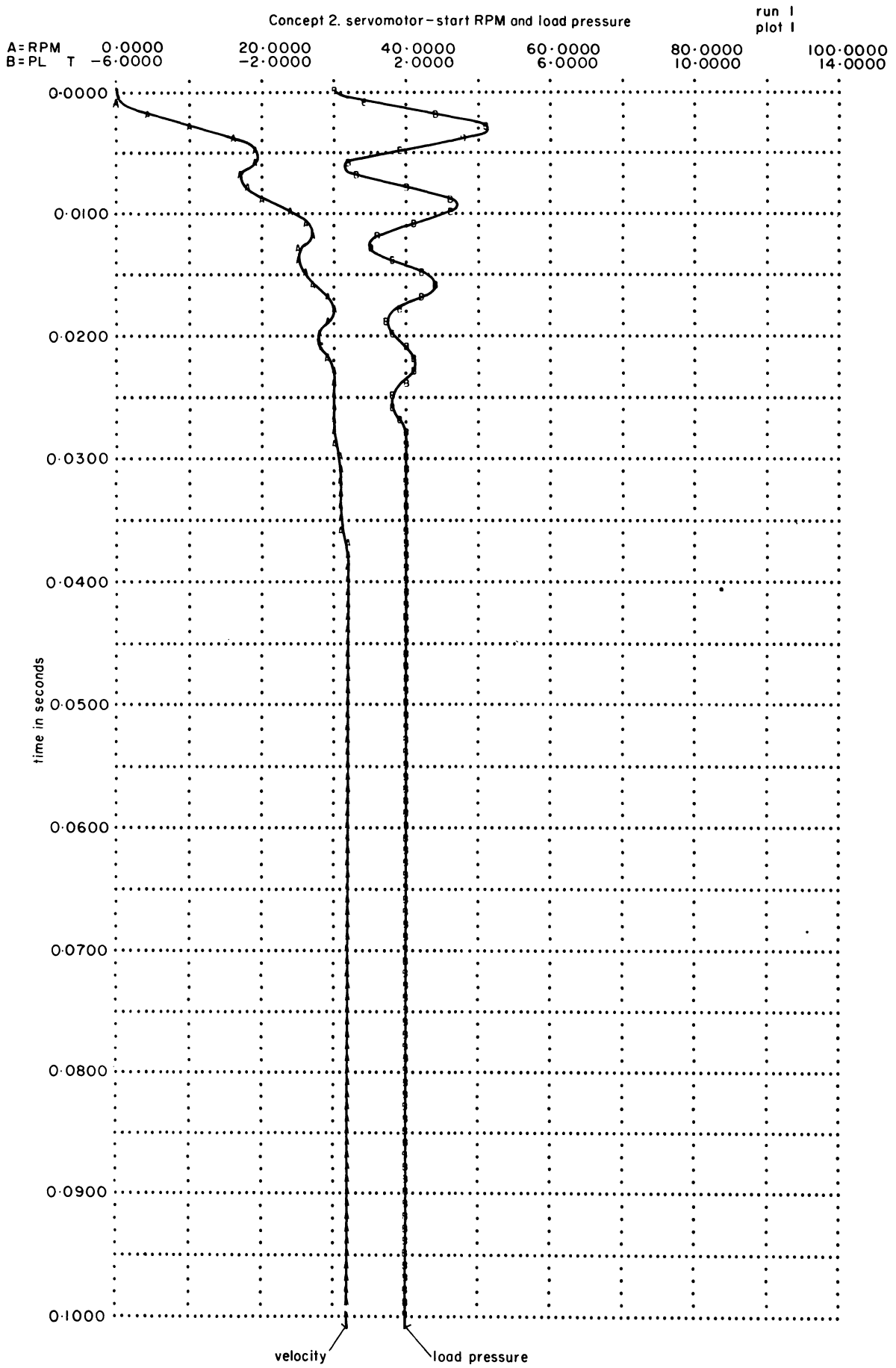


Figure 8. Polar diagram of transfer functions $M(s)$ related to the actual dynamics linearized by 32 rev/min.

170 CHARACTERISTICS OF ELECTRO-HYDRAULIC FEED DRIVES UNDER DYNAMIC LOAD CONDITIONS

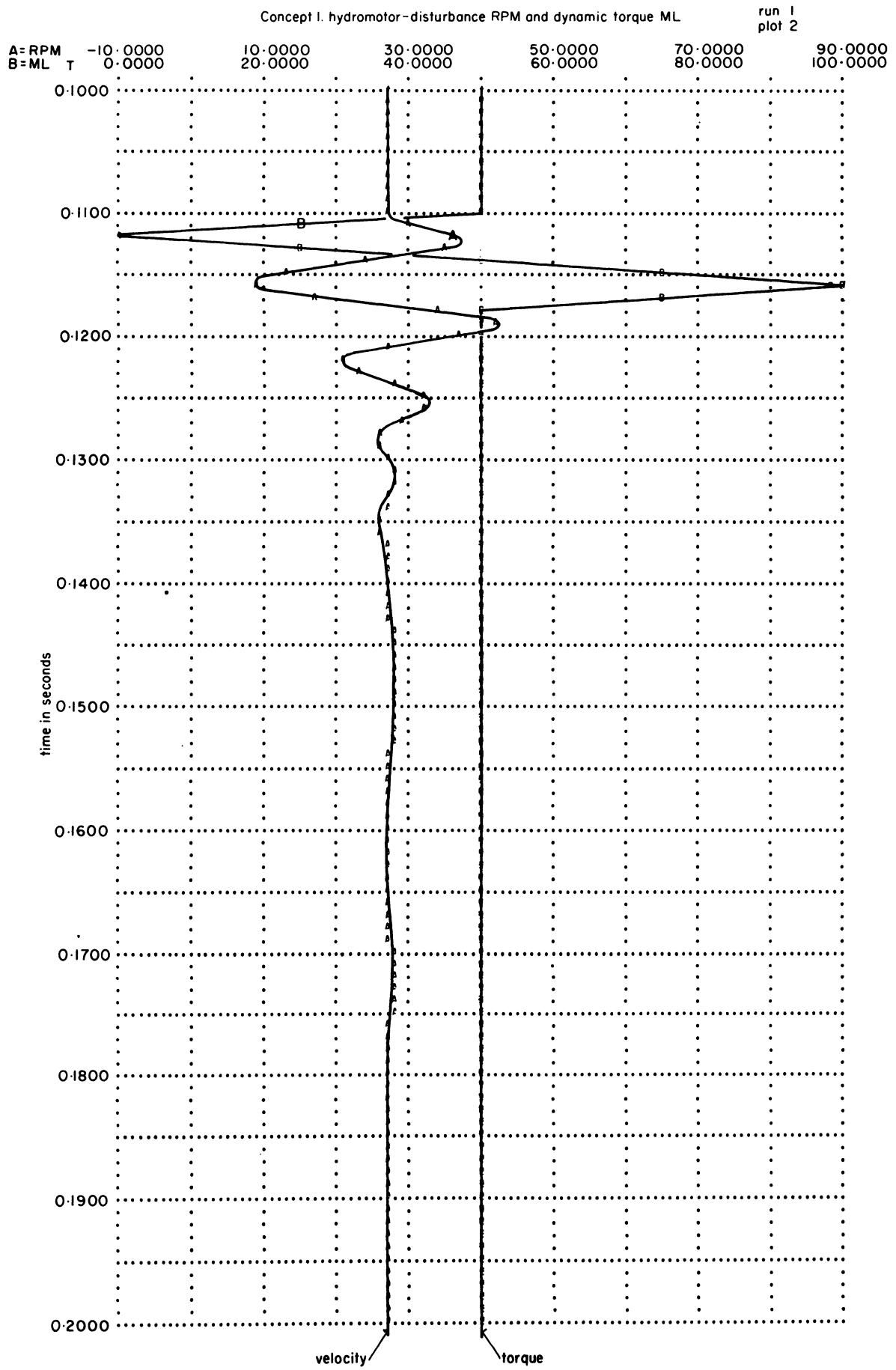


(a)

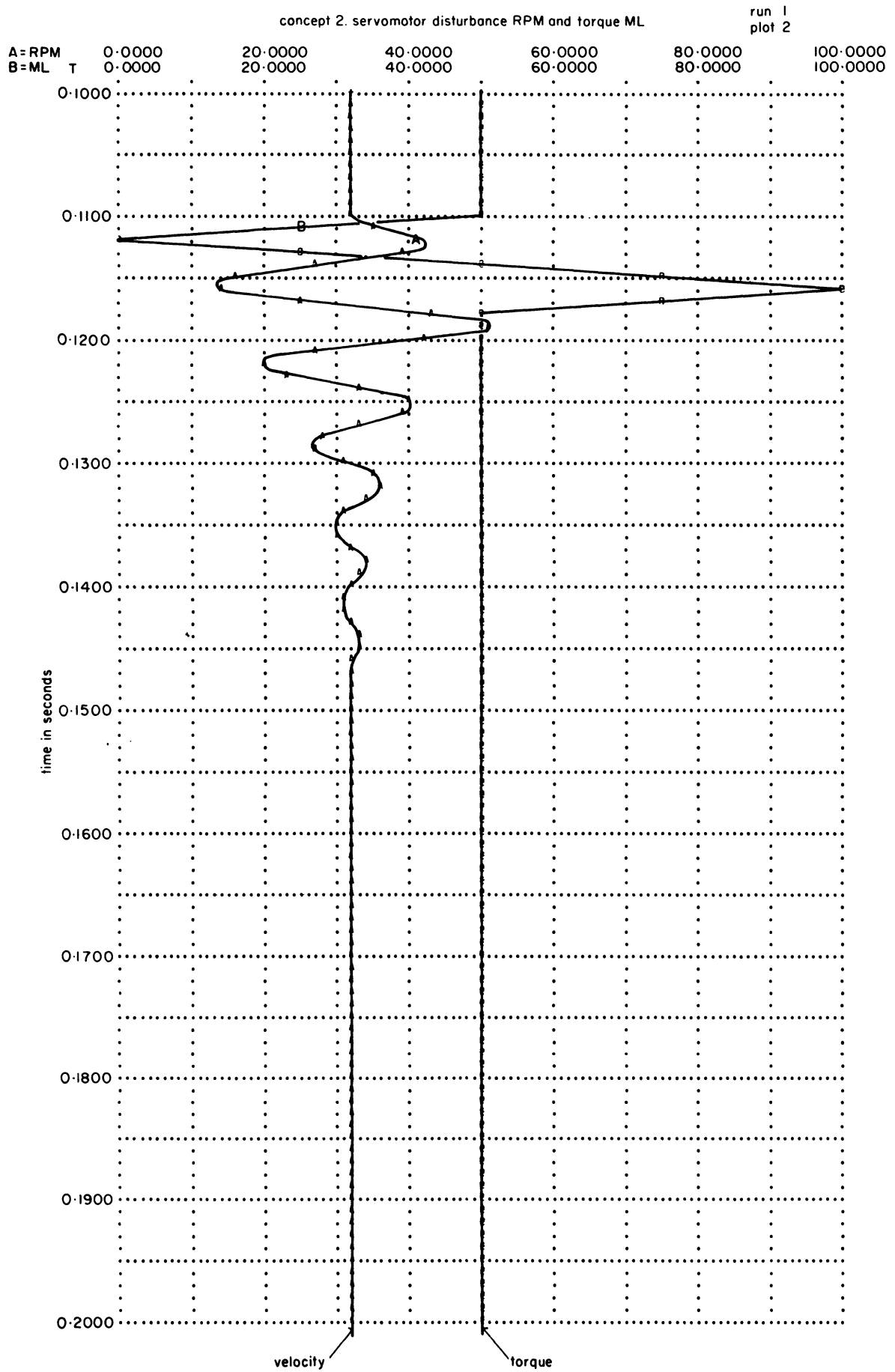


(b)

172 CHARACTERISTICS OF ELECTRO-HYDRAULIC FEED DRIVES UNDER DYNAMIC LOAD CONDITIONS



(c)



(d)

Figure 9. Simulated time response plots. (a), (b) Starting-up process. (c), (d) Disturbance response.

The measured results were compared with the computer simulations; they corresponded quite well. In the computer study⁵, the non-linear characteristics of the valve and the actual dynamics of the system were added to the simplified models. For this investigation a zero underlap valve was used. The flow Q can be described as a function of valve displacement ϵ and pressure P_L ¹¹.

$$Q = K_Q \cdot |\epsilon| \cdot \sqrt{\left| \left[\text{sign}(\epsilon) - \frac{P_L}{P_s} \right] \right|^{\frac{1}{2}} \cdot \text{sign} \left[\text{sign}(\epsilon) - \frac{P_L}{P_s} \right]} \quad (8)$$

The operator 'sign' represents the algebraic sign of the argument. The non-linear part of the friction and the elasticity of the structure were neglected in this first preliminary study. The plots given in figure 9 show some typical results of both concepts.

None of these concepts has a structural stability. Due to the non-linearity of the systems, the gain is a function of the amplitude as well.

The gain, which determines the stability of the systems, rises in the case of smaller servovalve strokes and of larger load pressures.

The minimum velocity at maximum torque load represents the limits of stability. A sustained velocity deviation will always be present when servo drive concept 1 is fitted with a proportional controller. The open loop gain being limited by stability conditions, a proportional-integral controller must be used in order to improve the accuracy. Electrical networks can be used to facilitate this improvement. However, setting the controller requires some experience, and a few trials will always be necessary.

The use of a PID controller will improve the dynamics somewhat, but arriving at an optimum setting will be more difficult because of the pronounced non-linearity. The feedback of concept 2 consists of an integral action (from velocity to position). The follow-up loop used for velocity control is equivalent to an integral controller. Accordingly there is no velocity error in a steady-state condition.

This system does not allow a setting of the controller. The stability must be ensured through a sufficient damping of the feed drive. Such a robust drive will require little maintenance but it cannot be set to optimum values.

CONCLUSIONS AND FUTURE STEPS

The most important difference between the two concepts lays in the type of feedback.

Concept 1 fitted with PI controller will give the most satisfactory results. Fine setting of the controller will provide the best control function, but it will always require some trials on the machine for

which, on account of non-linearity, some knowledge of hydraulic control systems is necessary.

Concept 2. Follow-up control should always be used when the machine possesses sufficient damping, when optimum control function is not a strict condition, and when a simple and robust drive is required. A torque amplifier with a built-in velocity gain is to be preferred wherever there is little experience of control technology available either in the workshop or in the engineering office.

For a better understanding of the problem both systems were designed as simply as possible.

Future work will therefore be to consider the actual operation of such types of feed drives. The results of the digital simulation show that the effects of non-linearities in the system must be fully considered in the investigation.

REFERENCES

1. W. Oppelt. *Kleines Handbuch technischer Regelvorgänge*, Verlag chemie, Weinheim, 1964.
2. H. E. Merritt. *Hydraulic Control System*, John Wiley and Sons, New York, 1967.
3. J. Thoma and D. Avramovic. Zwei Verfahren für die Untersuchung von hydraulischen Nachlaufsystemen, *Fertigung 5*, 1973.
4. D. Avramovic. Der Einfluss elektrohydraulischer Servoantriebe auf die Konstruktion im Maschinenbau, *oelhydraulik + Pneumatik* (1968) 12.
5. P. Fort. Störverhalten der Vorschubantriebe. Diplomarbeit am Institut für Werkzeugmaschinenbau und Fertigungstechnik ETHZ, Zürich, 1973.
6. H. Beer and L. Audolensky. *Messungen an den Kopier- und Vorschubsystemen der Drehmaschinen Dubied Typ 522 und 517*, Versuchsbericht SIG, Neuhausen, 1972.
7. D. C. Clark. Auswahl und Leistungskriterien von elektrohydraulischen Stellantrieben, *die Maschine* (1970) 2.
8. F. Cellier. Vorschubregelung für Errosionsmaschinen, Diplomarbeit am Lehrstuhl für Errosionmaschinen, Diplomarbeit am Lehrstuhl für Automatik der ETHZ Zürich, 1972.
9. FIDES: MIMIC Program Reference Manual, for CDC 6.500.
10. A. Schaffer. *Vorschubsystem des Hauptschlittens der Kopierdrehmaschine*, Versuchsbericht Dubied, Marin, 1973.
11. J. J. Schacht and J. G. Bollinger. Analog computer simulation techniques for analysis of a machine tool hydraulic drive, University of Wisconsin, 1972.
12. W. Jorden. Untersuchungen an einem Lageregelkreis für Werkzeugmaschinen unter besonderer Berücksichtigung des Spieles, *Fortschritt Bericht* (1969) VDI-Z, 2/20.
13. G. Stute and D. Schmid. Optimization of position control loops for NC Machine tools, *Proc. 14th Int. M.T.D.R. Conf.* 1973.

COMPUTER-AIDED DESIGN OF STANDARDIZED HYDROSTATIC JOURNAL BEARINGS

by

J. M. SANSINENEA* and R. M. BUENO†

SUMMARY

This paper contains the results obtained with the application of two design computer programs to a sample of twenty design proposals of hydrostatic journal bearings. The first computer program has been developed with the aim of achieving a design when the design variables—diameter, axial length, etc.—are only allowed to take some discrete values. The second program is intended for designing bearings with design variables of a continuous type. The results of the utilization of geometric ISO series as standard series for hydrostatic journal bearings are also discussed.

INTRODUCTION

The optimization of hydrostatic journal bearings is of the utmost importance. Very poor results can be expected from a design if it is not based on some optimization criteria. The situation gets even worse in cases of high-speed journal bearings. The optimization of hydrostatic journal bearings has been analysed by many authors such as Stansfield¹, Rowe and O'Donoghue², Favareto³ and Bueno⁴. The differences existing between the design procedures presented by these authors consist of the different design variables and constraints considered by them.

On the other hand, the standardization of hydrostatic journal bearings has been the subject of some publications, such as that presented by Hallstedt⁵.

In the present work, two computer programs are presented. Both are intended to carry out a fully-automatic design of optimum hydrostatic journal bearings. In the first one, named OPTIM 1, same design variables are only allowed to take some discrete values, so that OPTIM 1 can be used to obtain the optimum design of a bearing when the design variables can only be assigned standardized values. In the second program, OPTIM 2, all the design variables can vary continuously between two limits.

The aim of this research work was the setting up of the above-mentioned design computer programs and their application to a set of real design examples presented by a group of machine tool manufacturing firms. At the same time, the comparison of the results obtained with the application of OPTIM 1 and OPTIM 2, has produced some conclusions concerning the standardization of hydrostatic journal bearings.

NOTATION

a	width of circumferential flow land
b	width of axial flow land
D	bearing diameter
F_{\max}	maximum load
FB	Fibonacci number
h	radial design clearance
L	bearing length
n	number of recesses
N'	Rotational speed
P_s	supply pressure
Re	Reynolds number
W_B	pumping power
W_R	friction power generated in the lands
W_{Rc}	friction power generated in the recesses
W_T	total power
β	dimensionless design pressure ratio
λ_{\min}	minimum stiffness
η	oil viscosity
ρ	mass density of oil

DESIGN COMPUTER PROGRAMS

In every problem involving optimization it is always necessary to define clearly the objective function, the design variables and the constraints to be considered. The objective function employed in the optimum-finding process must be unique and any change of objective function during the process must be ruled out. On the other hand, the values assigned to the design variables should, at any moment be changed in order to obtain the fulfilment of the objective.

The first difficulty which presents itself is the necessity of choosing an objective function from

* Asociación de Investigación Industrial de la Máquina Herramienta (INVEMA), San Sebastián, Spain

† Centro de Investigaciones Técnicas de Guipuzcoa, San Sebastián, Spain

among the several existing. It is clear that without considering objective functions involving economic aspects such as manufacturing costs, several objectives are still available. From a strictly technical point of view, three objectives are worth mentioning: minimum consumption of power, minimum rise of temperature, and maximum stiffness. If the possibility of forming an objective function by combining two or three of the single objectives already mentioned is overlooked—a possibility which in some cases can be considered—a choice must be made by singling out one objective and transforming the other possible objectives into constraints.

The minimization of the consumption of power is the objective most commonly considered. This is because of the need to reduce the amount of heat dissipated by the bearing, which otherwise would cause a substantial increase in the temperature level of the whole system into which the hydrostatic bearing is fitted. Should this level—as a consequence of an improper design—attain an intolerably high value, the use of cooling devices would become inevitable.

The temperature rise between the oil inlet and outlet of system is usually small and can be neglected. The consideration of the rise of temperature can be of importance in the case of bearings working with very high-speed spindles. This can be included in the optimization process by adding another constraint.

Although having a high operating stiffness in the bearing is always a goal, importance of designing the bearing stiffness in accordance with that of the rest of the structure must be borne in mind.

From a dynamic point of view, a too high degree of stiffness could be bad if it should cause low damping⁶. For these reasons, it seems to be clear that stiffness must be present as a constraint.

The objective function which has been considered in both optimization programs, OPTIM 1 and OPTIM 2, is the total amount of power, w_T , dissipated by the bearing. There are three single terms which make up the total power. These are: pumping power, W_B ; friction power dissipated in the bearing lands, W_R , and friction power generated in the recesses, W_{Rc} . Hence

$$W_T = W_B + W_R + W_{Rc} \quad (1)$$

Expressions of the different terms of Equation (1) appear in Appendix I.

As far as the last term of Equation (1), W_{Rc} , is concerned, the theory presented by Shinkle and Hornung⁷ has been followed because of its approach to a possible turbulent regime in the recesses, as can be found in high-speed journal bearings.

In both OPTIM 1 and OPTIM 2, seven design variables have been employed: diameter; length-diameter ratio; axial land-length ratio; circumferential land diameter ratio; viscosity; design pressure ratio; and design clearance. Figure 9 shows the geometric dimensions of the bearings. All the variables are considered to be affected by min-max constraints. The constraint represented by a minimum operating stiffness is also taken into account. The design conditions under which the bearing will have to work are the spindle speed and the maximum load to be

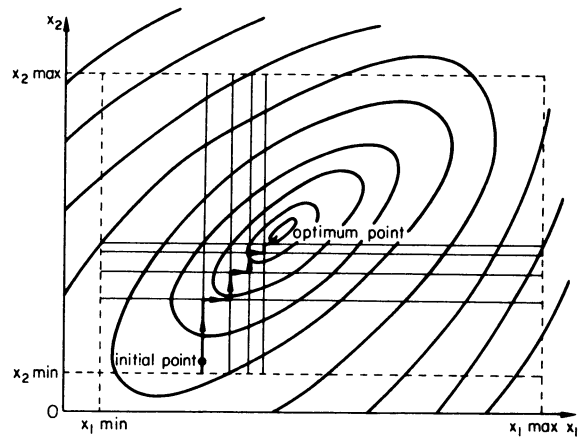


Figure 1. Alternative unidirectional search pattern.

withstood—with a maximum permitted offset of 30 per cent of the design clearance. The supply pressure is obtained either from the constraint of the maximum stiffness or from the maximum load depending upon which one of the two conditions is more severe. As optimum search method, a unidirectional alternative search algorithm along the direction of every single design variable has been used throughout this study (see figure 1).

The differences between OPTIM 1 and OPTIM 2 can be found in the different algorithms employed for the one-dimensional optimum finding along the direction of every design variable—a finding carried out with all the design variables fixed except one of them which is not fixed.

In OPTIM 1, as the design variables are of a discrete type and can only take values belonging to some standard series, the Fibonacci search technique was selected as the most appropriate technique for the special characteristics of the problem involved. Some details of the optimum seeking algorithm employed in OPTIM 1 are given in Appendix II.

In OPTIM 2, due to the continuous character of the variables, the golden section optimum finding technique was chosen. (See Appendix III for details.)

SOME EXAMPLES OF APPLICATION TO REAL CASES

To verify the efficiency of the design programs, a survey was carried out among machine tool manufacturing firms requesting data of real cases of designs of hydrostatic journal bearings. Adding other cases, prepared by the authors, to those supplied by these firms, a sample consisting of 20 design proposals was formed to which the design programs OPTIM 1 and OPTIM 2 were applied. The results obtained are shown in Table 1.

The following conclusions can be drawn from these results:

- There is a wide range of variation for the optimum supply pressure, depending upon the load, the stiffness and the dimensional limitations of the bearing.
- Many of the design variables often take extreme values (min or max).
- In most cases, the minimum value of the

diameter is the optimum. Only in some cases of low where spindle speeds were the optimum diameter values higher than the minimum.

- (d) As in the case of the supply pressure, the interval of values in which the optimum length-diameter ratio can be found is relatively large.
- (f) As far as b/L and a/D ratios are concerned (a and b being the circumferential and axial land) optimum values as low as 0.05 or 0.1 have been recorded.
- (g) The optimum design pressure ratio β is in most cases very near to 0.5, but it must not necessarily coincide with this value.
- (h) It is very unlikely for the design clearance to take as optimum values other values than the minimum ones. However, with very high spindle speeds, there is a trend towards higher values for the optimum design clearance.

The results obtained using the design computer programs were compared with those resulting from the application of other design methods. A comparison of both a design method presented by Bueno⁴ and the computer aided (C.A.) method proposed in this paper is shown in Table 1, which indicates the efficiency of the C.A. method. Different design procedures were applied to the sample (shown in Table 1) and they resulted in most cases, in poorer designs (from the point of view of the consumption of power) than those obtained with the C.A. method. These procedures have the additional drawback of not taking into account the limitations or constraints on the design variables which are likely to exist in some cases.

To emphasize the importance of the optimization, the consequences of the deviations of the design variables from their optimum values were analysed.

The results corresponding to one of the cases included in the sample are plotted in figures 2 to 6. It may be seen how, in this case, a non-optimized design

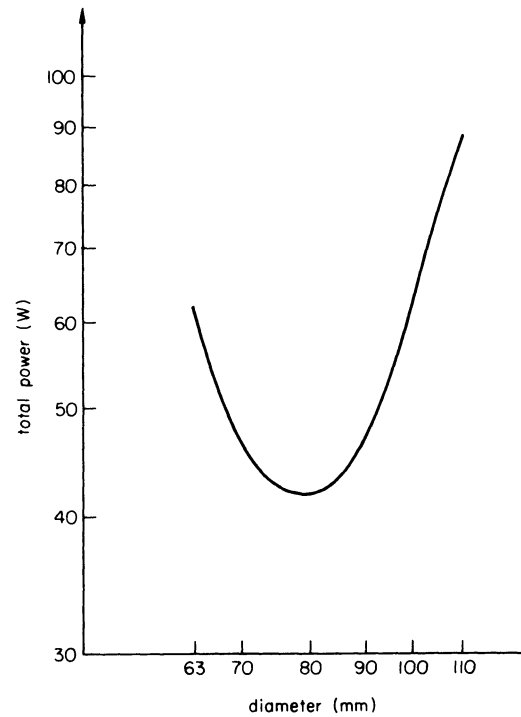


Figure 2. Effect of diameter variations on total power.

TABLE 1 Comparison of results obtained with both design program OPTIM 2 and Bueno's optimization method (SI units).

Variable	D	L/D	b/L	a/D	η	β	h	P_S	W_T
Case: FAB 3		rev/min = 750			$F_{MAX} = 0.588 \times 10^4$		$\lambda_{MIN} = 0.294 \times 10^9$		
MIN	0.55×10^{-1}	0.80	0.05	0.1	0.5×10^{-1}	0.5	0.15×10^{-4}		
MAX	0.70×10^{-1}	1.50	0.1	0.166	0.75×10^{-1}	0.7	0.25×10^{-4}		
OPTIM 2	0.55×10^{-1}	1	0.05	0.1	0.5×10^{-1}	0.510	0.15×10^{-4}	0.741×10^7	56.31
BUENO	0.55×10^{-1}	0.926	0.05	0.103	0.5×10^{-1}	0.5	0.15×10^{-4}	0.771×10^7	56.76
Case: FAB 4		rev/min = 600			$F_{MAX} = 0.980 \times 10^4$		$\lambda_{MIN} = 0.196 \times 10^{10}$		
MIN	0.060	0.50	0.05	0.05	0.25×10^{-2}	0.355	0.1×10^{-4}		
MAX	0.120	2	0.25	0.2	0.4×10^{-1}	0.710	0.3×10^{-4}		
OPTIM 2	0.065	2	0.05	0.2	0.46×10^{-2}	0.557	0.1×10^{-4}	0.541×10^7	37.15
BUENO	0.060	*2	0.05	*0.2	0.25×10^{-2}	0.500	0.1×10^{-4}	0.677×10^7	55.35
Case: FAB 5		rev/min = 200			$F_{MAX} = 0.294 \times 10^5$		$\lambda_{MIN} = 0.196 \times 10^{10}$		
MIN	0.180	0.40	0.05	0.05	0.5×10^{-1}	0.5	0.5×10^{-4}		
MAX	0.180	0.63	0.1	0.2	0.1	0.8	0.1×10^{-3}		
OPTIM 2	0.180	0.63	0.1	0.143	0.83×10^{-1}	0.5	0.5×10^{-4}	0.507×10^7	323.81
BUENO	0.180	*0.63	0.05	0.154	0.5×10^{-1}	0.5	0.5×10^{-4}	0.438×10^7	480.32
Case: FAB 7		rev/min = 1500			$F_{MAX} = 0.490 \times 10^4$		$\lambda_{MIN} = 0.294 \times 10^9$		
MIN	0.060	0.8	0.05	0.1	0.2×10^{-1}	0.5	0.2×10^{-4}		
MAX	0.080	1.0	0.15	0.2	0.5×10^{-1}	0.8	0.3×10^{-4}		
OPTIM 2	0.060	1.0	0.05	0.1	0.2×10^{-1}	0.51	0.2×10^{-4}	0.522×10^7	130.29
BUENO	0.060	1.0	0.05	0.119	0.2×10^{-1}	0.5	0.2×10^{-4}	0.504×10^7	131.49
Case: CIT 10		rev/min = 1500			$F_{MAX} = 0.392 \times 10^4$		$\lambda_{MIN} = 0.294 \times 10^{10}$		
MIN	0.042	0.80	0.05	0.05	0.01	0.1	0.2×10^{-4}		
MAX	0.140	1.70	0.1	0.25	0.07	0.9	0.4×10^{-4}		
OPTIM 2	0.068	1.70	0.06	0.25	0.01	0.529	0.2×10^{-4}	0.938×10^7	373.36
BUENO	0.042	*1.70	0.05	*0.25	0.01	0.5	0.2×10^{-4}	0.237×10^8	1416.62

* This variable has been assigned an extreme value due to the fact that the optimum design after Bueno's method was outside the permitted limits for the variable involved.

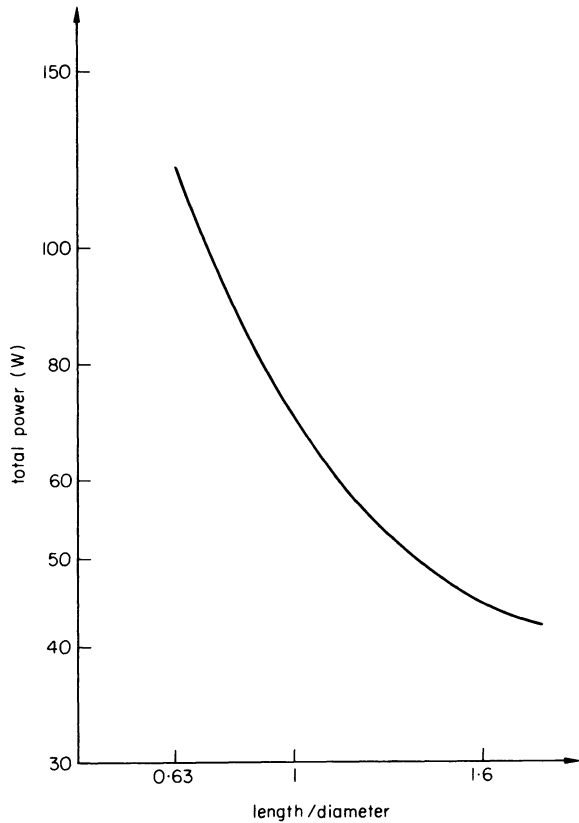


Figure 3. Effect of length-diameter ratio variations on total power.

presents an extra-consumption of power which must not be overlooked.

The consumption of power against the bearing diameter is represented in figure 2 when the rest of the design variables are kept at their optimum values. This graph shows an optimum diameter different from the minimum; this can occur when the spindle speed and the minimum diameter are low and the required operating stiffness or working load are high.

INTRODUCTION OF STANDARD SERIES

Once the design program OPTIM 1 was developed and its correct operation thoroughly checked, it was time

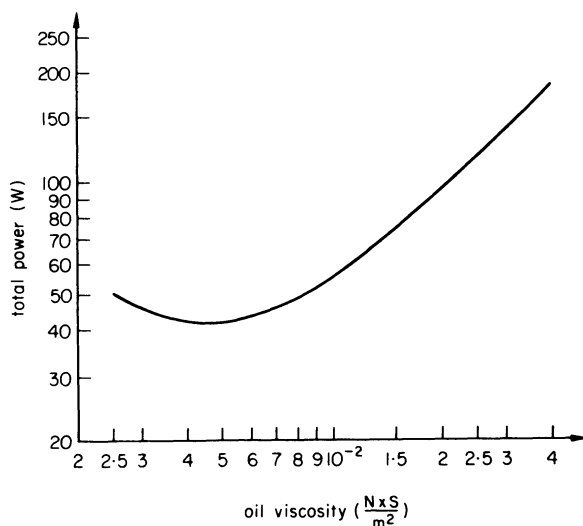


Figure 4. Effect of oil viscosity variations on total power.

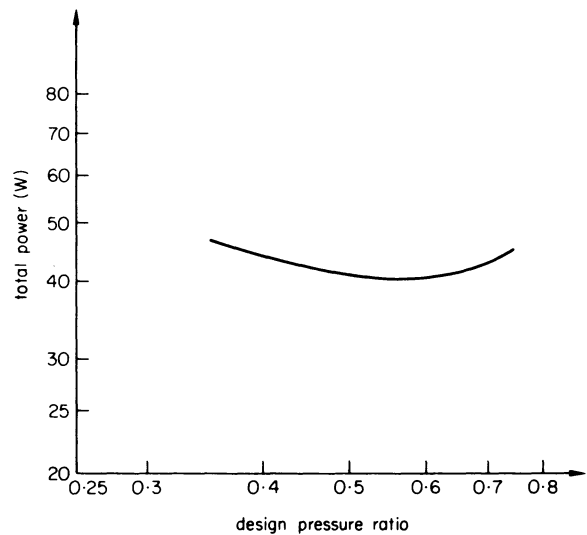


Figure 5. Effect of design pressure ratio variations on total power.

for the introduction of standard series for the design variables. It was thought to be convenient to choose geometric ISO series for the seven variables. Although the scope of OPTIM 1 is so far restricted to geometric standard series, it would be very easy, by modifying the program slightly, to deal with other types of series.

Different series have been chosen for each variable. With regard to the diameter, the corresponding series are ISO R20 and ISO R10. The second design variable, the L/D ratio, can only be assigned three values belonging to the R5 ISO series: 0.63, 1 and 1.60. b/L and a/D ratios, which account for the bearing lands, are in all cases kept fixed. They can be either 0.05 and 0.1 or 0.1 and 0.2. As for the rest of the variables—oil viscosity; design pressure ratio and design clearance, the highest ISO series have been chosen, so that they are treated as if they were of a

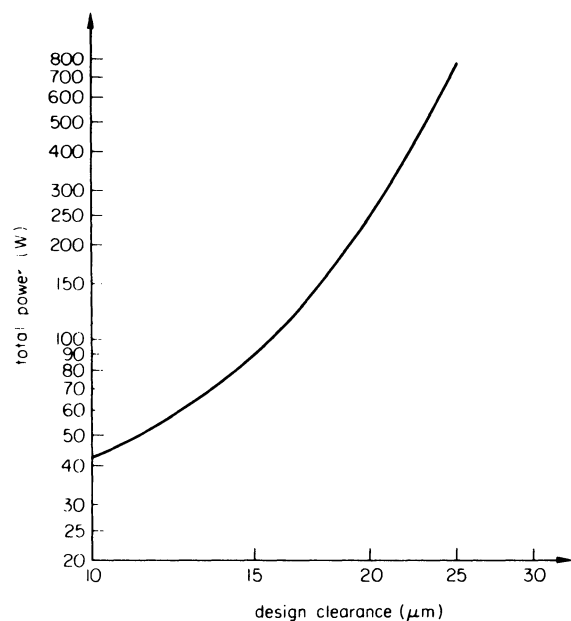


Figure 6. Effect of design clearance variations on total power.

continuous type. The reason for this choice is clear; it makes little sense speaking about standardizing a design variable, such as the design pressure ratio, which must always be of a continuous type.

By combining the different single standard series, four standard groups were set up. In the first one, the diameter belongs to R20, the length-diameter ratio to R5, being $b/L = 0.05$ and $a/D = 0.1$. In the second group, the diameter belongs to R10 and the remaining variables are as in the first group. As for the third standard group, the diameter belongs to R20 and L/D to R5, while $b/L = 0.1$ and $a/D = 0.2$. In the last group, the bearing diameter belongs to R10 and the other variables are as in the third group. These four standard groups, were then applied to the above-mentioned sample. In this manner, it can be said that every single design proposal of the sample has been resolved five times: four times as belonging to each standard group (using the design program OPTIM 1) and once, assuming the design variables to be of a continuous type, with the OPTIM 2 program.

Conclusions concerning the possible interest in using standard bearings have to be derived from a comparison of the results obtained. This comparison is shown in Table 2. The optimum consumption of power corresponding to every single design proposal of the sample can be seen in this table as well the percentages of extra power consumption when designing as belonging to standard groups, over the consumption which results with design variables of the continuous type (namely, values situated under the heading CONTINUOUS).

INTERPRETATION OF RESULTS

Some conclusions can be drawn from the results shown in Table 2. In the first place (and as expected)

these results clearly demonstrate that the optimum designs corresponding to continuous type design variables are, in general, better than those obtained with standardized variables. Designs giving the same consumption of power for both types of variables were recorded only in some cases. This occurs when the optimum is found with most of the design variables taking extreme values.

A closer study of Table 2 shows how standard groups using higher ISO series produce better optimum designs than those which can be reached with low series (series with few standard values). As a result of this, differences in power consumption can be high.

If the results given in columns 3 and 6 (column 1, appears with the heading CASE) are compared with those in columns 5 and 9, the conclusion is drawn that the designs calculated with $b/L = 0.05$ and $a/D = 0.1$ are, in most cases, better than those obtained with $b/L = 0.1$ and $a/D = 0.2$. On the other hand and as far as the average percentage of extra power consumption is concerned (this figure appears at the bottom of the corresponding column) a 10.61 per cent average increase of power was found for the first standard group. The second group has an average percentage of 16.42 per cent. The average percentage figures for the third and fourth groups are 19.5 per cent and 27.17 per cent, respectively.

It can be concluded from these results that, from the point of view of the minimum power consumption, the first standard group (with ISO R20 series for the diameter and $b/L = 0.05$ and $a/D = 0.1$) is the best in the four groups. The second group, with an average increase of 16.42 per cent in the power consumed is consequently less suitable than the first group; however, in some instances it may be preferred to the first group for the sake of reducing by half the allowed diameter values. The third and fourth groups

TABLE 2 Comparison of the optimum designs as belonging to the standard series with the absolute optimum designs (as continuous)

Case	Continuous			DR20; $b/L = 0.05$ $a/D = 0.1$		DR10; $b/L = 0.05$ $a/D = 0.1$		DR20; $b/L = 0.1$ $a/D = 0.2$		DR10; $b/L = 0.1$ $a/D = 0.2$	
	W_T	W_T	%	W_T	%	W_T	%	W_T	%	W_T	%
FAB1	325.79	325.79	0	327.56	+ 0.54	356.30	+ 9.36	356.30	+ 9.36		
FAB 2	599.41	617.10	+ 2.95	714.93	+ 19.27	600.71	+ 0.21	783.86	+ 30.77		
FAB 3	56.31	57.40	+ 1.94	72.57	+ 28.88	77.01	+ 36.76	98.39	+ 74.73		
FAB 4	37.15	41.98	+ 13.00	41.98	+ 13.00	42.24	+ 13.69	42.24	+ 13.69		
FAB 5	323.81	346.43	+ 6.98	346.41	+ 6.97	327.94	+ 1.27	328.08	+ 1.32		
FAB 6	932.54	1142.62	+ 22.52	1142.62	+ 22.52	1510.11	+ 61.93	1510.11	+ 61.93		
FAB 7	130.29	135.35	+ 3.88	135.35	+ 3.88	177.00	+ 35.84	177.00	+ 35.84		
FAB 8	42347.23	48594.75	+ 14.75	58947.35	+ 39.20	42620.89	+ 0.64	49602.43	+ 17.13		
CIT 1	52.36	61.17	+ 16.82	68.49	+ 30.81	67.25	+ 28.43	66.56	+ 27.12		
CIT 2	187.70	218.74	+ 16.53	218.74	+ 16.53	214.60	+ 14.33	214.60	+ 14.33		
CIT 3	200.42	209.18	+ 4.36	209.18	+ 4.36	212.85	+ 6.19	294.90	+ 47.13		
CIT 4	299.65	355.99	+ 18.80	355.99	+ 18.80	354.63	+ 18.34	354.63	+ 18.34		
CIT 5	250.93	268.60	+ 7.04	268.60	+ 7.04	270.58	+ 7.83	270.58	+ 7.83		
CIT 6	18.30	20.57	+ 12.41	20.57	+ 12.41	20.20	+ 10.39	20.20	+ 10.39		
CIT 7	535.60	543.68	+ 1.50	543.68	+ 1.50	535.60	0	535.60	0		
CIT 8	90.51	99.97	+ 10.45	100.06	+ 10.54	109.06	+ 20.49	109.09	+ 20.49		
CIT 9	509.99	617.06	+ 20.99	617.06	+ 20.99	877.43	+ 72.04	877.43	+ 72.04		
CIT 10	373.36	416.43	+ 11.53	494.77	+ 32.51	406.61	+ 8.90	450.16	+ 20.56		
CIT 11	62.28	72.02	+ 15.62	72.02	+ 15.62	70.86	+ 13.77	70.85	+ 13.74		
CIT 12	185.10	203.89	+ 10.15	227.80	23.06	226.76	22.50	271.50	46.67		
Average %			+ 10.61		+ 16.42		+ 19.15		+ 27.17		

Note: figures given under the headings W_T represent the total power consumed by the bearing.

seem to be of little interest because they show an excessive average power increase.

It must be pointed out that for the first standard group, the average percentage of extra power is only 10.61 per cent, while the total number of standard values considered in this group is small. In the case of the diameter, for example, seldom were more than ten numbers used which belong to the ISO R20 series. There are even smaller standard values for the length-diameter ratio, and the third and fourth design variables have been kept fixed.

CONCLUSIONS

Two computer programs were developed which, with reasonable computing costs, are able to achieve an optimized design of a hydrostatic bearing, starting from an initial design.

The results obtained from the application of both programs to a sample of designs showed how, in most cases, it is impossible to predict accurately the optimum values for the design variables by empiric rules, as stated in some design procedures. In support of this it must be brought out the design examples where the optimum values for the design variables were neither the maximum nor the minimum—as suggested in some design methods—but intermediate values. On the other hand, it was observed that the final design is to a great extent influenced by the bearing working conditions and as bearing load, operating stiffness and spindle speed. Not all the variables have to be determined with the same degree of accuracy when an optimum design is sought. Some of these, such as the design clearance, the bearing diameter and the oil viscosity are in a higher degree responsible for the exact determination of the optimum.

Finally, as regards the standard series presented in this paper, it was shown how with a standard series consisting of a very small number of values it was possible to obtain a great variety of designs which were in most cases comparable in every aspect with those achieved upon the assumption of a continuous nature of the design variables. The authors consider this to be an important reason for a full introduction of standardized hydrostatic journal bearings, particularly in the field of machine tool design.

ACKNOWLEDGMENTS

The authors wish to thank the Asociación de Investigación Industrial de la Máquina-Herramienta (INVEMA) for its financial support and also for permission to publish this paper.

APPENDIX I

The total amount of power consumed by a hydrostatic journal bearing can be written as:

$$W_T = W_B + W_R + W_{Rc}$$

where W_B is the pumping power, we obtain the following equation:

$$W_B = \frac{\pi \cdot D}{6 \cdot b} \times \beta \frac{P_s^2 \cdot h^3}{\eta}$$

W_R is the amount of friction power dissipated in the bearing lands and can be described as follows:

$$W_R = [n \cdot (L - 2b) \cdot a + 2\pi \cdot D \cdot b] (\pi \cdot D \cdot N')^2 \frac{\eta}{h}$$

W_{Rc} stands for the friction power consumed in the bearing recesses and, after Shinkle and Hornung⁷, can be written as:

$$W_{Rc} = fp(Re) [(\pi D - na)(L - 2b)] \rho \frac{(\pi \cdot D \cdot N')^3}{2}$$

$$\text{where } fp(Re) = \begin{cases} \frac{8}{Re} & \text{if } Re < 1000 \\ \frac{1}{24.6} Re^{-0.203} & \text{if } Re \geq 1000 \end{cases}$$

APPENDIX II

The Fibonacci search method is resorted to in cases where the maximum of a unimodal function with discrete variables is required. Basically, the method consists of the setting up of a correspondence between the actual discrete values of the variables—which will generally be unequally spaced and which will form the interval of uncertainty—and a series of equally spaced whole numbers. The search is carried out among the numbers belonging to the last series by a sequential placing of experiments in the points defined by Fibonacci numbers.

The sequence of Fibonacci numbers is defined as:

$$FB(0) = FB(1) = 1.$$

$$FB(N) = FB(N - 1) + FB(N - 2) \quad N = 2, 3, 4, \dots$$

Once the numbers of the equally spaced sequence were chosen, one must turn to the corresponding points in the original sequence where the actual evaluation of the objective function is to be made. Depending upon the results obtained, three situations can occur—see figures 7(c), (d) and (e). By further repetitions of the procedure, the interval of uncertainty is narrowed down until the optimum is found out.

APPENDIX III

The search by golden section is an optimum finding technique by which an interval of uncertainty can be cut down to a small final interval. The number of experiments involved in the search is small if compared with other methods.

As in the case of the Fibonacci search, three situations can occur after placing two experiments. The objective function is assumed to be unimodal within the interval of uncertainty. The three situations are illustrated in figures 8(a), (b) and (c).

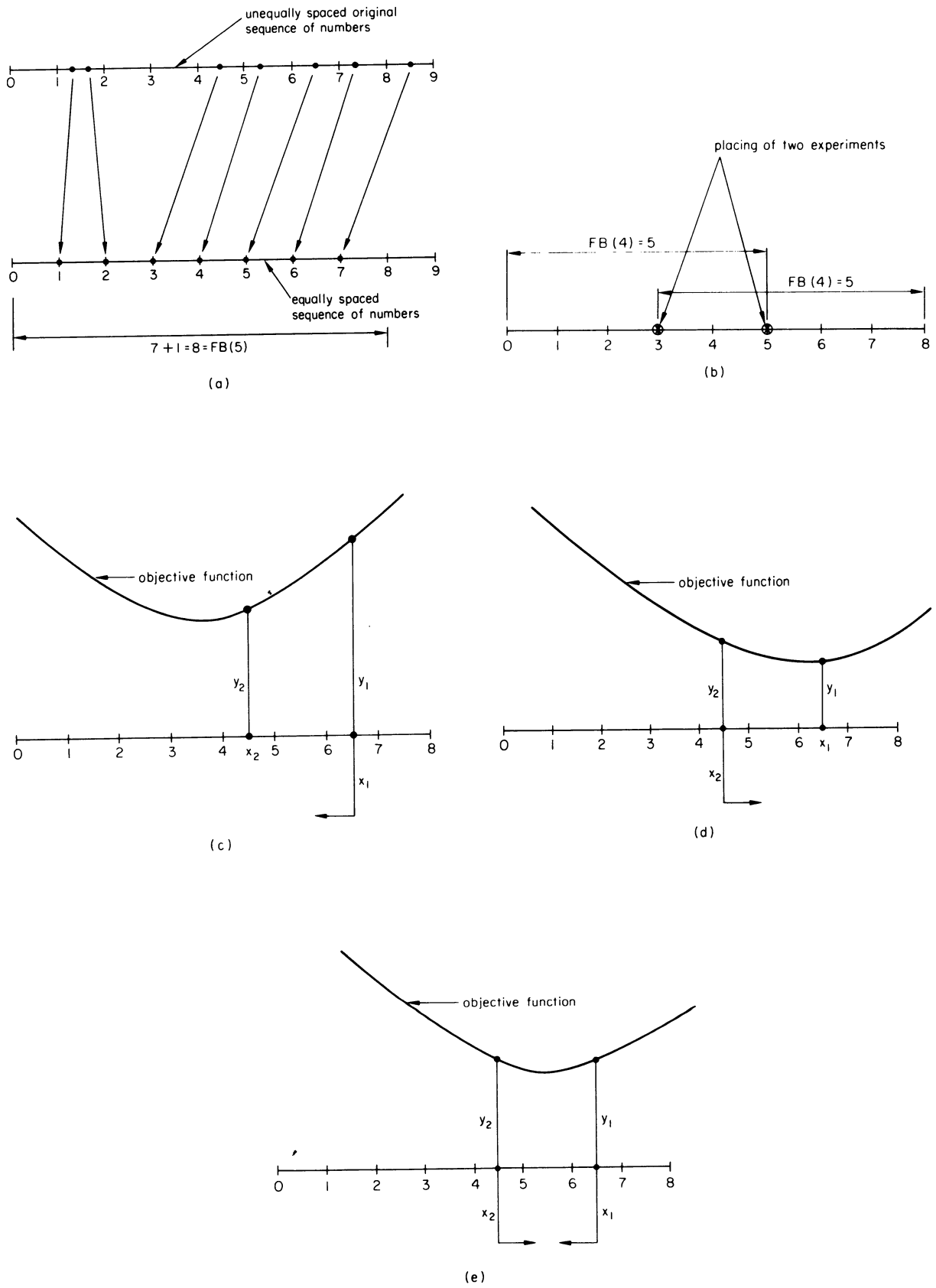


Figure 7. (a)–(e) Fibonacci search.

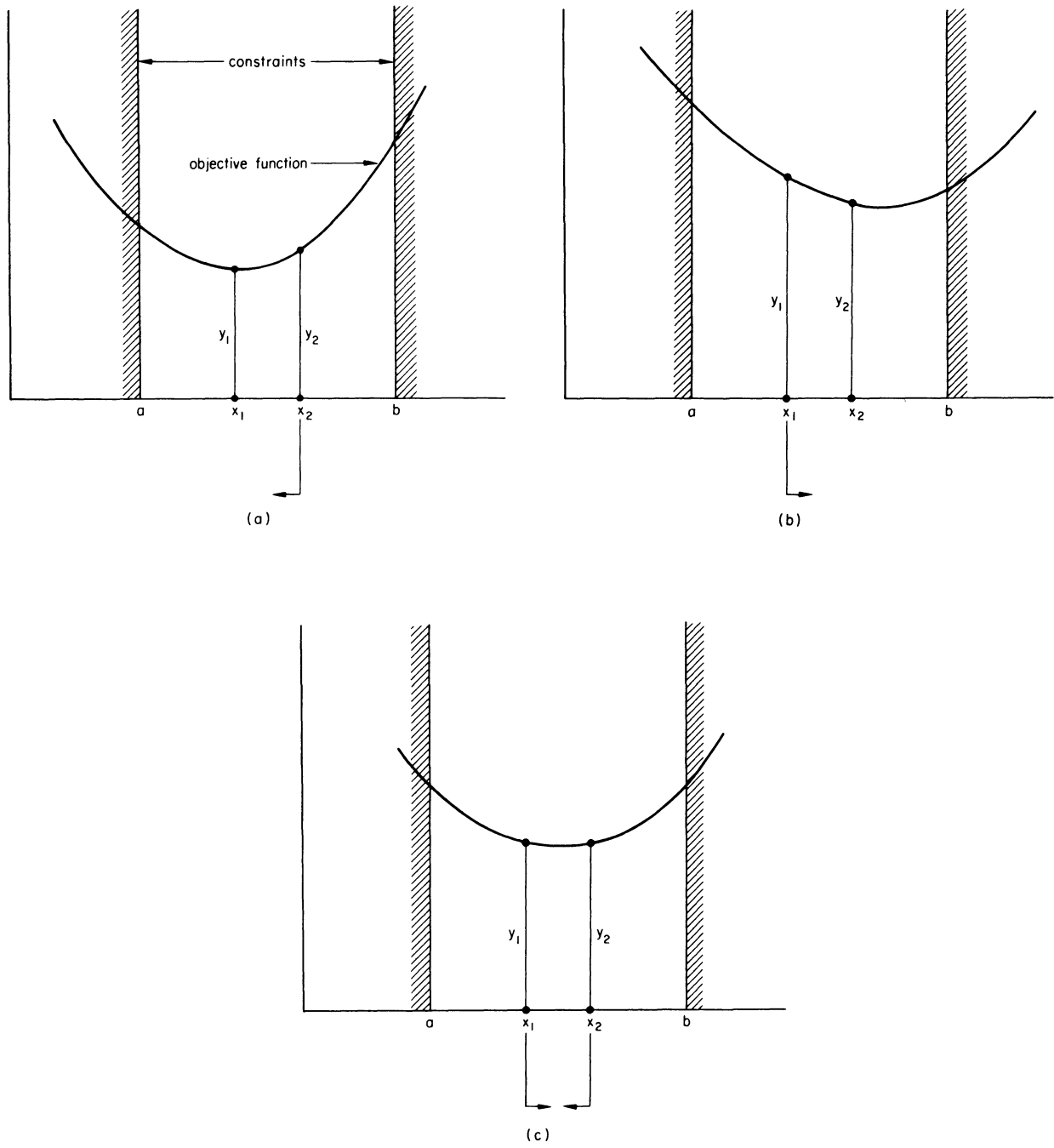


Figure 8. (a)-(c) Golden section search.

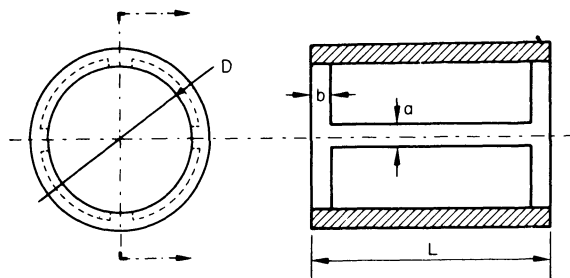


Figure 9. Bearing geometrical dimensions.

In figure 8(a): if $y_1 < y_2$, the new interval of uncertainty will be: $[a, x_2]$.

In figure 8(b): if $y_2 < y_1$, the new interval of uncertainty will be: $[x_1, b]$.

In figure 8(c): if $y_1 = y_2$, the new interval of uncertainty will be: $[x_1, x_2]$. If the interval of uncertainty at any stage of the search is represented by Δx , it can be written

$$b - a = \Delta x$$

If x represents the situation of an experiment, it follows that

$$b - x_1 = x_2 - a = \bar{x}$$

According to the golden section method, there is a relation between \bar{x} and x and this relation is given by:

$$\bar{x} = \alpha \Delta x, \quad \text{where } \alpha = \frac{(5)^{1/2} - 1}{2}$$

The reduction of the initial interval of uncertainty depends only on the total number of experiments which are to be placed. For example, if a reduction to 1 per cent of the initial interval is required, the number of trials will be 11.

REFERENCES

1. F. M. Stansfield. The design of hydrostatic journal bearings. *Proc. of the 8th Internat. MTDR Conference*, 419-445, 1967.
2. W. B. Rowe and J. P. O'Donoghue. *Design Procedures for Hydrostatic Bearings*. The Machinery Publishing Co., Ltd. pp. 9-14, 1971.
3. M. Favareto. Optimization methods for hydrostatic bearing systems. *The First World Meeting on Machine Tools*. Milan, 1971.
4. R. Bueno. Diseño y optimización de cojinetes hidrostáticos radiales. Publicación *INVEMA*. Lubricación hidrostática III, pp. 84-99, 1972.
5. G. Hallstedt. Standardized hydrostatic bearing units. *Joint Conference on Externally Pressurized Bearings*. Coventry, pp. 422-430, 1971.
6. A. Cowley and A. K. Kher. The dynamic characteristics of a hydrostatically supported spindle bearing system. *Proc. of the 10th Internat. MTDR Conference*, pp. 437-451, 1969.
7. J. N. Shinkle and K. G. Hornung. Frictional characteristics of liquid hydrostatic journal bearings. *J. Basic Engineering* (March, 1965) pp. 163-169.

METAL CUTTING

COOLANTS AND CUTTING TOOL TEMPERATURES

by

E. F. SMART* and E. M. TRENT*

SUMMARY

A recently developed metallographic method of estimating temperature gradients in cutting tools has been used to demonstrate the influence of coolants on the distribution of temperature in tools used to cut nickel and iron at relatively high cutting speeds. The results demonstrate that the temperature gradients in tools used to cut these two materials are very different. In particular, there is a cool zone at the cutting edge of the tools used to cut iron which is absent when cutting nickel. The most effective coolant action is shown to be through the tool body and streams of coolant were directed to the hottest and most vulnerable, accessible faces, which were different for iron and nickel. The action of coolant only slightly reduced the temperature at the interface between tool and chip but greatly steepened the temperature gradient into the tool and at the surfaces to which the coolant was directed. It is suggested that further investigation of temperature gradients in tools would provide a more logical basis for coolant application when cutting different materials.

INTRODUCTION

A metallographic method has recently been described for determining the temperature gradients in high speed steel cutting tools¹. This method is effective on tools used to cut higher melting point metals, such as iron and nickel and their alloys, at relatively high rates of metal removal. It depends on the fact that a heat treated tool steel is in a metastable condition and changes to a more stable condition when heated to elevated temperatures. Fully heat treated high speed steel is normally tempered at about 560°C before use and the structure and properties are practically stable up to this temperature. Over 600°C, however, the structure is modified and the hardness and other properties change to an extent dependent upon temperature and time. If the time of heating is known, the temperature can be estimated either from the observed structures or from the hardness to within an accuracy of $\pm 25^\circ\text{C}$ within the temperature range 650–900°C.

Under suitable cutting conditions this method can be used to give much more information about temperature distribution in the tools than has been obtained by other methods. In the present work the changes in metallographic structure have been used to investigate the influence of a coolant on the temperature distribution when cutting iron and nickel at relatively high speeds.

EXPERIMENTAL TECHNIQUES

All the cutting was carried out on a 13 hp lathe, the operation being continuous turning with a constant depth of cut of 1.25 mm (0.050 inch). The tools were in the form of tips clamped in a tool holder. The dimensions of the tool tip were 12.7 mm x 6.3 mm x 9.5 mm (0.5 in x 0.25 in x 0.375 in), the rake and flank surfaces being prepared to a metallographic finish. The cutting angles were:

Top rake	+6°
Side clearance	6°
End clearance	3°
Trail angle	10°
Approach angle	90°
Nose radius	0.37 mm (0.015 inch)

Tool tips were made from BM34 high speed steel (composition shown in Table 1) purchased as fully heat treated bars with a hardness of 840 HV.

The composition and hardness of the work materials are shown in Table 1. The iron was in fact a very low carbon steel in which very small areas of cementite and pearlite were visible among the ferrite grains. The nickel was a bar of commercially pure nickel which had been cast and extruded to a 2.5 inch square section.

The coolant used was a 30 : 1 dilution of Shell Dromus B soluble oil, one part oil, 30 parts water.

* Department of Industrial Metallurgy, University of Birmingham

TABLE 1 Composition of the high speed steel (BM34), iron and nickel used

	HSS (per cent)	Iron (per cent)	Nickel (per cent)
Carbon	0.87	0.04	0.15
Silicon	0.32	0.017	0.35
Manganese	< 0.30	0.14	0.35
Chromium	3.75		
Molybdenum	9.5		
Tungsten	1.65		
Vanadium	1.15		
Cobalt	8.25		
Nickel	< 0.25		99.0
Iron	Bal.	Bal.	0.4
Sulphur	< 0.025	0.010	0.01
Copper	< 0.20		0.25
Phosphorus	< 0.025	0.010	
Hardness VPN	830	83	160

The cutting time was normally 30 seconds, as experience had shown that a stable temperature gradient was established in a tool in a time considerably shorter than this and changed relatively slowly with further extended cutting times. Cutting was stopped by feeding the tool out rapidly by hand. In some cases the chip remained, adhered to the tool. To determine temperature gradients in three dimensions, at least two tools were used under each set of conditions, one for sections normal to the cutting edge and the other for sections parallel to the rake face. Metallographic sections were carefully prepared at definite locations in both directions, the ones parallel to the rake face being as near to the original surface as possible. The structures in the tool were revealed by etching in a 2 per cent solution of nitric acid in alcohol for 30 seconds.

EXPERIMENTAL RESULTS

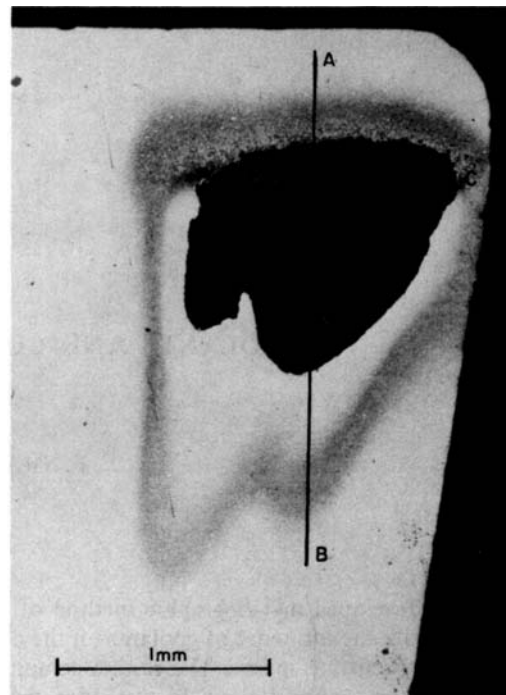
Tools used to cut iron and nickel in air

The etched surfaces of tools used to cut the iron at 183 m/min (600 ft/min), 0.25 mm/rev (0.010 in/rev) feed for 30 seconds are shown in figures 1(a) and 2(a). Figure 1(a) is the rake face of the tool and figure 2(a) a section normal to the cutting edge at a distance of 0.62 mm (0.025 inch) from the nose, A-B in figure 1a. The corresponding temperature gradients estimated from the structures are shown in figures 1(b) and 2(b).

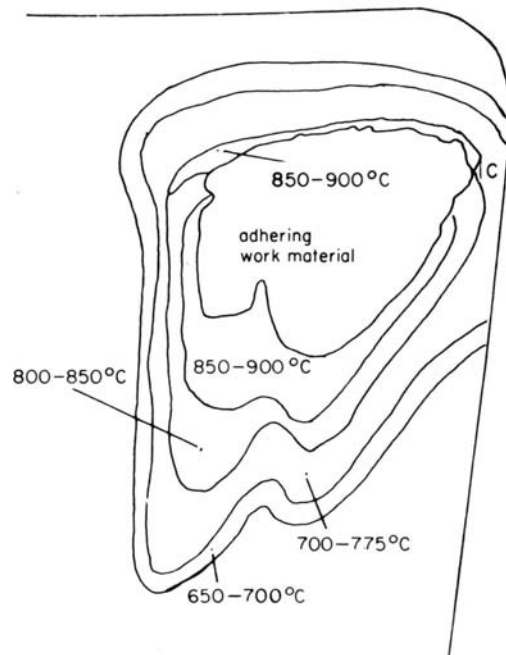
The main characteristics are:

(1) A zone of relatively low temperature extending along the whole cutting edge including the nose radius. In this region no structural changes were detected and the temperature was therefore less than 650°C. This zone was approximately 0.3 mm (0.012 inch) wide, i.e., somewhat wider than the feed. It is in this region that the maximum compressive stress is imposed by the cutting force. This stress falls off rapidly with distance from the edge^{2,3}. There was no visible deformation of the tool edge.

(2) A very steep temperature gradient further from the cutting edge leads to a maximum temperature of 950°C at approximately 1.5 mm (0.060 inch) from the edge. At this very high temperature the



(a)



(b)

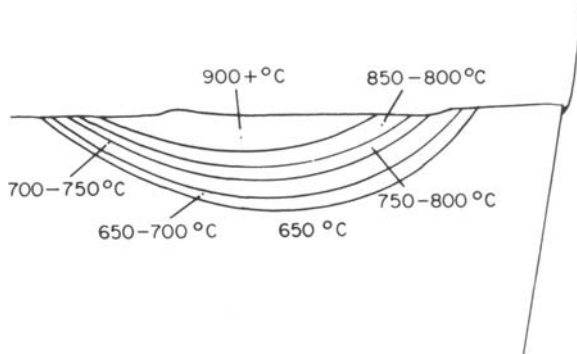
Figure 1. (a) Rake face of tool used to cut iron in normal dry atmosphere at 183 m/min (600 ft/min) 0.25 mm (0.010 in/rev) feed for 30 seconds. Etched 2 per cent Nital. (b) Isotherm diagram derived from metallurgical structures in figure 1(a).

yield strength of the tool steel is greatly reduced and it tends to be sheared away to form a crater⁴. The black area in the centre of the heated region (figure 1(a)) is such a crater which was filled with work material darkened by etching.

(3) The high temperature region extended to the end clearance face (C in figures 1(a) and 1(b) and this was the only part of the tool edge which was seriously weakened by heat. The compressive stress in



(a)



(b)

Figure 2. (a) Section normal to cutting edge of tool used to cut iron under same conditions as figure 1(a). (b) Isotherm diagram derived from metallurgical structures in figure 2(a).

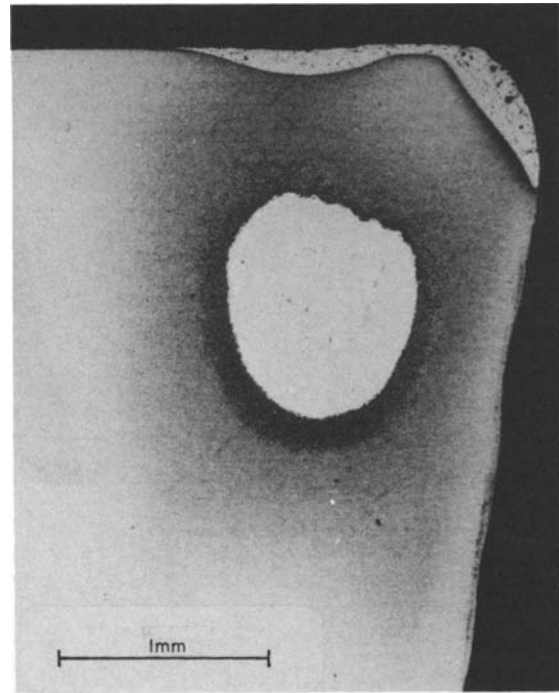
this region is considerable and failure may start at this position, the tool edge being depressed and bulged out on the end clearance face.

(4) As shown in figures 2(a) and 2(b) the 'heat affected region' (i.e., above 650°C) extended into the tool to a maximum depth of approximately 0.5 mm (0.020 inch) in a crescent shaped zone.

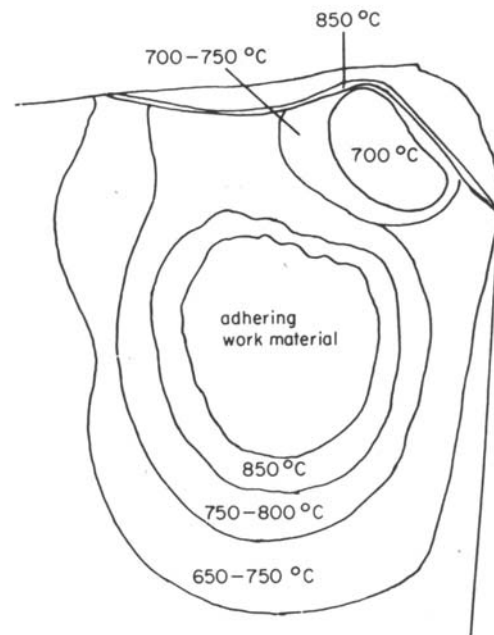
All these features of the thermal gradients are easily reproduced. Successive cuts on the same material under the same conditions reproduce with precision the features (1) to (4) above. The least reproducible feature is the outer limit of the heat affected zone, furthest from the cutting edge, which tends to be uneven and vary from one cut to the next.

Figures 3(a) and 4(a) are the corresponding etched surfaces of tools used to cut the commercial purity nickel at 46 m/min (150 ft/min), with other conditions the same as those for cutting the iron. The temperature gradients estimated from these structures are shown in figures 3(b) and 4(b). The temperature distribution is greatly different and the characteristic features are as follows:

(1) The temperature gradients are much less steep and there is no low temperature region close to the cutting edge. As a consequence, the cutting edge was unable to support the compressive stress imposed by the cutting force and the tool edge was deformed



(a)

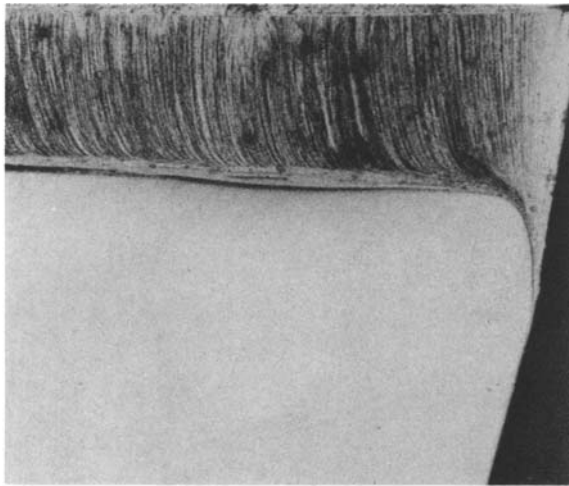


(b)

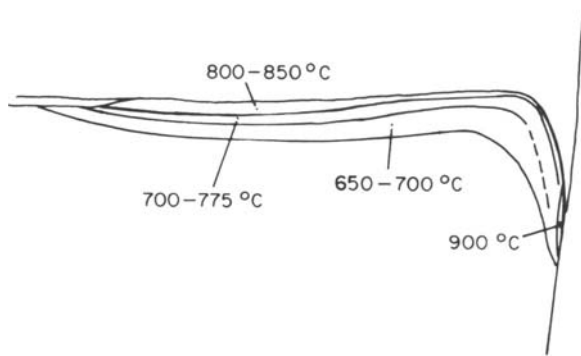
Figure 3. (a) Rake face of tool used to cut nickel in normal dry atmosphere at 46 m/min (150 ft/min) 0.25 mm/rev (0.010 in/rev) feed for 30 seconds. Etched 2 per cent Nital. (b) Isotherm diagram derived from metallurgical structures in figure 3(a).

downward as shown in figure 4(a). The layer of white material at the edge and around the nose radius in the rake section (figure 3(a)) is nickel filling the space left by the deformed edge. This tool would soon have collapsed completely.

(2) There is only a moderate increase in temperature from the cutting edge in the direction of chip flow, and in the region of maximum temperature there is a shallow crater filled with nickel (white)



(a)



(b)

Figure 4. (a) Section normal to cutting edge of tool used to cut nickel under same conditions as figure 3(a). (b) Isotherm diagram derived from metallurgical structures in figure 4(a).

rather further from the edge than when cutting the iron.

(3) The temperature at the end clearance face was not so high as when cutting the iron, and the tool did not tend to deform and collapse at this position.

(4) As shown in figures 4(a) and (b) the heat affected zone extended into the tool all along the rake surface and including the worn and deformed cutting edges.

As when cutting the iron these major features of temperature distribution were repeatable in successive tests under the same conditions.

Temperature gradients in tool with coolants

Coolants were applied in two ways. The first was to flood the rake surface and forming chip with a stream of coolant and the second was to direct a jet of coolant at that part of the exposed tool surface which was at the highest temperature when cutting in air. In the case of the tools used to cut iron this was at the end clearance face near the nose radius, and for tools used to cut nickel the jet was directed into the side clearance and nose (figure 5).

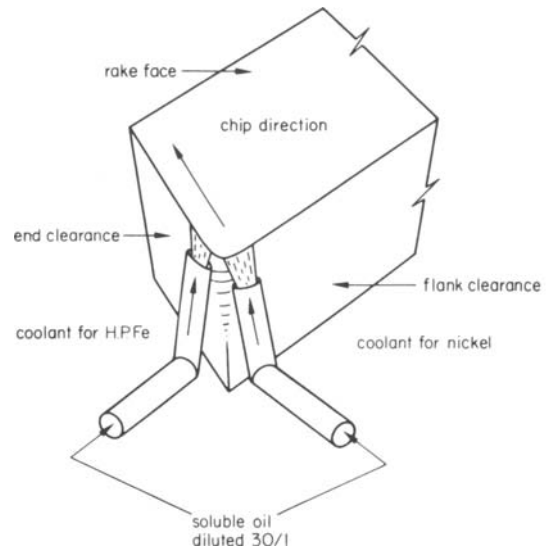


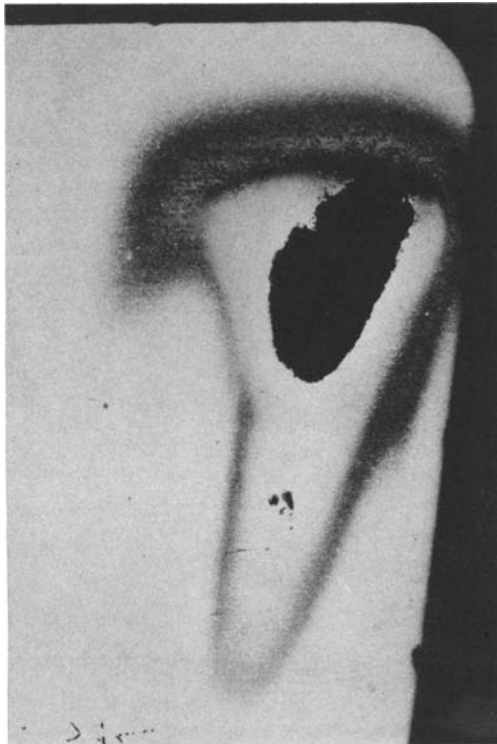
Figure 5. Illustration of selective application of coolant when cutting iron and nickel: Iron-jet applied below end clearance face; Nickel-jet applied below flank clearance face.

Figures 6(a) and (b) show the rake face of a tool used for cutting the iron, cooled by flooding on the rake face. Figures 7(a) and (b) show the tool cooled by the jet on the end clearance face. With the conventional flooding of the rake face, there is some reduction of size of the hottest region, although at one point this extends further from the cutting edge. The crater is somewhat smaller but there is no reduction in temperature at the end clearance face where the tool is beginning to deform. The low temperature region around the cutting edge is roughly the same size.

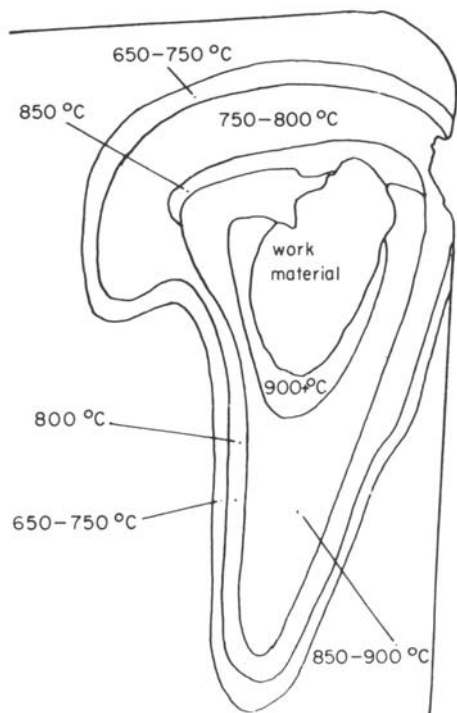
With jet cooling on the end clearance face there is a much greater effect on the temperature distribution (figures 7(a) and (b)). The central hot area is much smaller and there was no sign of a crater in this section. The coolant action was very effective in reducing the temperature at the end clearance to below 650°C and the width of the cool region all around the cutting edge was increased.

Sections through the cutting edge half way along the depth of cut (figures 8(a) and (b), 9(a) and (b)) show that the coolant has decreased the depth of the hot zone below the tool surface, and the influence of jet cooling (figure 9) is very great in this respect, although the surface temperature at the hottest position on the rake face was very little lower than when cutting in air. In figure 9 there is some shearing of the tool surface to form a shallow crater in spite of the coolant action. Figure 10 shows the influence of cooling on the temperature gradient into the tool from the hottest part of the rake surface.

The influence of coolant when cutting the nickel bar is shown in figures 11 and 12, which should be compared with figure 3. The coolant applied by the conventional flooding of the rake face (figure 11) reduced the temperature to a small extent over the whole surface. The temperature is higher than elsewhere at the cutting edge and in an area well back



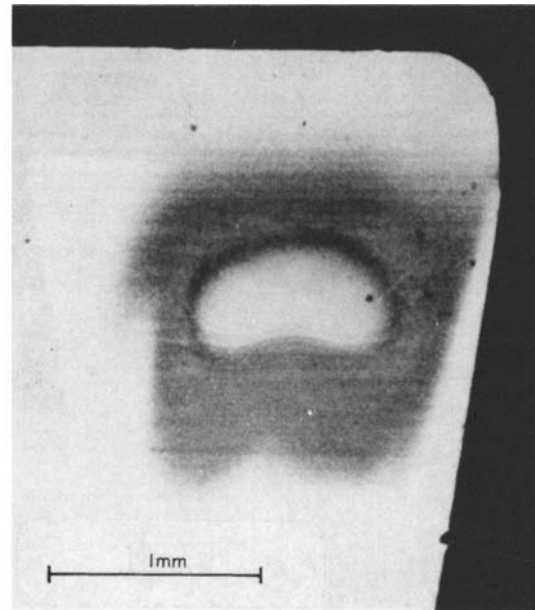
(a)



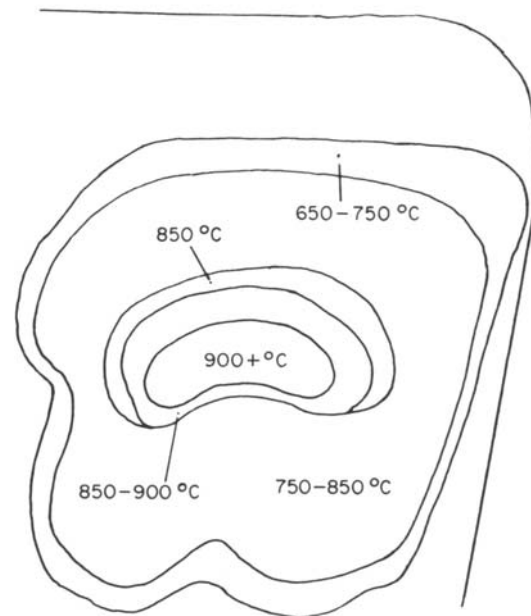
(b)

Figure 6. (a) Rake face of tool used to cut iron under a continuous flow of coolant, directed from above the rake face. Cutting conditions as figure 1(a). Etched 2 per cent Nital. (b) Isotherm diagram derived from metallurgical structures in figure 6(a).

from the edge but the temperature differences are small. No cratering was observed and, although the cutting edge and nose radius were still deformed downwards this deformation is much reduced in the cooled tool. Jet cooling in the side clearance (Figure



(a)



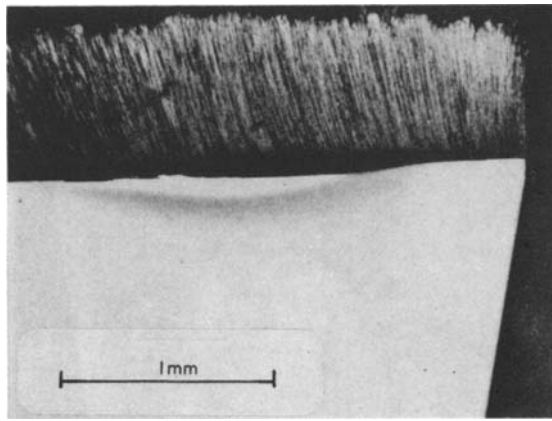
(b)

Figure 7. (a) Rake face of tool used to cut iron under a continuous flow of coolant applied from below the end clearance face (see figure 5). Cutting conditions, as for figure 1(a). (b) Isotherm diagram derived from metallurgical structures in figure 7(a).

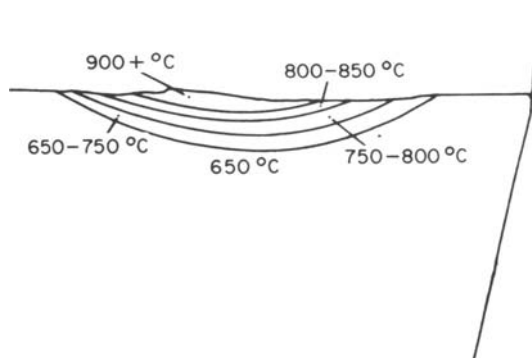
12) still further reduced the temperature, which was nowhere much higher than 750°C and the amount of deformation at the tool edge was much smaller. The sections through the tools (figures 13 and 14) show the considerable reduction in deformation and wear on the tool flank brought about by use of the coolant, particularly the jet in the side clearance.

DISCUSSION OF RESULTS

In considering the results the nature of the heat source must be appreciated. Under the cutting conditions used in these tests, as in most high speed

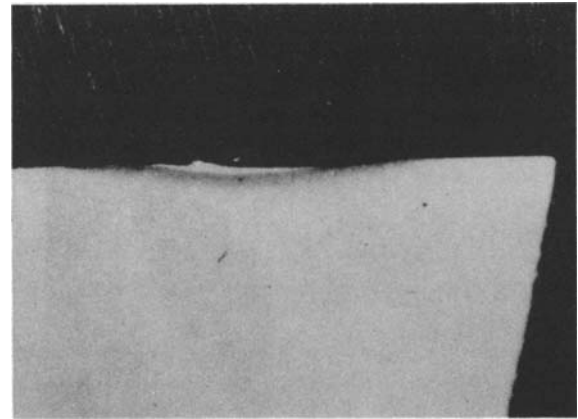


(a)

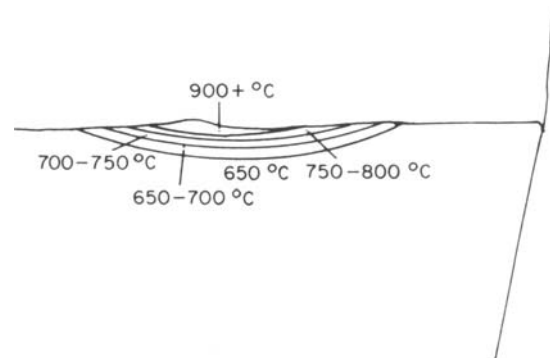


(b)

Figure 8. (a) Section normal to cutting edge of tool used to cut iron under a continuous flow of coolant directed from above the rake face. Cutting conditions, as for figure 1(a), etched 2 per cent Nital. (b) Isotherm diagram derived from metallurgical structures in figure 8(a).



(a)



(b)

Figure 9. (a) Section normal to cutting edge of tool used to cut iron under a continuous flow of coolant applied from below end clearance face (see figure 5). Cutting conditions as for figure 1(a). (b) Isotherm diagrams derived from metallurgical structures in figure 9(a).

cutting, the work material is effectively seized to the tool and movement of the chip involves intense shear in a very thin layer of the work material at the tool surface (called the 'flow zone'). This is most clearly seen in the sections through the tools used to cut the nickel (figures 4, 13 and 14). The work done on the shear plane to form the chip is converted into heat and largely carried away with the chip with little influence on the temperature of the tool. The heating of the tool is largely done by the shearing of the metal in the flow zone. Typically 20-25 per cent of the work of cutting is expended in the flow zone, but as the volume of metal in this zone is small (usually about 5 per cent of the volume of the chip) much higher temperatures are achieved. Thus in the iron chips under the cutting conditions used here, the flow zone reached a temperature of nearly 1000°C while the chip itself was never red-hot, probably about 300°C. The high temperatures reached are not readily observable but have a major influence on the performance of tools as is clear from the observations on the wear and change of shape of tools at positions of high temperature and stress.

Observations suggest that, over most of the seized interface, the bonding is metallic in character and

heat is therefore readily conducted from the flow zone into the tool and the temperature in the tool at any point must be very close to that of the part of the flow zone with which it is in contact. Since there are no gaps or capillary channels at this part of the interface to permit the coolant to penetrate, the interface can only be cooled by conduction of heat through the tool or the body of the chip. The only

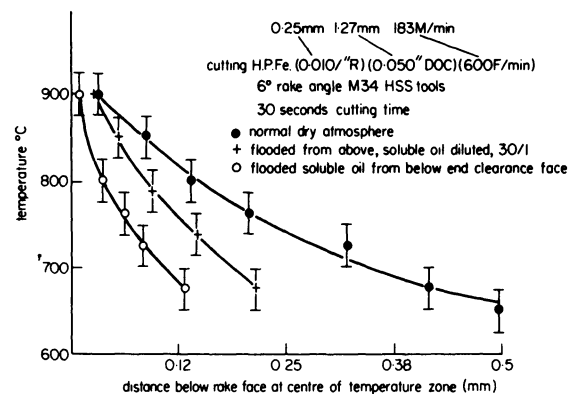
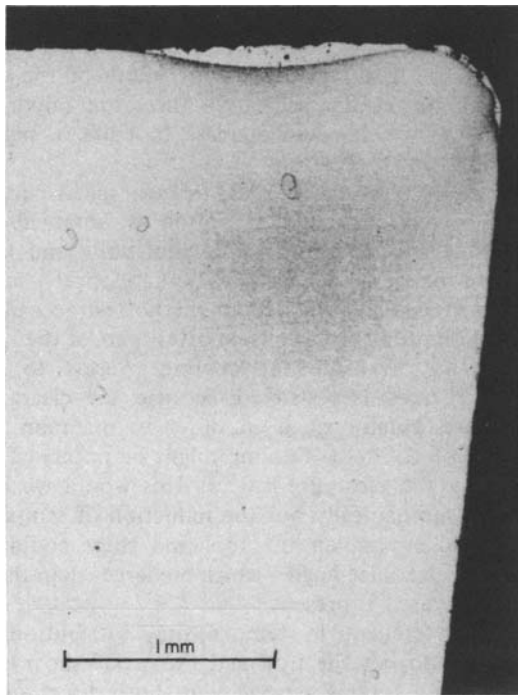
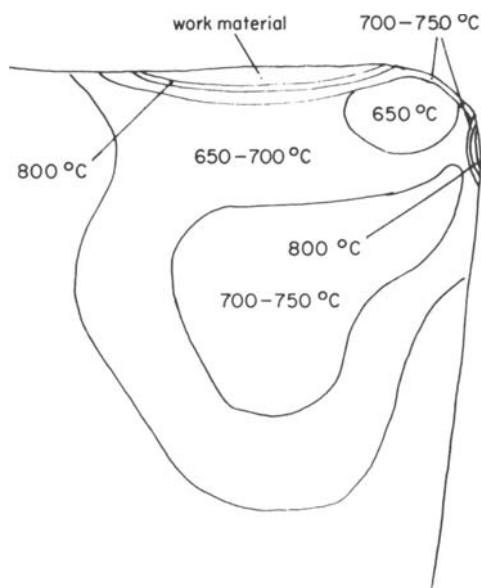


Figure 10. Influence of coolants on temperature gradients into tool from rake face.



(a)

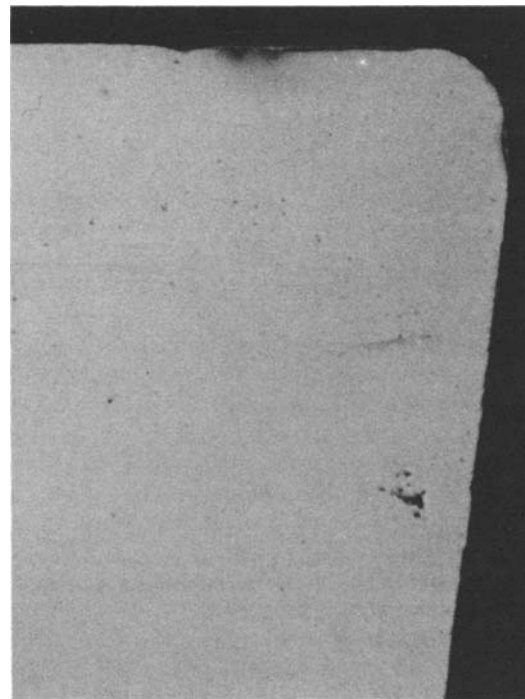


(b)

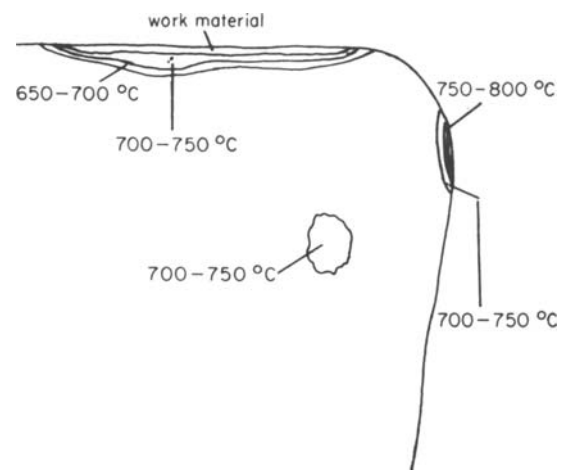
Figure 11. (a) Rake face of tool used to cut nickel under a continuous flow of coolant directed from above the rake face. Cutting conditions, 46 m/min (150 ft/min) 0.25 mm rev (0.010 in/rev) feed for 30 seconds, etched 2 per cent Nital. (b) Isotherm diagram derived from metallurgical structures in figure 11(a).

exception to this is at the edges of the chip where the edge of the flow zone is exposed to direct action by the coolant. This view of the cutting conditions existing in high speed cutting operations suggests that, in using a coolant, the aim should be a maximum reduction of temperature of those parts of the tool which suffer damage as a result of high temperature and stress.

The drastically different temperature distribution in tools used for cutting iron and nickel leads to damage to different parts of the tool. With the nickel, high temperatures were achieved at much lower cutting speeds and the tool was damaged at lower temperatures than when cutting the iron. The experimental results show that, when the rake face of the tool is flooded with coolant, there is some reduction of temperature in the tool when cutting both metals. At the interface the reduction of maximum temperature seems to be small in both cases. In the tool used for cutting the iron it is doubtful whether the maximum interface temperature was reduced at all,

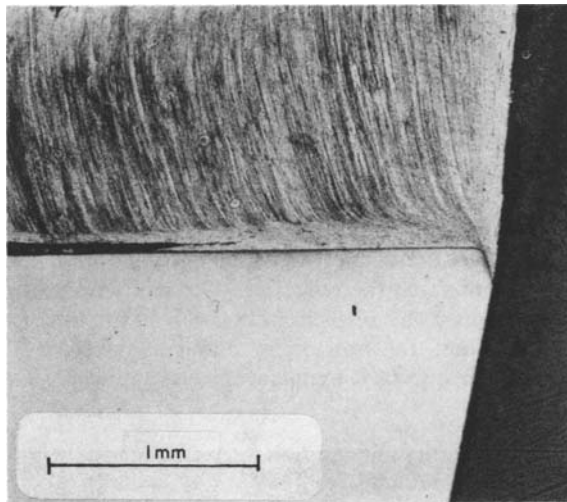


(a)

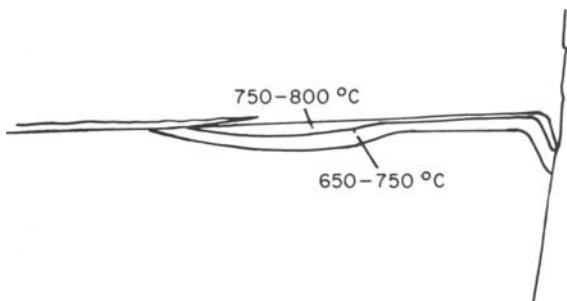


(b)

Figure 12. (a) Rake face of tool used to cut nickel under a continuous flow of coolant applied from below the flank clearance face. (see figure 5). Cutting conditions as for figure 11(a). The section is made a short distance below the level of the temperature contours in figure 14. (b) Isotherm diagram derived from metallurgical structures in figure 12(a).



(a)



(b)

Figure 13. (a) Section normal to cutting edge of tool used to cut nickel under a continuous flow of coolant directed from above the rake face. (b) Isotherm diagram derived from metallurgical structures in figure 13(a).

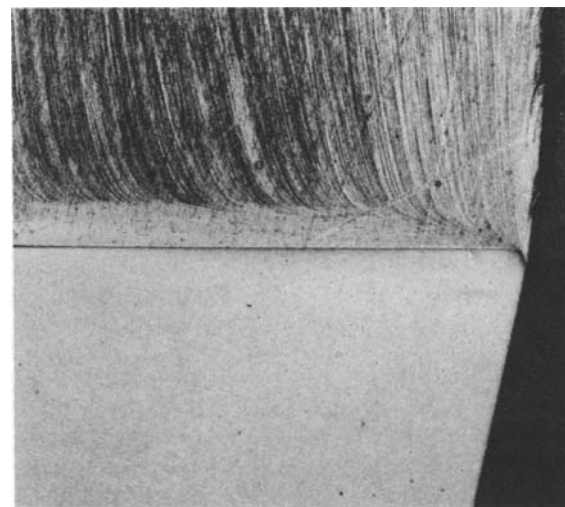
although the area at high temperature may have been somewhat smaller. With the nickel the reduction of temperature at the interface was greater, no crater was observed and there was less deformation and flank wear at the edge. The greater cooling effect when cutting nickel may be due to the lower cutting speed, which gives more time for cooling through the chip body. Comparison of figure 2 and figure 8 suggests that, while the surface temperature at the rake face is relatively unchanged, the cooling of the tool body has led to a steeper temperature gradient into the tool.

Direction of the coolant to the hottest and most vulnerable parts of the tool surface appears to be considerably more effective than flooding the rake surface. In tools used to cut the nickel, the maximum interface temperature seems to have been reduced, although the fall in temperature is relatively small, and the cooling by a jet directed at the cutting edge was more effective in reducing deformation and wear of the edge than flooding the rake face with coolant (figures 3, 11 and 12). When cutting iron the results are more emphatic. While cooling of the end clearance face did not greatly reduce the maximum temperature, it restricted the area of cratering and prevented it spreading to the end clearance and nose with consequent breakdown. It greatly steepened the

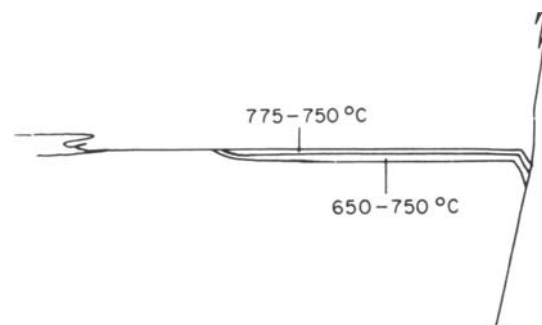
temperature gradient into the tool and reduced the depth to which the tool was overheated (figures 2, 8, 9 and 10). It also increased the width of the cool zone at the cutting edge. All these are important advantages which could increase tool life or permit higher cutting speeds.

These results suggest that, in high speed cutting, the most effective cooling action is obtained by promoting cooling through the tool body and that this can be achieved by use of a relatively small volume of coolant directed to the hottest accessible parts of the tool surface. The hottest part of the rake face is not available for cooling. Access to the clearance faces is restricted because the clearance angles are usually small in order to maintain the strength of the tool. Cooling might be improved by increasing the clearance angles. This would weaken the tool mechanically but the reduction of temperature could strengthen the tool and there could be optimum clearance angles which are larger than those commonly used at present.

The differences in temperature distribution in tools used to cut the iron and the nickel show that the optimum cooling arrangement (both direction of coolant and shape of tool) are likely to be dependent on the material being cut. This raises the question of



(a)



(b)

Figure 14. (a) Section normal to cutting edge of tool used to cut nickel under a continuous flow of coolant applied from below the flank clearance face (see figure 5). Cutting conditions, as for figure 11(a). (b) Isotherm diagram derived from metallurgical structures in figure 14(a).

the significance of the tests carried out on these two materials. Work done so far on steels, including carbon and austenitic stainless steels, has shown that the temperature gradients in tools used to cut these steels at high cutting speeds are similar in character to those in tools used to cut the iron demonstrated here. The addition of alloying elements greatly reduces the cutting speeds at which high temperatures appear but the temperature pattern remains much the same. Work on nickel based alloys has been more limited, but so far this suggests that tools used on the main nickel-based alloys show temperature gradients like those of tools used on the nickel with high temperatures near the cutting edge and shallow temperature gradients. Why iron and nickel and their alloys should behave in such fundamentally different ways during cutting is not clear and is the subject of continuing research.

The temperature gradients and cooling effects demonstrated here are likely to apply to steels and nickel-based alloys under conditions where the heat source is a flow zone on the rake face of the tool. In general this is in the high speed cutting range. It has not been possible to carry out similar tests using carbide tools because they do not undergo structural changes like those of high speed steel which can be used to monitor temperatures. However, observations suggest that the heat source is of the same character when machining with carbide tools as high speed steel⁵ and the general character of the temperature distribution should therefore be the same. The values of temperature must vary with the thermal conductivity of the tool, i.e. with the rate at which heat is conducted from its source in the flow zone. The present experiments suggest that the influence of the thermal conductivity of the tool material on the interface temperature might not be very great. It has been shown that greatly increasing the temperature gradient in the tool by cooling it has had relatively little effect on the temperatures at the interface.

Where cutting conditions are such that a flow zone on the tool rake face is not present, the general character of the temperature gradient in the tool could be greatly different from those demonstrated here. In particular the present experiments do not relate to conditions where a built-up-edge exists or where, as at very low speeds, the work material may slide over the tool surface. Conclusions regarding coolant application would be very different for these cutting conditions.

Because of the great variety of conditions used in industrial machining operations, no simple experiments can produce results directly applicable to all conditions. What has been put forward here is the proposal that the study of temperature gradients in tools can provide a more logical basis for the application of coolants in high speed cutting. This cannot be a substitute for long time tool wear testing and industrial experience, but it could provide a guide to development of cooling techniques. The simple experimental work recorded here should be supplemented by long time tests and further studies of temperature gradients with a variety of different work materials, tool shapes, cutting conditions and coolant arrangements.

CONCLUSIONS

The results of tests in which the temperature gradients have been estimated in high speed steel tools used for cutting iron and nickel with and without coolant, have led to the following conclusions for relatively high speed cutting conditions where the heat source is a flow zone at the tool rake surface.

- (1) The temperature gradients in tools used for cutting iron and nickel were very different in character, the temperature gradients being much steeper when cutting iron, with a cool zone near the cutting edge which was absent when cutting nickel.
- (2) The major cooling effect was through the body of the tool and the coolant acted most effectively when directed to the hottest and most vulnerable exposed part of the tool surface.
- (3) Effective use of the coolant reduced only slightly the maximum interface temperature between tool and workpiece but greatly increased the temperature gradient into the tool and near the cooled surface.
- (4) Coolant directed to the most vulnerable part of the tool surface which is accessible reduced the damage to the tool.

It is suggested that further investigation of temperature gradients could lead to a more logical guide to the applications of coolants when cutting different materials at high cutting speeds.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support given for the work by the Science Research Council. The work was made possible by the use of the facilities at the Department of Industrial Metallurgy at the University of Birmingham. We should like to thank Mr G. J. Gunnell for his work in preparing the illustrations. The authors are grateful to the International Nickel Co. Ltd. for supplying the nickel bar used in the machining tests.

REFERENCES

1. P. K. Wright and E. M. Trent. *Journal of the Iron and Steel Institute* (May, 1973) 211, Part 5, 364-68.
2. N. N. Zorev. *International Research in Production Engineering* (1963) ASME, New York 42-9.
3. G. W. Rowe and A. B. Wilcox. *Journal of the Iron and Steel Institute* (March, 1971) 209, no. 3, 231-2.
4. P. K. Wright and E. M. Trent. Metallurgical appraisal of wear mechanisms *Metals Technology* (January, 1974) 1, no. 1.
5. E. M. Trent. Conditions of seizure at the tool work interface *I.S.I. Special report* 94, 11.

TOOL WEAR OR TOOL DESIGN (A case study)

by

R. A. ETHERIDGE* and A. J. A. SCOTT†

SUMMARY

There are a group of problems which on face value appear to be associated with tool-wear mechanisms, but which, after a careful examination are either design based or metallurgical in their nature.

This paper deals with one such case and is concerned with the problems of uneven flank wear on a large, helically fluted form milling cutter. The problem caused great concern because considerable delays in production occurred because of premature failure of the cutting edge.

The fundamental problem is defined and discussed and a series of controlled tests are described in order to establish the criterion contributing most to the wear mechanism. It is shown, using a non-dimensional wear-rate that the rake angle has the predominant effect upon wear-rate for the particular tool-workpiece combination.

After careful measurements made upon the cutter it is shown that the manufacturing techniques used for the form cutter lead to incorrect rake-face geometry and consequently varying rake angle over the cutting face.

An analysis of the geometry forms part of a larger programme of research but an indication of the mechanism is given in the paper.

INTRODUCTION

There are a group of problems that, on face value appear to be predominantly associated with the tool-wear mechanism, but which after careful examination are shown to be either metallurgical in their nature or, more occasionally, design based.

Many problems connected with the machining of, for example, non-ferrous alloys, fall into the former area owing to the formation of complex intermetallic inclusions¹, but the case examined in this paper is one where a cutter's design contributed more to the failure mechanism than any other single factor.

The investigation centred upon a large, helically gashed form milling cutter for the production of steam-turbine blades machined from solid pieces of stainless iron.

STATEMENT OF THE PROBLEM

Turbine blades are made by several methods depending upon the blade geometry, work material, or the batch size.

In the particular case of milling from the solid, blade geometry is produced by the combined effects of forming and generating; the blade sectional form being produced by the form cutter while the twist and thickness is controlled by use of an inclined

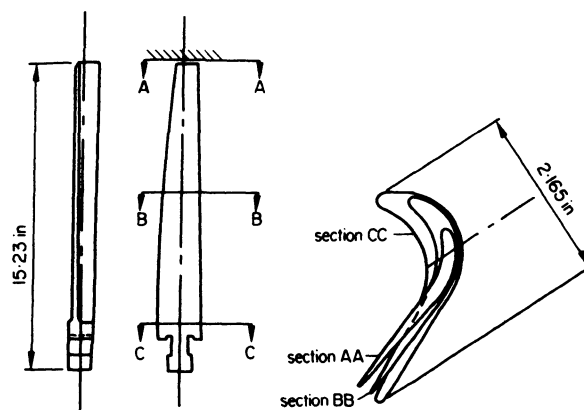


Figure 1. Blade geometry.

rotating fixture on the milling machine table. Figure 1 shows the blade geometry and figure 2 the cutter geometry.

The economics of such heavy forming operations are largely dependent upon the high replacement cost of the special form cutters and, more precisely, the tool life between regrinds. It had been found by experience that the maximum wear land at failure varied from batch to batch and further that the wear land around the cutter profile also varied. (The

* Department of Production Engineering, University of Aston in Birmingham, Gosta Green, Birmingham B4 7ET
† G.E.C. Turbine Generators Limited, Rugby.

practical wear criterion used was the inability of the tool to cut owing to excessive chatter.)

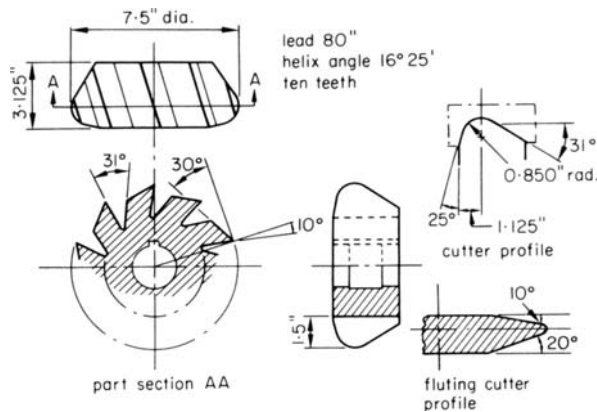


Figure 2. Form cutter geometry.

From production experience it had been found the best tool life was obtained using an arbor speed of 60 rev/min, a feed of 2.25 in/m while climb milling. The cutters were made from high-speed steel and were 7.5 in diameter and 3.125 in thick, whilst the depth of form was of the order of one-half the cutter radius.

With existing cutters it was usual, even so, to regrind at least once per batch of blades, and therefore any improvement in useful cutter life must reduce machine down time with a consequent saving in work-in-progress inventory. It was estimated that an increase of 50 per cent in tool life between regrinds would allow a complete batch of blades to be machined without stripping down the tooling.

The purpose of the investigation was therefore to determine the nature and cause of such uneven tool wear and to establish new design and/or manufacturing procedures in order to improve tool life.

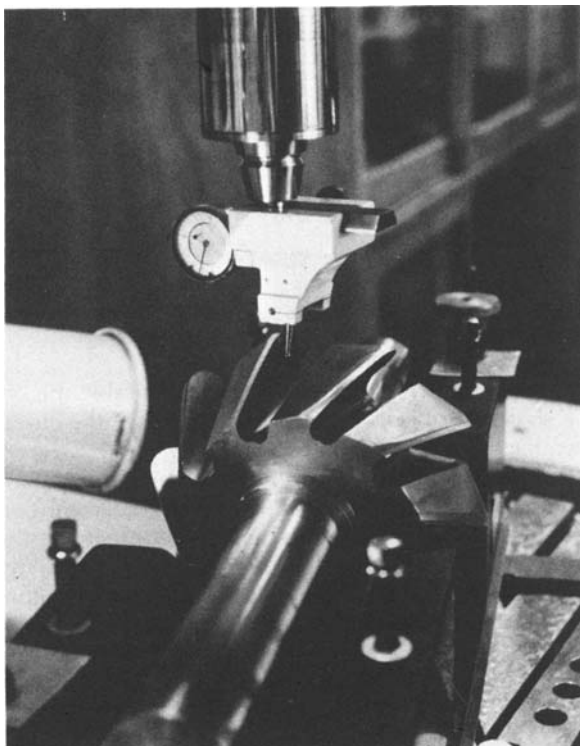


Figure 3. Set-up for measuring rake angle.

PHYSICAL EXAMINATION OF THE PROBLEM

The rake face on an unused cutter

Before any consideration was given to the design of experiments it was necessary to investigate the rake-face geometry of the cutter because helical fluting and helical grinding during manufacture had produced a complex profile over the cutting face.

Owing to the large physical size of the cutter it was impossible to use existing measuring equipment, and so the only practical alternative was to adapt a jig borer to such a purpose. The arrangement is shown in figure 3.

A Societe Genevoise D'Instruments de Physique No. 2P optical machine was used which had a two co-ordinate horizontal table movement with a guaranteed accuracy of 0.0001 in. Accurate vertical displacements of the spindle were obtained by a built-in depth-measuring device which consisted of an adjustable support for slip gauges and a dial indicator reading to 0.0001 in.

Before any measurements could be made it was necessary to establish a system of datums in order to provide a reference for cutter radius and rake angle at any point across the section. In order to do this the cutter was mounted on a purpose-made mandrel which in turn was carried on two micro-flat 'Vee' blocks. The whole fixture was clamped to the machine table at the helix angle specified on the cutter drawing. The helix angle was set by means of a 10-in sine bar, slip gauges, and parallels, while the correct end-positioning of the mandrel was ensured by means of a ground thrust-plate and a 0.375 in precision steel ball, locating in dimples.

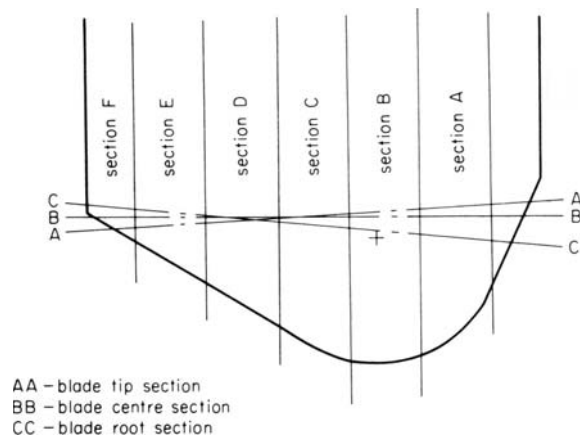


Figure 4. Transverse cutter reference sections.

Measurements were made on each tooth at six normal sections through the tooth profile (see figure 4) by means of an accurate locating dial gauge used as a fiducial indicator. Thus by using the vertical movement of the quill and careful table setting it was possible to establish the shape of the rake-face in the specified normal plane for each section. The normal radial rake was established by constructing a radial line from the cutter centre, intersecting with a tangent to that point on the cutter tip profile under consideration. Both the rake-face geometry and associated normal rake angle for a particular tooth at

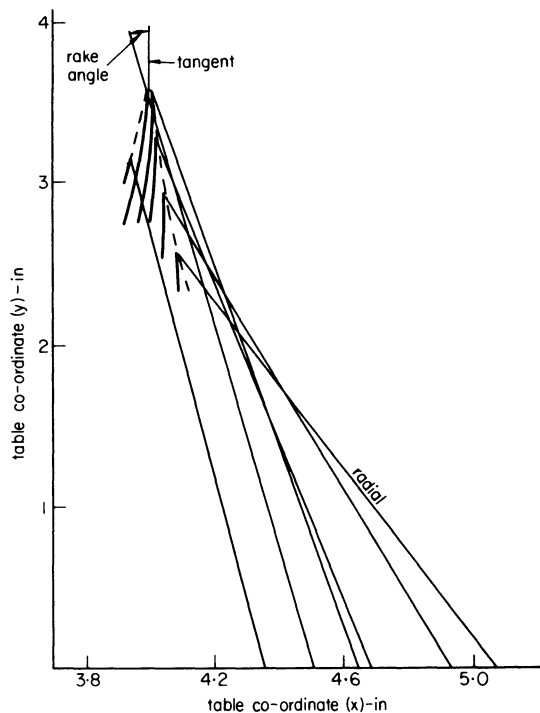


Figure 5. Rake angles at reference sections.

each section is shown in figure 5. Once the rake-face geometry had been established for each tooth the cutter was returned to the works for production use.

Wear characteristics in use

After being returned for a production run the cutter produced 78 blades before failure occurred due to the criterion previously stated. The wear land width was established by means of a locating microscope in the jig borer and using the original fixture. Measurements were made at each transverse section on every tooth, and the mean wear pattern was established. After a further regrind and in-service use the cutter failed after producing a further 69 blades. The results for both tests are shown in figure 6. Clearly although the pattern of wear was similar for both runs, the degree of wear was smaller in the second case, thus reflecting the somewhat arbitrary nature of the failure criterion.

Examination of the results from this section of the investigation indicated that the largest contributory factor to tool failure was rake angle and rake-angle variation over the cutter section.

EXPERIMENTAL DESIGN

The problems of flank wear in milling cutters had previously been examined under conditions of controlled experiment by Pera² and Sabberwall and Fleisher³, the latter clearly establishing that a rake angle associated with minimum wear-rate existed. The physical shape of the cutter under examination presented certain practical difficulties from the point of view of controlled experimentation because both the cutter diameter and rake angle varied over the section, thus causing variable cutting speeds from point to point around the profile. A variation in cutting depth was also present due to the superposed work-piece movement.

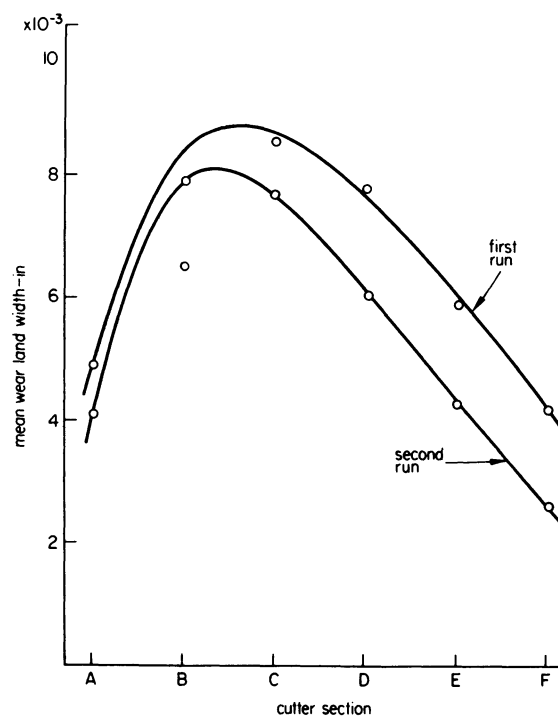


Figure 6. Wear-land widths at reference sections.

The complexity of interaction between these various parameters ruled out a diagnostic exercise using the production cutter, and a solution using simulation techniques was proposed. Such a simulation was made easier because certain parameters were fixed by the design specification, namely the depth of cut and the fixed radius of the cutter at any point. The latter obviously determined the velocity relationship over the profile.

The major difficulty posed in the simulation situation was the comparison of wear rates from point to point on the physical cutter owing to the velocity variation and change in depth of cut. The Pera² results had stressed the differences which existed between wear results considered against the base of time and the other alternative base of volumetric removal rate. This particular problem contained elements of both, and some other method of examining wear rate was necessary. This problem was overcome by using a non-dimensionalized wear-rate originally suggested by Hsu⁴ where the wear-rate represents the slope of the graph of wear-land width against tool-contact rubbed length. The main advantage of such a wear-rate was that the contact rubbed length could be calculated for any point on the cutter profile and could also be used in the simulated system.

Martellotti's⁵ analysis of the milling process showed that tooth-contact length was dependent upon the cutter radius, depth of cut, and feed per revolution of the cutter. Therefore if the contact-path length was known for the real cutter at any point, then a simulated cutting length for a given depth of cut could be obtained on a standard milling machine in the simulated system. The resulting feed rates were generally higher than in the real situation, but this gave the benefit of accelerated testing if the Pera² results were accepted.

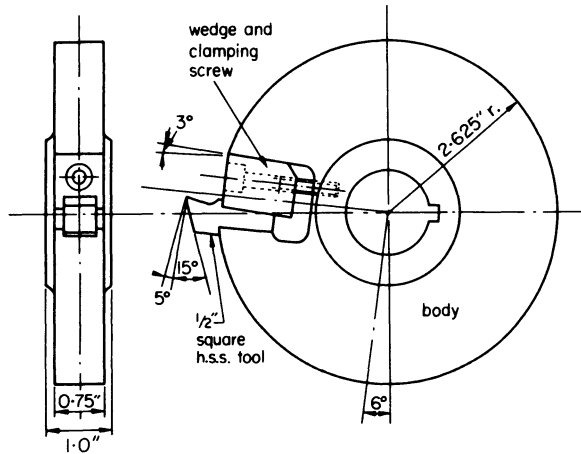


Figure 7. Single-tooth cutter.

Therefore the final experiment design was to use the single-tooth cutter shown in figure 7, with inserted teeth of similar material specification to the real cutter and having provision for varying the rake angle. The simulated workpiece was to be made from offcuts from the real work material and to be of such section that orthogonal cutting could take place. All tests were carried out on a Cincinnati plain milling machine, type 3P using Edgar Vaughn Ho-Cut 237 coolant. Wear measurements were made with a standard travelling microscope fitted with a 7x eyepiece and a scale reading to 0.0005 in.

The simulated path lengths were calculated by computer over a wide range of cutting depth for the standard feeds available on the Cincinnati machine. Observations were made of wear-land width over a range of rake angles for each path length relating to the real cutter.

EXPERIMENTAL RESULTS

A typical set of results for tool flank wear plotted against rubbed length over a range of rake angles are shown in figure 8. It is clearly seen that for a varying rake angle the non-dimensional wear rate dB/dL is not constant over the range of rake angles, and further that the end of useful tool life also occurs at different values. (The criterion for the latter being a

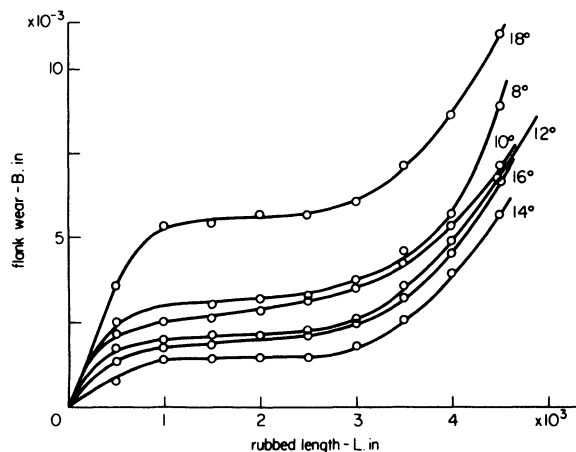


Figure 8. Non-dimensional wear curves for variable rake angles.

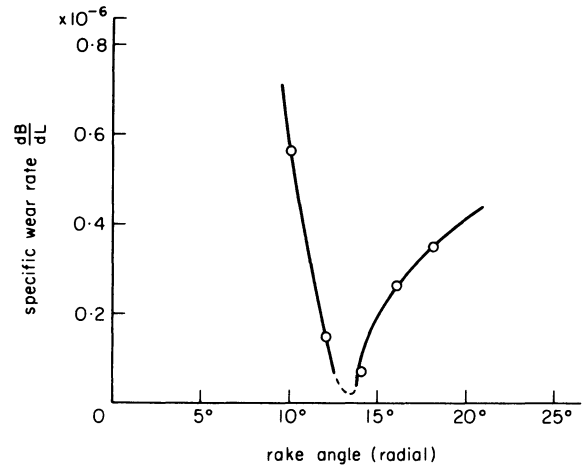


Figure 9. Variation of wear-rate with rake angles.

marked change in the slope of the wear curve, i.e. rapid deterioration of the cutting edge.) These results also show that for the group of rake angles 12° , 14° , 16° , the wear rates are fairly similar for both the constant and varying portions of the individual wear curves, while the other rake angles show quite significant variation. The 8° and 18° tools show a much greater rate of wear in the tertiary portion of their particular wear curves.

In order to examine the relationship between rake angle and wear-rate the non-dimensionalized wear-rate dB/dL for the constant (secondary) section of each wear curve was plotted against rake angle. These results are shown in figure 9 and indicate that for the particular tool-workpiece combination used, there was a tool rake angle which was associated with minimum wear-rate (non-dimensional). This rake angle was approximately 13° .

DISCUSSION OF RESULTS

The tool-life problem forming this discussion had two essential elements, namely, the wear characteristics of the cutter and those problems associated with the design and manufacture of the cutter.

Analysis will, therefore, be carried out separately under these headings.

Analysis of wear characteristics

By non-dimensionalizing the wear rate it was possible to investigate the complex wear problems associated with the production cutter and further, it was shown that although P.E.R.A.² had suggested no relationship existed between rake angle and tool life; for the high-speed steel/stainless iron combination used in these tests, it was shown that a relationship did exist for flank wear. Sabberwall and Fleisher³ had in fact suggested such a relationship in face milling of EN28 alloy steel and had found a rake angle of approximately 27° associated with minimum wear. (For a more detailed discussion of both specific wear-rate and rake angle see Etheridge and Hsu⁶.)

If the results shown in figure 6 for the production cutter's wear behaviour are examined, maximum flank wear occurred between sections B and C, where the rake angle was about 12° , minimum flank wear was found at section F with a rake angle of 18.75° .

These sections related to average depths of cut of 0.97 in and 0.15 in respectively, and cutter radii of 3.55 in and 2.70 in. The percentage differences were 24 per cent for velocity, 84.5 per cent for depth of cut and 66.4 per cent for the path lengths at these particular sections. Maximum wear on the production cutter occurred therefore at the maximum radius with a rake angle approximating to the optimum value suggested by the controlled test results. The magnitude of the flank-wear land on the real cutter corresponded to the end value found in the simulated trials for the 12° rake angle, thus suggesting the tip wear on the form cutter was consistent with the end of useful life. The flank wear at section D was of the order of 0.007 in (average) with a measured normal rake angle of 14° and a shorter path-contact length. The simulated tests suggested that the end of useful tool life was nearer to 0.005 in for this rake angle. It was clear from these tests that the variations in rake-face geometry was in a large part contributory to the low production figures obtaining before cutter failure. An examination of the cutter's design and manufacture was clearly necessary.

Design analysis

The design of the production cutters called for a helically gashed flute of 80 in lead and a helix angle of $16^\circ 25'$ based on the outside diameter of 7.5 in. The gashing cutter was to be the double-angle disc cutter shown in figure 2, with a nose radius of 0.130 in, and the finished flank was to be ground with a dish wheel on a special-purpose flute grinder to 10° rake angle.

These manufacturing processes create difficulties in maintaining a specified form on the flank of the helical slot owing to interference between the cutter and workpiece. In gear manufacture great efforts are made to eliminate such interference by designing cutters for the production of involute helicoids which have conjugate actions and tangential contact. (Dudley and Poritsky⁷.)

In tool and cutter manufacture, however, the large diversity of cutter types usually precludes individual machining techniques, and a compromise is usually made. Typical shapes produced in milling helical slots, together with an approximate analysis for a working profile are given by Etheridge⁸; it is also shown that the helix setting angle as well as cutter geometry has some bearing on the final flank profile.

In order to minimize the interference problem the tool designer usually specifies a helix angle based on the outside diameter of the workpiece, since this tends to produce positive radial top rake and to offset the fluting cutter which again tends to give positive top rake. For the cutter considered in this paper it was clearly evident that the rake-face geometry was always positive but variable owing to the blade profile shape intersecting the helicoidal surface produced by the fluting cutter. Figure 10 shows the problem.

Clearly for the cutter under consideration the controlled tests indicated that an optimum rake angle of 13° was desirable, but the manufacturing parameters in fact produced rake angles varying between 12° and 18° . The problem, therefore, becomes one of deciding what setting parameters will produce a more nearly constant normal radial rake angle over the

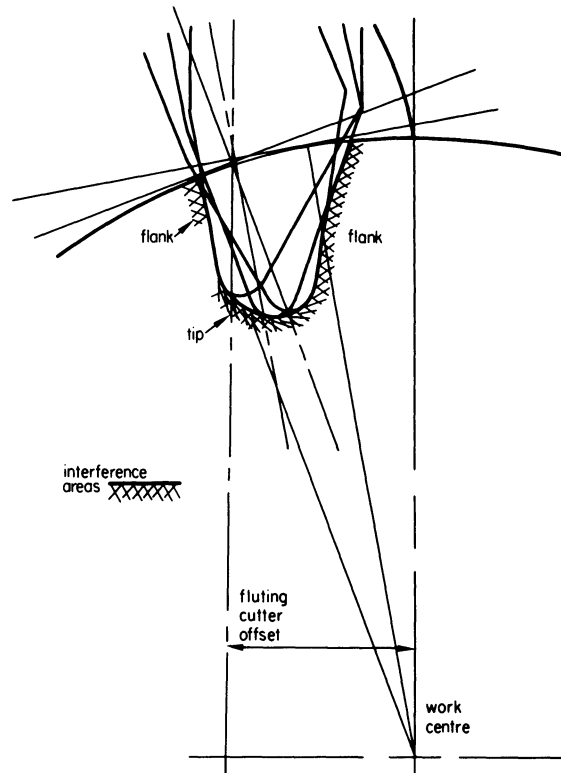


Figure 10. Schematic for fluting cutter movement.

cutter's profile. The method of doing this is difficult, but possible, by using vector differential geometry and the computer to solve the resulting equations for the helicoidal surface and the true normal radial rake at the intersection with the blade profile.

A method has been developed (Etheridge⁹) and is being used in further investigations into cutters of more complex shapes such as helically gashed conical cutters where the surface is related to either Archimedian or logarithmic spirals.

CONCLUSIONS

There are certain areas of tool engineering where the manufacturing procedures lead to problems, usually of early failure or, in the worst cases, of total failure as a cutting tool. This is almost always due to interference in the case of helically fluted cutters, and is highlighted by this particular case.

Clearly not all problems are what they appear at face value, and any steps towards a better understanding of the interference mechanism can only lead to improved tool design and tool life.

ACKNOWLEDGMENTS

The authors would like to thank all those at English Electric whose patience and co-operation helped in an understanding of the problem. In particular they thank the company for permission to publish.

REFERENCES

1. R. A. Etheridge. Machinability of LM2 alloy, University of Aston, Internal Report, 1973.

2. P.E.R.A. Milling alloy steel with high-speed steel cutters, Report no. 98, 1961.
3. A. J. P. Sabberwall and P. Fleisher. The effects of material and geometry on the wear characteristics of cutting tools during face milling. *Int. J. Mach. Tool Des. Res.* (1964) **4**, 47.
4. T. C. Hsu. A study of wear on cemented carbide cutting tools. *A.S.M.E. Paper No. 68-WA/Prod-9*, 1969.
5. M. E. Martellotti. An analysis of the milling process, Part 2.—Down milling. *Trans. A.S.M.E.* (1945) **67**, 233.
6. R. A. Etheridge and T. C. Hsu. The specific wear rate in cutting tools and its application to the assessment of machinability. *Annals of C.I.R.P.* (1970) XVIII, 107.
7. D. W. Dudley and W. Poritsky. On cutting and hobbing gears and worms. *A.S.M.E. J. Applied Mech.* (1943) **10**, no. 4, A139.
8. R. A. Etheridge. An analysis of the interference produced when milling a helical slot with disc type milling cutters. *Int. J. Mach. Tool Des. Res.* (1970) **10**, 143.
9. R. A. Etheridge. Unpublished work, 1973.

ULTRA-HIGH-SPEED MACHINING: NOTES ON METAL CUTTING AT SPEEDS UP TO 7300 Ft/s

by

G. ARNDT*

SUMMARY

An earlier study of UHSM ($V > 500$ ft/s) by the author showed that metal cutting at such speeds should theoretically be beneficial for certain workpiece materials, and described experimental cuts in the range $1770 < V < 5200$ ft/s. In the present study this range is extended to 7300 ft/s by cutting aluminium with HSS tools at depths of cut from 0.003 to 0.070 in, and 0.250 in width. To do this the suitably modified ballistic test system at Monash University was used to its operational limits. It was found that the extreme dynamic effects at the tool tip cause severe chipping and/or fracture. However, good cuts are still achieved, as was the case in the previous tests, whereas chip analysis becomes increasingly difficult. These effects are discussed and some comments on the practical application of UHSM are given.

INTRODUCTION

The continuous improvement in cutting-tool materials and machine tools has given rise to higher and higher cutting speeds, and will no doubt continue to do so in the future. Apart from the obvious advantage of higher output rates, other beneficial effects may prevail if this trend is carried to its limits, that is if machining speeds well above those used presently are considered. Cutting speeds several times the speed of sound in air have already been achieved, and for this reason the term Ultra-high-speed Machining (UHSM) has been coined^{1,2}. One of the benefits of UHSM may be seen in the change in cutting mechanism caused by the high-strain-rate effects prevalent in the shear zone, which may under certain conditions result in cutting forces below those

at conventional cutting speeds, but more importantly in the ability to cut metals 'conventionally' which cannot be cut at all at low speeds³. Previous fundamental and experimental investigations into UHSM, described elsewhere^{3,4}, have shown the feasibility of cutting metals at such speeds although, mainly on account of the high dynamic forces present, it is unlikely that this technique will ever become a widely-used practical proposition. Nevertheless, a knowledge of the behaviour of the cutting process at such high speeds is useful not only for its own sake but also because insights gained during such investigations may lead to measures which can often be usefully applied to the more conventional metal-working processes. Since strain rates of up to 10^9 s⁻¹ (for example at a cutting speed of 8000 ft/s)³ are achievable during UHSM,

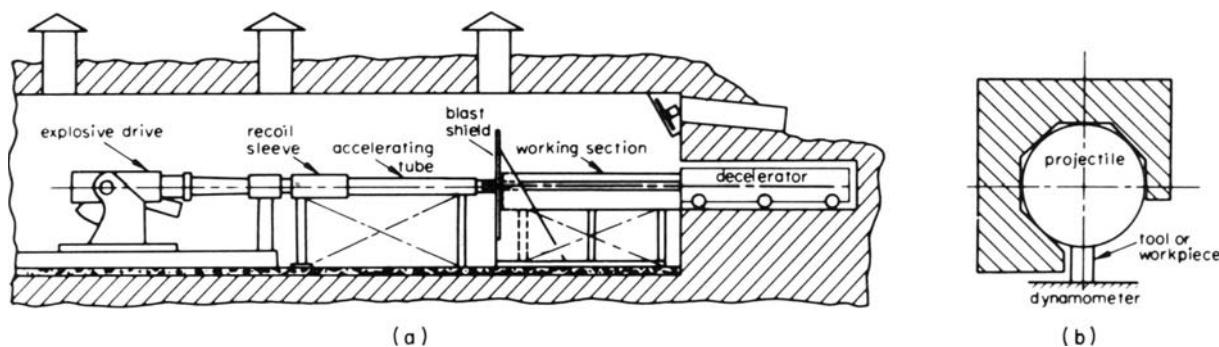


Figure 1. General arrangement of the ultra-high-speed test system at Monash University. (a) Bunker cross-section, (b) cross-section through the working section.

* Department of Mechanical Engineering, Monash University, Victoria, Australia

this process also constitutes a relatively simple way of studying materials behaviour at extreme rates of strain.

This paper presents some of the results obtained in a test series aimed at utilizing the full potential of the UHSM-test facility at Monash University. This facility uses a 40 mm Bofors gun as the drive unit, and is shown diagrammatically in figure 1. It has, together with some of the machining results obtained in earlier tests, already been described in some detail⁴, supplemented by the theoretical and experimental ballistic aspects associated with it^{5,6}.

EXPERIMENTAL SYSTEM

One of the means of increasing muzzle velocity in conventional gun systems is to increase the length of the gun barrel. In the previous test series⁴ (subsequently referred to as stage 1) the barrel length was 101 in and the bulk of the projectile velocities achieved was situated between 2000 and 4500 ft/s. The accelerating tube shown in figure 1 was then not yet used, but was incorporated into the present test series (stage 2), to give an overall barrel length of 180 in. This entailed a rearrangement of the working section and the projectile decelerator as shown in figures 2 and 3 in which the decelerator, originally resting on a wooden table, was transferred into a larger funnel where it could be opened and closed on site. Other modifications in the set-up of stage 2 relative to stage 1 were the redirection of the chip tube used for catching and decelerating the chips in a drum, (figure 2) and the use of eight pencil-lead contact probes to give two projectile (that is cutting) velocity readings before and two after the cut (figure 3). The firing conditions used in stage 2 were thus as follows:

barrel length	180 in
(smooth) bore diameter	1.625 in (nominal)
chamber volume	42 in ³
propellant	cordite type W/T 120-040
charge weights C	0.400 < C < 0.726 lb
projectile weights w	0.25 < w < 1.2 lb

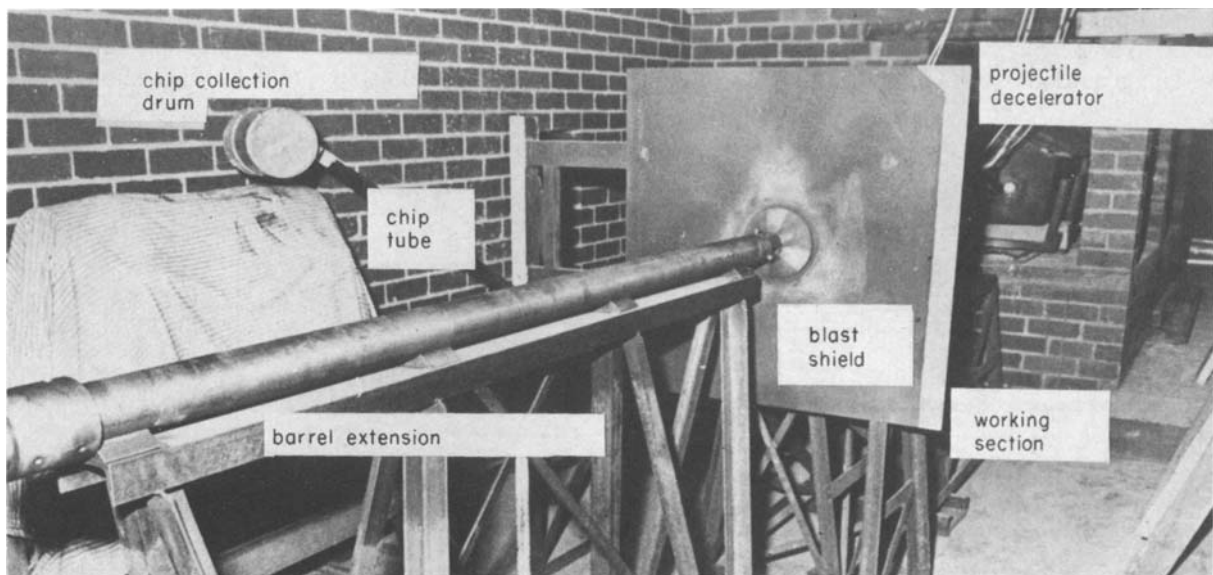


Figure 2. Stage 2—set-up showing extended barrel (accelerating tube), chip collection and decelerator arrangement.

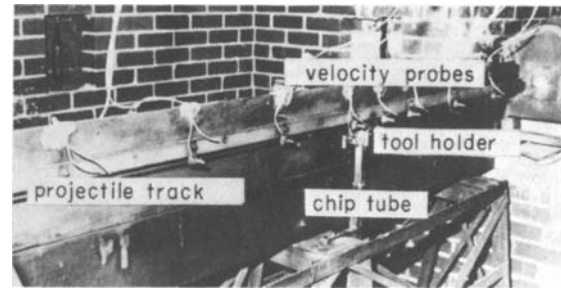


Figure 3. View of stage 2—working section (obscured by blast shield in figure 2).

As was concluded earlier⁴ there was no need to use the high charge weights predicted theoretically⁵ to achieve the required projectile velocities of 7000–8000 ft/s. Without the installation of massive blast or pressure-wave absorbers at the muzzle this would in fact be impossible, since even at the present maximum charge weight of $C = 0.726$ lb blast-wave conditions prevailed which bordered on structural damage of the bunker.

Based on previous experience the aluminium projectiles were mostly of the integral type. Length-to-diameter ratios less than one were dictated by the low projectile weights ($w = 0.25$ lb), but did not cause any special difficulties in tracking.

All other aspects of the experimental system were identical to those of stage 1.

RESULTS AND DISCUSSION

Cutting speed

Since the deceleration of the projectiles caused by frictional resistance (and, in fact, the cut itself) was too small to be detected by the impact-probe and electronic counter arrangement (accuracy ± 5 per cent), the projectile speed may also be taken as the cutting speed.

In figure 4 the cutting speeds attained in all shots of both stages 1 and 2 have been plotted. For the latter case the points shown represent the mean values of the four readings obtained per shot (full

points). The apparent scatter is caused by the differences in projectile weights and 'driving band' thicknesses. (The driving band is a ring of aluminium which must be sheared through before the projectile starts to move; in stage 2 this varied from 0.100 to 0.250 in.) Apart from showing the effects of barrel length and projectile weight on cutting speed, the graph illustrates the versatility of the test system.

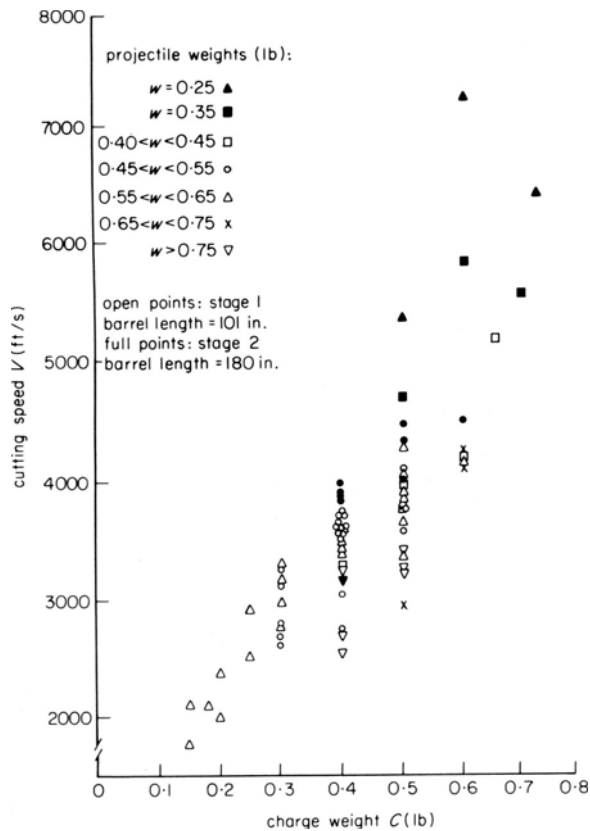


Figure 4. Variation of cutting speed with charge weight, projectile weight and barrel length, obtained on ballistic UHSM test system.

Speeds in the range 1770 to 7290 ft/s have been achieved. The lower limit could be further reduced by suitable adjustment of projectile and charge weights, but this was not the objective of the present study. Likewise, the upper limit could, from the viewpoint of the firing system, be increased by probably another 2000 ft/s, although to do so would require additional constructional features in the bunker. However, it is unlikely that such an increase would yield much more information concerning the cutting process than can be gained at the speeds already attainable.

Workpieces

All workpiece projectiles consisted of aluminium (composition: Cu 5.5, Pb 0.51, Bi 0.44, Fe 0.15, Si 0.08). Rectangular steps of $\frac{1}{4}$ in width were milled in to give true orthogonal cutting. For one shot (number 103) a steel insert was incorporated into the aluminium 'carrier', which separated on impact with the sand. The projectiles after recovery from the sand decelerator and the respective cutting speeds are shown in figure 5, together with one original blank as

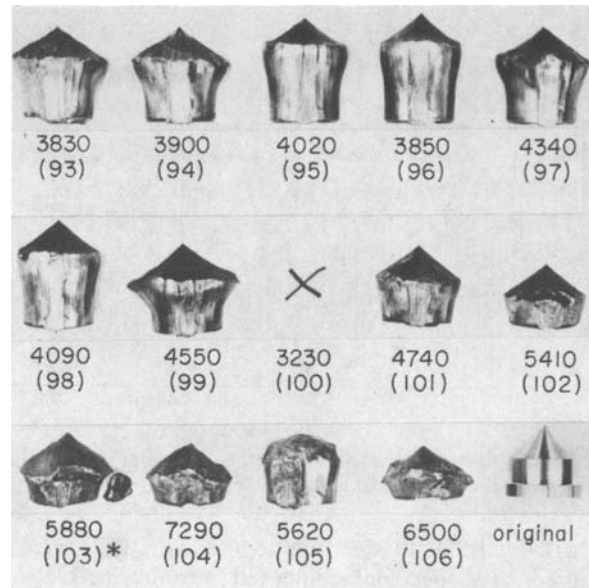


Figure 5. Projectile workpieces and cutting speeds (ft/s) for stage 2—tests. Shot numbers are given in brackets. No. 100: hollow projectile, shattered on impact with sand; no. 103: mild steel workpiece insert.

used in shots 102, 104 and 106. Despite the 'mushrooming' effect caused on impact with the sand (that is *after* the event of interest, namely the cut), it is seen that on the whole good, full-length cuts were achieved, although obviously some tools failed in the process of cutting. Some of the cuts are not aligned with the milled steps. This was attributed to a turning effect induced by uneven shearing of the driving band, and was remedied by the insertion of a new shear die into the explosive chamber.

Depth of cut

The nominal depths of cut, d_n , in stage 2 were as follows:

shots 93, 94: 0.010 in

shots 95–99, 101, 102: 0.005 in

shot 100: 0.007 in

shots 103–106: 0.003 in.

As can be seen from some of the projectiles in figure 5 the depth of cut increased with distance along the projectile. This digging-in effect has been described previously^{2,4}, and was found to be much more pronounced with the projectiles of stage 2 than those of stage 1. (This follows logically since the momentum force causing it is proportional to V^2 .) This effect may be expressed by the 'dig-in ratio', obtained by dividing the actual (mean) depth d_a of cut along the projectile by the nominal or set depth d_n , as shown in figure 6. All the points shown apply for one and the same type of tool-holder construction (shown later in figure 8), whose design, stiffness and support features determine the actual tool deflection during cutting to which the projectile deflection must be added. As indicated in figure 6 a rather continuous increase in dig-in ratio with cutting speed occurs for both the stage 1—and stage 2—results. Some of the scatter is explained by the different projectile lengths.

The whole concept of 'constant' or 'set' depth of

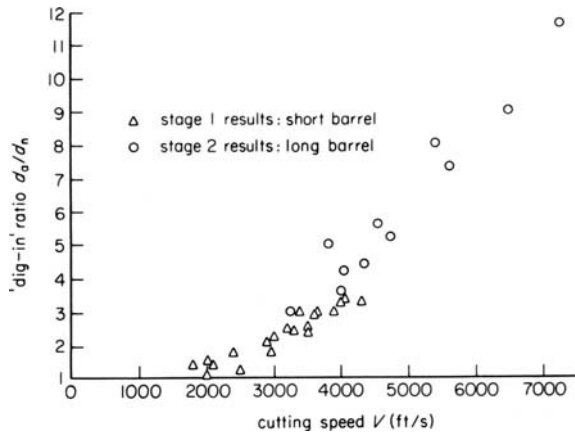


Figure 6. Effect of ultra-high cutting speeds on tool dig-in. (for actual depths of cut see text.)

cut in UHSM, at least for the tool-workpiece system used here, becomes somewhat meaningless on the basis of figure 6, since depth of cut effectively becomes a function of cutting speed. This of course has practical implications in that machining in its true sense, that is the achievement of a defined geometry of cut, will only be possible at ultra-high speeds either if workpiece and tool-deflections can be prevented, which implies an extremely rigid construction using high-strength materials, or if dynamic effects can be balanced out, for example by simultaneously cutting with two or more opposed cutting edges, resulting in the mutual cancellation of the momentum forces.

Cutting tools

All cuts in stage 2 were orthogonal, and performed by HSS tools whose cutting edges exceeded the width of

cut by $\frac{1}{8}$ in. each side. Rake angle was kept constant at 0° (except for shot 106, with $\alpha = -20^\circ$), clearance angle at 3° . Some of the tools after cutting are shown in figure 7, together with cutting speed and actual (mean) depth of cut values. The latter were obtained from

$$d_m = (d_n + d_e)/2$$

where d_n is the nominal depth of cut set and d_e is the actual depth of cut measured at the end of the projectile. Figure 7 should be observed in conjunction with figure 8 showing the tool holder (and the

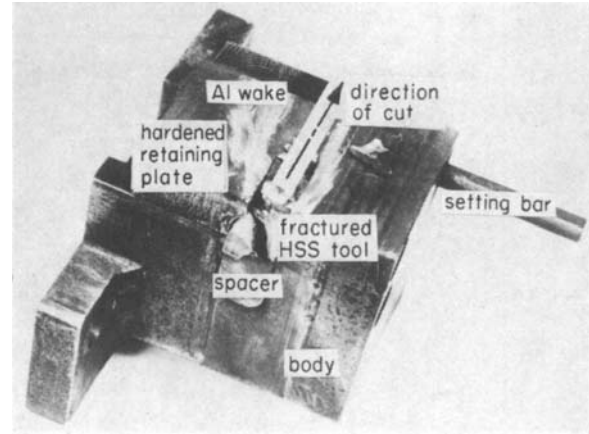


Figure 8. Tool-holder assembly (after shot 106).

remains of the negative-rake tool used in shot 106) in more detail than is visible in figure 3: the tools themselves had a close fit in their seat in the hardened steel retaining plate, through which they were also

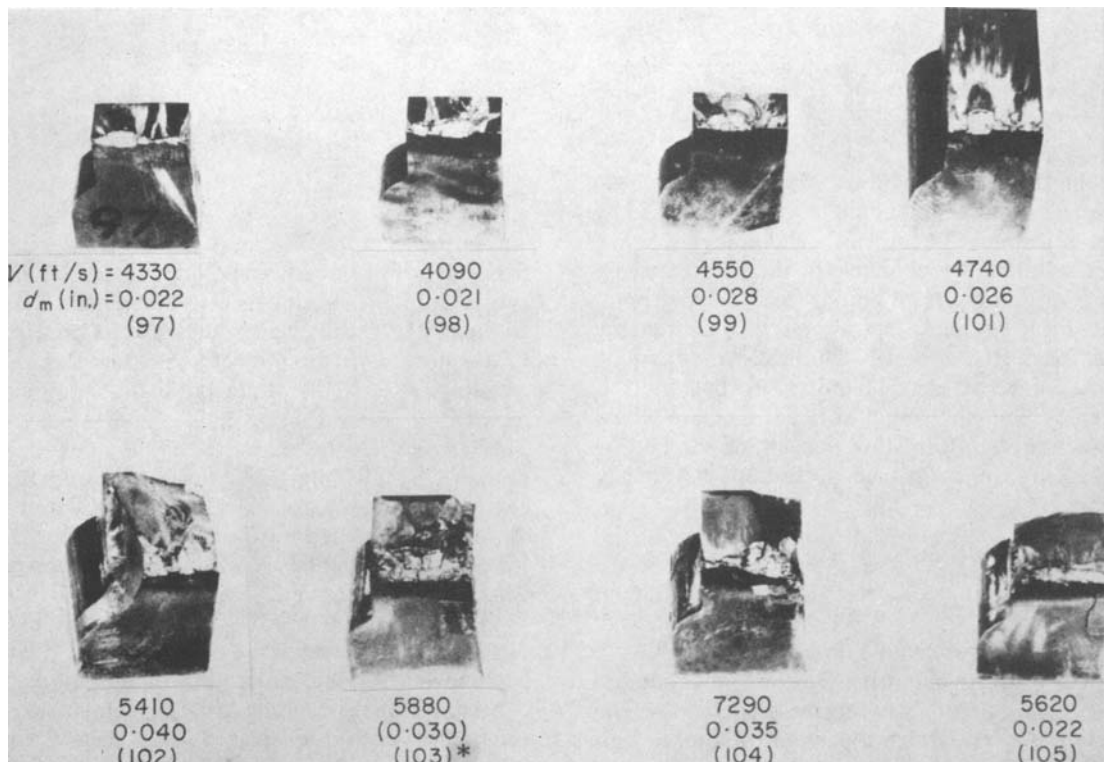


Figure 7. Stage 2—cutting tools. V is the cutting speed, d_m the average depth of cut over length of projectile, and the numbers in brackets are the shot numbers. No. 103: cut steel; all others aluminium.

clamped in position. Depth of cut was set by moving the entire holder up or down.

As can be seen from figures 7 and 8, most of the cuts resulted in severely chipped cutting edges ('lower' speeds: shots 97-101), while some cuts also led to complete fracture of the tool ('higher' speeds: shots 102-106). This is to be expected firstly because of the impact loading (more pronounced for steel, shot 103, than for aluminium), and secondly because the momentum force acts entirely at the tool tip (whereas the resultant 'mechanical' force, which at these speeds becomes negligible in comparison with this dynamic force, in fact acts further down the tool rake face). The tool tip thus becomes very prone to fracture, even at negative rake angles. This effect could probably be counteracted by providing the cutting edge with a considerable nose radius.

Two more features in figures 7 and 8 are of interest—the dark discolourations at the rake (front) faces, and the light discolourations at the clearance (top) faces of the tools. The former represents the chip-tool contact length l_i , which in the present tests resulted in contact ratios d_m/l_i between 0.1 and 0.3. The latter are traces of molten aluminium which in fact extended far across the tool holder, as shown in figure 8. They explain why in some cases no chips at all were found after the cut, and in other cases far less than were to be expected from the volume of material removed from the workpiece. At speeds around 6000-7000 ft/s, one gram of aluminium in fact has a kinetic energy of 2×10^{10} erg or 480 cal, which is twice that needed to melt it from 25°C⁷.

Chips

The proven method of catching the chips relatively undamaged by means of a long tube leading directly from the cutting edge to a spinning drum⁴ was still found to be adequate. The chips thus obtained are

shown in figure 9. With the exception of shot 104 (7290 ft/s), chips as such were only detected at cutting speeds below 5000 ft/s. For shots 102, 3, 5 and 6, with speeds from 5410 to 6500 ft/s, one must conclude that the considerable volume of workpiece material removed during projectile-tool impact 'disappeared' both due to melting (or vaporization) as well as possibly sideways scattering of the cut material during impact (not to be confused with chip side flow). Since the melting process takes a finite time it is unlikely that all of the chip material melts during the short projectile-tool encounter in UHSM. This also explains why a considerable quantity of chips was collected at the lower speeds.

The chips shown in figure 8 all appear to be discontinuous. However, closer inspection indicates that they must have left the tool as a substantially continuous chip which was then broken up during its travel to, or within, the collection drum. This is substantiated by the long chip 'slivers', found for example for shots 93-98 (3830-4330 ft/s). Some of these even exhibited the characteristic ripple formation at the free chip surface, indicating the usual shear cutting mechanism.

It appears from these tests that in the cutting speed range 4500-5500 ft/s, gross chip melting/scattering commences, with some solid chip material still present (shots 99, 101), while from 5500 ft/s upward the former mechanism and hydrodynamic effects predominate. The fact that these statements cannot be taken as conclusive is stressed by the solid chip formed at 7290 ft/s in shot 104 (which could more aptly be described as a 'fragment of the projectile'). Definitive statements on chip length (or thickness) can similarly not be made, although by inspection the order of cutting ratio observed previously (r_c values up to 3)⁴ were confirmed. In the limit, the concept of cutting ratio will in any event become somewhat meaningless,

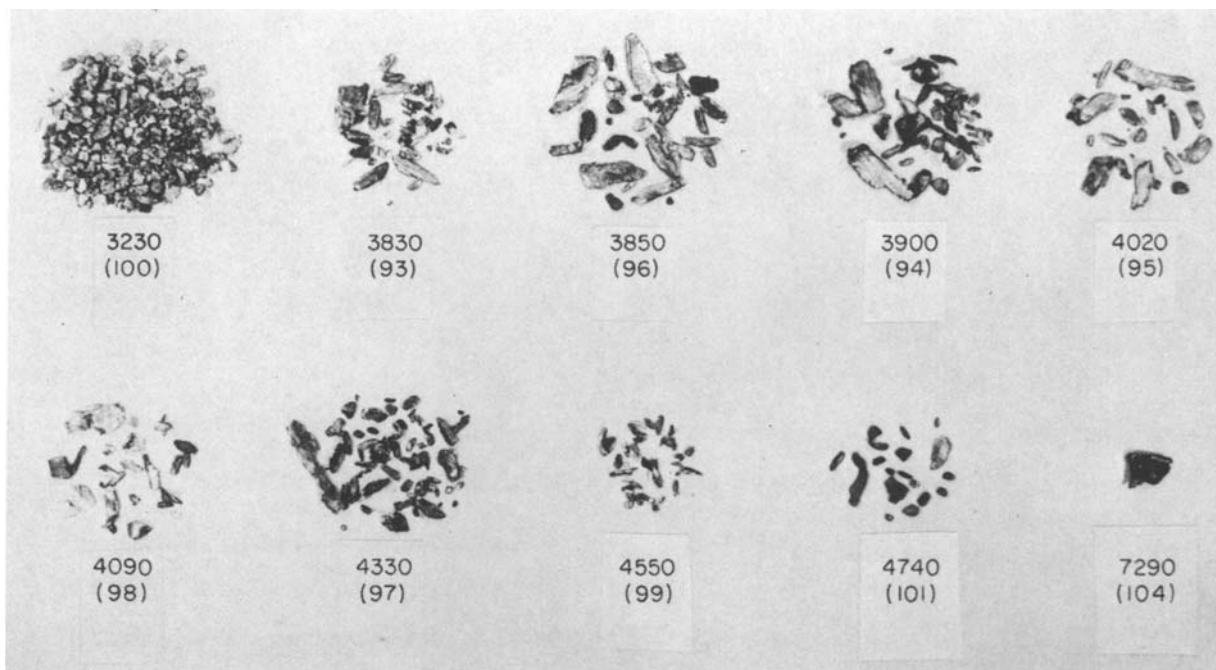


Figure 9. Stage 2—chips. The cutting speeds are shown in ft/s. Shot numbers are given in brackets.

namely at deformation speeds where chips as such no longer exist.

CONCLUSIONS

The results of a brief series of machining tests with cutting speeds from 3830 to 7290 ft/s were described. It was shown that cutting at and above the onset of the hydrodynamic region is possible with conventional HSS tools although severe tool wear occurs which, as expected, is more attributable to shock than to temperature effects. Chip analysis became increasingly difficult with increasing cutting speed, but originally-continuous chips were obtained at speeds up to 5000 ft/s. Thereafter the proportion of chip material lost on account of melting or scattering increased rapidly. Severe dig-in of the tool during cutting, caused by the high momentum force acting at the tool tip, was found to be a major problem during UHSM since it effectively leads to the situation where the effective depth of cut is a function of cutting speed.

It is suggested that both the problems of tool wear and dig-in could be alleviated by constructional measures, such as cutting with multiple cutting edges arranged so that they mutually balance out each other's dynamic forces, and the provision of a controlled tool nose radius. This effectively means that the use of single-edge cutting at 'ultra-high' speeds is unlikely ever to become a practical proposition, whereas the application of UHSM to such operations as piercing, blanking and spline cutting would seem possible. The process should then perhaps better be described as 'controlled ballistic penetration', in which context the large amount of knowledge on this subject obtained in investigations of a military kind could then, and

perhaps more usefully, be applied to the industrial scene. Further work along these lines would seem a fruitful undertaking.

ACKNOWLEDGMENTS

The assistance of Mr R. R. Peach in the tests and the financial support of this project by the Harold Armstrong Memorial Fund is acknowledged. Special thanks are also due to the A.-v.-Humboldt Foundation, who enabled the author's personal presentation of this paper during his stay at Stuttgart University as a Humboldt Research Fellow.

REFERENCES

1. R. L. Vaughn. Ultra-high-speed machining. *AMC technical Report* 60-7-635(1), 1960.
2. G. Arndt. Ballistically induced ultra-high-speed machining. *Ph.D. Thesis*, Monash University, Melbourne, Australia, 1971.
3. G. Arndt. Ultra-high-speed machining: a review and an analysis of cutting forces. *Proc. Inst. Mech. Eng.* (1973) **187**, 625-634.
4. G. Arndt and R. H. Brown. Design and preliminary results from an experimental machine tool cutting metals at up to 8000 feet per second. *Proc. 13th Int. MTDR Conf.*, Macmillan, 1972, pp. 217-223.
5. G. Arndt and J. T. McHenry (1970). A computerised internal ballistic analysis of conventional gun systems with muzzle velocities of up to 8000 ft/s. *Explosivstoffe* (1970) **18**, 253-267.
6. G. Arndt. The construction of, and experimental ballistic results from, a high-velocity gun system used for metal cutting. Accepted for publication in *Explosivstoffe* (1973).
7. R. H. Wentorf, Jr. Private communication.

ON LINE MEASUREMENT OF BENDING AND TORQUE IN MILLING

by

R. BEDINI* and P. C. PINOTTI†

SUMMARY

In order to realize a constraints-based adaptive control of the cutting process, an experimental investigation on bending and torque in milling was carried out on a numerically controlled milling machine.

A mini-computer operating in real-time and connected to the modified control unit of the machine-tool sets a wide range of spindle speeds and feeds while machining various materials at different depths and widths of cut. The signals received from bending and torque transducers are digitalized for each machining condition and statistically processed.

The relationships obtained between bending and torque and the cutting conditions show good agreement with those already available in the technical literature.

The results, moreover, confirm the reliability of the whole system, which has proven to be a powerful tool for the acquisition of cutting data.

NOTATION

M_b	Bending moment acting on the tool.
M_t	Torque.
W	Power.
h_b, h_t, h_w	Constants in cutting relationships, equation (1).
v_f	Feedrate.
n	Spindle speed.
x	Feedrate exponents.
y	Spindle-speed exponents.

INTRODUCTION

At the Centro per l'Automatica of Pisa University a research devoted to the study and development of a constraints-based adaptive control of milling machines using either a real-time mini-computer or a suitable A/C unit connected to the N.C. system is being carried out. In this paper the results of the experimental work in the field of on-line measurement of the mechanical stresses, torque, bending, and power acting on the tool in rough milling operations are reported. A test installation consisting of a digital process computer, operating in real-time, a machining centre and of an appropriate interface equipment was set up^{1,2}.

The use of the computer greatly simplifies the testing of the control algorithms which, being software implemented, can be easily modified and

improved in their performances simply by changing the programs. Moreover the computer can be used for collection and processing of cutting data.

In order to define an adaptive control strategy, an experimental work for determining the characteristics

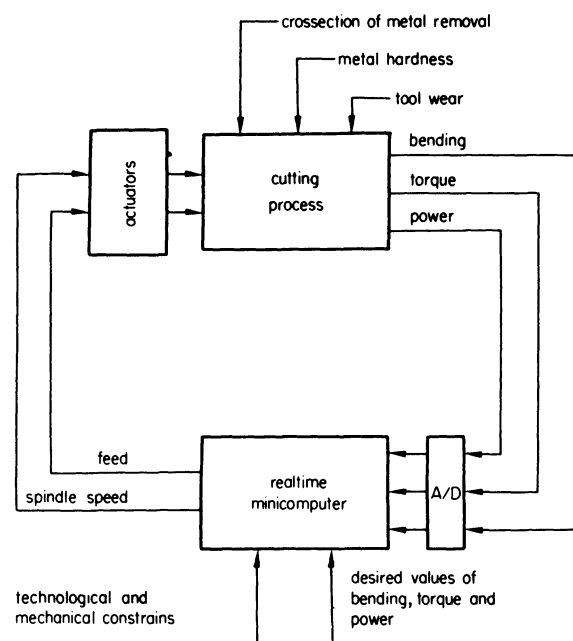


Figure 1. Adaptive control loop.

* Istituto di Ingegneria Meccanica, Università di Firenze, Italy

† Centro per l'Automatica 'E. Piaggio', Università di Pisa, Italy

of the milling process, considered as the controlled process (figure 1), was necessary. From this point of view the tests described in this paper had mainly the aim of investigating the relationships existing between the input variables, feed-rate and spindle speed, and the outcoming bending, torque and power for constant values of the secondary inputs as cross-section of metal removal and workpiece characteristics. The results obtained have confirmed that for describing the milling process of various materials (cast iron, aluminium, steel) the following relationships can be adopted:

$$\begin{aligned} M_b &= h_b v_f^x b n^y \\ M_t &= h_t v_f^x t n^y \\ W &= h_w v_f^x w n^y \end{aligned} \quad (1)$$

The knowledge of the above relationships, i.e. of the shape of the constraints in the control plane, v_f , n , (figure 2) made it possible to develop a constraints-based optimization strategy defining the direction and magnitude of the commands, during milling, of both feedrate and spindle speed in order to obtain the desired bending, torque and power allowing the cutting process to attain the optimum region^{3,4}.

Basically the adaptive controller, either software or hardware implemented, samples the actual values of the bending, torque and power, compares them with those which are assumed to be the optimal ones for the best machine utilization and, taking into account all the technological and mechanical constraints, changes the feed-rate and spindle speed in order to keep the bending, torque and power at the desired values.

EXPERIMENTAL SET UP

Machine tool

The machine tool used is an Olivetti numerically controlled machining centre capable of milling, drilling, boring and tapping operations.

Process computer

The computer, operating beside the machine tool, is a digital PDP11/40 with 16K, 16 bits words of core memory, 900 ns cycle time. The peripherals include a 1.2-million-word cartridge disk, a teleprinter, a high-speed paper-tape reader and punch and a cabinet for the analogue and digital interfacing with the process.

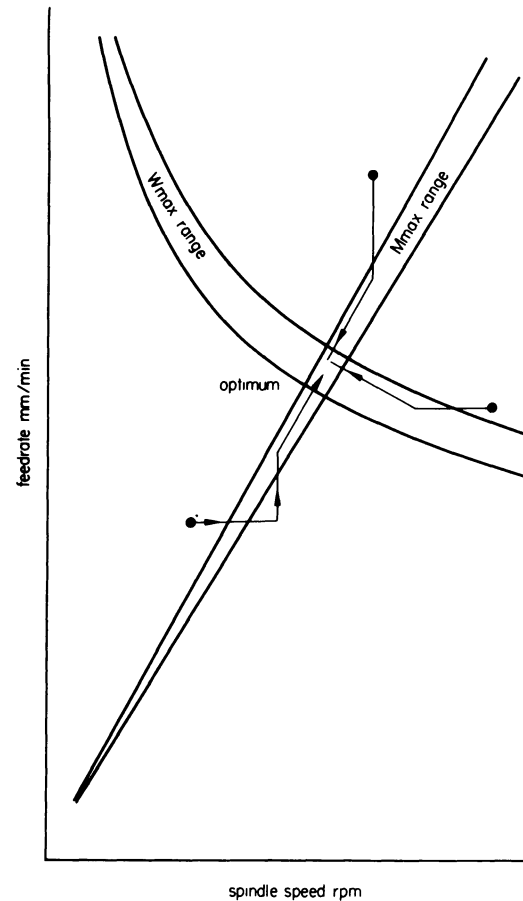


Figure 2. Adaptive control trajectories.

Computer-machine-tool interface

The control unit of the machine tool was modified in order that it could accept the digital outputs of the process computer providing independent variations of spindle speed and feedrate. In order to perform feedrate variations the computer has been interfaced to the existing feedrate override of the N/C unit. In this way sixteen feedrates ranging from 1/16 to 16/16 of the maximum selected feedrate during cutting were available.

In figure 3 the control system of the d.c. spindle motor is schematically shown. This control uses five on/off outputs of the computer for performing thirty-two different speeds on each of the two spindle-speed ranges available on the machine.

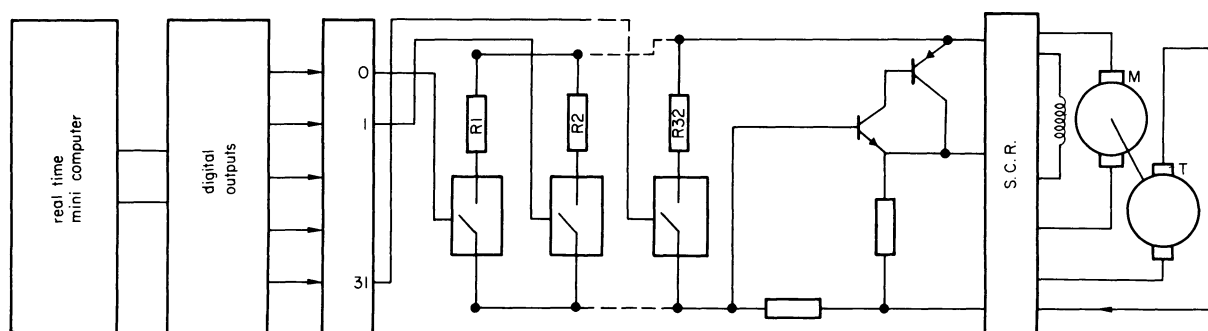


Figure 3. Control system of d.c. spindle motor.

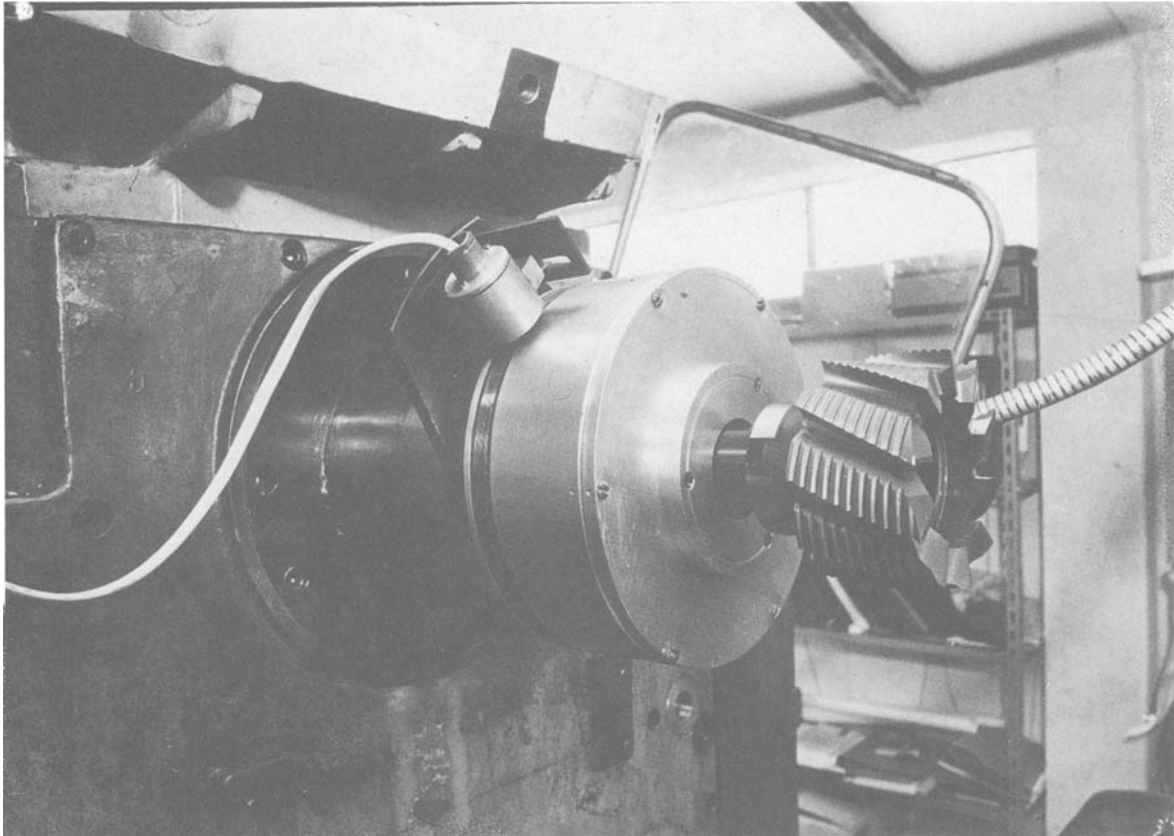


Figure 4. Bending transducer.

Milling process-computer interface

The measurement of the bending moment acting on the cutter was done by using wire strain-gauges, inserted into the four arms of a bridge and mounted on the tool-holder⁵. The oscillator rotates with the tool-holder itself and a frequency modulated unbalance signal is transmitted, through an inductive coupling, to a demodulator (figure 4).

In figure 5 a typical signal coming from the demodulator can be seen. The amplitude of this signal is proportional to the bending moment acting on the cross-section at which the strain-gauges are mounted. Since the cutter rotates, the effective value of bending can be measured twice per revolution. For this reason an instrument capable of detecting and holding the peaks was designed and set up (figure 6).

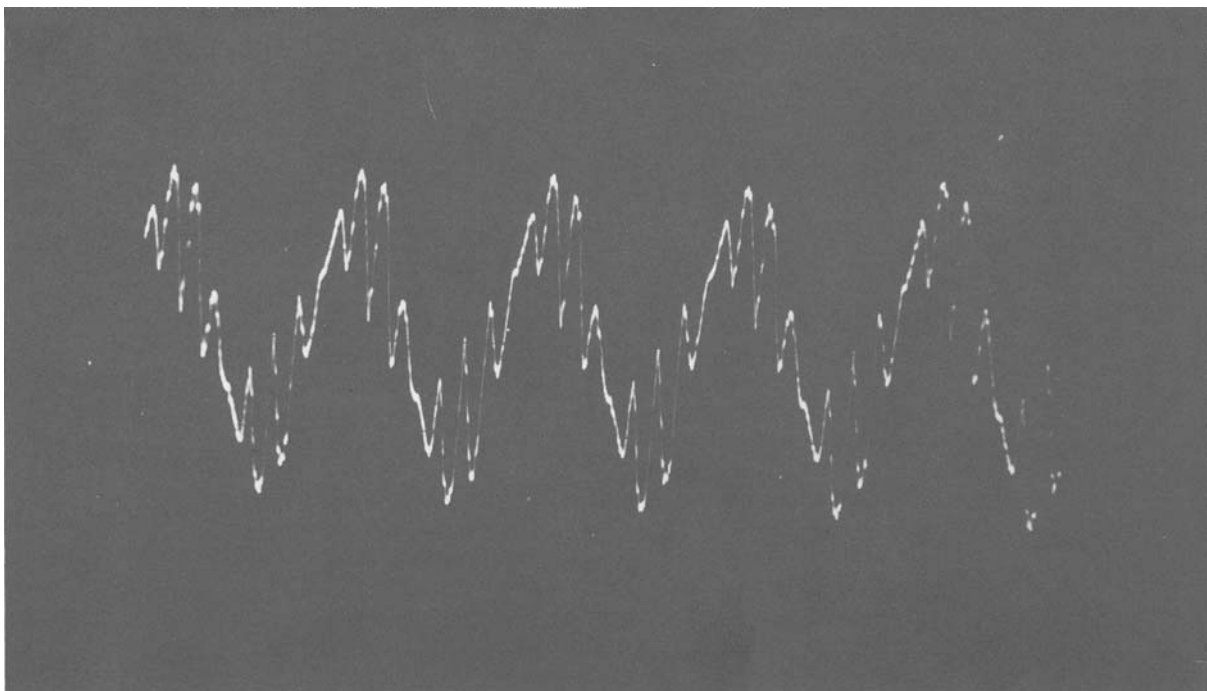


Figure 5. Bending signal.

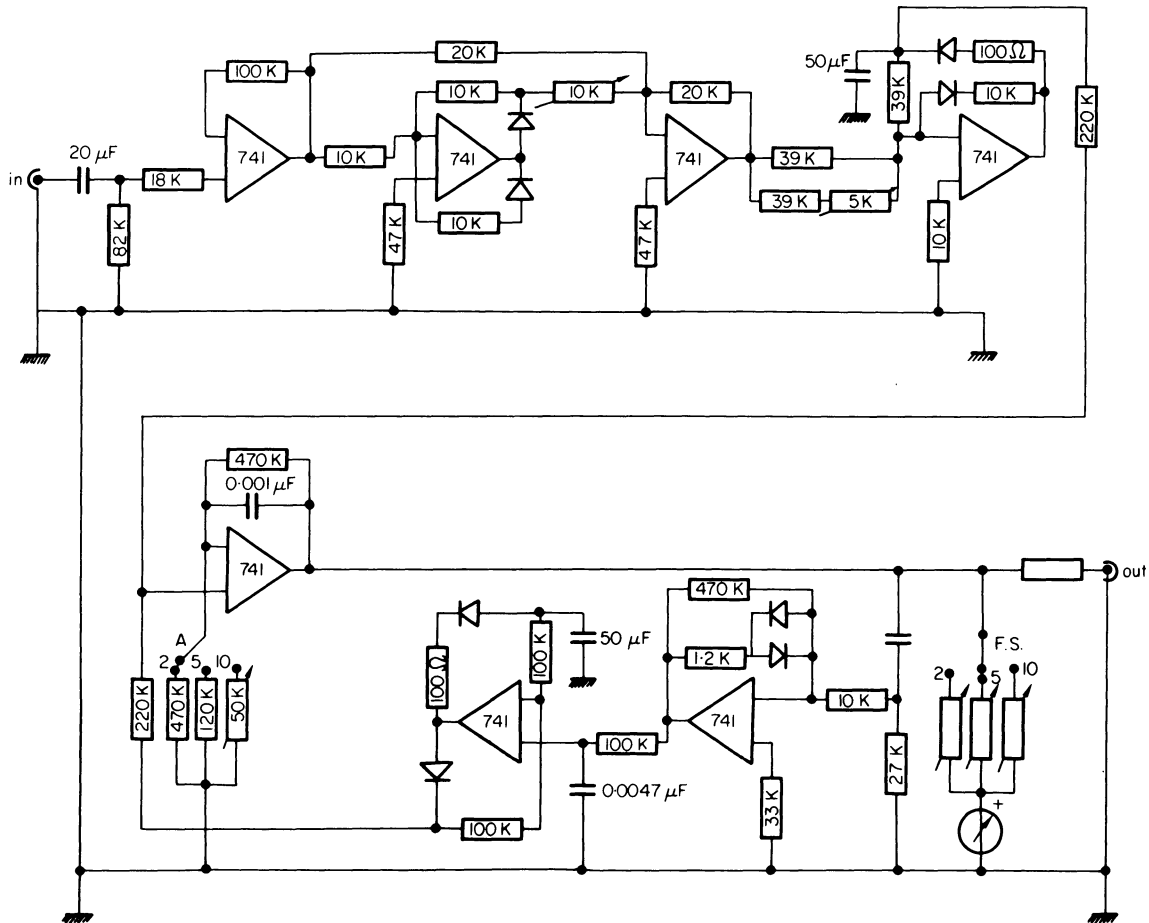


Figure 6. Peak detector.

The information is updated at each peak and maintained up to the next one. The output of the instrument was fed into an analogue input of the computer. The static performances and the linearity of the whole instrumentation were satisfactory for the data acquisition tests carried out. This bending measuring device does not reduce the stiffness of the machine and, moreover, could be used as a torque-meter, changing the strain-gauges connection in the bridge.

The torque measurement was performed using a Vibrometer TG10 torquemeter mounted on the shaft of the d.c. spindle motor. Because the torquemeter is remotely located from the cutter (four gears are also interposed) cutting torque dynamics cannot be accurately determined. In other words the shafts and the gears act as a mechanical filter on the signal. This fact does not, however, present significant problems, because in most adaptive control schemes only the mean value of the torque which is directly correlated to the current and heating of the d.c. motor are of interest⁶.

COMPUTER PROGRAMMING

Computer programming for the process computer-machine-tool installation was in two major areas: data collection programs and adaptive control routines. All programming was done in a 'building block' fashion; separate subroutines have been written for each function required and then added to the computer library.

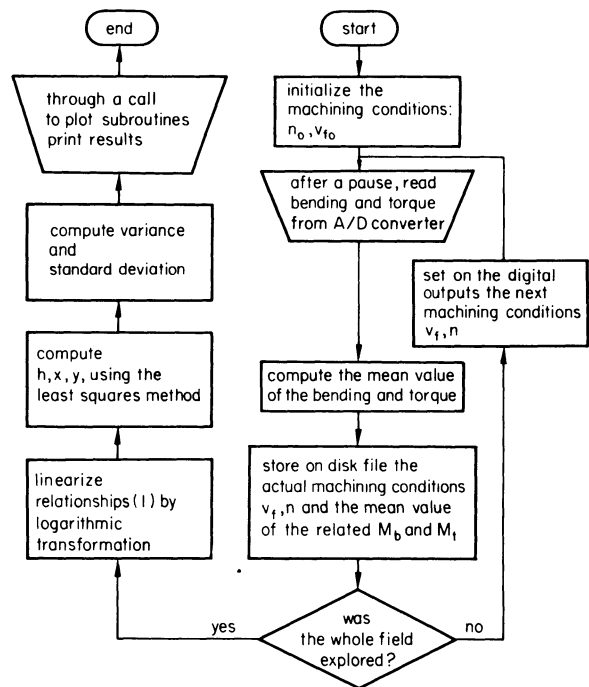


Figure 7. Flow chart of data collection and processing program.

In the area of data collection the basic functions to be performed by the computer were as follows: tests description, machine-tool driving, data sampling,

permanent data storage on disc files, data retrieval from the files, data tabulation in printed form after statistical elaborations, and final data plotting.

Programs have been written under Digital PDP11 Disc Operating System using Fortran IV. For peripherals interfacing Macroassembler has been used.

The main program (figure 7) sequentially sets in a prefixed range all possible values of feedrate and spindle speed while cutting. For each working condition several samples of the bending and torque signals are read and stored. The processor computes the cutting power from the torque and the set spindle speed. As soon as all the cutting conditions, for the given material at constant depth and width of cut, have been considered, the program processes the stored data using the multiple regression analysis. The relationships given in Equation (1) have been brought

into linear form by logarithmic transformation. The statistical results have demonstrated the validity of the above relationships for describing the laws of cutting.

EXPERIMENTAL PROCEDURE AND RESULTS

Several cutting tests have been carried out on steels (UNIC40, Fe 42 UNI 5334), cast iron (G20 UNI 5007) and aluminium (Al UNI 3044).

Each material has been machined at various depths of cut and tools having different amounts of wear have been used in order to determine the influence of these parameters on the constants and the exponents of the relationships (1).

In Table 1 the results of some of the tests performed are shown.

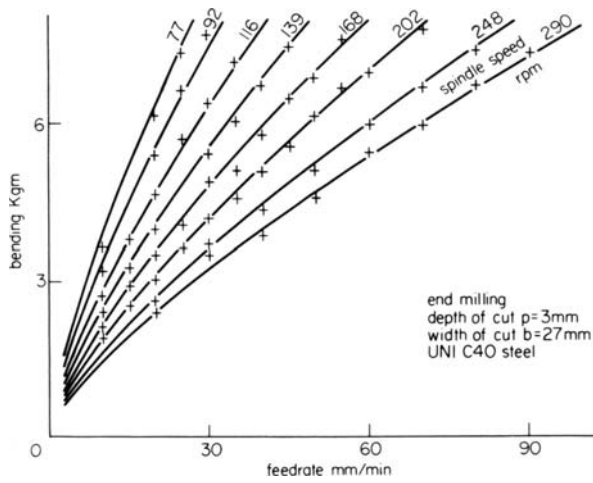


Figure 8. Experimental results (cutter: $\Phi 50$ mm, 6 teeth).

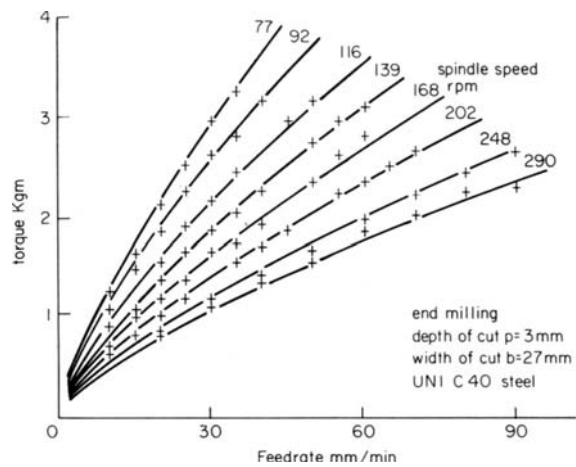


Figure 9. Experimental results (cutter: $\Phi 50$ mm, 6 teeth).

TABLE 1

	N	p	h_b	x_b	y_b	$\sigma_b\%$	h_t	x_t	y_t	$\sigma_t\%$	Range v_f	Range n
<i>Cutter</i> $\Phi 50$ 6 teeth Fe42 UNI5334	143	4	31	0.62	-0.72	0.16	4.0	0.40	-0.40	0.26	30 ÷ 90	103 ÷ 333
<i>Worn cutter</i> $\Phi 50$ 6 teeth Fe42 UNI5334	105	4	31	0.44	-0.58	0.61	3.4	0.25	-0.18	0.36	10 ÷ 90	81 ÷ 254
<i>Cutter</i> $\Phi 50$ 6 teeth C40 UNI2954	58	4	22	0.77	-0.80	0.24	4.3	0.76	-0.80	0.21	10 ÷ 70	77 ÷ 290
<i>Worn cutter</i> $\Phi 50$ 6 teeth C40 UNI2954	55	4	29	0.58	-0.71	0.54	4.5	0.55	-0.63	0.58	10 ÷ 80	77 ÷ 290
<i>Cutter</i> $\Phi 50$ 6 teeth C40 UNI2954	48	6	23	0.79	-0.78	0.23	5.6	0.75	-0.78	0.25	10 ÷ 70	77 ÷ 290
<i>Worn cutter</i> $\Phi 50$ 6 teeth C40 UNI2954	36	6	34	0.58	-0.69	0.75	6.1	0.61	-0.68	0.44	10 ÷ 60	77 ÷ 290
<i>Cutter</i> $\Phi 50$ 6 teeth G20 UNI5007	31	3	37	0.93	-1.10	0.43	3.2	0.57	-0.58	0.18	40 ÷ 90	57 ÷ 115
<i>Cutter</i> $\Phi 80$ 10 teeth AL UNI3044	49	6	22	0.78	-0.96	0.17	0.9	0.58	-0.41	0.15	150 ÷ 240	299 ÷ 368

N = number of cutting conditions for each test.
 p = depth of cut.
 σ = standard deviation.

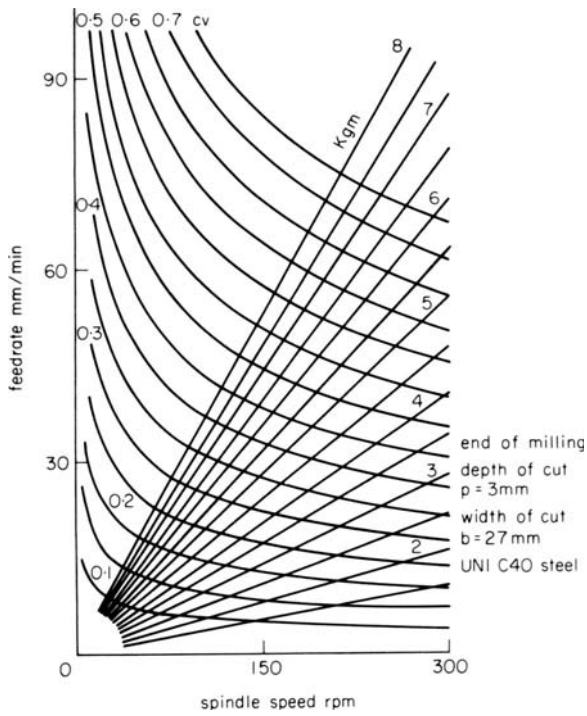


Figure 10. Experimental results (cutter: $\Phi 50$ mm, 6 teeth).

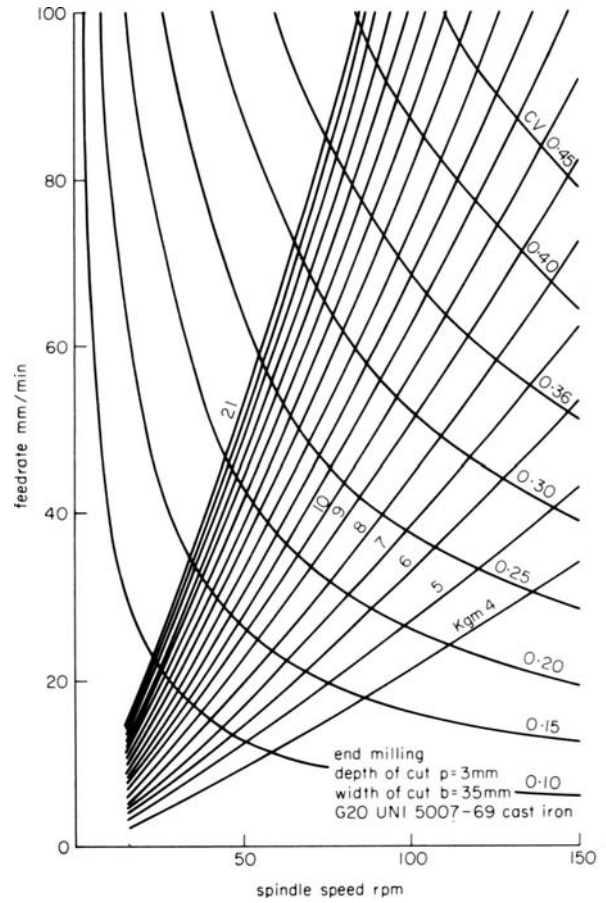


Figure 12. Experimental results (cutter: $\Phi 50$ mm, 6 teeth).

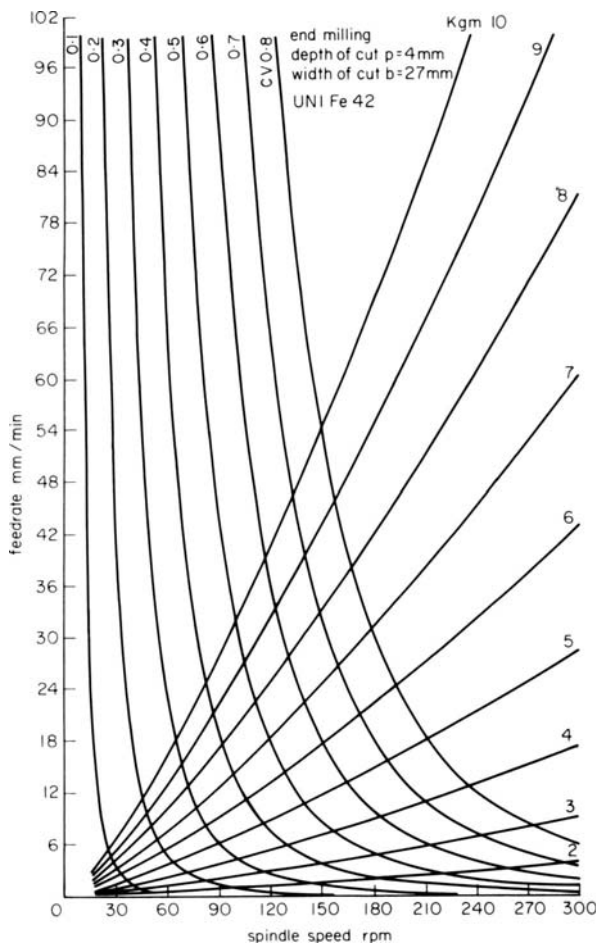


Figure 11. Experimental results with worn cutter ($\Phi 50$ mm, 6 teeth).

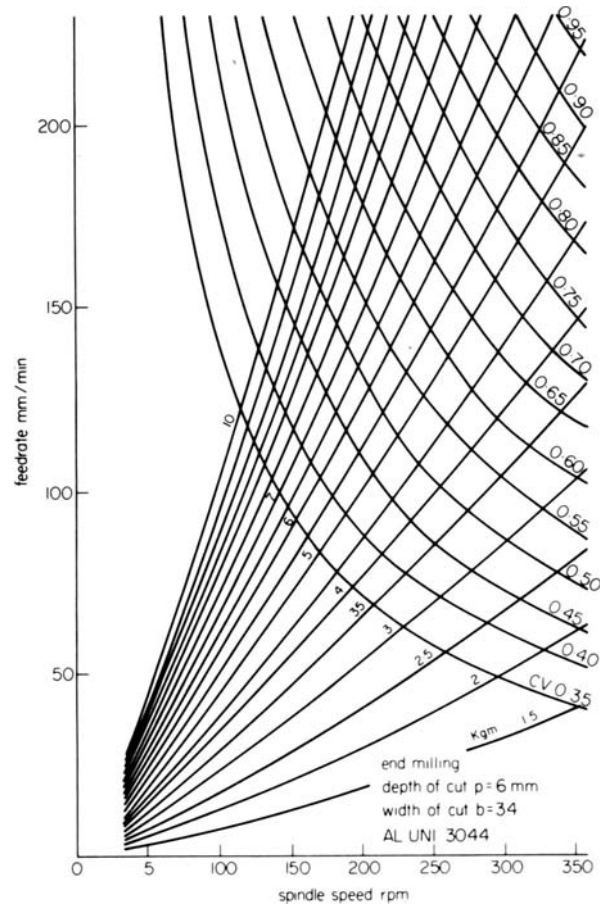


Figure 13. Experimental results (cutter: $\Phi 80$ mm, 10 teeth).

In figures 8, 9 and 10 the resulting plots for a single test on steel can be seen.

All the obtained results allow the following considerations:

- (1) The exponents in the bending and torque relationships (1) have reasonably similar absolute values. This means that the bending and the torque are functions of the feed per tooth⁷.
- (2) For worn cutters the exponents tend to decrease. This is justified considering that in this case the power grows more rapidly with the spindle speed (figure 11).
- (3) The standard deviation assumes small values for sharp cutters. For worn cutters it increases: the cutting conditions being worse.

Similar conclusions can be drawn for face milling. Figures 12 and 13 report the results of cutting tests carried out on cast iron and aluminium.

CONCLUSIONS

The process digital computer provided with analogue and digital interfaces and connected to an automated machining centre provides a versatile tool for conducting and recording machining tests. Versatility is due to the easiness with which programming for different tests and analysis of the results can be performed. Moreover since generally the computer requirements for machine-tool control and data acquisition are not excessive, it is noted that newly available mini-computers already incorporated in

DNC and CNC installations should prove adequate for implementation of data collection algorithms without the need for complex modifications of these machining systems or large investments in special-purpose controls.

ACKNOWLEDGMENTS

The authors wish to express their sincere appreciation to the Olivetti Machine Tool Corporation for the support under which the present work was carried out and to G. Sandrelli, L. Luchi and A. Poggialini for their help during the experimental work.

REFERENCES

1. P. C. Pinotti et al. Adaptive control in milling, *Proc. C.I.R.P. Seminars on Manufacturing Systems*, vol. 2, no. 3, 1973.
2. R. Bedini et al. Control strategies in metal cutting, Paper presented at IV C.I.R.P. Seminar on Manufacturing systems, 1972.
3. R. Bedini and G. Lisini. Computer control of milling machine, *S.M.E. Technical Paper MS73-173*, 1973.
4. R. Bedini et al. Digital adaptive control of a milling machine, *A.S.M.E. Paper 73-WA/Aut-12*, 1973.
5. P. C. Pinotti et al. A bending transducer for a A/C milling machine (to be published 1974).
6. B. R. Beadle and J. G. Bollinger. Computer adaptive control of a machine-tool, *C.I.R.P. Annals*, vol. XXIV, p. 61, 1971.
7. A. Hornung and G. F. Micheletti. *La Fresatura*, Edizioni Pem. (1969).

POWER HACKSAWING

by

P. J. THOMPSON* and M. SARWAR*

SUMMARY

Investigations into the mechanics and economics of power hacksawing have been carried out. Factors which control the metal removal rate of the blade are indicated, and some relationships obtained. The influences of the machine and blade characteristics on the cost per cut are discussed, together with the problems associated with blade testing.

INTRODUCTION

The problem of cutting-off material to size is common to practically every industry. Therefore, it is surprising that so little work has been done to understand the problems of this common operation. Many reasons have been given for this lack of interest; the common occurrence of cutting-off machines, the fact that these machines are frequently housed in stores away from main production areas, the machines may be operated by low grade labour, the operation of the machines appears simple, etc. The fact remains that cutting-off operations can be very costly. Later it will be shown that the cost of cutting-off an austenitic stainless steel bar, 75 mm diameter, can be as high as 39 p per cut.

Previous published works on cutting-off materials have been primarily concerned with circular and band sawing¹⁻³, and cost comparisons between alternative processes⁴⁻⁶. An aim of this work has been to understand the mechanics and economics of power hacksawing with a view to specifying changes in machine design which are likely to result in a reduction in cutting costs.

One reason for undertaking this work is the growing realization, on the part of manufacturers of both blades and machines, particularly in the Sheffield area, that the factors which control the mechanics and economics of power hacksawing are complex. Also, power hacksawing has been receiving increased competition from other cutting processes, such as band and circular sawing. Whilst these alternative processes are frequently quicker than power hacksawing their costs are in many applications higher. Also, when large sections of difficult-to-machine metals are to be cut-off, power hacksawing can be both quicker and cheaper⁵. Whilst the impact

of these alternative processes on the application of power hacksawing cannot be denied, there remains a significant field of applications for power hacksawing which is likely to remain unchallenged. In addition, if the costs of power hacksawing can be reduced, its potential fields of application will obviously be increased; this is of prime interest to manufacturers.

Another reason for undertaking this work is that some difficulty has been experienced in obtaining consistency in blade test data, obtained on hydraulic machines⁷. In the discussion which follows comments are made about the significance of load measurement during blade testing.

HACKSAWING MACHINES

Power hacksawing machines may be classified according to the method used to develop load between the blade and the workpiece during the cutting stroke. In all types of machines, this load is removed during the return stroke. The following are the three main classifications.

Hydraulic machines

Force between the blade and the workpiece, the thrust force, is developed in this type of machine by a hydraulic device. Pressure may be developed in the loading cylinder by either a restricted back-flow system, described and analysed in Appendix A, or the pressure may be supplied from a separate pump. These machines are the most common, and develop greater thrust loads than machines of other types. The thrust load developed by any hydraulic machine varies considerably throughout the cutting stroke, and the variation is characteristic of the hydraulic system used (figure 1).

* Department of Mechanical and Production Engineering, Sheffield Polytechnic

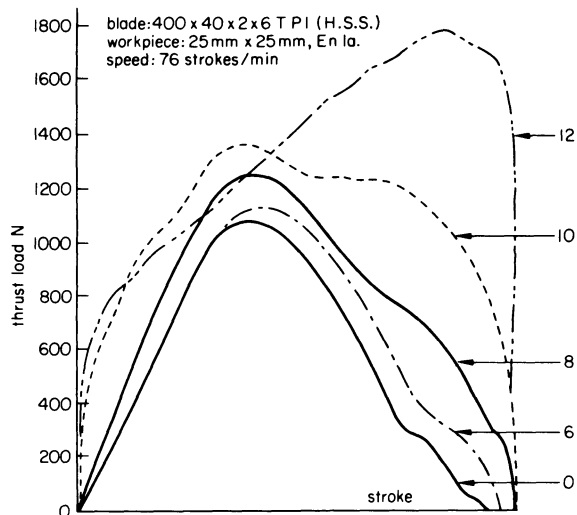


Figure 1. The load developed during the cutting stroke between the blade and the workpiece by a 200 mm hydraulic saw on various load settings.

Gravity fed machines

In this type of machine the thrust load is developed by the weight of the blade bow. The thrust load again varies throughout the cutting stroke, due to the reciprocating motion of the blade bow, and the action of the cam operated lift-off device at the beginning, and the end, of the cutting stroke. In many of these machines, the magnitudes of the thrust loads are fixed, although some machines are provided with adjustable weights on the over-arms for thrust load adjustment.

Positive displacement machines

In the positive displacement machine, the downward movement of the blade and, therefore, the material removal rate is controlled directly by a feed screw mechanism. This type of machine is not common since it can lead to overloading of the blade and premature blade fracture, particularly when the blade is worn. This type of machine will not be discussed further, but it is not prone to the same variation in thrust load during the cutting stroke as other types of machine.

For a given blade and workpiece, the material removal rates, achieved by hydraulic and gravity fed machines, are controlled solely by the thrust loads developed. Therefore, hacksawing may be said to be a process in which the material removal rate is force controlled; this is unlike most other material removal processes.

In the tests to be discussed a mean thrust force has been used. This and other thrust force parameters are defined in Appendix B. It is believed that, when these machine load parameters are used in blade wear and metal removal rate tests, the machine characteristics are fully described and the resulting cutting data are only characteristic of the blade.

Blade characteristics

Hacksaw blades are usually classified according to their principal overall dimensions, teeth pitch, the teeth set pattern and the material from which the blade is made. For most blades the major cutting

angles are as shown in figure 2. The principal overall dimensions, and the teeth pitch, are normally shown as follows

length of blade between hole centres (mm)
x overall blade width (mm) x blade thickness (mm)
x number of teeth per inch

This BS 1919 system⁸ of size designation is used throughout the paper.

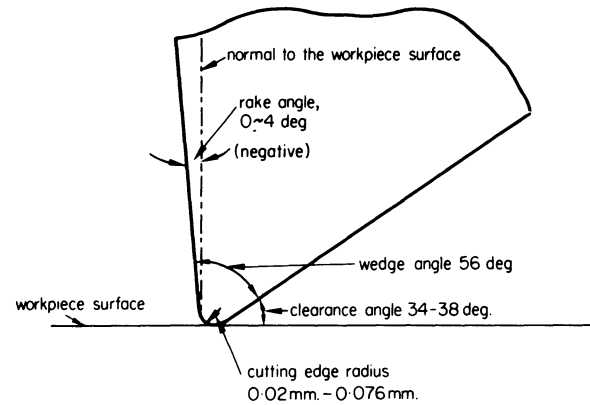


Figure 2. A typical tooth profile showing principal angles.

Some of the teeth which make up the cutting edge of the blade are bent laterally, or set, during manufacture, so that clearance is achieved between the blade and the slot produced during cutting. The following set patterns⁸ are common

- (1) Alternative set, right-left-right etc.
- (2) Raker set, right-centre-left etc.
- (3) Balanced set
- (4) Wavy set

Examinations of the cutting edges of hacksaw blade teeth have shown them to be irregular. This is due to the traditional method of manufacture in which the tooth profile is produced by gang-milling followed by heat-treatment. These processes produce badly formed cutting edges. The actual profile of the cutting edge differs considerably from one tooth to the next. By describing the cutting edge profile as a radius, some measure of the variation of tooth profile has been made. These measurements, carried out on an optical projector, show that this radius varies from 0.02 mm (0.0008 in) to 0.076 mm (0.003 in). Taking the statistical average of this profile radius for all teeth on a blade, and comparing it with the average for other blades having the same teeth pitch, small variations are observed. Comparing the average of blades having different pitch shows a more significant variation.

Another geometric variation, introduced during manufacture, which produces changes in the metal removal rate of the blade, is the variation in the overall thickness of the blade after setting. The overall thickness is measured across the set teeth, from the extreme edge on one side of the blade to the extreme edge on the other. This dimension controls the width of the slot produced and, therefore, the volume of material removed per cut.

Instrumentation

The prime requirement of the instrumentation was to enable the load, developed between the blade and the workpiece and the other components of the cutting force, to be measured against the position in the stroke. A secondary requirement was a need for this instrumentation to be easily adaptable to most hacksawing machines.

The three mutually normal components of the cutting force were measured, via the workpiece, by a dynamometer clamped in the workpiece vice of the machine. The load measuring element of the dynamometer consisted of the Kistler 9257A piezo-electric, three component dynamometer and its associated charge-amplifiers. The displacement of the blade was measured by a linear transducer and transducer meter. At slow rotational speeds, the outputs from these instruments were displayed on an X-Y plotter; loads against blade displacement. At high rotational speeds the outputs were displayed on a multi-channel oscilloscope, and recorded with a polaroid camera.

The cutting action of new blades

Cutting tests have shown that the average depth of cut achieved per tooth, defined in appendix B, is very small in most applications. These measurements, together with those of the cutting edge radius, show that in the majority of applications the depth of cut is less than the cutting edge radius. Therefore, the hacksaw blade may be classified as a blunt cutting tool.

Figures 3(a) and 3(b) show the deformation

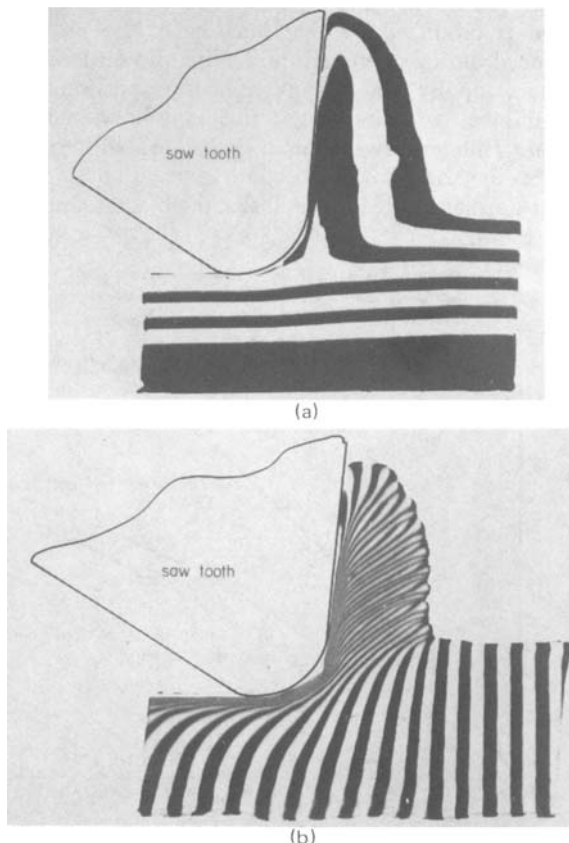


Figure 3. Deformation produced by a tool having a cutting edge radius larger than the depth of cut, (a) horizontal grid (b) vertical grid.

produced by a cutting edge having a radius larger than the depth of cut. These photographs were obtained using a perspex model of a blade tooth and striped plasticine workpieces. Whilst great accuracy is not claimed for the deformation patterns produced by these model tests, they are believed to be sufficiently accurate for some conclusions to be drawn. It is seen that a built-up edge develops in the front face of the cutting edge radius. This leads to a shear mode of chip formation, and a thin layer of heavily deformed material in the surface of the workpiece. With a work hardening metal workpiece this would produce a hard surface skin which would make tooth penetration difficult on subsequent strokes of the saw.

Observations of steel chips, produced by hacksawing under the action of a light thrust load, confirm that the chips are produced by a shear type

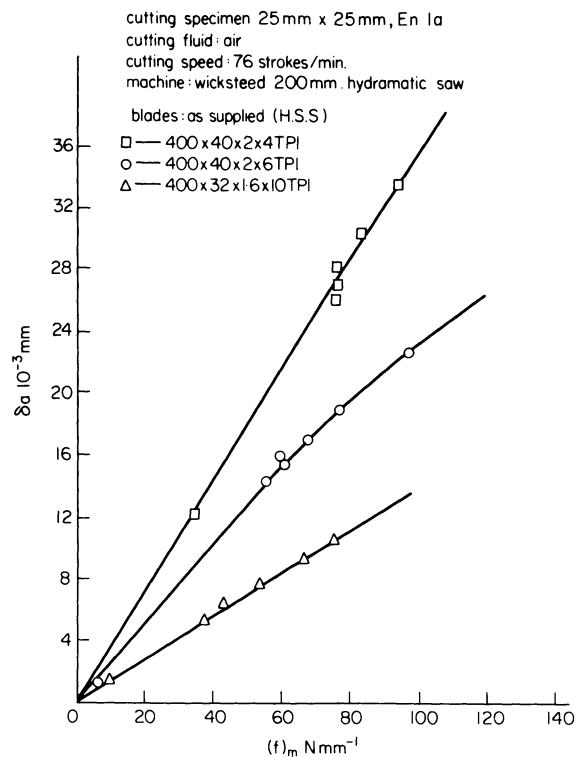


Figure 4. The average depth of cut per tooth against the thrust load per tooth per unit thickness for blades having different teeth pitch.

mode of deformation. The chips and debris vary considerably in size. Microscopic examinations of these chips show evidence of surface shear lines, produced as the material 'piles-up' in front of the blunt cutting edges. Frequently the chips show evidence of surface oxidation, indicating high cutting temperatures. When cutting a soft material, under the action of a large thrust load, the chips produced are continuous curls of material similar to chips produced by a sharp tool. This change in chip formation is brought about by the depth of cut achieved being significantly larger than the cutting edge profile radius on the blade teeth.

Figure 4 shows the variation in the average depth of cut per tooth against the mean thrust load per tooth per unit thickness for a En 1a workpiece and blades of different pitch. It can be seen that these parameters are directly proportional to each other for a given

blade pitch. Both parameters are defined in appendix B. This result may be written for a given workpiece material, workpiece size and blade pitch thus

$$\delta_a = K \cdot f_m \tag{1}$$

where

- K = cutting constant
- δ_a = average depth of cut per tooth
- f_m = mean thrust load per tooth per unit thickness

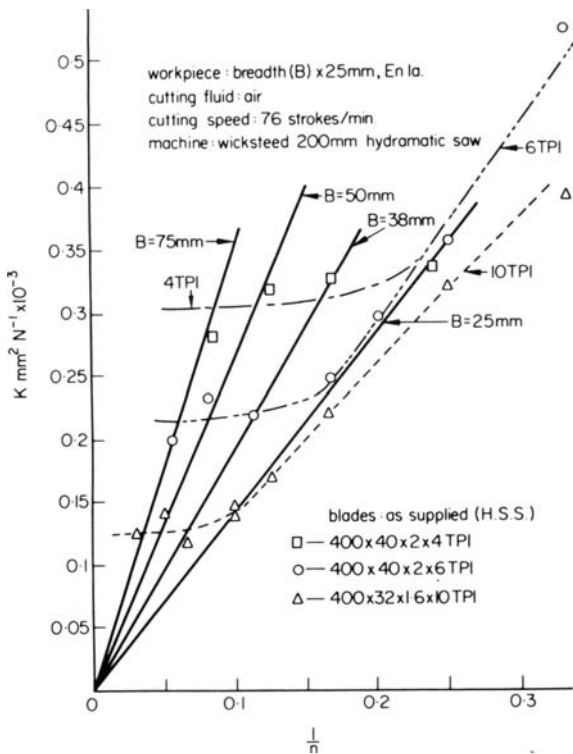


Figure 5. The cutting constant against the reciprocal of the number of teeth in contact with the workpiece.

Figure 5 shows the variation in the cutting constant against the reciprocal of the number of teeth in contact with the workpiece. These results show a significant size effect. These effects may be summarized as follows:

when $B \geq 25$ mm for En 1a

$$K = a + b \cdot p \tag{2}$$

when $2p < B < 25$ mm for En 1a

$$K \cdot n = c \tag{3}$$

where

- B = breadth of the workpiece
- n = number of teeth in contact with the workpiece
- p = pitch of the blade teeth
- a, b, c = material constants for a particular blade type.

The number of cutting strokes needed to cut through a rectangular workpiece may be found by combining Equation (1) with the expressions for the average depth of cut per tooth and the mean thrust force per tooth per unit thickness, given in Appendix B. Thus

$$N \cdot S \cdot F_m = \text{constant} \tag{4}$$

where

- N = number of strokes to cut through the workpiece
- S = stroke of the saw
- F_m = mean total thrust load

The constant in this equation depends on the workpiece material, workpiece size and blade.

The cutting action of worn blades

The teeth are worn in such a way that a wear flat is produced at the tip of each tooth, and the outer corners of the set teeth are rounded. The effect of this wear is to increase the thrust load needed to achieve a given depth of cut.

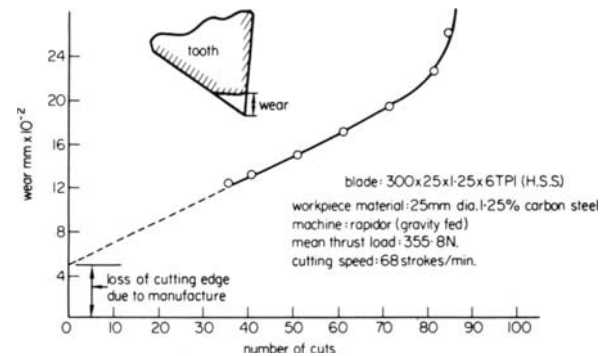


Figure 6. The tooth wear against the number of cuts achieved.

Figure 6 shows the loss in height of a tooth with increases in the number of cuts made. The shape of this graph is similar to that obtained by measuring flank wear on a single point cutting tool, and it shows primary, secondary and tertiary regions. The primary region is produced by the rounding of the profile, produced during manufacture. During the earlier part of the blade life, the loss in height is proportional to the number of cuts made; this is the secondary region. The tertiary region is associated with rapid increase in wear. Wear can become so pronounced that under normal applied thrust loads, the blade ceases to cut.

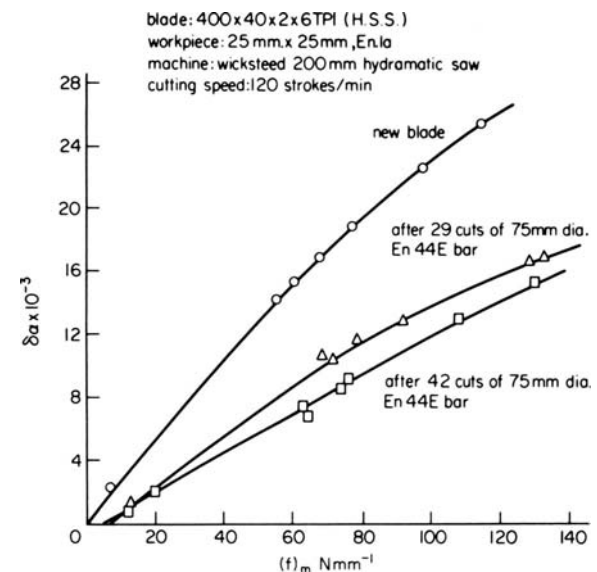


Figure 7. The average depth of cut per tooth against the mean thrust load per unit thickness for a blade in three states of wear.

TABLE 1 Cost per cut for a En 58J austenitic stainless steel workpiece, 75 mm diameter

Time utilization of the machine (%)	C ₁ (%)	C ₂ (%)	C ₃ (%)	C ₄ (%)	Total cost in pence per cut
20	4	61	18	17	39.0
40	3	45	27	25	26.2
60	2	36	32	30	22.0
80	2	30	35	33	19.9
100	1	26	38	35	18.6

Figure 7 shows the variation between the average depth of cut per tooth and the mean thrust load per tooth per unit thickness, for a blade in three stages of wear. It is seen that wear reduces the cutting constant and shifts the curve along the thrust force axis. This indicates that some initial load is needed before material is removed, when the blade is worn.

ECONOMICS OF HACKSAWING

Considerable cost data are available⁴⁻⁶ for power hacksawing, band sawing and circular sawing. However, in many cases, the methods of arriving at the cost figures quoted are not stated fully, and the data used are not given in sufficient detail to enable a breakdown of cost to be undertaken. Therefore, two cost analyses have been undertaken, based on present day costs and typical operating conditions for workpiece materials which are known to be difficult to cut. The data and details of the costing method are given in appendix C. The cost per cut has been divided under the headings given below, and are shown in tables 1 and 2 against the percentage time utilization of the machine.

- C₁ = non-productive time cost per cut
- C₂ = cutting time cost per cut
- C₃ = total blade cost per cut
- C₄ = scrap material cost per cut

DISCUSSION

Process characteristics

Some factors and relationships which govern the metal removal rate of a power hacksaw blade have been outlined.

A Power hacksaw blade is a blunt cutting tool whose metal removal rate is controlled, for a given workpiece, by the pitch of the teeth and the load applied to it by the hacksawing machine. The average depth of cut per tooth has been shown to be directly proportional to the mean thrust load per tooth per unit thickness for a given workpiece material, work-

piece size and teeth pitch. The ratio of these parameters has been defined as a cutting constant, and is shown to depend on the number of teeth in contact with the workpiece. For broad workpieces the cutting constant increases with an increase in tooth pitch, all other factors remaining constant. For thin workpieces the cutting constant is increased by a reduction in the number of teeth in contact with the workpiece and is less dependent on the pitch of the blade teeth. For all sizes of workpiece, work materials and blades, the number of cuts needed to cut through the workpiece is shown to be reduced by an increase in the total mean thrust load and an increase in the stroke of the saw.

Process costs and the machine

From the costs given in Tables 1 and 2, it is possible to draw some important and general conclusions about the requirements of the machine. For the occasional user whose time utilization is low the predominant cost is the time cost; this is because the cost rate is very high. The change most likely to benefit this type of user is to reduce the cutting time. For the frequent user whose time utilization is high, the predominant cost is the total blade cost with the cutting time cost also significant; this situation offers more scope for change. It would be possible to increase the cost of the machine to this type of user and still produce an overall cost benefit which could be significant, provided the higher cost machine could reduce the cutting time and increase the blade life.

The cutting time can be reduced by either reducing the number of cutting strokes needed to saw through the workpiece or by increasing the rate of cutting strokes. Equation (4) shows that the number of strokes needed to cut through the workpiece can be reduced by either increasing the mean thrust load or by increasing the length of the stroke.

Present machine designs enable the thrust load to be varied. The maximum load that can be satisfactorily used depends ultimately on the breaking strength of the blade; a limit which can be exceeded with existing machines. However, the thrust load has

TABLE 2 Cost per cut for a En 44E steel workpiece, 75 mm diameter

Time utilization of the machine (%)	C ₁ (%)	C ₂ (%)	C ₃ (%)	C ₄ (%)	Total cost in pence per cut
20	6	69	20	5	24.2
40	5	56	31	8	15.0
60	4	47	39	10	12.0
80	3	40	46	11	10.5
100	3	35	50	12	9.5

been shown to undergo considerable variation and the mean thrust load is considerably less than the maximum thrust load developed. For a given blade strength, the number of cutting strokes needed could be reduced if the mean thrust load is increased relative to the maximum thrust load. To achieve this, the thrust load would need to be uniform throughout the cutting stroke.

The other possible way to reduce the number of cutting strokes would be to increase the machine stroke for any workpiece size. Most existing machines operate with a fixed stroke, the size of which is determined by the blade size and the maximum workpiece size to be accommodated, e.g. for a blade length of 400 mm, say 375 mm usable length, and a required maximum workpiece capacity of 225 mm, the fixed stroke possible is 150 mm. In use, such a machine would only rarely cut workpieces as large as the maximum machine capacity. For smaller workpieces, 75 mm diameter being a reasonable average, the machine discussed above would only use 225 mm of the blade; its remaining length would be discarded with the worn blade unused. It is believed that this is an argument for using a machine with an adjustable stroke.

Another way to reduce the cutting time is to increase the stroke rate, 70–120 strokes per minute is typical of present practice. A few tests have been carried out on existing hydraulic machines at high speed. It has been found that, at these speeds, there is some variation in the thrust load applied between the blade and the workpiece during the cutting stroke. For high speed use, the dynamic stability and the speed of response of the hydraulic system of the machine will have to be considered.

Factors which would lead to an increased blade life are complex. One possibility lies in the recent development of a bi-metal blade. The main motivation in doing this is to produce a more flexible blade to meet American standards. However, the bi-metal blade gives the opportunity to use more wear resistant metals for the teeth, and this may ultimately lead to a significant increase in blade life.

Blade testing

To conclude, the difficulties in blade testing will be discussed. The object of present blade tests is to act as a quality control test for blade manufacturers.

For hand hacksaw blades, a gravity fed machine is used in a wear test, together with a standardized test bar; BS 1919 contains the test details. For a blade to be satisfactory it must perform a standard number of cuts on the test bar and the cutting time must not exceed a given standard. At present there is no British Standard test for power hacksaw blades. However, most manufacturers have their own test which they use for internal purposes. Again it usually takes the form of a wear test carried out on a standardized bar using a hydraulic machine. A typical criterion used is the number of cuts needed for the cutting time to be increased to twice its initial value. Variations in the results of such tests have been found even when the type and make of machine is standardized⁷. The points made in Appendix A emphasize how important load measurement is when using hydraulic

type machines. The absence of load measurement is the most likely explanation of the variation previously obtained.

ACKNOWLEDGMENTS

The author would like to thank James Neill (Services) Ltd for sponsoring this work and Mr J. Brimacombe for undertaking work which led to figures 3(a) and 3(b), as a final year project.

APPENDIX A

The slow speed thrust load characteristics of a hydraulically loaded power hacksaw machine, operating on the restricted back-flow principle

Figure 8(a) shows a simplified diagram of the principal mechanical and hydraulic arrangements of a typical hydraulic machine, operating on the restricted back-flow principle. The saw bow is carried in a slideway, housed in the swing-arm assembly. This assembly rotates about the pivot shown. The slideway in the swing-arm assembly, and the fixing points for the blade on the saw bow, are arranged so that a small taper exists between the cutting edge of the blade and the swing-arm slideway. The effect of this taper is such that, during the inward cutting stroke, the blade and the swing-arm assembly rotates clockwise about the pivot point. This motion causes the piston in the hydraulic cylinder to rise, and displaces oil through the flow control valve. This back-flow develops pressure in the cylinder, and a torque about the pivot point which opposes the

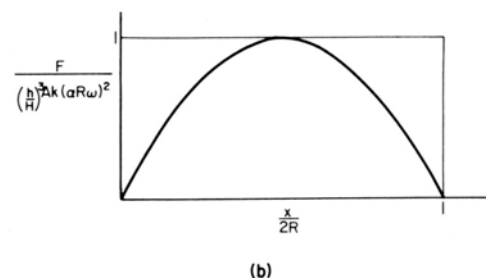
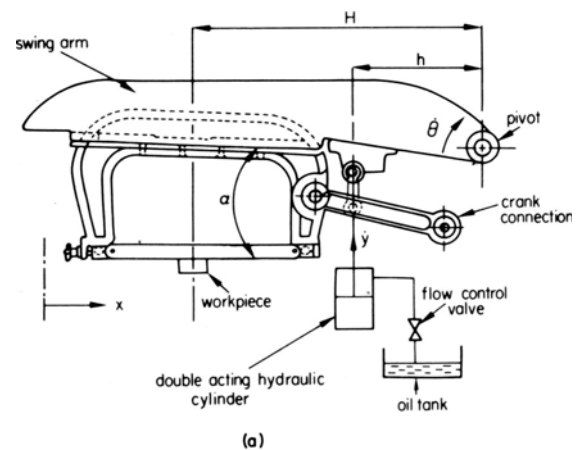


Figure 8. (a) A hydraulic machine showing the main features of the restricted 'back flow' method. (b) A theoretical thrust load against blade displacement diagram for such a machine and rigid, non-cutting blade.

motion of the swing-arm. It is this torque which develops the thrust load between the workpiece and the blade. The following expression is the result of a simple analysis of this mechanism which ignores inertia forces, as speeds are small.

$$F = \left(\frac{h}{H}\right)^3 \cdot A \cdot k [\dot{x}(\alpha - \phi - \beta)]^2 \quad (A1)$$

where

- F = instantaneous thrust load acting between the blade and the workpiece
- h = moment arm to the centre of the hydraulic cylinder
- H = moment arm to the centre of the workpiece
- A = cross section area of the hydraulic cylinder
- k = flow valve constant, when $p = k(\dot{y})^2$
- p = pressure developed in the hydraulic cylinder
- \dot{y} = velocity of the piston in the hydraulic cylinder
- \dot{x} = cutting velocity
- α = angle between the cutting edge of the blade and the slideway in the swing-arm assembly, when the blade is unloaded
- ϕ = change in the wedge angle due to blade deflection
- β = change in wedge angle due to the modified motion of the blade when removing material

It is useful, in understanding the characteristics of this mechanism, to make further assumptions. Consider the blade to be infinitely rigid, and not to be removing material. This can be approximately achieved by replacing the blade with a rigid bar without a cutting edge. This further simplification gives $\lambda = \alpha$, since $\phi = \beta = 0$. Also, consider the motion of the blade to be sinusoidal. This is not quite true since most sawing machines give the blade motion by means of a crank and connecting rod mechanism. For these conditions

$$F = \left(\frac{h}{H}\right)^3 \cdot A \cdot k \cdot (\alpha \cdot R \cdot \omega)^2 \sin^2(\omega t) \quad (A2)$$

where

- ω = angular velocity of the crank
- R = crank radius = $S/2$
- S = saw stroke

Figure 8(b) shows the variation in the thrust load predicted by the above expression.

Several important conclusions may be drawn from this simple analysis. Firstly, the general shape of the thrust load curve is such that no load is developed by the machine at the beginning or the end of the cutting stroke. Whilst this makes it easy to engage and disengage the blade without suddenly loading or unloading the blade teeth, it does mean that, for part of the stroke, the blade is only lightly loaded, and, as a consequence, the material removed during this part of the stroke is small.

The simple expression predicts that the maximum load is developed at the mid-stroke position. Whilst the general shape of the curve predicted is typical of those measured experimentally, when the blade is loaded lightly (figure 1), high loads cause considerable blade deflection and high material

removal rates which cause large variation in the effective wedge angle throughout the cutting stroke. The effect of blade deflection and material removal rate is to displace the point of maximum load from the mid-stroke position, usually towards the end of the cutting stroke, and to affect the magnitude of the loads developed. These effects in particular make this mechanism prone to strong interaction between the characteristics of the blade and its cutting action, and the load developed by the machine for a given flow control valve setting.

The only way in which the load developed can be controlled by the operator, for a given set of conditions, is by adjustment of the flow control valve. The above demonstrates that this is a crude method for load adjustment, and it cannot be assumed that doubling the load setting in the machine doubles the mean load acting between the blade and the workpiece defined in Appendix B.

APPENDIX B

Determination of the average depth of cut per tooth

The volume of material removed during a small displacement of the blade in the cutting stroke (figure 9) is

$$\Delta(\text{Vol}) = t \cdot n \cdot \delta \cdot \Delta x$$

where

- $\Delta(\text{Vol})$ = volume of material removed during the small displacement of the blade Δx
- t = thickness of the blade
- n = number of teeth in contact with the workpiece = Bp for broad workpieces
- B = breadth of the workpiece
- p = pitch of the blade teeth
- δ = instantaneous depth of cut achieved per tooth

As the thrust load varies during the cutting stroke, the instantaneous depth of cut achieved per tooth, also, varies. Hence, the volume of material removed during one cutting stroke $(\text{Vol})_s$ becomes

$$(\text{Vol})_s = t \cdot n \cdot \int_0^s \delta \cdot dx$$

As the instantaneous depth of cut is difficult to measure experimentally, it is convenient to introduce

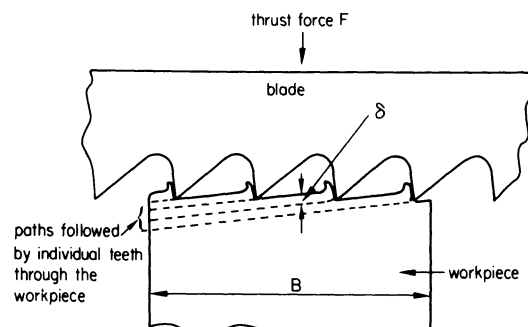


Figure 9. Cutting action of a hacksaw blade showing the method of material removal.

the average depth of cut, achieved during the cutting stroke per tooth, defined as

$$\delta_a = \frac{1}{S} \int_0^s \delta \cdot dx \quad (\text{B1})$$

where

δ_a = average depth of cut per tooth
 S = stroke of the saw

Combining these expressions, when the workpiece is broad compared with the pitch of the blade teeth, gives

$$(\text{Vol})_s = t \cdot B \cdot p \cdot \delta_a \cdot S \quad (\text{B2})$$

The volume of material removed during the cutting stroke may also be found from the loss in volume of the workpiece. For a rectangular workpiece this becomes

$$(\text{Vol})_s = W \cdot B \cdot \delta_s \quad (\text{B3})$$

where

W = width of the slot produced
 δ_s = increase in the slot depth produced per cutting stroke = D/N
 D = depth of a rectangular workpiece
 N = number of cutting strokes needed to cut through the rectangular workpiece

Combining these expressions gives

$$\delta_a = \left(\frac{w}{t}\right) \cdot \frac{D}{N} \cdot \frac{1}{S \cdot p} \quad (\text{B4})$$

The ratio (w/t) is the ratio of the slot width, or the thickness of the blade measured from the outside edges of the set teeth, to the thickness of the blade material, and is known as the thickness ratio, z .

The number of teeth which make contact with the workpiece during the cutting stroke, N_c , is

$$N_c = (S + B)p \quad (\text{B5})$$

However, the teeth which are in contact with the workpiece at the beginning and at the end of the cutting stroke do not travel across the full breadth of the workpiece. The equivalent number of teeth which may be said to make one complete traverse of the workpiece, N_e , may be shown to be

$$N_e = S \cdot p \quad (\text{B6})$$

Therefore, the expression for the average depth of cut per tooth may be written

$$\delta_a = \frac{D}{N} \cdot \frac{z}{N_e} \quad (\text{B7})$$

Determination of the mean thrust load per tooth per unit thickness

The mean thrust load per tooth per unit thickness may be defined as the mean load acting between the blade and the workpiece during the cutting stroke per tooth per unit blade thickness; assuming that each tooth in contact carries equal load.

The mean total thrust load acting between the blade and the workpiece is given by

$$F_m = \frac{1}{S} \int_0^s F dx \quad (\text{B8})$$

where

F = instantaneous thrust load after a displacement x of the blade from the beginning of the cutting stroke
 F_m = mean total thrust load

The mean thrust load per tooth per unit blade thickness, f_m , is given by

$$F_m = f_m \cdot n \cdot t \quad (\text{B9})$$

For broad workpieces this may be written

$$f_m = \frac{F_m}{B \cdot p \cdot t} \quad (\text{B10})$$

As the load developed between the blade and the workpiece, and its variation during the cutting stroke, depends on the type and characteristics of the sawing machine, the mean thrust load per tooth per unit blade thickness is a machine parameter which is controlled by the machine settings etc.

APPENDIX C

Cost data

The cost data shown in Tables 1 and 2 are based on a hydraulic power saw, a 400 mm x 40 mm x 2 mm x 4 TPI blade and two workpieces both 75 mm diameter, one in En 44E and the other in En 58J, austenitic stainless steel. The blade life and cutting time were obtained on a 6-7 load setting at a rotational speed of 104 strokes per minute.

It is assumed that the cost of the sawing machine is written-off over a period of 10 years and no interest is paid on the capital involved. The operators hourly labour rate is based on current semi-skilled rates, and it is assumed that he is operating two machines simultaneously; this assumption is believed to represent current trade practice. The workpiece is assumed to be manually clamped in the vice. All power and cutting fluid costs have been ignored.

Cost of machine	= £710
Cost of blade	= £1.12
Cost of workpiece material	
(i) En 44E	= £145 per ton
(ii) En 58J	= £790 per ton
Operators labour rate	= 60p per hour
Annual weeks worked	= 48
Hours worked by the plant	= 40
Non-productive time per cut	= 30 sec
Cutting time per cut	
(i) En 44E	= 6 min
(ii) En 58J	= 8.5 min

Time to change blade	= 75 sec
Width of cut	= 2.35 mm
Blade life in cuts	
(i) En 44E	= 24
(ii) En 58J	= 16

REFERENCES

1. R. A. Anderson. Band machining and friction sawing. *I. Mech. E. Proc. Conf. Technology of Engineering Manufacture* (1958).
2. G. Pahlitzsch and A. Willemit. Cutting-off steel with carbide-tipped saws. *Werkstattstechnik* (April, 1968) **58**, no. 4, 150-56.
3. E. Barz and K. Oberwinter. Phenomena in fusion sawing. *Werkstattstechnik* (January, 1966) **56**, no. 1, 13-21.
4. Fabrication estimation, Part 3. *Cold sawing, welding and Metal Fabrication* (Feb. 1966) **43**, 161-62.
5. R. E. Nelson, Bandsawing or Hacksawing? *American Machinist* (November, 1965) **109**, no. 24, 90-3.
6. J. L. Remmerswaal. On the economics of the cutting-off of metals. *Int. Institution for Prod. Eng. Res. Microtechnic XV*, no. 4.
7. A survey undertaken by NEL to assess the feasibility of a cutting test for power hacksawing blades. *Divisional Report, NEL* (1968).
8. *British Standard 1919: Specification of Hacksaw Blades*. British Standards Institution (1967).

OPTIMIZATION OF DRILL GEOMETRY FOR ORTHOPAEDIC SURGERY

by

G. H. FARNWORTH* and J. A. BURTON†

SUMMARY

With the advent of internal fixation a surgeon can assist the natural process of bone repair by ensuring that a perfect reapposition of parts is achieved. The rigidity of internal fixation determines the success of open fracture treatment, which most commonly depends on the correct use of plates and screws. This paper outlines the present state of knowledge on the performance testing of orthopaedic drills. It surveys the results of an experimental programme to minimize the forces developed in the drilling of bone, and to ensure accurate location of the holes produced, in order that maximum control of the drilling process can be maintained by the surgeon.

INTRODUCTION

The fixation of fractures represents a great compromise between biological principles and engineering principles¹. Since union cannot be imposed upon a fractured bone, it becomes the surgeon's responsibility to utilize natural healing tendencies. These tendencies are encouraged by observing the following principles. Firstly a reduction of the fractured bone must be gained and retained. A satisfactory reduction must include some degree of apposition of the bone fragments, proper alignment in both linear and rotary planes, and adequate length for that limb if such restitution does not compromise the first two requirements. Secondly, the fractured bone must be adequately immobilized until a certain amount of healing has taken place. Thirdly, contact-compression embodies the pathophysiology of bone resorption at the opposed cortical margins of reduced shaft fractures especially, and the desirability of permitting such fragments to move towards each other as such resorption takes place.

The rigidity of internal fixation therefore determines the success of open fracture treatment. For this form of treatment plates and screws are the devices most commonly used. The successful outcome of this method depends on fundamental mechanical principles, and failures are usually caused by defective materials or improper technique of application.

The majority of holes drilled preparatory to the insertion of screws are made in the shafts of long bones. These shafts are in effect tubes, there being a hard outer cortex and a central medullary cavity. The

point of entry of the drill is called the near or proximal cortex and the point of exit the far or distal cortex.

Of necessity the holes are drilled freehand, which is difficult even with a drill guide. With an inefficient cutting action there is considerably more effort required by the surgeon, and his control over the drilling process is affected. If the torque on the drill is too great there is a tendency for the drill to jam, or even to break in the hole. Should excessive thrust be required, the drill may plunge through the far cortex, damaging surrounding soft tissue.

The importance of selecting particular tool shapes for different materials is generally recognized in engineering, and it is usual to grind the tools to the angles determined by extensive investigation. This investigation endeavours to establish the effects of drill point form, helix angle and web thickness on drill performance when machining bone for orthopaedic purposes.

BONE STRUCTURE

Bone consists of a highly organized system of cells, surrounded by a matrix composed of a fibrous protein, collagen, encrusted with tiny calcium phosphate crystals. Cells occupy only about 3 per cent of the total volume, but are evenly distributed through the bone and interconnected by a series of canaliculi, so that no volume of interstitial bony material is more than 5 to 10 μm from a cell or its processes, and no bone cell is more than 200 μm from a capillary. In fully mineralized bone the protein

* Loughborough University

† Preston Polytechnic

matrix accounts for 35 per cent of the volume of the intercellular material, and the crystals occupy most of the remainder.

The collagen fibres in adult human bone are spatially highly organized, forming interlacing bundles and layers; this arrangement appears to be related to the direction and character of the mechanical forces which the organ must withstand. Furthermore, the intimate association of fibrous protein and small crystals constitute what, in materials engineering, has been termed a 'two-phase' material. The breaking strength and elasticity of bones depend not only on their mass and shape, but also on the intrinsic mechanical properties of this two-phase system and on changes that occur in those properties as the material ages.

The mineral component of bony material is a calcium phosphate generally considered to possess the crystal structure of hydroxyapatite. However, the extremely small size of bone crystals and their associated high surface-to-volume ratio have frustrated attempts to resolve this question. There is evidence to suggest that newly deposited mineral may exist in a different crystalline form that spontaneously changes to hydroxyapatite after it is deposited². This matter is of more than academic importance because the solubility properties and mechanical characteristics of the calcium phosphate minerals differ widely for different crystal configurations.

Bone is deposited on free surfaces by osteoblasts that synthesize collagen and secrete it onto the bone-forming surface. Soon after secretion the soluble collagen molecules aggregate into spatially oriented insoluble fibrils. The factors that control their orientation are not well understood, but bioelectrical fields of the magnitude known to be produced in bone by applied mechanical stress have been shown to influence the orientation of collagen fibrils in vitro. The layer of bone matrix thus deposited on a bone-forming surface is known as an osteoid seam, and in normal bone is about 10 μm thick. These seams are not immediately mineralizable. Under control of the parent osteoblast the matrix undergoes a complex series of biochemical changes, which ultimately render it capable of initiating crystal nucleation and growth. Crystal growth displaces the water that originally occupied 70 to 90 per cent of the volume of unmineralized matrix. There is almost no free water in mature bone, so that interaction between bone and body fluids is largely limited to free surfaces.

REVIEW OF FIXATION AND BONE DRILLING

Fixation

This requires the perfect reapposition of parts first of all, followed by the treatment of the gap between the broken ends as if it were an incised wound in which the edges are brought together not by sutures but by screws. The principle of rest is secured by fixation derived from screws in place of splints and bandages.

The initial use of internal fixation appears to be unknown, but in 1862 Gurlt reported a number of cases in which the freshening and reduction of

fractured bone ends together with nailing, screwing and wiring was attempted.

In 1868 Lord Lister operated successfully on the fractured neck of a femur. His method varied according to the particular bone involved, but sawing to allow locking together of fragments, drilling, and binding with silver wire, which was twisted and hammered to secure the ends, was typical of the steps taken to deal with, say, the shaft of a femur.

By 1894 Arbuthnot Lane³ reported on his method of screwing together fractures, but the operative treatment of simple fractures met with violent opposition, which continued until the British Medical Association appointed a committee to 'report on the ultimate results obtained in the treatment with or without operation'. The conclusions of this committee were presented in 1912, and supported Arbuthnot Lane's operative method. This method of treatment was also accepted in Germany a few years later.

The most commonly used orthopaedic implant is the surgical screw, which requires that the quality of the drilling process should be at such a level as to ensure that the subsequent holding power of the screws is sufficient to maintain stable fixation. Mismatch between hole size and screw diameter may leave insufficient bone to engage the screw or alternatively may demand excessive pressure for insertion of the screw in a hole that is too small, with a danger of breaking the screw or damaging the bone.

Compression fixation, in particular, demands a high level of precision with regard to the position and stability of fixating screws. The recommended levels of compression⁴ are achieved with an axial movement of 10 to 20 μm after contact between the fracture fragments. Consequently, a loss of bone at the screw site could easily negate any beneficial effect of this type of device.

Ansell and Scales⁵ considered the relative merits of self-tapping and non-self-tapping screws and also the relationship of the drill size and the screw core diameter. Screw breakage or bending was found to be associated with an incorrect ratio between drill and screw minor diameter, the insertion of screws into incorrectly tapped holes, a lack of alignment of the screw in a hole, particularly when the distal cortex is engaged, and a seizing of a screw when the distal cortex is engaged without clearance in the hole through the proximal cortex.

They found also that less torque is required to insert self-tapping screws than non-self-tapping screws, and that intermittent rotation increased the torque more noticeably for non-self-tapping screws. The visco-elastic nature of bone and the lack of cutting during insertion of non-self-tapping screws result in the mainly frictional torque between screw and bone increasing as the bone relaxes onto the thread when rotation stops.

Bone drilling

The drills and screws first used in orthopaedic surgery were described by Lane in 1914. The drills were of a flat pattern used by carpenters and corresponded in size to the screw which was to be used, except where

abnormally soft bone was involved, when a smaller drill was used. Standard wood screws were used with the addition of a thread up to the head, so that they held only the proximal layer of the compact bone, the screw not being allowed to enter the distal cortex.

By 1957 orthopaedic surgeons became interested in the advantages claimed by Ernst and Haggarty for the spiral-point drill. Using this type of point, they demonstrated the resistance to slipping of the drill point when it enters the slippery round surface of the bone.

With an included angle of 72° , bone entry at an angle of 45° to the surface was achieved, and a spiral-point surgical drill with a 72° point was used in a hip-pinning operation at the Drake Memorial Hospital in Cincinnati, in December 1957.

Working with the Greenfield Tap and Die Company in 1959, Bechtol, Ferguson and Laing⁶ established design recommendations for orthopaedic drills. They stated that the drill should not 'walk' on the bone surface, so that accurate placement in the centre of the holes in metal implant plates could ensure that the plate was undamaged and that the screw was accurately centred. This factor again influences the process after the drill has penetrated the first cortex of the bone and has commenced entry of the second cortex. By this time it is firmly held in the near cortex and point wander may cause drill breakage.

These workers also stated that since bone is similar to a plastic material, the 90° drill point used to prevent chipping of plastics should be used for the drilling of bone. They claimed that such a point would allow contact with the second cortex of a bone at a greater angle than is possible with the more common 118° point used in drilling metals.

They drew attention to the advantage of having zero rake on the cutting face of the drill, which was claimed to prevent the drill from catching as it emerged from the cortical surface of the bone and thus chipping the surface.

Dulling of the sharp edges of drill flutes was recommended to minimize the reaming of holes by drill wobble to an unacceptable oversize.

Finally, they recommended a drill which was not corroded in the sterilization process in an autoclave, and which was not brittle. Stainless steel and vitalium were considered to be acceptable drill materials. Their recommendations were endorsed by the Association for the Study of the Problems of Internal Fixation (ASIF).

The majority of the twist drills used in bone surgery today are similar to the standard chisel point drill used in the drilling of metal, often with the reduction of the point angle to 90° and the removal of the cutting lands, as recommended by Bechtol. To avoid corrosion, they are usually made from stainless steel or a cobalt chrome alloy.

In 1965 R. S. Sneath⁷, who was dissatisfied with existing orthopaedic drills, carried out research with groups of drills, having four-facet points and spiral points, with and without lands.

The point angle was varied from 60° to 160° , with a 15° clearance angle, and hand-drilling of beef bones, using various motive power was followed by measurement of the rate of penetration through the beef

bone with a constant drill speed and thrust force. The experiment was repeated with human bone, on the assumption that the most efficient drill would penetrate a given thickness of bone in the shortest time.

From this work, Sneath was of the opinion that cutting lands are better than no lands, that the spiral-point shape is better than the four-facet shape, and that an included point angle of 130° to 140° gives optimum results.

The distance travelled by the drill after completion of the hole depends on the surgeon. The smaller this distance is, the less the unnecessary tissue damage and risk to vital structures. Sneath considers this distance to be proportional to the thrust force necessary to advance the drill so that a low thrust force, compatible with a satisfactory rate of drilling, ensures that the surgeon's control over the drill's direction and distance of travel is optimum.

Sneath assumed that heat production in the bone was proportional to thrust force and that any excessive heat would char the organic matter of the bone, blocking the drill flutes and preventing the removal of debris.

However, extensive work on the effect of drill performance on temperature has been carried out by Matthews and Hirsch⁸, drilling only the first cortex of the bone since no change in temperature was noted when the second cortex was drilled.

Drilling speeds ranging from 345 to 2900 rev/min and drilling thrust forces from 19.6 to 117.7 N were used with 3.2 mm twist drills.

The highest average maximum temperature was 93.1°C at a drilling speed of 2900 rev/min and a force of 19.6 N. In every case an increase in the drilling thrust reduced the average maximum temperature. There were no significant differences in maximum temperature attributed to drill speed. The use of water as a cutting fluid was very effective in reducing the cortical temperature.

Bonfield and Li⁹ have produced evidence that damage to bone occurs at 50°C . They found that the mechanical properties of bone which had been heated to this temperature were irreversibly altered. The association of an increase in the penetration rate, with a reduction in heat, also changed the distribution of heat, with a greater proportion entering the chips and less in both the drill and the bone specimen.

It was concluded that the heat generated when drilling was due to friction and to fragmentation of particles of bone at the cutting edge of the drill. This was said to be related to the total number of revolutions required to complete the hole. When drilling with a higher thrust force, the drill cut deeper, penetrated more rapidly and completed the hole in fewer revolutions. The smaller movement of the cutting edge, with less cutting at a higher force level, together with the redistribution of heat into the chips, was assumed to explain the decrease in temperature and the decreased duration of temperature rise when the drilling force was increased.

Contrary to these results, in 1971 Eichler and Berg¹⁰ found that an increase in the thrust force produced an increase in the heat generated when drilling compact bone. An increase in drill diameter

achieved a substantial reduction in the temperature rise.

EXPERIMENTAL PROCEDURE

The test drills were 6.35 mm diameter high-speed steel with fast, normal and slow helices, and comprised two groups. Three forms of drill point geometry were used in the first group, these being radial-relief, four-facet and spiral-point. All the drills in this group had a web thickness of 17.5 per cent of drill diameter. The drills of the second group had radial-relief points and the web thickness at the point varied from 0.74 mm to 1.60 mm. The first group had point angles ranging from 80° to 150° on the three point shapes and the second group, all with 118° radial-relief points, were eventually point-thinned.

Torque and thrust measurements were carried out on an air bearing dynamometer (see figure 1). The workpiece material was pig femora (figure 2), which is similar in many respects to human bone. This could

be obtained fresh in large quantities, the animals being approximately all the same age and size. The animals were killed at regular intervals, and as soon as the bone was available it was stored in Ringer's physiological saline solution at room temperature. This is a solution of the chlorides of sodium, potassium and calcium. A number of bones were sectioned to obtain the mean thickness of the cortex, which was 5 mm.

Drilling was confined to the centre portion of the shaft of each bone, and a different drill was used for each hole to avoid the over-influence of a particular drill on a set of results.

Hand-operated drilling machines produced good chip formation with fast-helix drills, but there was a tendency to snatch on break-through. With spiral-point drills the thrust was greatly reduced and the torque decreased as the point angle increased. The locating ability of the spiral point was good up to a point angle of 140° .

A hand-operated electric drilling machine resulted in conventional and four-facet drills snatching on

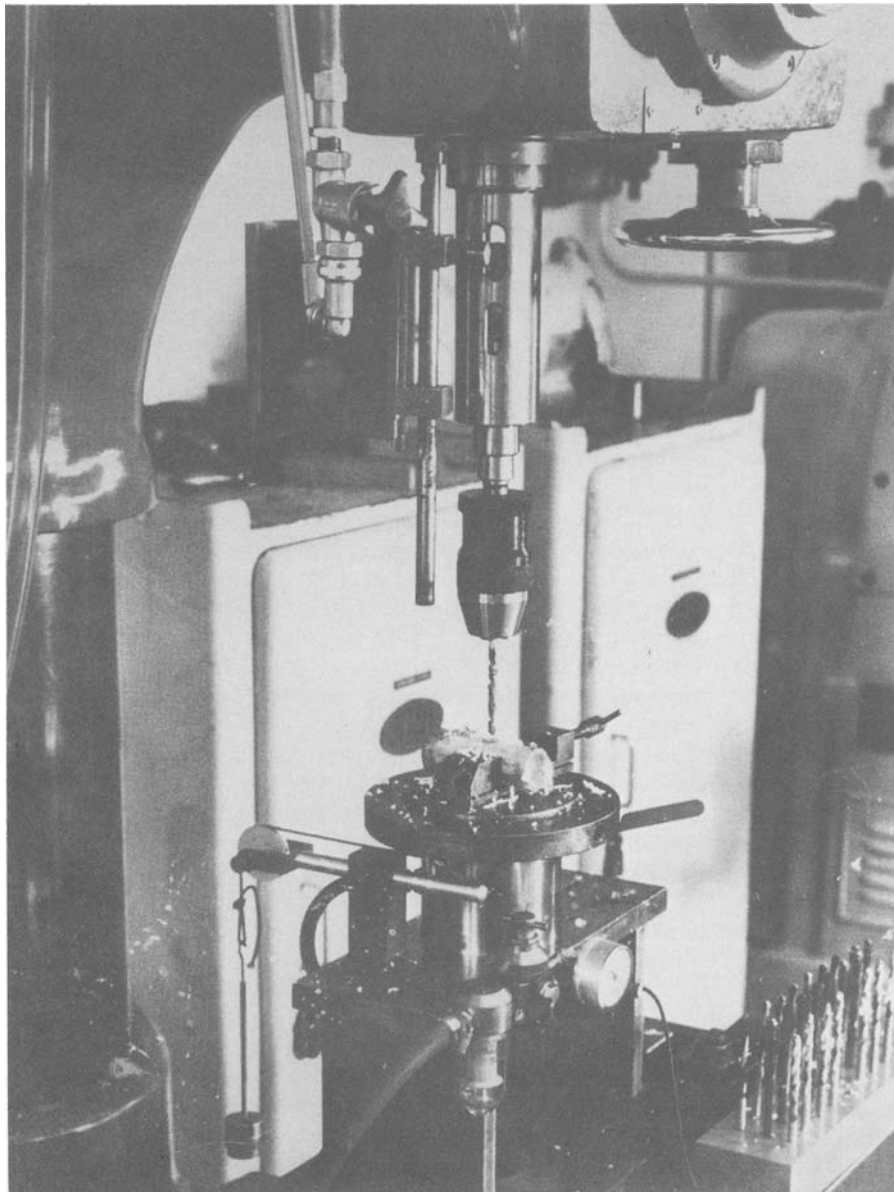


Figure 1. Air bearing dynamometer with section of bone material.



Figure 2. The workpiece material was pig femora of the size indicated.

break-through, and feed control was difficult. The locating ability decreased as the point angle increased. Spiral-point drills located reasonably well at all point angles, and at 120° to 140° break-through was not so severe and snatching was small, particularly with slow-spiral drills. All subsequent tests were carried out on pillar drilling machines.

Constant load tests enabled the rate of penetration of the drill to be measured in passing through the proximal cortex, when subjected to a thrust force of 19.6 N. With radial-relief points the time required for fast-helix drills to penetrate the bone cortex increased rapidly with point angle from 0.25 min for an 80° point, to over 3.5 min for a 150° point when penetrating a distance of 1 mm.

The normal-helix drill achieved a more rapid penetration. The time for 1 mm ranging from 0.12 min for an 80° point to 2 min for a 150° point. Slow-helix drills could be compared with the fast-helix drill, with a penetration time of over 0.25 min at the 80° to 90° end of the point angle range, to almost 3.5 min at a 150° point angle.

Fast-helix 80° four-facet points took 0.12 min for 1 mm penetration and this increased to nearly 2.25 min for a 150° point. The normal helix drill penetrated 1 mm in 0.12 min, and the 150° point entered the same distance in 1.75 min.

Again, 0.12 min achieved 1 mm penetration when using an 80° slow-helix drill, and the 150° point took less than 1.25 min for the same depth of hole.

When using a fast-helix spiral-point drill, the point

angle had only a slight effect on the penetration rate. The majority of the holes achieving a depth of 1 mm in not more than 0.12 min. Normal-helix drills required 0.25 min with an 80° point, 0.12 min at 90° and very much less time for larger angles. The time remained constant up to 150° .

A slow-helix pattern of drill, showed some variation in penetration time from 0.37 min for an 80° point angle, down to 0.12 minutes for 90° and 100° points. The time required again fell with increasing point angle to 0.12 min at 150° . (See figure 3.)

Constant feed tests were carried out at a feed of 0.102 mm/rev and a speed of 456 rev/min, the drills passing through both proximal and distal cortices.

The torque required by radial-relief points fell steadily as the point angle was increased, as shown by figure 4. The effect of helix angle was negligible. In drilling with four-facet points the torque remained constant with changes in point angle and the helix angle had little effect. Fast and normal-helix spiral-point drills produced approximately the same value of torque, which fell slightly as the point angle was increased.

The thrust force developed by radial-relief points increased as the point angle increased, the pattern being similar for the three helix angles tested. Again, as in the case of the radial-relief points, the four-facet drills indicated an increase in the thrust force required to feed the drill through the bone cortex as the point angle increased. A change in helix angle had little effect.

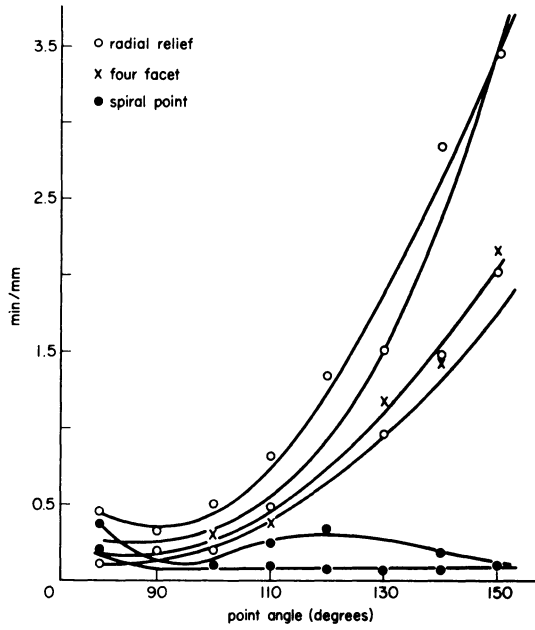


Figure 3. Rate of penetration of drills under constant load, with different point forms and point angles.

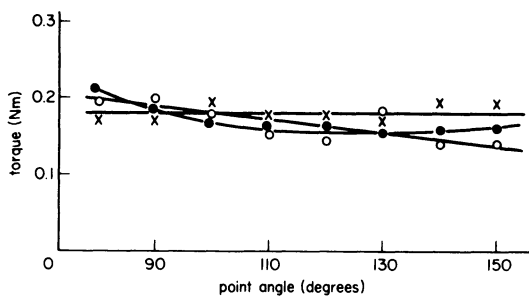


Figure 4. Influence of point angle on torque, with different point forms and helix angles.

Changing the helix angle from fast through normal to slow on drills with spiral points resulted in an increase in thrust. Point angle changes appeared to

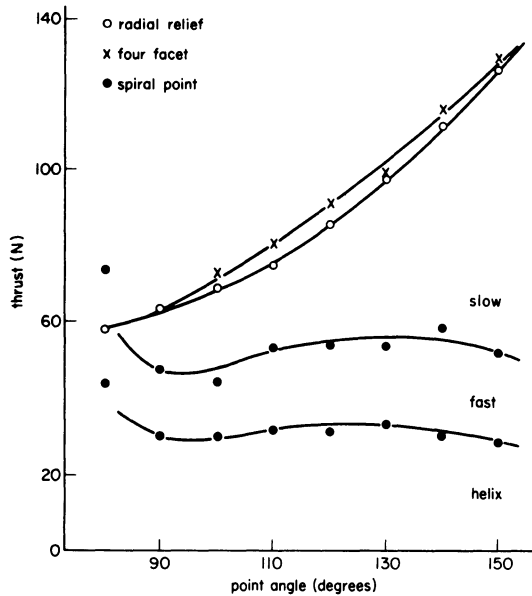


Figure 5. Change in thrust force due to point angle, helix angle and point form.

produce a maximum thrust in the region of 120°, with a small decrease at higher and lower point angles, together with a big increase in the region of 80° to 90°; see figure 5.

During the tests with varying web thickness, as expected an increase in the web thickness produced an increase in the thrust force, the fast-helix drill of

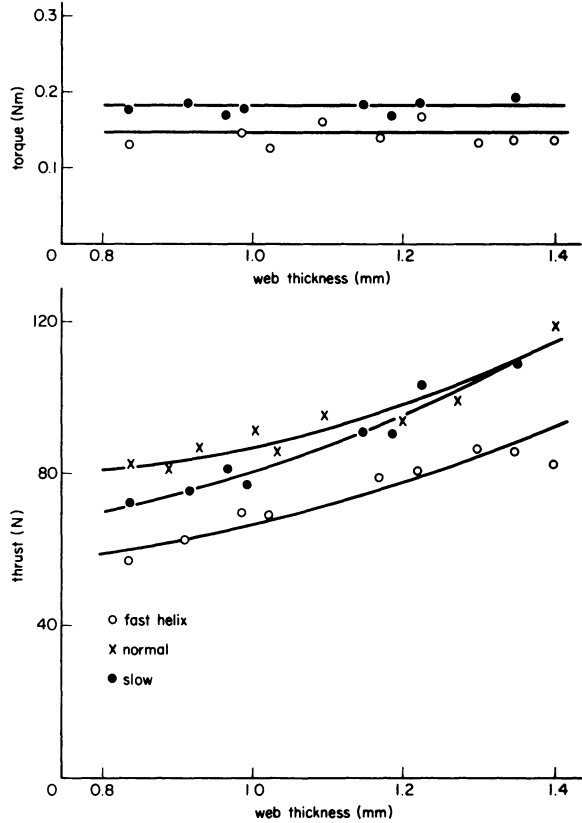


Figure 6. Thrust and torque changes resulting from web thickness and helix angles.

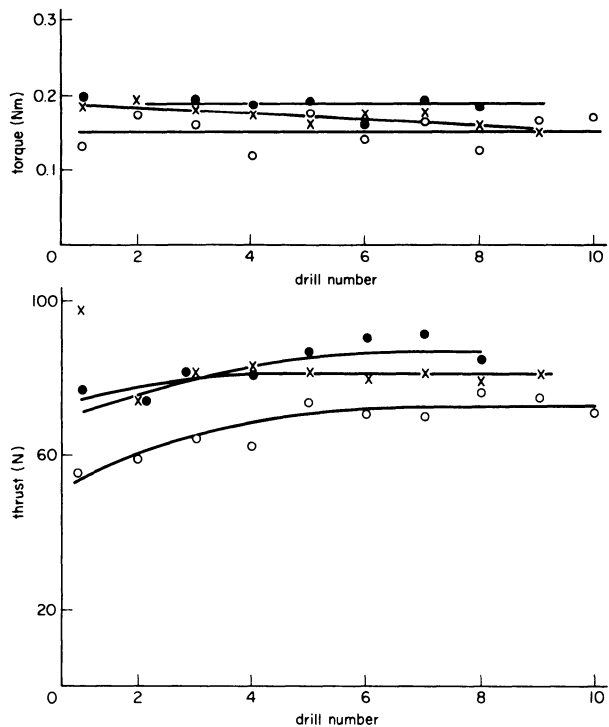


Figure 7. Point-thinning effects on thrust force and torque.

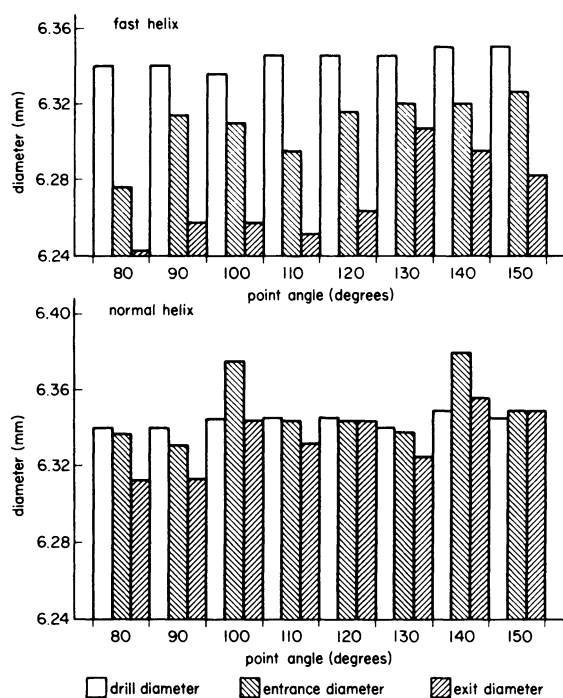


Figure 8. Influence of point angle and helix angle on the diameter of the finished holes.

course producing a lower thrust than either the normal-helix or the slow-helix drills. The torque was unaffected by the change in web thickness, but increased as the helix angle was reduced from fast to slow, as shown by Figure 6.

Point thinning becomes effective only over a certain web thickness. There appears to be little reason to use extremely thin webs, as below a web thickness of 1.13 mm a further thinning of the web had only a slight effect on the cutting forces. Above this value there was a considerable reduction in thrust, brought about by the decrease in web thickness, and in most cases a slight increase in torque as a result from the lengthening of the cutting edges. A point-thinning operation increases the possibility of inaccurate drill grinding affecting the performance of a drill. The graphs in figure 7 show the effect of point thinning on torque and on thrust.

The helix angle of the drills had more effect on the diameter of the finished holes, than the size of the point angle; this is shown by the graphs in figure 8. The fast-helix drill cut the bone most effectively, the holes closing in after drilling, because of the visco-elastic property of the bone. The normal-helix drill produced a hole similar in size to the drill diameter, and the slow-helix drill tended to cut oversize when using four-facet and spiral-point geometry

CONCLUSIONS

From the results it may appear that fast-helix drills would have the most satisfactory performance in

bone. However, drilling experience shows that this helix tends to catch on break-through rather more than either the normal or the slow-helix drills, even though slow-helix drills had the worst cutting action of the three helix types tested. Again, the tendency for drills with small point angles to jam as they pass through the cortex is more obvious when four-facet points are used than is the case with radial-relief points.

Further problems of jamming occurred in passing through the cortex of the bone when using small point angles in the region of 80° to 90° , even though this point angle is normally recommended for surgical drills.

From a consideration of the thrust force values, the rate of penetration, the torque and the quality of the holes produced, it is recommended that a point angle of between 120° and 140° is the most satisfactory angle, associated with a spiral point, for this type of surgery.

The helix angle should be normal, that is about 27° , the web thickness up to 18 per cent of the drill diameter, and the clearance angle 15° . These recommendations reinforce those of R. Sneath and differ from existing standards.

REFERENCES

1. C. O. Bechtol. *Engineering Principles Applied to Orthopaedic Surgery*, *American Academy of Orthopaedics* (1952) 9.
2. P. B. Beeson and Walsh McDermott. *Textbook of Medicine*. Vol. 2, W. B. Saunders Co., USA, 1967.
3. W. A. Lane. *The Operative Treatment of Fractures*. The Medical Publishing Co. Ltd., 1914.
4. S. M. Perren, et al. The reaction of cortical bone to compression. *Acta Orthopaedica Scandinavica Supplementum* 125, 1969.
5. R. H. Ansell and J. T. Scales. A study of some factors which affect the strength of screws and their insertion and holding power in bone. *J. Biomechanics* (1968) 1.
6. C. O. Bechtol, A. B. Ferguson and P. G. Laing. *Metals and Engineering in Bone and Joint Surgery*. Bailliere, Tindall and Co. Ltd., 1959.
7. R. S. Sneath. The determination of optimum twist drill shape for bone, in *Biomechanics and Related Bio-Engineering Topics*, Pergamon, 1965.
8. L. S. Matthews and C. Hirsch. Temperatures measured in human cortical bone when drilling. *J. Bone and Joint Surgery* (1972) 54A, 2.
9. W. Bonfield and C. H. Li. The temperature dependence of the deformation of bone. *J. Biomechanics* (1968) 1.
10. J. Eichler and R. Berg. *Temperatureinwirkung auf die Kompakta beim Bohren*. Gewindefschneiden und Eindrehen von Schrauben, Orthopadische Klinik, Giessen, 1971.

MACHINE TOOL DYNAMICS

THE DYNAMIC CHARACTERISTICS OF EPOXY RESIN BONDED MACHINE TOOL STRUCTURES

by

M. I. CHOWDHURY*, M. M. SADEK* and S. A. TOBIAS*

SUMMARY

A new way of fabricating machine tool structures was developed as an alternative to welding. It involves the use of epoxy resin as a bonding agent between structural components. The overarm of a horizontal milling machine was manufactured according to this technique and its performance under dynamic tests and cutting conditions was compared with, respectively, those of a cast and of a welded overarm. The bonded overarm permitted a substantially higher metal removal rate before the onset of chatter, this being due to its very much higher damping capacity. Production cost estimates of the bonded overarm were considerably below those of the welded type. The fabrication technique developed can be used for modifying existing structures. An example is discussed.

INTRODUCTION

Welded machine structures are expensive to manufacture and require heat treatment for stress relieving before machining. Welded structures, once completed, cannot be modified since this leads to distortions.

In this paper, the authors discuss a new method for fabricating machine tool structures, which involves the use of epoxy resin as a bonding agent between the structural components. The advantages of this technique are a reduction of production costs not only as far as the labour content is concerned, but also by eliminating the necessity for subsequent stress relieving. The technique permits the modification of existing welded or cast structures after heat treatment and machining, by the incorporation of stiffening ribs or damping or sound absorbing layers.

The basic idea was first applied in a pilot study to the manufacture of a model Warren beam. This showed that, in comparison with a welded beam, a bonded component of identical construction generally has a lower stiffness, but its damping capacity is substantially higher. However, the stiffness of the bonded beam was easily raised far above that of the welded type by the introduction of stiffening ribs.

The feasibility of the technique having been established, it was applied to the manufacture of the overarm of a small horizontal milling machine. The dynamic performance of this, in relation to a cast and welded component, was assessed by using forced vibrations as well as chatter tests. The bonded overarm proved to perform very substantially better than the other two types, permitting a very much

higher rate of metal removal before the onset of chatter. Further improvement was achieved by bonding sound deadening steel sheets to the top surface of the bonded overarm. Moreover, an estimate of production costs showed that a bonded overarm is substantially cheaper than a welded type, mainly because of the very much lower labour content.

EPOXY RESIN USED AND PREPARATION OF BONDED SURFACES

The epoxy resin used in this investigation was Araldite AY 103 + HY 951 hardener. This is a cold curing bonding agent which, after mixing, remains usable for 3 to 4 hours and becomes fully cured, at room temperature, after 24 hours. Its modulus of elasticity, when fully cured, is $0.75 \times 6 \text{ lb/in}^2$ and its shear modulus is 2600 lb/in^2 .

Surfaces to be bonded were prepared by sand blasting and subsequent degreasing with carbon tetrachloride.

The bonded components were manufactured by applying the bonding agent to the cleaned surfaces and assembling the component in a simple jig, which pressed adjacent plates and rolled sections to one another. After the component was fully assembled, small tack welds were applied to join adjacent outside plates. The component was then removed from the jig and allowed to cure.

The tack welding was applied to permit the removal of the component from the jig before it was fully cured. It did not contribute in any measurable extent to the stiffness of the component and it is in no way an essential feature of the technique used, having been applied solely to facilitate handling.

* Department of Mechanical Engineering, University of Birmingham

PILOT STUDY

The general outline of the encastré Warren beam used is shown in figure 1(a). This was manufactured by welding and by bonding, the design in both cases being identical.

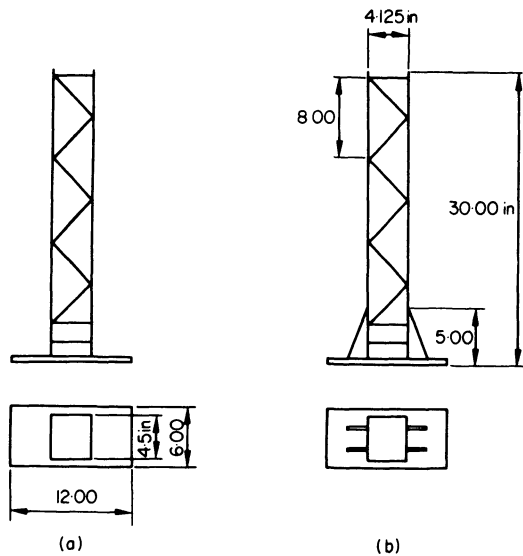


Figure 1. Welded/bonded beam without stiffeners (a) and modified bonded beam with stiffeners (b).

The two beams were tested by applying a static or dynamic load at the free end and the following results were obtained:

	Welded Beam	Bonded Beam
Static stiffness for load at free end	19×10^4 lb/in	13.1×10^4 lb/in
Natural frequency	155 Hz	135 Hz
Damping ratio	0.039	0.041

As expected, the bonded beam had a lower stiffness (and a lower natural frequency) than the welded beam, this no doubt being due to the lower shear modulus of epoxy resin (2600 lb/in^2) with respect to mild steel ($16\,000 \text{ lb/in}^2$). However, by bonding two pairs of small ribs to the base end of the bonded beam, in the manner shown in figure 1(b), the conditions were reversed to the extent that the stiffness of the bonded beam was increased to be 200 per cent above that of the welded type.

DESIGN OF BONDED OVERARM

The box overarm of a small horizontal milling machine was chosen on two grounds. Firstly, a cast iron and a fully welded component of the same type were already available, which made a comparison of the three manufacturing methods possible. Secondly, the particular horizontal milling machine had been tested on previous occasions¹ and its dynamic behaviour under tests and cutting conditions was fully explored.

It was known from the experience gained in the pilot study that welding cannot be substituted with bonding without some substantial redesign. Bonding requires large contact areas to avoid peeling of the epoxy resin. Moreover, for keeping material costs low, it is clearly an advantage to construct the bonded overarm from readily available, standard channel and I section beams. Thus, the original cast

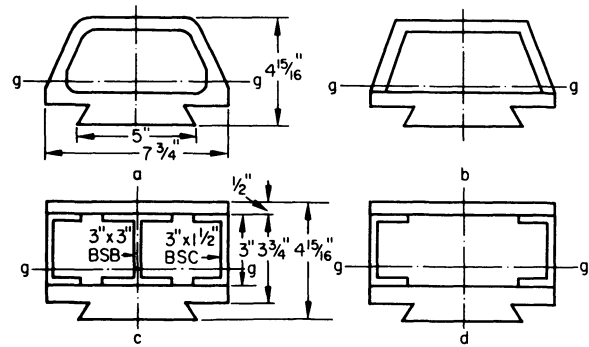


Figure 2. Cross-section of overarms

	Area (in ²)	I _{gg} (in ⁴)
(a) Cast iron	13.09	27.71
(b) Welded	13.58	26.4
(c) Bonded	17.44	49.65

cross-section of the overarm, shown in figure 2(a), was modified into the form shown in figure 2(c). The figure also contains the cross-section of the welded overarm, see figure 2(b). The cross-sectional area *A* and its second moment *I_{gg}* of the three designs are also given in the figure. Note that, for the bonded component, both were substantially larger than for the welded design, mainly because of the central I section in figure 2(c). This was quite deliberate since, on the basis of the pilot experiment, it was considered to be desirable to compensate for the loss of stiffness due to the lower shear modulus of the bonding agent. It will be seen that this precaution was unnecessary from the point of view of the relative dynamic performance.

EVALUATION OF DYNAMIC PERFORMANCE WITH FORCED VIBRATION TESTS

(a) Method of evaluation

The dynamic performance of the three overarm designs (figures 2(a) to 2(c)) was evaluated by determining their coefficient of merit (*CoM*) for a range of values of the angle between the representative cutting force *P* and the normal to the machined surface *X* using a technique which was described in detail in a previous publication¹. As shown there, the *CoM* is a measure of the chatter resistance of the machine, being proportional to the maximum width of cut which is stable at all speeds. The factor of proportionality is a constant dependent only on the work piece material, cutter design, cutting conditions (excluding cutting speed), but not on the machine characteristics.

The *CoM* is derived from the operative receptance of the system, this being the cross receptance between the representative cutting force *P* and the normal to the machined surface *X*. The operative receptance is derived from the direct and cross receptances of the machine measured by forced vibration tests. The technique of measurement was discussed in reference 2.

The representative cutting force *P* acts in the middle of the arc of contact and is equivalent to the dynamic action of all the cutting force elements generated. The normal to the machined surface *x* connects the centre of the arc of contact with the

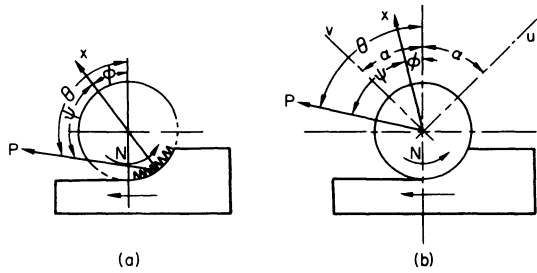


Figure 3. Schematic representation of a milling operation. (a) Orientation of the cutting force P and the common normal x in up-milling operation and (b) Orientation of the directions of excitation u and v with respect to the directions of the cutting force P and common normal x .

centre of the arbor. The orientation of the directions P and x is shown in figure 3(a), which also defines the angle between them ψ . The orientation of P and x is dependent on the depth of cut d , and the cutter radius R (on d/R), which affect the angles θ and ϕ .

Once the operative cross receptance for a particular condition was found, the CoM is defined

$$CoM = \frac{1}{2XO} \quad (1)$$

where XO is the maximum negative in-phase component of the operative receptance. As a typical example, figure 5 shows the operative receptance of the cast, welded and bonded overarms for a particular condition (defined by ψ and d/R). The meaning of XO is clear from this figure.

When assessing the relative improvement achieved by design changes in a particular machine, it is convenient to measure this in terms of the relative improvement coefficient RIC , which is defined as

$$RIC_{AB} = \frac{CoM_A - CoM_B}{CoM_B} \times 100 \text{ per cent} \quad (2)$$

the subscripts A and B refer to the two machines compared. The RIC , giving the improvement of machine A with respect to machine B , is plotted as a function of the depth of cut ratio d/R for a series of values of the angle ψ between P and x . Using this technique, the final results of the dynamic tests will be presented in terms of the RIC , giving the improvement achieved by the bonded overarm in relation to the welded and cast types.

(b) Results

Figure 4 shows the direct and cross receptances of the milling machine, measured in the directions u and v as defined in figure 3(b), for the three types of overarm used. From these receptances, the operative receptance loci are calculated by the use of the method given in ¹, a set of corresponding to angles $\theta = 87^\circ$ and $\phi = 22^\circ$ and $d/R = 0.3$ (see figure 3(a)) having been presented in figure 5. From a series of curves of this type, the CoM is found by using Equation (1) and this can be plotted as a function of the depth of cut ratio d/R for a series of values of the angle $\psi = \theta - \phi$ (see figure 3(a)). Combining two such graphs yields the relative improvement coefficient RIC and the variation of this for the bonded overarm relative to the cast iron type is presented in figure 6(a) and that of the bonded type in relation to the welded component in figure 6(b).

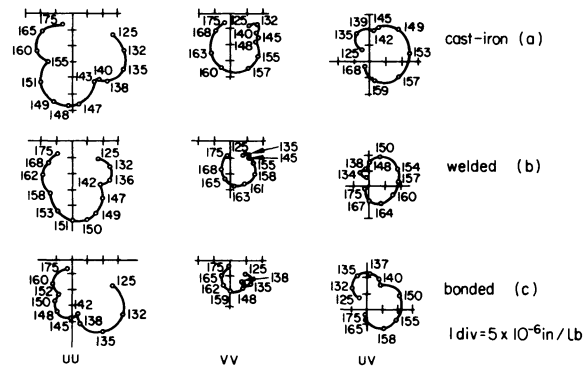


Figure 4. Direct and cross receptance loci for milling machine. (a) Cast iron; (b) welded and (c) bonded overarms.

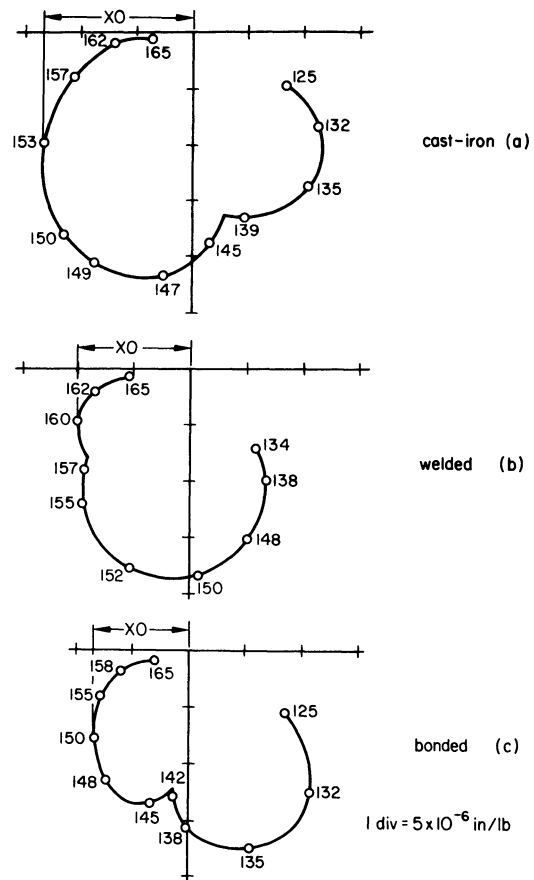


Figure 5. Operative receptance loci for milling machine. (a) Cast iron overarm; (b) welded overarm, and (c) bonded overarm. Cutting condition represented by $d/R = 0.3$ and $\psi = 65^\circ$

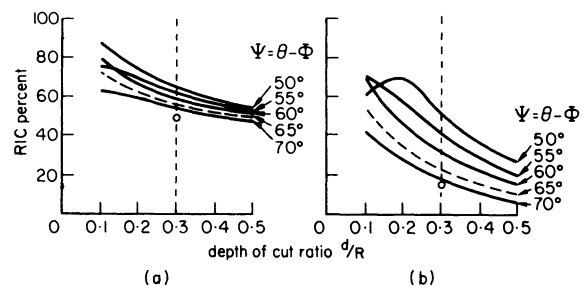


Figure 6. Relative improvement coefficient, RIC . Curves of bonded against (a) cast iron and (b) welded overarm.

As can be seen from figure 6(a), the technique predicts that the bonded overarm should have a CoM which is 46–88 per cent above that of the cast type, depending on the cutting conditions. In other words, with the bonded overarm, the maximum width of cut which is free of chatter at all speeds is 46–88 per cent above that arising with a cast type. The analogous result for the bonded overarm in relation to the welded type (see figure 6(b)) indicates that the former should allow a chatter free width of cut which is 8–70 per cent greater than that obtained with the latter, depending on the cutting conditions.

EVALUATION OF DYNAMIC PERFORMANCE WITH CHATTER TESTS

(a) Experimental technique

The verification of the predicted relative performance, as summarized in figures 6(a) and 6(b), is achieved by chatter tests determining stability charts of the machine with one of the three overarms in situ for particular cutting conditions. From these stability charts, the maximum width of cut w_{mo} which is stable at all speeds, is extracted and these values, in conjunction with Equations (1) and (2), can then be used to verify the appropriate curves in figures 6(a) and 6(b).

Chatter tests were carried out with an inserted carbide blade cutter with the following specification

cutter diameter $2R = 4$ in. width of cutter = 0.5 in.
 number of blades $z = 16$ spiral angle = 0°
 rake angle = 10° relief angle = 5°

The depth of cut d chosen for the tests was $d = 0.6$ in, which resulted in a depth of cut ratio of $d/R = 0.3$, corresponding to an angle $\phi = 22^\circ$, as defined in figure 3(a). For these cutting conditions, the angle between the representative cutting force P and the normal to the machined surface x was $\theta = 87^\circ$, which means that $\psi = \theta - \phi = 65^\circ$. The curve appropriate for this value of ψ has been chain dotted in figures 6(a) and 6(b).

Cutting tests were performed on strips of blank mild steel 6 in long, with a cross-sectional dimension of $\frac{1}{2}$ by 2 in, which were tapered with an angle 2° , varying the width of cut in accordance with the needs of the test.

(b) Experimental results

The stability chart is obtained by plotting the width of cut w , corresponding to the threshold of stability as a function of the cutter speed¹. The stability charts thus determined for the three types of overarms are presented in figures 7(a) to 7(c).

The maximum width of cut stable at all speeds w_{mo} is obtained by drawing a horizontal asymptote to the boundary line between the unconditionally stable and conditionally stable region¹, which is a hyperbolic curve tangential to the unstable lobes, which is denoted by w_m in the figures. The value of w_{mo} for each of the three types of overarms is given

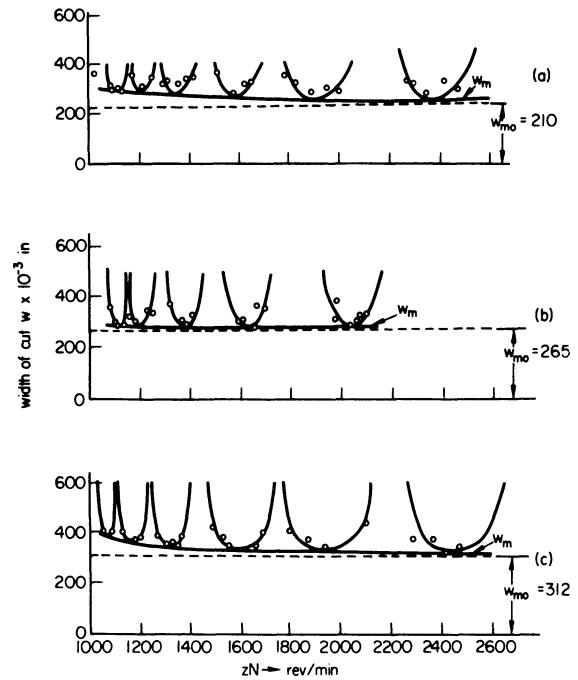


Figure 7. Experimental stability charts. (a) Cast iron overarm; (b) welded overarm and (c) bonded overarm.

in the appropriate figures. The relative improvement coefficient relating the bonded with the cast overarm is now obtained as

$$RIC_{AC} = \frac{w_{moA} - w_{moC}}{w_{moC}} \quad (3)$$

where A relates to the bonded and C to the cast overarm. A similar relationship gives the RIC for the bonded against the welded overarm. The RIC values thus obtained have been plotted in figures 6(a) and 6(b).

The RIC values obtained experimentally and those predicted from the operative receptances are summarized in Table 1.

TABLE 1 Relative improvement coefficient RIC

	Bonded/welded	Bonded/cast iron
Experimental RIC	18%	48.2%
Predicted RIC	24%	57%

Correlation between experimental and predicted values, for this particular cutting condition, appears to be reasonable. It is of the order of magnitude that can be expected¹ when forced vibration and cutting tests are involved.

The full verification of figures 6(a) and 6(b) requires much more extensive chatter tests, covering a range of values of the depth of cut ratio d/R and possibly also different cutters (to change the value of ψ). The experimental effort required for this is very large indeed, time consuming as well as costly in view of the large amount of material that has to be machined. Moreover, the repeatability of such tests is relatively poor in view of the variability of the machining characteristics. A relative evaluation based on forced vibration tests requires only a fraction of the experimental effort and gives reasonable results.

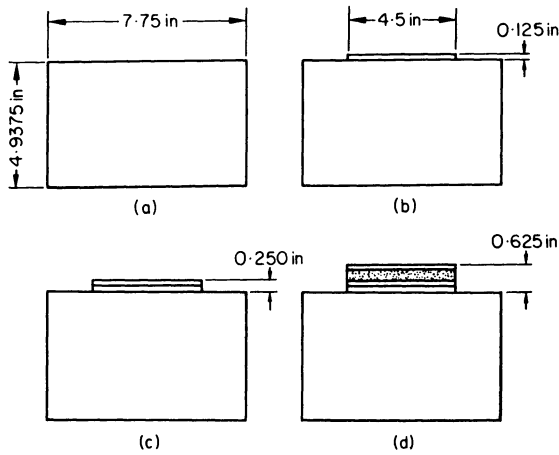


Figure 8. Bonded overarm with stiffening and damping layers. (a) Original bonded overarm, without any stiffening or damping layer. (b) Overarm with one layer of sound deadening steel. (c) Two layers of sound deadening steel. (d) Three layers of sound deadening steel and a PVC layer.

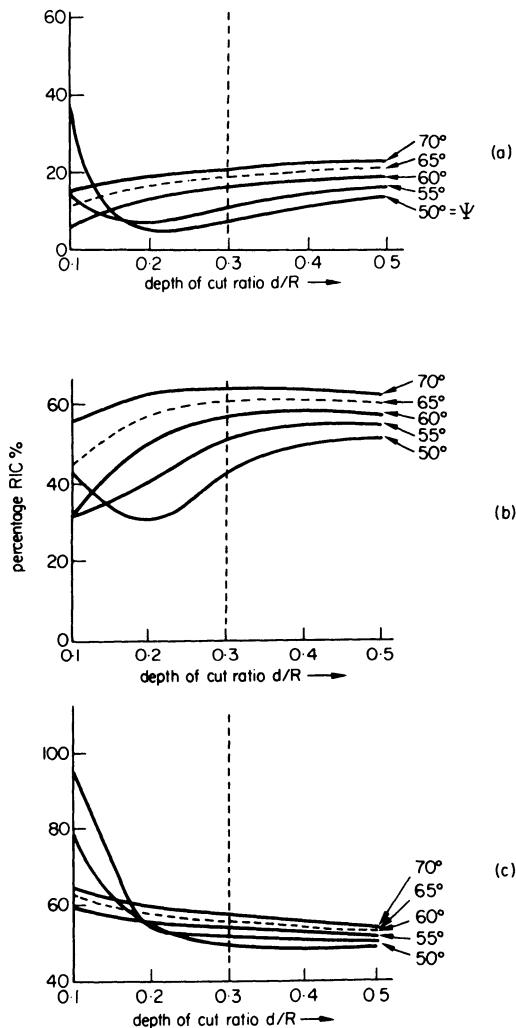


Figure 9. Relative improvement coefficient *RIC* curve against original bonded of that: (a) with one layer of sound deadening steel; (b) with two layers of sound deadening steel and (c) with three layers of sound deadening steel and a PVC layer.

MODIFICATION OF EXISTING STRUCTURES

One of the main attractions of the bonding technique is that it permits the modification and improvement of existing structures. This was demonstrated by modifying the bonded overarm in three stages in which composite steel and visco-elastic laminates were bonded to its top surface, symmetrically about the longitudinal axis (as indicated in figure 8) in the following manner:

- Stage 1.* Addition of one layer of sound deadening steel (see figure 8(b)). This was 0.125 in thick and contained two outer steel layers sandwiching a 0.005 in thick visco-elastic layer.
- Stage 2.* Addition of a second layer of the same material (see figure 8(c)).
- Stage 3.* On top of the previous two layers, a layer of flexible PVC compound 9503 visco-elastic material was bonded and, on top of this, a third layer of sound deadening steel was cemented, making the total thickness of the sandwich 0.625 in (see figure 8(d)).

The *RIC* for the stage 1 modification, determined with forced vibration tests and the methods of calculation already described, is shown in figure 9(a). As can be seen the improvement in the dynamic performance, achieved in relation to the original bonded design was between 5 and 33 per cent, depending on the cutting conditions, that is, the depth of cut ratio *d/R* and the cutting angle ψ . For the type of milling cutter used in the chatter tests previously discussed ($\psi = 65^\circ$) the increase in the chatter free width of cut would vary between 10 per cent (shallow cuts) and 20 per cent (deep cuts).

The *RIC* for the second stage of modification is presented in figure 9(b). Improvement with respect to the original bonded overarm was within 30 and 65 per cent. For the third stage of modification, the improvement was between 50 and 98 per cent.

COMPARISON OF THE DYNAMIC PROPERTIES OF BONDED, WELDED AND CAST OVERARMS

It has already been mentioned that the cross-section and therefore the weight of the bonded overarm was greater than that of the other two, as shown in Table 2.

TABLE 2 Weight of overarms

	Cast iron	Bonded	Welded	Modified bonded
Total weight lb	155	119	84	95

It might be concluded from this that the reason for the bonded overarm performing better than the other two was its greater static stiffness. This was not the case. Quite to the contrary, its equivalent static stiffness, which is the stiffness that counts from the vibration point of view, was actually lower than that of the other two. However, at the same time, its equivalent damping factor was very much higher and because of the interplay of these two, a better performance was obtained.

This conclusion is arrived at by an examination of the dynamic characteristics of that mode of vibration which is responsible for the chatter behaviour of the machine. The principle direction of this mode was practically the same for all three types of overarms, forming an angle of between 110° and 115° with the direction normal to the feed (vertical direction). If now the parallel receptance (due to a force acting in the principle direction) is determined from the direct and cross receptances presented in figure 4, the modal characteristics can be found by using the Kennedy and Panceau method³. The results of this evaluation are given in Table 3.

TABLE 3 Modal characteristics

	Type of Overarm			
	Cast iron	Welded	Bonded	Modified bonded
Natural frequency Hz	148	150	144	145
Equivalent stiffness lb/in × 10 ⁵	3.5	3.2	2.6	2.5
Equivalent weight lb	156	140	120	118
Equivalent damping factor × 10 ⁻²	2.85	3.59	4.86	4.14

It will be noted that the natural frequencies of the three overarms were very close to each other, in spite of their differing static weight, as given in Table 2. The equivalent static stiffness of the bonded overarm was 74 per cent of the cast and 81 per cent of the welded type. However, at the same time, the damping factor of the bonded was 171 per cent of the cast and 135 per cent of the welded component.

For a machine tool system which can be represented by a simple single degree of freedom system, the maximum width of cut which is stable at all speeds w_{mo} , is equal to twice the product of the equivalent stiffness and equivalent damping factor. Using the data contained in Table 3, this product is highest for the bonded overarm and lowest for the cast iron type. Such an elementary argument is valid for a complex vibratory system (such as the milling machine used) only as a first approximation but, nevertheless, it explains the superior performance of the bonded overarm.

With the aim of proving that the superiority of the bonded overarm was not due to its larger weight (and hence higher material cost) a modified version was fabricated, the cross-section of which is shown in figure 2(d). By eliminating the central I section, its weight was reduced to the value given in Table 2. This did not affect its modal characteristics appreciably, as can be seen from the last column in Table 3.

The relative improvement coefficient *RIC* of this modified design in relation to the original bonded type is presented in figure 10(a). The *RIC* values fall very largely into the negative range and, hence, the modified design is inferior to the original one, as was expected. Figures 10(b) and 10(c) show the *RIC* in

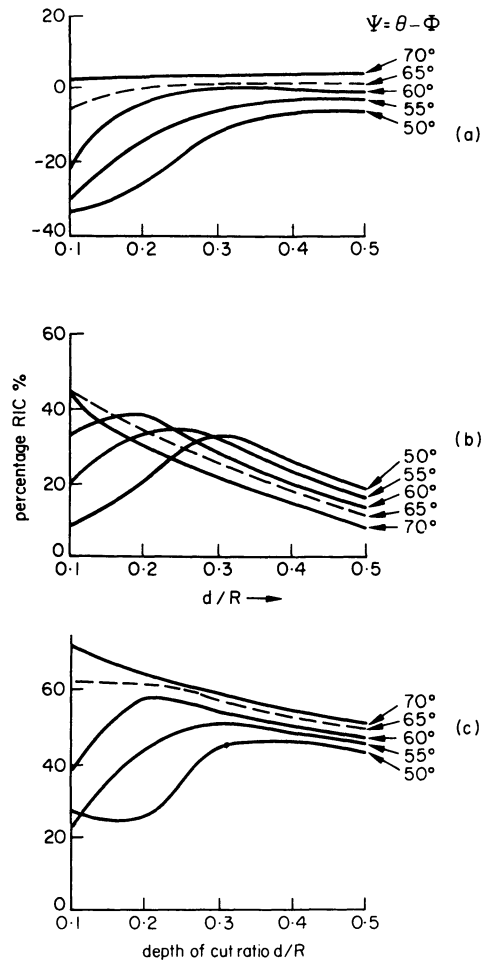


Figure 10. Effect of cross-sectional modification on the relative improvement coefficient *RIC* of: (a) modified overarm against original bonded overarm; (b) modified against welded overarm, and (c) modified against cast overarm.

relation to the welded and cast types. The modified version was still substantially better than the cast and welded components, the improvement being 22–72 per cent with respect to the former and 7–44 per cent as compared to the latter. The explanation is that the equivalent stiffness and damping of this design is only slightly lower than that of the original version.

RELATIVE COSTS

Realistic costs figures are difficult to arrive at, since they depend on the scale of production, that is, whether or not special jigs are justified. On a one-off basis, the production costs of the four types of overarm are estimated in Table 4.

TABLE 4 Estimated production costs of the four types of overarm

Cast iron	Welded	Bonded	Modified bonded
£77	£68	£45	£40

These figures cover material, labour costs and overheads. For the cast iron component, the cost of the pattern (about £20) was excluded.

When produced in batches, the cost of the cast iron overarm drops of £33. That of the welded type is not very much affected, since this is a highly labour intensive process, in which relatively little can be saved by special jigs. Roughly the same considerations apply also to the bonded type, although the preparation of surfaces for bonding and the application of the cementing agent can be mechanized.

The cost figures given cover production only. They do not take into consideration stock levels required for avoiding delays in delivery, which are particularly high for cast components and low for bonded ones. Bottlenecks in the supply of highly qualified labour, as required for welding, may also constitute an important consideration. The stress relieving heat treatment required for welded components may also be a bottleneck because of limited furnace capacity, either of throughput or of size. This suggests that bonding will be particularly attractive when the components involved are very large and/or when a large number of them is required at short notice.

CONCLUSIONS

The bonded overarm performed substantially better in both vibration and chatter tests than both the cast and the welded types. In addition to this, it was considerably cheaper to produce than the welded component.

The potential of the bonding technique was fully demonstrated in this paper. Its greatest potential is not with small components and for modifying or improving existing ones, but for large structures, which require a very large furnace for stress relieving.

It offers some new possibilities by allowing the manufacture of structures with varying plate thicknesses, as required by the Mitchell optimum design concept⁴.

ACKNOWLEDGMENTS

The original idea of fabricating machine tool structures by the use of bonding was proposed by Mr H. S. Royce. Thanks are due to Adcock and Shipley Limited for their assistance and continuing co-operation; the British Steel Corporation, and to B.P. Chemicals International Ltd., for providing the sound deadening steel and the visco-elastic damping material, respectively. The work is being supported by a grant from The Science Research Council.

REFERENCES

1. M. M. Sadek and S. A. Tobias. Comparative dynamic acceptance tests for machine tools applied to horizontal milling machines, *Proc. Inst. Mech. Engrs* (1970-71) **185**, 319.
2. M. I. Chowdhury, M. M. Sadek and S. A. Tobias. Determination of dynamic characteristics of machine tool structures. *Proc. Inst. Mech. Engrs.* (1969-70) **184**, 943.
3. C. C. Kennedy and C. D. P. Pancu. Use of vectors in vibration measurement and analysis. *J. Aeron. Sci.* (1947) 603.
4. A. S. L. Chan. The design of Mitchell optimum structures. College of Aeronautics, Cranfield Report 142, 1960.

ON-LINE IDENTIFICATION OF MACHINE TOOL RECEPTANCES FROM CUTTING EXPERIMENTS

by

N. H. HANNA* and A. W. KWIATKOWSKI*

SUMMARY

An on-line method for simultaneous four-channel data acquisition is developed and used for the identification of lathe receptances. Receptances obtained under simulated and actual cutting conditions were found to be different. The causes of this difference are discussed and it is concluded that a more accurate chatter prediction is achieved by using receptances obtained from cutting tests.

INTRODUCTION

In existing methods of identification of machine tool structure receptances, the cutting conditions are simulated using an exciter to provide the exciting forces. These forces may be either sinusoidal or random. The simulation, however, is usually incomplete due to design and/or measurement difficulties and, as a result, inaccurate receptances are obtained, as shown by Hanna and Kwiatkowski¹. Since ideal conditions exist only when a machine tool is performing a cutting operation, it is obvious that a more realistic determination of its structure receptances would be obtained from experiments carried out under these conditions. The investigation reported here is thus concerned with the identification of lathe receptances from cutting experiments, using the random components of the cutting forces as inputs to the system.

This investigation is part of a more comprehensive project aiming at establishing the technique of random force analysis as a reliable method of machine tool systems identification. It passed through the following stages of development: (1) checking the receptances obtained by random force excitation against those obtained by sinusoidal excitation; (2) testing the random force method under more realistic simulation of cutting conditions; (3) improving the technique of data acquisition by using an on-line digital computer and (4) dispensing completely with the exciter and using the random components of the cutting forces to provide the necessary excitation. The first and second stages were discussed in earlier publications^{1,2}; therefore only the third and fourth stages of development will be discussed in this paper.

The cutting forces and the resulting translational

response are, in general, three-dimensional space vectors having components measured in three mutually perpendicular directions. These reference directions are usually chosen in such a way as to simplify the design of the force and response measuring elements. In the case of a lathe, for instance, these directions are usually taken to coincide with the tangential, radial and axial directions with respect to a cylindrical workpiece. The model (or system) describing the machine structure would then have three inputs (the force components) and three outputs (the response components) and nine impulse response functions or equivalently receptances (one direct and two cross for each direction).

Although the three-dimensional (three input—three output) model is more general, its analysis requires a large computer. However, in practice some special cases may exist which can be exploited to reduce the required core capacity of the computer used. For instance: (a) if the three reference directions coincide with three principal axes of the structure, that is, the cross-responses between these three directions are negligibly small compared to the direct responses, then only three impulse responses are needed to identify the system. Moreover, each direction can be treated in isolation of the other two and the problem is simply reduced to three, one-dimensional systems, identified separately. (b) If one of the reference directions coincides with a principal axis, then five impulse responses (three direct and two cross) will represent the system. In this case, one-dimensional analysis may be used to obtain the direct receptance in this direction and two-dimensional analysis to obtain the four receptances in the other two directions and (c) if the response (both direct and cross) in one of the directions is very small,

* Department of Mechanical Engineering, University of Birmingham

then a two-dimensional model suffices. It was found that, in the case of the lathe under consideration, the response along the workpiece axis was negligibly small and therefore the latter case was applied.

DEVELOPMENT OF A METHOD FOR DATA ACQUISITION

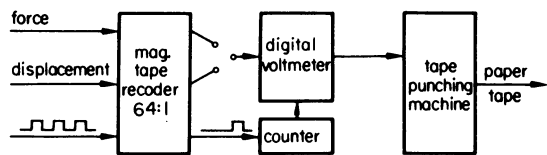
Figure 1 shows the stages of development of the data acquisition procedure. Stage 1 (figure 1(a)) was used in conjunction with a two directional exciter in which one force and one displacement were measured at a time. The two signals were recorded on two channels of a magnetic tape recorder and an identification square wave, at the required digitization frequency, was subsequently recorded on a third channel. Either of the force or displacement channels was played back at a much slower rate and the output, together

with the square wave of the third channel, were fed to a digital voltmeter which operated a tape punching machine. Since each signal was digitized separately, the square wave served as means for identifying the starting point on the tape as well as for controlling the rate of sampling. The number of samples taken was controlled by a mechanical counter. The reduction of the playback speed of the tape recorder was necessary because of the slow response of the tape punching machine. Since the reduction of tape speed was limited to 64 : 1 and the number of punched characters per second to 10, the maximum digitization frequency was, therefore, limited to 640 samples per one second of real time. Consequently, the maximum identifiable frequency was limited to 320 Hz. Moreover, misspunching of characters and/or miscounting of the number of samples was frequent and every time a defect occurred a useless tape was produced. This could only be discovered after the tape had been read and rejected by a digital computer which was a time wasting process.

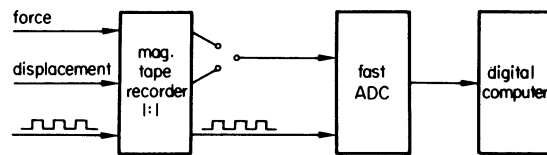
In the second stage, the force and displacement channels were also played back one at a time, but at the recording speed and into a fast analogue to digital converter (ADC) and the sampled data were stored in the core of a digital computer. Two main advantages were gained by this development: (a) the maximum digitization frequency was raised to 15 kHz, and (b) the errors in punching and counting the data have been eliminated. However, two main disadvantages still remained—the first was the contamination of signals by magnetic tape noise and the second was the need to digitize the signals separately, which was time consuming.

To overcome these disadvantages, the magnetic tape recorder was replaced by four simultaneous digitization channels, as shown in figure 1c. With this arrangement, two force signals and two displacement signals could be simultaneously digitized and stored in the computer in an on-line manner. The forces and displacements signals can be derived from either simulation or actual cutting experiments. The sequence of sampling is indicated in figure 1(d) and is explained as follows.

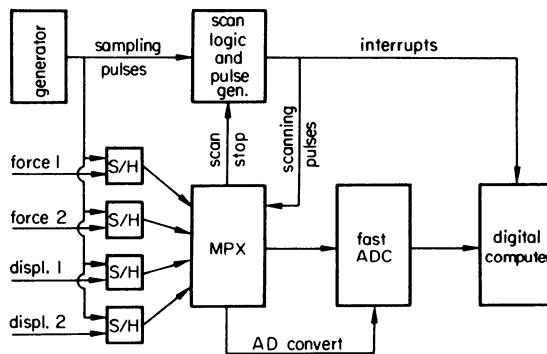
As shown in figure 1(c), the four signals are fed to four S/H amplifiers which are operated, that is, switched to the tracking or holding modes, by means of a train of pulses, from a pulse generator, at the required sampling frequency. Consider the first pulse of this train (figure 1(d)). The leading edge simultaneously sets all four S/H amplifiers to the tracking mode and the trailing edge resets them back to the holding mode in which they remain until the next sampling pulse arrives. The trailing edge also triggers off another train of channel scanning pulses at a predetermined frequency much higher than the sampling frequency. The scanning pulses perform two functions. They index a multiplexer cyclically through a set number (in this case four) of analogue channels and also each of them interrupts the digital computer, causing it to execute the service routine which reads the ADC data. The analogue to digital conversion is started automatically after each indexing of the multiplexer. The timing is so arranged that, by the time the computer routine reads the data, the



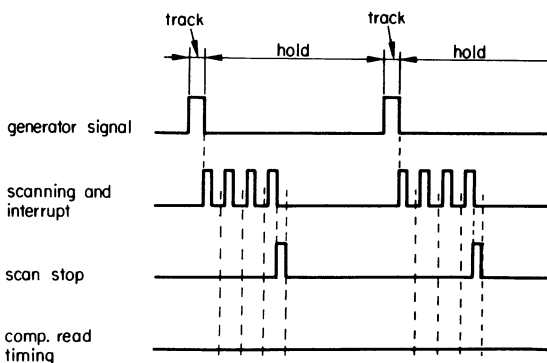
(a) stage 1



(b) stage 2



(c) stage 3



(d) sampling sequence

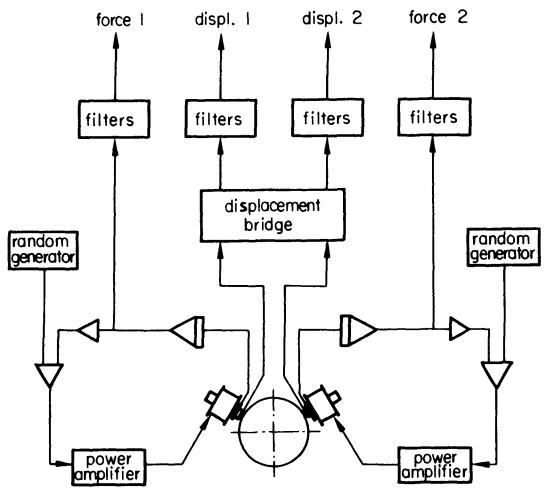
Figure 1. Stages of development of the data acquisition system.

ADC had sufficient time to convert the output voltage of the S/H amplifier, which is then connected through the multiplexer to the ADC. After receiving four scanning pulses, the multiplexer is returned to the first analogue channel and a special logic circuit suppresses further generation of scanning pulses until the sequence is restarted by the arrival of the next sampling pulse. It is seen that the analogue channels are sampled sequentially, but as a result of the S/H action, the values read in any scanning cycle are those which existed during the sampling pulse which triggered the scanning pulses.

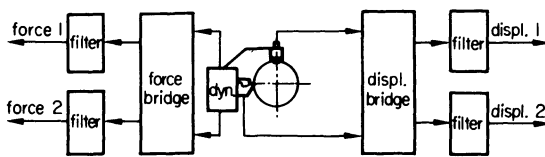
The condition which must be satisfied to allow this system to work is that the time of scanning, converting and storing all the channels must lie within the period during which the amplifiers are in the 'hold' mode, as shown in figure 1(d). This condition limits the sampling frequency as well as the numbers of channels. In the existing installations, the duration of the sampling and scanning pulses is 20 μ s and 10 μ s, respectively, and the repetition period of the latter is 75 μ s. In the present application, the sampling rate was 1600/s at which the maximum possible number of analogue channels is $(10^6/1600 - 20)/75 = 8$, which is twice the number required.

MEASURING EQUIPMENT

A mobile installation was developed which can be used, on the machine site, for measuring force and displacement components during either simulation tests (see figure 2(a) or actual cutting tests (see figure 2(b)). During simulation tests, the outputs of two random generators are fed to two power amplifiers which energize the driving coils of a two directional, electro-magnetic exciter³ mounted on the tool post.

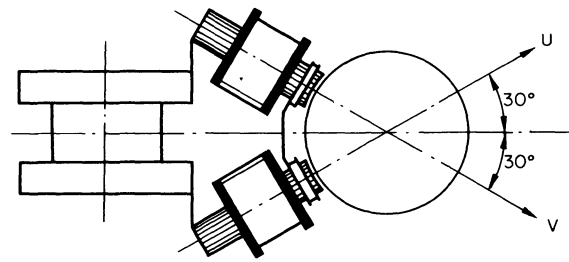


(a) case of simulation

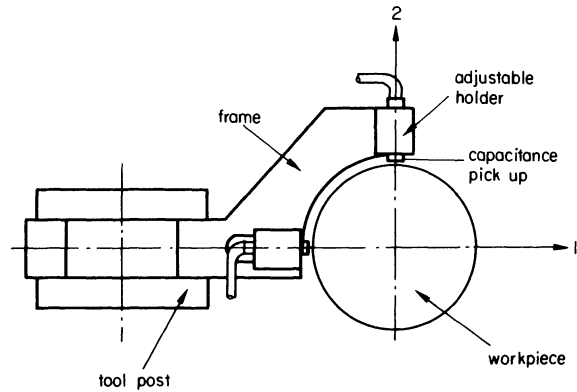


(b) case of cutting

Figure 2. Force and displacement measuring installation.



(a) case of simulation



(b) case of cutting

Figure 3. Directions of measurements.

The outputs of its search coils are integrated and used as a measure of force components. A fraction of each integrated signal is fed back negatively to the appropriate power amplifier to equalize the frequency characteristics of the exciter force over a 0-500 Hz range. Two capacitance pick up units, in conjunction with a two-channel bridge, are used to measure the relative displacement between the exciter and a cylindrical workpiece in the direction of excitation. The force and vibration signals are passed through identical band-pass filters to remove both low frequency components due to workpiece eccentricity and the unnecessary high frequency components lying outside the machine bandwidth.

On the other hand, during cutting tests, two cutting force components are measured by a strain gauge dynamometer, in conjunction with a twin bridge. Two capacitance pick-up units, mounted on a special holder shown in figure 3(b), are also used for measuring the relative displacement between the tool post and the workpiece in the horizontal and vertical directions. The signals are filtered in the same way as in case of the simulation experiment.

EXPERIMENTAL RESULTS

(a) Simulation experiments

Simultaneous, two direction random excitation experiments were carried out on a lathe using an electro-magnetic exciter. The directions of excitation are shown in figure 3(a). In these experiments, a 3.3 in diameter steel workpiece was used and the spindle speed was 280 rev/min.

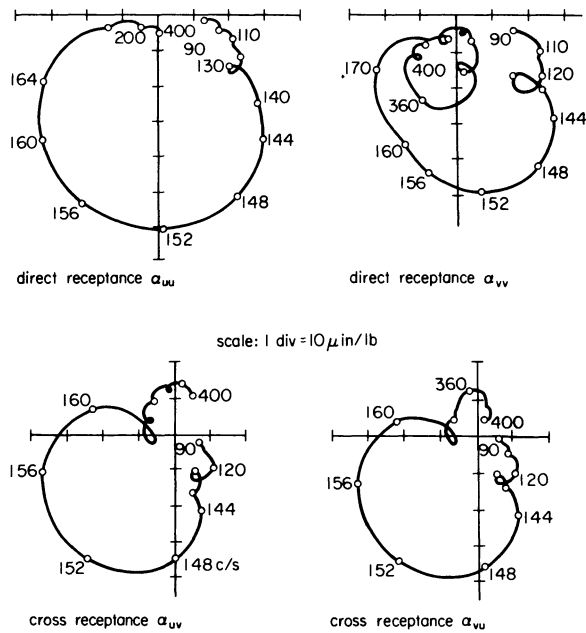


Figure 4. Receptances from simultaneous random excitation test.

The sampling rate was 1000/s and the numbers of samples collected per channel was 1600. Correlation functions were calculated and time domain analysis, using the deconvolution method, was carried out to obtain four impulse response functions up to 80 delay points. Upon Fourier transforming these responses, four receptances were obtained. The computer results were directly plotted on an x-y plotter, as shown in figure 4.

(b) Cutting experiments

A mild steel workpiece, 3.5 in diameter, 28 in long, was used in these experiments. The turning operation was carried out with a 3/8 in square, carbide tool tip, mounted on a two directional strain gauge dynamometer clamped in the tool post. The cutting conditions were chosen as: cutting speed = 200 ft/min,

feed = 0.008 in/rev and depth of cut = 0.030 in. The radial (horizontal) and tangential (vertical) force components as well as the relative displacements between the tool post and the workpiece in the horizontal and vertical directions were measured during cutting.

The forces and displacement signals were filtered, sampled and processed in the same way as in case of excitation experiments, but with a sampling rate of 1600 samples/s. The results are shown in figure 5.

DISCUSSION AND CONCLUSIONS

It is pertinent to point out that the aim of the simulation experiments was (1) to test the data acquisition system; (2) to test the two-input, two-output system analysis program before using it for actual cutting tests and (3) to provide some means of comparison with the receptances obtained from cutting test experiments.

The accuracy of machine tool receptances identification is customarily judged by the degree of similarity between the two cross receptances. This criterion is based on the Maxwell reciprocity principle which is valid for linear mechanical systems and when the lines of application of the force and of the measurement of displacement defining one cross receptance coincide with those defining the other. In simulation tests, the geometry of excitation and of vibration measurement complies with this requirement. Subject to the condition of linearity being satisfied, one should expect the cross receptances to be equal. In cutting tests, however, the line of application of the vertical force (tangential to the workpiece surface) is not the same as that of the measurement of the vertical displacement (vertical line passing through the workpiece centre) which affects the vertical receptances. In view of this, one should not expect equality of the cross receptances under cutting conditions. The inequality of the cross receptances should not, however, be regarded as symptoms of error in either the measurement or in the results.

Figure 4 shows the receptances of the machine structure as determined from simultaneous excitation in two directions. It is obvious that the two cross receptances, which are independently calculated, are very similar, which satisfies the first two objectives of the test, i.e., that the data acquisition system and the computer program were working satisfactorily.

As expected, in the results obtained from the cutting experiments shown in figure 5, the degree of similarity between the two cross receptances is less than in the case of simulated excitation. It is also noticed that these cross receptances are very small in comparison with the direct receptances. The method of identification is sensitive to the contamination of signals by noise and, in the present application, the smallness of the cross receptances must have caused the effective signal to noise ratio, involved in their determination, to be unfavourable. However, the cross receptances have a recognizable pattern and the most pessimistic view one can take is that the signal to noise ratio, involved in their determination, was on the discrimination threshold of the method. The signal to noise ratio, effective in the determination of the direct receptances, is better by a factor equal to

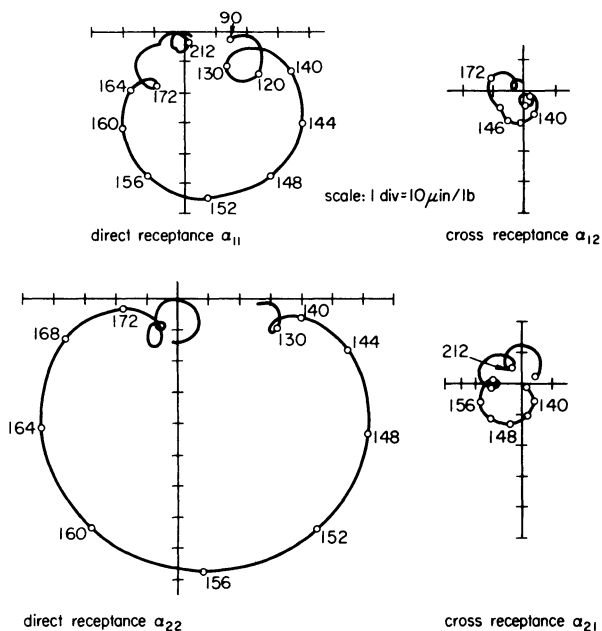


Figure 5. Receptances from cutting test.

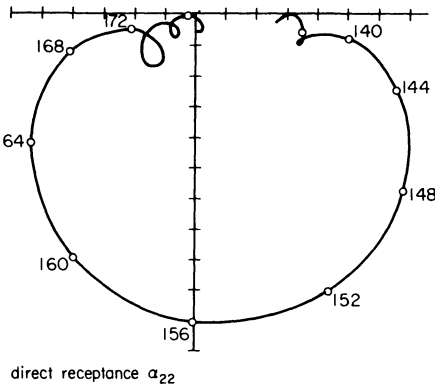
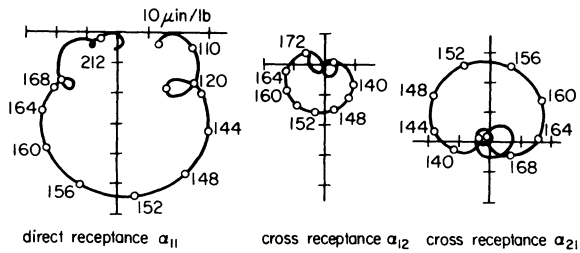
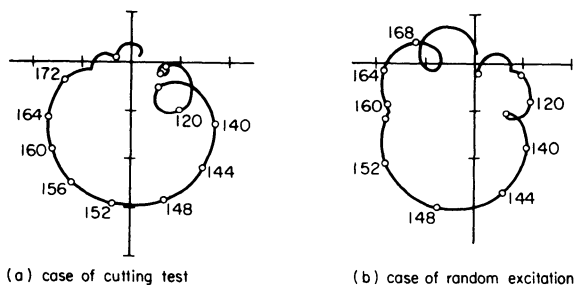


Figure 6. Sample of repetition tests.

the ratio of direct to cross receptance magnitudes which is large. This was confirmed experimentally by repeating the cutting experiments and comparing the results. The results of another cutting test are shown in figure 6. Comparing these results with those shown in figure 5, it is obvious that the corresponding direct receptances are very similar and that the maximum discrepancy appears in the cross receptances. The conclusion is therefore reached that the worst inaccuracy of a direct receptance point is no larger than a representative dimension of the cross receptance.

Before comparing the results of the cutting tests with those of the excitation tests, it is relevant to mention that the reference directions of both cases are different. Therefore, a receptance along a common direction has to be calculated for both cases and taken as a basis of comparison. Since one of the



(a) case of cutting test (b) case of random excitation

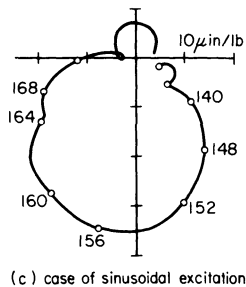
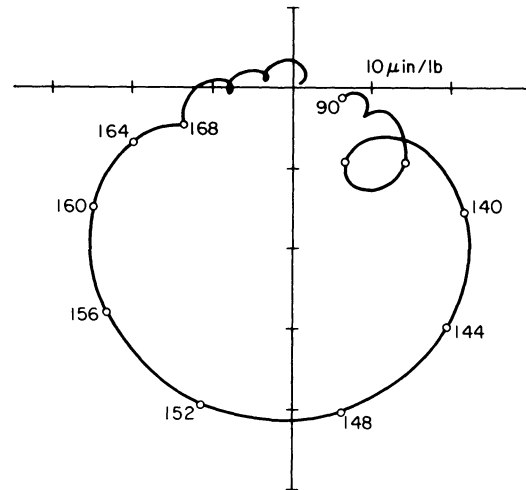
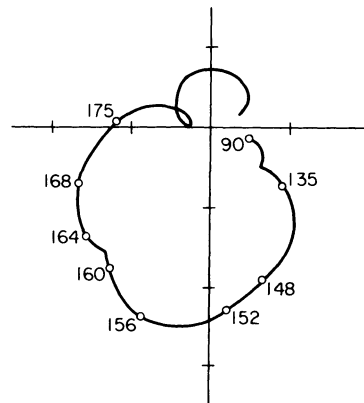


Figure 7. Operative receptances.

most important receptances of a machine tool structure is its operative receptance between the cutting force direction and the normal to the workpiece surface, which is used in stability predictions, it will, henceforth, be taken as the basis of comparison. Fortunately, in the calculation of this operative receptance, the effect of moving the line of measurement of the vertical displacement in the case of the cutting experiments does not introduce any errors since the vertical receptances are not involved. The calculated operative receptances (the direction of the cutting force is assumed to be 66° with the horizontal) are shown in figure 7. For the sake of



(a) case of cutting test



(b) case of sinusoidal excitation with feed motion

Figure 8. Mean operative receptances.

completeness, the operative receptance (obtained from sinusoidal excitation tests under similar conditions as for the random excitation) is also included in figure 7.

It is obvious from figure 7 that the three operative receptances obtained by the three methods are different both in magnitude and in phase. It may be argued, however, that a comparison between these three receptances is not quite valid for two reasons. Firstly, as shown by Hanna and Kwiatkowski⁴, the sinusoidal method receptance represents a mean or an average value, whereas the random method (or cutting test) receptance represents an almost instantaneous value and, secondly, the excitation tests were

carried out under stationary carriage conditions, whereas the carriage was moving during the cutting tests.

To clarify this point, cutting tests were repeated several times (16 times) and the operative receptance was calculated from the results of each test. The mean operative receptance was then computed. On the other hand, sinusoidal excitation tests were carried out under moving carriage conditions, as described by Hanna and Kwiatkowski¹ and four receptances were obtained from which the operative receptance was calculated. The results of these tests are shown in Figure 8.

It is seen from figure 8 that the two receptances are still different. The difference may, therefore, be attributed to the conditions on the sliding surfaces prevailing during cutting, which cannot be achieved by the existing simulation techniques. As a consequence, chatter predictions based on receptances obtained by the two methods will be different. It is concluded that a more accurate prediction will be accomplished by using receptances obtained from cutting tests.

ACKNOWLEDGMENTS

This investigation was carried out under a contract from the S.R.C. The authors wish to express their gratitude to Professor S. A. Tobias for his support and encouragement throughout the work.

REFERENCES

1. N. H. Hanna and A. W. Kwiatkowski. Effect of feed motion during dynamic tests on the chatter prediction of a lathe. *Proc. of the 13th Int. MTDR Conference*, p. 39.
2. N. H. Hanna and A. W. Kwiatkowski. Identification of machine tool receptances by random force excitation. *Int. J. Mach. Tool Des. Res.* (1971) 11.
3. A. W. Kwiatkowski. A magnetic vibrator for determination of machine tool characteristics. *Proc. of the 9th Int. MTDR Conference*.
4. N. H. Hanna and A. W. Kwiatkowski. Some problems and effects associated with measurement of machine tool receptances in the presence of feed motion. *ASME Design Engineering Technical Conference*, Cincinnati, Ohio, September 1973.

INFLUENCE OF LATHE TOOL WEAR ON THE VIBRATIONS SUSTAINED IN CUTTING

by

P. MARTIN*, B. MUTEL* and J.-P. DRAPIER*

SUMMARY

The adaptive control of machine tools (for example, optimum choice of cutting conditions at each instant of time) requires a criterion for the state of tool wear or its derivative.

In order to establish an easily measurable magnitude which is characteristic of the state of wear, we have investigated a relation between the vertical vibrations of a lathe tool and its flank wear.

The investigation shows that the vertical vibrations of the tool in the course of stable machining are almost sinusoidal, with frequency equal to the natural frequency of the tool. For XC 38 steel and for cutting speeds of 75 to 150 m/min, the power of the acceleration signal obtained by spectral analysis is a linear function of the cutting speed and of the tool wear, and varies in the ratio of 1 : 10 between the new tool and the worn tool.

These experimental results are justified theoretically with the aid of a closed-loop model composed of a second-order system with two non-linear elements. The first non-linear element characterizes the cutting process and describes the variation of the main cutting force as a function of the effective cutting speed. The second characterizes the limiting of the displacement of the tool.

This model permits an explanation of the phenomenon of sustained vibrations in the neighbourhood of the natural frequency of the tool and, in the range of values investigated, results in a variation of signal power which conforms well with the experimental results.

NOTATION

a (mm)	depth of cut
s (mm/rev)	feed rate
V_c (m/min)	cutting speed
u (mm)	flank wear
k_i, h	constants

The relation between the vibrations of a lathe tool and the wear has been pointed out by several authors^{4,5,6}. Moreover, the measurement of the vibrations is simple and the flank material classic. Thus we decided to investigate this relation. The horizontal and radial vibrations have often been studied in order to reduce them, so we examined the behaviour of the vertical vibrations (in the direction of the main cutting force) of the lathe tool during normal cutting as a function of the cutting and wear parameters, with the aid of their amplitude and power spectra.

1 INTRODUCTION

Among the systems of automation of production processes, numerical control of machine tools has been particularly developed in recent years, and the need has arisen for a more sophisticated automation by adaptive control, optimizing, for example, cutting in the course of machining with respect to tool wear, which must be measured continuously.

The wear sensor chosen must be usable in the industrial environment and must deliver a signal which varies significantly with wear. The continuous wear sensors currently proposed^{1,2,3} are not used industrially. It is also desirable that preliminary setting runs after a change of tool or of material be reduced as much as possible.

2 EXPERIMENTAL RESULTS

2.1 Choice of variables

We have limited our experiments to a single type of tool and to two machined materials. The principal variables are the cutting speed V_c , the feed rate s and the state of the tool wear determined by the height u of the flank wear, which is more characteristic, more visible and more easily measurable than the crater wear. The depth of cut has been kept constant.

We have accomplished a series of experiments with several lengths of tool overhang. The position of the tool with respect to one end of the workpiece undergoing machining was also recorded.

* CUCN Laboratoire d'Electricité et d'Automatique, Université de Nancy I, France

2.2 Equipment used

(1) Lathe

The runs were made on a Cazeveuve G 36 16 horsepower regulator engine lathe. This machine has a very rigid frame and a power at the spindle much greater than the cutting power required. Although relatively old, it is satisfactory for our experiments.

(2) Tool

Throw-away negative-rake carbide square inserts, gradation P40, were used. These were mechanically fastened to a lead angle turning tool.

(3) Workpiece material

The tests were accomplished by the machining of bars of XC 38 steel with 187 Brinell Hardness Number. The bars were 900 mm long and 120 mm in diameter. A second series of tests was made with 35CD4 steel of 160 BHN.

(4) Measuring device

A Bruel and Kjaer piezo-electric accelerometer was fastened by a magnet or pasted to the upper surface of the tool. Fastening by adhesive is clearly preferable provided that a product is used which is resistant to the elevated temperatures of the tool. A magnet permits easy regulation of the position of the sensor but risks being accidentally displaced.

The accelerometer gave an output signal of 22 picoCoulomb/g which an amplifier converted into voltage. (g is the acceleration of the mass.) The signal obtained was displayed on an oscilloscope and recorded on magnetic tape.

The natural frequency of 49 kHz of the vibration transducer and the natural frequency of the order of 20 kHz of the fastening system were much greater than the observed frequencies (2.5 kHz) and did not perturb them.

The accelerometer which was mounted on the tool as close as possible to the cutting edge and thus to the source of vibrations, did not perceptibly modify the vibratory behaviour of the system, and particularly its natural frequency.

(5) Flank material

The spectral analysis of the acceleration signal was accomplished on samples of several seconds and gave results between 0 and 5 kHz with the aid of filters with a pass band of 10 Hz, using a Federal Scientific analyser.

Furthermore we measured the power spectral density (PSD), which gives the part of the average power of the signal $x(t)$ contained in a frequency band B infinitesimally narrow, centred at frequency f , the average value being determined over an infinitely long time. By integration it gives the total power.

For a purely sinusoidal signal $A \sin \Omega t$, the spectral density W and the power P are given by

$$W(\omega) = \frac{A^2}{2} [\delta(\omega + \Omega)]$$

$$P = \frac{A^2}{2}$$

2.3 Operative mode

We carried out 100 recordings of approximately 30 seconds, corresponding to a turning operation without cutting fluid, with fixed cutting parameters and with wear assumed constant. The inserts were either new or with known wear. The change of cutting edge of the insert was rapid and did not cause displacement of the tool, which therefore kept the same position in the course of a series of measurements. The value of each of the cutting parameters varied around the values advised by the tool manufacturer: $V_c = 125$ m/min, $s = 0.4$ mm/rev.

Thus s varies from 0.2 to 0.6 mm/rev, and with $76 < V_c < 152$ m/min, we were outside the zone of formation of a significant built-up edge.

For every couple (s, V_c) the flank wear of the tools used was in the range 0 to 0.8 mm (although in practice the surface state and the tolerance on the piece do not permit this maximum value).

A test of continuous cutting was also undertaken in order to show agreement of results with the preceding intermittent tests.

The power and amplitude spectra obtained exhibit, for the most part, a very pronounced peak between 2000 and 3000 Hz.

We therefore studied the frequency, the amplitude and the spectral density corresponding to this peak.

2.4 Frequencies

(1) Natural frequency of the tool

The acceleration signal of the tool obtained by striking the free end of the tool was approximately 2500 Hz. Its double integration gives as the damping factor $\xi = 0.1$.

The tool, likened to a beam rigidly fixed at one end and free at the other, had a natural frequency of 7900 Hz, a value much greater than the observed frequencies. The oscillating system was not composed of the tool alone but of the toolholder and the tool.

Furthermore, the specific flexion of the bar was 2×10^{-8} m/N.

(2) Frequency of forced vibrations in cutting

With constant cutting parameters and wear the fundamental frequency, obtained with the aid of amplitude spectra, increased as the tool overhang decreased, and varied between 2000 and 2800 Hz. As the spectra present a very pronounced peak at this frequency, the observed signal is relatively sinusoidal and its period is, moreover, very much in the neighbourhood of the natural period of the free system. These results agree with those of Pankine⁶, Kronenberg⁵ and Weller⁴.

Pankine⁶ points out the existence of two vibration zones, one at the high cutting speeds corresponding to vibrations of the tool itself and in which our experiments lie, and the other observed at the low speeds which coincide with the vibration domain of the piece machined.

Kronenberg⁵ specifies that the natural frequency of the tool, approximately 2200 Hz, is that of the vibrations recorded at different cutting speeds. However when the machine lacks rigidity, the vibrations of the other parts are superposed on the vibration of the tool.

Weller⁴ indicates also that the fundamental frequency of the vibrations is the resonant frequency of the system excited by the friction of the cutting edge on the workpiece.

(a) *Influence of the cutting parameters.* As in our experiments the frequency observed varied from 300 Hz to the maximum, it is impossible to separate from the experimental errors the variations due to the cutting parameters or to the wear.

For the very low frequencies, the amplitude spectra exhibit weak components of frequencies between 70 and 120 Hz, proportional to the speed of rotation. The spindle rotation is certainly the cause of this.

(b) *Influence of the workpiece.* The frequency of the vibrations of the workpiece, of the order of 400 Hz in our experiments, can affect that of the tool. But this frequency does not appear in the observed spectra, the cutting speeds used being without doubt sufficiently high to prohibit the vibrations of the workpiece. Also there does not appear to be a direct relation between the position of the tool with respect to one end of the workpiece and the vibrations.

2.5 Amplitude

The orders of magnitude of the actual amplitude were determined on the photographs of the signals obtained with the oscilloscope. The amplitude of the acceleration varies between 40 and 200 g according to the cutting conditions; the displacement of the sensor, assumed to be purely sinusoidal, is between 1.6 and 8 μm and corresponds to a displacement *A* of the tip of the tool of 3 to 13 μm.

If ω is the period of the vertical vibration of the tool, the tip speed *A*ω varies between 2 and 12 m/min, which is 3 to 12 per cent of the corresponding cutting speed *V_c*.

As the variation, as a function of the cutting parameters, of the amplitude of the fundamental observed in the amplitude spectra is smaller than that of the corresponding power, we concerned ourselves with the latter (figures 1 and 2).

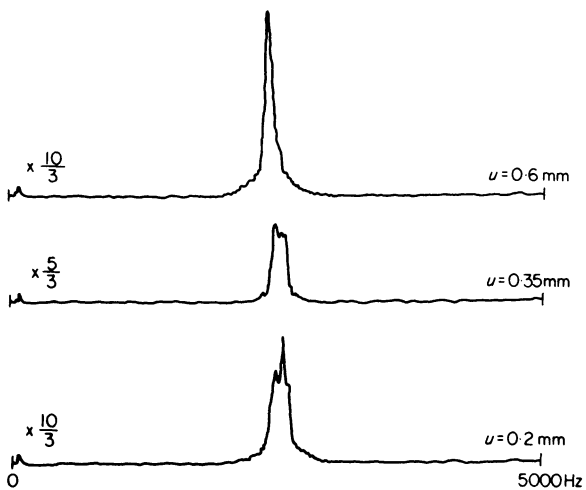


Figure 1. Power spectrum for XC 38 steel with *a* = 2 mm, *V_c* = 113 m/min, *s* = 0.2 mm/rev.

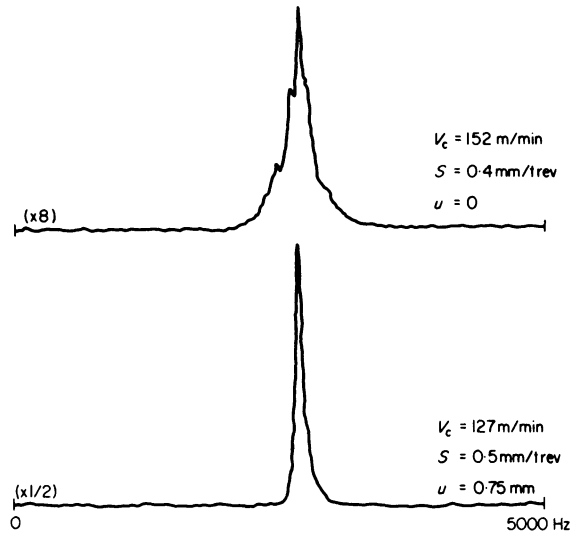


Figure 2. Power spectrum for XC 38 steel with *a* = 2 mm.

2.6 Power

The power *P* of the acceleration signal is obtained by the integration of the power spectrum, itself determined by the area under the curve. The value obtained is proportional to the real power, the coefficient of proportionality being imposed by the gain of the measurement chain. The most significant imprecisions are due essentially to the inertia of the recording table for the spectrum, the influence of which was limited by maintaining the height of the peaks at the same order of magnitude. However, a relative precision of at most 15 per cent permits us to determine the variations of power with the cutting parameters and the wear. For the two materials studied, we have plotted curves of *P*(*sV_c*) for constant wear (figure 3), and *P*(*u*) for constant machining conditions (figure 4).

In the range of values studied, *P* increases linearly with *u* and *sV_c* from 1 to 5 as the flank wear varies from 0 to 0.75 mm, and from 1 to 2 as the product *sV_c* changes in the same ratio (figures 3 and 4).

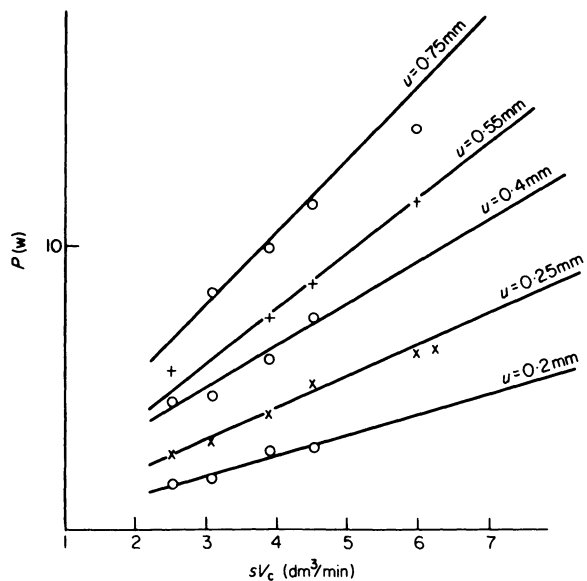


Figure 3. Variation of the experimental acceleration signal power *P* with *sV_c*.

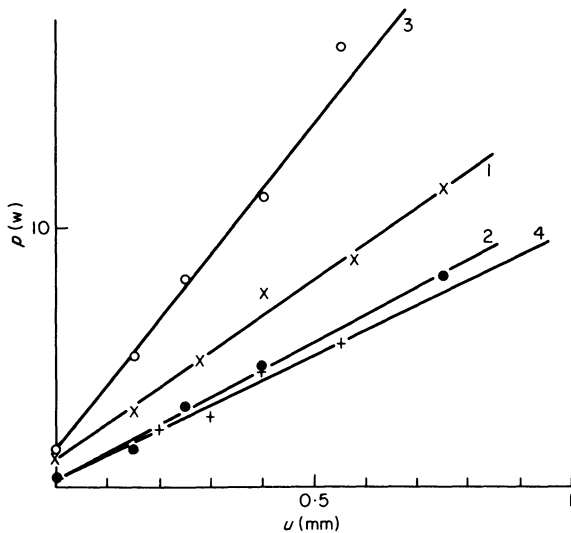


Figure 4. Variation of P with u : (1) $V_c = 114$ m/min, $s = 0.4$ mm/rev, (2) $V_c = 76$ m/min, $s = 0.4$ mm/rev, (3) $V_c = 127$ m/min, $s = 0.6$ mm/rev, (4) $V_c = 127$ m/min, $s = 0.2$ mm/rev.

Moreover, a good agreement of results is obtained between the continuous cutting test and the short tests of 30 seconds.

3 DEFINITION OF A MODEL OF THE CUTTING SYSTEM

3.1 Introduction

From Pankine⁶ and Weller⁴ the vibrations of a lathe tool in stable machining are caused by the friction of the flank face of the tool against the workpiece.

According to Kronenberg⁵ the vibrations of a carbide tool can be sustained by the true rake angle with the displacement of the tool, the main cause being, however, the vibration of the relative speed of the tool with respect to the workpiece.

The excitation of the tool by 'stick-slip' cannot be considered because the relative tool-workpiece speed is never zero, as experimental results prove.

Shaw⁷ attributes the sustaining of vibrations at low cutting speed to the formation and disappearance of built-up edge; this ceases at more than 1000 Hz.

Bjorke⁸ has proposed and simulated on a computer a closed-loop system composed of a non-linear cutting model and a tool likened to a damped element of the second order. He found that the speed of the vibrations must be greater than the cutting speed and that the amplitude of the vibrations varies only slightly with the cutting conditions. In the light of this author, we will admit that the vibrations are sustained in a non-linear closed-loop system, and are well defined in the absence of chatter.

3.2 Forces put into play

In the planar model of orthogonal cutting, the forces applied to the chip and the tool are:

- (1) the tool-chip friction Q on the tool face,
- (2) the shearing force F acting in the shear plane,
- (3) the ploughing R at the level of the tool nose,
- (4) the tool flank-workpiece friction L ,
- (5) the compressive force C on the shear plane.

Only the first two terms appear in the classical theory of Merchant. The force R has been brought to light by Albrecht⁹ and permits an explanation of the behaviour of the curve of the main cutting force F as a function of the cutting speed by the formation of built-up edge (figure 5).

The main cutting force F obtained by projection of the forces in the direction of the cutting velocity is given by the relation

$$F = k_1 as + k_2 R + k_3 au$$

When the chip cross-section varies, the curve in figure 5 is displaced as a whole in the direction of the F -axis. However, as the characteristic $F(V_c)$ conveys only the average phenomena, to analyse the process in order to study the vibrations it is necessary to introduce the dynamic characteristic of the cutting force.

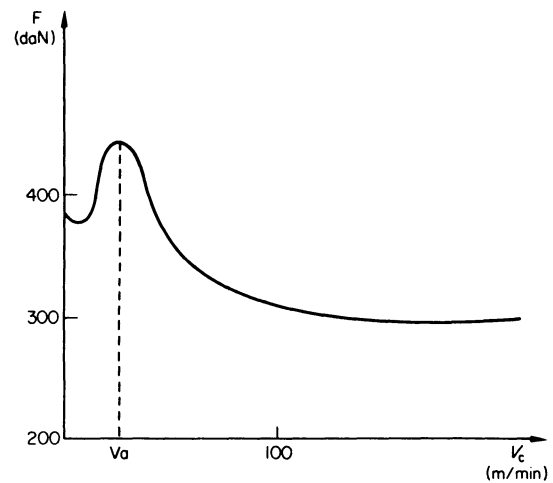


Figure 5. Relation between main cutting force and cutting speed.

3.3 Dynamic characteristic of the cutting force

(1) General form

Designating by y the relative position of the tool tip with respect to its position at rest, considered positive in the direction of the cutting velocity, the actual speed of the tool is

$$V = \bar{V}_c - \frac{dy}{dt}$$

where \bar{V}_c is the average value of the cutting speed.

In the range of values considered ($V_c > V_a$), the dynamic curve $F(V)$ by analogy with $F(V_c)$ can be approximated by two lines, one with constant negative slope $-n$, the other horizontal with ordinate F_0 . We will assume that the slope n of the dynamic curve $F(V)$ is of the same order of magnitude as that of the average characteristic $F(V_c)$.

Moreover, the main cutting force is limited by the obligatory shearing of the chip. By admitting that in the dynamic characteristic $F(V_c)$ is representative of the average force, the shearing force $F(V)$ of the chip is greater than $F(V_c)$. In the course of the shearing the main cutting force remains constant, and the dynamic characteristic $F(V_c)$ of the cutting process will be described in figure 6.

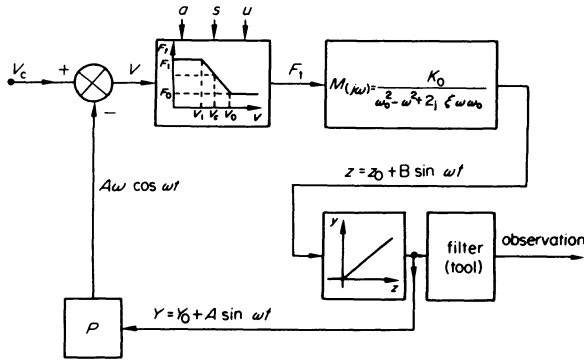


Figure 6. Equivalent diagram of the system.

(2) Quantitative description

The variation of the cutting force is slight *vis-à-vis* the average force and F_0 (proportional to the chip section), which is only considered in general. Thus $F_1 - F_0$ is very small compared with F_0 , which further justifies the hypothesis regarding $F(V_c)$. The difference $F_1 - F(V_c)$ can be attributed to the variation in the dynamic region of the compressive stress which intervenes linearly in shearing limit stress (Caquot Bridgman's formula⁹). As one can reasonably admit that this compression in the dynamic regime becomes zero when V_c does, let us write

$$F_1 - F(V_c) = kV_c^{-1}$$

with $k < n$. This hypothesis joins the ideas of Pomey¹⁰, who asserts that the compressive stress is an increasing function of V_c .

Besides, as the tool nose enlarges with the wear u we will suppose that $F_1 - F(V_c)$ contains a term proportional to u , whence

$$F_1 - F(V_c) = kV_c + hu$$

Finally, by analogy with the average characteristic, we will admit that the depth of cut intervenes only in the case of F_0 to the exclusion of the difference $F_1 - F(V_c)$.

3.4 Action of the cutting force on the tool

We will represent the action of the cutting force on the tool by the transfer function

$$M(j\omega) = \frac{K_0}{\omega_0^2 - \omega^2 + 2j\xi\omega\omega_0}$$

where ξ is the damping factor, ω_0 is the undamped natural frequency, and K_0/ω_0^2 is the static gain factor.

Under the effect of a sinusoidal force $F_T(j\omega)$, the displacement of the tool tip, which for the moment we will name z , is given by

$$z = M(j\omega) F_T(j\omega)$$

3.5 Limiting of the displacement

The tool tip cannot go beyond the horizontal plane containing the axis of the workpiece, and the preceding displacement z is susceptible to being limited. The real displacement y of the tool tip is related to z by the characteristic of $y(z)$ (figure 6).

3.6 Equivalent diagram of the system

The displacement y which appears in the expression

$$V = V_c - \frac{dy}{dt}$$

of the effective velocity of the tool tip with respect to the workpiece causes the looping of the system following the model of figure 6.

Moreover, between the tip and the sensor mounted on the tool there exists a system of the second order $M(j\omega)$ having essentially the role to filter the first harmonic of y . In order to determine y , we will use the method of harmonic linearization which applies to this system.

4 STUDY OF THE LINEAR MODEL

4.1 Principle

In a system like ours which includes a non-linear process followed by a linear part with transfer function $M(j\omega)$, it is generally supposed that the second element uniquely passes the fundamental and filters the higher harmonics, and the method of harmonic linearization is applied. We will use this method in our case, where the vibrations possess a very significant first harmonic.

The required output signal y , signifying the position of the tool, is reduced to

$$y = y_0 + A \sin \omega t$$

and the velocity to

$$V = V_c - A\omega \cos \omega t$$

4.2 Results

The development of the method permits us to determine the conditions of oscillation, their stability, the amplitude A and the frequency ω as functions of the parameters of the system.

A is defined by the intersection of the curves $H(A)$ and $D(A)$ in figure 7, with

$$H(A) = \frac{2\pi\omega_0\xi}{nK_0} \left(2 - \frac{4z_0}{\pi A} \right) \quad \text{(portion 2)}$$

$$D(A) = \pi - \arccos x + x\sqrt{1-x^2} \quad \text{(portion 1)}$$

$$D(A) = \arccos y - y\sqrt{1-y^2} - \arccos x + x\sqrt{1-x^2} \quad \text{(portions 2 and 4)}$$

$$D(A) = \arccos y - y\sqrt{1-y^2} \quad \text{(portion 3)}$$

$$\text{with } x = \frac{V_c - V_1}{A\omega} \quad \text{and} \quad y = \frac{V_c - V_0}{A\omega}$$

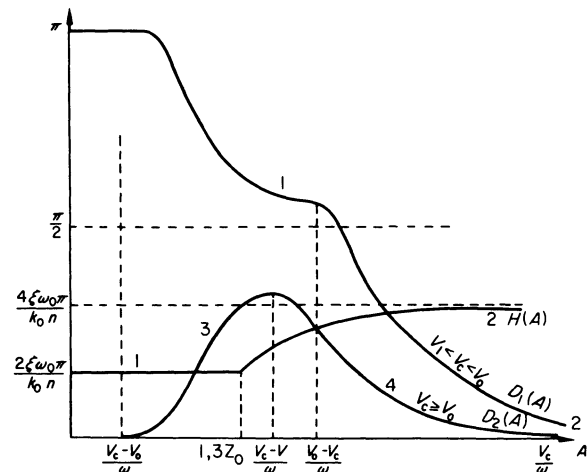


Figure 7. $H(A)$ and $D(A)$ curves.

TABLE 1 Variation of $\pi\omega_0 A/K_0$ from intersection of $D(A)$ and $H(A)$ curves

	$V_1 < V_c < V_0$ and $V_c - V_1 < A\omega < V_0 - V_c$	$V_1 < V_c < V_0$	$V_0 < V_c$
Without limiting of the displacement	$\frac{4(kV_c + hu)}{4\xi\omega_0 - nK_0}$	$\frac{(k-n)V_c + nV_0 + hu}{\xi\omega_0}$	$\frac{kV_c + hu}{\xi\omega_0}$
With limiting	$\frac{4(kV_c + hu + 4k_1as\xi)}{8\xi\omega_0 - 4K_0}$ (1)	$\frac{(k-n)V_c + hu + 4k_1\xi as}{2\xi\omega_0}$	$\frac{kV_c + hu + 4k_1\xi as}{2\xi\omega_0}$

Based on the stability criterion for closed-loop non-linear systems, the oscillations obtained will be stable for every point of intersection of $H(A)$ with the decreasing part of $D(A)$.

Furthermore, the amplitude of the oscillations when they exist is always a linear increasing function of u and an increasing or decreasing function of V_c . Moreover, if $A > 1.3z_0$ the amplitude is also an increasing function of the chip cross-section as , as shown in Table 1.

Frequency ω is equal to the undamped natural frequency ω_0 .

The power, which at the level of the system $M(j\omega)$ corresponds to the damping of the vibrations in z , is

$$\frac{B^2\omega_0^3}{K_0}$$

We admit that for the actual signal y one has $\xi A^2\omega_0^3/K_0$ in every first approximation, the amplitude A being given by the different preceding expressions, all of which contain V_c (with + or - sign) and some u and as .

Experimentally, measuring the power proportional to the square of the amplitude A of the displacement of the tool, we have made conspicuous the effects of wear and cutting velocity.

In the framework of the preceding theory we find ourselves in the presence of limitation of the displacement and encroachment of the dynamic region on the single horizontal shearing portion of the cutting characteristic ($V_c - V_1 < A\omega < V_c - V_0$).

Numerically, K_0/ω_0^2 is just the static specific flexion already calculated: 2×10^{-8} m/N. Furthermore, we have $\omega_0 = 2500 \times 2\pi$ and $\xi = 0.1$. The average cutting characteristics suggest that we take $n = 400$ N/m s⁻¹, and using tool manufacturers' data we take $k_1 = 1000$ N/mm².

By choosing $k = 150$ N/m s⁻¹ and $h = 1000$ N/mm, using expression (1) in Table 1 we obtain, in the range of cutting parameters considered, the curves of figures 8 and 9 which are almost rectilinear in this range of values and which are comparable to the corresponding experimental curves of figures 3 and 4.

The numerical connection between the experimental and theoretical results is correct because with the values chosen for k and h in expression (1), with $V_c = 120$ m/min and $as = 0.8$ mm², an amplitude of $15 \mu\text{m}$ is obtained, only slightly greater than the values measured directly.

Under these experimental conditions, we obtain for the power $\xi A^2\omega_0^3/K_0$ a value of approximately 15 W, an order of magnitude which appears correct but which is obviously not compatible with a measurement in a global measurement of the power.

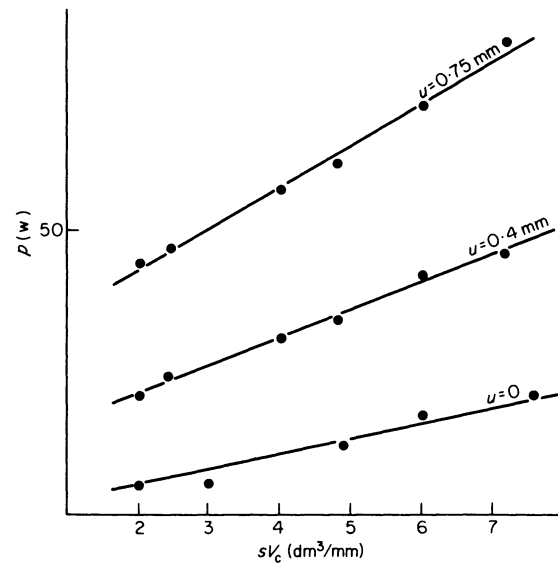


Figure 8. Variation of the theoretical acceleration P_1 with sV_c .

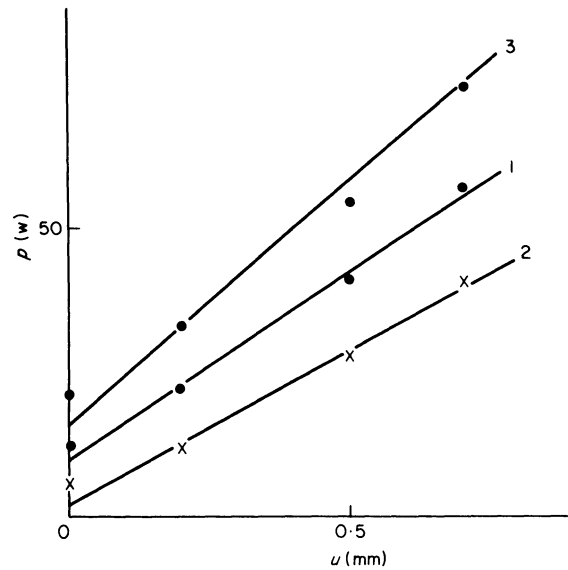


Figure 9. Variation of P_1 with u : (1) $V_c = 120$ mm/min, $s = 0.4$ mm/rev, (2) $V_c = 76$ mm/min, $s = 0.4$ mm/rev, (3) $V_c = 125$ mm/min, $s = 0.6$ mm/rev.

CONCLUSION

The experiment shows that the vertical vibrations of a lathe tool in the course of stable machining are almost sinusoidal, with frequency perceptibly equal to the natural frequency of the tool. In the range of values studied, the power of the acceleration signal determined by spectral analysis is a linear function of the cutting speed and of the tool wear, the signal increasing in the ratio of 1:10 between the new tool and the worn tool.

Similar results were obtained starting from a closed-loop model composed of a second-order system with two non-linear elements. The first non-linear element characterizes the cutting process and describes the variation of the main cutting force as a function of the actual cutting speed by defining a saturation due to the shearing of the chip as well as a descending portion. The second non-linear element characterizes the limiting of the displacement of the tool in the horizontal plane containing the axis of the workpiece.

This model explains the sustaining of vibrations at the natural frequency of the tool and, in the range of values studied in the case of the limiting of the tool displacement, results in a variation of signal power which conforms well with the experimental results.

The relation obtained between vibrations and wear in the case of turning allows the consideration of a wear sensor which is simple, sturdy and responsive with a view to adaptive control. We are thinking of generalizing this relation to other types of machining.

ACKNOWLEDGMENTS

The authors wish to express their thanks to Professor A. Fruhling, Directeur du Laboratoire d'Electricité et d'Automatique for his advice and encouragement in

accomplishment of this work. They gratefully acknowledge the support received from Professor A. Charles, Directeur du Département Génie Mécanique de l'Institut Universitaire de Technologie de Nancy.

REFERENCES

1. F. Guisti. Regulation, en process, des paramètres de coupe, en fonction de la température de coupe. *Meeting of the CIRP*, Nottingham, 1968.
2. J. R. Jaeschke, R. D. Zimmerly and S.-M. Wu. Automatic cutting tool temperature control. *Int. J. Mach. Tool Des. Res.* (1967) 7, 465-75.
3. A. de Filippi and R. Ippolito. Adaptive control in turning : cutting forces and tool wear relationship for P10, P20, P30 Carbides. *Annals of the CIRP* (1969) 17, 377-85.
4. E. J. Weller, H. M. Schrier and B. Weichbrodt, What sound can be expected from a worn tool? *J. Engng. Industry USA* (1969) 91, no. 3, pp. 525-34.
5. M. Kronenberg. *Machining Science and Applications*. Pergamon Press, 1966, pp. 306-14.
6. A. V. Pankine. Usinage des métaux par la coupe. *Ecole Supérieure Moscou* (1965), pp. 114-20.
7. M. C. Shaw and S. R. Sanghani. On the origin of cutting vibrations. *Meeting of the CIRP*, Hague, 1962.
8. O. Bjorke. An analysis of the cutting process dynamics with reference to errors in cutting force measurements. *Meeting of the CIRP*, Paris, 1966.
9. J.-P. Bouillet. La coupe des métaux. *Dunod. tome 1* (1964), pp. 137-51.
10. J. Pomey. Phénomènes de mécanique physique qui régissent la coupe des métaux. *Mécanique matériaux électricité* (1971) no. 256, pp. 8-35.
11. P. Martin. Influence de l'usure d'un outil de tour sur les vibrations entretenues en coupe. *Thèse de Docteur-Ingénieur*, Université de Nancy I, 1973.

INVESTIGATION OF PARAMETRIC INSTABILITY IN MILLING SIMULATED BY SCREW CHASING OPERATION

by

G. DROUBI* and M. M. SADEK*

SUMMARY

The parametric instability in milling, resulting from the equivalent periodic stiffness variation of the machine structure is investigated, using a special machining set-up which eliminates regeneration. A theoretical model has been developed to predict this type of instability for a specially designed workpiece of a single degree of freedom. The principal axis of vibration of this system rotates in relation to the cutting force orientation. The analysis led to a Mathieu's equation incorporating a damping term. Good correlation has been achieved between the predicted results and those experimentally obtained from cutting tests.

NOTATION

c	damping coefficient
dF	component of the incremental cutting force in the direction of vibration
F	resultant cutting force
k_1	chip thickness coefficient per unit width of cut
m	mass of vibratory system
N	rotational speed of workpiece in rev/min
Q	dynamic magnification factor
t	time
w	width of cut
x	instantaneous vibrational motion
\dot{x}	dx/dt
λ	stiffness of vibratory system
ψ	angle between steady state resultant cutting force and the normal to the cut surface
Ω	angular speed of workpiece
ω_0	angular natural frequency of vibratory system.

INTRODUCTION

The susceptibility to chatter of machine tools with stationary cutters, such as lathes, can easily be predicted from the dynamic characteristics of the machine structure, using the method developed by Sweeney and Tobias². However in the case of horizontal milling machines where the cutter rotates, a difficulty arises by virtue of the fact that the direction of the cutting thrust changes as the cutter tooth sweeps across the arc of contact. This complexity has been overcome by assuming a representative position of each cutter tooth, this being that which coincides with the middle of the arc of con-

tact. Sadek and Tobias¹ verified this technique experimentally on three milling machines with the average number of teeth in contact with the workpiece higher than unity ($z_c > 1$). In this case, the periodic variation of the direction coefficients, representing the influence of the tooth position is negligibly small as compared to the mean value corresponding to the mid-position of the arc of contact³.

The experimental investigation³ carried out for a milling process with the number of teeth in contact $z_c = \text{unity}$, showed that the experimental stability chart obtained from chatter cutting tests is in good agreement with that predicted by the theoretical model of Sweeney and Tobias. On the other hand for the case of $z_c = 0.5$, chatter cutting tests performed under the same cutting conditions showed that the experimental stability chart differs considerably from that predicted by the conventional single tooth model, thus confirming the invalidity of this model for cutting, when the number of teeth in contact with the workpiece is less than unity, as previously suggested¹.

This discrepancy is possibly caused by the periodic variation of both the direction coefficients and the dynamic cutting coefficients along the arc of contact. As far as the variation of the dynamic coefficients is concerned, it has been shown^{1,4} that this has an insignificant effect on the stability boundary for cutting conditions attainable practically. It is therefore adequate to use a set of dynamic cutting coefficients which correspond to the uncut chip thickness in the middle of the arc of contact.

In order to predict the stability criterion for cases

* Mechanical Engineering Department, University of Birmingham

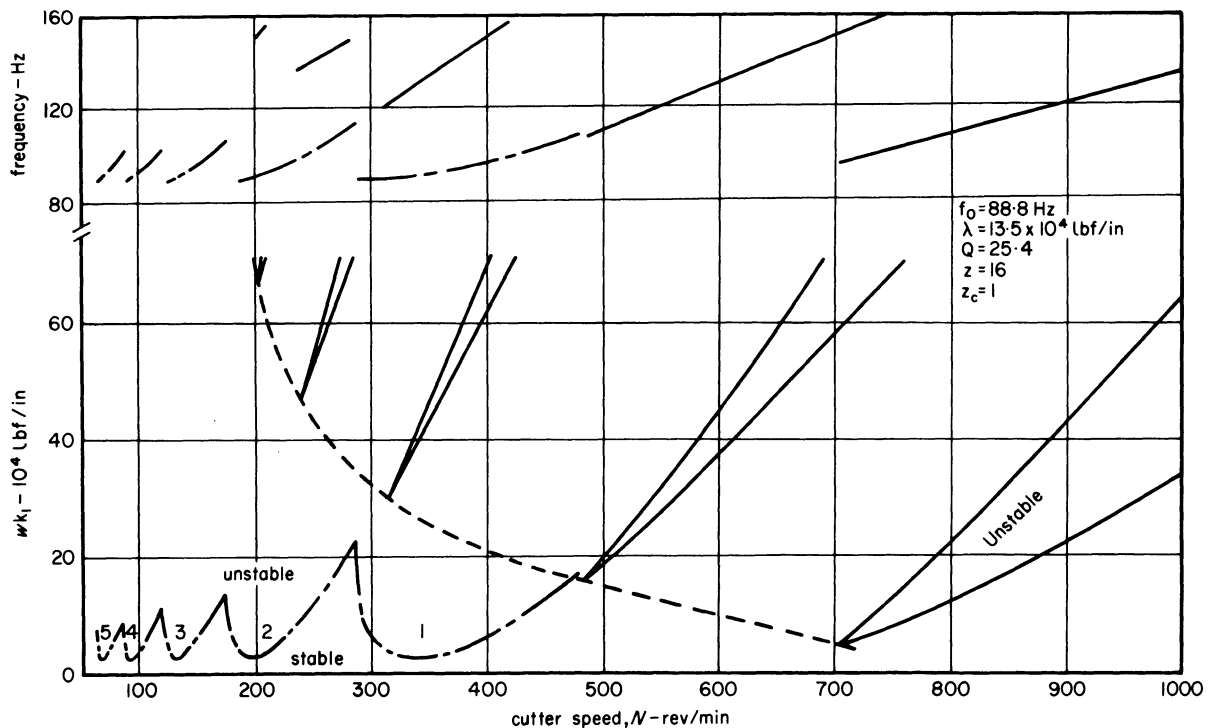


Figure 1. Overall stability chart of the milling process.

of $x_c \leq 1$ a theory has been developed based on the solution of a difference differential equation of motion³. The periodic direction coefficients inserted in this equation, describing the dynamics of the milling process, have the equivalent effect of a corresponding variation of the stiffness and damping of the machine tool structure. This analysis rendered two stability criteria, one being regenerative and the other being the parametric instability resulting from the periodic variation of the direction coefficients. These two criteria are shown in the stability chart illustrated in figure 1, predicted by the above mentioned equation for the milling process previously investigated, when the number of teeth in contact equals unity³. In this chart, the dynamic chip thickness coefficient is plotted versus the cutter speed showing two sets of unstable lobes, the chain dotted lobe corresponding to the conventional regenerative instability and that represented in continuous curve is that corresponding to the parametric instability. The latter results from the periodic variation of the direction coefficients of the process. The instability arising from the conventional regenerative chatter occurs at considerably low values of the dynamic chip thickness coefficients as compared with those corresponding to the parametric instability. This explains the reasons for the unattainability of the parametric instability boundary from chatter cutting tests, since the milling process will become unstable due to regenerative chatter before the parametric instability comes into effect. This shows that the instability in this speed range will be predominantly due to regenerative chatter and justifies the use of the conventional chatter theory for the range of low cutting speeds. For higher speeds the effect of the periodic variation of the direction coefficients will be predominant. Since these cutting speeds are beyond those practically used in milling steel workpieces, it is

experimentally impossible to determine the parametric stability boundary.

This paper is aimed towards devising a special machining set-up in which regenerative chatter is non-existent, thus facilitating the experimental investigation of the parametric instability. For this set-up a theoretical model will be developed to predict this type of instability. The stability chart experimentally obtained will be compared with the predicted one.

MACHINING SET-UP

For the experimental determination of the parametric instability boundary a special machining process, simulating the milling process without regenerative chatter, has been developed. In this set-up a screw chasing operation is performed, using, as illustrated in figure 2, a specially designed workpiece supported between centres on a centre lathe. This workpiece consists of a threaded cylindrical mid-part with a rectangular bar on each side. The stiffness of these rectangular bars in one direction is considerably higher than that in the other direction, thus simulating a single degree of freedom system. The screw chasing operation eliminates the regenerative effect, whereas the rotation of the principal axes of the rectangular sections of the workpiece introduces the periodic variation of the workpiece stiffness, which is equivalent to the periodic variation of the direction coefficients in the case of a milling process.

Such a machining set-up had been previously dealt with by Sridhar et al.⁵, though their model failed to predict the unstable lobes as those illustrated in figure 1, but gave only a single stability borderline almost parallel to the horizontal axis of the stability chart.

Workpiece design

The main factor which influenced the design of the test workpiece illustrated in figure 2, and used in the screw chasing process, was that its natural frequency should be sufficiently low to enable the cutting tests to be performed at practical cutting speeds. The increase in the natural frequency ω_0 will cause a

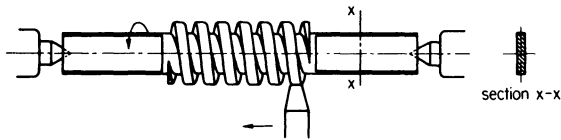


Figure 2. Test workpiece used in the special machining operation.

corresponding increase in the rotational speed Ω at which the unstable lobes occur. This condition is also essential in order to ensure that the dynamic characteristics of the lathe structure on which the workpiece is clamped, do not influence the stability conditions. The dynamic characteristics of this workpiece as obtained from the impulse testing technique are as follows:

Natural frequency $f_0 = 35$ Hz
 Damping factor $D = 0.00191$

THEORETICAL ANALYSIS

The screw-chasing operation is represented by the theoretical model illustrated in figure 3 where the principal axis of the major mode of the workpiece is indicated by the direction x along which the stiffness and damping factor are given by λ and c respectively. The principal axis rotates at an angular speed of Ω and thus the instantaneous direction of vibration is defined by the angle Ωt . The steady state cutting force F is inclined to the normal to the cut surface by an angle Ωt .

The vibrational motion of such a system is given by the following differential equation

$$m\ddot{x} + c\dot{x} + \lambda x = dF \tag{1}$$

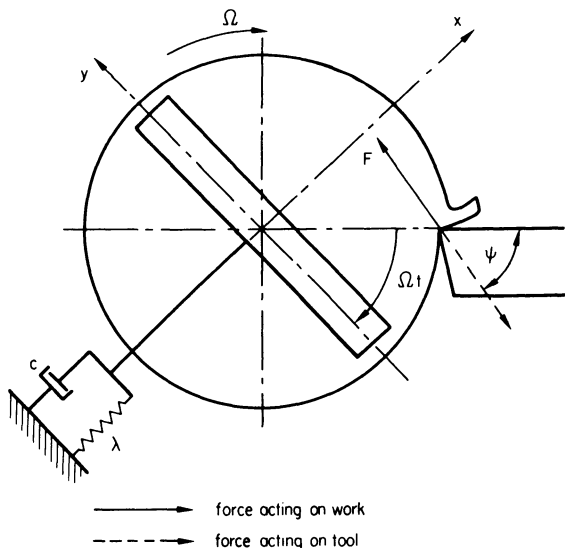


Figure 3. Schematic representation of vibratory model.

where dF is a self-induced incremental cutting force. If the penetration rate effect is neglected due to its insignificance at high cutting speeds¹ the expression for the incremental cutting force will be as follows, representing only the chip thickness dependent component

$$dF = -wk_1 x \sin \Omega t \sin (\Omega t - \psi) \tag{2}$$

Manipulating Equations (1) and (2) will render the following

$$\ddot{x} + \frac{\omega_0 \dot{x}}{Q} + \omega_0^2 \left(1 - \frac{wk_1}{2} (\cos (2\Omega t - \psi) - \cos \psi) \right) x = 0 \tag{3}$$

This equation is transformed into the form of Mathieu's equation by introducing the co-ordinate u defined by

$$x = e^{-(r/2Q)z} u \tag{4}$$

where

$$r = \frac{\omega_0}{\Omega} \quad z = \Omega t$$

thus leading to:

$$\frac{d^2 u}{dz^2} + (\alpha - 16\delta \cos (2z - \psi)) u = 0 \tag{5}$$

where

$$\alpha = r^2 \left(1 + \frac{wk_1}{2} \cos \psi - \frac{1}{4Q^2} \right), \delta = \frac{wk_1}{32\lambda} r^2 \tag{6}$$

and

$$z = \Omega t.$$

The stability conditions of such an equation have already been established and given in the form of the Ince-Strodt diagram⁶⁻¹¹. As it stands, this diagram is unsuitable for predicting the stability of Equation (3) unless the damping term is eliminated. Since the stability borderlines for Equation (3) lie within the unstable regions of this diagram⁶ it is, desirable to determine the unstable solutions of Equation (5) from which, using the transformation (4), the stability conditions of the basic Equation (3) can be derived.

Several methods are proposed for determining these unstable solutions, amongst which the one introduced by Whittaker^{9,10} is applicable for small values $|\delta|$. These solutions, as pointed out by Whittaker, are in the form of

$$u = e^{\mu z} \phi(z) \tag{7}$$

in which $\phi(z)$ is a quasi-periodic function containing harmonic terms, the coefficients of which are powers of δ . The substitution of this solution in Equation (5) gives the following final results for the first unstable region, after lengthy mathematical manipulation^{3,8}

$$\begin{aligned} \mu &= 4\delta \sin 2\sigma - 12\delta^3 \sin 2\sigma \dots \\ \alpha &= 1 + 8\delta \cos 2\sigma + \delta^2 (-16 + 8 \cos 4\sigma) \\ &\quad - 8\delta^3 \cos 2\sigma + \dots \end{aligned} \tag{8}$$

where σ is a phase angle between the motion u and the stiffness variation represented by the second term in the coefficient of u in Equation (5).

As far as the stability is concerned, only the characteristic exponent μ needs to be evaluated. For known values of α and δ the value of the phase angle σ is calculated from the second equation of (8), and by substituting this into the first, the value of μ is obtained. It can thus be decided whether the system is stable or otherwise depending on $\mu \leq 0$. If $\mu > 0$, the vibrational amplitude will grow with time, thus the system is unstable. On the other hand $\mu < 0$ indicates that the system is stable and the stability threshold of stability is obtained by equating μ to zero.

Similarly, the unstable solutions associated with the second and third unstable regions are given as follows^{8,11}

$$\mu = -4\delta^2 \sin 2\sigma + \dots \quad (9)$$

$$\alpha = 4 - \delta^2(16/3 - 32 \sin^2 \sigma) \dots$$

and

$$\begin{aligned} \mu &= \frac{4}{3}\delta^3 \sin 2\sigma \dots \\ &= 9 + 4\delta^2 + 8\delta^3 \cos 2\sigma + \dots \end{aligned} \quad (10)$$

Thus the stability boundary conditions of the original Equations (3) will be predicted using the above-mentioned equation. Inserting Equation (4) into (7) results in the following solutions

$$x = e(\mu \pm r/2Q)z\phi(z) \quad (11)$$

indicating that the condition of the threshold of stability is as follows:

$$\mu^2 = \left(\frac{r}{2Q}\right)^2 \quad (12)$$

For a known rotational speed Ω , the value of the characteristic exponent μ , satisfying the stability threshold condition, can be determined. Inserting this value in the sets of Equations (8)-(10) the parameters α and δ , corresponding to each point on the borderline of stability for each region, can be calculated by the simultaneous solution of each set. This is achieved by eliminating σ from the above mentioned equations, thus giving the first, second and third unstable regions respectively.

$$1 - 16\delta^2 - \alpha + A_1 + b_1 = 0 \quad (13)$$

$$1 - \frac{\mu^2}{16\delta^4} - A_2 = 0 \quad (14)$$

$$1 - \left(\frac{3\mu}{4\delta^3}\right)^2 - A_3 = 0 \quad (15)$$

where

$$A_1 = 2(1 - \delta^2) \frac{(4\delta - 12\delta^3)^2 - \mu^2}{1 - 3\delta^2}$$

$$B_1 = \frac{8\delta^2}{(4 - 12\delta^3)^2} ((4\delta - 12\delta^3)^2 - \mu^2)^2 - \mu^2$$

$$A_2 = \frac{4 + (32/3)\delta^2 - \alpha}{16\delta^2}$$

$$A_3 = \frac{(\alpha - 9 - 4\delta^2)^2}{8\delta^3}$$

Since α and δ are functions of wk_1/λ as given in Equation (6), the three Equations (13)-(15) can be converted to functions of a single variable wk_1/λ which can be solved numerically. Consequently α and δ can be calculated giving a stability chart in the α - δ plane in the same manner in which the stability conditions of Mathieu's equation are displayed.

However, if wk_1/λ , being proportional to the width of cut, is plotted against the workpiece speed, $N = 60\Omega/2\pi$, the theoretical stability chart is obtained in the practical form shown in figure 4. This

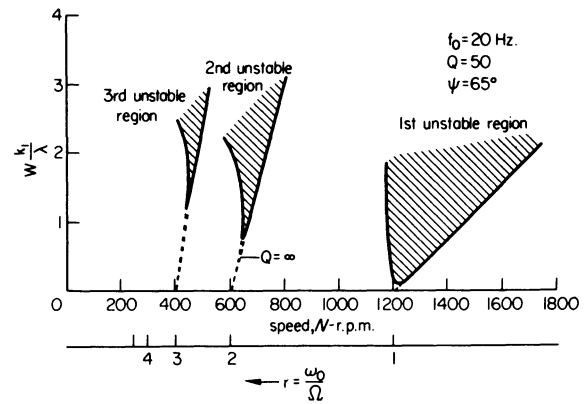


Figure 4. Predicted stability chart for a theoretical system.

is obtained for a hypothetical single degree of freedom system, for a cutting force direction inclined at an angle $\psi = 65^\circ$. In this figure the first three unstable lobes are shown shaded. It is also noticeable that the area of the first unstable lobe is considerably wider than the others whose areas decrease by the reduction of the cutting speeds. In addition the stable region separating two consecutive unstable lobes considerably diminishes as the cutting speed is reduced. This chart shows also that the level of stability increases at lower cutting speeds and is minimal at the first unstable lobe where the rotational speed of the workpiece, Ω , coincides with the natural frequency ω_0 . The increase in stability at low cutting speeds is caused by the damping in the system. In the absence of damping, Equation (3) would take the form of Mathieu's equation, whose boundary of stability meets the α axis at the points 1, 2, 3 as shown in figure 4 where the first, second and third unstable lobes for $Q = \infty$ touch the horizontal axis at points 1, 2, 3 respectively.

CHOICE OF CUTTING CONDITIONS

Equation (6) shows that the angle ψ between the cutting force and the normal to the cut surface will have a considerable influence on the stability boundary conditions of the machining process investigated. This effect is illustrated by figure 5, in which the third unstable lobe is plotted for a series of

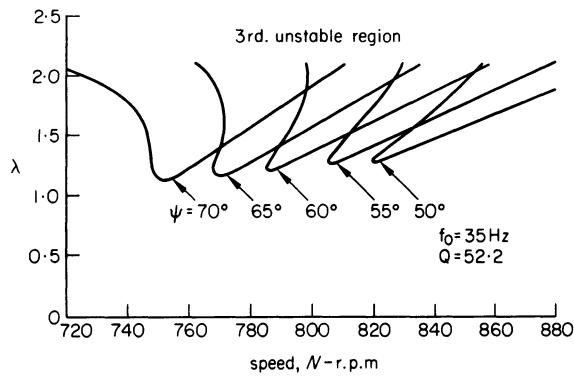


Figure 5. Effect of variation of ψ on the boundary of stability.

values of the angle ψ within the practical range $50^\circ < \psi < 70^\circ$. This figure shows that the increase of the angle ψ causes a considerable increase in the area of the unstable area and shifts its position to lower cutting speeds. It is therefore desirable to choose a large value of the angle ψ so that the cutting tests could be carried out at practical cutting speeds and the unstable lobe covers a wide speed range.

As this angle is primarily dependent on the rake angle¹³ increasing with the increase of the rake angle, a special tool holder, allowing the variation of the rake angle, was used in this investigation. Preliminary cutting tests were carried out, using different rake angles, to find out the appropriate one which should be used in the chatter cutting tests. This was found to be a rake angle of 1° corresponding to an angle ψ of 57° .

PREDICTED STABILITY CHART

For the workpiece used in this investigation, the dynamic characteristics of which were given earlier and for the above mentioned cutting conditions, the third unstable lobe has been calculated from Equation (10). The resulting lobe is plotted in figures (6) and (7) in the α - δ and (wk_1/λ) - N co-ordinate planes respectively. The stability borderlines of Mathieu's equation for this region chain-dotted in figure 6 show, as expected, that the level of stability is raised by introducing a damping term in Mathieu's equation. The variation of μ along the borderline of stability, plotted at the top of this figure shows that this variation is small in the vicinity of 0.025 and hence it is expected that the resulting stability borderline would be very close to the iso- μ curve⁸ shown dashed in the figure, for this specific value of μ . This verifies the accuracy of the analysis carried out to predict the stability of the damped system.

CHATTER CUTTING TESTS

Cutting tests were carried out on the workpiece illustrated in figure 8 at a constant nominal uncut chip thickness of 0.00125 in. At each selected speed, cuts were taken at various widths and the vibration of the workpiece, as sensed by a strain gauge bridge cemented on one of the workpiece end bars, was recorded. The operation was considered to be stable during a certain pass if the records obtained during that cut showed no identifiable build up of the

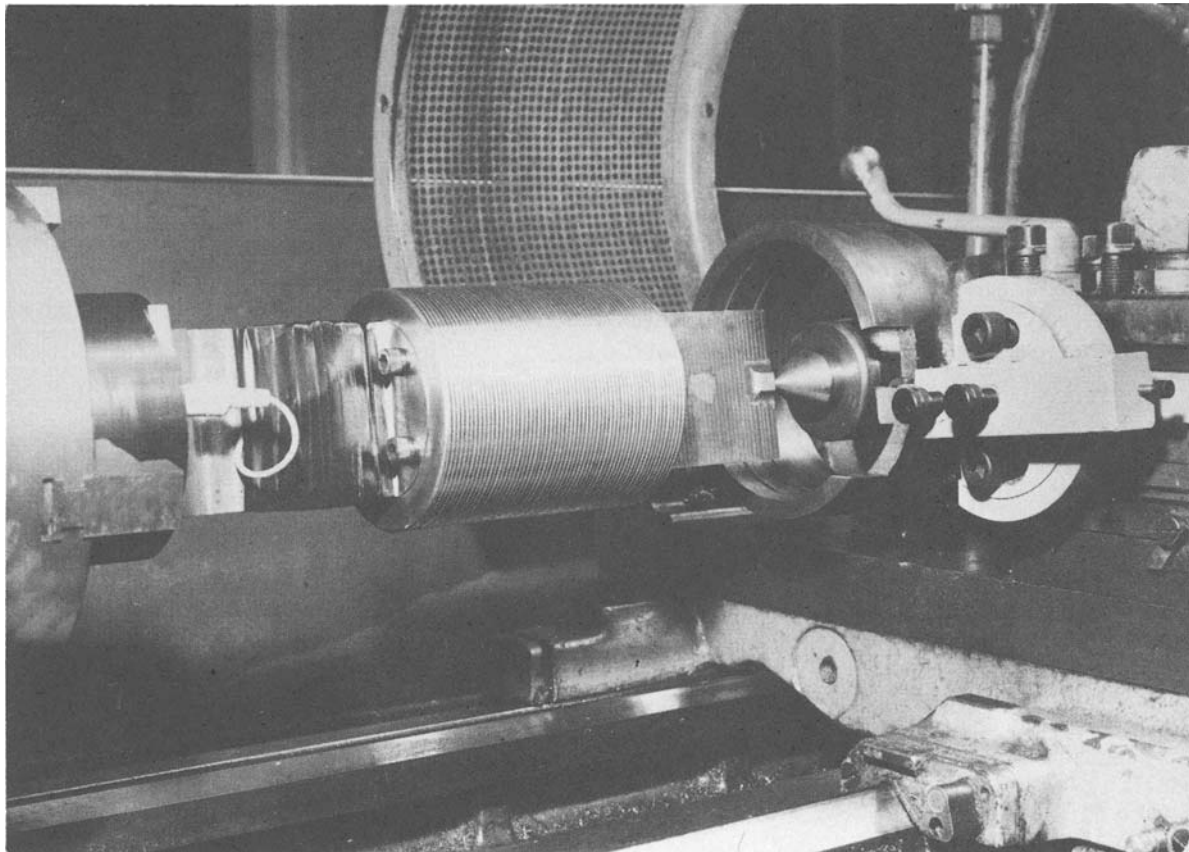


Figure 6. General view of the experimental set-up.

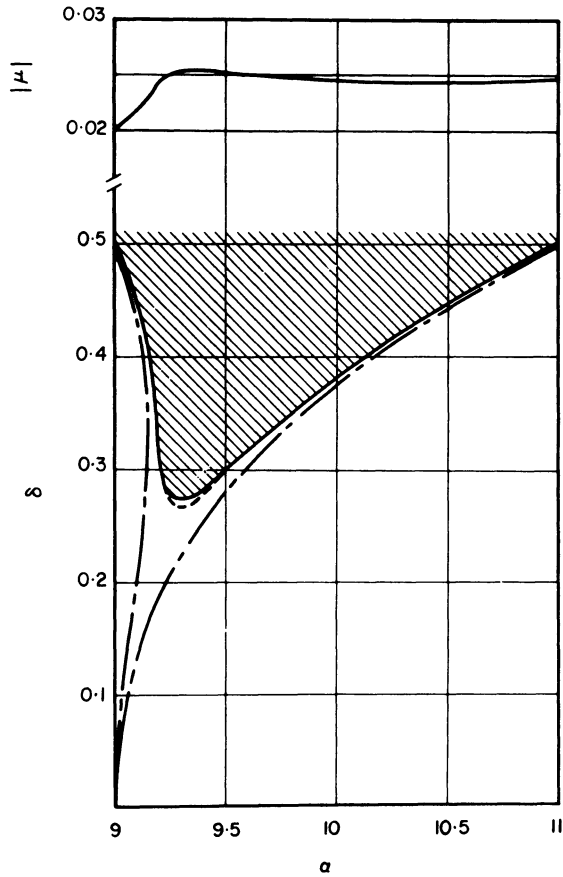


Figure 7. Predicted third unstable region in the $\alpha - \delta$ plane.

vibrational motion as the tool travelled along the workpiece, as illustrated by record *a* in figure 9. On the other hand, violent build-up of vibration indicated that the operation was unstable. However the build-up of vibration growth was shown by the records to take place only during the first few revolutions of the workpiece after the start of the tool engagement, followed by random oscillations with a drop in the vibration amplitude, as illustrated by record *b* of figure 9. This drop of vibrational motion is caused by the workpiece starting to leave the tool at large amplitudes of vibrations, thus leading to a machining operation accompanied by impacts. These impacts will disturb the vibrations which have started, thus resulting in large deflections of the

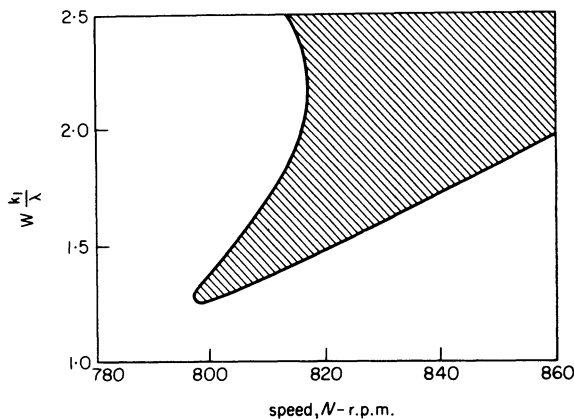


Figure 8. Predicted third unstable region in the conventional co-ordinate system.

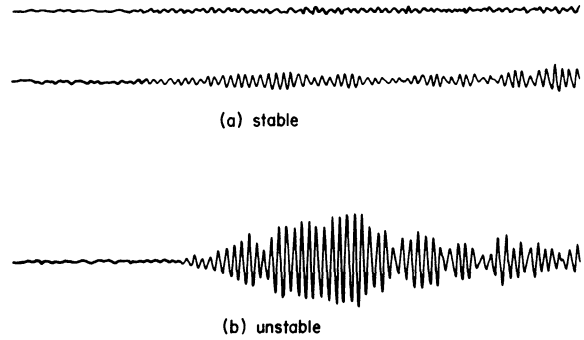


Figure 9. Vibration records.

workpiece. The signals resulting from these deflections were filtered out by a band-pass filter inserted in the measuring circuit to cut off the low frequency signals resulting from the deflection of the workpiece due to its own weight.

The widths of cut in the vicinity of the transition between stable and unstable cutting conditions determined in the manner described above, are plotted versus the cutting speed, forming the third lobe of the experimental stability chart as illustrated in figure 10. The good correspondence between this chart and the predicted one, illustrated in figure 8, verifies the developed theoretical model for predicting parametric instability.

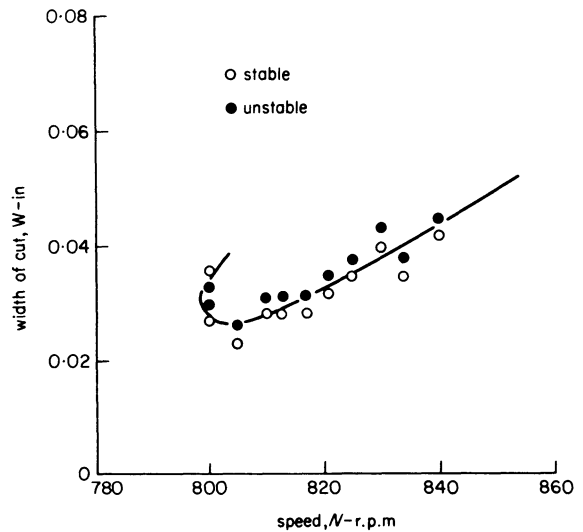


Figure 10. Experimental third unstable region.

For quantitative assessment of the theoretical analysis, the chip thickness coefficient k_1 for the workpiece was determined from a matching process between the experimental and theoretical stability results. The stiffness of the workpiece, being 3.57×10^3 lbf/in, as measured from static loading tests, was used and it was found that k_1 is of the order of 16×10^4 lbf/in unit width of cut. This value coincides with that found during the turning operation of the same material under regenerative chatter conditions¹².

CONCLUSIONS

Dynamic instability of the milling process is either of the regenerative type whose cause is well known or of

the parametric type resulting from the cyclic variation of the dynamic characteristics of the vibratory system. The analysis carried out to predict the regenerative instability has been confirmed for any average number of teeth simultaneously engaged with the workpiece⁴. The theoretical analysis presented in this work for predicting the conditions of parametric instability has been verified by the good agreement obtained between the predicted and experimental charts.

At low cutting speeds it has been found that the parametric instability occurs at higher widths of cut as compared to those at which regenerative chatter occurs. This illustrates the insignificance of the parametric effect in this low speed range. However, at higher cutting speeds than those corresponding to the first lobe in the regenerative stability chart, the parametric instability may occur at comparatively low widths of cut. It is thus necessary, when machining light alloys at high cutting speeds, to take the parametric instability into consideration.

The analysis developed in this work to predict the instability of the special screw-chasing operation is equally applicable for non-regenerative machining operations accompanied by a periodic stiffness variation of the vibratory system, such as cutting bolts from hexagonal bars.

ACKNOWLEDGMENT

The authors wish to express their gratitude to Professor S. A. Tobias, Head of the Mechanical Engineering Department, University of Birmingham, for his continuous guidance and advice given throughout the investigation. Thanks are due to Mr A. G. Petter for his assistance in the experimental tests.

REFERENCES

1. M. M. Sadek and S. A. Tobias. Comparative dynamic acceptance tests for machine tools applied to horizontal milling machines, *Proc. Inst. Mech. Engrs.* (1970-71) 185, London.
2. G. Sweeney and S. A. Tobias. An algebraic method for the determination of the dynamic stability of machine tools, *Inst. Res. in Prod. Eng. Conf.*, Pittsburg, p. 475, 1963.
3. G. Droubi. *Instability During Continuous and Interrupted Horizontal Milling*, Ph.D. Thesis, The University of Birmingham, 1972.
4. G. Droubi, M. M. Sadek and S. A. Tobias. The effect of the instantaneous force orientation on the stability of horizontal milling, to be submitted to the *Inst. Mech. Engrs.*, London, 1974.
5. R. Sridhar, E. R. Hohn and G. W. Long. A stability algorithm for a special case of the milling process—contribution to Machine Tool Chatter Research, Part 6, Journal of Engineering for Industry, *Trans. A.S.M.E. Series B* (1968) 90, 325.
6. N. W. McLachlan. *Theory and Application of Mathieu Functions*, Oxford University Press, 1947.
7. N. W. McLachlan. *Ordinary Non-linear Differential Equations*, 2nd Edition, Oxford Press, 1956.
8. C. Hayashi. *Non-linear Oscillations in Physical Systems*, McGraw-Hill Book Co. New York, 1964.
9. E. T. Whittaker. General solution of Mathieu's equation, *Proc. Edinburgh Math. Soc.*, (1913-14) 32, 75.
10. E. T. Whittaker and G. N. Watson. Cambridge University Press, London, 1935.
11. A. W. Young. Quasi-periodic solutions of Mathieu's Equation, *Proc. Edinburgh Math. Soc.* (1913-14) 32, 81.
12. W. A. Knight and S. A. Tobias. Torsional vibrations and machine tool stability, *Proc. 10th Int. M.T.D.R. Conf. Manchester*, p. 299, 1969.
13. M. M. Nigm, M. M. Sadek and S. A. Tobias. Prediction of dynamic cutting coefficients from steady state cutting data, *Proc. 13th M.T.D.R. Conf. Birmingham*, 1972.
14. G. Droubi, M. M. Sadek and S. A. Tobias. Determination of the dynamic cutting coefficients for milling. *Proc. 13th M.T.D.R. Conf. Birmingham*, 1972.

ELASTIC WAVE TRANSMISSION AND REFLECTION IN LONG UNIFORM BARS DUE TO THE END IMPACT OF A RIGID MASS

by

W. JOHNSON*

SUMMARY

The manner in which the elastic exponential stress wave, generated by the impact of a *rigid* mass on the end of a long elastic bar, is related to that which would be generated by the impact of an equal *elastic* mass, is demonstrated. The stress wave systems, generated by the impact of a rigid mass on the end of a bar or bars which possess a rigid mass at the other end, is also investigated, and stress wave reflection is analysed when the distal end of a bar is only partly fixed.

NOTATION

A	cross-sectional area
E	Young's modulus
G	torsion modulus
I	moment of inertia
J	second polar moment of area
M	mass of rigid body
T	time to travel to end of bar and return
T_Q	torque
z	$A_1 \rho_1 c_1 / (A_1 \rho_1 c_1 + A_2 \rho_2 c_2)$
a	radius of shaft
c	longitudinal wave speed
c_T	torsional wave speed
e	strain
l	length of bar or shaft
i_m	moment of inertia of shaft about its axis
m	mass of bar
t	time
v	speed
ϕ	shear strain
ρ	density
ρc	impedance
σ	compressive or tensile stress
ω	angular speed

INTRODUCTION

Impact is a feature in metal forming processes, for example in forging presses. The consequences of impact for the equipment at large do not always seem to be fully anticipated, and this is especially the case regarding simple elastic stress wave propagation between one element and another. The original and classical analytical determination of the elastic stress

wave in a long stress free bar, when struck by a mass at one end, is due mainly to Boussinesq and St. Venant: their aim was to determine how the stress at a given section of the bar varies with time. The problem becomes especially complex when multiple reflections from a fixed end are considered. *Familiarity with this work is rather uncommon among mechanical engineers, but since it is important and in order that it be better known, a number of situations somewhat similar to those originally treated are presented below.* It is hoped that the cases analysed will be of value to design and research engineers in as far as they extend the applicability of the earlier authors' approach and because they present some new and useful features.

Idealized elements (i.e. uniform bars (or springs) and rigid masses) are dealt with rather than real machine elements (e.g. tie rods, bolts and bolsters), but this should not detract from an appreciation of analyses involving the latter. We consider the simple stress states in a long bar in tension or compression and no buckling or bending is envisaged. First, derivations of the most elementary expressions for stress intensity due to impact and stress wave speed are given, followed by an introduction to the classical problem.

Section 3 tries to illustrate a rigid striker in elastic impact situations. We shall see that the term *rigid* comes to refer to the neglect of stress wave effects in the mass concerned; *rigid* also implies an infinite Young's modulus and, since wave speed equals $(E/\rho)^{1/2}$ it also implies an infinite wave speed.

Sections 4, 5 and 6, to different degrees, re-examine the classical problems with some attention to multiple wave reflection; first with a rigid mass

* Department of Mechanical Engineering, U.M.I.S.T.

attached to the free end, and then when a chain of bars is involved. In the final section 8, an attempt is made to put into useful graphical form some of the equations previously derived in order to facilitate rapid calculation, or the speedy appreciation of the consequences of change in cross-section shape and bar end condition. This latter section also deals with partially fixed end conditions.

1 ELEMENTARY BASIC EQUATIONS

Two identical, long, uniform bars moving with the same speed, v_0 , impinge head-on (figure 1). At time $t = l/c$ after impact first took place (where l is a bar length, and c the elastic stress wave speed in it), both bars will have come to rest and all their kinetic energy will have been converted into elastic strain energy; the compressive stress being σ . We have then

$$\frac{1}{2}(A\rho l)v_0^2 = \frac{1}{2}(Al)\frac{\sigma^2}{2E} \tag{1.1}$$

A is the cross-sectional area of a bar, ρ is its density and E its Young's modulus. Hence,

$$\rho v_0^2 = \frac{\sigma^2}{E} \text{ or } \sigma = v_0(E/\rho)^{1/2} \tag{1.2}$$

The strain in each bar will be

$$e = \sigma/E = v_0 t/ct \text{ or } \sigma = Ev_0/c \tag{1.3}$$

By the time knowledge of the impact at the common face of the bars has been propagated to the opposite end, the latter will have travelled a distance $v_0 t$ and hence each bar will be compressed by an amount $v_0 t$.

From Equations (1.2) and (1.3)

$$c = (E/\rho)^{1/2} \tag{1.4}$$

and eliminating E between them gives

$$\sigma = \rho c v_0 \tag{1.5}$$

The Equations (1.2) to (1.5) are fundamental for elementary elastic stress wave analysis.

If we repeat the above exercise with the homologous system in which a pair of identical shafts, each rotating about its axis but in opposite directions, are

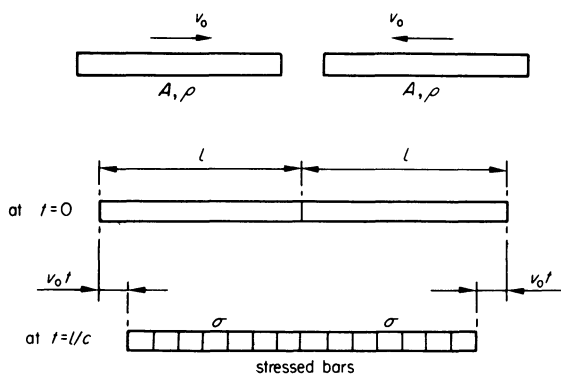


Figure 1.

brought together so that their common interface is immediately arrested and torsional stress waves are thereafter propagated away and into each, we have

$$\frac{1}{2}\left(A\rho l \cdot \frac{a^2}{2}\right)\omega_0^2 = \frac{1}{2}T \cdot \frac{Tl}{JG} \tag{1.6}$$

where $A\rho l a^2/2$ is the moment of inertia of a shaft about its axis. Hence

$$T/J = \omega_0(G\rho)^{1/2} \tag{1.7}$$

The torsional strain in a shaft is

$$\phi = r\omega_0 t/c_T t = \tau/G \tag{1.8}$$

where c_T is the torsional wave speed, so that

$$G\omega_0/c_T = \tau/r = T/J \tag{1.9}$$

From (1.7) and (1.9)

$$G\omega_0/c_T = \omega_0(G\rho)^{1/2}$$

or

$$\left. \begin{aligned} c_T &= (G/\rho)^{1/2} \\ T &= JG\omega_0/c_T \end{aligned} \right\} \tag{1.10}$$

and

2 THE RIGID STRIKER

A rigid mass M impinges on the proximal end of a long, uniform elastic bar with speed v_0 and generates in it an elastic stress wave (figure 2(a)). Let the speed

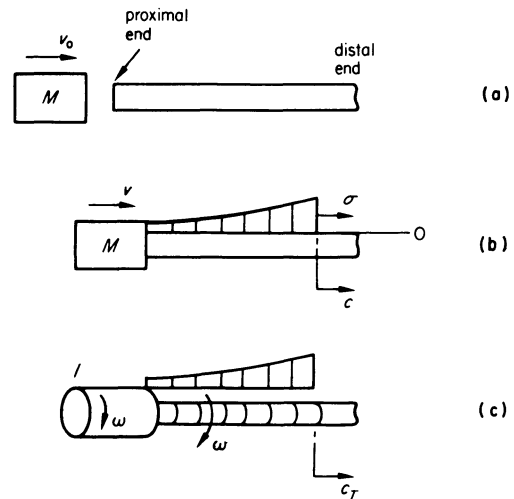


Figure 2.

at the striker-bar interface be v at time t after impact takes place and let the interfacial stress be σ .

The equation of motion for M is

$$M \frac{dv}{dt} = -A\sigma \tag{2.1}$$

Substituting for σ , from Equation (1.5), into Equation (2.1)

$$M \frac{dv}{v} = -A\rho c \cdot dt \tag{2.2}$$

Integrating (2.2), and noting that $v = v_0$ at $t = 0$

$$\left. \begin{aligned} v &= v_0 \exp\left(-\frac{A\rho ct}{M}\right) \\ \sigma &= \rho cv_0 \exp\left(-\frac{A\rho ct}{M}\right) \end{aligned} \right\} \quad (2.3)$$

For a bar of finite length l , we might more usefully write

$$\left. \begin{aligned} v &= v_0 \cdot \exp\left(-\frac{m}{M} \cdot \frac{t}{l/c}\right) \\ \sigma &= \sigma_0 \exp\left(-\frac{m}{M} \cdot \frac{t}{l/c}\right) \end{aligned} \right\} \quad (2.4)$$

where m is the mass of the bar.

The form of the stress wave thus propagated in the elastic bar has the exponential form shown in figure 2(b).

Starting from Equations (2.4), St. Venant and others have carried out a very thorough analysis of elastic stress wave reflection in a *finite length* elastic bar fixed at the distal end.

As before, considering a rigid mass with moment of inertia I about its axis and possessing an initial angular velocity ω_0 , we may repeat the analysis immediately above. This time, however, consider the circumstance in which the mass suddenly impinges on, or is firmly applied to, the end of a long, uniform, stationary cylindrical shaft, and so causes a torsional stress wave to be propagated along it (figure 2(c)). (Such a situation arises when a jaw clutch is engaged). The equation of rotational motion for the mass is,

$$T \frac{d\omega}{dt} = -T_Q \quad (2.5)$$

where ω is its current angular speed about the axis at time t after the commencement of impact, and T_Q is the applied torque due to the torsional stress wave. Substituting from Equation (1.10) for T_Q , (2.5) becomes

$$T \frac{d\omega}{dt} = -\frac{JG}{c_T} \omega \quad (2.6)$$

Integrating (2.6) and noting that $\omega = \omega_0$ at $t = 0$, it is found that

$$\omega = \omega_0 \cdot \exp\left(-\frac{JG}{Ic_T} \cdot t\right) \quad (2.7)$$

$$= \omega_0 \cdot \exp\left(-\frac{i_m}{I} \cdot \frac{c_T t}{l}\right) \quad (2.8)$$

where $i_m = \pi a^2 l \rho / 2$ is the inertia of the whole shaft about its axis. Clearly Equations (2.4) and (2.8) are formally identical, the mass ratio of Equation (2.4) being replaced by the inertia ratio of Equation (2.8).

Now it may strike the reflective reader as somewhat curious, if not inconsistent, that this latter discussion should combine the notions of a *rigid* striker and an *elastic* bar. At least this prompts a

re-consideration of the limitations to ideal rigidity in this context. We shall see, as we may have anticipated, that rigidity is a valid notion when the impedance (ρc) of the striker is very much greater than that of the bar, but we shall also note the manner in which that notion is realised.

3 THE RIGID STRIKER AS THE LIMIT OF A HIGH IMPEDANCE BAR

To investigate the implications attached to the notion of the rigidity of the striker in the preceding section, we must re-examine the situation (figure 3(a)). Let us

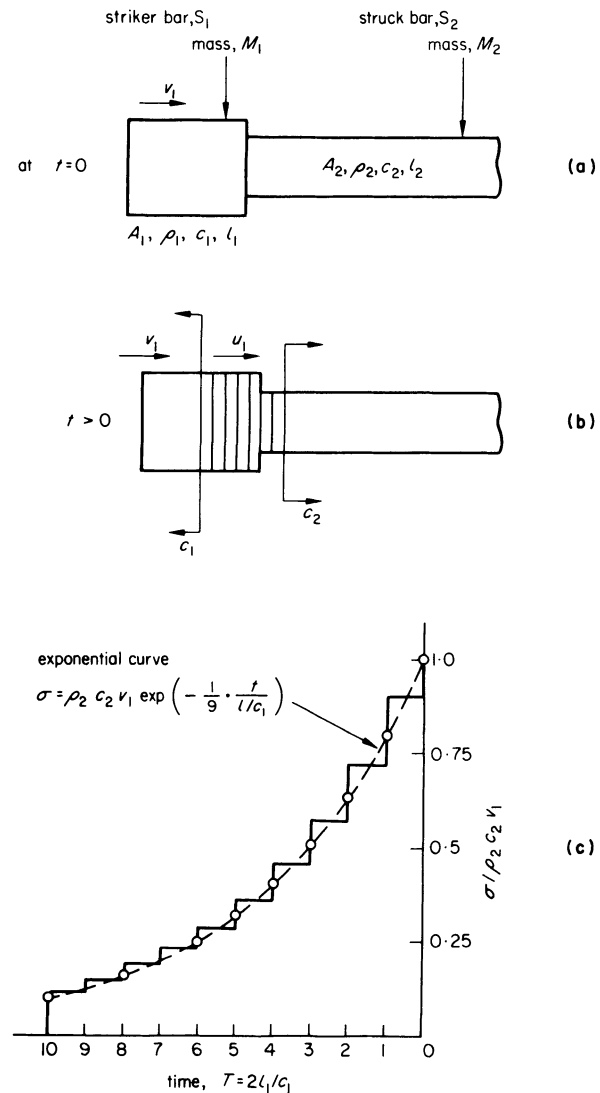


Figure 3.

suppose that $A_1 \rho_1 c_1 / A_2 \rho_2 c_2$, i.e. the ratio of the total impedance of the striker and the struck bar, is large. The striker S_1 is treated as an elastic bar which is continuously loaded and unloaded, its speed decreasing as energy is fed forward and into the struck bar, S_2 .

After impact has taken place, S_1 is initially unstressed and moving with speed v_1 . Let the speed of the portions of S_1 and of S_2 into which stress waves have travelled as a consequence, be u_1 (figure 3(b)).

Then, at the interface, since the actions on the end of both bars must be equal

$$A_1 \rho_1 c_1 (v_1 - u_1) = A_2 \rho_2 c_2 u_1$$

and

$$u_1 = \frac{A_1 \rho_1 c_1}{A_1 \rho_1 c_1 + A_2 \rho_2 c_2} \cdot v_1 = z \cdot v_1 \quad (3.1)$$

where z is a constant for given bars S_1 and S_2 .

Since our assumption was that $A_1 \rho_1 c_1 / A_2 \rho_2 c_2$ is large, then z is slightly less than unity and the stress at the interface will be $\sigma_1 = \rho_2 c_2 u_1 = z \rho_2 c_2 v_1$.

At time T after impact first took place (where $T = 2l_1/c_1$) the initiated (loading) stress wave will have travelled to the free end and been reflected (unloading) back to the interface. The whole of bar S_1 will then be stress free, but its translational speed will have been reduced to $(2u_1 - v_1)$. This is now the 'new' impact speed of S_1 so that for the period of time $2l_1/c_1 < t < 4l_1/c_1$, or $T < t < 2T$ the interface will have a speed of u_2 where

$$u_2 = z(2u_1 - v_1) = zv_1(2z - 1) \quad (3.2)$$

Using Equation (3.1), the stress will be

$$\sigma_2 = z\rho_2 c_2 v_1(2z - 1). \quad (3.3)$$

Similarly, during the period $4l_1/c_1 < t < 6l_1/c_1$ or $2T < t < 3T$

$$u_3 = v_1(2z - 1)^2 \quad (3.4)$$

and

$$\sigma_3 = z\rho_2 c_2 v_1(2z - 1)^2 \quad (3.5)$$

Thus, generalizing the above results, during the period $(n - 1)T < t < nT$

$$u_n = v_1(2z - 1)^{n-1}$$

and

$$\sigma_n = z\rho_2 c_2 v_1(2z - 1)^{n-1} \quad (3.6)$$

Evidently, S_1 reduces its speed so that at times $0, T, 2T, \dots$ the whole bar has speeds given by u_1, u_2, u_3, \dots ; i.e. successive interfacial speeds are reduced by a factor of $(2z - 1)$. In bar S_2 there will be a stress wave, the length of each of whose steps is $c_2 T$, and the stress intensity which reduces discontinuously with time is successively given by $\sigma_1, \sigma_2, \sigma_3, \dots$

Clearly

$$\frac{\sigma_{n+1}}{\sigma_n} = \frac{2z - 1}{1}. \quad (3.7)$$

Now, if we had assumed S_1 to be rigid, by adapting Equation (2.3), we should have had

$$\sigma = \rho_2 c_2 v_1 \exp\left(-\frac{A_2 \rho_2 c_2 t}{A_1 \rho_1 l_1}\right) \quad (3.8)$$

where c_2 is the wave speed in S_2 .

Thus using $t_n = n.T$ and $t_{n+1} = (n + 1).T$, equation (3.8) provides

$$\begin{aligned} \frac{\sigma_{n+1}}{\sigma_n} &= \frac{\exp\left(-\frac{A_2 \rho_2 c_2}{A_1 \rho_1 l_1} \cdot (n + 1)T\right)}{\exp\left(-\frac{A_2 \rho_2 c_2}{A_1 \rho_1 l_1} \cdot nT\right)} \\ &= \exp\left(-\frac{A_2 \rho_2 c_2}{A_1 \rho_1 l_1} \cdot T\right) \\ &= \exp\left(-\frac{A_2 \rho_2 c_2}{A_1 \rho_1 l_1} \cdot \frac{2l_1}{c_1}\right) \\ &= \exp\left(-2 \frac{A_2 \rho_2 c_2}{A_1 \rho_1 c_1}\right) \\ &= \exp - 2 \left(\frac{1 - z}{z}\right) \end{aligned} \quad (3.9)$$

Since z is slightly less than unity, $(1 - z)/z$ is small and hence

$$\frac{\sigma_{n+1}}{\sigma_n} \simeq 1 - \frac{2(1 - z)}{z} \quad (3.10)$$

By writing $z = 1 - y$, where y is a small quantity, from Equation (3.10) we obtain

$$\frac{\sigma_{n+1}}{\sigma_n} = 1 - \frac{2y}{1 - y} \simeq 1 - 2y \quad (3.11)$$

This is the same as is given using Equation (3.7).

Evidently the stress wave, generated by a mass such as S_1 for which $A_2 \rho_2 c_2 / A_1 \rho_1 c_1 = y$ is small, or alternatively $A_1 \rho_1 c_1 / A_2 \rho_2 c_2$ is large, (and therefore the impedance of the striker to that of the rod is large), generates a stress wave in S_2 which tends towards the exponential as y becomes smaller. In particular, if S_1 is said to be rigid, then $E \rightarrow \infty$, $c_1 \rightarrow \infty$, and $y \rightarrow 0$.

In figure 3(c), at times $t = 0, T, 2T \dots 10T$ (where $T = 2l_1/c_1$), the stepped stress wave introduced into S_2 by the relatively rigid S_1 is shown. Using Equations (3.3), (3.5), etc., when $z = 0.9$ or $A_2 \rho_2 c_2 / A_1 \rho_1 c_1 = 1/9$ we have,

$$\sigma_1 / \rho_2 c_2 v_1 = z = 0.9$$

$$\sigma_2 / \rho_2 c_2 v_1 = z(2z - 1) = 0.9 \times 0.8 = 0.72$$

and

$$\sigma_3 / \rho_2 c_2 v_1 = z(2z - 1)^2 = 0.9 \times 0.8^2 = 0.576$$

etc.

From Equation (3.8), the form of the exponential stress wave for the corresponding rigid striker is

$$\frac{\sigma}{\rho_2 c_2 v_1} = \exp\left(-\frac{1}{9} \cdot \frac{t}{l_1/c_1}\right) \quad (3.12)$$

This, at non-dimensional times $t/l_1/c_1 = 2, 4, \dots$, provides

$$\frac{\sigma}{\rho_2 c_2 v_1} = \exp\left(-\frac{2}{9} n\right) \quad (3.13)$$

and this is shown in figure 3(c).

4 THE IMPACT WITH INITIAL SPEED v_0 OF A RIGID STRIKER OF MASS M ON TO THE PROXIMAL END, A , OF A LONG LINEAR ELASTIC BAR (OR SPRING) OF MASS m TO THE END, B , OF WHICH IS ATTACHED AN EQUAL STATIONARY RIGID MASS M

To analyse the problem we assume the following sequence of events (figure 4(a)):

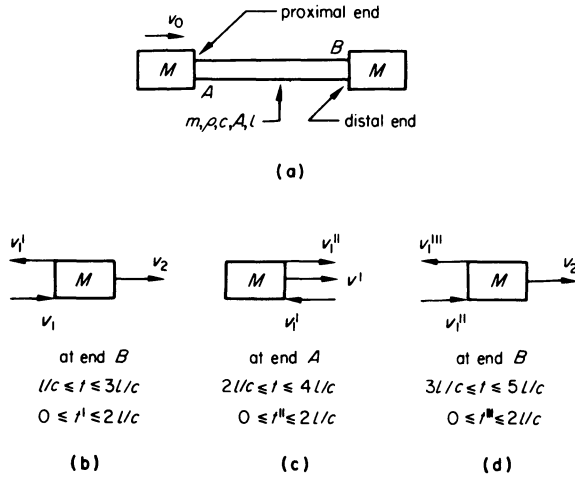


Figure 4.

- (i) When impact at A occurs at time $t = 0$, a stress wave, exponential in form, is initiated and travels along the bar to reach the mass at B after time l/c .
- (ii) This incident wave is reflected by the mass at B , and, because it is initially stationary, the *head* of the reflected wave is of the same intensity as that of the incident wave. This is the same as saying that at $t = l/c$, the mass at B behaves as if it were a fixed mass.
- (iii) For $t > l/c$, the mass at B moves and effectively allows a *release* wave to be propagated into the bar towards A . If the M/m ratio is less than 2, the total force on the mass at B falls to zero before $t = 2l/c$. Thus if the mass is not positively fastened to the bar, contact will cease and the mass will be effectively ejected away from B .
- (iv) The wave reflected by the mass at B after time $t = 2l/c$, is reflected by the mass at A so that its motion is further retarded, and subsequently it comes to rest. Restitution then occurs, and, at some later time, the force between the bar and the mass at A is reduced to zero so that contact ceases and it is ejected.

At time t after the initial impact at A , the stress intensity at the striker-bar interface is σ_1 and the corresponding striker speed is v_1 . From Equation (2.4) we have

$$\sigma_1 = \sigma_0 e^{-m/M \cdot ct/l}$$

and

$$v_1 = v_0 e^{-m/M \cdot ct/l}$$

These relationships apply for $0 < t < 2l/c$.

First reflection from the distal end B

After time l/c , the wave generated at impact reaches B and is thereafter reflected. After a further time t' , let the absolute speed of the mass at B be v_2 , and the speed due to the wave reflected from the mass at B be v_1' (figure 4(b)). The equation of motion for the mass at B , assuming that the reflected wave is compressive and of intensity σ_1' , is

$$M \frac{dv_2}{dt'} = A(\sigma_1 + \sigma_1')$$

or

$$\frac{dv_2}{dt'} = \frac{mc}{Ml} (v_1 + v_1') \quad (4.1)$$

However,

$$v_2 = v_1 - v_1'$$

or

$$v_1' = v_1 - v_2 \quad (4.1a)$$

so that substituting for v_1' in Equation (4.1)

$$\frac{dv_2}{dt'} + \frac{mc}{Ml} v_2 = 2 \frac{mc}{Ml} \cdot v_1 \quad (4.2)$$

and Equation (4.2) becomes

$$\frac{d}{dt'} (v_2 \cdot e^{mct'/Ml}) = 2 \frac{mc}{Ml} \cdot v_0 \quad (4.3)$$

after introducing Equation (2.4) in Equation (4.3) but necessarily replacing t with t' . Hence

$$v_2 e^{mct'/Ml} = 2 \frac{mc}{Ml} \cdot v_0 t' + \text{constant} \quad (4.4)$$

The constant in Equation (4.4) is zero because at $t' = 0$, $v_2 = 0$; mass B must have zero speed at $t' = 0$, if only a finite force is applied to it through the bar. Thus

$$\frac{v_2}{v_0} = 2 \frac{mc}{Ml} \cdot t' \cdot e^{-mct'/Ml} \quad (4.5)$$

From Equation (4.1a) we have, using Equations (2.4) and (4.5)

$$\frac{v_1'}{v_0} = \frac{v_1}{v_0} - \frac{v_2}{v_0} = e^{-mct'/Ml} \left(1 - 2 \frac{mc}{Ml} \cdot t' \right) \quad (4.6)$$

The total stress between the bar and mass B is

$$\begin{aligned} \frac{T\sigma_2}{\sigma_0} &= \frac{v_1}{v_0} + \frac{v_1'}{v_0} = 2 \frac{v_1}{v_0} - \frac{v_2}{v_0} \\ &= 2e^{-mct'/Ml} \left(1 - \frac{mc}{Ml} \cdot t' \right) \end{aligned} \quad (4.7)$$

Stress $T\sigma_2$ becomes zero when

$$\frac{t'}{l/c} = \frac{M}{m} \quad (4.8)$$

This condition can apply provided $0 \leq ct'/l \leq 2$, i.e. for $\infty > m/M \geq \frac{1}{2}$. If this condition does apply, then the speed which the distal mass at B has acquired

when the force between it and the spring has fallen to zero is, from Equation (4.5)

$$\frac{v_2}{v_0} = 2e^{-1} \tag{4.9}$$

Also

$$\left. \begin{aligned} \frac{v'_1}{v_0} &= -e^{-1} \\ \frac{v_1}{v_0} &= e^{-1}, \text{ at the end } B \end{aligned} \right\} \tag{4.10}$$

and

$$\frac{v}{v_0} = e^{-(1+m/M)} \text{ at the end } A.$$

If the fixing of the mass to the end at *B* was unable to sustain tension, then separation would also take place at this time. Separation implies a free end bar, and the speed of the end *B* would thereafter be

$$2v_1 = 2v_0 \exp(-mct'/Ml) < 2v_0e^{-1}$$

The fraction of the input kinetic energy carried off by the ejected mass would be $4/e^2 \approx 0.54$.

If the rigid impinging mass *M* was conceived to collide with a stationary rigid bar of mass *m* which then collided with another stationary rigid mass *M*, and if the coefficient of restitution at each impact was unity, then the speed of the second mass would be reduced by a factor of $4Mm/(m+M)^2$. The result in Equation (4.9) shows this final speed to be independent of mass ratio when a stress wave approach is adopted.

First reflection from the proximal end *A*

Later, the wave reflected from mass *B* acts on the striker at *A* and is itself reflected (figure 4(c)), so that the absolute speed of the striker becomes v' . Its equation of motion is

$$M \frac{dv'}{dt''} = -A(\sigma'_1 + \sigma''_1)$$

σ''_1 is the stress in the wave reflected from the mass at *A*, and t'' is the time measured from first arrival of the wave reflected from end *B*.

Thus

$$\frac{dv'}{dt''} = -\frac{mc}{Ml} (v'_1 + v''_1) \tag{4.11}$$

But

$$v' = v''_1 - v'_1$$

or

$$v''_1 = v' + v'_1 \tag{4.12}$$

and substituting into Equation (4.11) we have

$$\frac{dv'}{dt''} + \frac{mc}{Ml} \cdot v' = -2 \frac{mc}{Ml} \cdot v'_1 \tag{4.13}$$

Hence

$$\frac{d}{dt''} \left(\frac{v'}{v_0} \cdot e^{(mc/Ml)t''} \right) = -2 \frac{mc}{Ml} \left(1 - \frac{2mc}{Ml} t'' \right) \tag{4.14}$$

after using Equation (4.6). Thus

$$\frac{v'}{v_0} \cdot e^{mct''/Ml} = -2 \frac{mc}{Ml} \left(t'' - \frac{mc}{Ml} \cdot t''^2 \right) + \text{constant} \tag{4.15}$$

Now at $t'' = 0$, $v' = v_1$ which is obtained using Equation (1.1) at $t = 2l/e$, i.e. $v' = v_0e^{-2m/M}$. Hence the constant is $e^{-2m/M}$ and Equation (4.15) becomes

$$\begin{aligned} \frac{v'}{v_0} &= -2 \frac{mc}{Ml} \cdot t'' \cdot e^{-mct''/Ml} \left(1 - \frac{mc}{Ml} \cdot t'' \right) \\ &\quad + e^{-m(2+ct''/l)/M} \end{aligned} \tag{4.16}$$

Also

$$\frac{v''_1}{v_0} = e^{-mct''/Ml} \left[2 \left(\frac{mct''}{Ml} \right)^2 - 4 \left(\frac{mct''}{Ml} \right) + 1 + e^{-2(m/M)} \right] \tag{4.17}$$

Further, it may easily be found that the total stress on the striker mass *A* during the time $0 \leq t'' \leq 2l/c$ is

$$\begin{aligned} \frac{T\sigma'_1}{\sigma_0} &= \frac{v'_1}{v_0} + \frac{v''_1}{v_0} \\ &= 2e^{-mct''/Ml} \left[1 - 3 \left(\frac{mct''}{Ml} \right) + \left(\frac{mct''}{Ml} \right)^2 \right. \\ &\quad \left. + \frac{1}{2} e^{-2(m/M)} \right] \end{aligned} \tag{4.18}$$

Table 1 gives some corresponding values of m/M and ct''/l for $T\sigma'_1 = 0$; the permissible range for t'' is of course $0 \leq ct''/l \leq 2$.

TABLE 1

<i>m/M</i>	0.36	0.48	0.50	1.00	2.00	10.00
<i>ct''/l</i>	2.00	1.04	0.94	0.44	0.19	0.04
<i>v'/v₀</i>	0.03	0	-0.08	-0.31	-0.31	-0.32

From Equation (4.18), if $T\sigma'_1 = 0$

$$2 \left(\frac{mct''}{Ml} \right)^2 + e^{-2(m/M)} = 2 \left[3 \left(\frac{mct''}{Ml} \right) - 1 \right]$$

and substituting this into Equation (4.16), it will be found that v'/v is positive if $mct''/Ml > 1/2$. The impact stress may thus fall to zero and at the same time the striker *A* may still be travelling forward.

Second wave cycle for the mass at B

When the stress wave reflected from the mass at A reaches the mass at B, it will be reflected for a time $0 < t''' < 2l/c$; this is the second wave cycle.

Adapting Equation (4.2) (Fig. 4(d)), we have

$$\frac{dv_2'}{dt'''} + \frac{mc}{Ml} v_2' = 2 \frac{mc}{Ml} v_1'' \quad (4.19)$$

Hence

$$\frac{d}{dt'''} (v_2' \cdot e^{mct''/Ml}) = 2 \frac{mc}{Ml} v_0 \left[2 \left(\frac{mct'''}{Ml} \right)^2 - 4 \left(\frac{mct'''}{Ml} \right) + 1 + e^{-2(m/M)} \right] \quad (4.20)$$

and thus,

$$\frac{v_2'}{v_0} \cdot e^{mct''/Ml} = 2 \frac{mc}{Ml} \left[2 \left(\frac{mc}{Ml} \right)^2 \cdot \frac{t'''^3}{3} - 2 \left(\frac{mc}{Ml} \right) \cdot t'''^2 + t''' + t''' e^{-2m/M} \right] + c \quad (4.21)$$

When $t''' = 0$, $v_2' = v_2$ at $t' = 2l/c$, or, using Equation (4.5)

$$\frac{v_2'}{v_0} = 4 \frac{m}{M} e^{-2m/M} = c$$

so that

$$\frac{v_2'}{v_0} = 2 \frac{mc}{Ml} e^{-mct''/Ml} \left[2 \left(\frac{mc}{Ml} \right)^2 \cdot \frac{t'''^3}{3} - 2 \left(\frac{mc}{Ml} \right) t'''^2 + t''' + t''' e^{-2(m/M)} \right] + 4 \frac{m}{M} e^{-2(m/M)} \cdot e^{-mct''/Ml} \quad (4.22)$$

Now, the stress on the mass at B, $T\sigma_2'$ is zero when

$$\frac{T\sigma_2'}{\sigma_0} = \frac{v_1'' + v_1'''}{v_0} = \frac{2v_1''' - v_2'}{v_0} \quad \text{or} \quad v_2' = 2v_1''' \quad (4.23)$$

Hence, using Equations (4.17) and (4.22), we require

$$\frac{2}{3} \left(\frac{mc}{Ml} \right)^3 \cdot t'''^3 - 4 \left(\frac{mc}{Ml} \right)^2 \cdot t'''^2 + \frac{mc}{Ml} (5 + e^{-2(m/M)}) t''' + e^{-2(m/M)} \left(1 + 2 \frac{m}{M} \right) - 1 = 0 \quad (4.24)$$

Putting $mct'''/Ml = \tau$, Equation (4.24) becomes,

$$\frac{2}{3} \tau^3 - 4\tau^2 + \tau(5 + e^{-2(m/M)}) = 1 - e^{-2(m/M)} \left(1 + 2 \frac{m}{M} \right) \quad (4.25)$$

If $ct'''/l = 2$, then $\tau = 2m/M$, and a limiting value of m/M is given by

$$2\tau^3 - 12\tau^2 + 15\tau - 3 = -3e^{-\tau} (1 + 2\tau) \quad (4.26)$$

a solution to which is $\tau \approx 1.6$.

5 THE EARLY MOTION AND STRESS WAVES IN A LONG LINEAR ELASTIC BAR IMPACTED BY A RIGID MASS AT ONE END AND POSSESSING AN ATTACHED RIGID MASS AT THE OTHER

The problem and the analysis of this system are very similar to that treated in the previous section. With the obvious notation of figure 5, we have for the

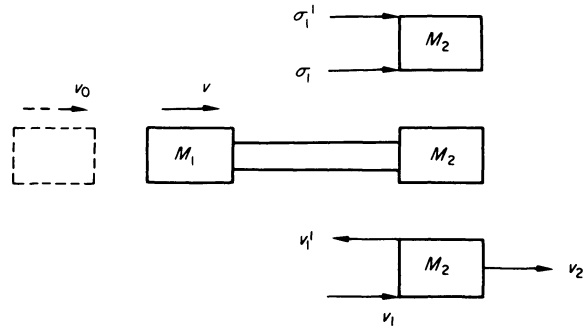


Figure 5.

equation of motion for mass M_2 , at time t from the moment of arrival of the stress wave due to the initial impact of M_1 on the bar,

$$M_2 \frac{dv_2}{dt} = A(\sigma_1 + \sigma_1')$$

where σ_1 is the intensity of the incident compressive wave and σ_1' that of the reflected wave.

$$\text{i.e.} \quad \frac{dv_2}{dt} = \frac{mc}{M_2 l} (v_1 + v_1') \quad (5.1)$$

But $v_1' = v_1 - v_2$, so that Equation (5.1) becomes

$$\frac{dv_2}{dt} = \frac{mc}{M_2 l} (2v_1 - v_2)$$

Thus

$$\frac{d}{dt} (v_2 e^{(mc/M_2 l)t}) = 2v_0 \frac{mc}{M_2 l} \cdot e^{-(mct/l)(1/M_1 - 1/M_2)}$$

so that,

$$\frac{v_2}{v_0} \cdot e^{(mc/M_2 l)t} = 2 \cdot \frac{e^{-(mct/l)(1/M_1 - 1/M_2)}}{(1 - M_2/M_1)} + c$$

At $t = 0$, $v_2 = 0$ and $c = 2/(M_2/M_1 - 1)$.

Thus

$$\frac{v_2}{v_0} = \frac{2}{(M_2/M_1 - 1)} \cdot [e^{-(m/M_2)(ct/l)} - e^{-(mct/M_1 l)}] \quad (5.2)$$

Also

$$\frac{v_1'}{v_0} = \frac{1}{(1 - M_2/M_1)} \cdot \left[2e^{-(m/M_2)(ct/l)} - \left(1 + \frac{M_2}{M_1} \right) e^{-(m/M_1)(ct/l)} \right] \quad (5.3)$$

TABLE 2

μ_2/μ_1	1	1.2	1.5	2	3	4	5	6	10	20	100
ct/l	1	0.91	0.81	0.69	0.55	0.46	0.40	0.36	0.255	0.158	0.046
v_2/v_0	0.74	0.80	0.89	1.00	1.15	1.27	1.37	1.43	1.55	1.72	1.91

The total stress acting on M_2 is

$$\begin{aligned}
 T\sigma_2 &= \frac{v_1}{v_0} + \frac{v_1'}{v_0} \\
 &= e^{-mct/M_1 l} + \frac{2e^{-mct/M_2 l}}{1 - M_2/M_1} - \frac{1 + M_2/M_1}{1 - M_2/M_1} \cdot e^{-(m/M_1)(ct/l)} \\
 &= \frac{2}{1 - M_2/M_1} \cdot \left[e^{-mct/M_2 l} - \frac{M_2}{M_1} \cdot e^{-mct/M_1 l} \right] \quad (5.4)
 \end{aligned}$$

Now $T\sigma_2/\sigma_0$ becomes zero when

$$\frac{M_1}{M_2} = e^{-(mct/l)(1/M_1 - 1/M_2)} \quad (5.5)$$

or

$$\frac{ct}{l} = \frac{\ln M_1/M_2}{m/M_2 - m/M_1} = \frac{\ln \mu_2/\mu_1}{\mu_2 - \mu_1}, \quad (5.6)$$

putting $\mu_1 = m/M_1$ and $\mu_2 = m/M_2$.

Also when $T\sigma_2 = 0$, by substituting from Equation (5.6) into Equation (5.2),

$$\begin{aligned}
 \frac{v_2}{v_0} &= \frac{2}{1 - M_2/M_1} \cdot [e^{-(m/M_1)(\ln M_1/M_2)/m(M_1 - M_2)/(M_1 M_2)} \\
 &\quad - e^{-(m/M_1)(\ln M_1/M_2)/m(M_1 - M_2)/(M_1 M_2)}] \\
 &= 2 \left(\frac{\mu_1}{\mu_2} \right)^{(\mu_1/\mu_2)/(1 - \mu_1/\mu_2)} \quad (5.7)
 \end{aligned}$$

When $M_1 = M_2$ or $\mu_1 = \mu_2$, then $v_2/v_0 = 2e^{-1}$. Some typical results are given in Table 2.

6 SOME RESULTS FOR ELASTIC WAVE PROPAGATION ALONG A CHAIN COMPOSED OF IDENTICAL LINEAR ELASTIC BARS OF MASS m , BETWEEN IDENTICAL RIGID MASSES, M , AND WHEN IMPACTED ON THE PROXIMAL END BY ONE MASS, M , MOVING AT SPEED v_0

The situation to be analysed is shown in figure 6(a). Mass M impinges with speed v_0 on one end of a bar

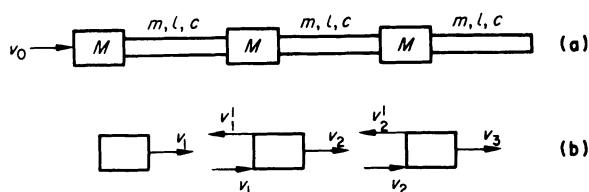


Figure 6.

and a wave of stress intensity σ_1 is propagated through it. At time t after impact first occurs, as before, we have

$$\sigma_1 = \sigma_0 e^{-(mc/Ml)t} \quad \text{and} \quad v_1 = v_0 e^{-(mc/Ml)t} \quad (2.4)$$

where σ_1 is the stress between the first mass and the bar and v_1 the speed of the first mass. σ_0 is the stress initiated at the first moment of impact.

At time l/c after first impact, the initiated wave is both reflected from, and transmitted by, the second mass, the equation of motion for which (figure 6(b)) is

$$M \frac{dv_2}{dt} = \frac{mc}{l} (v_1 + v_1' - v_2) \quad (6.1)$$

But

$$v_1 - v_1' = v_2$$

or

$$v_1' = v_1 - v_2$$

so that

$$M \frac{dv_2}{dt} = 2 \frac{mc}{l} (v_1 - v_2) \quad (6.2)$$

Rearranging Equation (6.2) and introducing Equation (2.4) we have

$$\frac{dv_2}{dt} + 2 \frac{mc}{Ml} v_2 = 2 \frac{mc}{Ml} v_0 e^{-mct/Ml} \quad (6.3)$$

Integrating Equation (6.3) but noting that $v_2 = 0$ when $t = 0$, where t is measured from the moment at which σ_1 reaches the second mass, it is found that

$$\frac{v_2}{v_0} = 2 \left[e^{-mct/Ml} - e^{-2(mct/Ml)} \right] \quad (6.4)$$

Similarly for the third mass

$$\frac{v_3}{v_0} = 4 \left[e^{-mct/Ml} - e^{-2(mct/Ml)} \left(1 + \frac{mct}{Ml} \right) \right] \quad (6.5)$$

And

$$\frac{v_4}{v_0} = 8 \left[e^{-mct/Ml} - e^{-2(mct/Ml)} \left(1 + \frac{mct}{Ml} + \left(\frac{mct}{Ml} \right)^2 \frac{1}{2} \right) \right] \quad (6.6)$$

Evidently

$$\begin{aligned}
 \frac{v_n}{v_0} &= 2^{n-1} \left[e^{-mct/Ml} - e^{-2(mct/Ml)} \right. \\
 &\quad \times \left(1 + \frac{mct}{Ml} + \frac{1}{2} \cdot \left(\frac{mct}{Ml} \right)^2 + \dots \right. \\
 &\quad \left. \left. + \frac{1}{(n-2)!} \cdot \left(\frac{mct}{Ml} \right)^{n-2} \right) \right] \quad (6.7)
 \end{aligned}$$

Now, the expression for the total stress on the second mass for the period l/c after it is first actuated by a stress wave is given by

$$\begin{aligned} T\sigma_2 &= \frac{v_1}{v_0} + \frac{v_1'}{v_0} - \frac{v_2}{v_0} = \frac{2(v_1 - v_2)}{v_0} \\ &= 2[e^{-mct/Ml} - 2(e^{-mct/Ml} - e^{-2(mct/Ml)})] \\ &= 2[2e^{-2(mct/Ml)} - e^{-mct/Ml}] \end{aligned} \quad (6.8)$$

Similarly

$$\begin{aligned} \frac{T\sigma_3}{\sigma_0} &= \frac{2(v_2 - v_3)}{v_0} \\ &= 2^2 \left[(e^{-mct/Ml} - e^{-2(mct/Ml)}) \right. \\ &\quad \left. - 2 \left\{ e^{-mct/Ml} - e^{-2(mct/Ml)} \left(1 + \frac{mct}{Ml} \right) \right\} \right] \\ &= 2^2 \left[2e^{-2(mct/Ml)} \left(1 + \frac{mct}{Ml} \right) \right. \\ &\quad \left. - (e^{-mct/Ml} + e^{-2(mct/Ml)}) \right] \end{aligned} \quad (6.9)$$

Generally

$$\begin{aligned} \frac{T\sigma_N}{\sigma_0} &= 2^{n-1} \left[\left(\frac{x^{n-2} \cdot e^{-2x}}{n-2} - e^{-x} \right) \right. \\ &\quad \left. + \left(1 + x + \dots + \frac{x^{n-2}}{n-2} \right) e^{-2x} \right] \end{aligned} \quad (6.10)$$

where $x = mct/Ml$.

The total stress on a given mass immediately a stress wave arrives at the next mass when $m/M = 1$ is

$$\begin{aligned} \text{Mass 1: } \frac{T\sigma_1}{\sigma_0} &= e^{-1} = 0.368 \\ \text{Mass 2: } \frac{T\sigma_2}{\sigma_0} &= 2(2e^{-2} - e^{-1}) = -0.194 \\ \text{Mass 3: } \frac{T\sigma_3}{\sigma_0} &= 4(3e^{-2} - e^{-1}) = 0.152 \\ \text{Mass 4: } \frac{T\sigma_4}{\sigma_0} &= 8(3e^{-2} - e^{-1}) = 0.304 \\ \text{Mass 5: } \frac{T\sigma_5}{\sigma_0} &= 16\left(\frac{11}{6}e^{-2} - e^{-1}\right) = 0.256 \\ \text{Mass 6: } \frac{T\sigma_6}{\sigma_0} &= 32\left(\frac{11}{4}e^{-2} - e^{-1}\right) = 0.128 \end{aligned}$$

7 ELASTIC WAVE PROPAGATION ALONG A CHAIN OF UNEQUAL LINEAR ELASTIC BARS BETWEEN UNEQUAL RIGID MASSES IMPACTED ON THE PROXIMAL END BY A GIVEN RIGID MASS

The analysis is entirely similar to that carried out immediately above. In place of Equation (2.4) we have (figure 7)

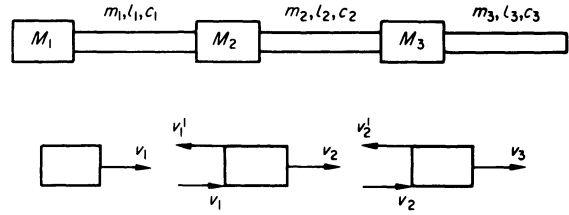


Figure 7.

$$\sigma_1 = \sigma_0 e^{-(m_1 c_1 / M_1 l_1) t} \quad \text{and} \quad v_1 = v_0 e^{-(m_1 c_1 / M_1 l_1) t} \quad (7.1)$$

and instead of Equation (6.1) we now have

$$M_2 \frac{dv_2}{dt} = \frac{m_1 c_1}{l_1} (v_1 + v_1') - \frac{m_2 c_2}{l_2} v_2$$

so that substituting $v_1' = v_1 - v_2$

$$\frac{dv_2}{dt} + 2 \frac{m_2 c_2}{M_2 l_2} v_2 = 2 \frac{m_1 c_1}{M_2 l_1} v_1 \quad (7.2)$$

Solution then proceeds as before after substituting for v_1 from Equation (7.1).

8 ELASTIC WAVE PROPAGATION, TRANSMISSION AND REFLECTION IN A LONG BAR SUBJECT TO IMPACT BY A RIGID MASS AT ONE END WHEN THE DISTAL END IS ONLY RESTRAINED IN PART

In the classical problem studied by St. Venant and others, the distal end of the impacted bar is perfectly fixed; the case of a bar which has a free end has also been studied, and the continuous bar which constitutes perfect transmissibility lies between these two cases. We refer here to cases of partial end restraint, or end restraint which is less than perfect, i.e. to the circumstance in which there is some degree of stress wave transmission beyond, or into, the region of fixity and some degree of reflection.

Consider a striking mass impinging on to the end of a long bar in which there is a simultaneous change of cross-sectional area and mechanical impedance at XX (figure 8(a)). It is easily shown, from considerations of force equilibrium and identity of speed in the plane at XX, that the transmitted stress σ_T and the reflected stress σ_R , in terms of the current incident stress σ_1 at XX, is

$$\left. \begin{aligned} \frac{\sigma_T}{\sigma_1} &= \frac{2A_1 \rho_1 c_1}{A_1 \rho_1 c_1 + A_2 \rho_2 c_2} = 2 \frac{A_1}{A_2} \cdot \frac{(i_2'/i_1')}{(i_2'/i_1') + 1} \\ &= 2 \frac{A_1}{A_2} \cdot \frac{x}{x + 1} = K_T \\ \text{and} \\ \frac{\sigma_R}{\sigma_1} &= \frac{A_2 \rho_2 c_2 - A_1 \rho_1 c_1}{A_2 \rho_2 c_2 + A_1 \rho_1 c_1} \\ &= \frac{(i_2'/i_1') - 1}{(i_2'/i_1') + 1} = \frac{x - 1}{x + 1} = K_R \end{aligned} \right\} \quad (8.1)$$

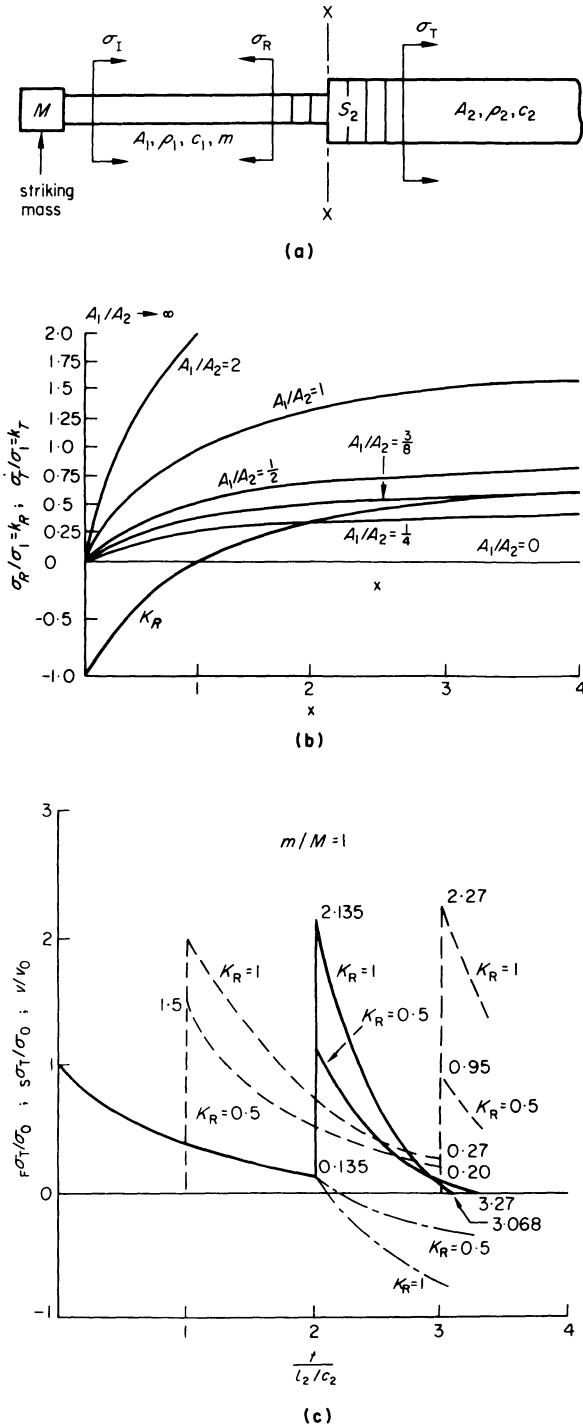


Figure 8. $F\sigma_T$ ----- stress at fixed end
 $S\sigma_T$ ————— stress at struck end
 v/v_0 ----- velocity at struck end

where $i'_1 = A_1\rho_1c_1$ and $i'_2 = A_2\rho_2c_2$ so that x is the ratio of these impedances.

For x between 0 and 4, figure 8(b) shows values of σ_R/σ_1 or K_R , the reflection coefficient, and values of σ_T/σ_1 or K_T , the transmission coefficient, for different values of the area ratio A_1/A_2 . Note that $K_R = -1$ implies a free-end bar, and $K_R = +1$ implies that it is firmly fixed. When $x = 1$ and $K_R = 0$, there is perfect transmission, and when $K_R = K_T$, i.e. $\sigma_R = \sigma_T$, then $A_1/A_2 = (x - 1)/2x$ and, if $\rho_1c_1 = \rho_2c_2$, then $A_2 = A_1/2$.

Multiple reflexions in a bar partially restricted at one end

If a rigid striking mass, as shown in figure 8(a), impinges on the end of a bar (the distal end of which is partially restrained in the manner shown), then the coefficients defined above may be used to study the consequences for elastic wave propagation.

The reflected wave stress intensity, σ_R , is given by

$$\sigma_R = K_R \sigma_0 \exp\left(-\frac{m}{M} \cdot \frac{t}{l/c}\right) \quad (8.2)$$

where time t is measured from the first moment of arrival.

The total stress at the restrained end is

$$(1 + K_R)\sigma_0 \exp\left(-\frac{m}{M} \cdot \frac{t}{l/c}\right)$$

and at the first reflection, i.e. for $ct/l = 0$, it attains a value of $(1 + K_R)$.

The expression for the speed ratio for particles in the plane of change of section XX is

$$v_T/v_1 + v_R/v_1 = 1$$

or

$$\frac{\sigma_T/\sigma_1}{\rho_2c_2/\rho_1c_1} + \frac{\sigma_R}{\sigma_1} = 1 \quad (8.3)$$

thus

$$\frac{K_T}{(A_1A_2)x} - K_R = 1 \quad (8.4)$$

Note that a negative sign appears before K_R since the speed associated with σ_R is in the negative direction.

For example, with $m/M = 1$ and by taking $K_R = 0.5$, we have for the total stress at the ends of the bar, the values shown in Table 3 (when t is measured from the moment of impact).

TABLE 3

	$\frac{ct}{l}$	0	1	2
Stress/ σ_0	Impacted End	1	0.37	0.20
	Distal End	1.5	0.55	0.30

To determine some of the consequences of introducing K_R for the reflected wave in the second cycle, i.e. for

$$2l/c \leq t \leq 4l/c$$

or

$$0 \leq t_2 \leq 2l/c$$

we easily find, after setting up the equation of motion for M and integrating (figure 8(c)), that

$$\frac{Ml}{mc} \sigma_2 \exp\left(\frac{m}{M} \cdot \frac{t_2}{l/c}\right) = -2K_R \sigma_0 t_2 + c \quad (8.5)$$

At $ct_2/l = 0$, or at the moment of commencement of the second cycle, the striker speed is $v_0 \exp(-2m/M)$. Hence the particle speed in the reflected second wave cycle is then

$$v'_2 = v_0 [-K_R + \exp(-2m/M)] \quad (8.6)$$

and the stress is

$$\sigma'_2 = \sigma_0 [K_R + \exp(-2m/M)] \quad (8.7)$$

so that in Equation (8.5)

$$c = \frac{M}{m} \cdot \frac{l}{c} \sigma_0 [K_R + \exp(-2m/M)]$$

Thus the stress intensity in the reflected second wave cycle is

$$\sigma'_2 = \sigma_0 \exp\left(-\frac{m}{M} \cdot \frac{t_2}{l/c}\right) \left[K_R + \exp(-2m/M) - 2K_R \frac{m}{M} \cdot \frac{t_2}{l/c} \right] \quad (8.8)$$

The total stress on the striker, then, for $2l/c \leq t \leq 4l/c$ or $0 \leq t_2 \leq 2l/c$, is

$$\tau \sigma_S = \sigma_0 \exp\left(-\frac{m}{M} \cdot \frac{t_2}{l/c}\right) \left[2K_R + \exp(-2m/M) - 2K_R \frac{m}{M} \cdot \frac{t_2}{l/c} \right] \quad (8.9)$$

With $K_R = 0.5$ for $m/M = 1$, then for ct_2/l equal to 0 and 1, $\tau \sigma_S = 1.35 \sigma_0$ and $0.05 \sigma_0$ respectively.

If from Equation (8.9) $s \sigma_T = 0$, then

$$\frac{t_2}{l/c} = \frac{2 + \frac{1}{K_R} \cdot \exp(-2m/M)}{2m/M} \quad (8.10)$$

and, with the same constants as before, $ct_2/l = 1 + 2e^{-2} \approx 1.27$ or $ct/l \approx 3.27$. This result is expected since, if $K_R < 1$ (and for $K_R = 1$ $ct_2/l \approx 1.135$), the reflected stress will be less than

what it would be for $K_R = 1$, and hence, when it encounters the striking mass, it will take a longer time to arrest it.

If $s \sigma_T = 0$ at the end of the second wave cycle, i.e. when $ct_2/l = 2$, then the required value of $K_R = e^{-2}/2 \approx 0.0675$ and $x \approx 0.86$. The striker forward speed in the second wave cycle is

$$v_2 = v_0 \exp\left(-\frac{m}{M} \cdot \frac{t_2}{l/c}\right) \left[\exp\left(-\frac{2m}{M}\right) - 2K_R \frac{m}{M} \cdot \frac{t_2}{l/c} \right] \quad (8.11)$$

and this is zero when

$$\frac{t_2}{l/c} = \frac{1}{2K_R} \cdot \frac{M}{m} \cdot \exp\left(-\frac{2m}{M}\right) \quad (8.12)$$

so that for $K_R = 0.5$, this speed is zero when $ct_2/l = 0.135$. This can happen in the second wave cycle only when

$$\frac{M}{m} \exp\left(-\frac{2m}{M}\right) \leq 4K_R \quad (8.13)$$

and for $K_R = 0.5$, $M/m \leq 3.54$. Complete arrest occurs in this cycle when the end fixity coefficient is less than unity, e.g. 0.5. This is obviously only possible when the mass ratios involved are less than those which apply for a coefficient of unity.

When $s \sigma_T = 0$, the striker speed is

$$-2K_R v_0 \exp\left[1 - \frac{\exp(-2m/M)}{2K_R}\right]$$

and for $K_R = 0.5$, this is $-0.32 v_0$.

The stress at the partially fixed end, at $ct/l = 3$ or $ct_2/l = 1$, is

$$(\sigma_2 + K_R \sigma_2)$$

or

$$\sigma_0(1 + K_R)(K_R + \exp(-2m/M))$$

When $K_R = 0.5$, this stress is $0.9525 \sigma_0$.

Many of the above results for $K_R = 0.5$ have been included in figure 8(d) where they are compared with those for perfect end fixity or $K_R = 1$.

SELF-EXCITED VIBRATION OF CIRCULAR SAWS CUTTING ALUMINIUM

by

D. S. DUGDALE*

SUMMARY

A theoretical discussion is given of how lateral vibration is excited in a saw during cutting. Waves travel in the rim of a saw in the direction of rotation of the saw, and also in the reverse direction. There is necessarily an integral number of complete waves in the whole circumference of the saw. In any given saw, each train of waves travels at a speed, relative to the saw, which is dictated by its wavelength. From these speeds and the number of teeth in the saw, it is possible to deduce the rotational speeds for resonance.

The surface profile, set up by vibrating teeth on the side wall of the saw-cut, is described, and the force due to the side-cutting action of each tooth is shown to supply positive excitation. Experimental methods for cutting tests are described, and results give support to the theory which is therefore proposed as a guide for saw design.

NOTATION

D	saw diameter
N	saw speed, revolutions per sec.
Q	number of saw vibrations per saw revolution
S	total feed rate of saw
V	peripheral speed of saw
f_n	natural frequency of saw in n th mode
n	integral number of waves in rim of saw
p	axial component of specific cutting pressure
t	time
u_0	amplitude of lateral displacement
x	tangential distance of tooth travel
α	fractional advance of tooth path wavelength
η_0	side clearance angle of tooth
λ	wavelength of tooth path or surface profile

INTRODUCTION

When a circular saw vibrates during cutting, the width of cut produced, or kerf, is larger than it needs to be. If the vibration amplitude varies as the cut proceeds, the sawn surfaces will not remain flat and parallel, and often, but not always, the surface finish is inferior. Stresses in the saw are increased, but noise emitted may or may not increase with vibration amplitude. Generally, one would wish a saw to remain perfectly flat and free from vibration.

Attention is restricted to vibration of the saw blade itself, and problems of noise generated by the machine frame, or work-piece, due to out-of-balance force or tooth impact are disregarded. Noise may occasionally be generated by axial vibration of the

spindle and amplified by the saw. A high-pitched noise is sometimes emitted by a carbide-tipped saw while it runs freely without cutting. This has been attributed to a high-frequency vibration in the rim of the saw induced by air flow around the teeth^{1,2}. It is most in evidence with saws having fine-pitched teeth. The vibration is usually of the stationary wave, fixed-node type, and the vibration mode is not directly linked with the precise number of teeth in the saw.

Vibration sometimes occurs when sawing a section which is thin relative to the tooth pitch, and a transient vibration sometimes occurs as a saw enters the corner of a rectangular section. This type of vibration is not discussed here and is the subject of further research. The present report is concerned with the steady-state vibration observed at certain saw speeds while a long cut is being made in a deep section, although the section need not be of uniform depth. This work was stimulated by problems in sawing large aluminium ingots, but the findings may be relevant to high-speed circular saws of all kinds.

An obvious impact force is exerted by the saw as each tooth enters the cut. It is easy to see that if the frequency of tooth impact coincides with one of the natural frequencies of the machine frame, a resonance may occur. When extending this idea to the saw blade, it must be remembered that the speed of a train of waves in the saw rim, relative to a fixed work-piece, is equal to the wave speed in a stationary saw compounded with the rotational speed of the saw.

Leaving aside frequency relationships, which are

* Department of Mechanical Engineering, University of Sheffield, England

fairly simple, it is not immediately obvious that tooth impact forces can excite lateral vibration. First, it is noted that tangential cutting forces, acting on un-bevelled teeth, have no first-order component, acting axially, which might supply energy to a vibration. Further, it is a matter of observation that when some object is pressed in an axial direction against the teeth of a rotating saw, no vibration is set up. At any instant, those points on the rim of a vibrating saw which happen to be in a prominent position, are those which instantaneously have maximum lateral displacement and therefore zero lateral velocity. It follows that lateral impacts, imposed on these prominent points by a fixed obstacle, are not able to supply energy to an initially existing vibration. Hence it will be necessary to explain how a special kind of prominence on the side wall of the cut can be generated, and sustained, by the cutting action of vibrating teeth, and how it can then exert an inward force on each tooth during the time when the tooth is returning from its point of maximum lateral displacement.

CONDITIONS FOR RESONANCE

Suppose a train of n complete waves travels round the rim of a disc. With the disc stationary, let the number of waves per second passing a fixed point be f_n . This is the same as the natural frequency of the disc vibrating with a stationary wave system having n nodal diameters position-fixed relative to the disc. Now suppose that the train of waves travels in the direction of rotation of the disc. Its speed relative to the disc remains unchanged, so that one complete revolution of the disc will allow n extra waves to pass a point fixed in space. If the waves travel in the direction opposite to that of rotation, n fewer waves will pass. Hence, if the rotational speed is N revs/s, the frequency with which travelling waves pass a fixed point is given by $f_n \pm nN$, the upper sign being taken when the waves are forward-travelling.

It is useful to introduce the number of vibrations Q which occur in the saw per complete revolution of the saw. Noting that $f_n = QN$, the time between the

passing of successive waves in a rotating saw becomes $1/N(Q \pm n)$. Further, it is clear that successive teeth pass a fixed point in time interval $1/NT$, where T is the total number of teeth in the saw. A resonance condition is associated with a regularly occurring conjunction of teeth with wave peaks. Figure 1 shows the wave peaks as arrows drawn from the saw centre. After an initial conjunction of wave 0 with tooth 0 on the horizontal datum line, the simplest sequence is that each tooth should subsequently meet each wave on this line. By equating the time intervals given above, the number for resonance Q is found to be

$$Q = T \mp n \tag{1}$$

This kind of interaction may be labelled *primary*. A *secondary* kind occurs when the waves pass by at twice the rate of the teeth, when we have $Q = 2T \mp n$. Here, each tooth continues to trace out the same path. Another kind of interaction between teeth and waves occurs when teeth pass the datum point at twice the frequency at which the waves pass. As illustrated in figure 1, this leads to two tooth paths, each followed by alternate teeth and a half-wave profile is generated on the side wall of the saw-cut. However, this last type of interaction was not observed in any steady vibration, as mentioned later. In general, the wavelength λ , of the profile generated on the work-piece, is given by $\lambda = \pi D/Q$, where D is saw diameter.

EXCITATION STRENGTH

Firstly, the profile of the side of the saw-cut, formed by oscillating teeth, must be defined. Each tooth advances in the y -direction by amount s , this being the feed per tooth. At the same time, the tooth path of wavelength λ is assumed to advance a small distance $a\lambda$ in the x -direction, this being the direction of cutting. Figure 2 is a view of the x - y plane; that is,

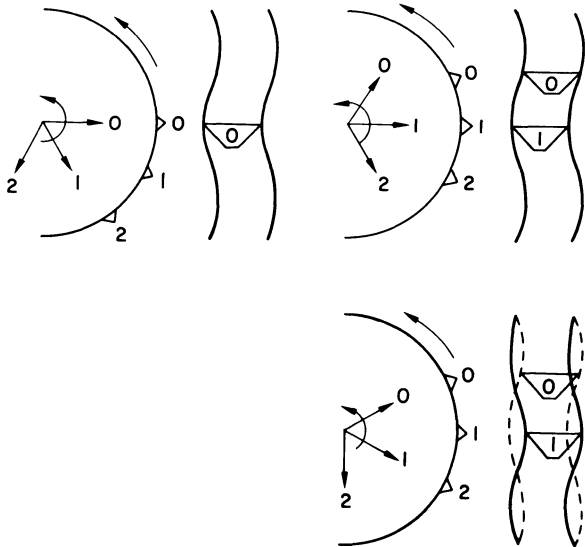


Figure 1. Conjunctions between teeth and travelling waves.

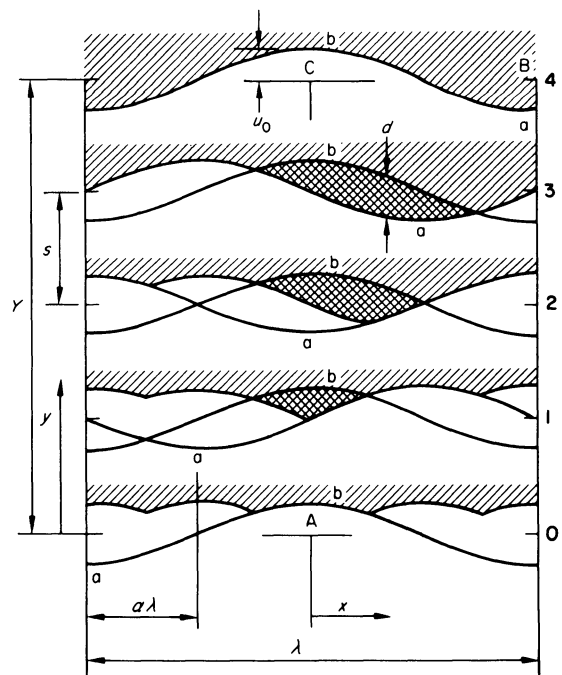


Figure 2. Steady-state geometry of chip produced by oscillating teeth.

the plane of the side of the saw-cut, and successive perpendicular sections are indicated, each stage representing the passing of one tooth. An excessively large value of $\alpha = 1/4$ is taken so as to limit the number of stages during which the tooth path shifts through a whole wavelength, and the current path at each stage is marked (a).

At stage 1 the relatively long radial cutting edge extends radially inward over stages 0 and 1 to remove all material below the current tooth path. Similarly at stages 2 and 3 further material is removed from all the earlier stages, so that by the time the tooth passage of stage 3 is completed, the only remaining material is that shown with single shading. Over the surface of the single strip in final stage 4, all material is removed by the axial cutting edge of the tooth. However, material is simultaneously removed from all the preceding stages by the side-cutting action of the radial cutting edge. Therefore, the final tooth path marked (b) is superimposed on all the previous stages to reveal the material removed by the passage of this tooth. This material is shown by double shading, and represents the steady-state shape of the chip removed by the side-cutting action of each and every tooth. The prominence left on the side wall of the cut extends radially inward a distance $Y = s/\alpha$. Since contact ceases along the straight line AB, the side cutting takes place over triangle ABC. When a tooth is at a position defined by the x co-ordinate, the width in the y -direction of the chip removed is given by

$$Y - y = (s/\alpha)(1 - 2x/\lambda)$$

If advance α is small, the depth of cut d is uniform in the y -direction and is given by

$$\begin{aligned} d &= u_0 \cos \frac{2\pi x}{\lambda} - u_0 \cos 2\pi \left(\frac{x}{\lambda} + \alpha \right) \\ &= 2\pi\alpha u_0 \sin \frac{2\pi x}{\lambda} \end{aligned}$$

This, multiplied by the width, gives the cross-sectional area A of the uncut chip,

$$A = 2\pi s u_0 \left(1 - \frac{2x}{\lambda} \right) \sin \frac{2\pi x}{\lambda}$$

which is independent of advance α . Now, let the specific cutting pressure for the work-piece material have a component p acting normal to the cut surface, tending to repel the tooth away from this surface. The force F acting on the tooth in a direction which is outwards, relative to the saw, is therefore $F = -pA$. Outward tooth displacement is $u = u_0 \cos 2\pi x/\lambda$. Hence energy U supplied to the lateral motion of the saw during the interaction of one tooth with one surface wave is

$$U = \int F du = \frac{1}{2} \pi^2 p s u_0^2 \int_0^{\lambda/2} \frac{8}{\lambda} \left(1 - \frac{2x}{\lambda} \right) \sin^2 \frac{2\pi x}{\lambda} dx$$

The value of this last integral is unity.

It is further noted that each tooth, in passing through a slab of depth h will encounter h/λ surface waves, and NT teeth pass per second. Also, the cut

has two sides, leading to an expression for the rate W at which energy is supplied to a saw vibration

$$W = \pi p h S Q u_0^2 / D \quad (2)$$

where S is the total feed rate for the saw. This applies to *primary* and *secondary* types of tooth interaction, the latter having higher Q -values. However, with the half-wave profile of the cut surface (figure 1), much of the excitation is lost through the tops of the waves being absent, and the above value for W should be multiplied by a calculated factor 0.3. This may explain why continuous vibration of this type was not observed in tests.

In order that the side-cutting force can supply positive excitation, a positive advance α is needed. This means that the saw speed must be slightly higher than resonance speed $N = f_n/Q$, whether the wave motion is in the forward or reverse direction. Experimental results indicated strong excitation for values of α not exceeding 0.07, so this value was used for marking upper limits to resonance ranges in figure 6.

LIMITING AMPLITUDE

Figure 3 shows a surface profile set up at a resonance speed. A tooth moves at speed V in direction x , tangential to the saw. The height of the surface above

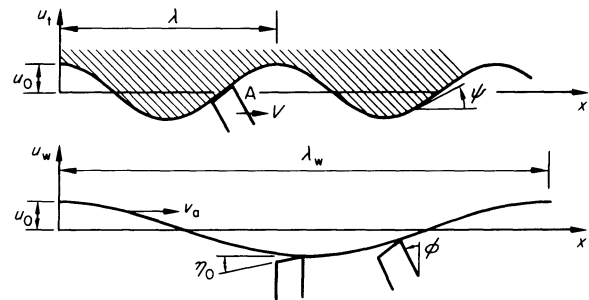


Figure 3. Surface profile of workpiece and wave in rim of saw.

datum is given by $u_s = u_0 \cos 2\pi x/\lambda$. Due to the passage of a wave of absolute velocity v_a , the tooth receives a lateral displacement $u_w = u_0 \cos 2\pi (x\lambda_w v_a t)/\lambda_w$, where wavelength $\lambda_w = \pi D/n$. It may be verified that $u_s = u_w$ for any x and t . Gradient ψ of the surface profile is given by

$$\psi = \frac{\partial u_s}{\partial x} = -\frac{2\pi u_0}{\lambda} \sin \frac{2\pi x}{\lambda}$$

Rotation of the tooth is given by

$$\phi = \frac{\partial u_w}{\partial x} = -\frac{2\pi u_0}{\lambda_w} \sin 2\pi \frac{(x - v_a t)}{\lambda_w}$$

It is noted that $v_a = \pm v_r + V$, where v_r is the wave velocity relative to the saw disc, the upper sign for a forward-moving wave. A resonant condition implies a frequency relation $v_r/\lambda_w = V/\lambda$, which leads to the revised expression for rotation

$$\phi = \pm \frac{2\pi u_0}{\lambda_w} \sin \frac{2\pi x}{\lambda}$$

Suppose that the tooth is initially ground so that it has side clearance angle η_0 , measured in a plane normal to a radial line. At any instant, the effective side clearance angle η will depend on tooth rotation and inclination of cut surface, such that $\eta = \eta_0 - \psi + \phi$. Writing $\lambda = \pi D/Q$ and $\lambda_w = \pi D/n$, this becomes

$$\eta = \eta_0 + 2(Q \pm n)(u_0/D) \sin 2\pi x/\lambda$$

This value is a minimum at position A in figure 3, where $\sin 2\pi x/\lambda = -1$. It is supposed that the amplitude of saw vibration is limited when the side of the tooth rubs on the side of the saw-cut. This occurs when side clearance angle η is reduced to zero. For an interaction of teeth and waves of the primary type, Equation (1) may be used to obtain the maximum amount of kerf widening

$$2u_0 = \eta_0 D/T \quad (3)$$

For example, with $\eta_0 = 3^\circ$ (0.052 radians), $D = 800$ mm, $T = 24$, we have $2u_0 = 1.73$ mm. For the secondary type of tooth interaction, half this value is taken.

EXPERIMENTAL METHODS

The saw was mounted on the spindle of a horizontal milling machine fitted with a variable-speed d.c. drive. To provide sufficient cutting time for vibrations to develop in the saw, slices were sawn off slabs of wide cross-section, as shown in figure 4. These slabs were

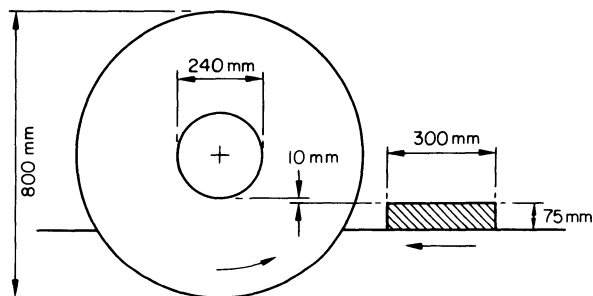


Figure 4. Relative position of saw and workpiece.

of chill-cast aluminium with 3 per cent copper, 5 per cent silicon. Slabs of double depth, slabs of magnesium alloy and bars of extruded aluminium alloy were additionally used without finding any obvious difference in results. Also, full-depth cuts were made with the slab passing beneath the saw, and resonance speeds were found to be unchanged. This indicated that the exact vertical distance of the slab beneath the saw centre was unimportant.

When cuts were made in a slab of pure aluminium, and also in a casting of aluminium with 5 per cent magnesium which had an extreme degree of porosity, no vibration was set up due to chips re-entering the saw-cut and adhering to the walls of the cut so as to rub on the saw body. Further, when various hardwoods were sawn, no vibration was observed, presumably because cutting forces were too low. Feed rate was increased with increasing rotational speed,

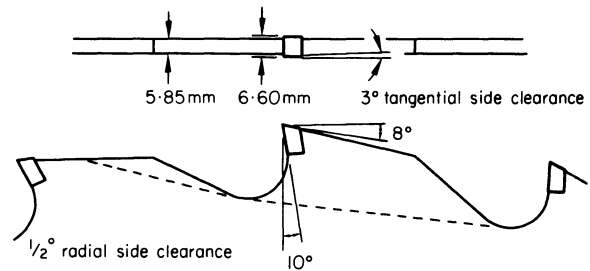


Figure 5. Saw tooth geometry.

and ranged from 250 to 750 mm/min. This approaches the values adopted in practice for this type of saw.

The saw had carbide-tipped teeth of the form shown in figure 5. It should be noted that the tangential side clearance angle of 3° adopted here is about twice that commonly employed in saws used on non-ferrous metals. The saw originally had 24 teeth. On completion of tests with the saw in this condition, alternate teeth were cut away to give a 12-tooth saw, as indicated in figure 5. Such a small number of teeth is not used in practice. The steel clamping collars used were of 25 mm thickness and of various diameters. As the diameter most used was 230 mm, the results obtained with these collars were selected for presentation. Static vibration modes were investigated by clamping the saw horizontally on a heavy iron base. Although modes having one or more nodal circles can be excited in this way, they appear to have no practical significance. Natural frequencies used for calculations were obtained by applying a vibration generator to the saw while it was mounted on the machine spindle. Values of these frequencies f_n for various numbers of nodal diameters n are shown in Table 1.

TABLE 1

n	Natural frequencies (Hz)	
	24-tooth saw	12-tooth saw
1	47.1 to 48.8	48.2 to 50.3
2	65.0 to 67.1	68.5 to 71.2
3	119.8 to 120.2	127.2 to 127.9
4	202.8 to 203.0	217.0 to 217.5
5	307	330
6	429	441

Natural frequencies rise with rotational speed. This increase may amount to about 3 per cent at a speed of 10 revs/s, and depends on collar diameter and vibration mode. Although this increase can be both calculated and detected, it is not large enough to affect the essential arguments put forward here, and has not been taken in the calculations presented. Spindle speeds were obtained from an electronic counter receiving impulses over a 10 second period from a cam-operated switch.

As previously argued, the frequency of waves passing beneath a probe fixed in space is given by $f_n \pm nN$. At a selected saw speed N , the probe frequencies can be calculated for each n -value and for forward-travelling and reverse-travelling waves. When these values are superimposed on the experimentally

obtained spectrum of probe frequencies, the types of wave present in the saw can be identified.

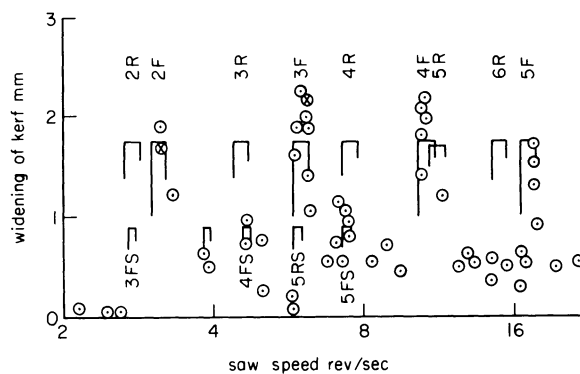
Two methods for obtaining probe frequencies were used. A flat-ended microphone was placed at a radius a little less than that of the tooth gullets, and within 5 mm from the surface of the saw. A frequency analysis was then carried out on a tape-recording of the microphone output. Although convenient, this method suffers from the disadvantage that sound pressure depends jointly on vibration amplitude and frequency, so relative heights of voltage peaks do not directly indicate relative amplitudes of vibration at different frequencies³. Another method was to place an insulated metal capacitor disc of 50 mm diameter adjacent to the rim of the saw, with an air gap of 3 mm. On applying a voltage of frequency 25 kHz to this disc, variations in capacitance appeared as voltage variations across a bridge circuit, which were amplified and recorded. In this recording, heights of voltage peaks truly indicated relative amplitudes of vibration.

It may happen that the probe frequencies for two waves, e.g. the $n = 3$ (forward) and $n = 4$ (reverse) waves, lie close together, so that the observed probe frequency is not, by itself, sufficient to indicate which of these two waves is present in the saw. This question was decided as follows. At the end of a saw cut, the feed motion was disengaged and the machine motor was switched off, so that the saw slowed down without further contact with the workpiece. Meanwhile, the tape recording of the probe signal was continued. On subsequently setting the analyser to the frequency of a steady-state peak, it was observed whether this frequency increased or decreased as the saw slowed down. A decreasing frequency identified the wave as being forward-moving.

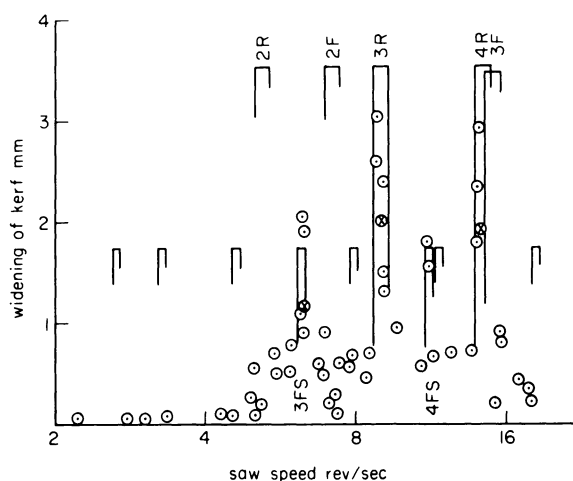
DISCUSSION OF RESULTS

Slices about 8 mm thick were cut from aluminium slabs using a number of steady saw speeds. As the saw approached the end of the section being cut, it was stopped and the width of cut at this point was measured with slip gauges. After deducting the cutting width of the tooth (6.60 mm), the excess kerf width was plotted. Figure 6(a) shows results for the 24-tooth saw. Calculated resonance ranges are marked and designated by the n value followed by F (forward), R (reverse), S (secondary).

In this saw, forward-travelling waves for $n = 2, 3, 4$ and 5 were excited within their predicted speed ranges. For some of these, the upper limit of kerf widening is seen to be rather higher than the value calculated from Equation (3). At a resonant speed, with maximum amplitude of vibration, the extreme end of the saw cut appeared to meander from side to side in a periodic time of the order of one second. This was interpreted as a slow precession of the tooth path wave in the direction of cutting, as expected. Results subsequently obtained for the 12-tooth saw (figure 6(b)) showed well-developed resonances for reverse-travelling waves of $n = 3$ and 4. Secondary resonances for forward-travelling waves of $n = 3$ and 4 were also established.



(a) 24-tooth saw



(b) 12-tooth saw

Figure 6. Measurements of excess kerf. Vibration analysis is given for tests marked with cross.

Frequency distributions obtained from analysis of tape recordings of probe signals for selected tests are shown in figure 7. An analogue analyser was used and proved to be quite satisfactory. Referring to the distribution for a nominal 2F vibration, it is seen that while this frequency predominated, a 2R wave of appreciable amplitude was also present. A similar frequency distribution was obtained on a previous occasion by exciting a 2F wave in a freely rotating saw using an electromagnet³. The presence of an oppositely directed wave is attributed to partial reflection of the excited wave at one or more irregularities of internal stress in the saw disc. The existence of such irregularities may be inferred from the various natural frequencies, shown in Table 1, which were obtained as the exciter was moved to various points around the rim.

The variation for mode $n = 1$ is also seen to be large, and waves of this kind were not found to be continuously excited in any of the tests. As also shown in Table 1, the variation of natural frequency in modes $n = 3$ and higher was quite small. This may account for the relatively pure vibration in mode $n = 3$ shown in figure 7(a). Similarly pure waves were recorded for the 4F and 5F resonances, but are not shown. Vibrations in mode $n = 6$ and higher modes were not found, presumably because of increased flexural stress in relation to vibration amplitude, and hence higher hysteresis damping relative to excitation strength.

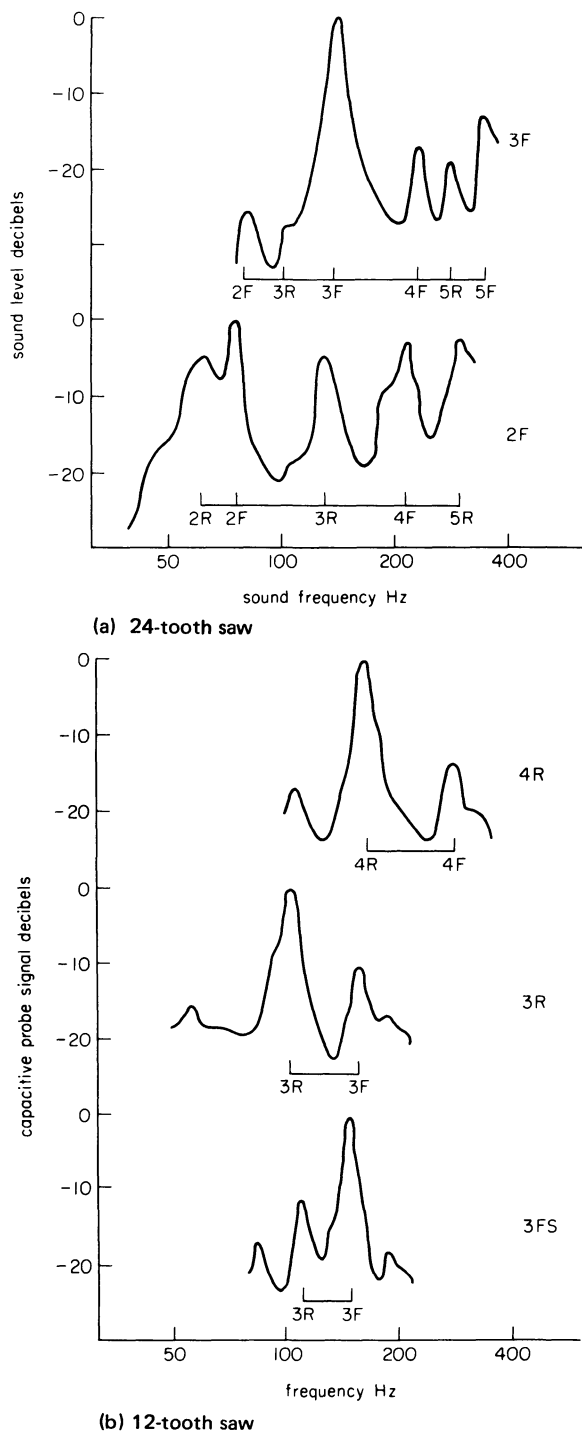


Figure 7 Frequency analysis of probe signal. Calculated values are marked.

Frequency distributions obtained from the 12-tooth saw can now be examined. It is perhaps not surprising that waves of $n = 2$ were not found, as the excitation strength for this saw is lower in all modes (Equation (2)). Figure 7(b) confirms the presence of fairly pure waves of the expected kind at three of the resonant speeds (one diagram for a further resonant speed is not shown). In each a dominant wave is shown, accompanied by an oppositely moving wave of smaller amplitude which tends to fade out as the amplitude of the dominant wave increases. This suggests a possible explanation for the effect of slots in the rim of a saw in alleviating vibration. These will

ensure that if any particular wave develops and sets up its distinctive profile on the surface of the workpiece, an oppositely moving wave will soon reach sufficient amplitude to destroy this profile and remove the excitation.

There is not enough evidence to explain why certain resonances were not found. From analysis of probe signals at speeds where an expected resonance failed to develop, two or more waves of small amplitude were found to be present. It appears likely that if two resonance ranges overlap, a weak vibration consisting of both waves is generated. Neither can set up its appropriate profile on the surface of the workpiece without interference from the other. In particular, a resonance of the *secondary* type, with its high Q -value and high available excitation as indicated by Equation (2), is likely to be set up preferentially. Thus the development of any interaction of the *primary* type which happens to have a coincident range of resonance is inhibited. This may explain the non-appearance of 3R and 4R waves in the 24-tooth saw (figure 6(a)).

If this idea is valid, it might elucidate the effects of certain practical expedients. Re-smithing to change internal stress or changing collar diameter will certainly change natural frequencies in the lower modes. This changes relative positions of resonance ranges, with the possible effect of excluding a previously available resonance.

CONCLUSIONS

1. Tests on a saw with two successive arrangements of teeth revealed resonances during cutting, at speeds associated with the presence of identified and predictable travelling waves. These waves may be forward-travelling or reverse-travelling.
2. The proposed mechanism of excitation leads to an appreciation of the factors controlling vibrations of the kind which are continuously sustained during cutting. The limiting amplitude appears to be proportionate to side clearance angle and tooth pitch. It is realized that requirements for efficient cutting may not allow these variables to be changed at will. Increasing the number of teeth is expected to increase excitation strength but reduce the attainable amplitude of vibration.
3. If the continuity of the rim is destroyed by inserting slots or holes, these will tend to convert a travelling wave to a stationary wave for which no excitation is available. It is thought that for the best effect, the number of evenly spaced slots should equal the number of radial nodal lines in the vibration to be suppressed.
4. Obviously, a working speed should be chosen so that it is not in the neighbourhood of a resonance range. When reasons are better understood as to why certain resonances appear to be excluded, it might be possible to arrange the exclusion of a difficult resonance by, for example, adopting a particular collar diameter.

5. As the vibrations examined here are of low frequency, that is, of low n -value, the introduction of damping capacity through the clamping collars may possibly be effective⁴.

ACKNOWLEDGMENT

The work described was commissioned and supported by Firth Brown Tools Ltd., Sheffield, and is published with their permission.

REFERENCES

1. D. S. Dugdale. Discrete frequency noise from free-running circular saws. *J. Sound Vib.* (1969) **10**, 296.
2. C. D. Mote, Jr. Noise generation by saw teeth. *Proc. 4th Wood Machining Seminar* (University of California) (1973) (*In press*).
3. D. S. Dugdale. Experimental analysis of circular saw vibration. *Proc. 4th Wood Machining Seminar* (University of California) (1973) (*In press*).
4. A. E. Grub et al. Damping lateral vibrations of circular saws. *Lesn. prom-st'* (1969) **2**, 18.

NOISE AND ITS REDUCTION

NOISE GENERATION FROM AN IMPACT FORMING MACHINE STRUCTURE

by

A. E. M. OSMAN*, W. A. KNIGHT* and M. M. SADEK*

SUMMARY

The noise, vibration and load generated during the operation of an impact forming machine are determined for a fixed input energy and the sound field around the machine is investigated. It is found that the major proportion of the noise energy is generated in the impact area and the dominant source of noise is the free vibration of the structure with a further contribution due to excitation of the structure by the impact load. The effect of varying the ratio of the billet deformation energy to the total input energy shows that the noise and vibration levels increase as the billet height is reduced.

INTRODUCTION

Recent concern with the effect of industrialized society on the human environment has focused attention, among other things, on the hazardous effect of excessive noise levels, which may result in permanent physical and psychological damage to exposed persons. Legislation passed in the U.S.A.¹ stipulates a maximum permissible noise level in factories of 90 dBA for a one hour exposure period and permits a peak value of 140 dB for exposure to impulsive noise. In Britain, the Wilson Committee² recommended similar noise levels for industrial processes and the introduction of a bill is only a matter of time.

Metal forming operations are a particularly hazardous source of noise, because of the high-intensity, short-duration noise pulses which are generated repetitively. Although a significant reduction of the workshop noise level can be obtained by means of enclosures and sound-absorbent materials, in the long term, the noise problem has to be remedied in the design stage of the machines used. In order that this may be achieved, it is necessary to study noise generating and radiating mechanisms in forming machines. This paper contains the results of an investigation into the role that structural vibrations play in generating noise in an impact forming machine. The tests described were carried out on a Petro-Forge machine, but there is no reason to suspect that the results are untypical of any impact forming machine.

It is recognized that there are other sources of noise in impact forming machines³, in particular pressure pulses due to the ejection of air from rapidly approaching surfaces^{3,4} and to the sudden arrest of

impacting bodies⁵. Both these effects, however, result in noise pulses of very short duration and it may be expected that the greatest overall noise level will be due to structural vibrations.

NOISE AND VIBRATION MEASUREMENTS

Measurements were taken on a Petro-Forge Mk 11 C machine sited in a conventional laboratory, with other machines located nearby, which made it possible to take measurements on one side only. However, because of the symmetry of the machine along its vertical axis, it would not be expected that the sound field would vary much around the machine and, for this reason, noise measurements were taken along a critical axis facing different structural members of the machine. In order to avoid near field effects, the microphone was placed 550 mm from the machine centre.

In addition to the noise measurements, vibrations of the two main structural members, the platen and anvil, were monitored with piezo-electric accelerometers. The forming load was measured by means of a strain gauge load cell placed directly under the lower die.

All signals were recorded on an instrumentation tape recorder with a band width of d.c. to 80 kHz after transmission through long cables. The analogue signals were played back at a speed of $1\frac{7}{8}$ in/s, digitized through an analogue to digital converter into a computer and subsequently processed using a Fourier transform program to give frequency spectrums of the signals in terms of the amplitude density.

* Department of Mechanical Engineering, University of Birmingham

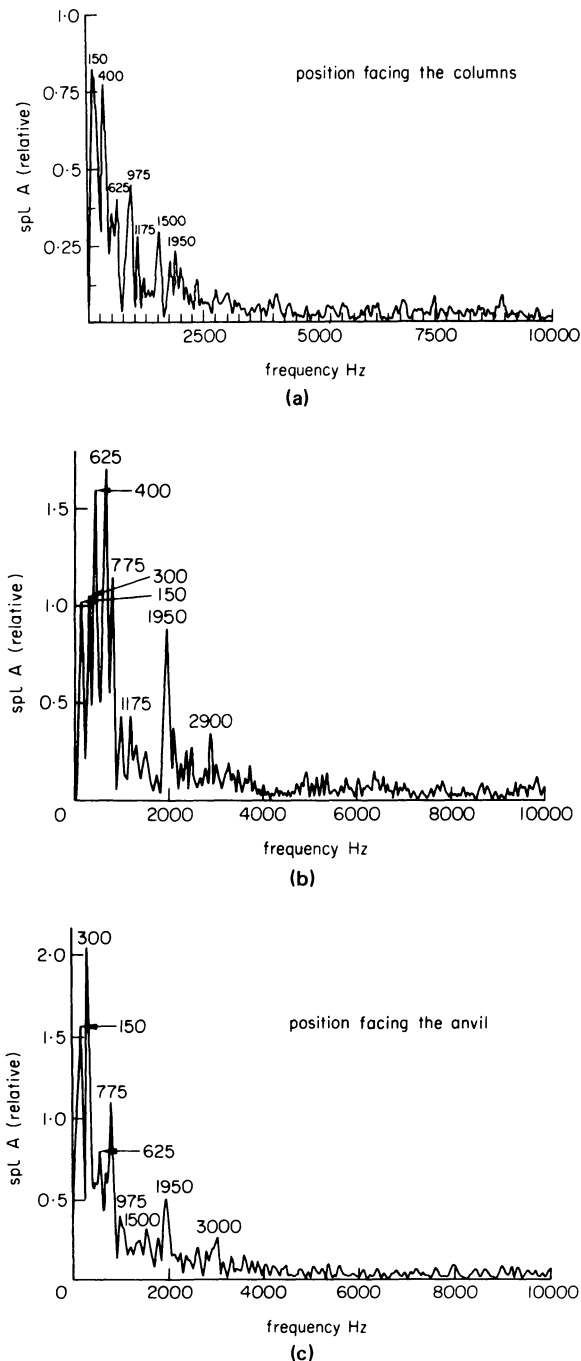


Figure 1. Sample noise spectra along the vertical axis of the machine. (a) Microphone facing columns. (b) Microphone facing billet. (c) Microphone facing anvil.

Noise field around the machine

Tests were carried out on the machine for a simple upsetting operation on cold steel 1 in diameter billets, with a fixed input energy of 10 000 ft lb. Measurements were taken for various vertical positions of the microphone. Figure 1 shows sample noise spectra for three positions of the microphone, i.e. opposite the billet, facing the anvil and facing the columns of the machine. A number of pronounced peaks in the spectra are apparent below 3 kHz, with some frequencies common to all spectra, but with the relative magnitudes of the peaks varying with the microphone position. This can best be illustrated by reference to figure 2, which shows the magnitudes of the common frequency components plotted against the micro-

phone position. Also shown is the overall peak noise level contour.

The highest overall peak noise level occurs at the position facing the billet, which also exhibits the highest number of peaks in the spectrum below 800 Hz (see figure 1). Most of the frequency peaks also show the highest amplitude at the level of the impact zone and anvil, with the main exception of the peak at 775 Hz, which has its greatest amplitude when the microphone position is facing the legs of the machine. Opposite to the billet, the 400 Hz frequency has the

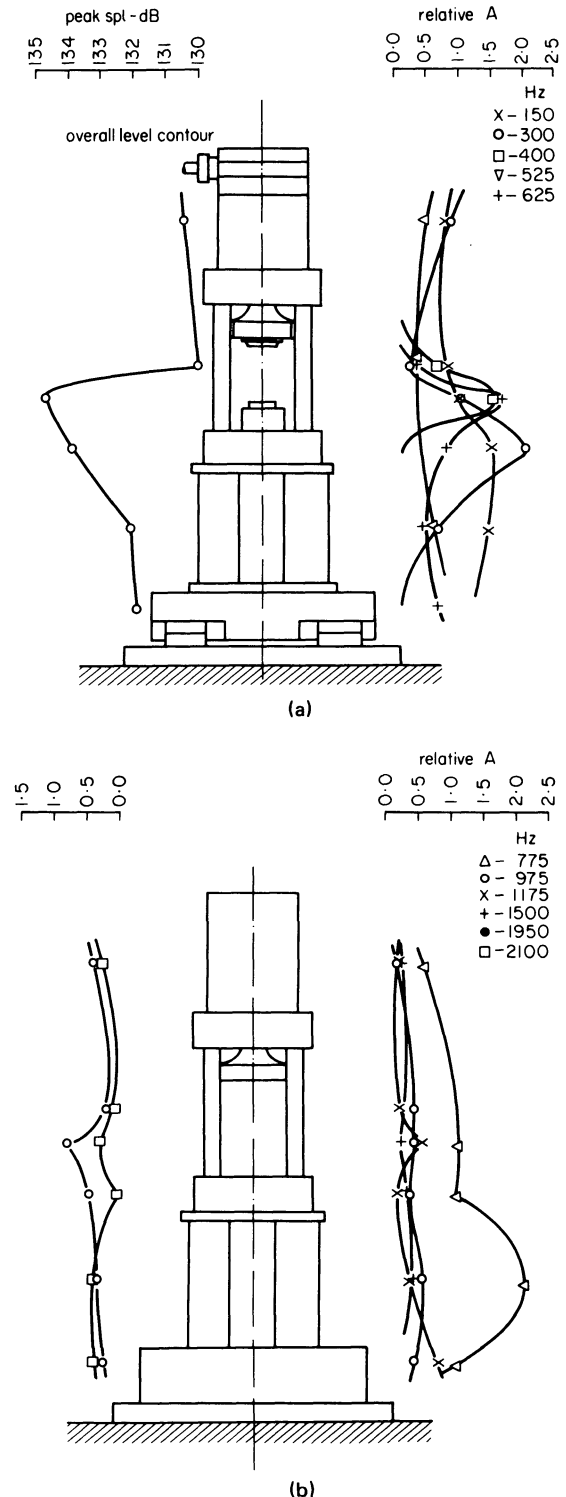


Figure 2. Frequency contours of noise peak amplitude density. (a) 150-625 Hz. (b) 775-2100 Hz.

highest noise level, whereas opposite the anvil, the 300 Hz peak is the largest.

In conjunction with the noise measurements, acceleration measurements were carried out on the platen and anvil of the machine. Sample spectra shown in figure 3 indicate a pronounced low fre-

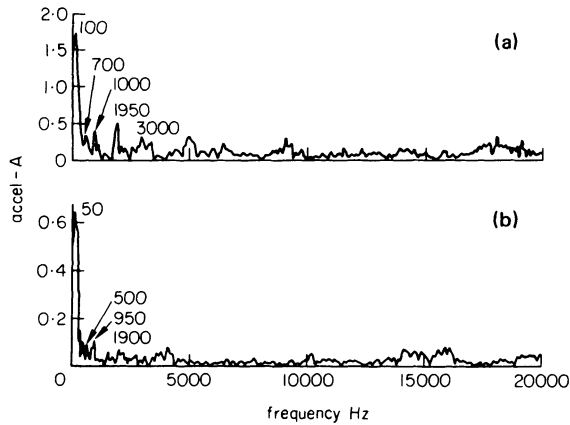


Figure 3. Samples of acceleration spectra. (a) Platen vibration. (b) Anvil vibration.

quency content. For the platen, several peaks occur, with the lowest frequency peak at 100 Hz being of the greatest magnitude. From zero to 500 Hz a downward slope of the spectrum is apparent and the frequency content is similar to that of the forming load spectrum illustrated in figure 4; this is apparently due to the impact force. This range of frequencies represents the forced response of the platen to the predominant frequencies of the impact force which mainly lie in the same range. The spectrum also shows a number of peaks which vary in their relative amplitude according to the test position. Frequencies at 600, 700, 950, 1150, 2300, 3050, 3400, and 500 Hz, respectively, represent the natural frequencies of the platen.

The spectrum for the anvil shows an identical pattern that of the platen in the frequency band up to 200 Hz, indicating a peak value at zero frequency and, further, a large number of frequency components ranging 300 Hz up to 13 kHz. It is rather difficult to identify similar frequencies in other spectra apart from the 300, 800 and 3600 Hz components. Components of lower frequency band up to

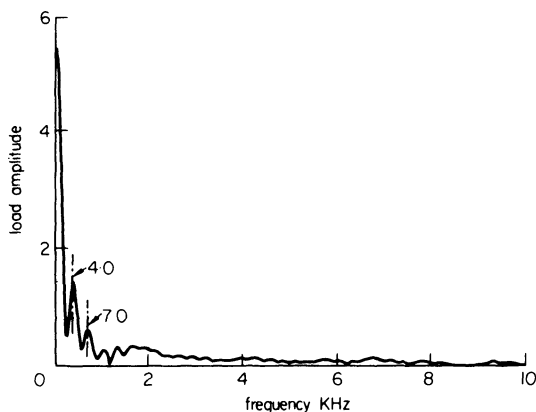


Figure 4. Forming load spectrum for upsetting $\frac{3}{4}$ in billet.

200 Hz are due to the exciting force; it is not clear, however, whether the 300 Hz component is due to a forced or natural vibration. Other spectra suggest that it represents a normal mode which is super-imposed on the forced response at the position close to the shock excitation.

In order to correlate the structurally borne noise with the vibration of the two impacting bodies, a visual inspection of the relevant spectra is discussed here. Noise spectra of the two microphone positions facing the impact zone and the anvil are compared to the acceleration spectra of the platen and the anvil. The first peak in the noise spectra at 125 Hz or 150 Hz could be attributed to the forced vibration of the platen and anvil, which is indicated in their respective spectra by a peak value in the frequency band 0 to 150 Hz. The noise spectrum of the position facing the impact zone shows other peaks at the following frequencies: 300, 400, 625, 775, 975, 1175, 1325, 1500, 2100, 2350, 2450 and 2900 Hz. In the noise spectrum of the anvil position, peak frequencies occur at 300, 825, 775, 975, 1500, 1950, 3000 Hz. The frequency component of 300 Hz, which is more dominant in the anvil position, is due to the forced and natural vibration of the bottom die and the anvil. The frequency components at 625, 975, 1175, 1325, 1950, and 2900 Hz are comparable to the natural frequencies of the platen at 600, 700, 950, 1150, 1350, 1950, and 3050 Hz. Hence, it could be stated that those frequency components in the noise spectrum are caused by the natural vibration of the platen. These frequencies also appear as dominant components in the noise spectrum at the position facing the columns. The frequency components of 775 Hz, which also appears as the highest peak in the spectra of the two positions facing the legs and the base plate, could be due to a bending mode of vibration for the anvil and the base plate. In particular, acceleration measurement on the lower die shows a dominant component at 800 Hz. The remaining component in the noise spectrum of a frequency 400 Hz appears clearly in the vibration spectrum of the bottom die and thus could be attributed to the forced motion of the platen in response to the impact force, which acts simultaneously on both the platen and the bottom die.

The higher frequency components in the vibrations of the platen and the anvil do not radiate any appreciable noise and should be disregarded.

ELASTIC STRAIN ENERGY AND ITS INTERACTION WITH NOISE GENERATION

When the platen of the machine, travelling at high speed, impacts with the billet, its kinetic energy is transformed into deformation energy absorbed by the billet and elastic strain energy in the machine structural members. Subsequent to deformation, the strain energy is dissipated in the damping allocated with the natural mode of vibration of the structure. Thus, if the energy of deformation is altered, for the same energy input a greater or smaller amount of energy is available to excite the natural mode of vibration with the resulting effects on the noise levels generated. One way of achieving a variation in the deformation

energy is to upset billets of different heights for the same energy input. This was done for EN8 billets of 1 in diameter with heights ranging from 0.75 to 2 in with a fixed input energy to the machine of 11 200 ft

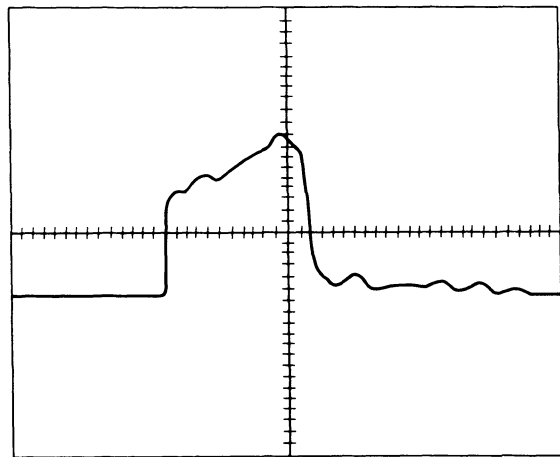


Figure 5(a) Variation of load during deformation.

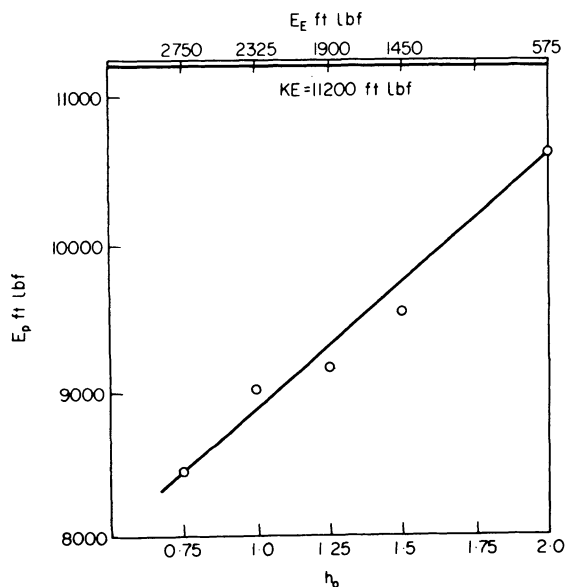


Figure 5(b) Variation of deformation energy with billet weight in constant platen kinetic energy.

lbf, the corresponding impact velocity being 50 ft/sec. In each case, the forging load platen vibration and noise for a microphone position of 20 in from its billet were measured. Figure 5 shows a typical load time curve for the upsetting process. In each case, the energy of deformation was determined from measurements of the amount of deformation in each billet.

The energy absorbed by the plastic deformation of the billet can be given by the relationship

$$E_P = \int_0^H F_C ds$$

where

E_P = energy of deformation of the billet

F_C = current forming load

s = current deformation stroke

H = total deformation stroke

A theoretical analysis of the effect of billet dimensions on the forming load and the energy of deformation is beyond the scope of the present work and only experimental data are used here. The energy of deformation would be obtained by measuring the load and displacement time histories and calculating the area under the plotted load-displacement diagram. However, while carrying out the experimental work, no displacement transducer was available and an alternative approach was followed. It can be shown readily that

$$E_P = A_0 h_0 \int_{h_f}^{h_0} \sigma de$$

where

A_0 = initial cross-sectional area

h_0 = initial height

h_1 = final height

σ = natural stress

e = natural strain

Using a natural stress-strain diagram obtained under similar conditions, it is possible to construct another diagram where the new parameters are E_P/V_0 against the engineering strain e . The procedure followed is to divide the natural strain axis into small values ($\epsilon_0, \epsilon_1, \dots, \epsilon_n$) and calculating the elemental areas contained by these values. The first elemental area would thus represent the value of E_P/V_0 for an engineering strain equal to ϵ_1 . Summing up the elemental areas up to a value ϵ_m would yield the magnitude of E_P/V_0 for an engineering strain equal to ϵ_m . Thus, if the initial and final heights of the billet were measured, it would be possible to obtain an estimate of the energy of deformation. It is of importance to note that the method is applicable only to billets of initial diameter equal to that for which the natural stress-strain diagram was obtained. Due to the effects of friction and to other factors, the natural stress-strain diagrams will probably show a marked difference for billets of different diameters.

The value E_P for different heights is plotted in figure 5(b) and a datum representing the kinetic energy of 11 200 ft lbf is shown at the top, together with the values of elastic strain energy, E_L , for every billet height. The relation between the peak load, peak sound pressure level and the elastic strain energy is shown in figure 6(a). Further, the peak values of the load were plotted against the peak values of sound pressure levels in figure 6(b). The two figures clearly indicate that the increase in the forming load is proportional to the increase in elastic strain energy and, consequently, in the generated noise. This would have been expected, as the increase of strain energy

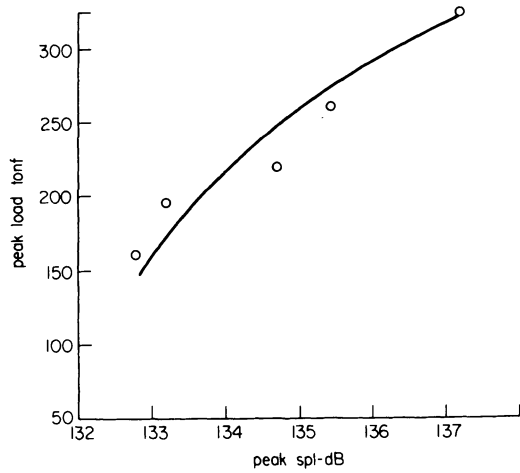


Figure 6(a). Variation of peak load and peak sound pressure level with elastic strain energy of the machine structure.

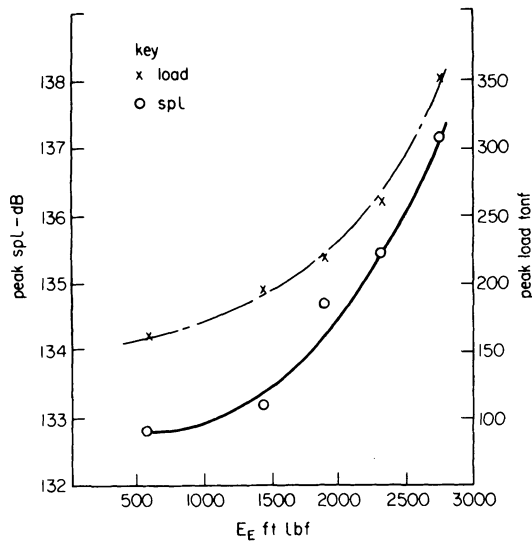


Figure 6(b). Relation between peak sound pressure and peak load.

and exciting force would yield an increase in the vibrational energy of the system in terms of both the forced and free vibrations.

In order to identify the effect of increasing the elastic strain energy on the generated noise, the relative amplitudes of the sound pressure level at different peak frequencies were plotted against the strain energy values, as shown in figure 7. Similarly, those frequency contours are plotted for acceleration levels in figure 8. The frequency contours in both curves indicate an increase in acceleration and sound pressure levels with the increase of strain energy. The rate of increase of acceleration for the different frequencies does not seem to differ much, apart from the 250 Hz contour. However, the rate differs widely for the sound pressure levels, in particular the 400 Hz contour. This frequency was marked as peak frequency during noise measurements in the microphone position facing the impact area and was attributed to the force motion of the platen. It seems that the increase of the existing force largely affects the system response at this frequency. In the acceleration contours, this frequency was not plotted as the spectra generally showed a broad peak sloping down-

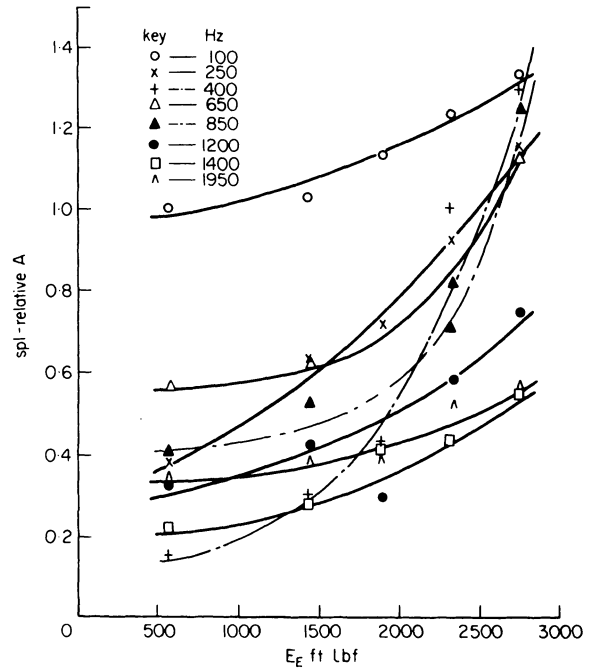


Figure 7. Variation of sound pressure amplitude density with elastic strain energy at various frequencies.

ward up to a frequency in the range of 500-600 Hz. Figure 9 shows a plot of the relative acceleration and sound pressure amplitudes at the common frequencies that it was possible to detect. Frequency analysis of the load waveforms were carried out and frequency spectra for the two extreme billets used—the tests (0.75 and 2 in) are shown figures 4 and 10, respectively. The spectrum for the short billet exhibits a higher number of loops and the frequency of 400 Hz appears as a dominant component thus corresponding to a

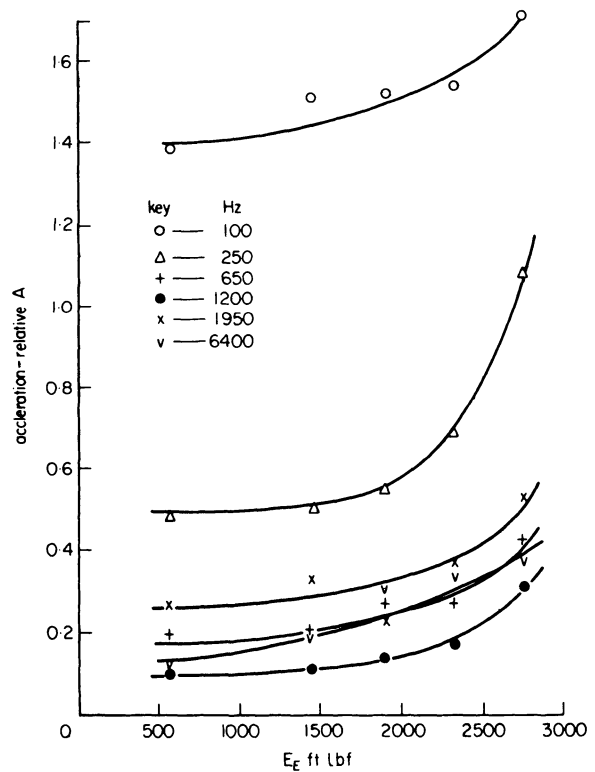


Figure 8. Variation of platen vibration acceleration with elastic strain energy at various frequencies.

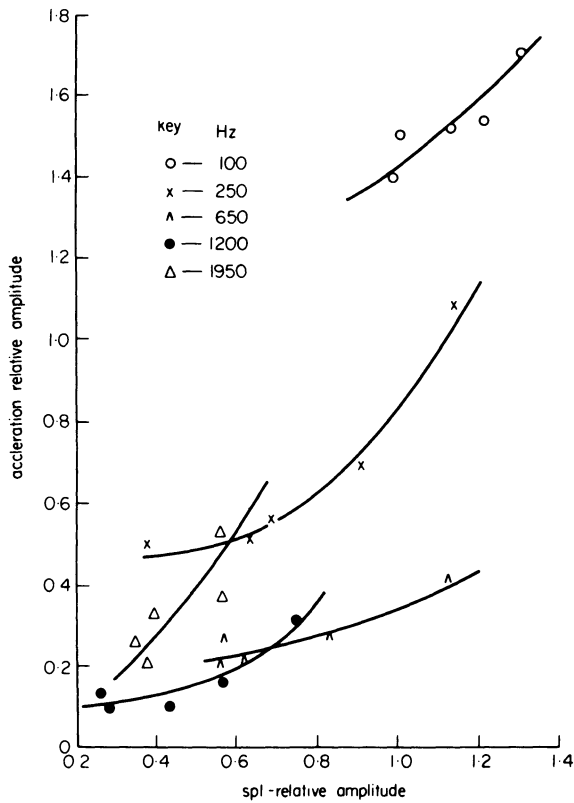


Figure 9. Relation between sound levels and platen acceleration at common frequencies.

peak frequency in the noise spectrum, as shown above. This again shows that the particular frequency results from forcing by the components of the input force.

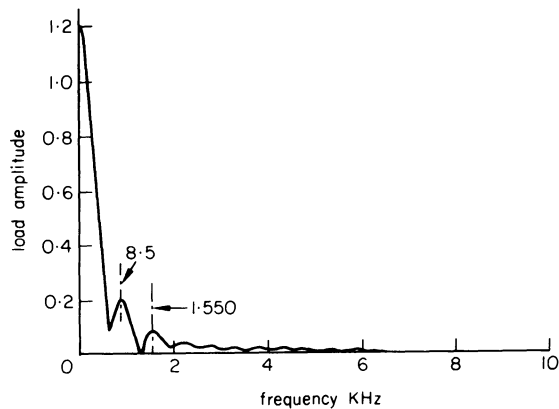


Figure 10. Frequency spectrum of load for upsetting 2 in billet.

CONCLUSIONS

Frequency analysis of the noise and vibration of the machine indicates that the noise is generated by both the forced motion of the platen and the natural vibrations of the impacting bodies, that is, the platen, the lower die and the anvil and, to a lesser extent, by the natural vibrations of other machine members. Vibration frequencies above 3 kHz do not radiate appreciable noise, while most attention should be directed to the platen and the lower die vibrations up to 2 kHz. Natural vibrations of the anvil are of relatively small amplitude due to the high structural stiffness and most of the noise radiated from this part is due to its forced response to the predominant frequencies of the shock.

The noise generated is clearly dependent on the force and the amount of energy expended in elastically straining the machine member. In addition, the increase in forming load results in a change in its frequency content and the increase in the generated noise does not occur at the same rate for every frequency component. However, the increase in the vibrational energy will result in an increase of the forced motions and free vibrations of the system.

ACKNOWLEDGMENT

This work was carried out in the Department of Mechanical Engineering, The University of Birmingham. The continuous advice and encouragement of Professor S. A. Tobias, Head of the Department, is gratefully acknowledged. The project is financed by a grant from The Science Research Council.

REFERENCES

1. Walsh-Healey Public Contracts Act (1969). *U.S. Federal Register*, 34, 96, Rule 50-204-10.
2. Anon. Noise (Final Report). *H. M. Stationery Office*, London, 1963.
3. V. Gregorian, M. M. Sadek and S. A. Tobias. The fundamentals of noise generated in impact forming machines. *Proc. of the 13th Int. M.T.D.R. Conference*, Birmingham, 1972.
4. G. H. Trengrouse and F. K. Bannister. Noise due to an air ejection from clash surfaces of impact forming machines. *Machine Design Conference*, A.S.M.E., Cincinnati, Ohio, U.S.A., 1973.
5. G. Nishimura and K. Takahashi. Impact sound by mutual collision of two steel balls. *Bull. The Japan Soc. of Prec. Engg.* 1.2 (1963).

NOISE REDUCTION IN A HERF CROPPING MACHINE

by

G. DROUBI*, M. M. SADEK* and S. A. TOBIAS*

SUMMARY

Noise generated in a HERF cropping machine is investigated with the aim of providing data for the design of an improved larger version. The main noise sources are identified and the effectiveness of an acoustic shield is predicted. An experimental acoustic shield is described and the actual sound level reduction achieved with this is compared with the prediction made.

The experimental acoustic shield produced a substantial reduction of the noise level emitted but by itself proved insufficient. Further tests were made by enclosing the stand of the machine with a polyurethane skirting with the aim of eliminating a secondary noise source.

On the basis of measurements made, general design principles were evolved for the improved machine. The acoustic features of this machine are described.

INTRODUCTION

It is well known that in impact-forming operations the intensity of noise emitted increased with increasing 'severity' of the blow, the worst possible condition arising with a clashing of solid die surfaces.

The operation of high-speed cropping machines involves the clashing of three masses; that of a striker/ram assembly impacting on a cropping tool (moving blade) and that of the cropping tool with a decelerator. The dynamics of this system may result in a multiple clashing of these components and hence it is by no means surprising that the maximum noise intensity generated can reach a level of 120 dB (A).

The present paper is concerned with the impact noise generated in a particular high-speed cropping machine, that of a Petro-Crop Mk. ID hammer. The primary aim of the project was to establish the major sources from which this noise is radiated and the development of means for its reduction to an acceptable level. The secondary aim was the derivation of general design data for a new, large and improved machine.

With the technique described, it was established that the major part of the high-frequency noise was emitted from the impact zone. This suggested the use of a sound-isolating enclosure of the die area. An estimate was formed of the effectiveness of such a shield and this was experimentally verified. Although a substantial reduction of the noise level was thus achieved, it appeared that with this particular machine there existed a secondary noise source in the

stand of the hammer. Enclosing this with a skirting reduced the noise to an acceptable level. Further improvement was achieved by coating the impact area with a thin oil film.

On the basis of these results the acoustic design of the larger, improved machine was evolved and this is briefly discussed. The work reported is part of a major investigation on impact noise generated in forming machines, covering both the fundamental and the applied aspects.

GENERAL DESIGN FEATURES OF PETRO-CROP Mk. ID

This machine is a fully automatic cropping installation suitable for cropping bars in the diameter range 0.8 in to 1.5 in, at a rate of about 30 billets per minute depending on the length of the off-cut. The system, a photograph of which is shown in figure 1, consists of the following sub-units:

- (1) Petro-Forge Mk. ID-5K hammer and control consol,
- (2) cropping die set,
- (3) bar rack and feed unit.

The operation of these three units is integrated with suitable interlocks permitting the automatic production at a high rate of output of high quality billets in a wide range of materials.

(2) *Petro-Forge Mk. ID-5K*. This is a general-purpose high-speed hammer suitable for the production of components by hot, warm, and cold

* Department of Mechanical Engineering, University of Birmingham, Birmingham

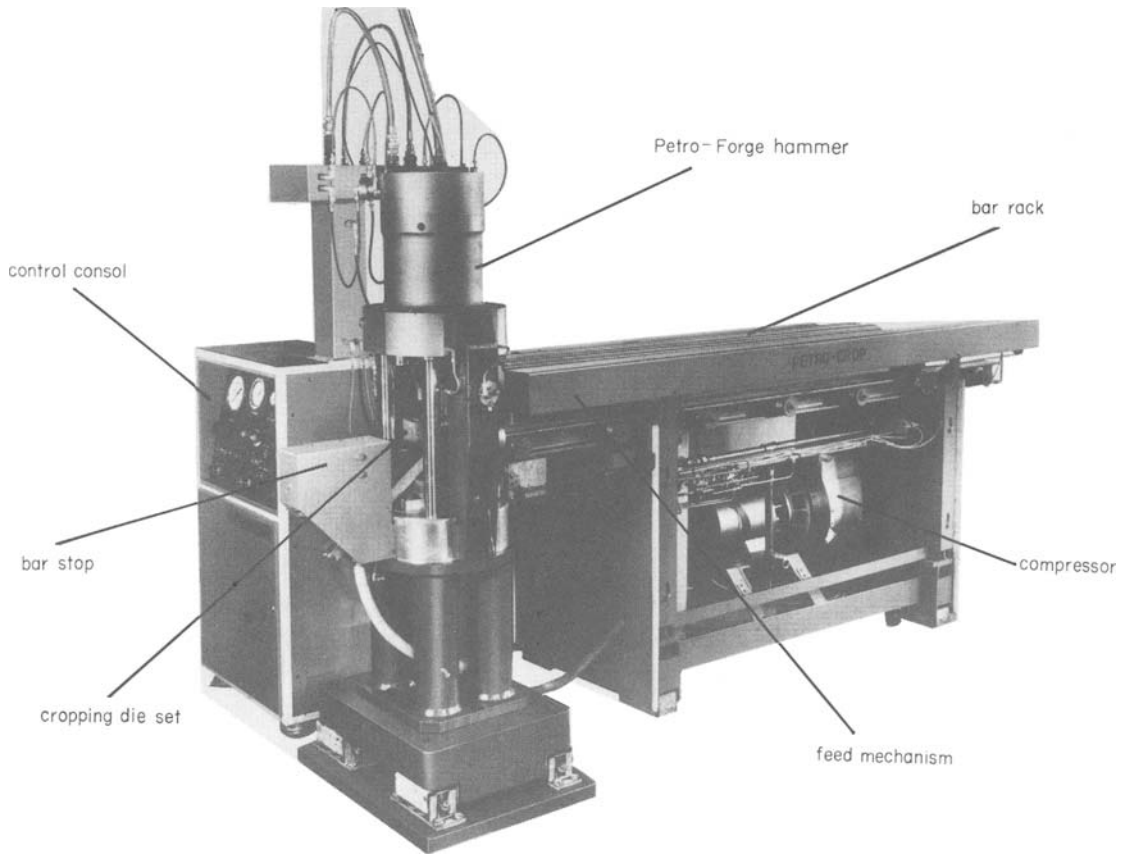


Figure 1. Petro-Crop Mk.ID-5K.

forming, piercing, blanking, powder compaction, sinter forging coining, rubber forming, and cropping, depending on the type of tooling used. This versatility of the machine is clearly of considerable practical importance but it has the disadvantage of requiring design features which are not ideal from the point of view of the impact noise arising in cropping, as will be seen later.

The nominal output of the hammer is 5000 ft lbf achieved at a die closing speed of about 50 ft/s. Maximum cycling rate is about one blow per second.

(2) *Cropping die set.* This is a self-contained unit die set, a cutaway drawing of which is shown in figure 2. During cropping the bar, which is clamped by the

pneumatically-operated collet A, passes through the stationary blade B into the moving blade C. The platen (striker) of the Petro-Forge hammer strikes the moving blade C, imparting to it kinetic energy. The major part of this is absorbed by the cropping operation and the remainder is taken up by the decelerator D, which is a pneumatically supported piston contained in the base plate E of the die set.

The principal moving components partaking in a cropping operation are the piston/ram/platen assembly of the power unit of the hammer, the moving blade carrier and the decelerating piston, as shown diagrammatically in figure 3(a). In the figure these components are drawn to the same scale and the weight of each of them is specified. Note that the ram/platen forging has a mushroom like shape, this

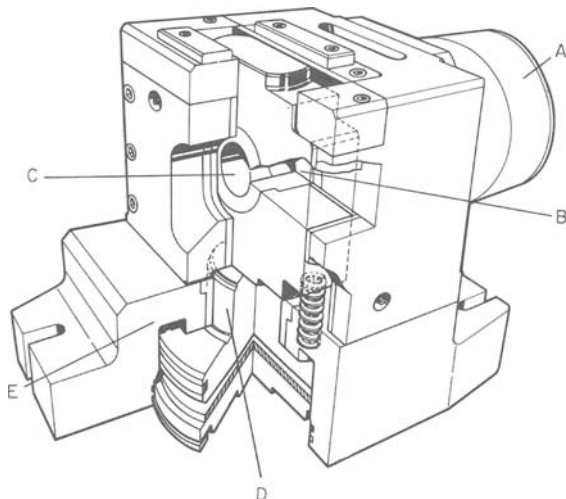


Figure 2. Cutaway drawing of the cropping die set.

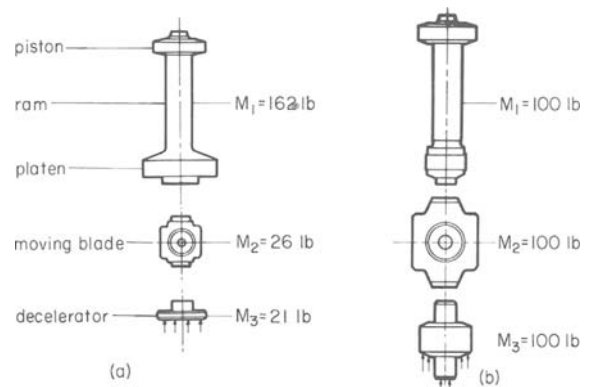


Figure 3. Moving components of hammer and die set (a) original version, (b) new design.

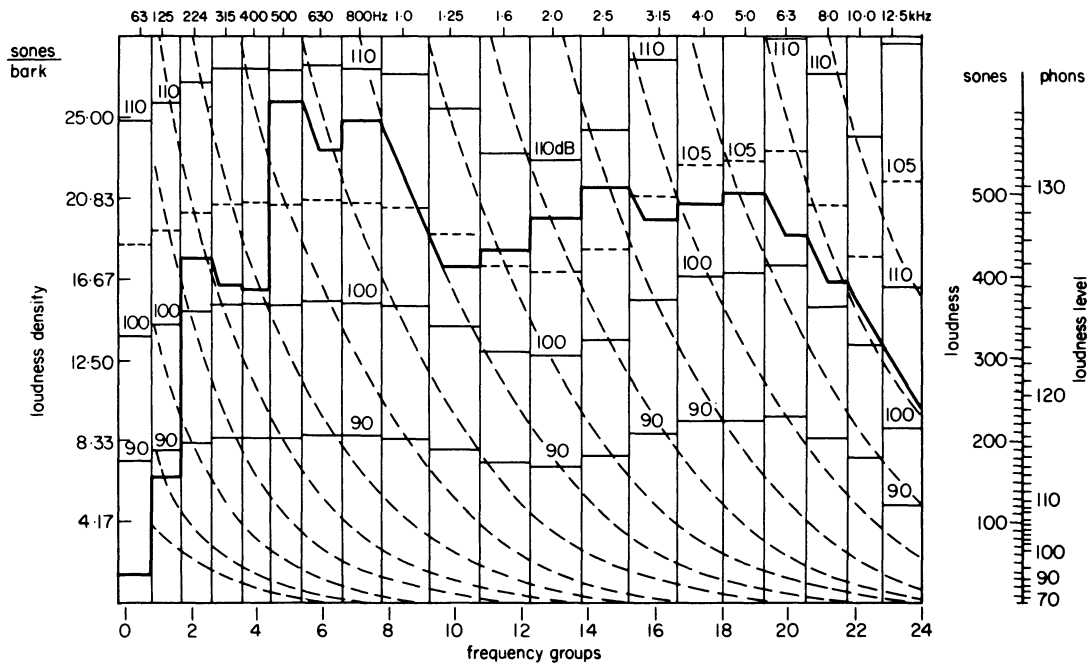


Figure 4. Example of loudness chart.

being determined by the need for a large platen surface for the attachment of tools required for forming operations other than cropping.

Impact noise is generated as a result of the clashing of the surfaces between the piston/ram/platen assembly and the moving blade carrier, and between the latter and the decelerating piston. An investigation of the dynamics of this three-mass system, involving direct measurement and high-speed cine photography¹ has shown that energy is being transferred from one mass to the next by multiple impacts.

As already mentioned, the noise generated in high-speed forming machines is dependent on the 'severity' of the blow. When energy is converted into forming work over a long forming stroke (forward or backward extrusion) then the noise level generated is low. When the energy is absorbed over a short stroke, as in the case of coining, then the noise intensity is high. Cropping falls into the second category.

The simplest method for reducing the intense noise is by enclosing the whole hammer in a sound-isolating booth. As a matter of fact, high-speed hammers are eminently suitable for such treatment since they are, by virtue of the high kinetic energy of their moving masses, very much smaller and more compact than machines operating at low speeds. However, acoustic enclosures of sufficient size are expensive, even for high-speed hammers. They are also cumbersome and space-consuming and therefore alternative solutions are preferable.

Previous work on a similar though somewhat larger machine² suggested that a simpler solution may be adequate. Acoustic measurements taken showed that in ordinary forming operations the major portion of the noise energy originates in the die area, the stand of the machine making the next significant contribution. The question arose, therefore, as to what can be achieved by enclosing the die area with

noise absorbing/isolating panels. The following sections describe the procedure by which an estimate was formed.

EXPERIMENTAL EQUIPMENT

The measuring instruments used consisted of a $\frac{1}{4}$ in condenser microphone connected via an amplifier to a magnetic tape recorder, the output of which is analysed by a loudness analyser*. This instrument contained twenty filters of one-third octave width (frequency range 45 Hz to 14 kHz) analysing the sound in terms of these. Each component is weighted in accordance with the Zwicker method³ and finally the output data of the instrument provides an x-y plot which gives the loudness density, that is, the subjective loudness as perceived by the ear, as a function of the centre frequencies of the one-third octave filters. The output is plotted on special graph paper which also permits the identification of the unweighted sound pressure levels of the filter bands. A typical example of such a graph is presented in figure 4, and will be discussed later.

The total loudness of the sound is automatically computed by the analyser as the integral of the area under the loudness density curve. This gives an overall measure of the intensity of the sound as sensed by the ear, and it is measured in phons.

The analyser is essentially an integrating instrument which computes the loudness density (and the overall intensity of the sound) over the period of measurement. In order to exclude the contribution of background noise caused by auxiliaries and so on, the tape recorder storing the impact noise signal also recorded a pulse signal indicating the instant of impact. This pulse was then used for triggering the loudness analyser, fixing the starting point of the

* Hewlett-Packard Model 8051A Loudness Analyser.

analysis and computation. The pulse was produced by a photocell attached to the frame of the hammer and actuated by the platen passing in front of it.

It is essential to realize that the impact sound signal recorded and analysed comprised the whole impact process, that is, it included noise generated by the impact of the platen with the blade carrier and the impact of the blade carrier with the decelerating piston, as well as the multiple impact of these components, the whole process taking a period of about 10 ms.

All tests to be discussed in the following sections correspond to an impact energy level required for cropping a 1 in diameter mild steel bar.

EXPERIMENTAL TECHNIQUE

The technique adopted involved the measurement of the impact noise at a level of the point of impact between platen and moving blade, at distances varying between 10–100 in from the impact zone, as shown diagrammatically in figure 5. On the assumption that the noise is generated by a point source radiating in spherical waves, the frequency components emitted from the impact area ought to vary as a function which is inversely proportional to the radial distance from the point source. This means that for the components originating in the impact

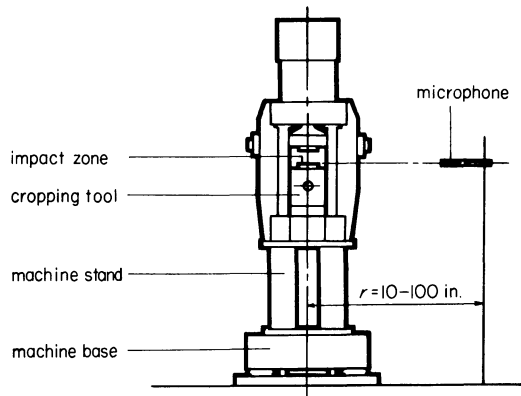


Figure 5. Microphone position facing the impact area.

zone, the sound pressure, as a function of the distance of the microphone from that zone, ought to be a hyperbolic function. This can be represented in a log-log diagram by a straight line inclined at 45° . Frequency components which originate outside the impact area should vary in a different manner.

At each microphone position, fixed by the radial distance r , the tests were repeated five times at a constant energy level set on the control consol. In each of these the recorded signal was analysed, yielding a loudness diagram of the type shown in figure 4. From each of these graphs the average sound pressure of each frequency band was extracted and this was plotted on a log-log basis as a function of the radial distance. The resulting graphs are shown in figure 6.

The curves presented in figure 6 indicate that the frequency components below 630 Hz (figures 6(a) to 6(e)) do not show the inverse linear variation

expected from components generated in the impact zone. This means that these are very largely (though not entirely) emitted from other parts of the machine. Most of the remaining curves show the expected tendency and these correspond to sound components which originate very largely from the impact zone.

Once the frequency components originating from the impact area have been identified, an estimate can be made of the contribution of these components to the total sound level generated.

PREDICTION OF NOISE REDUCTION

It is clear from the previous section that the noise level of the hammer can be substantially reduced by shielding the impact area with an absorbing material which has a high loss factor at the high-frequency range. The particular material chosen was therefore an acoustic laminate system combining the properties of lead cladding together with the absorption characteristics of polyurethane foam. This material had the following construction:

- (1) An acoustic membrane made of lead cladding, containing two lead layers sandwiching a steel layer for reinforcement to withstand the impulsive loading of the machine.
- (2) An absorption layer made from 2 in thick polyurethane foam, intended to absorb random incidence noise and to prevent the build-up of acoustic energy inside the enclosure.

The transmission loss factor of this material, obtained in standard tests in which a panel is mounted in an infinite wall, is shown in figure 7. As can be seen, its transmission loss factor increases with frequency, reaching a value of 40 dB for frequencies beyond 6 kHz. Under practical conditions, when the material does not form a part of a cut-out in a single wall but is used for surrounding some area, the transmission loss will be smaller than that given in figure 7, the loss of efficiency being accounted for by a factor of between 0.5 and 0.6 as specified by the manufacturers. In all subsequent calculations it was assumed that this loss factor is 0.6.

In figure 8 the full curve shows the loudness density spectrum of the hammer without shields, measured at a distance of 40 in from the impact area. It has already been established that frequency components emitted from the impact area are those which lie above 630 Hz, and hence for determining the effectiveness of the shield these need reducing by subtracting the transmission loss for each frequency band as given by figure 7 and taking into consideration the loss of efficiency factor of 0.6. The resulting line is the broken line in figure 8.

The sound pressure for each band is thus determined from the following equations:

$$\text{SPL}_1 = 20 \log_{10}(P_1/P_0) \quad (1)$$

where SPL_1 is the sound pressure level at frequency band F_1 ,

P_1 is the sound pressure at this band,

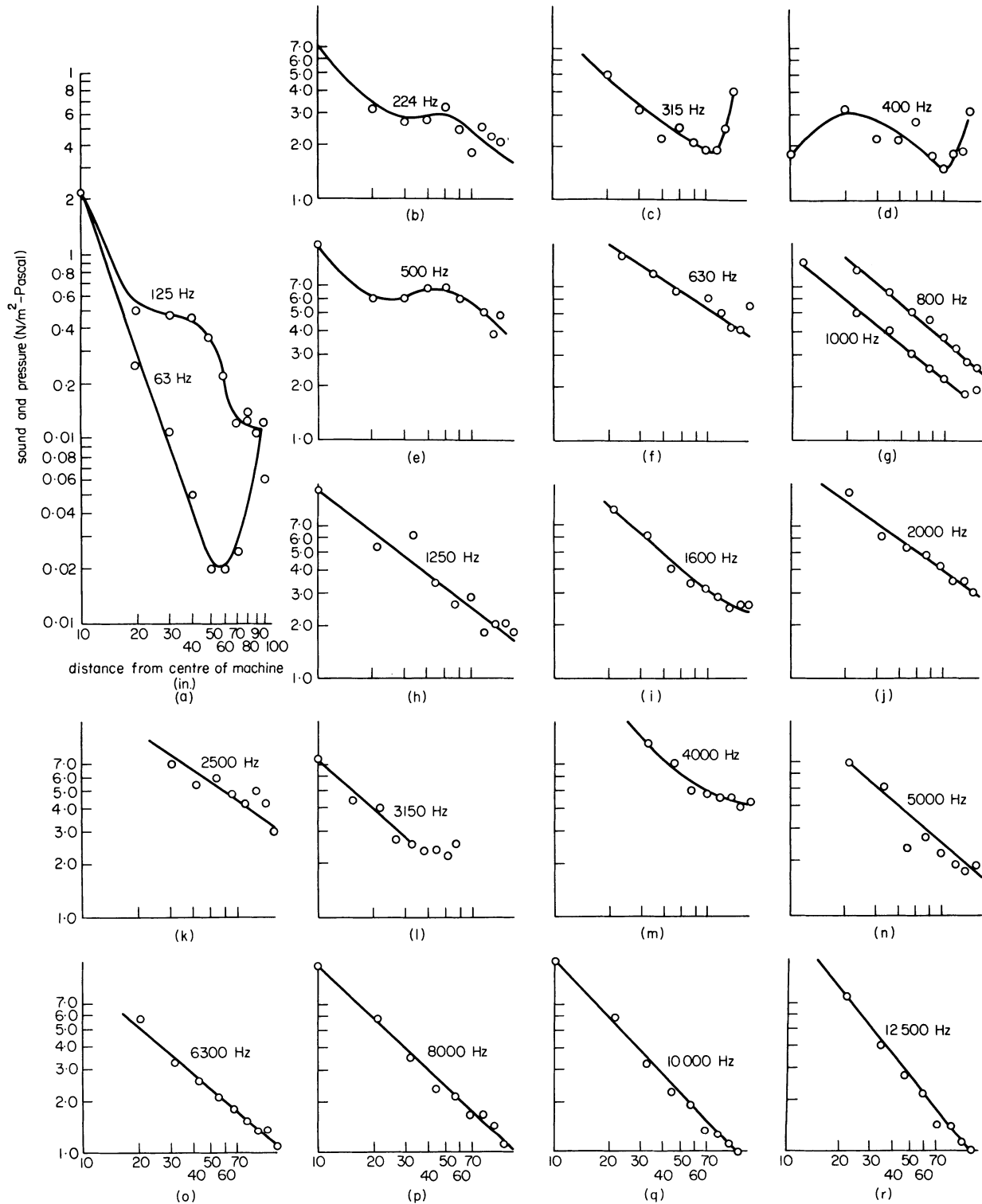


Figure 6. Variation of sound pressure level (SPL) in various frequency bands with the microphone position.

P_0 is the reference sound pressure;

$$P_t = \sqrt{(P_1^2 + P_2^2 + P_3^2 + \dots + P_n^2)} \quad (2)$$

where P_t is the total sound pressure,
 $P_1, P_2, P_3, \dots, P_n$ are the individual pressures
of the different frequency bands $f_1, f_2,$
 $f_3, \dots, f_n,$

thus resulting in the total sound pressure level

$$SPL_t = 20 \log_{10}(P_t/P_0) \quad (3)$$

This procedure leads to a predicted total sound pressure level of 107 dB which is 11 dB below that measured without the shields.

CONSTRUCTION OF SOUND-ISOLATING ENCLOSURES

The acoustic laminate was incorporated in a two-piece welded frame, the structure of which is shown in figure 9. The frame surrounding the whole impact area when clamped to the machine was isolated from

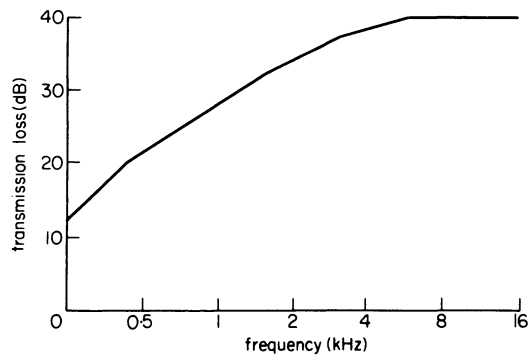


Figure 7. Variation of transmission loss factor with frequency.

it by visco-elastic strips $\frac{1}{8}$ in thick. The acoustic laminate panels were bonded to the inside of the frame, and were isolated from this also by isolating strips. The complete shield, when clamped to the machine, formed an air-tight enclosure, all precautions having been taken to close any openings. The enclosure attached to the hammer can be seen on the photograph in figure 10.

EFFECTIVENESS OF SOUND-ISOLATING ENCLOSURE

The effectiveness of the sound-isolating enclosure can be seen from figure 11. This shows the loudness density spectrum without the shield (solid line) identical with figure 8 and with the shield (broken line), measured under otherwise identical conditions, that is 40 in from the centre of the impact area. As expected the shield had no effect on frequencies below 450 Hz, partly because these were generated mainly outside the impact area and partly because of the characteristics of the acoustic laminate. Integrating the area below the two curves, as provided automatically by the analyser, gave a total loudness level of 129 phons without and 117 phons with the shield respectively. This corresponds to an unweighted sound-pressure level of 118 dB and 110 dB, respectively, a reduction of 8 dB. It will be recalled that the predicted reduction was 11 dB.

The reduction achieved is substantial, particularly at high frequencies which have a high annoyance value, but not adequate as far as industrial requirements are concerned.

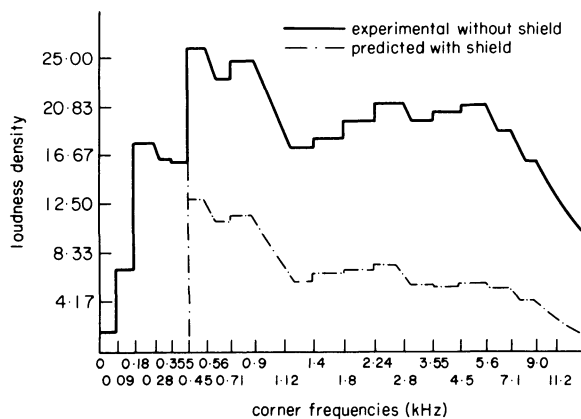


Figure 8. Predicted loudness density spectrum of the hammer with shield in relation to original loudness.

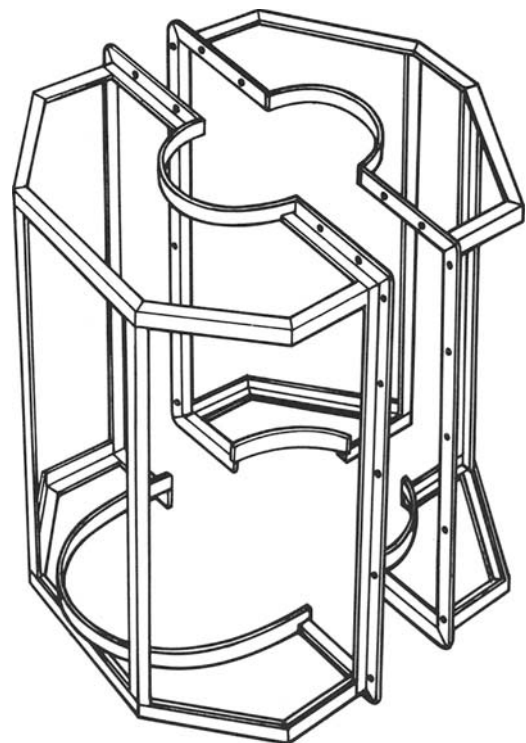


Figure 9. The welded shield frame.

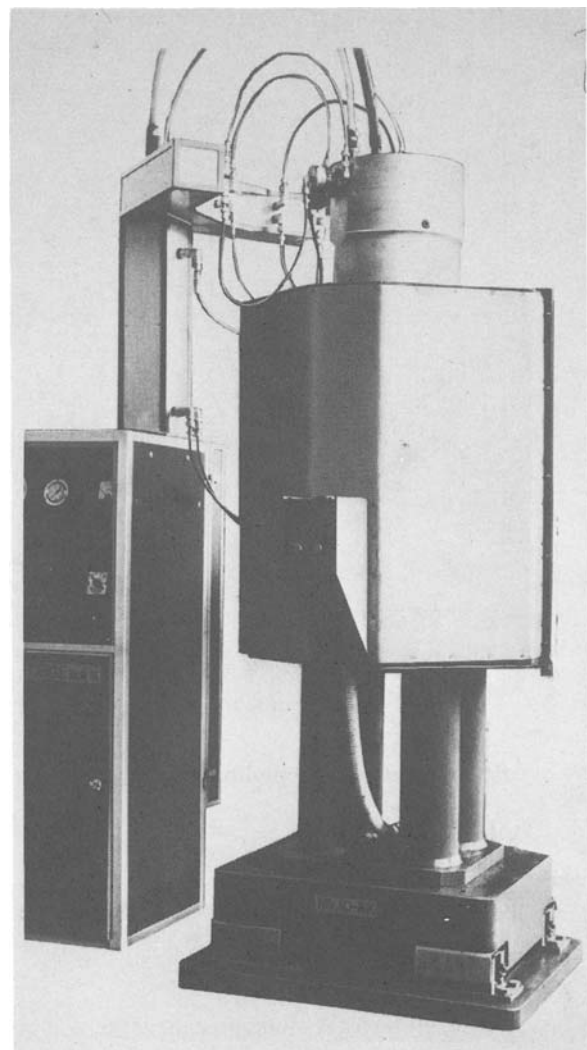


Figure 10. Petro-Crop Mk.ID with shield installed.

ISOLATION OF NOISE EMITTED BY MACHINE STAND

Further improvement was achieved by treating that area of the machine which, after the impact zone, makes the major contribution to the overall noise level generated, that is the surface of the machine stand. This stand is of welded construction, consisting of four circular columns welded to top and bottom plates. Thus, it is a lightly damped structural element which can be expected to 'ring' as a result of impact loading.

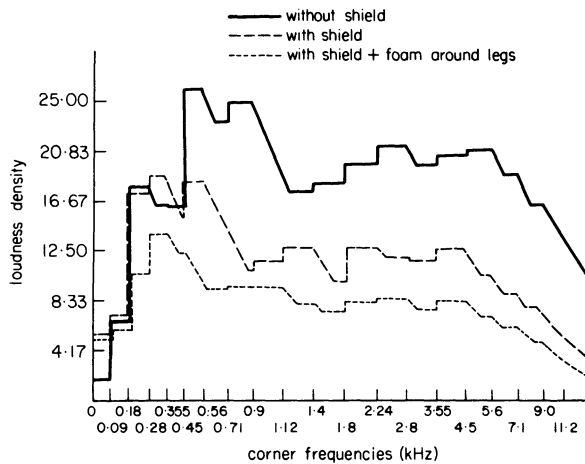


Figure 11. Loudness density spectrum of the hammer.

After surrounding the machine base with a skirt of 4 in-thick polyurethane foam, the loudness level shown by the *broken* line in figure 11 was obtained. From this diagram the loudness level was found to be 112 phons, which corresponds to an unweighted sound level pressure of 100 dB, that is a reduction of 10 dB. As far as the sound level pressure perceived by the ear is concerned, this corresponds to 93 dB (A).

EFFECT OF OIL FILM ON IMPACT SURFACES ON SOUND LEVEL PRESSURE

In the course of the cropping tests it was observed that the sound level pressure is substantially reduced by covering the impact surfaces with a thin film of oil. The reduction increased with the viscosity of the oil, best results being achieved with grease. It also increased with the thickness of the oil film (grease) layer. Mainly the higher frequencies were affected, which is important since this decreases the annoyance level of the sound. Noise components due to the stand of the machine were practically unchanged.

This observation is being investigated further with the aim of establishing the mechanism involved and also of optimizing the effect. However, even with the very limited understanding of the phenomenon it appears that by flooding the impact area with oil, which is quite feasible with the machine considered, the noise level can be reduced by 3 dB.

Thus if all three means for the reduction of the noise level are applied—an enclosing of the die space, a surrounding of the machine stand by a skirting and a covering of the impact surface with a thin film of oil—a total reduction of the order of 21 dB can be expected.

DESIGN OF THE PETRO-CROP MK.IE FROM THE ACOUSTIC POINT OF VIEW

The investigation has shown the way for a redesign of the Petro-Crop Mk.ID machine (see figure 1) from the acoustic point of view. Two major modifications have been introduced:

- (1) The die space, containing the impact area, has been surrounded by sound isolating panels, those on the two sides of the hammer being integral with the machine and permanently attached to the structure, and the other two (front and back) removable. The suspension of all panels on isolating strips completely isolates them from the vibration of the main structure.
- (2) The machine stand has been completely redesigned to be manufactured from cast iron integral with the base of the hammer. This unit is completely closed and the hollow inside has been filled with sound absorbing material. The

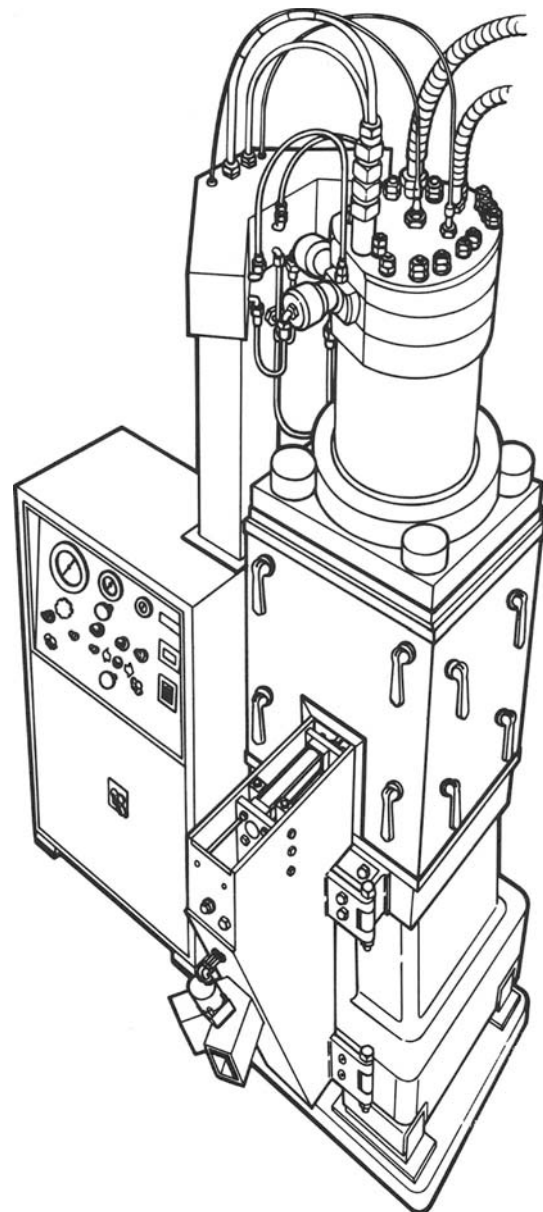


Figure 12. Diagrammatic front view of Petro-Crop Mk.IE showing acoustic shield.

stand and the base can be surrounded by an acoustic shield should this be necessary.

The changes introduced are essentially of a piecemeal nature. They are the only feasible ones (short of enclosing the whole hammer in an acoustic booth) if the hammer is to continue to be suitable for a variety of forming processes other than cropping, a feature which has some definite advantages.

Acoustic considerations played a dominant part in the conception of the Petro-Crop Mk.IE installation. Since this is intended to be a special-purpose machine suitable for cropping only, there was complete freedom as to solutions adopted in the light of acoustic (and dynamic) considerations. The machine was therefore evolved not as a result of piecemeal improvements but on the basis of a fundamental rethinking of the requirements.

From the acoustic point of view, the design of the Petro-Crop Mk.IE hammer differs from the earlier model in the following features:

(1) *Design of moving masses.* With the hammer investigated (figure 1) the major part of the impact noise was generated in the impact zone (die space). One of the reasons for this is the geometrical shape of the moving mass. This consists of a piston and a

mushroom-like forging constituting in the ram/platen, as shown in figure 3(a). As already mentioned, the reason for this design lies in the need for a large platen surface for attaching the bulky tools required for, say, hot forging. However, the shape of this is far from ideal from the vibration point of view. Tests showed that, in spite of its stiffness, during impact the platen executes flexural oscillations and these make a substantial contribution to the sound level generated.

When designing from the acoustic point of view, it is generally desirable that components liable to flexural vibrations should be avoided since these are much more easily set up than extensional oscillations. In view of this the piston/ram/platen assembly was redesigned to take the form shown in figure 3(b). This design is adequate for a special-purpose cropping machine since in this the only tool attached to the face of the ram is a striker to impact on the moving cropping blade.

(2) *Optimization of the moving masses.* It has already been mentioned that in the course of a single cropping cycle, the piston/ram/platen assembly impacts on the moving blade several times and the moving blade impacts on the energy-absorbing piston (figure 3). Although during these multiple impacts

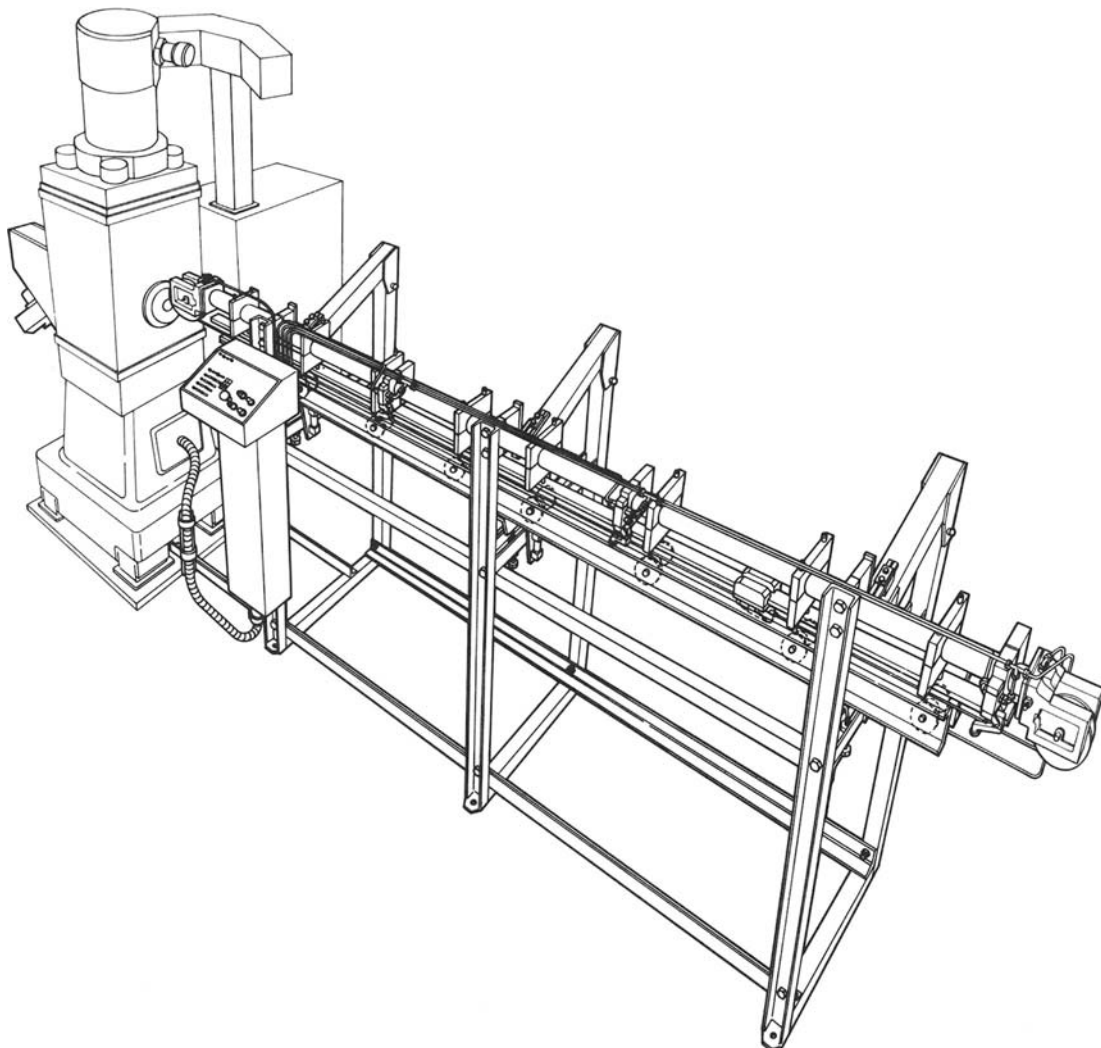


Figure 13. Diagrammatic rear view of Petro-Crop Mk.IE showing feeding mechanism.

the energy transferred diminishes and hence progressively less impact noise is generated, the overall effect is to increase the impact noise level.

It can be shown by elementary considerations that these multiple impacts can be avoided simply by ensuring that the three moving masses are equal. If this is the case then when the piston/ram/platen assembly impacts on the moving blade, the momentum of the former is transferred 100 per cent to the latter. Hence the platen assembly becomes stationary, and is returned to its starting point by the back-pressure in the drive unit. The blade is rapidly accelerated to attain the original velocity of the platen and, after performing the cropping operation, hits the decelerating piston. At that stage all the residue energy left in the blade, that is initial kinetic energy minus cropping energy, is totally transferred to the decelerator. Hence the blade is arrested, and is returned to its original position by springs. The decelerator continues in free flight and its energy is absorbed by appropriate means. Thereafter it is returned to its starting position by springs and the cycle can start afresh.

It should be noted that the optimization of the moving masses has other advantages in addition to a reduction of the impact noise generated. The avoidance of multiple impacts is clearly desirable from the point of view of the life of the components concerned. The efficient energy transfer allows a better utilization of the blow energy available and hence the cropping of larger bar diameters.

(3) *Isolation of impact zone.* The two impact zones, that between the platen and the moving tool and between the moving tool and the decelerator, are contained in air-tight enclosures. These are the first of two acoustic barriers for suppressing the noise emitted. Both these zones are flooded with lubricant required for washing away any debris produced in cropping and for reducing the noise intensity.

(4) *Enclosure of die area.* The whole die area is completely enclosed by sound absorbing/isolating panels. These are easily removed for handling the die. The appearance of the hammer of the installation is shown in figure 12.

(5) *Redesign of stand and base.* The stand/base is

an integral casting which is fully closed and filled with sound-absorbing material. This unit can be surrounded by a further enclosure should this become necessary. The general appearance of the hammer of the machine is shown in figure 12 and that of the bar rack/feeding unit in figure 13.

CONCLUSIONS

The intensity of sound emitted with the particular machine considered could easily have been reduced by enclosing its hammer in a sound-isolating booth. However, this was held to be expensive and cumbersome and therefore a series of alternative means were adopted, including the enclosure of the die area, the use of skirting surrounding the machine stand and the application of a thin oil film on the impacting surfaces. With all these measures an overall reduction of the sound level of the order of 21 dB to about 90 dB (A) can be expected.

Thus with the existing machine a series of piecemeal improvements were required for achieving the desired result. However, with the design of the new improved machine a much more fundamental approach was adopted. This involved an improvement of the dynamics interaction of the moving parts, that is the avoidance of multiple impacts, design changes aiming at an elimination of a 'ringing' of some of the machine components, the use of double sound-isolating enclosures of the impact areas and a flooding of the impacting surfaces with lubricant.

REFERENCES

1. M. K. Das and S. A. Tobias. Dynamics of a HERF cropping machine. *Proc. 15th Int. MTDR Conf.*, Birmingham, Macmillan, 1974.
2. E. Zwicker. Sub-division of the audible range into critical bands. *Institute of Acoustics Society of America* (1961) 33, 248.
3. A. E. M. Osman, M. M. Sadek and W. A. Knight. Noise and vibration analysis of an impact forming machine. *The Design Engineering Technical Conference, Cincinnati, Ohio, Trans. A.S.M.E. Jnl. of Engineering for Industry*, New York, 1973, p. 1.

THE EFFECT OF BLADE WEAR ON NOISE LEVELS OF POWER HACKSAWS

by

G. J. McNULTY*

SUMMARY

The parameter of blade wear on the noise levels of power hacksaw blades has been investigated. The results of the effect of blade thrust force on the wear and its influence on the noise level has been outlined.

The main stages of wear in relation to noise levels and time of cut are given. This leads to recommendations on optimum blade life based on these factors.

INTRODUCTION

In view of impending legislation the problem of industrial noise is of contemporary interest. The papers given at the P.E.R.A. conference¹ recommend that the main means of noise attenuation of machine tools should be either concentrated in the design stages or screens should be erected. Both of the above means may be beneficial but costly. McNulty² has shown that by selection of the forces and blades in power hacksaws the noise output can be reduced. To date there is no published work of the effect attributable to blade wear on the emitted noise of power hacksaws.

The aim of this paper is to show how the noise levels are affected by a carefully controlled wear pattern. Recommendations as to the optimum blade life with respect to the parameters of noise levels and cutting times are given.

EXPERIMENTAL PROCEDURE

The measurements were made on a Wicksteed Power Hacksaw machine. Previous results² have shown that the background noise from the motors and ancillary equipment does not interfere with the blade cutting noise. The ambient background noise is also negligible, as shown in figure 2.

The prime requirement of the instrumentation was to enable the load developed between the blade and the workpiece with the corresponding noise level to be monitored at each position of the stroke. Figure 1 shows the instrumentation used.

The three mutually normal components of the cutting force were measured from the dynamometer clamped on the vice of the machine. The measuring element was a Kistler piezo-electric transducer.

A microphone placed at a set distance from the cutting action was fed to a B and K precision sound-level meter—which was used as a true r.m.s. meter. The d.c. output of this was fed to the first channel of the X-Y plotter. In the second channel the force was

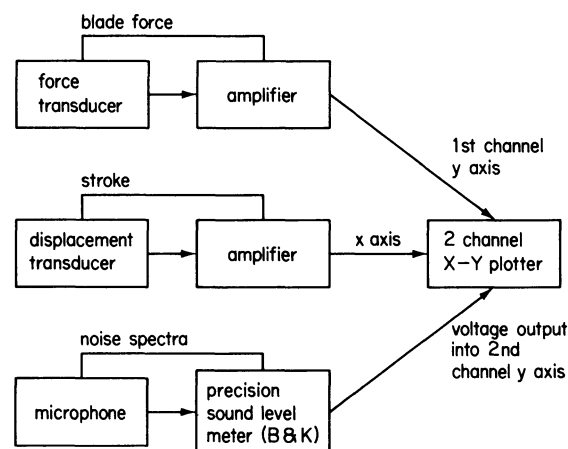


Figure 1. Instrumentation for measuring blade force and noise spectra versus stroke.

recorded corresponding to the sound pressure level. Tape recordings of the blade noise were also made and the later narrow-band frequency analyzed.

The displacement of the blade was measured by a linear transducer and transducer meter.

THREE STAGES OF WEAR IN POWER HACKSAW BLADES

Primary wear

When the incipient stages of this wear are observed the teeth are worn so that a flat is produced at the tip

* Acoustic Service Unit, Department of Mechanical Engineering, Sheffield Polytechnic, Sheffield S1 1WB.

of each tooth and the outer corners of the set teeth are rounded. The effect of this wear is to increase the thrust force needed to achieve a given depth of cut.

Secondary wear

In this region the blade loses height in proportion to the number of cuts made.

Tertiary wear

This region is associated with rapid increase in wear. Under the normal application of thrust loads the blade ceases to cut. The present work revealed a high-pitched rubbing noise during the advanced stages of this wear.

DISCUSSION OF RESULTS

Figure 2 shows the result of a narrow-band frequency analysis. The significance of the spectra is that the pattern is similar for all stages of wear, with peaks of 1 kHz and 8 kHz predominating in each case. Another feature of the spectra is that the sound

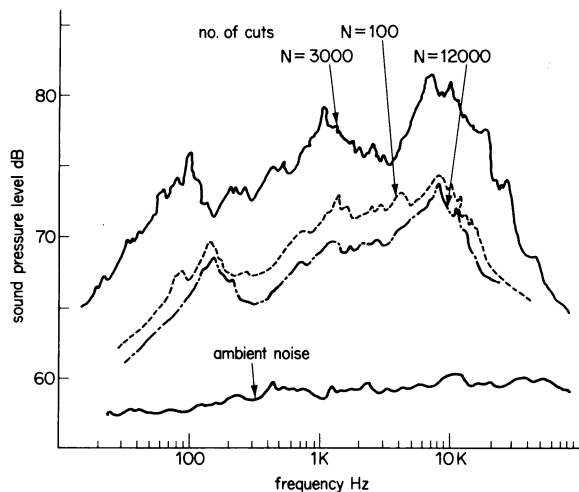


Figure 2. Narrow-band frequency analysis of noise spectra in relation to number of cuts of blade.

pressure level reaches a maximum for a number of cuts then falls steadily. The price which has to be paid for this reduction in noise output is a rapid increase in the time to cut a given specimen as outlined in Table 1.

Table 1 shows results similar to other blades being tested for wear at the present time Taylor³. The reason for the reduction in sound pressure level is that the blades have reached a stage of wear called 'tertiary wear' Thompson⁴. During the early part of wear (primary and secondary) the blade loses height which is proportional to the number of cuts made. The tertiary region is associated with a condition in which the blade does not cut but 'ploughs' through the specimen for a given thrust force. Thus the noise emitted during this stage of wear is reduced as the cutting content is lost, but the result is an increase in the time of cut.

The explanation for the reduction in level is not altogether clear but it may be due to the chips crowding round the blade during the ploughing action, reducing the flexural vibration of the blade. A

TABLE 1 Sound pressure levels and cutting times to cut a specimen EN 44E VPN 309 of 1 in². Mean thrust load 473 N. 68 strokes per min

Specimen piece no. from bar	No. of strokes required to cut specimen	Cutting time (min)	Sound pressure level: dB filtered at 8 kHz (max. peak)
1	335	4.93	72
2	362	5.31	75
3	394	5.19	78
4	420	6.17	78
5	426	6.28	78
6	453	6.66	81
7	476	7.0	81
8	534	7.86	83
9	560	8.23	78
10	590	8.68	78
11	605	8.9	75
12	650	9.58	71
13	672	9.89	71
14	723	10.62	71
15	810	11.81	92

reduction of 5 dB resulted when the blade was artificially damped², offering a similar condition to the above.

The predominant frequencies shown in figure 2 may be due to a combination of several factors as listed below.

- (1) The forcing frequency due to engagement and disengagement of the blade teeth on to the workpiece.
- (2) The natural frequencies of the blade during the cut.
- (3) The dimensions and type of material of workpiece.

Let us consider each in turn.

Forcing frequency

The frequency due to engagement of the teeth is given by

$$f = N\pi S n / 60$$

where N is the rotational speed of the machine = 68 rev/min

S is the stroke of machine = 0.127 m

n is the number of teeth per unit length of blade = 240 teeth per metre.

From above $f = 107$ Hz. For the three curves the first peak falls within the frequency range of 100 Hz and 130 Hz. In the curve for 3000 cuts there is evidence of repeated peaks lying between 200 and 300 Hz, 400 and 600 Hz, 800 and 1000 Hz. These may be harmonics of the fundamental of 107 Hz.

Natural frequency of blade during cut

From the Progress Report⁵ on Hacksaw Research, the transverse vibration of the blade at a pre-set tension of 8 kN is 200 Hz for the fundamental frequency. However, the blade is not likely to vibrate only at its fundamental frequency as the workpiece establishes a varying node along the blade's length. Thus the fundamental frequency when the workpiece is midway along the length is 400 Hz, and the blade may vibrate at multiples of 400 Hz. The combination

of the forcing frequency harmonics and the transverse vibration of the blade may therefore account for the peaks up to 1 kHz.

Dimensions and type of material of workpiece

The noise radiated is a function of workpiece dimensions and damping. The dimensions also affect the forcing frequency as there is a complex pattern of engagement and disengagement of teeth on the workpiece. These points are being studied in present research investigations conducted in the department. It is hoped that this will help to quantify the higher frequency noise levels emitted by power hacksaws (8 kHz).

Figure 3 shows the sound pressure level-stroke for different stages of wear. The figure shows that as the blade wears there is less evidence of ripple during the return stroke. This is accompanied by a reduction in the noise content of the stroke. At 2000 strokes the ripple is reduced to a minimum. The curve suggests

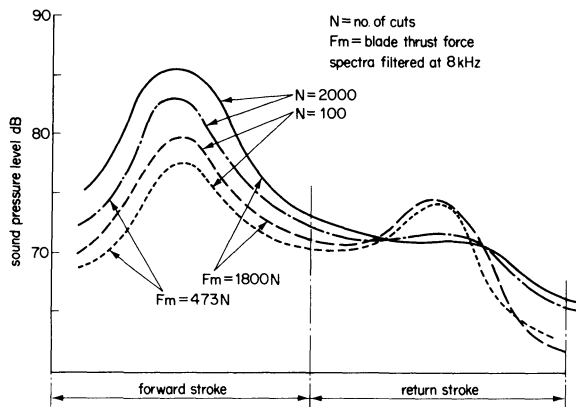


Figure 3. Noise spectra filtered at 8 kHz showing effect of blade thrust force and wear over stroke.

that the ripple is solely dependent upon wear. The number of strokes is associated with the tertiary wear region and the explanation given concerning increased damping at this region is partially substantiated by figure 3.

Another feature of figure 3 is the variation of the sound pressure level for two different thrust forces. A study of thrust on S.P.L. has been completed in Ref. 2. The main characteristic to note here is the reduction in S.P.L. with thrust force. However, the return-stroke ripple shows no dependency on the magnitude of the force.

RUBBING NOISE

Table 1 shows that after tertiary wear, there is a point where the blade reaches its lowest efficiency in cutting. This is the case after fifteen specimens in the present test where the time to 'cut' is rapidly increased. The S.P.L. has reached an offensive level of 92 dB at 8 kHz. It should also be noted that the screech persists throughout the cut.

CONCLUSIONS

The work on noise reduction in power hacksaw blades due to wear can be summarized from the investigation as follows.

- (1) A narrow-band frequency analysis (6 per cent) revealed that the predominant S.P.L. peaks occur at high frequencies. In the present case 8 kHz.
- (2) The reduction in thrust force reduces the S.P.L. at the peaks, but the level of the high thrust force shows the absence of return-stroke ripple. This ripple can add a considerable amount to the sound energy emitted over a large number of strokes. In this work the difference of 10 dB was recorded per stroke at 8 kHz. A cutting operation may take 100 strokes, thus the accumulated emitted sound energy is considerable.
- (3) The increase in wear causes an increase in the S.P.L. up to the limit of secondary wear. Thereafter during secondary wear the level falls. In this region of wear there is also an absence of ripple in the return-stroke noise spectra. A dangerous high peak in the noise spectrum is prevalent at a point in the tertiary wear region, this is caused by the rubbing of the blade on the workpiece. The level recorded in this work was 92 dB at 8 kHz, an increase of 20 dB over the previous cut.
- (4) The time to cut a given specimen increases by a factor of between 2 and 3 as the wear progresses from primary wear to tertiary wear.

RECOMMENDATIONS

In order to achieve minimum S.P.L. for optimum wear conditions it is recommended that the following points be observed.

- (a) When the blade reaches its tertiary wear region, (this is observed by its lack of ability to cut) there is a point some time after when there is rubbing action on the workpiece. Such a condition causes an offensive squeal for the prolongation of the cut. The blade should be discarded immediately tertiary wear is observed.
- (b) High thrust forces are advantageous, both to speed up the cut and to reduce return-stroke ripple noise.

ACKNOWLEDGMENT

The author is grateful to Mr P. J. Thompson for his information on Hacksaw Blade Wear a current research project conducted at Sheffield Polytechnic Mechanical Engineering Department.

REFERENCES

1. *Proceedings of Conference on Industrial Noise*, P.E.R.A., September, 1972.
2. G. J. McNulty. The Noise Emitted from Power

- Hacksaw Machines. *Proc. 14th Int. M.T.D.R. Conf.* September, 1973.
3. R. Taylor. Private Communication. Sheffield Polytechnic Department of Mechanical Engineering, January 1974. Investigations of Power Hacksaw Performance.
 4. P. J. Thompson. Power Hacksaws. Paper to be published, 1974.
 5. Sheffield Polytechnic Progress Report on Power Hacksaw Research April to July 1972. Department of Mechanical and Production Engineering.

DISCUSSION

Several points were raised at the conclusion of the paper, and the following four points summarize these questions:

- (1) What is the noise level of a hacksaw averaged over an eight-hour day?
- (2) What are the noise levels of a number of hacksaws working together? Is the contribution of mechanical handling noise significant?
- (3) Can tertiary wear be observed from a noise reading?
- (4) The noise level of a single hacksaw is small; is it worth the effort, therefore, to reduce it?

The discussion which followed the above points is summarized below:

- (1) The sound pressure level recorded on a noise dosimeter for a single hacksaw working through all stages of wear for several days was 83 dBA. This is known as the L_{eq} of the equivalent continuous sound pressure level averaged over an eight-hour day.
- (2) A slide was shown of ten hacksaws working together. The sound pressure level recorded for this situation was 88 dBA averaged from five

readings. The mechanical handling noise was included in this. However, it was small, as most machines were cutting when each reading was taken. Furthermore, the specimens being cut were 1.0 in diameter which contributed a negligible proportion to the overall noise level.

- (3) Tertiary wear can be observed from a subjective observation of the noise level, especially during rubbing noise. In the paper it is shown that there was an increase of 20 dBA in this wear region; such an increase is easily detected by the ear.
- (4) The philosophy presented in the paper (and also in ref. 2) is one in which a study of the noise levels is made against variation in machine parameters. The purpose of this study is to reduce the noise levels in the hacksaws by optimizing cutting parameters instead of resorting to expensive shielding. Moreover, although the emitted noise level of a hacksaw is relatively small it is desirable that it should still be reduced. Every noise-producing element in an area, if thus treated, would contribute to the overall diminution in emitted noise energy.

GRINDING AND SURFACE TOPOGRAPHY

OPTIMIZATION OF A GRINDING PROCESS AND CRITERIA FOR WHEEL LIFE

by

G. TRMAL* and H. KALISZER*

INTRODUCTION

The need for a more rational approach in determining the required grinding conditions and selecting suitable grinding wheels is becoming more and more urgent. In most other machining processes various optimizing techniques are already widely used to establish the best cutting conditions. One of the reasons that the grinding process still remains more of an art than a science is the difficulty of determining one of the basic wheel parameters—the wheel life.

Under the action of the cutting forces and the high temperatures generated during grinding, the grinding wheel loses its original geometrical form and its cutting ability. As a result the surface roughness and waviness of the ground component deteriorates, and thermal damage may also appear.

In order to maintain the correct geometrical form and cutting ability of the grains, the wheel requires periodic dressing. This however reduces the overall output, since dressing is a time-consuming operation. Such reduction in performance is particularly evident with materials which are difficult to grind, because the wheel life may be shortened fifteen to twenty times by these materials and the cost of dressing may reach 60 to 70 per cent of all machining costs per operation.

The wheel life reflects not only the physical aspects of grinding but also the economic and technical factors. It is for this reason that the wheel life is considered to be one of the basic parameters describing the characteristics of a grinding wheel.

The unproductive time required for dressing forms a significant proportion of the total grinding time. The cost of dressing contributes substantially to the total grinding costs, in spite of the fact that the cost of wheels is relatively small, at least in the widely used straight plunge grinding operations.

The purpose of this paper is to analyse various criteria of wheel life and to optimize the grinding conditions on the basis of the minimum overall time. The suggested method is basically applicable for straight plunge grinding of steels at conventional grinding speed and with medium finish requirements ($Ra = 0.3$ to $0.7 \mu\text{m}$).

EXPERIMENTAL PROCEDURE

All grinding tests were carried out on a Studer RHU 500 external grinder commonly used for finish machining.

The selected material and grinding conditions are given in Table 1.

TABLE 1 Material and grinding conditions

	Grinding wheel	Workpiece Steel 42 HRC
Speed	30 m/s	20 m/min
Diameter	300 mm	100 mm
Width		14 mm

The dressing was carried out with a single point diamond. The diamond lead (axial distance that the diamond moves per wheel revolution) was selected so as not to introduce rapid changes in surface roughness during the initial stage of grinding. Such dressing conditions were defined as optimum (figure 1).

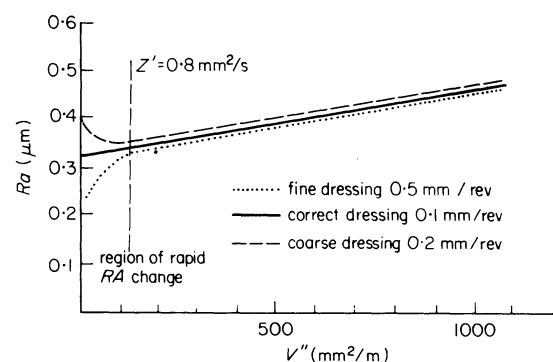


Figure 1. Surface roughness as a function of dressing conditions.

The parameter Z' used in the graph represents the amount of material removed per unit wheel width and per unit of time (mm^2/s) was first introduced by Hahn¹. The parameter V'' , which describes the volume of material removed by 1 mm of wheel width and 1 m of wheel periphery (mm^2/m), was initially

* Department of Mechanical Engineering, University of Birmingham

introduced by CIRP⁷. An advantage of both these parameters is that they express the volume of material removed, regardless of the workpiece speed and the wheel and workpiece dimensions.

The duration of the sparking out period was determined on the basis of maximum improvement in surface roughness. It was found that after six workpiece revolutions further improvement in surface roughness was not very significant. This procedure was thus considered acceptable and was adopted.

Some typical experimental results from grinding tests are plotted in figure 2. As the main output

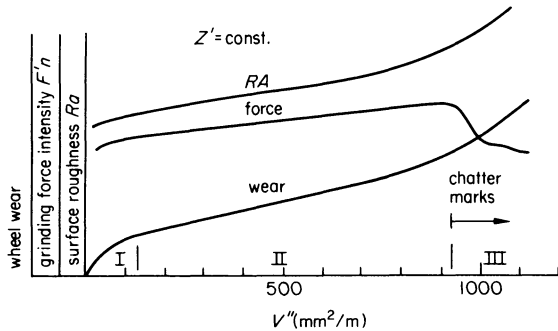


Figure 2. Wheel wear, grinding force and surface roughness as a function of volume of material removed.

parameters besides the wear rate, the following were plotted: surface roughness (*Ra* value) and normal force intensity *F'n* which can be defined as the normal component of the grinding force per 1 mm wheel width.

For most wheel and material combinations the chatter-free zone gives a reasonable linear relationship between the output parameters (*Ra* and *F'n*) and the material removal parameters (*Z'* and *V''*). On the basis of the results shown in figure 3, the following empirical relationships have been established.

$$Ra = A + Z' \times B + \frac{V''}{1000} (C + Z' \times D) \quad (1)$$

$$F'n = E + Z' \times F + \frac{V''}{1000} (G + Z' \times H) \quad (2)$$

In Equations (1) and (2) *A, B, C, D, E, F, G* and *H* represent constants which depend upon the various

wheel-material combinations used in the grinding tests.

NATURE OF WHEEL WEAR

The wear of grinding wheels depends upon the grinding-wheel/work material combination as well as the geometric and mechanical conditions existing in any particular grinding operation.

The curve expressing the volume of wheel wear against the volume of metal removed is quite similar in shape to a wear curve for any cutting tool. As may be seen from figure 2, it forms three distinct regions. In the first region, taking place during the initial contact between the wheel and workpiece after dressing, the rate of wear is very high. This is due to the fact that some of the grains which are shattered or cracked by the dressing action are quickly removed.

In the second, or stabilized, region the wear rate remains approximately constant. The grinding force slowly increases with time and so does the surface roughness. The increase in surface roughness indicates a decreasing number of cutting edges on the wheel surface. The increase in the cutting force can be explained by the attritious wear of the cutting edges. This produces an unfavourable cutting geometry of individual grains and therefore an increase in the forces.

If grinding is continued into the third region, self-excited vibrations usually appear and more extensive fracture wear takes place. When a large percentage of active grains is fractured, the number of cutting edges decreases causing a sharp increase in the wear rate. At this stage the surface topography deteriorates and in most cases the grinding force tends to decrease.

FACTORS DETERMINING THE WHEEL LIFE

From the practical point of view the useful wheel life ends when the dimensional and topographical parameters, or surface integrity, exceed the permissible

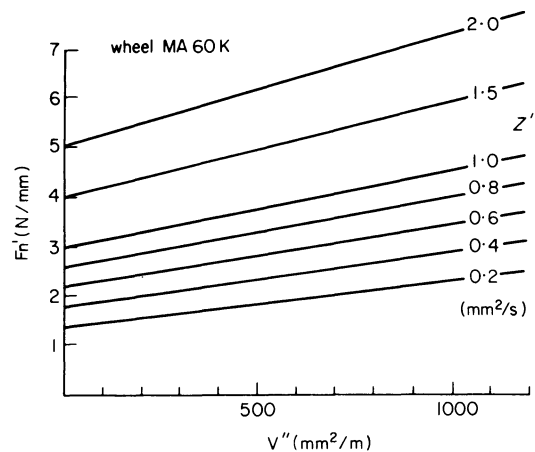
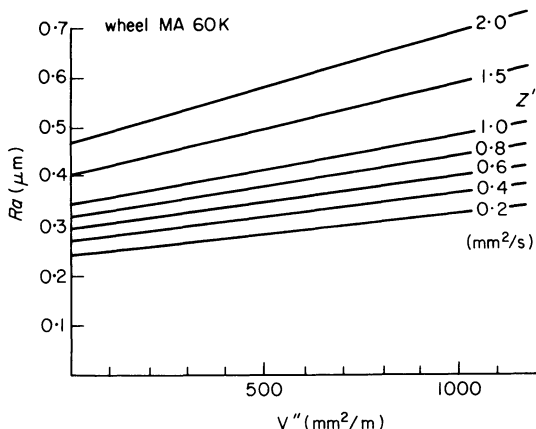


Figure 3. Surface roughness and force intensity as a function of rate of material removal and material removed.

limits, so that wheel redressing is required. Thus the wheel life is governed by the following factors:

- (1) the surface roughness,
- (2) the grinding force,
- (3) the radial wheel wear,
- (4) chatter marks.

Surface roughness

As follows from figure 3, the workpiece roughness is a function of the rate of material removal (Z') and the volume V'' of metal removed since the last redressing. On the basis of the data plotted on the graph it can be seen that the required surface roughness can be achieved with differing values of Z' . A higher removal rate is possible but it will involve shorter wheel life and more frequent dressing, increasing the ratio of dressing time to grinding time.

The same surface roughness can be produced with a lower material removal rate but longer wheel life.

The total grinding time for removing a unit of material volume per wheel unit width can be calculated as follows:

$$T_T = T_g + T_d \text{ (s/mm}^2\text{)} \quad (3)$$

where T_g is the grinding time, given by

$$T_g = \frac{1}{Z'} \quad (4)$$

and T_d is the dressing time, given by

$$T_d = \frac{T_{DR} \times 1000}{V_d'' \times D_s} \quad (5)$$

Here T_{DR} is the time required for a redressing operation (s),

V_d'' is the volume of metal removed by 1m wheel periphery 1mm wheel width and between two consecutive redressing (mm^2/m),

D_s is the wheel diameter (mm).

From Equations (1), (3), (4) and (5) the total grinding time can be computed on the basis of the selected wheel life (volume of metal removed between redressing) for various values of Ra and the chosen type of dressing.

To the first approximation the required time for dressing can be assumed as follows:

For the dressing tool fixed on the wheel head $T_{DR} = 1$ min.

For the dressing tool fixed on the work table, and considering that some re-adjustments in the position of the dresser will be necessary, $T_{DR} = 3$ min.

If a resetting in the position of the dresser must be performed before each dressing operation, the time for redressing must be increased to $T_{DR} = 6$ min.

On the basis of figure 4, where the total grinding time T_T is plotted as a function of volume of metal removed between redressing operations, the optimum wheel life can easily be determined. For fine grinding ($Ra = 0.35 \mu\text{m}$) the optimum wheel life which will fall within the minimum values of T_T will vary from $250 \text{ mm}^2/\text{m}$ to $400 \text{ mm}^2/\text{m}$ of metal removed. For medium grinding ($Ra = 0.5 \mu\text{m}$) the optimum wheel life will fall within V'' range varying from 400 mm^2 to

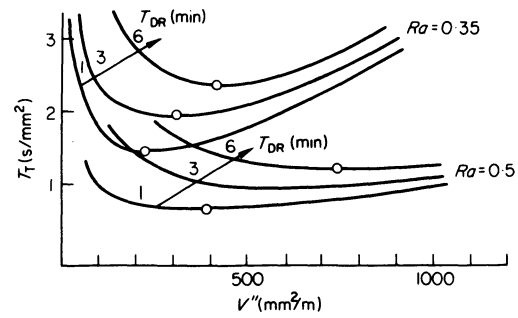


Figure 4. Total grinding time as a function of wheel life for Ra as a limiting factor.

$800 \text{ mm}^2/\text{m}$. The higher values of V'' will correspond to a longer time required for the dressing operation. When V'' and Ra are known the corresponding values of Z' can be selected from figure 3.

Grinding force

As mentioned before, for many grinding operations the grinding force should be considered as a significant constraining factor. This is particularly relevant in cases where any increased forces generated during grinding may be responsible for an unacceptable error of size, position or form. In addition, an excessive force intensity may be responsible for high residual stresses, thermal defects and so on. The full quantitative determination of these aspects of wheel life presents considerable difficulties and will require more investigations. To a first approximation, however, the limiting value of force intensity was assumed to be $F'n \approx 3 \text{ N/mm}$ for fine grinding and $F'n \approx 5 \text{ N/mm}$ for medium grinding.

On the basis of these values the total grinding time for an optimum wheel life can be determined in a similar way as for surface roughness.

It can also be seen from figures 4 and 5 that the

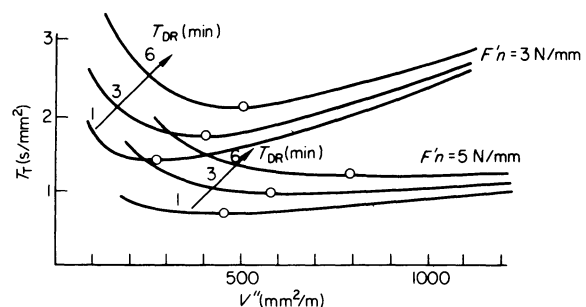


Figure 5. Total grinding time as a function of wheel life for force intensity as a limiting factor.

optimum wheel life which corresponds to the minimum time does not differ significantly for the output parameters Ra or $F'n$. In both these cases the curves in the minimum region are almost flat and small differences in V'' are not very significant.

When values of $F'n$ and V'' are known, the corresponding values of Z' can be determined on the basis of equation (2) or figure 3.

Radial wheel wear

The surface roughness of a workpiece ground with a

freshly dressed wheel is superimposed upon the workpiece macroform generated by the wheel profile, which in turn is created by the relative movement of the wheel and the dresser. With a correct dressing action the wheel cutting surface should be nominally straight, and this should lead to a straightness of the workpiece surface in the axial direction. However, during the wheel life the straightness of the wheel cutting surface cannot be fully maintained, due mainly to non-uniform wheel wear³. In general the irregularity of the workpiece profile 'across the lay' cannot be detected because most of roughness measuring devices use cut-off values (filter frequency) to suppress the low frequency signal generated by the waviness.

Our experience is that after the radial wheel wear has approached 15 μm the workpiece waviness cannot be maintained within the usually acceptable limits. It is for this reason that a radial wear of 15 μm has been adopted as a limiting value for medium plunge grinding. Such wheel wear will be defined as a limiting radial wear.

Self-excited vibration

If self-excited vibrations develop during grinding, their detrimental effect upon surface topography (chatter marks) will inevitably limit the period of wheel life. Although chatter marks can usually be distinguished by a naked eye they are seldom detectable by most surface-roughness measuring devices. Their formation on the workpiece periphery begins after the self-excited vibrations have reached a certain advanced level. This critical level, which is reached after the wheel and workpiece contact is interrupted by vibratory motion, is then considered as limiting the period of wheel life. The wheel life which is determined by self-excited vibration will be defined as chatter-limited wheel life.

OPTIMIZATION OF GRINDING CONDITIONS

On the basis of the experimental results discussed earlier an optimum wheel life curve can be drawn as shown in figure 6. This curve, which gives values of V_d'' as a function of Ra , can be obtained by superimposing the Z' lines from figure 3 with the results from figure 4. As can be seen from figure 6 for any given value of Ra the curve determines the optimum

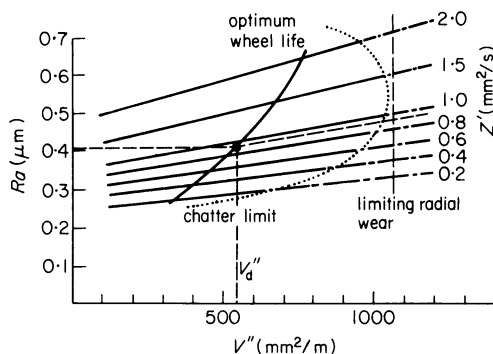


Figure 6. Optimum grinding conditions (Z' and V'') as a function of surface roughness.

grinding conditions, that is the rate of material removal Z' and the optimum wheel life V_d'' .

The other curves also show the 'radial wear limited wheel life' and the chatter-limited wheel life. It follows from figure 4, which gives the total grinding time as a function of V_d'' , that there is no justification for reducing the rate of removal and extending the wheel life beyond the economical volume of metal removal indicated by V_d'' in figure 6.

To optimize the selection of wheel characteristics and other grinder parameters a computer program was written for the IBM 1130. The flow chart of the optimization procedure is shown in figure 7.

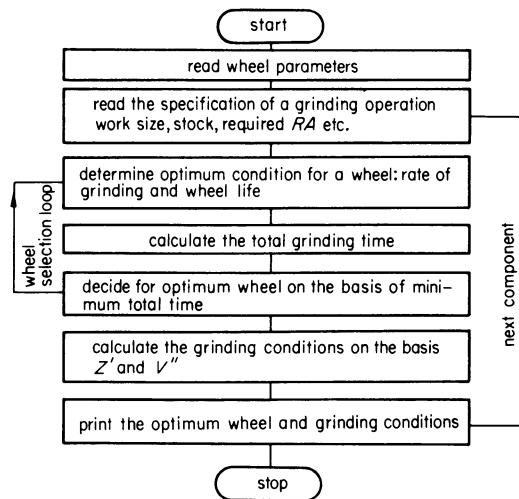


Figure 7. Flow chart for optimization of grinding.

input parameters

work dia	width	stock	wheel dia	max RA	force	force int
mm	mm	mm	mm	um	N	N/mm
35.00	20.0	0.40	300	0.40	120.0	4.0

output parameters

optimum wheel	WA	80	m 6 V
total grinding time	0.41	min	
number of components	18		
power required	1.1	kW	

Z	VF	sparking	dressing speed	depth	limit	NW
SQmm/s	mm/s	sec	mm/min	mm	RA	rpm
0.96	0.52	2	190	0.020	RA	180

Figure 8. Computer print-out of optimum grinding conditions (VF = input feed rate, NW = workpiece speed).

As can be seen from figure 8, the computer output gives all the necessary technological and economical data for the required grinding operation.

CONCLUSION

An attempt has been made to develop a method for optimizing the wheel characteristics and grinding conditions by considering the relevant technological and economical factors.

At present this method is applicable to steel materials ground at a conventional wheel speed (30 m/s), but it is hoped that in the next stage other important factors such as higher wheel speed, materials which are difficult to grind, surface integrity and so on, will be included.

The described optimization procedure is now tested in practical industrial conditions by selecting a mass-produced car component which is difficult to machine.

ACKNOWLEDGMENT

The authors wish to thank Professor S. A. Tobias for providing facilities for the work and the Science Research Council for financial support. They also wish to thank Mr M. Singh whose experimental work was used in this paper.

REFERENCES

1. R. P. Lindsay and R. S. Hahn. On the basic relationship between grinding parameters. *Annals of CIRP* (1970) **18**.
2. H. Opitz, W. Ernst and K. F. Meyer. Grinding at high cutting speeds. *6th MTDR Conference*, Manchester (1965).
3. H. Kaliszer and G. Trmal. Generation of surface topography on a ground surface. *14th MTDR Conference*, Manchester (1972).
4. H. Kaliszer. Analysis of chatter vibration during grinding. *Advances in Machine Tool Design and Research* (1971).
5. S. Malkin and R. B. Anderson. Active grains and dressing particles in grinding. *Proc. Int. Grinding Conference*, Pittsburgh (1972).
6. J. Peters. Economic selection of grinding wheels. Introduction of wheel efficiency parameter. *Proc. Int. Grinding Conference*, Pittsburgh (1972).
7. M. C. Shaw, A. Decneut, R. Snoeys, W. König and R. Bierlich. Unpublished materials of CIRP.

THE COLLECTION OF DATA FOR THE ASSESSMENT OF A GRINDING WHEEL DRESSING TREATMENT

by

E. J. PATTINSON* and J. LYON*

SUMMARY

Profilographs produced from grinding-wheel data which was collected by a specially developed chisel-edge stylus are compared with those obtained by a conventional pointed stylus. For a number of dressing feeds chisel-stylus profilographs are shown to exhibit clear evidence of helical grooving on the wheel surface and can be used as a visual indication of the permanence of a dressing treatment. Experiments are described in which dressing feed, dressing-tool geometry, and the number of passes of the dressing tool are varied. The relative importance of these parameters is discussed.

INTRODUCTION

The work described in this paper is part of an investigation which attempts to describe characteristics of a grinding-wheel surface and then to associate these characteristics with known attributes of the corresponding ground work surface. The primary variable throughout, has been the dressing given to the wheel prior to use; dressing has so far been restricted to the use of a single-point diamond. The problem has been to identify characteristics of the wheel surface with this dressing treatment. One possible approach to the problem is to collect the maximum amount of information from the surface of the grinding wheel and then attempt to filter out unwanted data by some form of mathematical treatment. It should be remembered, however, that a grinding wheel is composed of abrasive grits which are randomly spaced and of random geometry and, in addition, less than one-half of the volume of the wheel is attributed to the abrasive. The remaining half contains the bond, but more important, a substantial amount of voids. It will be appreciated that the profiles taken from the surface of grinding wheels will give information concerning the bond, the voids and the grit and any subsequent mathematical filtering will of necessity have to differentiate between these three elements.

For the purpose of this investigation the collection of data from the surface of a grinding wheel is an attempt to obtain some measure of the potential of the wheel surface as a metal removal tool, and since the investigation is primarily concerned with dressing the data should give some indication of the effectiveness or the permanence of the dressing treatment. In short, the stylus used has to be able to differentiate between the inherent voids in the wheel profile and

the semblance of the helical groove left by the dressing tool as it is traversed across the face of the rotating wheel.

The chisel stylus

An alternative to the mathematical filtering of data after its collection is to mechanically filter the data during collection. This can be done by means of a chisel-shaped stylus used instead of the more conventional pointed stylus. In use the chisel stylus is propelled across the wheel surface with its chisel edge perpendicular to the direction of movement.

A number of chisel styli were developed to investigate a range of grinding wheels. Wheels of a grading sieve size up to 0.014 in diameter and 50 per cent grit content were to be investigated. It was also considered that craters produced as a result of the removal of two neighbouring grits were to be bridged by the stylus, which meant a theoretical minimum chisel width of about 0.070 in. After making allowances for a certain tolerance in grit size and grit spacing as suggested in the literature^{1,2} a chisel edge width of 0.090 in was decided upon.

The work described in this paper is based on data collected by a chisel stylus of 0.090 in width, 60° apex angle and 0.001 in tip radius. In addition data collected by a 90° pyramid stylus of 0.001 in tip radius is presented for comparison.

Experimental equipment

The grinding wheel used throughout this work was of a plain cylindrical type and composed of white alumina abrasive of 100 grit size held together by a vitrified bond (WA 100 KV). The wheel dimensions were 8 in \times $\frac{3}{4}$ in \times 2 in. The grinding machine used

* Department of Mechanical Engineering, University of Salford, Salford, Lancs.

was of the surface grinding type with an additional attachment to allow known uniform dressing feed rates to be applied to the wheel. No cross-feed was used while grinding. Prior to each dressing test the wheel was prepared in such a way that the effects of previous use were completely removed and the surface of the wheel was left in a uniform condition. Details of this pre-dressing treatment have already been reported³.

The data collected by means of a stylus was obtained in the form of a profilograph as output from an inductive type transducer and rectilinear recorder. The two types of stylus described above were used, the 90° pointed pyramidal diamond and the specially developed hardened-steel, chisel-shaped stylus. Details of the chisel stylus are shown in figure 1.

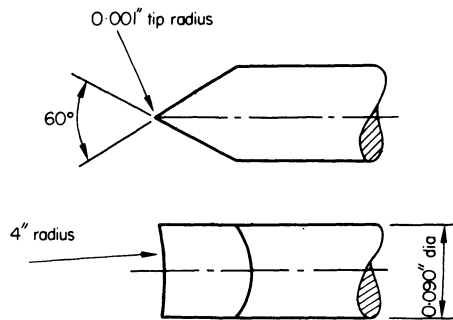


Figure 1. The chisel stylus.

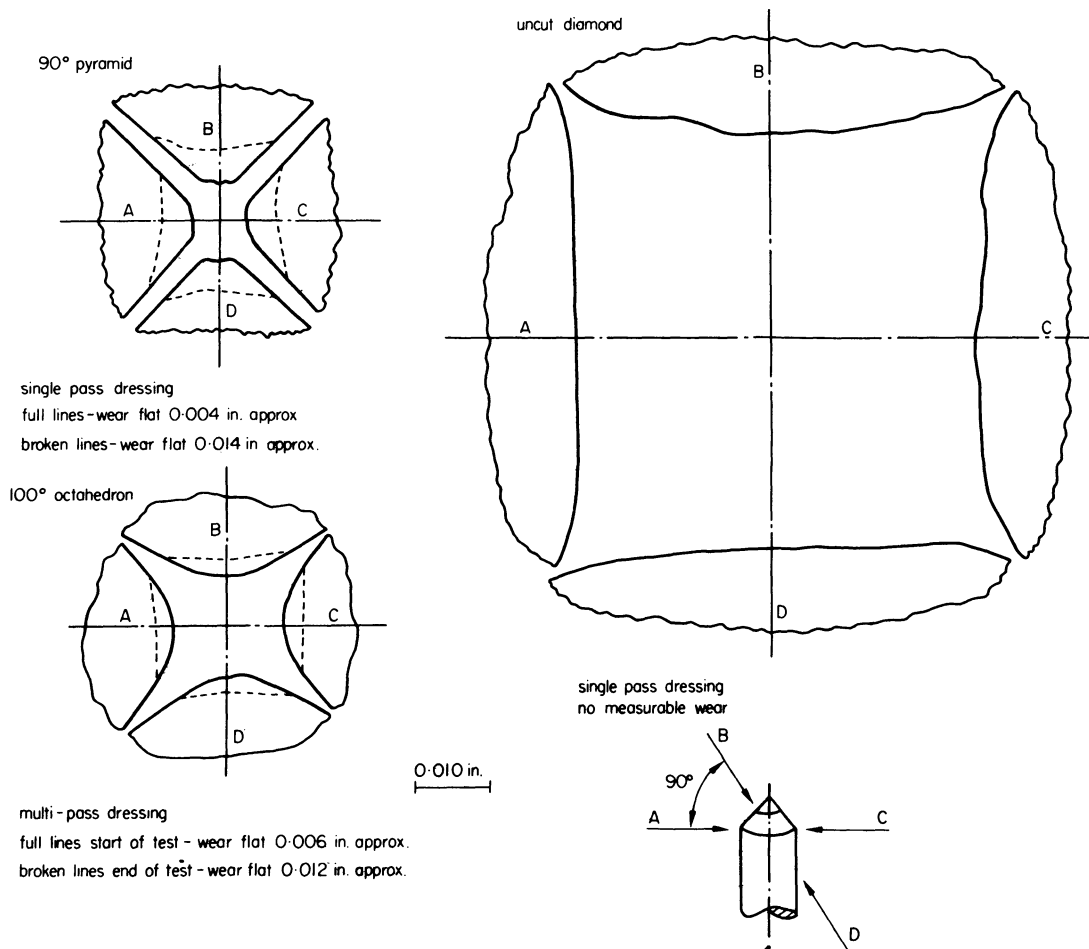


Figure 2. Profiles of the diamond-dressing tools.

The dressing tools used were a 90° pyramid diamond in two distinct states of wear, an uncut diamond of indefinite geometry with a very large wear flat, and a reasonably sharp octahedral diamond. The pyramid and uncut-diamond tools were used for single-pass dressing treatments only, the octahedral diamond was used for multi-pass dressing treatment. The geometry of each dressing tool used is shown in figure 2.

The grinding operation was interrupted at intervals when the wheel was used to grind a thin steel blade. The reproduction of the wheel's worn, stepped profile on the blade made it possible to measure the radial wheel wear during each interval. The height of the step on the measuring blade was determined in the form of a profilograph by making use of the inductive transducer described above for surface profile assessment.

A stereoscopic microscope was used to monitor the collection of data from the wheel surface.

TEST CONDITIONS AND EXPERIMENTAL RESULTS

Dressing feeds of 0.001, 0.010 and 0.017 in per revolution of the grinding wheel were used for each of the single-pass and multi-pass dressing treatments. The results of the wheel-wear measurements are

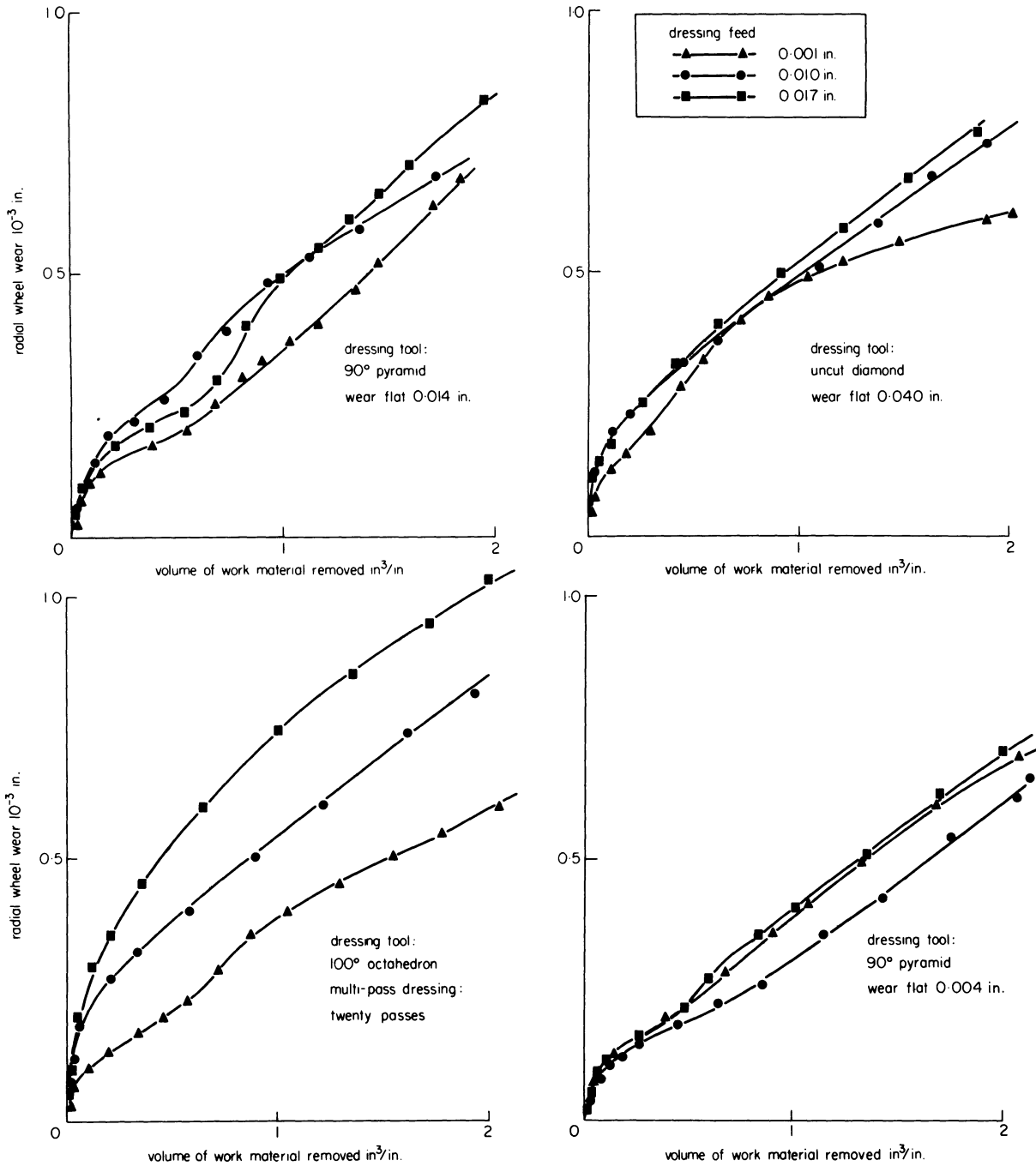


Figure 3. Wheel-wear characteristics.

presented in figure 3 in the form of graphs of radial wheel wear plotted against the volume of work material removed. Profilographs of the grinding-wheel surface were taken after a nominal amount of work material had been removed by grinding. These profilographs are presented in figure 4, based on data collected by both pointed and chisel stylii. In addition, for the single-pass dressing treatments, profilographs were obtained from data collected by both the stylii at three stages of wheel wear; these traces are presented in figures 5 and 6.

Finally the chisel stylus was used to collect data from the wheel after, 4, 8, 12, 16 and 20 passes of the octahedral diamond. The profilographs obtained, in each case for a freshly dressed wheel, are presented in figure 7.

A summary of the experimental conditions are given in Table 1.

TABLE 1

Grinding wheel	
Type	WA 100 KV 7G
Size	8 in \times $\frac{3}{4}$ in \times 2 in
Dressing	
Type	Single-point diamond
Infeed of the tool	0.001 in
Presentation of the tool	10° dragging with respect to the wheel radius
Wheel speed	4000 ft/min
Grinding	
Type	Surface
Wheel speed	5000 ft/min
Workspeed	60 ft/min
Wheel depth of cut	0.00025 in /pass
Crossfeed	None
Workpiece material	
Composition:	carbon 1.0%; manganese 1.3%; silicon 0.3%; chromium 0.5%; tungsten 0.5%; vanadium 0.2%;
Hardness	240 VPH

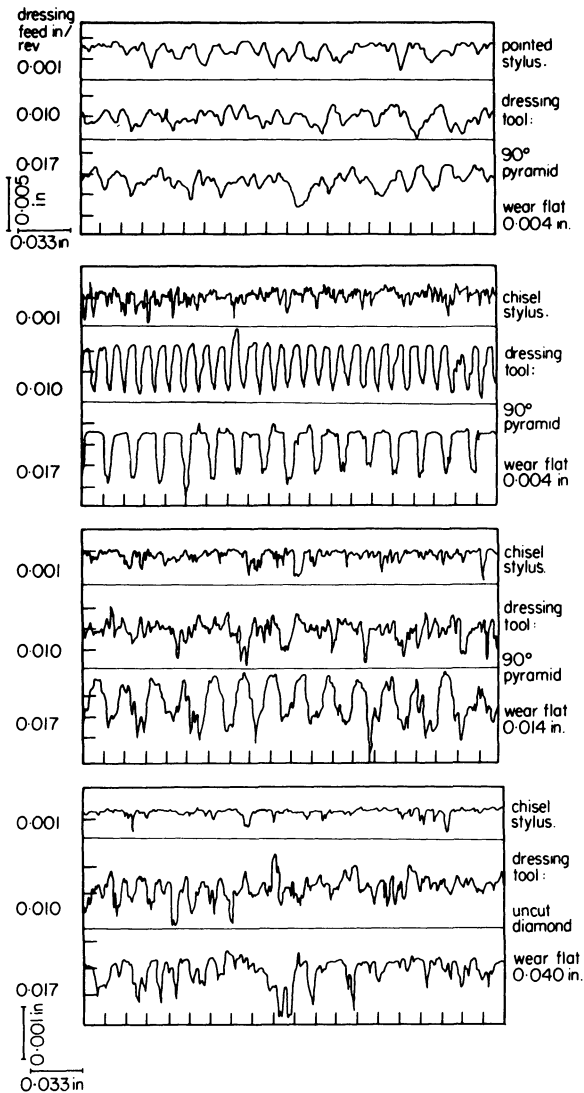


Figure 4. Profilographs of freshly dressed grinding-wheel surfaces.

DISCUSSION OF RESULTS

Chisel-stylus data

When discussing the collection of data by chisel styli from rough surfaces it is always useful to keep the correct scale of the profile in mind⁴, and also to obtain a realistic view of the scope of the chisel stylus as a profile measuring instrument. Figure 8 shows three chisel-stylus profilographs which have equal horizontal and vertical magnification. An outline of the appropriate dressing tool has been superimposed over the profile to the correct scale.

It is always useful to look at a profile which is not distorted with a high vertical magnification prior to turning to the remainder of the results which, of necessity, are presented in the conventional manner. Figure 8 shows immediately the very modest impression and gradient associated with the profiles of dressed grinding wheels and also how consistent and realistic the impression of the diamond appears in the profile produced from data collected by a chisel stylus. It will be realized that the inherent voids and the craters which are produced as a result of wholesale grit removal will not be shown up by this stylus.

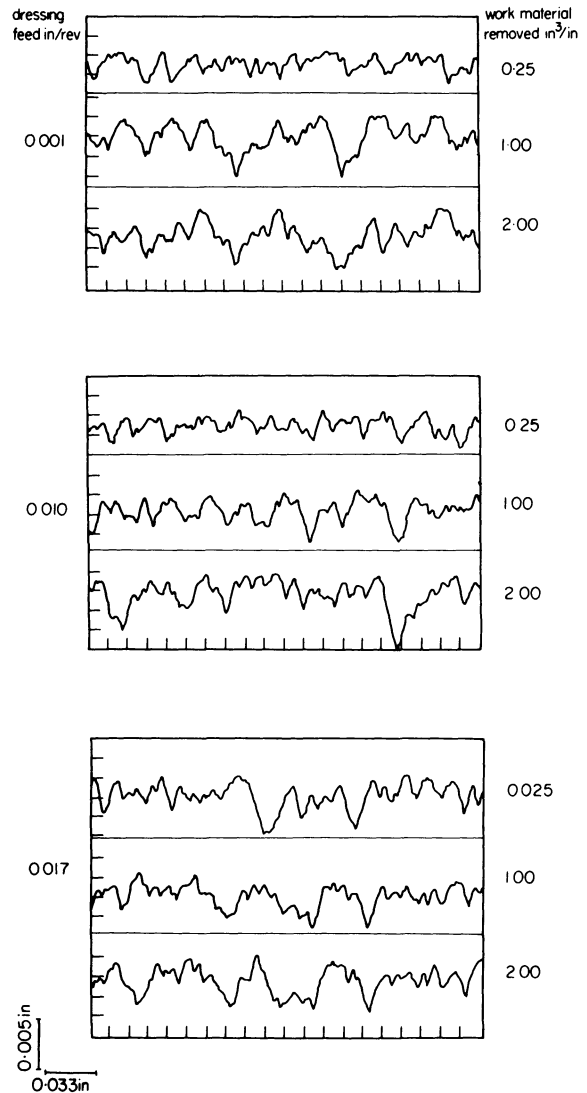


Figure 5. Pointed-stylus profilographs at various stages of wheel wear.

On the other hand, the lateral dimensions of peaks will be enhanced by a chisel stylus since any peak which occurs within the whole width of the chisel edge will be accounted for to its full height.

Wheel surface data

Figures 4, 5 and 6 contrast profilographs obtained by the chisel stylus with those obtained by the more conventional pointed stylus. The profilographs obtained by the two styli have visually very little in common. A chisel stylus profilograph exhibits a regular feed pattern provided the feed is greater than the wear flat on the diamond (figure 4), and this feed pattern is still visible after an amount of wheel wear has occurred equal to one-half the depth of cut of the diamond dresser (figure 6). No semblance of a regular groove is found in the profilographs obtained by the pointed stylus and virtually the only visual difference which can be seen between the various profilographs of figure 5 is an increase in peak-to-valley height with dressing feed.

The dressing feed, however, can still influence chisel stylus profilographs at values for less than those necessary to produce a regular feed pattern on the

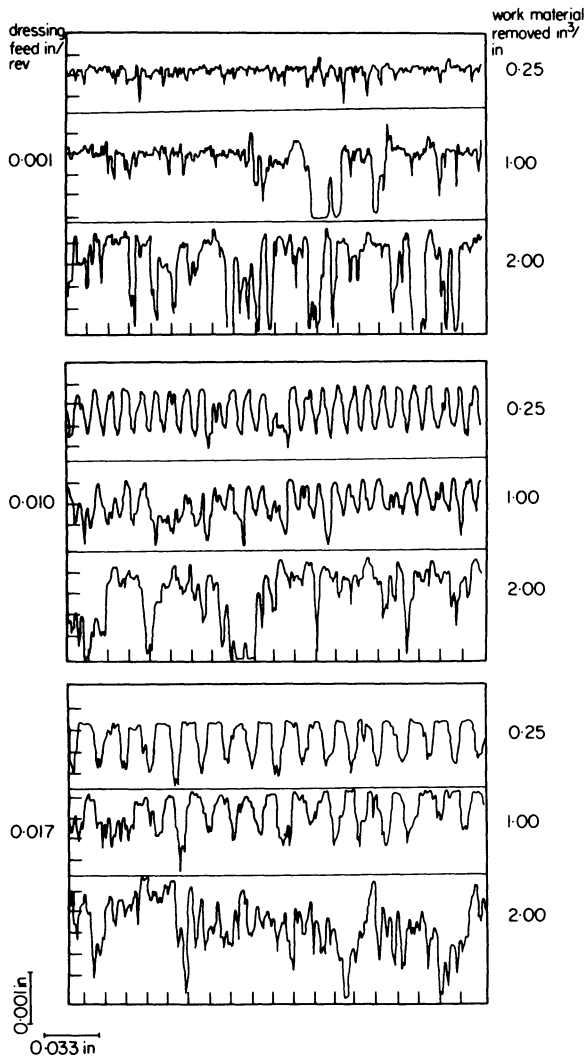


Figure 6. Chisel-stylus profilographs at various stages of wheel wear.

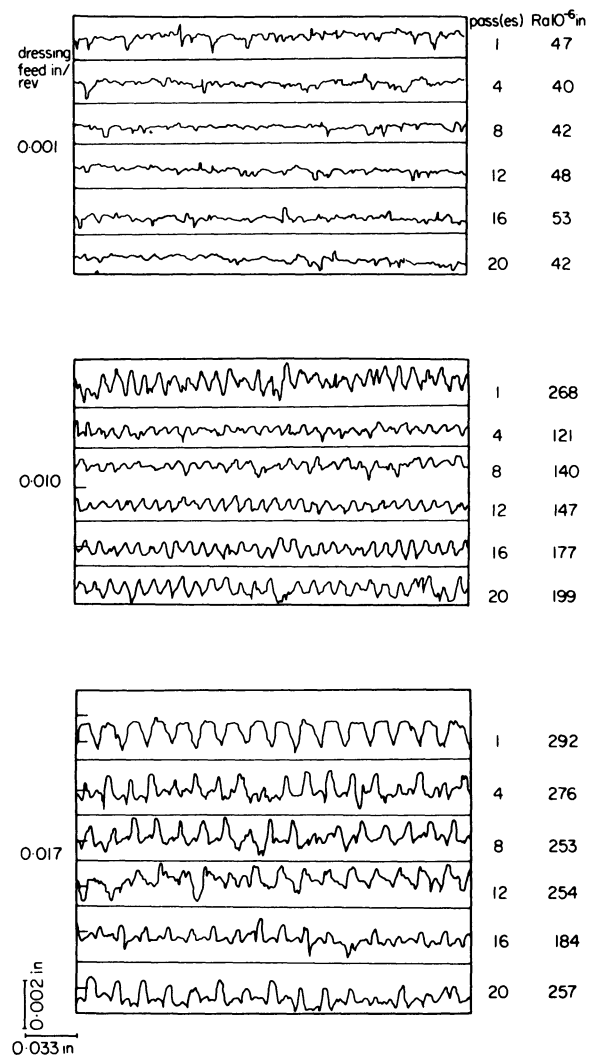


Figure 7. Profilographs of grinding-wheel surfaces after multi-pass dressing.

wheel surface. This influence is clearly seen in figure 4 for the case of the uncut diamond. A 0.001-in feed pattern produces a very different wheel surface from a 0.010-in feed even though the wear flat on the diamond is nominally 0.040 in. Now the time taken over dressing is of particular interest in industrial applications where automatic work cycles are used which involve a dressing treatment. The time required for this dressing may exceed the actual grinding time and pressures to reduce the time are great. An alternative dressing treatment may well be available, but from the results of figure 4 a faster feed of a diamond with a larger wear flat is not a straight substitute for a slow dressing feed.

Direct visual monitoring of the passage of the stylus across the grinding wheel using a stereoscopic microscope revealed that a pointed stylus rarely recorded the true extent of a peak height but rather was diverted across the side of the hill. Side tracking of this type was not seen to occur with the chisel stylus.

Now the passage of a diamond dresser across the surface of a grinding wheel would be expected to exercise a fairly rigid geometric constraint on the

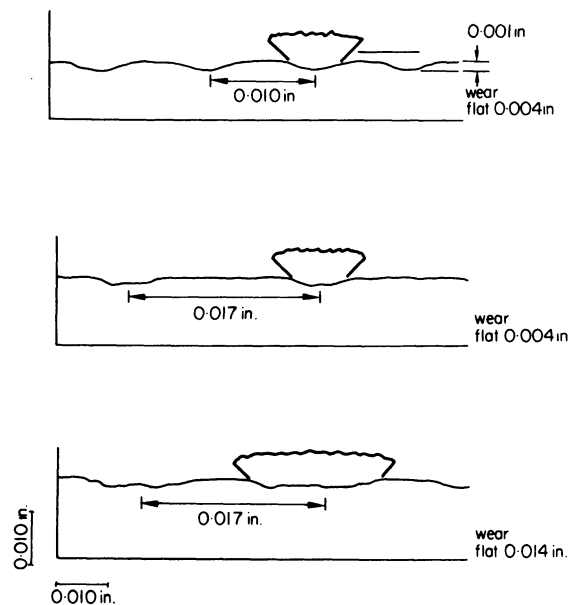


Figure 8. Profilographs of grinding-wheel surfaces with equal horizontal and vertical magnification.

height of the peaks within the profile. The profilographs of figures 4 and 6 produced from data collected by the chisel stylus show that the tops of the peaks of a reasonably sharp wheel do fall within a very small height band, on the other hand, the height levels of the peaks of figure 5 produced by the pointed stylus show far greater scatter. This scatter can be attributed to side tracking of the pointed stylus.

Side tracking of the pointed stylus was also seen to occur when passing through valleys within the wheel profile. Here of course, the tendency was to emphasize the lateral dimensions of the valley and to record its lowest depth. Efforts were made to locate the wheel in the same radial position for each profilograph at each wear stage. Figure 5 shows that this was achieved, valleys which appear in the track of the pointed stylus at one of the earlier wear stages can be seen to occur again later. Care should be taken, however, when assessing the reproducibility of a profile purely on the reoccurrence of similar valleys since side tracking of the stylus may flatter the reproducibility.

Multi-pass dressing

When dressing a grinding wheel it is common practice to pass the dressing tool over the grinding-wheel surface a number of times; the finer the dressing required, the greater the number of passes of the dresser. Since the second and subsequent passes of the dressing tool cannot be guaranteed to be in phase with previous passes, substantial variations may be expected in the topography of surfaces prepared by multi-pass dressing. Further consideration suggests, however, that variations may only occur if the interacting width of the dressing tool is substantially less than the feed of the diamond. This is rarely the case in an industrial dressing operation. Figure 7 shows chisel-stylus profilographs taken from grinding-wheel surfaces after multi-pass dressing. The wear flat on the hitherto sharp octahedral diamond increased from 0.006 in to 0.012 in during this series of tests, consequently only the (0.010 and) 0.017 in/rev feed was greater than the interacting width of the dressing tool throughout the whole test.

The profilographs of figure 7 show that a single-pass dressing produces a visually different profilograph even for the fine feed of 0.001 in/rev. Little, if any, difference can be detected, however, between four passes of the diamond and twenty passes, with the possible exception of the very large 0.17-in/rev feed. For the two smaller feeds it would seem that the character of the wheel surface is largely erased and replaced by each successive pass of the diamond. Certainly for the ratios of interacting width to dressing feed in common industrial use there would seem to be little merit in dressing the wheel for a great number of times. For the experiments considered here after about four passes of the dressing tool the only effect would seem to be additional wear of the diamond.

Chisel-stylus profilographs cannot in themselves give sufficient information concerning the nature of a grinding-wheel surface, nor for that matter can it be

assumed that a variation in a profilograph automatically represents a wheel of different wear characteristics. Accordingly the wear characteristics of the dressed grinding wheel will now be considered.

Wheel-wear characteristics

The amount of radial wheel wear represented by the curves of figure 3 is large for the type of grinding wheel used and in some cases corresponds to a radial depth far in excess of the depth of cut of the dressing tool.

Considering the three single-pass dressing treatments represented in figure 3; the rate of wheel wear which occurred in the second half of each test was surprisingly constant in view of the range of dressing feeds and geometries of the dressing diamonds used. The only exceptional curve was the low rate of wear which resulted from the smallest feed rate (0.001 in/rev) and the bluntest diamond (0.040-in wear flat). Turning now to the first half of the wear tests, for the majority of tests both the amount of radial wear and the rate of wear of the wheel showed some dependence on dressing feed, especially for the smallest dressing feed of 0.001 in/rev. It is also seen during this first half of the wear test that a number of wear curves pass through a dip or point of inflexion. This change in the rate of wheel wear, coupled with the final consistent wear rate for the wheel, can be attributed to a form of internal accommodation within the wheel profile prior to the setting up of a self-dressing mechanism which in turn will establish and maintain an inherent wear rate characteristic for the particular wheel-workpiece combination. A wear-rate characteristic which does not necessarily depend upon prior dressing treatment given to the wheel. The choice of dressing feed can hasten the establishment of this inherent wear-rate characteristic, but it is considered that it is unlikely to be responsible for its creation. Dressing treatments which may result in additional thermal damage combined with high dressing force such as one would expect with a 0.040-in wear flat and a 0.001-in feed may well come outside the scope of these remarks. In a similar manner multi-pass dressing treatments, with their extended possibility of wheel damage, could be expected to produce wear characteristics of a somewhat different shape. The curves of radial wheel wear against volume of metal removed for multi-pass dressing shown in figure 3 do exhibit a slightly different characteristic and their overall shape is similar to curves which have been obtained for crush dressing⁵.

CONCLUSIONS

Chisel-stylus profilographs have been shown to give a useful visual indication of the effectiveness of a dressing feed in the preparation of a grinding-wheel surface. When the feed of the dressing tool is greater than the wear flat on the diamond these profilographs exhibit a clear, regular feed pattern. For feeds which are less than the wear flat on the diamond variations in the profilographs are still obtained for different dressing feeds.

When wheels are dressed by badly worn dressing

tools variations in the surfaces of the wheel have been shown to exist which are a feature of the dressing feed. This would indicate that where a particular dressing feed has been found to produce an acceptable wheel surface the substitution of a faster feed rate of a diamond of larger wear flat to reduce dressing time is unlikely to achieve the desired result.

When multi-pass dressing is used for wheel preparation the character of the wheel surface is largely erased and replaced by each successive pass of the dressing tool. For dressing feeds and diamond tool geometries in common industrial use there would seem to be little merit in dressing a wheel for a great number of times.

When the wear characteristics of a grinding wheel are obtained over an extended period of wear, a wear-rate characteristic is obtained which does not necessarily depend upon prior dressing treatment. The choice of dressing tool and dressing feed can hasten the establishment of this inherent wear-rate characteristic but is considered unlikely to be responsible for its creation.

ACKNOWLEDGMENTS

The contribution of C. P. Bhateja of the Torrington Company, U.S.A., to some aspects of this work is gratefully acknowledged.

REFERENCES

1. J. Peklenik. Testing the grades of grinding wheels. *Microtecnic* (1960) XIV, no. 5, 233 and no. 6, 258.
2. M. W. Gormly. Technical aspects of vitrified grinding wheel manufacture. *Ceramic Bulletin* (1958) 37, no. 2, 77.
3. C. P. Bhateja, A. W. J. Chisholm and E. J. Pattinson. A computer aided study of the working surfaces of grinding wheels. *Proc. 12th Int. M.T.D.R. Conf.* (1971) 535.
4. J. B. P. Williamson. Microtopography of surfaces. *Proc. Inst. Mech. Engrs.* (1967) 182, 21.
5. E. J. Pattinson and A. W. J. Chisholm. Effect of dressing techniques on grinding wheel wear. *Conf. Manuf. Tech., Michigan* (1967) 601-616.

CHATTERING PHENOMENA AS THE CRITERION OF REDRESS LIFE OF GRINDING WHEEL: A STUDY ON THE ESTABLISHMENT OF OPTIMUM OPERATIONAL CONDITION IN PRECISION GRINDING OF HARDENED STEEL

by

H. MAKINO*

SUMMARY

The occurrence of chatter in the grinding process is adopted as one of the most important criteria of the wheel life. The chattering phenomena are examined experimentally by using many kinds of measurements such as grinding force, roughness of ground surface, profile of wheel face, etc., as well as mechanical vibration.

The generation of chattering is recognized as taking place in the region where the static component of normal grinding force saturates after increasing the grinding force gradually. This means that the resonant state of the vibration system of grinding is kept in that region of saturation because the coupling stiffness between wheel and workpiece is settled at a constant value.

The sub-microscopic conditions of the wheel face are observed. It is verified that there exists an optimum grinding condition with regard to the production rate. The saturation of grinding force appears lastly through the proper combination of attrition wear and fracture of cutting edges under that condition. It is confirmed that such a minute profile of the ground surface as shapes of valleys in a profilograph changes in connection with the grinding force as the grinding progresses.

INTRODUCTION

An experimental study has been carried out to obtain the optimum operational condition in cylindrical plunge grinding of hardened steel, from the viewpoint of both quality of the product and production rate. Effects of operational conditions such as dressing condition, grinding condition, rigidity of machine, types of grinding wheel, etc., upon the material stock removal between two dressings and roughness of the ground surface have been examined¹⁻⁴. From the results, an optimum operational condition for obtaining a medium degree of finish has already been proposed⁴ (Appendix I).

In the process of the study mentioned above, it is recognized that there are two ways in which the tool life of the grinding wheel terminates. One is the occurrence of chatter which is characterized by a steep increase of grinding sound and the appearance of a peak value, followed by wavy fluctuation of grinding power. Another is the appearance of burn marks on the workpiece surface accompanied by a simultaneous steep increase of grinding power. The former is termed the *chattering type* and the latter is termed the *burning type*.

The *chattering type* is the more important as the criterion of wheel life because chatter occurs inevitably in the grinding process, while *burning type* does not necessarily appear. Thus, individual properties

and mutual relations of such measured values as machine vibration, grinding force and circumferential profile of wheel have been examined, and it is clarified that the chattering occurs in the region where the static component of normal grinding force saturates, after the force is gradually increased during a grinding process.

These results suggest that there may be an optimum grinding condition (rate of stock removal) in view of the production rate, i.e. the maximum possible stock removal between two dressings. This should result in a grinding condition under which the saturation of grinding force appears last through the moderate combination of attritional wear and fracture of grains. The existence of such a grinding condition is verified by the use of a measuring apparatus that is able to observe the sub-microscopic conditions of the wheel face.

The paper deals also with the variation in minute profile of the ground surface corresponding to the variation of the grinding force in the grinding process as mentioned above.

EXPERIMENTAL PROCEDURE

Experimental condition of grinding

Types of grinding: cylindrical and plunge cut grinding—where workpiece is held between centres.

* Department of Mechanical Engineering, College of Science and Engineering, Aoyama Gakuin University, Tokyo.

Workpiece: bearing steel; hardness in Rockwell C scale, 63; diameter of ground portion, 60 mm.
Grinding wheel: white alumina grain; grain size, #60; grade of hardness, J; medium structure; vitrified bond; outside diameter, 355 mm; speed, 1740 rev/min.

Grinding fluid: water (soluble type).

Dressing condition: single point diamond dresser, apex of diamond formed conically with an apex angle of 120° ; wear of apex is strictly limited because of the important effect wear has on the grinding characteristic of the dressed wheel⁴; feed rate of dresser, 0.1 mm per wheel revolution; depth of diamond setting for finished dressing, 0.01 mm \times 5 passes after the remaining effects of the wheel face from the previous grinding experiment have been removed.

Grinding condition: grinding width, 20 mm; stock removal per workpiece, 0.4 mm, depth setting of wheel, $\Delta = 1, 2, 4, 8, 12, 16 \mu$ per workpiece revolution; speed of workpiece, $n = 54, 108, 216$ rev/min, without spark-out.

Measuring apparatus

In order to understand the various phenomena accompanied with grinding, it is necessary to adopt as many kinds of measurements as possible because of the complicated grinding process. The following measuring apparatus is used, besides the measurements of the radial wear of the wheel, grinding power and roughness of the ground surface, which have been adopted hitherto.

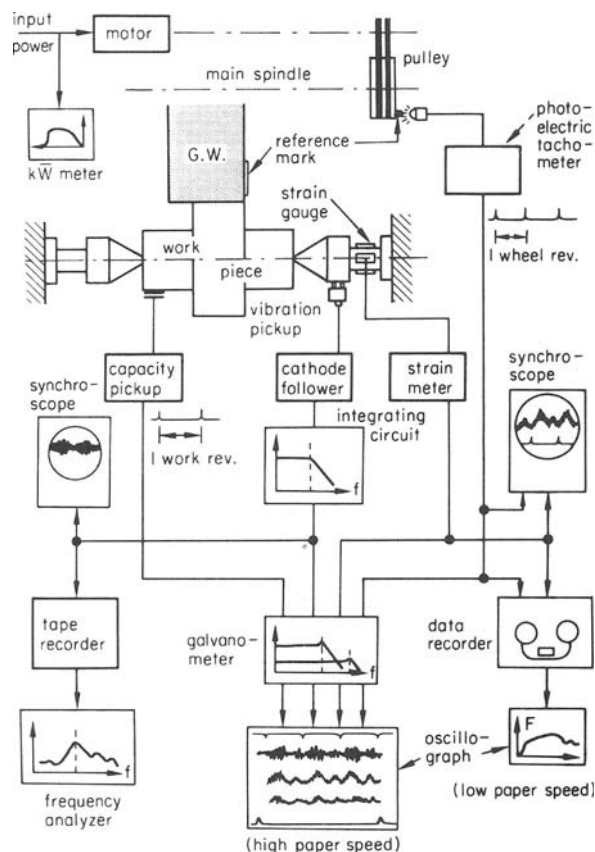


Figure 1. Measuring apparatus of grinding force and mechanical vibration.

Measurement of mechanical vibration and grinding force

The arrangement of the measuring apparatus is shown in figure 1. Mechanical vibration and grinding force (tangential force and normal force) are detected by the vibration pickup of piezoelectricity type and strain gauges respectively. Both are attached to the tailstock centre, and are recorded. The signals of each revolution of the workpiece and the grinding wheel are simultaneously recorded.

The reasons for the middle portion of both centres being reduced are:

- (i) to make it possible to approximate, closely, the grinding system of vibration to a *single degree of freedom* system by making the rigidity of the workpiece system much lower than that of the grinding wheel system (Appendix II)
- (ii) to make the various grinding phenomena, accompanied by the termination of wheel life, more obvious compared with the case of high rigidity centres⁴
- (iii) to increase the sensitivity of measurement reading of mechanical vibration and grinding force.

Measurement of grinding wheel profile

The waviness generated on the periphery of wheel by chattering vibration is detected by use of an electric transducer with its feeler running along the periphery of the grinding wheel which is driven at 3 rev/min through a reduction gear.

The material of the feeler is usually cemented carbides or hardened steel⁵ (aluminium was used in this particular experiment). Several revolutions of the wheel are allowed before a measurement is taken, and then the top of the feeler comes into reliable contact with the wheel face, i.e. the feeler measures over a broad effective area owing to its gradual wear. Then, the variation of high frequency caused by individual grains are eliminated and it is possible to record the waviness clearly. The recorded effect of feeler wear in a measurement is linear with regards to time, and does not effect the evaluation of waviness as shown lately in figure 4(c).

Measurement of plateau area of cutting edges and successive cutting point spacing

The principle of measurement is shown in figure 2. The light reflected from the plateau of worn cutting edges (figure 2(a)) is received by a photomultiplier through a slit. The output is amplified then sliced at a certain level (figure 2(b)) and is modulated into a square wave form (figure 2(c)). Summation of the individual lengths of square waves l_i is divided by the total scanning length L and gives the ratio of the plateau area of worn cutting edges to the total area of wheel face η . On the other hand, the total scanning length divided by the number of square waves n_c gives a mean value for the successive cutting point spacing a .

The measurements were carried out intermittently during a grinding process at a slow scanning speed.

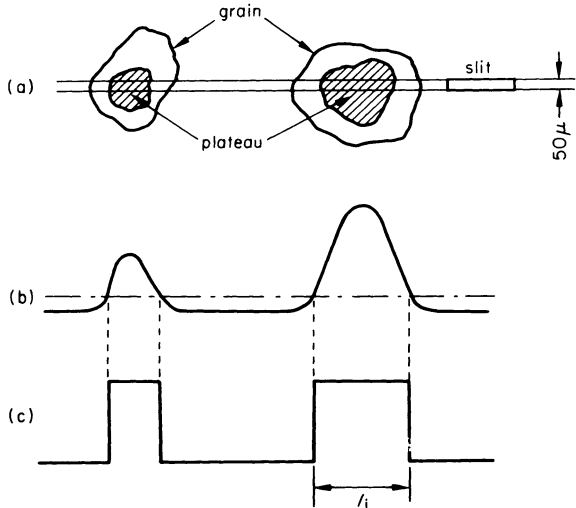


Figure 2. Measuring method of sub-microscopic conditions of wheel face.

However, new measuring apparatus has been developed, and in-process measurement apparatus is now available⁶. The practical measuring procedure will be described later.

PHENOMENA AT THE TERMINATION OF WHEEL LIFE

A typical example is shown in figure 3. The termination of a wheel life is indicated by the sudden

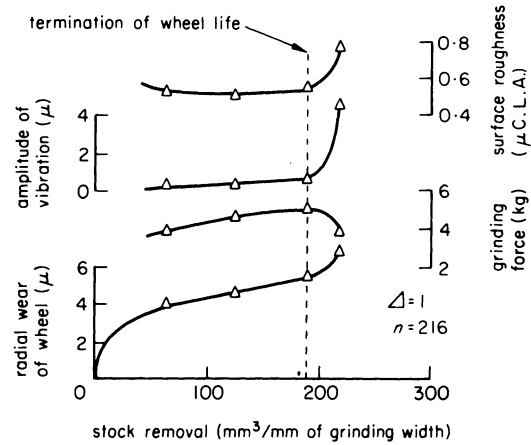


Figure 3. Phenomena at termination of wheel life (i).

increase in roughness of the ground workpiece surface, the amplitude of mechanical vibration and the radial wear rate of the wheel. The amplitude of waviness formed on the periphery of wheel also steeply increases simultaneously, though it is not plotted. At this point the normal grinding force decreases.

Normal grinding force

Figure 4(a) shows the variation of the static component of normal grinding force during a grinding process. The process appears to be composed of three regions, the *transient region*, the *steady region* and

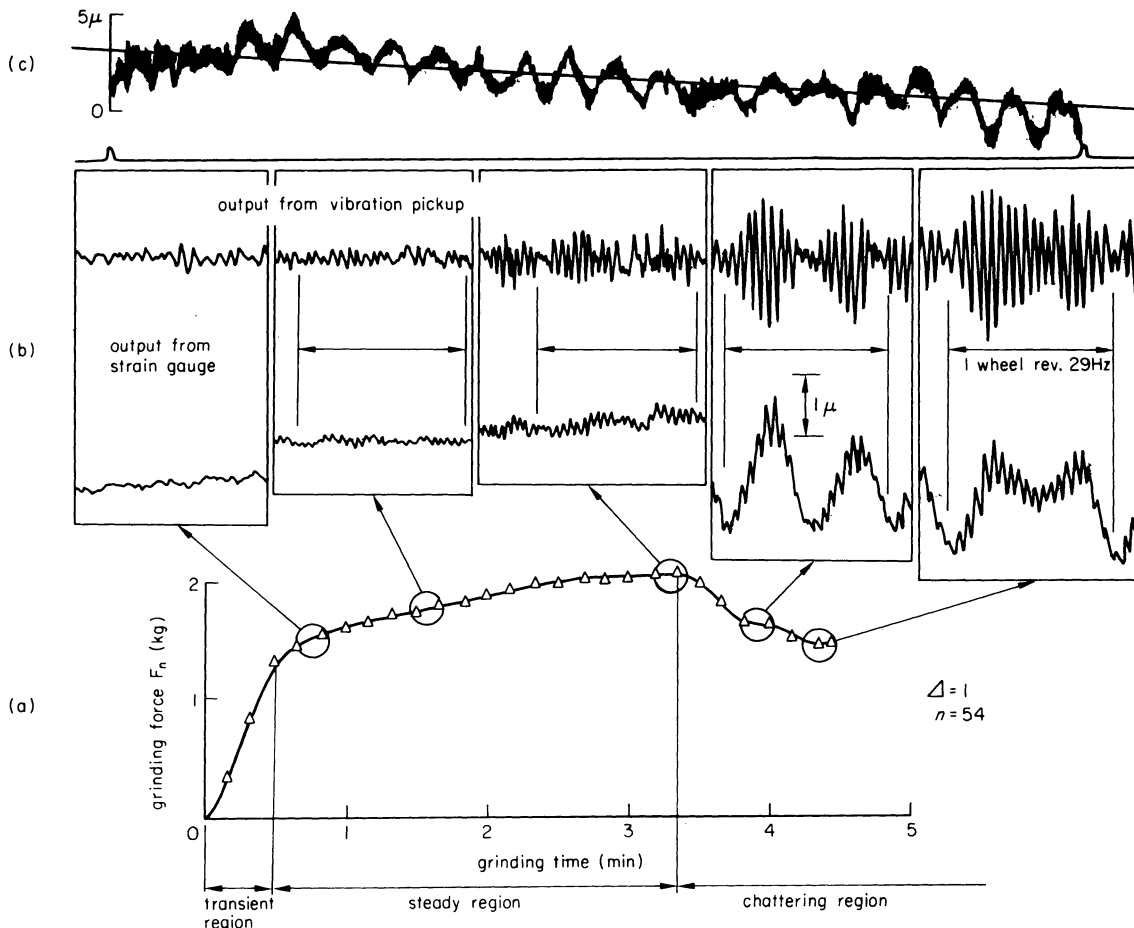


Figure 4. Variation of grinding force and mechanical vibration.

the *chattering region*. In the transient region, the actual depth of cut of wheel steeply increases but not instantly, owing to the elasticity of the workpiece system.

Though the grinding force increases gradually in the steady region, the rate of increase however, decreases and the grinding force finally saturates. In the chattering region it was found that the grinding force decreased suddenly and wavy fluctuations were indicated thereafter.

Mechanical vibration

The observation of the wave form and the frequency of mechanical vibration is carried out mainly by the output from the vibration pickup, and the observation of amplitude is mainly monitored by the strain gauges. It should be noted that the dynamic component of output from strain gauges does not necessarily indicate the correct dynamic component value of the grinding force. This is due to the close relationship between the chattering and the natural mode of the vibration system of grinding. The tailstock centre displacement is obtained from the output of the strain gauges.

Wave form and amplitude of vibration

Figure 4(b) shows the variation of mechanical vibration in connection with the normal grinding force. While random vibration appears in the earlier stages of the grinding process, the vibration becomes more periodic with the progress of grinding in the steady region, and the amplitude of vibration increases to some extent. When the grinding force begins to decrease after saturation, the amplitude increases suddenly and the wave form is modulated by hour-glass shaped waves which have long wavelengths corresponding to two cycles per wheel revolution in this experiment.

The output from the strain gauges shows the oscillogram similar to the one from the vibration pickup. There appears, however, a violent modulation in wave form by the waves with long wavelengths in the chattering region

Figure 4(c) shows an example of the profile of the periphery of the wheel in the chattering region. The profile shows a close similarity to the mechanical vibration oscillogram

- (i) the two waves of long wavelength in the wheel periphery correspond to the modulation in mechanical vibration
- (ii) the frequency and the variation in amplitude of the waviness of the wheel surface comprises a short wavelength corresponding in frequency and in form to the wave with short wavelength in mechanical vibration.

Figure 5 shows a recorded example of the output from the strain gauges in the latter stage of the steady region, followed by the chattering region at a slow oscillograph paper speed. Fluctuation in the no-load range is due to the wrong alignment of the centre holes of the workpiece. The steep increase in both amplitudes of vibration with long and short wavelengths occurs, showing the termination of wheel life during grinding.

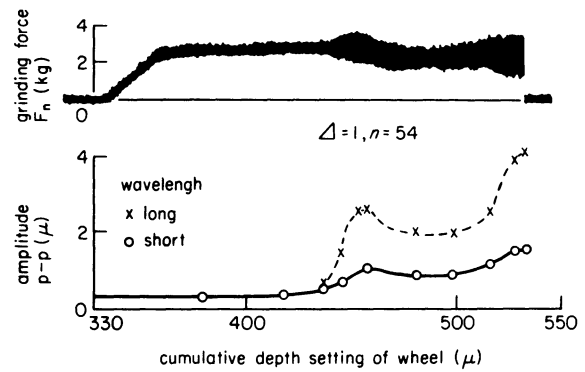


Figure 5. Phenomena at termination of wheel life (ii).

The occurrence of vibration with long wavelength seems to have certain connection with the growth of vibration with the short wavelength from the observation of a noticeable variation of wheel profile relative to time. The reason why such vibration occurs has not yet been clarified despite further sophisticated experiments. Therefore, only the vibrations with short wavelengths are dealt with hereafter.

Frequency of vibration

It is one of the peculiarities of grinding chatter that the frequency of vibration changes with time during grinding⁵. The frequency is analysed by using Type 2107 Frequency Analyser and Level Recorder Type 2305 made by Brüel Kjaer Co. An example is shown in figure 6(a). In the steady region, the frequency increases and the peak of the frequency spectrum

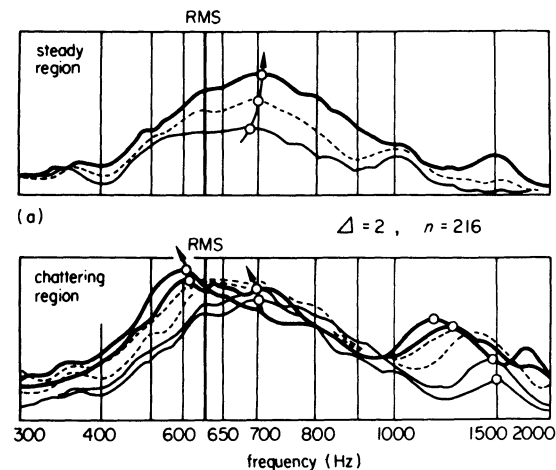
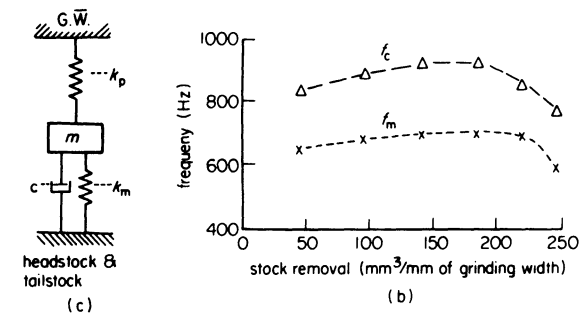


Figure 6. Frequency of mechanical vibration.

becomes more apparent. In the last stage of steady region, the peak of higher harmonics appears, showing the deformation of the wave form of vibration.

While the frequency at the peak decreases in the chattering region, the rate of decrease is not even. As the grinding progresses, the spectrum, whose frequency is lower than the peak that existed originally, grows (narrow line in figure 6(a)). The peak becomes rather flat (dotted line), then a new peak appears at a lower frequency (broad line). The frequency of these sharp-pointed peaks has the value of integer times the wheel revolution, showing that the waves formed on the periphery are also integer. As mentioned previously the normal grinding force increases in the steady region and decreases in the chattering region. Therefore, the frequency of mechanical vibration has, apparently, a close relationship with the magnitude of grinding force.

Figure 6(b) shows the calculated values f_c compared with the measured values f_m of the frequency of mechanical vibration using the formula of the natural frequency of vibration system

$$f_c = \frac{1}{2\pi} \sqrt{\frac{k_g + k_m}{m}} \quad (1)$$

This is achieved by applying an equivalent dynamic system (figure 6(c)) to the grinding system, where k_g is the static grinding stiffness, which is the static component of normal grinding force F_n divided by the depth setting of wheel per workpiece revolution Δ . k_m and m are the other constants of the system (Appendix II).

The coupling stiffness between wheel and workpiece, k_p , is thought to be a tandem in series of the grinding stiffness and contact stiffness. Since the dynamic values of these stiffnesses are difficult to obtain the static grinding stiffness k_g is used in the place of k_p as a test case. As seen in figure 6(b), both values have the same form though the f_c value is higher than the f_m value because of the over estimation of k_p due, principally, to neglecting the contact stiffness.

The contact stiffness takes a large value when the real area of contact between the wheel and workpiece is large⁷. In the grinding process, the magnitude of the grinding force is dominated by the distribution density and the dullness of the cutting edges (this is verified in the following section). Thus, it is thought that the coupling stiffness, and, accordingly, the frequency of the mechanical vibration, is closely related to the static component of normal grinding force.

From these results mentioned above, the phenomena at the termination of wheel life may be explained as follows: In a grinding process, a relative vibration between wheel and workpiece is generated at a frequency coinciding with the natural frequency of the grinding system, and waviness is formed on the wheel circumference. However, in the earlier stages, where the normal grinding force grows generally, the natural frequency of the system increases with time due to the increase in coupling stiffness between wheel and workpiece. Accordingly, the pitch of wheel

waviness tends to change, and the amplitude of waviness does not increase. The amplitude of mechanical vibration is also limited. Since the coupling stiffness shows an almost constant value in the succeeding region where the grinding force saturates. The natural frequency of the grinding system is fixed and resonant state is kept for a certain period, then the vibration becomes violent and waviness also develops accompanied by the sudden decrease in grinding force due to the fracture or breaking down of grains, and the appearance of chatter marks on the ground workpiece surface.

THE OPTIMUM GRINDING CONDITIONS

It has been already shown that there is an optimum grinding condition (regarding the rate of stock removal) under which the possible amount of stock removal is obtained between two dressings^{2,4}. The reason for this is not yet absolutely clear.

From the conclusions in the preceding section, it is presumed that in excessively light grinding, chattering will occur at an early stage of the grinding process because of the early saturation of the grinding force, through the fast increase of the attritional wear of grains. In excessively heavy grinding, occurrence of chatter will also be early in the process because the grinding force maintains a rather constant value from the beginning of the grinding process through the severe fracture of grains. Thus, there would be the moderate grinding condition under which the saturation of grinding force appears last through a suitable combination of attritional wear and fracture of grains.

In order to be sure of these presumptions the variation of η (the ratio of the plateau area of worn cutting edges to the total area of wheel face) and n_c (the number of cutting edges) are measured with the simultaneous measurement of grinding force, mechanical vibration, roughness of ground surface, etc., as described previously.

The grinding experiments are carried out under various depths of wheel setting while the workpiece revolution is kept constant. η and n_c are measured intermittently during each grinding process. Prior to each measurement, the wheel face is coated with a thin film of soot, then an extremely light grinding, without grinding fluid, is performed so as to remove the soot on the top of the cutting edges. This makes the distinct contrast between the plateau of cutting edges and the valleys.

Evaluation of η and count of n_c are measured longitudinally at the table speed of 60 mm/min (this being the maximum scanning speed) on which the measuring apparatus is set. This is because it is difficult to drive the wheel smoothly and steadily at a slow speed. Every measurement is repeated at twelve different positions on the periphery of the wheel and a mean value is plotted.

η in the grinding process

Several examples are shown in comparison with the static component of normal grinding force in figure 7. It is to be expected that the grinding force varies with η ⁸ though there appears some phase differences

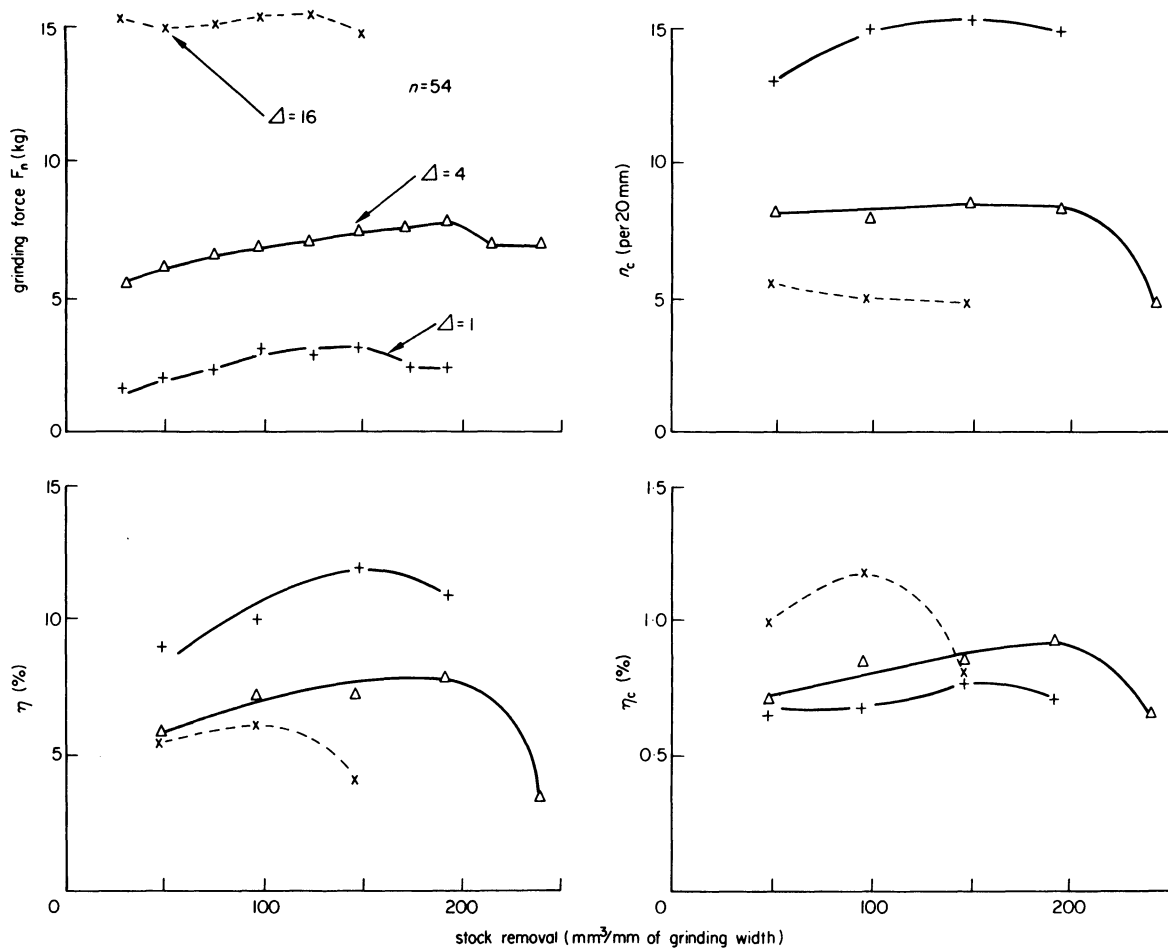


Figure 7. Variation of sub-microscopic conditions of wheel face.

between both variations, owing to the fewer points plotted of η than of the grinding force. Thus, the appearance of a peak value, followed by the sudden decrease of η , can be adopted as a criterion of a chattering type wheel life. However, in the case of the burning type, the magnitude of η is adopted as a criterion of wheel life⁸.

n_c in the grinding process

In figure 7 the values of n_c are also plotted. It is noticed that n_c shows, in light grinding, a similar variation of η meaning the predominant influence of n_c on η . The similarity between the curves becomes less in heavier grinding. η_c in figure 7 is η divided by n_c . This is the relative plateau area of the individual cutting edge. From the results obtained, the heavier the grinding condition is, the more apparent is the peak of η_c at the termination of wheel life.

From these experimental results, the existence of the optimum grinding condition is verified from the viewpoint of η which dominates the characteristics of a wheel with regard to the normal grinding force. η , however, is governed by different factors. In rather light grinding processes the n_c contributes to η , whilst η_c contributes to η in the heavier grinding processes.

Figure 8 shows the stock removal obtained until the grinding wheel life terminates under various depth settings of wheel. The conservative criterion of wheel life, described in the preceding section, is adopted. Under the moderate grinding condition ($\Delta = 4 \sim 8 \mu$ per workpiece revolution in this case), the stock removal is the largest. Of course, under this grinding condition the saturation of η appears last through the most suitable combination of attritional wear and break down of grains as shown in figure 7.

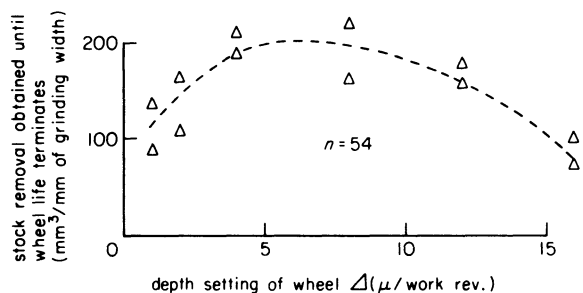


Figure 8. Optimum grinding condition.

PROFILE OF THE GROUND SURFACE

The roughness of the ground surface is closely related to the successive cutting point spacing⁹. This can be recognized also in figure 9 which is obtained from the experimental results in the preceding section, where the successive cutting point spacing a is L (the total scanning length) divided by n_c (the number of cutting edges), as previously described.

There is an operational condition under which the roughness of the ground surface keeps an almost

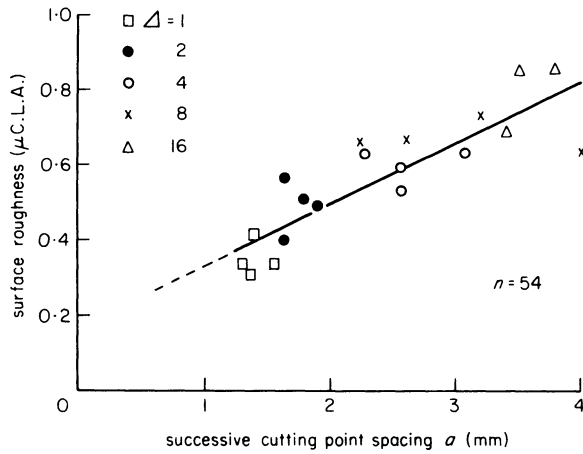


Figure 9. Relation between successive cutting point spacing and surface roughness.

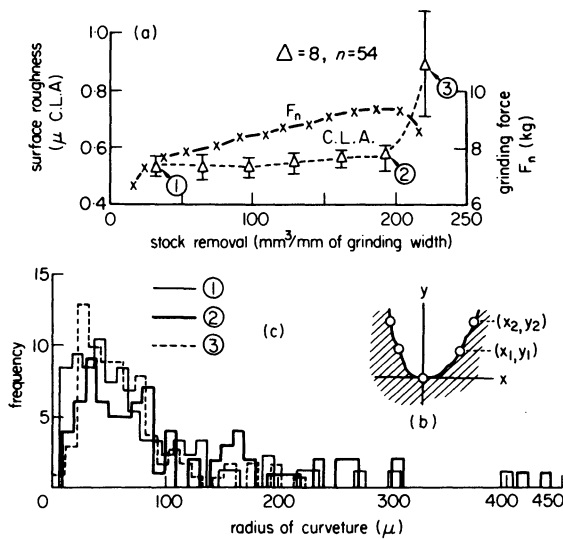


Figure 10. Variation of minute profile of ground surface.

constant value until the wheel life terminates by a suitably combined condition of dressing and grinding⁹. An example shown in figure 10(a) means n_c does not vary. However, the grinding force shows a certain increase in the same wheel life range. Therefore, there should be an increase of the plateau area of the individual cutting edge, according to the conclusion of the preceding section. Contrary to this, there would be the decrease of the plateau area in the chattering region through the fracture or the breaking down of grains.

It is supposed that the minute profile of the surface ground by the wheel having sharp or dull cutting edges would differ to some extent, but such a value of surface roughness, indicated by the centre line average, would show identical results. The shapes of valleys in a profilograph, which were formed by the individual cutting edges, were examined. Three workpieces were picked out; one in the early stage of grinding (figure 10(a) ①), one in the latter stage in the steady region (figure 10(a) ②) and one in the chattering region (figure 10(a) ③).

The profilograph of each workpiece is obtained by using Talysurf 4 at four positions of the workpiece surface, tracing in the direction parallel to the axis of the workpiece. From each profilograph twenty

valleys were picked out at random and the deepest point of each valley is set on the original point of an orthogonal coordinate as shown in figure 10(b).

The coordinates of two points x_1 at y_1 and x_2 at y_2 on the right hand side of a profile curve were measured by using a universal tool microscope (where y_1 and y_2 are $0.5 \mu\text{m}$ and $1 \mu\text{m}$ on the real scale, respectively). Then, x_1 and x_2 are converted to the values in the real scale, considering the magnification of the profilograph, and are substituted into a formula of a curve of the second order

$$x^2 + Ay^2 - By = 0 \quad (2)$$

with y_1 and y_2 in the real scale.

Measurement of the points on the profile curve and substitution into the formula are carried out also on the left hand side of the ordinate, and the mean value of both $B/2$'s is obtained as a mean radius of the curvature of the deepest point of each valley. Incidentally, if $A > 1$ or $0 < A < 1$, $A = 1$, $A = 0$ and $A < 0$ the curve is an ellipse, a circle, a parabola and a hyperbola respectively. Histograms of radii are shown in figure 10(c).

From the results, it is apparent that the valleys in the profile of ground surface are rather sharp in the early stage of the grinding. The shapes of valleys are, however, rounded with the progress of grinding and resharpened during chattering. Though the increase of the surface roughness during chattering is not appreciable in cases where machine centres of high rigidity are used⁴, the condition of resharpening of workpiece valleys during chattering can be recognized in other experiments.

CONCLUSIONS

In order to obtain the optimum grinding condition, the process of cylindrical plunge grinding of hardened steel is examined experimentally with the following results.

- (1) Chattering (the occurrence of which is one of the most important criteria of the termination of wheel life), generates and grows in the region where the static component of the normal grinding force saturates, after being gradually increased, because the resonant state of the vibration system of grinding has been kept there.
- (2) The ratio of the plateau area of the worn cutting edges to total cutting area of the wheel has a close relationship with the static component of the normal grinding force. There is also an optimum grinding condition from the viewpoint of the production rate, under which the saturation of grinding force appears last due to the ideal combination of the attrition wear and the fracture of cutting edges.
- (3) The minute profile of the ground surface changes with the variation of grinding force; though such a value of surface roughness, using the centre line average, does not show any appreciable changes.

ACKNOWLEDGMENT

The author wishes to express his sincere thanks to Professor Dr Y. Suezawa who gave him the opportunity to present this paper at the 15th International MTDR Conference. He also wishes to express his deepest appreciation to Mr Y. Suzuki for his assistance in the initial preparation of this study, especially in the measurements of vibration, and to Mr T. Suzuki for his assistance in the final stages, also to several undergraduates for their assistance in carrying out the intricate time consuming experiments.

REFERENCES

1. H. Makino. The life of the grinding wheel. *Bulletin of Japan Society of Grinding Engineers* (1961). 1, 41.
2. H. Makino, H. Noguchi, S. Suzuki, T. Mizushima, T. Suto, K. Kamakura and H. Inoue. The optimum condition of the metal grinding. *Journal of Mechanical Laboratory of Japan* (1962). 8, 1.
3. H. Makino. Roughness of finished surface in grinding operation of hardened steel. *Bulletin of the Japan Society of Precision Engineering* (1966). 1, 281.
4. H. Makino. The tool life of grinding wheel in precision grinding of hardened steel. *Bulletin of the Japan Society of Precision Engineering* (1966). 2, 47.
5. G. Sweeney. Grinding instability. *Proceedings of the 6th International M.T.D.R. Conference* (1965). 552.
6. T. Suto, T. Waida and T. Sata. In-process measurement of wheel surface in grinding operations. *Proceedings of the 10th International M.T.D.R. Conference* (1969). 171.
7. P. D. Singhal and H. Kaliszer. The effect of workpiece dimensions and wheel parameters on the waviness during grinding. *Proceedings of 6th International M.T.D.R. Conference* (1965). 629.
8. H. Yoshikawa. Criterion of grinding wheel tool life. *Bulletin of the Japan Society of Precision Engineering* (1963). 1, 29.
9. H. Makino, T. Suto, and E. Fukushima. An experimental investigation of the grinding process. *Journal of Mechanical Laboratory of Japan*. (1966). 12, 17.

APPENDIX I

In the case where vitrified grinding wheels of medium structure with conventional peripheral speed and fluid of water (soluble type) are used, the following operational conditions are recommended for cylindrical and plunge grinding of hardened steels.

Grinding wheel: single crystal alumina grain; grain size, varying from #80 to #100; grade of hardness I and J.

Dressing condition: feed rate of single point diamond dresser, 0.1 mm per wheel revolution; depth setting, 0.01 mm.

Grinding condition: rate of stock removal, 1 mm³/mm of grinding width/sec (for example, depth setting of wheel 3 μm per workpiece revolution and peripheral speed of workpiece 20 m/min); spark-out, 5 revolutions of workpiece.

With the combination of these conditions, maximum stock removal between two dressings is attained, keeping a surface roughness within 0.5 μm of centre line average value.

APPENDIX II

Stiffness of grinding wheel system is $k_w = 10 \text{ kg}/\mu\text{m}$.

Stiffness of workpiece system is $k_m = 0.63 \text{ kg}/\mu\text{m}$.

Equivalent mass of workpiece system is $m = 0.15 \text{ kg/s}^2/\text{m}$. This is estimated as the total mass of the workpiece itself, plus the mass of the centre's cones and the adjacent material between their two reduced portions.

Damping ratio of workpiece system is $\zeta = 0.021$ calculated from the record of damped oscillation which is obtained by an impulse response to the workpiece system.

The natural frequency of the workpiece system f is 327 Hz which is calculated by substituting the constants, mentioned above, into the formula of natural frequency of a vibration system, under single degree of freedom. The value shows an almost identical result with the measured value of 340~350 Hz obtained from the impulse response.

FINISHING AND RUNNING-IN OF PLAIN BEARINGS, WITH AN ECONOMIC APPRAISAL

by

G. TRMAL,* H. KALISZER* and G. W. ROWE*

1 INTRODUCTION

The main objective of this paper is a cost analysis for surface preparation of journals and bearings and an evaluation of the benefits of different surface roughness by considering their effect upon bearing life. From these it is possible to optimize the machining conditions for an expected lifetime of bearings. It is believed that the results should be of interest to users of various machines equipped with bearings as well as to those concerned with surface finishing processes.

The major emphasis was placed on investigating the behaviour of grease lubricated bearings during running-in, although some pilot tests which were conducted on a pin-and-disc Timken wear-testing machine have indicated that the results are similar for oil-lubricated bearings in the regime of boundary lubrication.

The process of running-in is essential to practically all moving parts, yet it is only in recent years that measuring equipment and technique has reached a degree of sophistication sufficient to allow detailed study of the changes occurring during running-in. Qualitatively it is easy to achieve steady smoothing and improvement in conformity of sliding surfaces by a variety of combined mechanical and chemical processes until some acceptable final condition is achieved. Accurate specification of this condition, however, has always required very expensive rig tests or even more expensive field trials. Basically it is recognized that a pair of properly run-in surfaces depend upon a balance, which may be very delicate indeed, between a topography which is smooth enough to avoid abrasion or ploughing and to allow uninterrupted elastohydrodynamic and boundary lubrications, and one which does not present sufficiently large smooth areas to encourage metal transfer. The latter phenomenon, in its various guises of local seizure, scuffing or pickup, can provide serious limitations in machine bearings, gears, slide-ways, rams, seals, cams, guides, piston rings and cylinders, and in all metal-forming processes.

2 EXPERIMENTAL PROCEDURE

It is important to recognize that the surface topography of both the journal and the bearing can influence the level of wear. To examine the various combinations of rough and smooth surfaces of bearings and journals under realistic conditions, a special wear-testing machine¹ has been used (figure 1). It consists of a shaft (1) which can be statically

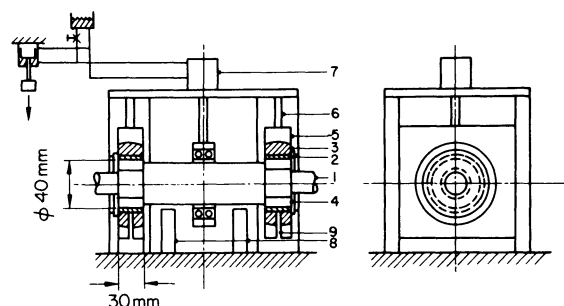


Figure 1. Wear-testing machine.

loaded from a hydraulic cylinder (7). The shaft runs in two identical bearings (5). To preserve the shaft the actual journals are expendable steel cylinders (4) firmly keyed on to the shaft. The rotational speed was set to 1200 rev/min and closely controlled at that level. The grease lubricant was fed at controlled pressure to each bearing. Special measures were taken to eliminate the end play in the shaft and to minimize bending and tilt of the journal relative to the bearing.

Each phosphor-bronze shell (2) is mounted in a special self-aligning spherical casting (3) from which it is never removed. For measurement purposes the whole casting is slid gently off the shaft and accurately located on a Talysurf platform for axial profilometry.

The profile of the bearing surface was measured with a stylus instrument over the whole length of the bearing surface along the line of maximum load, before assembly and at specified intervals of time

* Department of Mechanical Engineering, University of Birmingham

during each test. These measurements were made with high precision and exact 3 mm ball relocation². All data were digitized and stored on line for further computer analysis according to various topographical parameters⁴.

The absolute wear was measured by recording the full length of the bearing including two reference steps 0.03 mm deep, one on each end of the bearing as shown in figure 2. A Talysurf side-acting gauge

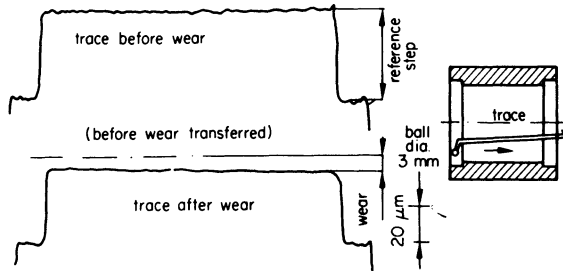


Figure 2. Measurement of absolute wear.

with 3 mm ball instead of stylus was used for measurement³.

A manufacturing technique was developed for producing bearings with a required dimensional tolerance and specified surface topography.

Complementary tests were made on a pin-and-disc Timken wear-testing machine kindly lent by the National Engineering Laboratory. As can be seen from figure 3 a steel ring (50 mm diameter) representing the journal runs against a phosphor-bronze

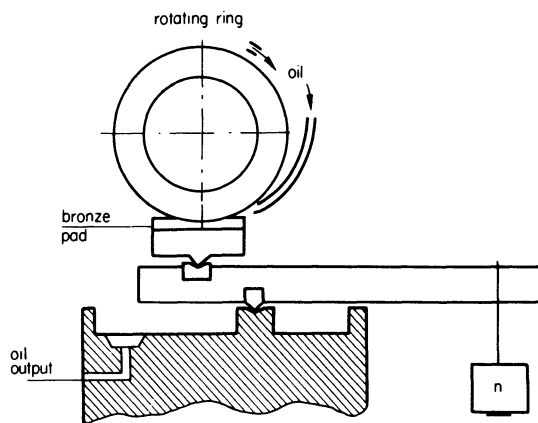


Figure 3. Schematic arrangement of the Timken wear-testing machine.

pad forming the bearing surface. A load is applied between them by a system of levers. A commercial general-purpose oil was used for lubrication. The selected running conditions were comparable with those used on the wear-testing machine shown in figure 1.

3 ANALYSIS OF SURFACE TOPOGRAPHY

All the experimental results show that the first stage of a bearing surface change is a smoothing of the machined surface (figure 4). The original bearing surface disappears and is replaced by a smoother

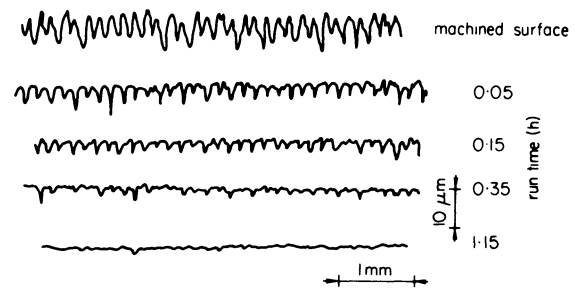


Figure 4. Stages of wear of bearing surface profile.

surface, generated by the rotating shaft. The influence of initial surface topography of the bearing does not appear to be very significant after the running-in period.

To investigate the influence of the journal roughness on the bearing surface a special test has been set up on the Timken machine. For the same fine roughness of the pad, four rings of different roughness were used for tests, each of which lasted 4 h. These rings were machined as follows:

rough ground:	$Ra = 1.7 \mu\text{m}$
medium ground:	$Ra = 0.75 \mu\text{m}$
fine ground:	$Ra = 0.16 \mu\text{m}$
polished:	$Ra = 0.05 \mu\text{m}$

The width of wear scar on the phosphor-bronze pad representing the bearing was measured and the total volume loss calculated. As can be seen from figure 5 (width of scar) and figure 6 (total volume loss), there is a pronounced influence of surface roughness of the hard ring, representing the journal, upon the wear of the bronze bearing material, despite the fact that the decrease of specific pressure with the size of wear scar has a reducing effect upon the wear.

The total volume loss due to abrasion for the medium ground ring ($Ra = 0.75 \mu\text{m}$) was about 100 times higher than for the polished ring ($Ra = 0.05 \mu\text{m}$) despite the fact that the load in the case of polished ring was four times higher.

The difference in specific pressure is even greater due to the difference in the size of the loaded area (wear scar). The tested rings with corresponding worn

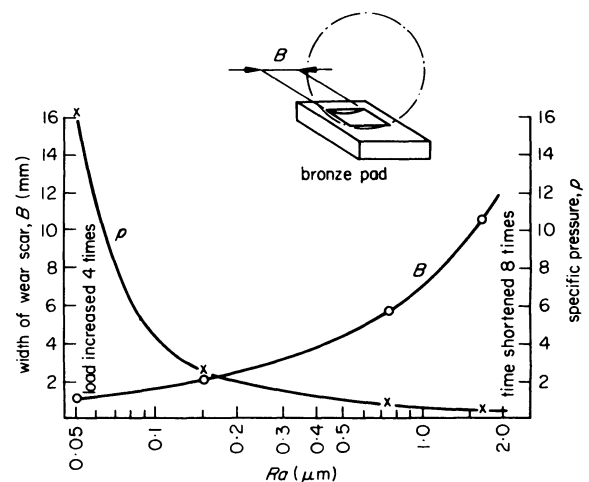


Figure 5. Wear as a function of surface roughness Ra of the ring (Timken machine).

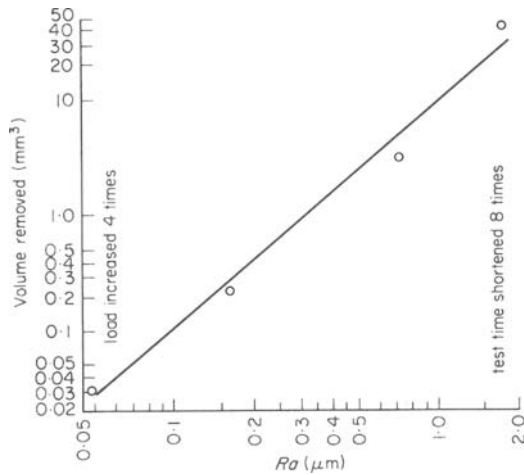


Figure 6. Volume loss as a function of surface roughness of the ring (Timken machine).

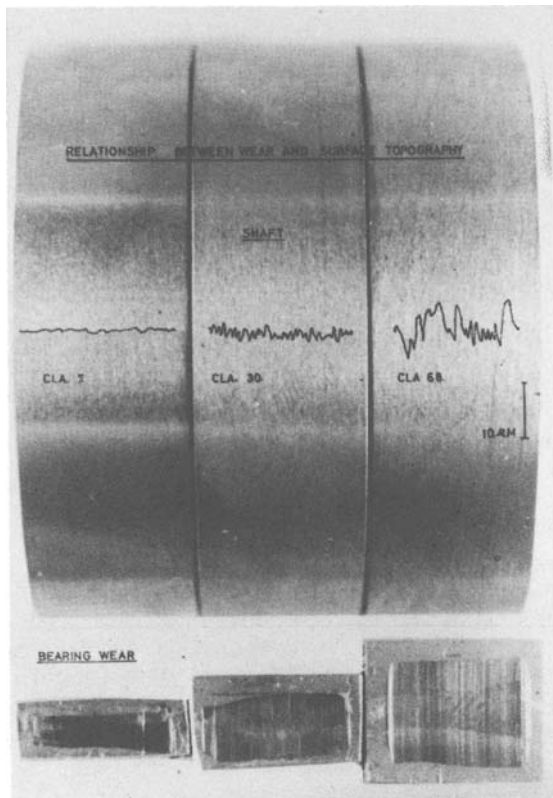


Figure 7. Tested rings with corresponding worn pads.

pads can be seen in figure 7. The scanning electron microscope photograph of ground ($Ra = 0.8 \mu\text{m}$) and polished ($Ra = 0.05 \mu\text{m}$) surfaces shown on figure 8 clearly illustrates the difference between these two surfaces.

To confirm these results, further tests were carried out on the special wear-testing machine shown in figure 1. Replaceable journals of three different surfaces, namely medium ground ($Ra = 0.75 \mu\text{m}$), fine ground ($Ra = 0.15 \mu\text{m}$) and polished ($Ra = 0.05 \mu\text{m}$), as well as two types of bearing surfaces, fine bored ($Ra = 0.8 \mu\text{m}$) and rough bored ($Ra = 6 \mu\text{m}$), were prepared and their combinations were tested.

The test run for wear measurement was selected to

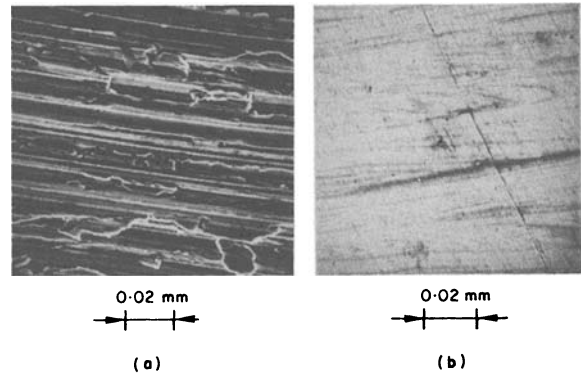


Figure 8. Scanning electron microscope photographs of the ring. (a) Ground, $Ra = 0.8 \mu\text{m}$; (b) polished, $Ra = 0.05 \mu\text{m}$.

last 16 h. The bearing wear was determined as described earlier using reference steps and evaluating the volume loss. Two testing procedures were used: the A-procedure when the 16 h of running time consisted of eight short periods starting with 5 min and doubling the time. After each period the bearing was removed from the machine and measured; and the B-procedure when the machine was run for two 8-h periods.

The results of these tests can be seen in figure 9,

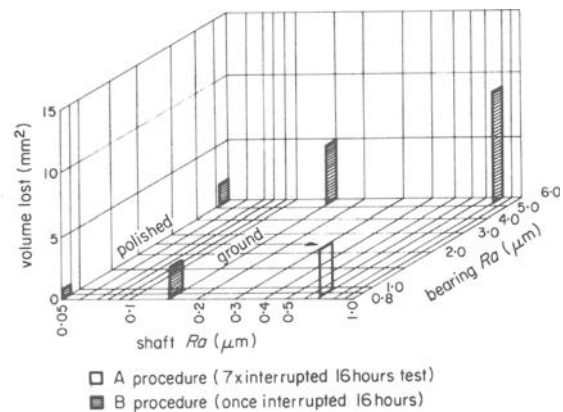


Figure 9. Bearing volume loss as a function of surface roughness of the bearing and the shaft.

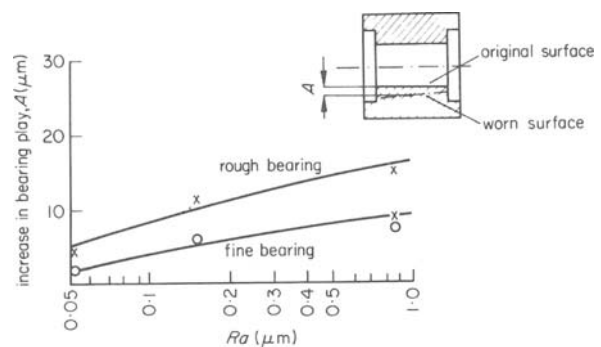


Figure 10.

where volume loss is plotted as a function of the bearing and shaft topography, and in figure 10 which shows the increase in bearing clearance as a function of bearing and shaft topography. The results indicate again a very pronounced effect of the surface roughness of the shaft upon the abrasive wear of the bearing.

Workpiece: bearing steel; hardness in Rockwell C scale, 63; diameter of ground portion, 60 mm.
Grinding wheel: white alumina grain; grain size, #60; grade of hardness, J; medium structure; vitrified bond; outside diameter, 355 mm; speed, 1740 rev/min.

Grinding fluid: water (soluble type).

Dressing condition: single point diamond dresser, apex of diamond formed conically with an apex angle of 120° ; wear of apex is strictly limited because of the important effect wear has on the grinding characteristic of the dressed wheel⁴; feed rate of dresser, 0.1 mm per wheel revolution; depth of diamond setting for finished dressing, 0.01 mm \times 5 passes after the remaining effects of the wheel face from the previous grinding experiment have been removed.

Grinding condition: grinding width, 20 mm; stock removal per workpiece, 0.4 mm, depth setting of wheel, $\Delta = 1, 2, 4, 8, 12, 16 \mu$ per workpiece revolution; speed of workpiece, $n = 54, 108, 216$ rev/min, without spark-out.

Measuring apparatus

In order to understand the various phenomena accompanied with grinding, it is necessary to adopt as many kinds of measurements as possible because of the complicated grinding process. The following measuring apparatus is used, besides the measurements of the radial wear of the wheel, grinding power and roughness of the ground surface, which have been adopted hitherto.

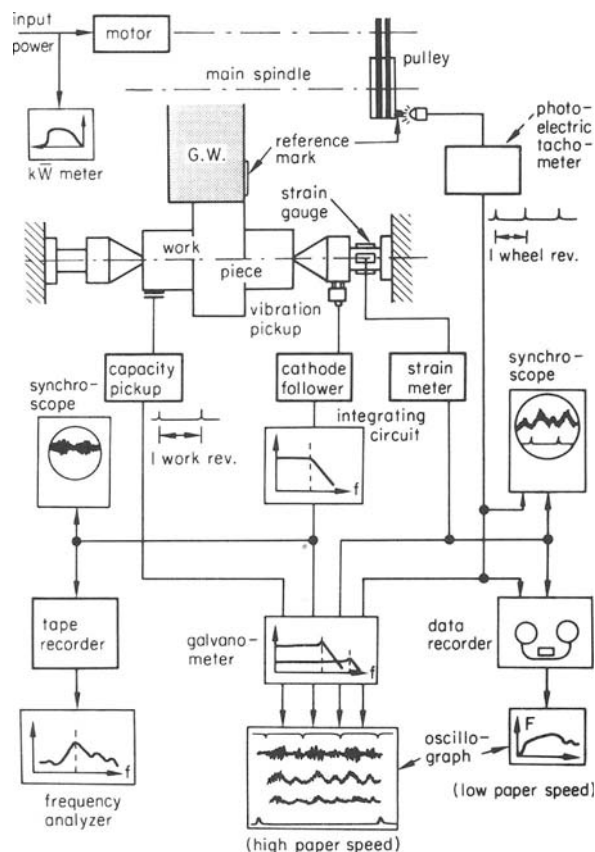


Figure 1. Measuring apparatus of grinding force and mechanical vibration.

Measurement of mechanical vibration and grinding force

The arrangement of the measuring apparatus is shown in figure 1. Mechanical vibration and grinding force (tangential force and normal force) are detected by the vibration pickup of piezoelectricity type and strain gauges respectively. Both are attached to the tailstock centre, and are recorded. The signals of each revolution of the workpiece and the grinding wheel are simultaneously recorded.

The reasons for the middle portion of both centres being reduced are:

- (i) to make it possible to approximate, closely, the grinding system of vibration to a *single degree of freedom* system by making the rigidity of the workpiece system much lower than that of the grinding wheel system (Appendix II)
- (ii) to make the various grinding phenomena, accompanied by the termination of wheel life, more obvious compared with the case of high rigidity centres⁴
- (iii) to increase the sensitivity of measurement reading of mechanical vibration and grinding force.

Measurement of grinding wheel profile

The waviness generated on the periphery of wheel by chattering vibration is detected by use of an electric transducer with its feeler running along the periphery of the grinding wheel which is driven at 3 rev/min through a reduction gear.

The material of the feeler is usually cemented carbides or hardened steel⁵ (aluminium was used in this particular experiment). Several revolutions of the wheel are allowed before a measurement is taken, and then the top of the feeler comes into reliable contact with the wheel face, i.e. the feeler measures over a broad effective area owing to its gradual wear. Then, the variation of high frequency caused by individual grains are eliminated and it is possible to record the waviness clearly. The recorded effect of feeler wear in a measurement is linear with regards to time, and does not effect the evaluation of waviness as shown lately in figure 4(c).

Measurement of plateau area of cutting edges and successive cutting point spacing

The principle of measurement is shown in figure 2. The light reflected from the plateau of worn cutting edges (figure 2(a)) is received by a photomultiplier through a slit. The output is amplified then sliced at a certain level (figure 2(b)) and is modulated into a square wave form (figure 2(c)). Summation of the individual lengths of square waves l_i is divided by the total scanning length L and gives the ratio of the plateau area of worn cutting edges to the total area of wheel face η . On the other hand, the total scanning length divided by the number of square waves n_c gives a mean value for the successive cutting point spacing a .

The measurements were carried out intermittently during a grinding process at a slow scanning speed.

produces much more bearing wear than a smoother one. The bearing tends to reach an equilibrium, since a smoother bearing surface has a tendency to roughen while rougher ones are abraded down.

(2) The economically justifiable journal roughness depends upon the life of a bearing assembly in relation to the life of the whole machine.

ACKNOWLEDGMENTS

The authors wish to thank Professor S. A. Tobias for providing facilities for the work and the Science Research Council for financial support. They also wish to thank Mr A. Cotter whose experimental work was largely used in this paper.

REFERENCES

1. H. Kaliszer, G. W. Rowe and G. Trmal. Surface Integrity after Severe Wear or Grinding. ASME, Int. Conf. on Surface Technology, Pittsburgh (USA), 1973.
2. Williamson, J. B. B. and Hunt, R. T. *J. Sci. Inst. (J. of Phys. E)* Sev. 11r 1, 1968.
3. Grieve, D. J., Ph.D. Thesis, Dept. of Mech. Eng. University of Birmingham, 1970.
4. H. Kaliszer and G. W. Rowe. Digital computation for surface topographical measurements applied in tribology. Proc. of Conf. on the Use of Digital Computers in Measurements. Inst. Elec. E.

IN-PROCESS INDICATION OF SURFACE ROUGHNESS USING A FIBRE-OPTICS TRANSDUCER

by

D. SPURGEON* and R. A. C. SLATER†

SUMMARY

For further progress to be made in the field of adaptive control of the finish turning process, considerable advances must be achieved in the acquisition of data associated with the process. The problems associated with the design and development of suitable transducers must therefore be overcome. The possibility of in-process indication of surface roughness during the finish turning process, by means of a fibre-optics transducer, is examined, and the operational principle of the transducer is briefly described.

It is shown that, when the transducer probe traces the same path as the cutting tool, the reflectivity of light from a newly turned surface varies inversely as the roughness of that surface. Polaroid film records, which illustrate the nature of the transducer output for a single revolution of the workpiece, are presented. However, it is suggested that the output, for a number of revolutions, can be analysed using an integrated circuit to give an average output which can be correlated with the conventional CLA value. The use of a fibre-optics transducer is also discussed in relation to an adaptive control loop, to continuously monitor the cutting conditions with special reference to surface roughness, and thus, indirectly, to tool wear.

It is concluded that such an optical transducer can give an indication of surface roughness when $1 \mu\text{m} < \text{CLA} < 3 \mu\text{m}$ during the finish turning process, and information can also be obtained for use in a control algorithm to adaptively control the process. Some suggestions are given for future investigation.

INTRODUCTION

There is a growing need to develop and use our manufacturing resources more efficiently and more economically. Merchant¹ has referred to this in terms of a new *optimization technology* which includes studies of computer-aided design, part-family design, numerical control and adaptive control. Of these, adaptive control becomes increasingly important as the efficiency of the machine tool increases, reaching a point when it becomes uneconomic to operate under anything other than optimum conditions. Since, by definition, adaptive control is a self-optimizing type control, considerable interest has been aroused in this recent and exciting area.

In general terms an adaptive system may be considered as one in which the performance of the process is continuously monitored and compared with a given index of performance. The process parameters are then adjusted in order to obtain an optimum set of values. For example, these values may be maximum rate of metal removal, minimum cost or maximum rate of return.

A major problem, presenting itself in work of this kind, is the design and manufacture of suitable sensors. These must be accurate, often delicate, and yet able to withstand the pollution and hazards associated with the manufacturing environment. Parameters that have been considered include cutting force, spindle torque², workpiece dimensions³, tool tip temperatures⁴ and surface finish. In some cases, it is not possible to obtain a direct reading, and, therefore, it is necessary to obtain an indication of the desired parameter by measurement of other parameters, and then determine the correlation which exists. In-process measurement of surface roughness is known to be a problem area. Given a particular part to be produced, within a defined tolerance band of surface roughness, it is required to manufacture parts at the most economical rate, whilst at the same time keeping to within the specification. This becomes particularly important when turning workpieces of large dimensions or when machining under numerical control. In the latter case, it may not always be convenient to change the tool in the middle of a tape-run.

* Department of Industrial Engineering, Hatfield Polytechnic

† Department of Mechanical Engineering, The City University

If in-process measurement of surface roughness is going to give a clear, meaningful indication of the current performance of the process, it is imperative that a continuous record of the surface roughness value be obtained. This implies assessing the condition of the workpiece surface just behind the cutting edge of the tool. To achieve this, using any form of stylus instrument, would probably result in the destruction of the stylus, due to the high contact pressure involved. Thus, a non-contact method is required, based upon either pneumatic, electrical-electronic or reflective techniques.

The main intentions of the present paper may now be summarized as:

- (a) to briefly describe the operational principle of the transducer employed in the present investigations;
- (b) to show that a transducer, based upon light reflection, may be used to obtain an indication of the quality of a newly turned surface;
- (c) to present the results, obtained when using the transducer, to give an in-process indication of surface roughness during finish turning;
- (d) to suggest the use of the transducer in relation to adaptive control of the finish turning process.

THE OPTICAL TRANSDUCER

The optical transducer used in the work was a 'Fotonic Sensor', Type KD 45 having a 1 mm diameter probe manufactured in the USA. It is of a similar type to that used by Jona^{5,6} who apparently first considered the adaptive control of the turning process from a geometrical point of view.

The sensor consists of a flexible sensing probe, detachable cartridge containing the photocell and light source, and the electronics package. The flexible probe consists of a number of fibres which have a core of high quality flint glass surrounded by a glass cladding to prevent 'cross talk' between the fibres. Light is transmitted through half of these fibres, reflected from the observed specimen, captured by the remaining fibres, and then translated into an output signal.

Since the amount of light emerging from the end of a fibre bundle is less than that applied at the other end, both the size and length of the bundle are important design considerations. Equally important is the distribution of the fibres. The most sensitive response is obtained by surrounding each receiving fibre with four transmitting fibres. In this way more and more receiving fibres are affected by light emanating from a single transmitting fibre as the distance between the optic probe and object increases. However, this technique results in costly manufacture and therefore it is customary to distribute the fibres in a random manner.

The light source consists of a pre-focussed, incandescent filament, 2.5 volt bulb which usually operates at between 1.0 V to 1.7 V to extend its life. The photocell receiver is a photo-diode which generates a current output dependent upon the intensity of the light received by the cell. Both the

light source and photocell are mounted in the cartridge.

The 'Fotonic Sensor' is marketed as an optical proximity sensor and can be used for a number of purposes of which surface roughness comparison is only one possibility. The likely areas of application may be appreciated from the response characteristic curve for the sensor presented in figure 1. By referring to this figure, it may be clearly seen that the curve has three distinct regions, 0 to 1, 1 to 2 and 2 to 3. In the regions, 0 to 1 and 2 to 3 the output of the sensor

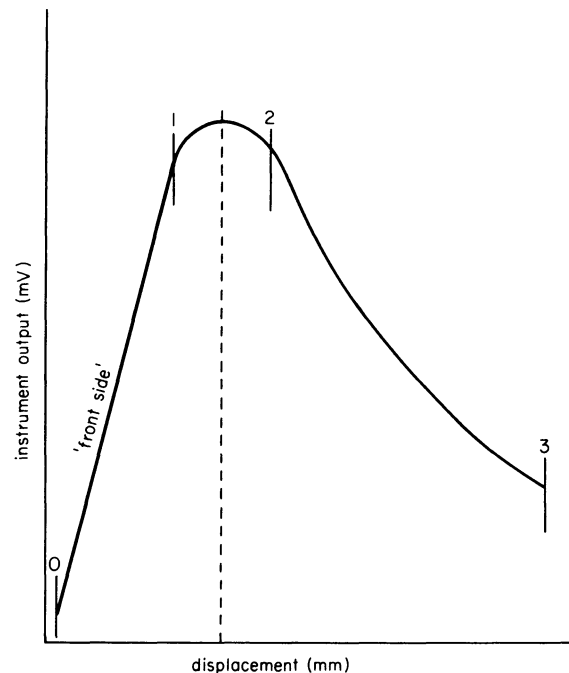


Figure 1. Response characteristic curve for the 'Fotonic Sensor'.

is dependent upon its position from the object. Of these regions the 'front side' is the normal working region for the measurement of displacement, since sensitive linear output is obtained over this region. The third portion of the curve (1 to 2) is the region adjacent to the peak which occurs at a particular value of displacement when all the receiving fibres are illuminated by reflected light. At this point the sensor is relatively insensitive to displacement but still very sensitive to reflectivity. For the instrument employed in this work the value of the critical gap between the optic probe and the workpiece was 0.2 mm.

DESCRIPTION OF EQUIPMENT AND MATERIAL USED

The work was carried out on a Lodge & Shipley 'Powerturn' lathe equipped with an infinitely variable speed motor giving spindle speeds ranging from 13 rev/min to 2000 rev/min. Available feed rates ranged from 0.019 mm/rev to 1.17 mm/rev.

Accurate cutting speeds were obtained using a tachometer which permitted a direct reading to 10 m/min, and estimation to within 5 m/min. Surface roughness was measured using a portable surface measuring indicator. A 'Talysurf' surface roughness measuring instrument was also used to check the

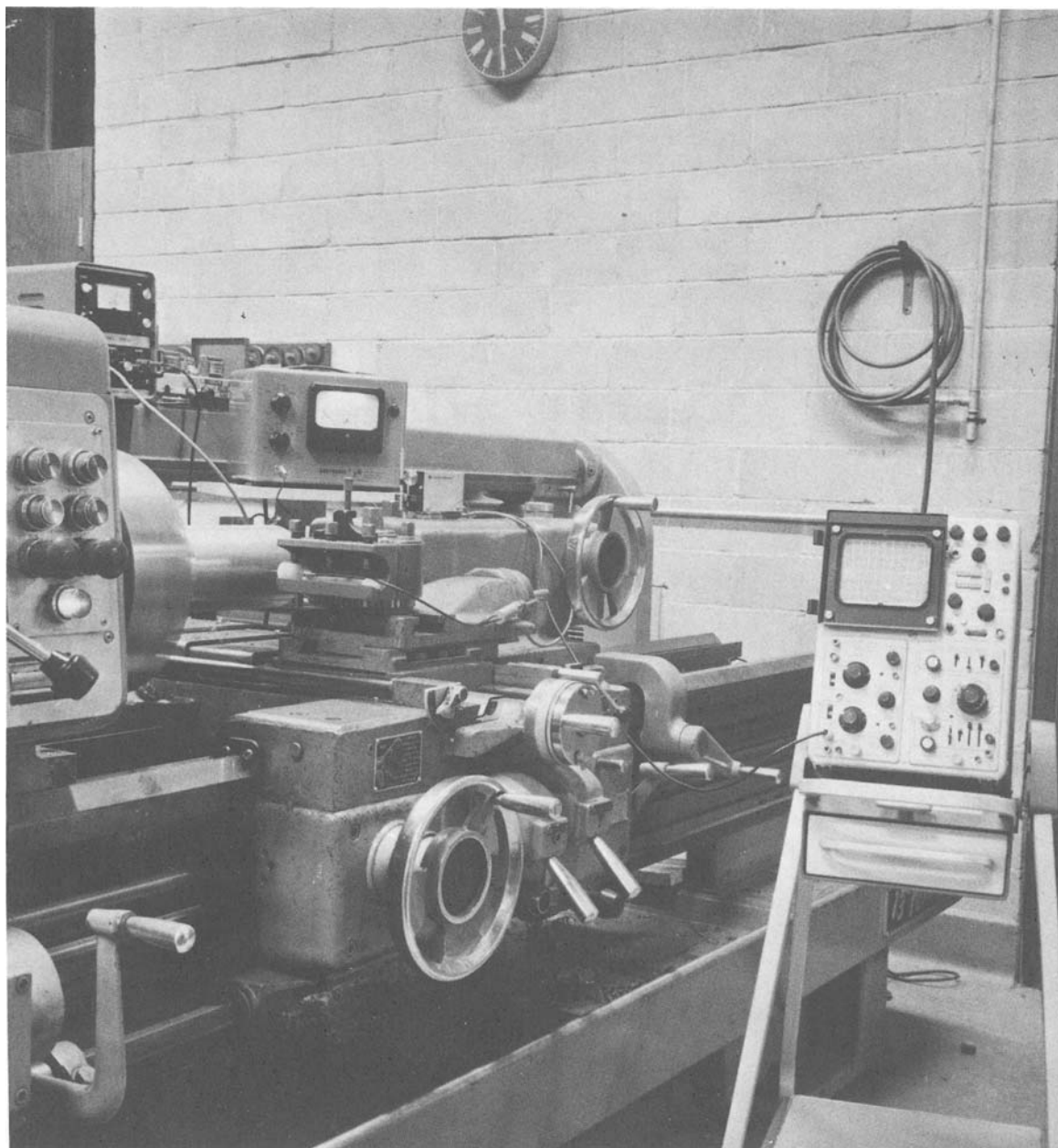


Figure 2. General arrangement of the experimental equipment.

accuracy of the results obtained using the portable instrument. Wear on the primary cutting edge was measured by means of a toolmaker's microscope fitted with a filar eyepiece and an *X-Y* co-ordinate table. Both direct and indirect lighting was fitted, enabling illumination of the tool clearance face. Results of these measurements are not reported in the present paper. Duration of the cut was measured using a time recording clock, capable of indicating to an accuracy of ± 0.5 seconds.

In-process measurement of surface roughness was obtained using the transducer previously referred to. A 1.0 mm diameter probe was used and the output passed to a storage oscilloscope, via a filter. Photographs of the oscilloscope traces were recorded using a camera fitted over the oscilloscope screen.

A photograph illustrating the general arrangement of the experimental equipment is shown in figure 2 and a close-up view of the transducer in the scanning position is presented in figure 3.

A Sandvik, positive rake toolholder type 174.2 with a built-in chip-breaker was used throughout the work. Sandvik, Coromant inserts were used of P30 grade type SPUN 12-03-04. Tool angles for these inserts were as follows:

Normal rake angle	+6°
Normal clearance angle	+5°
Cutting edge inclination	0
Cutting edge angle	75°
Included angle	90°

The corner radius used was 0.8 mm.

All tests were carried out on steel bars 100 mm diameter \times 1 m in length to specification En 8.

The actual composition of the steel used was as follows:

C: 0.46%, Si: 0.20%, Mn: 0.58%, S: 0.015%, P: 0.02%

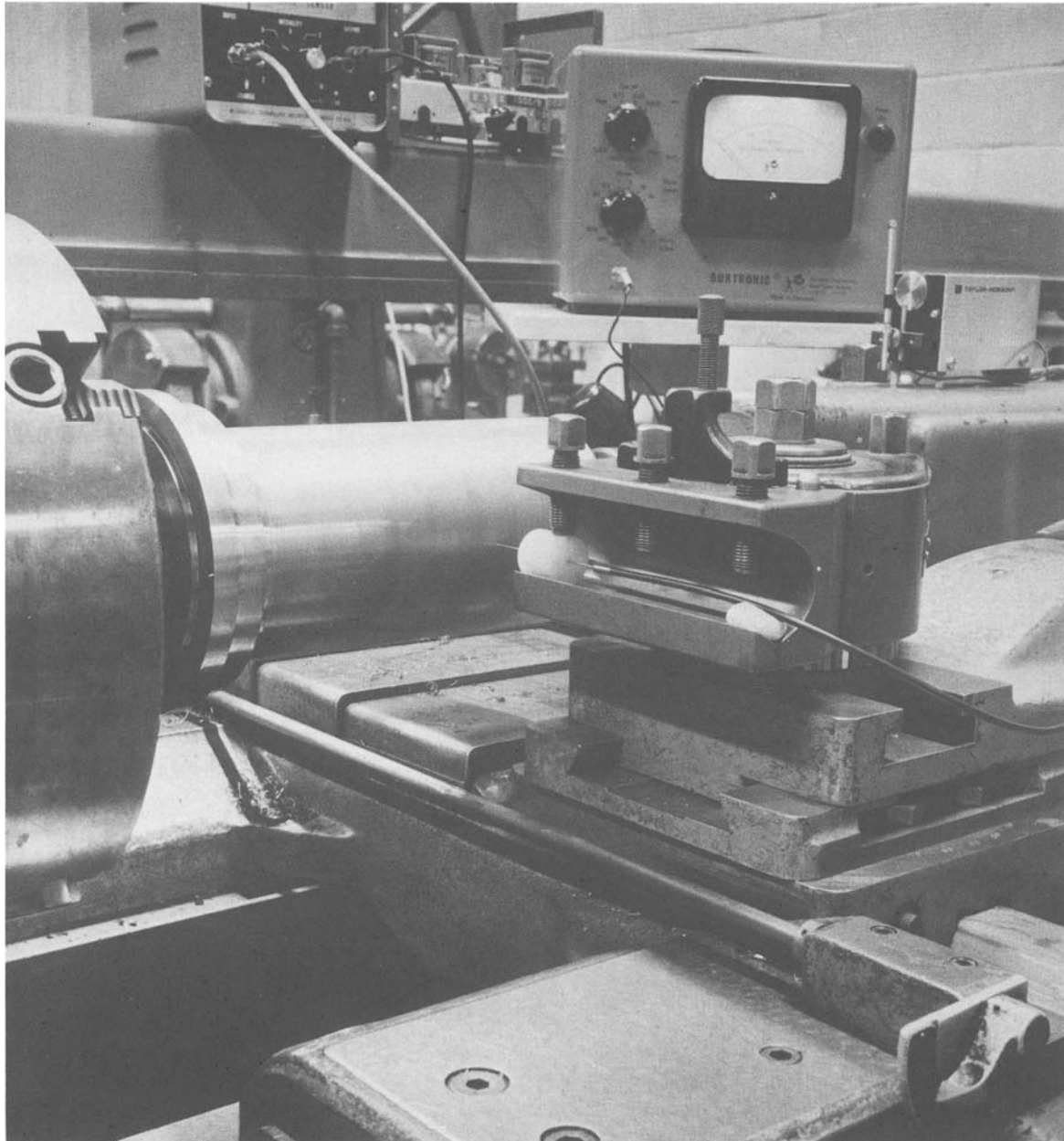


Figure 3. Close-up view of the transducer in the scanning position.

EXPERIMENTAL PROCEDURE

Initially a number of tests were performed in order to establish the range of speeds and feeds to be used. For the purpose of finish turning and to ensure that the surface roughness was not significantly affected by the cutting speed a minimum cutting speed of 250 m/min was adopted. Four cutting speeds were chosen from the R20 series, namely; 250 m/min, 280 m/min, 315 m/min, 350 m/min. To span the range of acceptable surface roughness values, values of feed rate were chosen such that the roughness obtained with a new tool did not exceed $3 \mu\text{m}$, a value which was considered to be extreme for finish turning.

Four feed rates were chosen as follows; 0.1 mm/rev., 0.16 mm/rev, 0.20 mm/rev, 0.25 mm/rev.

The depth of cut was maintained constant at 2 mm.

Readings of surface roughness were taken approxi-

mately every minute using the portable measuring instrument. All readings were taken on the same portion of bar, about 50 mm from the chuck, in order to eliminate possible errors due to vibration. Six readings were taken on each occasion at equal distances around the bar circumference and the mean of these readings were recorded. This procedure was repeated for each of the four cutting speeds and feed rates.

From these results, graphs were produced showing the relation between surface roughness and time for the four constant cutting speeds and constant feed rates of 0.1 mm/rev, 0.16 mm/rev and 0.02 mm/rev. No graphs were plotted for the fourth feed rate since deterioration of the workpiece surface was extremely rapid.

The surface produced after each test was scanned by the transducer in place of the cutting tool. The output from the return branch of the probe was

passed through a filter enabling frequencies between 100 Hz and 20 000 Hz to be selected. The time base of the oscilloscope was set on each occasion to record one revolution of the workpiece. Photographs were taken of each trace recorded using the camera set to a lens aperture of 5.6 and using an exposure value of 0.1 sec. The average value of the transducer output was obtained from the photograph for each pass and the results used to produce a graph showing the relation between the transducer output and feed rate. A regressional analysis was performed on these results to obtain the relation between the transducer output and feed rate. The validity of the results was assessed using statistical techniques.

RESULTS AND DISCUSSION

A typical set of results indicating the variation of surface roughness with time is presented in figure 4 for a constant cutting speed of 280 m/min and feed rates of 0.1, 0.16 and 0.20 mm/rev. There are three distinct regimes in which the surface roughness is shown to either increase with time or remain sensibly constant. Initially, an improvement in the surface roughness occurs but this improvement is of short duration to be followed by a prolonged period of increase in roughness at a sensibly constant rate. This rate of increase becomes greater at both higher feed rates and cutting speeds. A period of zero increase in surface roughness with time then follows although the cutting tool is still being worn on its major

cutting edge. In general, for a particular cutting speed the length of this plateau decreases as the feed rate increases. There does not appear to be any marked change in the size of the plateau as the cutting speed increases. After a considerable time the surface roughness again increases, indicating that the grooving wear on the tool nose radius is increasing to a condition when rapid breakdown of the cutting edge will occur. This effect was found to be more significant at higher cutting speeds. In all cases, the initial value of surface roughness was found to be greater than the theoretical value predicted by the well known equation

$$R_a = 32.1 s^2/r \tag{1}$$

where R_a = the surface roughness
 s = the feed rate
 r = the tool nose radius

The present results are in general agreement with those reported by Olsen⁷ at feed rates below 0.25 mm/rev to 0.30 mm/rev where a sharp tool produces scratching in the bottom of the profile and may also induce vibration leading to a deterioration of the surface. The values of the initial surface roughness were independent of cutting speed thus giving some substance to the validity of Equation (1). Support for the existence of a period of stable surface roughness has come from Pekelharing and Luttervelt⁸ and Lambert⁹. However, these authors indicate a more substantial plateau than was found during the current investigations. A typical graph showing the variation of initial surface roughness (after 15 seconds machining) with feed rate is shown in figure 5. This indicates a relationship similar to that predicted by Equation (1). It will be seen from figure 5, however, that the experimental values deviate from the theoretical curve, this deviation increases with feed rate. Cutting speed was found to have little effect upon this deviation.

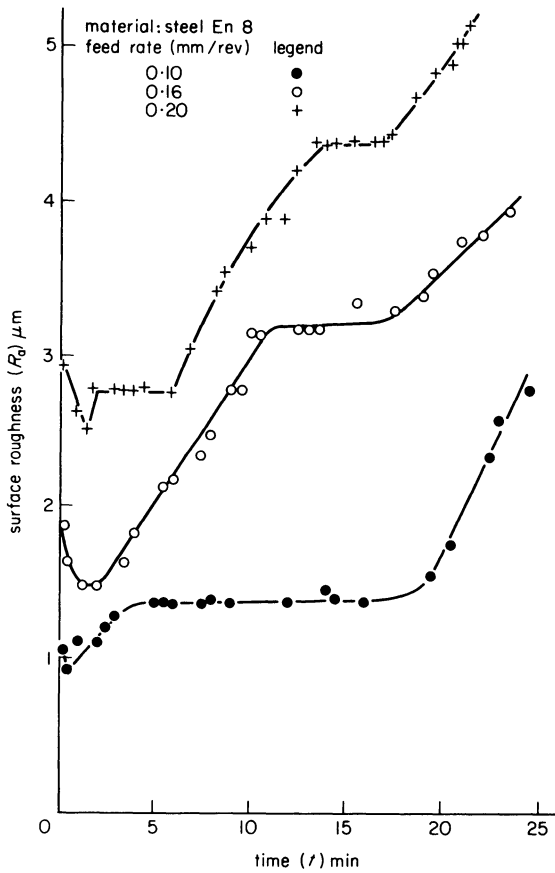


Figure 4. Relation between surface roughness (R_a) μm and time (t) min for a constant cutting speed of 280 m/min with varying feed rates.

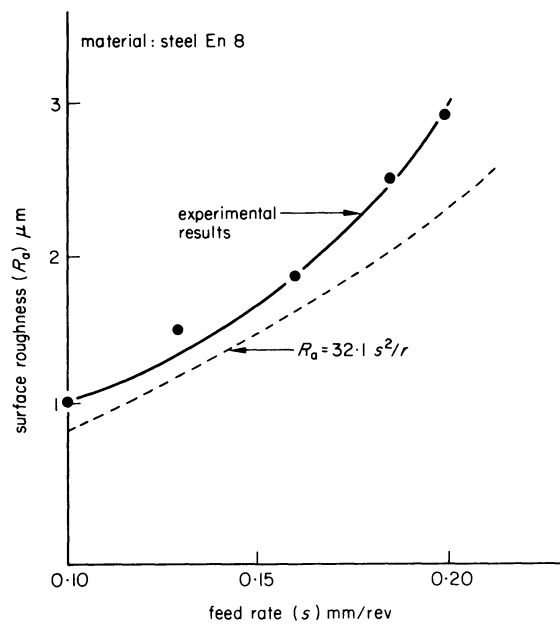


Figure 5. Relation between surface roughness (R_a) μm and feed rate (s) mm/rev after machining for a period of 15 seconds.

The results of the regression analysis of surface roughness upon feed rate yielded a relationship of the form

$$R_a = 24.8 s^{1.46} \quad (2)$$

Tests for significance on the results indicated that, for a 95 per cent confidence level, the results could be considered as being significant.

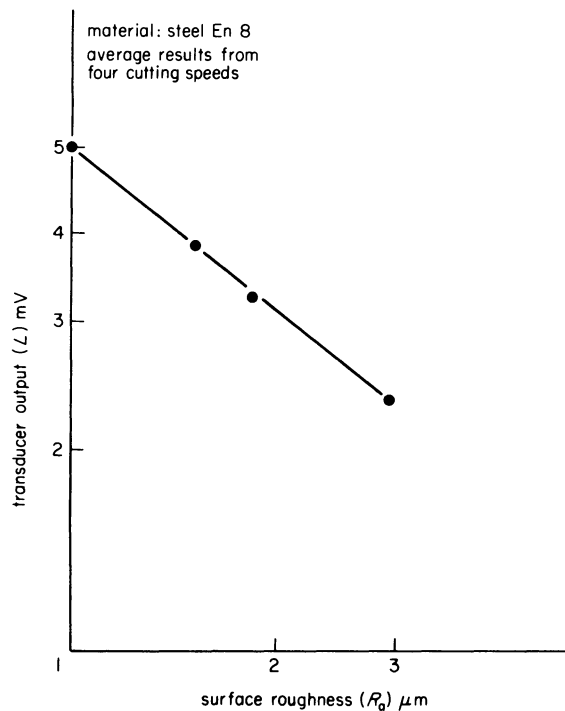


Figure 6. Logarithmic relation between the transducer output (L) mV and surface roughness (R_a) μm .

The variation of the transducer output with surface roughness is presented in logarithmic form in figure 6. The results of a regression analysis of the relation between the transducer output and surface roughness yielded the following equation

$$LR_a^{0.724} = 5.1 \quad (3)$$

where L = output from the transducer (mV)
 R_a = surface roughness (μm)

A statistical analysis, at a confidence level of 95 per cent, indicated that the results obtained were significant. It should not be expected that all values of surface roughness determined in this way are very accurate in inspection terms. However, they will be sufficiently accurate to provide in-process information for an adaptive control system.

Of the large number of photographic records obtained showing the waveforms representing the workpiece surface, a typical selection is given in figure 7. In all cases, the output from the transducer reduced as the surface roughness value increased. There was some variation in the output at different cutting speeds but this was so slight as to be

considered insignificant. It is not unexpected that some variation should exist since the results were derived over one revolution of the workpiece only. To obtain an average value, as is obtained conventionally using a 'Talysurf' surface roughness measuring instrument for example, results should be taken over a larger number of revolutions, typically 37 revolutions at a cutting speed of 280 m/min and a feed rate of 0.20 mm/rev.

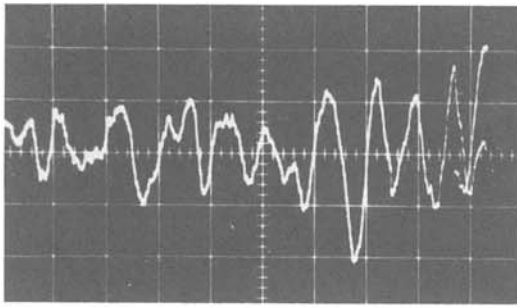
Figure 8 shows the relation between the experimental values of surface roughness and transducer output obtained whilst machining. Since the slope of the graph at the higher values of surface roughness is greater than the slope at the lower values it would seem that the use of this type of transducer should be limited to the lower range of surface roughness. This will tend to limit the use of the device to finish machining applications and this series of tests indicate that it should not be expected to give satisfactory results on surfaces for which the surface roughness exceeds, say, 3 μm .

It is therefore possible to monitor changes in the surface roughness of a machined surface using an optical transducer of the type used in the present study provided its use is limited to surfaces having a roughness value better than 3 μm . Therefore, such a transducer is particularly suitable for the finish turning process where it is required to maintain the surface roughness within limits over a long pass. In such circumstances, accuracy is best maintained if the work can be completed without changing the tool. Such a situation might exist in the case of NC turning where it is often advantageous to change the tool at the end of a tape run rather than at any other time.

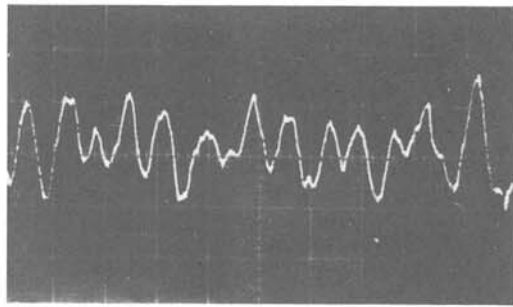
OPTIMIZATION OF THE PROCESS

As indicated in figure 4, the surface roughness of a turned specimen increases with time resulting from tool wear but the initial surface roughness is mainly dependent upon the tool geometry and the chosen feed rate. However, it does not follow that because the feed rate is limited by the surface roughness value, optimum or near optimum conditions cannot be obtained. True optimum conditions depend upon the proper choice of cutting speed, feed rate and depth of cut. Of these, depth of cut is limited in the finish turning process and therefore only cutting speed and feed rate have to be considered. Both of these parameters affect machining cost, and feed rate also affects the quality of the turned surface. Cutting speed also has some effect upon the surface quality in so far as it affects the wear of the tool and hence the rate at which the surface roughness increases. Thus, as the tool wears and the feed rate is reduced in an attempt to maintain the surface roughness value constant, it is possible to use the cutting speed as a means of obtaining better utilization of the machine.

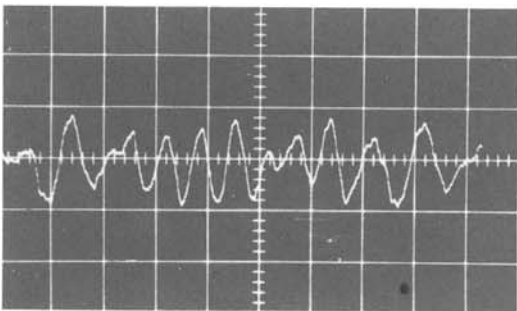
Choosing machining cost as a suitable index of performance an equation can be produced relating all the cutting parameters to this index. The equation is then optimized for cost, the values of cutting speed and feed rate being chosen such that machining may be carried out under conditions of maximum metal



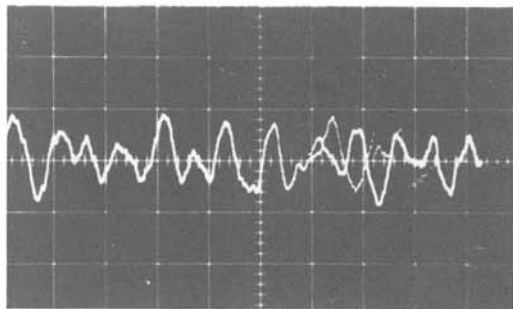
$V = 315 \text{ m/min}, s = 0.10 \text{ mm/rev}$
5 mV



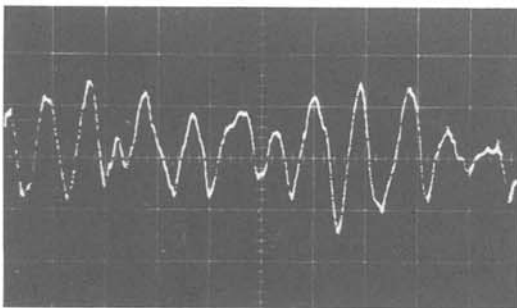
$V = 250 \text{ m/min}, s = 0.13 \text{ mm/rev}$
3.75 mV



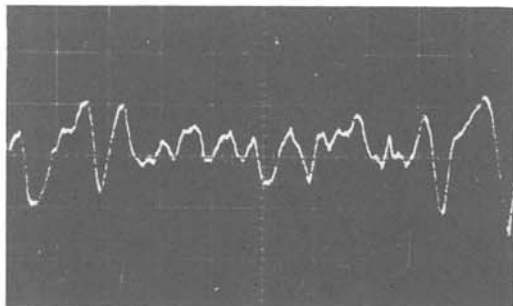
$V = 280 \text{ m/min}, s = 0.16 \text{ mm/rev}$
3.3 mV



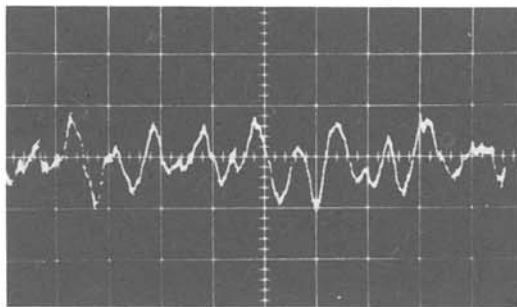
$V = 315 \text{ m/min}, s = 0.20 \text{ mm/rev}$
3 mV



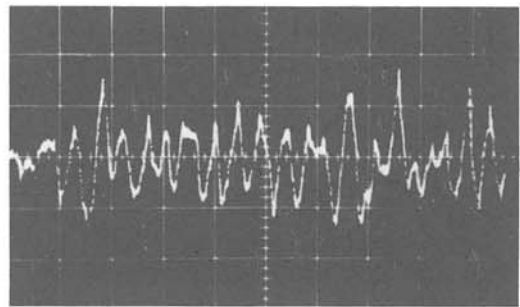
$V = 250 \text{ m/min}, s = 0.10 \text{ mm/rev}$
5.1 mV



$V = 350 \text{ m/min}, s = 0.16 \text{ mm/rev}$
3.2 mV



$V = 210 \text{ m/min}, s = 0.148 \text{ mm/rev}$
2.7 mV



$V = 210 \text{ m/min}, s = 0.164 \text{ mm/rev}$
3.1 mV

Figure 7. Typical selection of photographic records of waveforms representing the workpiece surface.

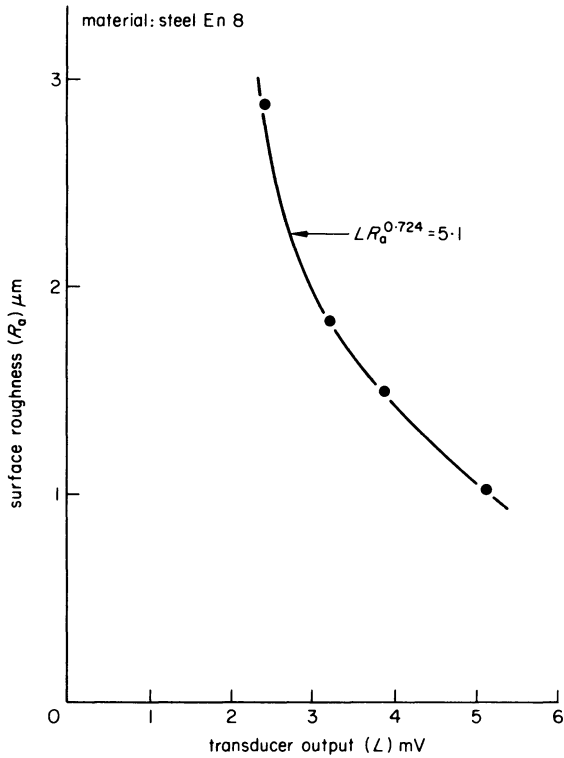


Figure 8. Relation between the experimental values of surface roughness (R_a) μm and the transducer output (L) mV.

removal for each chargeable minute. An appropriate equation is

$$R = Q / \{a + (at_1 + C/N)(1/T)\} \quad (4)$$

where R = metal removal rate

Q = volume removed per minute

T = tool life in minutes

C = tool cost in pence

N = number of edges

a = labour rate

t_1 = time to replace worn tool in minutes

A further constraint is imposed upon the value of the feed rate possible due to the specification for surface quality. This relationship has been expressed by Equation (2). It is therefore possible to produce an initial set of conditions which ensure optimum working from a cost point of view. A measure of adaptive control should be obtained by in-process measurement of the surface roughness and, by making alterations to the cutting speed and feed rate adopted, meet the surface roughness requirements whilst, at the same time, ensuring the most economic conditions possible.

To test the validity of the above hypothesis a computer program was formulated to simulate the work of such an adaptive control system. Such a system would have data concerning tool life, surface roughness and transducer details contained within the package as basic data. The information concerned with surface roughness variation can then be converted to digital form and compared with the reference information. Control algorithms would then process this information and compute new values of cutting speed and feed rate which would then be used to drive the servo motors of the machine tool. The

sensitivity of the system can be simply controlled by the algorithms used and, by the continuous scanning of the workpiece surface with the optical transducer, there need be no large or sudden changes in either the cutting speed or the feed rate.

A schematic diagram of a suggested adaptive control system using the above principles is given in figure 9. The system is extremely simple and low in cost provided a small digital computer is available. The results of the simulation are too numerous and beyond the scope of the present paper but the effects

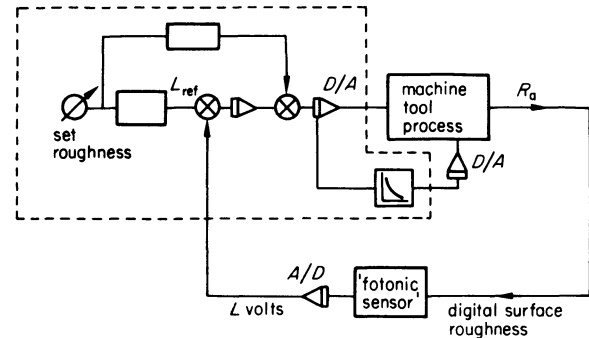


Figure 9. Schematic diagram of a suggested adaptive control system monitoring surface roughness.

may be simply stated in the following manner. During the short initial period when the surface roughness improves it is possible to increase the feed rate slightly. However, as the surface roughness increases the feed rate must be reduced in order to maintain the specified value. The cutting speed is also increased in order to ensure that the operation is carried out under conditions which are as near optimum as possible in terms of machining cost.

CONCLUSIONS

From the results of this study a number of conclusions may be drawn, supported by statistical analyses, which are expected to be helpful in the design of an adaptive control system for the finish turning process. These conclusions may be stated as follows

- (1) A relationship exists between initial surface roughness and feed rate for En 8 steel which is practically independent of cutting speed. This relationship, for the particular set of conditions employed, may be expressed by the equation

$$R_a = 24.8s^{1.46}$$

- 2. A transducer based upon light reflection may be used to monitor the variation in surface roughness of a newly turned surface. The output from such a device reduces as the surface roughness increases in value. For the particular transducer used a relationship between transducer output and surface roughness may be expressed by the equation

$$LR_a^{0.724} = 5.1$$

3. Using these results it is possible to produce a set of control algorithms for a computer simulation of an adaptive control device. The results from such a simulation can be shown to compare favourably with the experimental results obtained during these tests.
4. The indications are that a practical system could be produced quite cheaply using an optical transducer and feeding the output directly into a small, general purpose, on-line computer. The output from the computer can then be used as information to control the servo motors of the lathe.

Future investigations

This paper has attempted to show that it is possible to obtain meaningful information about the surface roughness of a finish turned surface by continuously monitoring the newly turned surface using an optical transducer. Before a more advanced control system can be designed and built further work is required to be carried out along the following lines

1. The present work has been limited to one material, namely, En 8, steel. Further work is required using other steels together with the light alloys and yellow metals. Since the method depends upon the reflectivity of the surface, the relationship between transducer output and surface roughness will undoubtedly change.
2. A modified circuit is required in the output line from the transducer in order to derive the mean voltage over a large number of revolutions. This mean voltage would then be used to represent the surface in a similar way to the conventional CLA value derived by stylus instruments.
3. Further work should be carried out using different size probes. The present work using a probe of 1 mm diameter meant that a number of turned grooves were scanned which, to some extent, results in an averaging effect. However, if in fact, the output from the transducer is governed by the amount of light absorbed in the valleys of the turned surface, then a larger diameter probe could be expected to improve the accuracy of surface roughness indication. In addition, the larger the diameter of probe that can be used the more robust it will be and hence the more acceptable for industrial application.

ACKNOWLEDGMENTS

The authors offer their thanks to Professor J. C. Levy, Head of Department of Mechanical Engineering, The City University, for making it possible for the study to be undertaken.

One of the authors (D.S.) is indebted to the Hatfield Polytechnic for financial assistance and to Mr Gerald Smith, Head of Department of Industrial Engineering at that Institution for his personal encouragement to participate in the study.

Sincere thanks are due to Mr H. V. Marsh of the Manufacturing Processes Laboratory, The City University, for his technical assistance during the experimental work, to Mr J. Mattick who prepared the illustrations and to Miss Susan Balding who was responsible for typing the manuscript.

REFERENCES

1. M. E. Merchant. Trends in manufacturing systems. *Proc. 10th Int. Mach. Tool Des. and Res. Conf.* University of Manchester. Pergamon Press, Oxford, p. 567 (1970).
2. Anon. Adaptive control package debuts *Metal-working Production* (1968) 112, no. 35, 49.
3. M. Bath and R. Sharp. In process control of lathes improves accuracy and productivity. *9th Int. Mach. Tool Des. and Res. Conf.* University of Birmingham. Pergamon Press, Oxford, Part II p. 1209 (1969).
4. N. F. Shillam. Closed loop control of lathe speeds and feeds using the tool as temperature and force sensor. *RAE Tech. Memo AD/W 20*, December (1969).
5. M. G. Jona. Surface finish measurements in turning and their mathematical interpretation. *Annual meeting of Energy Systems Exposition of Am. Soc. mech. Engrs.* Pittsburgh USA, Nov. 12-17 (1967).
6. M. G. Jona. Contribution to the development of geometrical adaptive control in turning. *Proc. 11th Int. Mach. Tool Des. and Res. Conf.* University of Birmingham. Pergamon Press, Oxford, Vol. A, p. 429 (1971).
7. K. V. Olsen. Surface roughness of turned steel components and the relevant mathematical analysis. *The Prod. Engr.* (1968) 47, no. 12, 593.
8. A. J. Pekelharing and C. A. Luttervelt. Wear of finish turning tools. *OECD Seminar on Metal Cutting*. Paris. Sept (1966).
9. H. J. Lambert. Two years of finish turning research at the Technological University of Delft. *CIRP Annalen* (1961-62) p. 246.

PERFORMANCE ANALYSIS OF THE STYLUS TECHNIQUE OF SURFACE ROUGHNESS ASSESSMENT: A RANDOM FIELD APPROACH

by

J. B. AGULLO* and J. PAGES-FITA*

SUMMARY

This paper deals with one of the weak points of surface roughness assessment procedures using the stylus technique—obtaining the signal representative of surface profiles.

The stylus geometry, as well as other mechanical characteristics of the stylus system, together with the surface geometry, may cause the signal obtained to be quite different from the real profile. This may affect the statistical characteristics obtained for the surface. The effect on several standard roughness parameters is studied by means of a random field approach.

Conclusions are drawn with respect to the surface textures and roughness ranges in which this effect is more important. Experimental results are included.

NOMENCLATURE AND TERMINOLOGY

Nominal geometrical surface. The surface determined by the design or by the process of manufacture, neglecting errors of form and roughness.

Effective surface. The representation of a real surface obtained by instrumental means.

Effective profile. The contour that results from the intersection of the effective surface by a plane conventionally defined with respect to the Nominal geometrical surface.

		$x'(u, v)$	Vector representation of the secondary effective surface
		u, v	position parameters
		x_1, x_2, x_3	Cartesian orthogonal coordinates, along the profile, transversal to the profile, and normal to the nominal geometrical surface respectively
		$f(u), f(u, v)$	Random process, random field, representing a rough surface
		dA	differential of area of $x(u, v)$
		λ	factor of area, defined in Equation (5)
Ra	arithmetical mean deviation from the mean line of the profile. Deviations are averaged without considering their algebraic sign	ϵ	level variation introduced by the geometrical transformation
Rq	root mean square average of the deviation from the mean line of the profile.	ρ	stylus tip radius
Rt or R_{max}	the distance between two lines parallel to the mean line and touching the profile at the highest and lowest points, respectively, within the sampling length	C_m	local mean curvature
		C_G	local Gaussian curvature
		N	local external normal versor of the rough surface
		v	external normal versor of the nominal geometrical surface
		σ	standard deviation of $f(u)$ or $f(u, v)$
		$\sigma_u, \sigma_{uu}, \sigma_{uv}$	standard deviations of f_u, f_{uu}, f_{uv}
		$R(u), R(u, v)$	autocorrelation function of $f(u), f(u, v)$
Rp	The distance between the mean line and its parallel touching the profile at the highest point within the sampling length	$g(x)$	Gauss function. $g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$
$x(u, v)$	vector representation of the primary effective surface. It also represents real surfaces such that all their points can be in contact with the stylus	$G(x)$	error function + $\frac{1}{2}$; $G(x) = \int_{-\infty}^x g(u) du$
		E	mathematical expectation
		F	normal stylus force
		α	shank angle

* Department of Mechanical Engineering, ETSIIB, Barcelona, Spain.

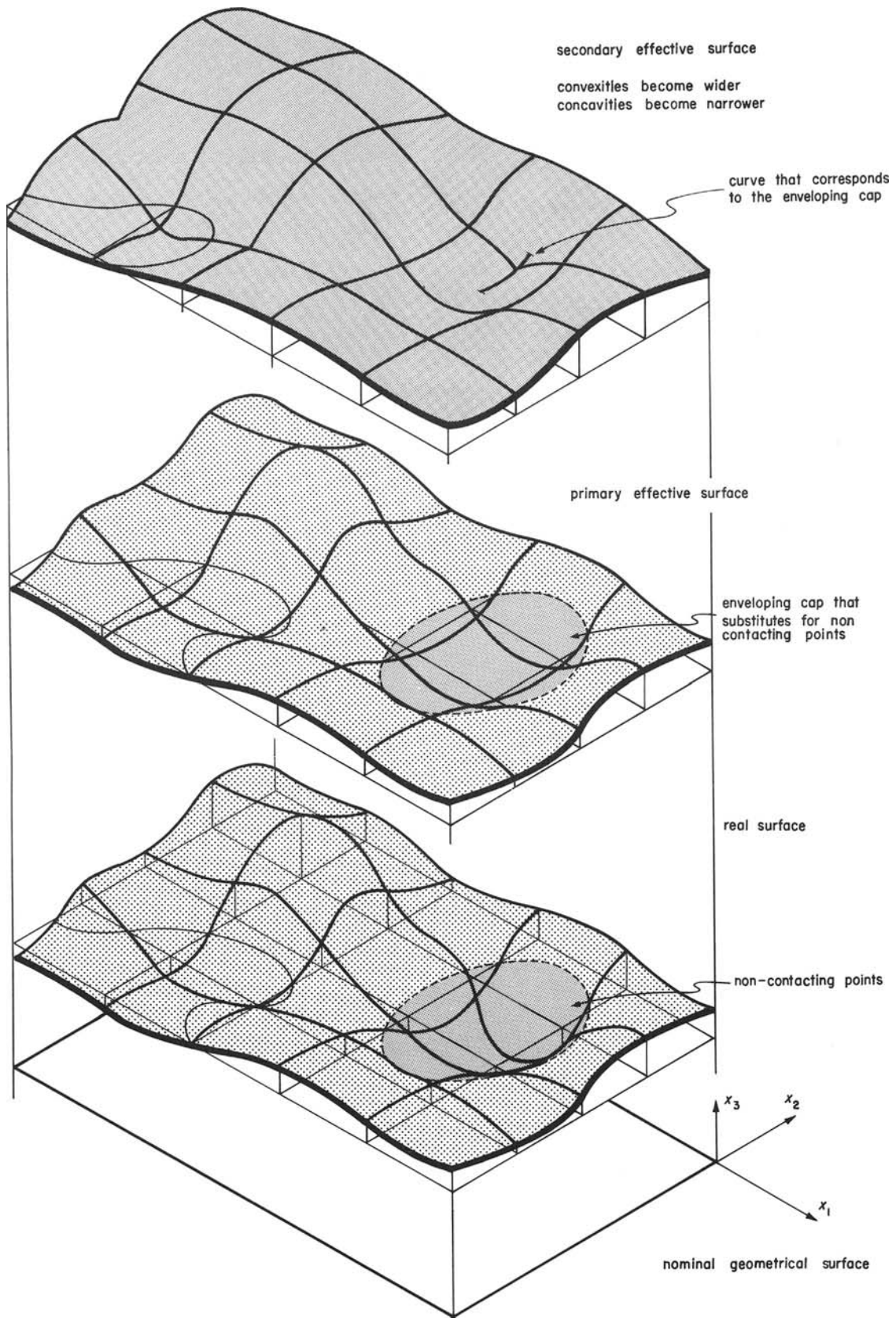


Figure 1. Effective surfaces associated to a rough surface and to a given stylus.

δ_H lateral deviation
 δ_v incremental vertical displacement due to lateral deflection of the stylus

INTRODUCTION

Although a variety of methods of assessing the surface roughness have been devised which employ optical, pneumatic, mechanical, acoustical and other physical phenomena, the most widely used method concerns the microgeometry of the surface and depends on the electrical amplification of the motion of a stylus perpendicular to the surface traversed by the stylus.

In normal practice observation is restricted to the exploration of profiles from which standardized parameters such as Ra , Rr , Rp , Rq , Rz , $t(c)$, are determined. For research purposes other parameters as well as functions are considered, such as the probability density function, autocorrelation function, and power spectral density¹⁻¹¹.

The motion of the stylus perpendicular to the nominal geometrical surface is considered to be the *effective profile*. This assumption allows time, instead of displacement, to be used in recording and to perform such operations as averaging, filtering and autocorrelation. Consequently substantial use can be made of electronic networks to implement them. Nevertheless the stylus motion differs from the so-called *real profile*, and also from the closest obtainable approximation to it for a given stylus, which is a profile of a surface all points of which can be in contact with the stylus.

Three surfaces can be conveniently defined (figure 1):

(1) *Real surface*. From a practical point of view the real surface can be defined as the surface that a much sharper stylus would detect if physically possible.

(2) *Primary effective surface*. This is defined as the envelope of the stylus surface as it moves all over the rough surface. It is the closest approximation to the real surface obtainable by means of a given stylus. It can be identical to the real surface, for example for a conical-spherical stylus this occurs if the slopes and principal curvatures of the real surface are lower than the slope of its conical part and the curvatures of its spherical part respectively.

(3) *Secondary effective surface*. This is defined as the geometrical locus of the positions of a specific point on the stylus, for example the lower point, as it moves all over the rough surface. This surface coincides only asymptotically with the primary effective surface as the tip radius tends to zero.

Profiles used correspond to the secondary effective surface, and its statistical characteristics may differ from the ones of the primary effective surface or from those of the real surface. Thus the stylus operation introduces a transformation of the surface over which it moves, and this can modify the roughness parameters. Since the transformation is non-linear, its effects may be important for certain ranges of heights and widths of the surface irregularities, and some operations such as frequential

filtering can be misleading. The main object of this paper consists in the analysis of the effects of this transformation, which is intrinsic to the geometry of the stylus method, on the statistical characteristics of the surface involved and on the standardized roughness parameters.

Other factors concerned with the stylus mechanics can affect its performance: lateral deflections due to the lateral flexibility of the stylus, elastic-plastic behaviour of the rough surface under the stylus force, and dynamics of the stylus movement. A brief introductory review is given on the lateral deflection problem.

GEOMETRICAL TRANSFORMATION

Let $x(u, v)$ be a real surface such that all its points can be in contact with the spherical part of the stylus tip, and let x have first and second-order derivatives $x_u, x_v, x_{uu}, x_{vv}, x_{uv}$. Its primary effective surface will be therefore identical to the real surface.

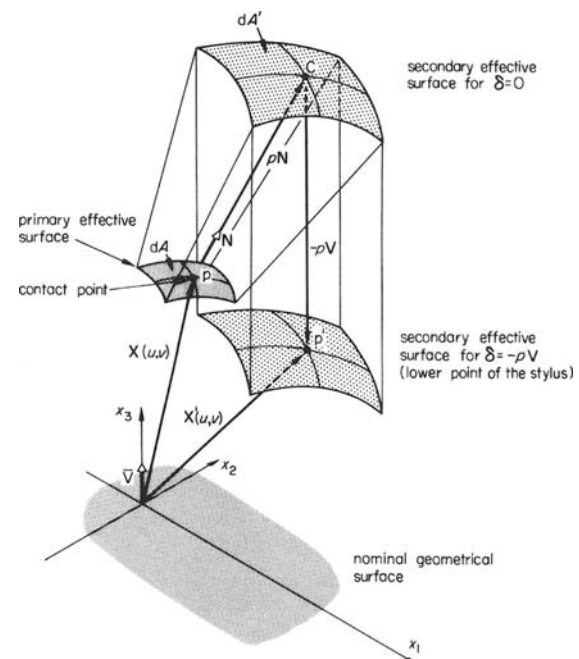


Figure 2. Geometrical transformation associated to stylus operation.

The centre of the spherical part of the stylus tip is at a distance ρ from the contact point P perpendicular to the surface, ρ being the radius of the sphere (figure 2). Any point P' of the stylus can be positioned from the centre C by the addition of a constant vector $\delta = \vec{CP}'$; for the lowest point of the stylus $\delta = -\rho v$, v being the external normal vector to the nominal geometrical surface. The secondary effective surface is described by the equation

$$x' = x + \rho N - \rho v \tag{1}$$

where N is the external normal vector to the primary effective surface. By means of this transformation (Appendix 1) each differential of surface dA of the primary effective surface is transformed to a differential of surface dA' of the secondary effective

surface, related to dA by the area factor

$$\begin{aligned}\lambda &= dA'/dA \\ &= 1 - 2\rho C_m - \rho^2 C_G^2\end{aligned}\quad (2)$$

where C_m and C_G are the mean and Gaussian local curvatures respectively. The level is modified by the amount

$$\epsilon = -\rho(1 - \mathbf{v} \cdot \mathbf{N})\quad (3)$$

Parameters u , v can be associated to orthogonal Cartesian coordinates on the nominal geometrical surface, which is assumed to be plane:

$$x_1 = u, \quad x_2 = v, \quad x_3 = f(u, v)\quad (4)$$

where $f(u, v)$ is a bidimensional random field. Equations (2) and (3) can then be expressed as

$$\begin{aligned}\lambda &= dA'/dA \\ &= 1 - \rho \frac{(1 + f_v^2)f_{uu} + (1 + f_u^2)f_{vv} - 2f_{uv}f_{uv}}{(1 + f_u^2 + f_v^2)^{3/2}} \\ &\quad + \rho^2 \frac{f_{uu}f_{vv} - f_{uv}^2}{(1 + f_u^2 + f_v^2)^2}\end{aligned}\quad (5)$$

$$\epsilon = -\rho \left[1 - \frac{1}{\sqrt{(1 + f_u^2 + f_v^2)}} \right]\quad (6)$$

Anisotropic surfaces with parallel-lay (PL) texture can be idealized by means of surfaces generated by parallel straight lines. In this case the local Gaussian curvature is zero. If x_2 is chosen in the direction of lay, Equations (5) and (6) reduce to

$$\left(\frac{dA'}{dA} \right)_{\text{PL}} = 1 + \rho \frac{f_{uu}}{(1 + f_u^2)^{3/2}}\quad (7)$$

$$\epsilon = -\rho \left[1 - \frac{1}{\sqrt{(1 + f_u^2)}} \right]\quad (8)$$

where $f(u)$ is a random process.

To ensure stylus contact all over the surface, the following conditions must be satisfied:

$$\sqrt{(f_u^2 + f_v^2)} \leq \cot \frac{\alpha}{2}\quad (9)$$

where α is the shank angle, related to the slopes, and

$$C_m + \sqrt{(C_m^2 - C_G^2)} \leq \frac{1}{\rho}\quad (10)$$

related to the curvatures.

In the case of surfaces generated by parallel straight lines, Equations (9) and (10) become

$$|f_u| \leq \cot \frac{\alpha}{2}\quad (11)$$

$$\frac{f_{uu}}{\sqrt{(1 + f_u^2)}} \leq \frac{1}{\rho}\quad (12)$$

FREQUENCY RESPONSE OF THE GEOMETRIC TRANSFORMATION

The geometric transformation defined by Equation (1) is non-linear, and consequently it does not verify the superposition property. Due to this fact the spectrum analysis carried out from the oscillograms

obtained from the stylus can be misleading. The spectrum obtained corresponds to profiles of the secondary effective surface and is not easily related to the spectrum of profiles of the real surface.

To illustrate this non-linear behaviour figure 3 shows computed spectra of profiles that would be obtained by means of a stylus orthogonally traversing parallel-lay rough surfaces with sinusoidal real profile of equal amplitude but different wavelengths. A stylus tip radius of $2.5 \mu\text{m}$ and amplitude (half peak-to-peak) of $0.157 \mu\text{m}$, which corresponds to rough grade N3, have been assumed. Figure 3 clearly shows high harmonics generation and progressive attenuation of the input profile.

RANDOM VARIABLE APPROACH FOR PARALLEL-LAY TEXTURES

A usual first approximation is to assume that the random process $f(u)$ used to characterize parallel-lay textures, is a stationary twice-derivable Gaussian process¹³. Let us also assume that the joint probability distribution of $f(u)$ and its first and second derivatives f_u and f_{uu} is also Gaussian. The distribution is characterized by the mean of $f(u)$, which can be taken equal to zero, and the standard deviations σ , σ_u , and σ_{uu} of f , f_u and f_{uu} respectively.

From this joint probability distribution the statistical characteristics of the random process f' obtained by means of the transformation (1) applied to $f(u)$ can be determined. The distribution function and parameters Ra and Rq are shown in Appendix 2 to be given by the expressions

$$\begin{aligned}P_{f'}(\mu) &= \frac{1}{\sigma_u} \int_{-\infty}^{\infty} g\left(\frac{\eta}{\sigma_u}\right) \left\{ G\left[\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma\sqrt{(1 + \eta^2)}}\right] \right. \\ &\quad \left. - \frac{\rho\sigma_u^2}{\sigma(1 + \eta^2)^{3/2}} g\left[\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma\sqrt{(1 + \eta^2)}}\right] \right\} d\eta\end{aligned}\quad (13)$$

$$\begin{aligned}R'a &= m' + \frac{2\rho}{\sigma_u} \int_{-\infty}^{\infty} g\left(\frac{\eta}{\sigma_u}\right) \left\{ \frac{\sigma}{\rho} g\left[\frac{\rho}{\sigma} + \frac{m'}{\sigma} - \frac{\rho}{\sigma\sqrt{(1 + \eta^2)}}\right] \right. \\ &\quad \left. + G\left[\frac{\rho}{\sigma} + \frac{m'}{\sigma} - \frac{\rho}{\sigma\sqrt{(1 + \eta^2)}}\right] \right. \\ &\quad \left. \times \left[1 + \frac{m'}{\rho} - \frac{1}{\sqrt{(1 + \eta^2)}} \left(1 + \frac{\sigma_u^2}{1 + \eta^2} \right) \right] \right\} d\eta\end{aligned}\quad (14)$$

$$R'q^2 = \sigma^2 - (\rho + m')^2 + \frac{\rho^2}{\sigma_u} \int_{-\infty}^{\infty} \frac{g(\eta/\sigma_u)}{1 + \eta^2} \left(1 + \frac{2\sigma_u^2}{1 + \eta^2} \right) d\eta\quad (15)$$

m' being the mean value of f' given by

$$m' = \frac{\rho}{\sigma_u} \int_{-\infty}^{\infty} g\left(\frac{\eta}{\sigma_u}\right) \frac{1}{\sqrt{(1 + \eta^2)}} \left(1 + \frac{\sigma_u^2}{1 + \eta^2} \right) d\eta - \rho\quad (16)$$

where g is the Gauss function and G is the error function $+ \frac{1}{2}$.

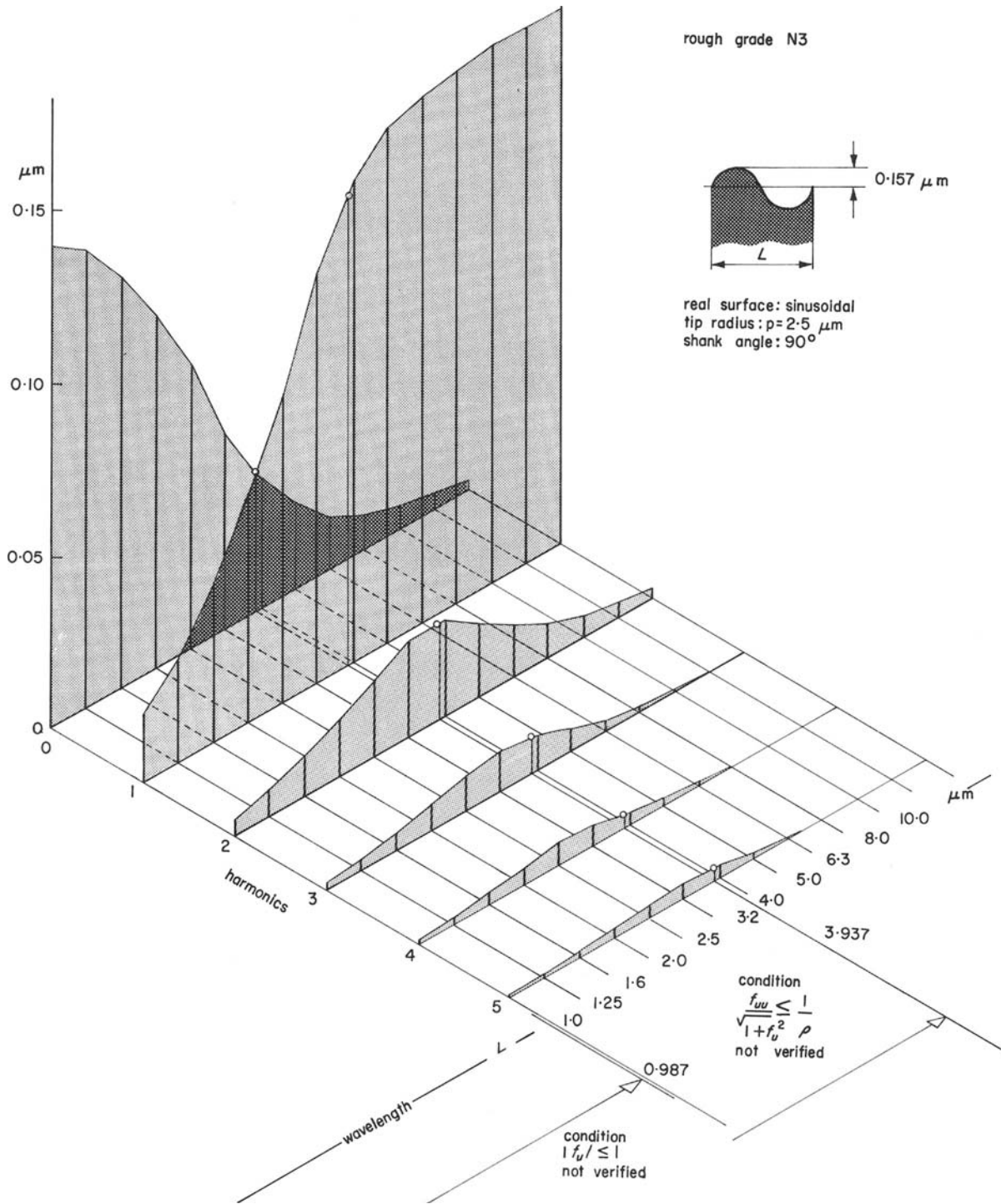


Figure 3. Frequency response of the non-linear geometric transformation.

It is important to note that the distribution function, and consequently all statistical characteristics, depend only on σ , σ_u and the tip radius ρ . Incidentally, the incremental value of Rq does not depend on σ . Unfortunately, the validity of the Gaussian approach, in spite of its advantage of allowing a complete analytical treatment, decreases as the interest of the corrections involved by the geometrical transformation increases. The random Gaussian process $f(u)$ does not verify conditions (11) and (12), which are required for primary effective surfaces at an increasing fraction of points.

To determine the statistical characteristics of the secondary effective surface from the statistical

characteristics of the real surface is a very hard problem, even if the real surface is assumed to be Gaussian. It involves finding the probability of a random process that remains under a given deterministic curve.

To overcome the complexity of the analytical treatment, a straightforward approach is envisaged through a direct simulation of the surface detection by the stylus. In the geometrical transformation approach simulation provides a means of relating the statistical characteristics of the primary and secondary effective surfaces. The primary effective surfaces used must strictly fulfil conditions (11) and (12).

To generate pairs of mating effective surfaces a

Monte Carlo technique has been used. By applying adequate weighting coefficients, points from a Gaussian process directly generate points of the primary effective surface, and by means of Equation (1) they also generate points of the secondary effective surface. In generating the secondary effective surface, weighting coefficients are equal to λ for points verifying conditions (11) and (12), λ being the factor of area defined by Equation (7). The weighting coefficient is equal to unity for points that verify (11) but not (12), and equal to zero otherwise. In generating the primary effective surface, weighting coefficients are respectively equal to 1, $1 - \lambda$ and 0 in the cases described.

To define the Gaussian process, values of σ have been chosen that correspond to a series of random processes with standardized preferred Ra values (ISO series of roughness grade number N), σ_u values equal to 0.1 and 0.2 have been used as representative of real surfaces of engineering interest such as those obtained by grinding. From the autocorrelation functions of figure 4 it can be seen that, as expected, surfaces with different roughness grades but with the same topography have similar σ_u values. For surfaces with roughness grade number below $N7$, the value of σ_u obtained by means of stylus exploration decreases as

the primary effective surface increasingly differs from the real surface, and consequently the secondary effective surface topography becomes increasingly shallower than that of the real surface. Convenient σ_{uu} values can be obtained from the autocorrelation function

$$\sigma_{uu}^2 = \left(\frac{d^4 R(u)}{du^4} \right)_{u=0}$$

Most autocorrelation functions of surface roughness can be approximated near the origin by means of

$$R(u) \cong \sigma^2 e^{-1/2 \frac{\sigma_u^2}{\sigma^2} u^2} \quad (17)$$

which for σ_{uu} gives the expression

$$\sigma_{uu} = \frac{\sigma_u^2}{\sigma} \sqrt{3} \quad (18)$$

Figure 5 shows distribution functions of pairs of mating primary and secondary effective surfaces obtained by means of Gaussian processes with $\sigma = 0.125 \mu\text{m}$ corresponding to a roughness grade $N3$, and $\sigma_u = 0.1$ and 0.2 .

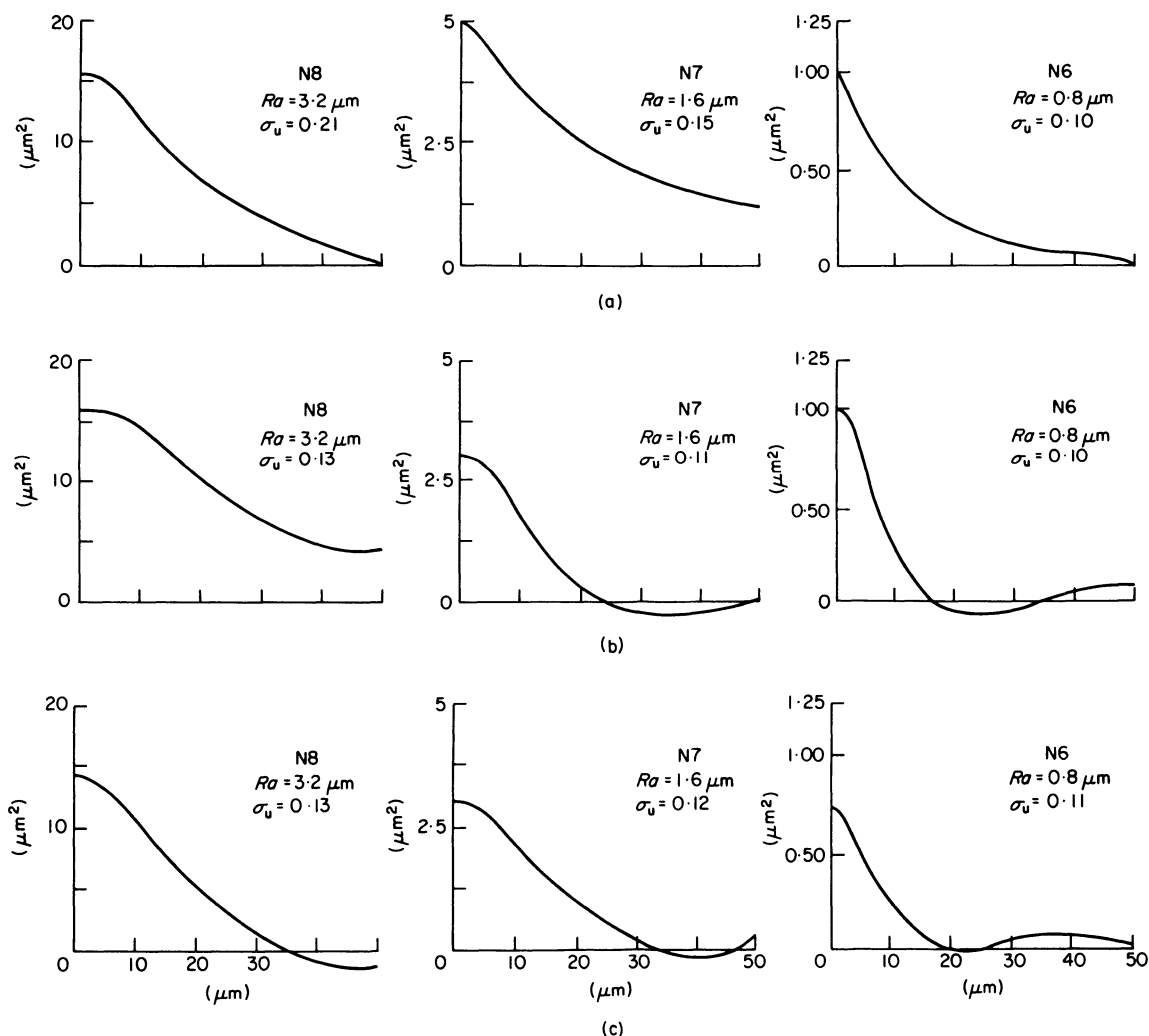


Figure 4. Autocorrelation functions for roughness comparison specimens representing (a) ground flat surfaces, (b) electrochemically ground surfaces, and (c) grit blasted surfaces.

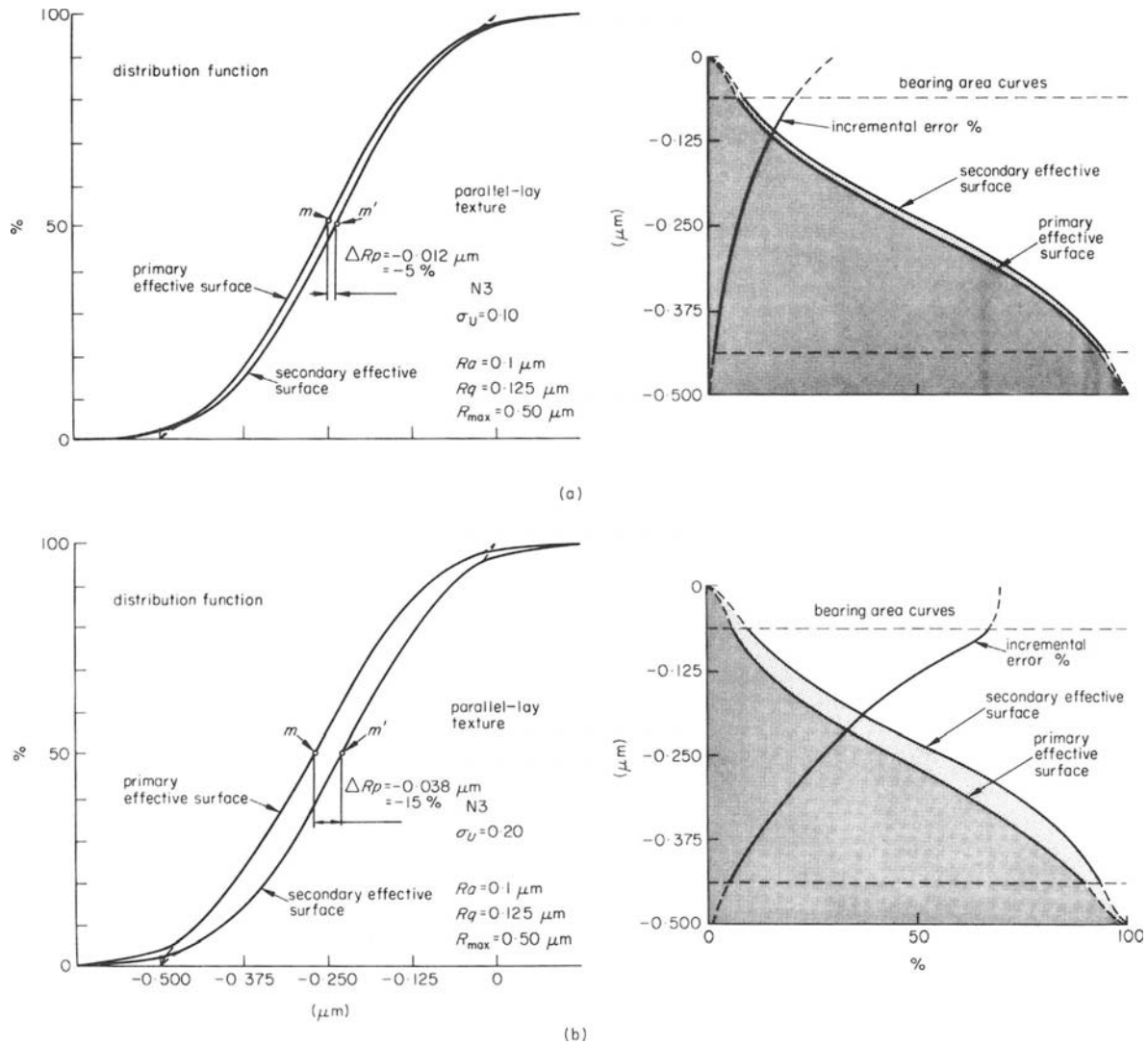


Figure 5. Variation of distribution functions and statistical parameters of primary and secondary effective surfaces for a rough surface with parallel-lay texture.

RANDOM FIELD APPROACH FOR ISOTROPIC TEXTURES

The bidimensional random field $f(u, v)$ used for rough surface characterization will be in this case homogeneous and isotropic¹²⁻¹⁴.

As in the case studied previously, the Gaussian assumption allows a complete analytical study. The required joint probability distribution is defined in the six-dimensional space of $f, f_u, f_v, f_{uu}, f_{vv}, f_{uv}$, but as before it depends, assuming zero mean for f , on a reduced number of parameters, $\sigma, \sigma_u = \sigma_v, \sigma_{uu} = \sigma_{vv} = \sqrt{3} \sigma_{uv}$ as shown in Appendix 3.

Distribution function and parameters Ra and Rq for the transformed surface f' are given by the expressions

$$P_{f'}(\mu) = \frac{\sqrt{2\pi}}{\sigma_u^2} \int_0^\infty wg \left(\frac{w}{\sigma_u} \right) \left\{ G \left[\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma\sqrt{1+w^2}} \right] - \frac{\rho}{\sigma} \frac{\sigma_u^2(2+w^2)}{(1+w^2)^{3/2}} g \left[\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma\sqrt{1+w^2}} \right] \right\} dw$$

$$\frac{\rho^2 \sigma_u^4 (\mu/\sigma + \rho/\sigma - \rho/\sigma\sqrt{1+w^2})}{\sigma^2 (1+w^2)^2} g \left[\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma\sqrt{1+w^2}} \right] \} dw \quad (19)$$

$$R'a = m' + \frac{2\rho}{\sigma_u^2} \sqrt{2\pi} \int_0^\infty wg \left(\frac{w}{\sigma_u} \right) \left\{ \frac{\sigma}{\rho} g \left[\frac{\rho}{\sigma} + \frac{m'}{\sigma} - \frac{\rho}{\sigma\sqrt{1+w^2}} \right] + G \left[\frac{\rho}{\sigma} + \frac{m'}{\sigma} - \frac{\rho}{\sigma\sqrt{1+w^2}} \right] \times \left[1 + \frac{m'}{\rho} - \frac{1}{\sqrt{1+w^2}} \left(1 + \frac{\sigma_u^2}{1+w^2} \right) \right] \right\} dw \quad (20)$$

$$R'q^2 = \sigma^2 - (\rho + m')^2 + \frac{\rho^2}{\sigma_u^2} \sqrt{2\pi} \int_0^\infty \frac{w}{1+w^2} \times \left(1 + 2 \frac{2+w^2+\sigma_u^2}{1+w^2} \sigma_u^2 \right) g \left(\frac{w}{\sigma_u} \right) dw \quad (21)$$

m' being the mean value of f' , given by

$$m' = \frac{\rho\sqrt{(2\pi)}}{\sigma_u^2} \int_0^\infty \frac{w}{\sqrt{(1+w^2)}} \left(1 + \frac{2+w^2}{1+w^2} \sigma_u^2 \right) \times g\left(\frac{w}{\sigma_u}\right) dw - \rho \quad (22)$$

All statistical characteristics depend only on σ , σ_u and ρ , and the incremental value of Rq does not depend on σ , as in the parallel-lay case.

Determination of statistical characterization of the secondary effective surface from the statistical characteristics of the real surface is harder in this case, as it involves finding the probability of a random field that remains under a given surface. Direct simulation of this problem involves random field generation techniques¹⁵.

As in the previous case, a Monte Carlo technique has been used to generate points of mating effective surfaces by the geometric transformation approach. In this case conditions (9) and (10) substitute for (11) and (12) in the isotropic random field involved.

Isotropic rough surfaces of engineering interest, such as those obtained by electrochemical grinding, blasting and power metallurgy, present typical σ_u values of the order of 0.1-0.2, as can be seen from the autocorrelation curves shown in figure 4.

$\sigma_{uu} = \sigma_{vv}$ and σ_{uv} can be obtained from the autocorrelation function as

$$\sigma_{uu}^2 = \left(\frac{\partial^4 R(u, v)}{\partial u^4} \right)_{u=0, v=0}$$

$$\sigma_{uv}^2 = \left(\frac{\partial^4 R(u, v)}{\partial u^2 \partial v^2} \right)_{u=0, v=0}$$

Most autocorrelation functions of isotropic surface roughness can be approximated near the origin by the expression

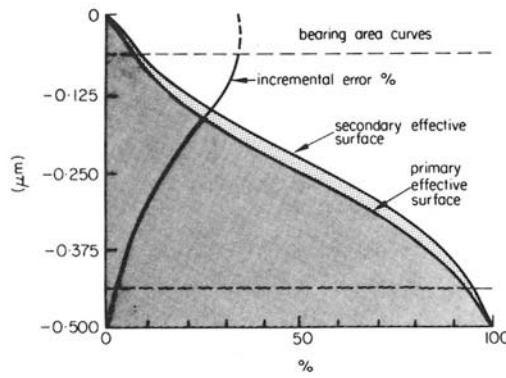
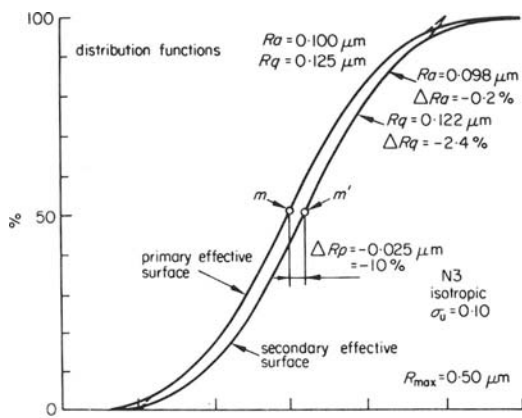
$$R(u, v) \cong \sigma^2 \exp \left[-\frac{1}{2} \frac{\sigma_u^2}{\sigma^2} (u^2 + v^2) \right] \quad (23)$$

which gives

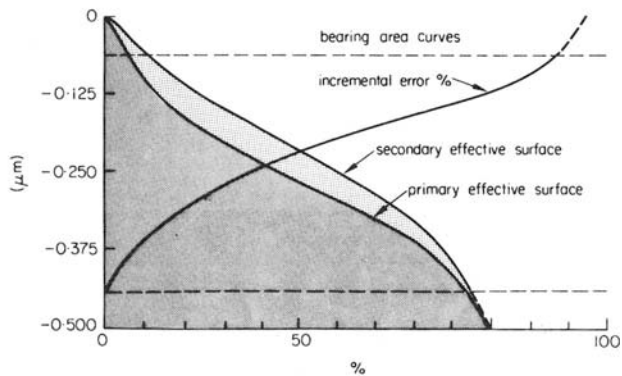
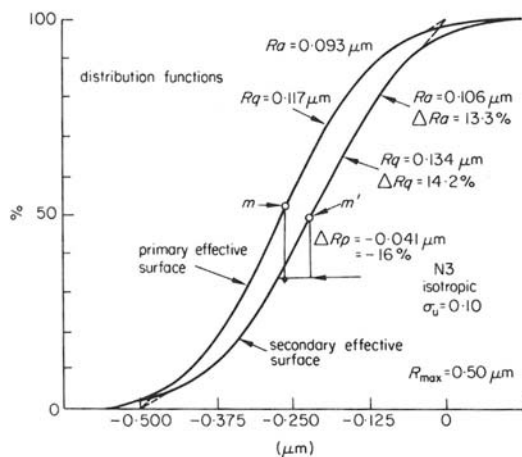
$$\sigma_{uu} = \sqrt{3} \frac{\sigma_u^2}{\sigma} \quad (24)$$

$$\sigma_{uv} = \frac{\sigma_u^2}{\sigma} \quad (25)$$

Distribution functions of pairs of mating primary and secondary effective surfaces shown in figure 6 have been obtained by means of random fields with $\sigma = 0.125 \mu\text{m}$ (roughness grade N3) and $\sigma_u = 0.1$ (figure 6(a)) and $\sigma_u = 0.2$ (figure 6(b)).



(a)



(b)

Figure 6. Variation of distribution functions and statistical parameters of primary and secondary effective surfaces for an isotropic rough surface.

LATERAL DEFLECTION OF THE STYLUS—AN INTRODUCTORY ANALYSIS

In the exploration of isotropic rough surfaces the stylus tends to deflect laterally as it contacts the surface through a point on one of its sides, due to its lateral flexibility. The observed profile does not correspond to the intersection of the secondary effective surface with x_1x_3 but to the projection on the x_1x_3 plane (figure 7) of a non-plane intersection of it.

Assuming that the lateral restoring force is proportional to lateral deflection, and that the normal force F from the stylus is constant and the friction force is negligible, the equilibrium equation gives at each x_1 value

$$x_2 = -\frac{F}{k} \left(\frac{\partial f'(x_1, x_2)}{\partial x_2} \right)_{x_1, x_2 = \delta_H} \quad (26)$$

where k is the stiffness coefficient.

In order to avoid unstable positions of the stylus, F/k must verify (Appendix 4)

$$\frac{F}{k} \leq \rho \sin^3 \frac{\alpha}{2} \quad (27)$$

where α is the shank angle. For $\rho = 2.5 \mu\text{m}$ and $\alpha = 90^\circ$,

$$F/k \leq 0.884 \mu\text{m} \quad (28)$$

The minimum lateral stiffness coefficient k required for $F = 0.07 \text{ g}$ (as set by BS 1134 and ISO 468) or for $F = 0.1 \text{ g}$ (as set by ANSI B46.1) is

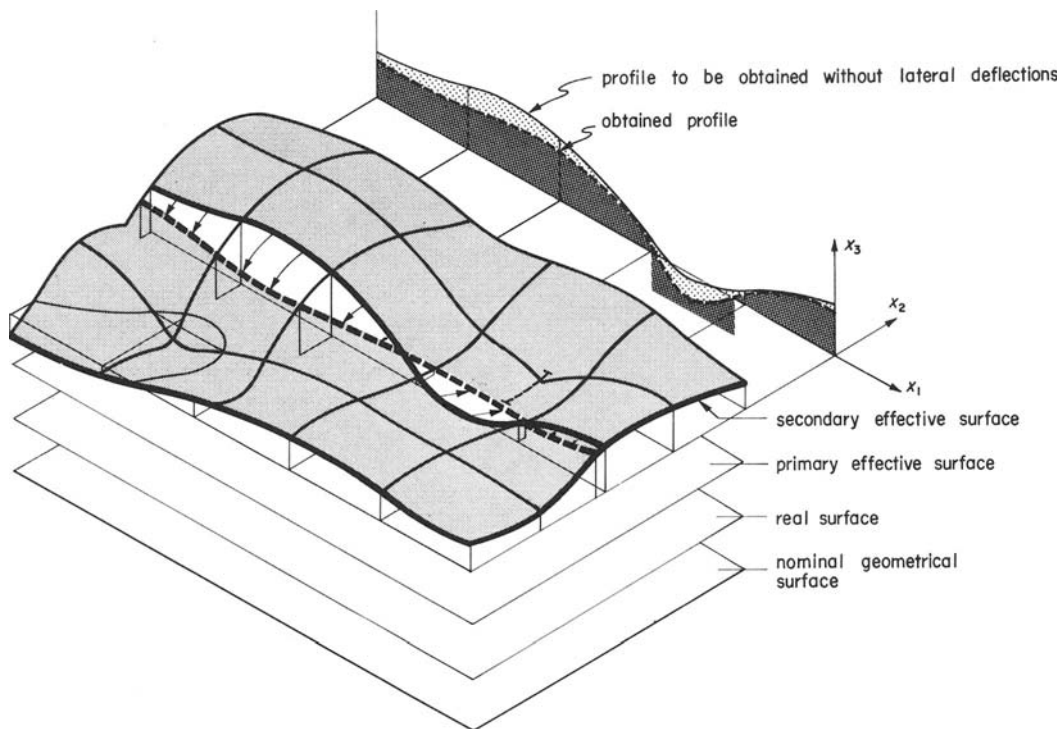
$$k_{\min} = 0.079 \text{ g}/\mu\text{m} \text{ for } F = 0.07 \text{ g}$$

$$k_{\min} = 0.113 \text{ g}/\mu\text{m} \text{ for } F = 0.10 \text{ g}$$

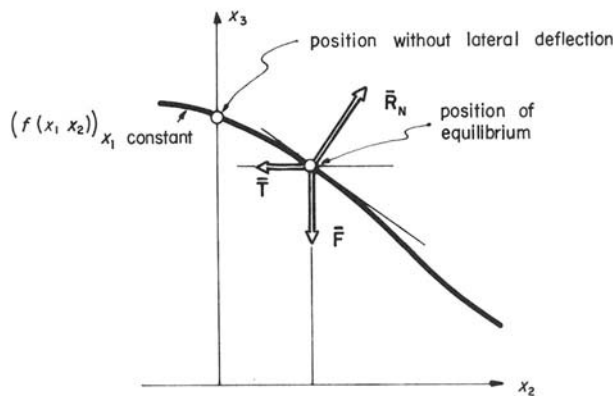
or considering maximum flexibility,

$$k_{\max}^{-1} = 12.6 \mu\text{m}/\text{g} \text{ for } F = 0.07 \text{ g}$$

$$k_{\max}^{-1} = 8.84 \mu\text{m}/\text{g} \text{ for } F = 0.10 \text{ g}$$



(a)



restoring force is assumed to be $|T| = k x_2$
friction force is neglected

(b)

Figure 7. Lateral deflections of the stylus.

If stability condition (27) is verified, lateral deflections imply negative vertical displacements (going down); thus the obtained profile will be lower than the profile that corresponds to a stiffer stylus. Contact points between both profiles—where lateral as well as vertical deflections are equal to zero—are tangential points. For $\alpha = 90^\circ$ the error δ_v introduced in the profile height at each point is related to lateral displacement δ_H by the equation

$$\delta_{v \min} \leq \delta_v \leq \delta_{v \max} = \delta_H$$

with

$$\delta_{v \min} = r \left\{ \cos \arctan \frac{\delta_H}{F/k} - \cos \left[\arcsin \left(\sin \arctan \frac{\delta_H}{F/k} - \frac{\delta_H}{r} \right) \right] \right\}$$

where $r = \rho \cos(\alpha/2) / \cos \theta_e$. Figure 8 shows the interval of δ_v for several values of F/k .

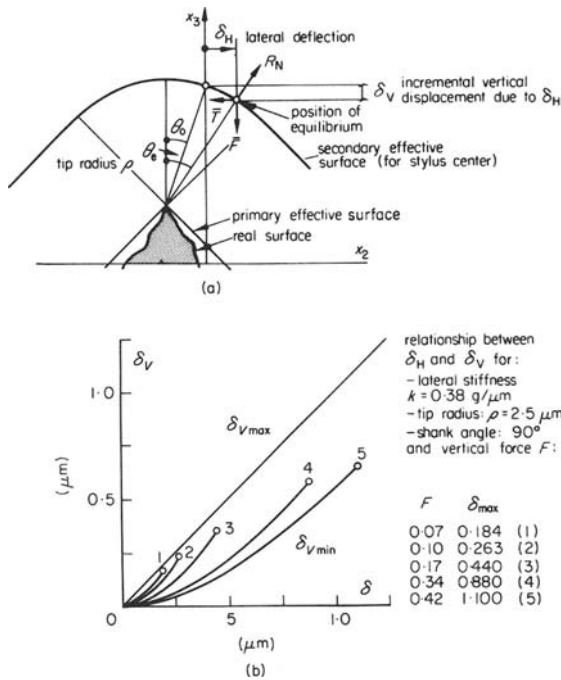


Figure 8. Stability of lateral deflections and vertical associated displacements.

Oscillograms of lateral deflections are shown in figure 9. A stylus with tip radius $\rho = 2.5 \mu\text{m}$, $\alpha = 90^\circ$ and $F = 0.4 \text{ g}$ was used. A displacement transducer of differential transformer type was used to measure lateral deflection, which increases the stiffness and applied normal force values to $k_T = 0.44 \text{ g}/\mu\text{m}$ and $F = 4.8 \text{ g}$. In this case $F/k = 1.1$ has a value slightly higher than the maximum value required for stability. However this will not sensibly affect the experimental results, since for a conical-spherical stylus instability can only appear at points from a narrow band of the stylus near the conical-spherical junction. $40.5^\circ \leq \theta \leq 45^\circ$ and in the worst case the additional horizontal and vertical displacements due to the instability will be $0.16 \mu\text{m}$ and $0.15 \mu\text{m}$ respectively.

Oscillograms of figure 9(a) correspond to lateral deflection and obtained profile for the stylus travers-

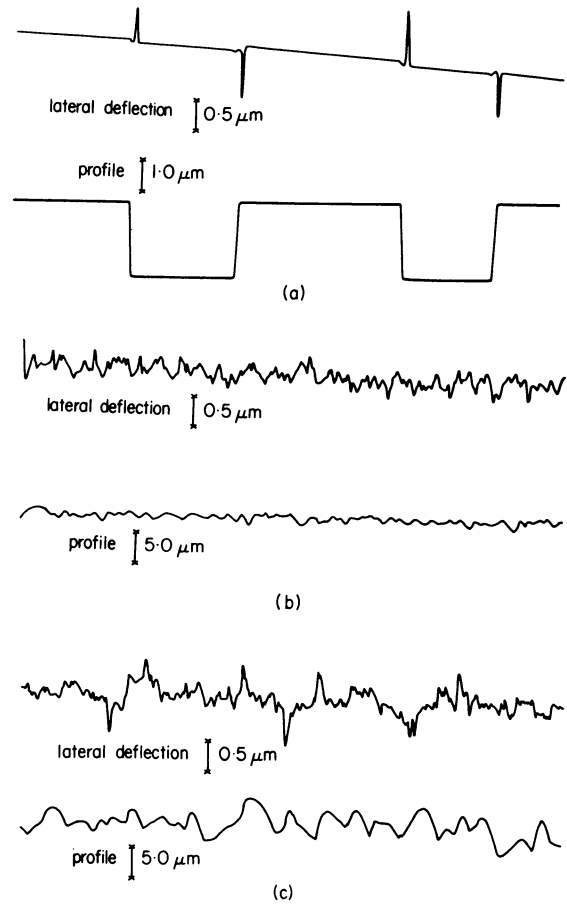


Figure 9. Oscillograms of lateral deflections for (a) height calibration specimen; (b) electrochemical grinding comparison specimen, roughness grade N5; (c) electrochemical grinding comparison specimen, roughness grade N7.

ing a height calibration specimen with three rectangular grooves of $2.4 \mu\text{m}$, in a direction that forms an angle of 13° with the grooves. The oscillogram clearly shows the lateral displacement with opposite sign due to the stylus 'sliding down'.

Oscillograms of figures 9(b) and 9(c) correspond to lateral displacements and obtained profiles, for two electrochemically-ground comparison specimens of grade N5 ($R_a = 0.4 \mu\text{m}$) and N7 ($R_a = 1.6 \mu\text{m}$) respectively. These oscillograms clearly show that lateral displacements are of similar magnitude, as can be expected from Equation (26), for real surfaces with similar roughness textures that nearly satisfy conditions (9) and (10), as is the case for the specimens used.

For a given texture, σ_u of the secondary effective surface decreases as the roughness quality of the real surface increases, and conditions (9) and (10) are not verified at an increasing fraction of points. Lateral deflections, as suggested by Equation (26), will decrease for the finer grades while maintaining its values $-F/k \leq \delta_H \leq F/k$ for the rough grades. The study of lateral deflection in the intermediate grades is in progress.

CONCLUSIONS

Profiles obtained using the stylus technique correspond to a surface—the secondary effective surface—of different statistical characteristics than both the real

surface and its closest approximation obtainable by means of a given stylus—the primary effective surface.

A geometrical transformation can be introduced which relates primary and secondary effective surfaces. It is a non-linear transformation, and in consequence operations such as spectrum analysis can be misleading.

Most rough surfaces of engineering interest are of parallel-lay texture or are isotropic. A random process provides a good means of representing the former case, while isotropic surfaces can be conveniently handled by means of a bidimensional random field.

Under a Gaussian assumption the distribution function, and consequently statistical parameters such as *Ra* and *Rq*, of secondary effective surfaces depend only on tip radius ρ and σ and σ_u of the Gaussian distribution assumed. Simulation of the geometric transformation provides a means of relating the statistical properties of primary and secondary effective surfaces.

To relate the statistical characteristics of the real and secondary effective surfaces is a very hard problem. However, a direct simulation of the surface detection by the stylus can be envisaged.

Lateral flexibility of the stylus is another source of error in the assessment of the roughness of isotropic surfaces. In this case the profiles obtained do not correspond to plane intersections of the secondary effective surface. As stability requires a minimum lateral stiffness, its value would be standardized.

These concepts, after a thorough study of the range of textures of engineering interest, may allow modification of the auxiliary electronic equipment associated with roughness measurement, in order to determine standardized parameters of the real or of the primary effective surface.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to the Fundación María Francisca Roviralta for an instrumentation grant, as well as to the Laboratorio de Automatica, ETSIIB for providing computing facilities, and to the Centro de Estudios e Investigación de la Máquina Herramienta for partial financial support.

APPENDIX 1 PROPERTIES OF THE GEOMETRICAL TRANSFORMATION DEFINED BY $\mathbf{x}' = \mathbf{x} + \rho\mathbf{N} - \rho\mathbf{v}$

1 Level variation in the direction of the versor \mathbf{v} , normal to the nominal geometrical surface

$$\epsilon \triangleq \mathbf{w} \cdot (\mathbf{x}' - \mathbf{x}) = -\rho(1 - \rho\mathbf{v}\mathbf{N})$$

2 Coefficients of the first fundamental quadratic form

$$E' \triangleq \mathbf{x}'_v \mathbf{x}'_v = E - 2\rho e + \rho^2(Ge^2 + Ef^2 - 2Fef)/(EG - F^2)$$

$$G' \triangleq \mathbf{x}'_u \mathbf{x}'_u = G - 2\rho g + \rho^2(Eg^2 + Gf^2 - 2Fgf)/(EG - F^2)$$

$$F' \triangleq \mathbf{x}'_u \mathbf{x}'_v = F - 2\rho f + \rho^2[Efg + Gef - F(eg + f^2)]/(EG - F^2)$$

where e , g and f are the coefficients of the second fundamental quadratic form of $\mathbf{x}(u, v)$.

3. Area factor

$$\lambda \triangleq \frac{dA'}{dA} = \frac{\sqrt{(E'G' - F'^2)}}{\sqrt{(EG - F^2)}} = 1 - 2\rho C_m + \rho^2 C_G^2$$

where

$$C_m = \frac{Eg + Ge - 2Ff}{2(EG - F^2)}$$

is the mean curvature, and

$$C_G^2 = \frac{eg - f^2}{EG - F^2}$$

is the Gaussian curvature.

4 Tangent planes to mating points are parallel since normal vectors are parallel

$$\begin{aligned} \mathbf{N}' &= \frac{\mathbf{x}'_u \wedge \mathbf{x}'_v}{|\mathbf{x}'_u \wedge \mathbf{x}'_v|} \\ &= \frac{(\mathbf{x}_u \wedge \mathbf{x}_v)(1 - 2\rho C_m + \rho^2 C_G^2)}{\sqrt{(E'G' - F'^2)}} \\ &= (\text{sign of } \lambda)\mathbf{N} \end{aligned}$$

5 Coefficients of the second fundamental quadratic form

$$e' \triangleq -\mathbf{x}'_u \mathbf{N}'_u = e - \rho(Ef^2 + Ge^2 - 2Fef)/(EG - F^2)$$

$$g' \triangleq -\mathbf{x}'_v \mathbf{N}'_v = g - \rho(Eg^2 + Gf^2 - 2Fgf)/(EG - F^2)$$

$$\begin{aligned} f' &\triangleq -\mathbf{x}'_u \mathbf{N}'_v \\ &= f - \rho[Egf + Gef - F(eg + f^2)]/(EG - F^2) \end{aligned}$$

6 Mean and gaussian curvatures of $\mathbf{x}'(u, v)$

$$C'_m \triangleq \frac{G'e' + E'g' - 2F'f'}{2(G'E' - F'^2)} = \frac{C_m + \rho C_G^2}{\lambda}$$

$$C_G'^2 \triangleq \frac{e'g' - f'^2}{E'G' - F'^2} = \frac{1}{\lambda} C_G^2$$

7 Principal curvatures

$$C'_1 = C'_m + \sqrt{(C_m'^2 - C_G'^2)} = \frac{C_1}{1 + \rho C_1}$$

$$C'_2 = C'_m - \sqrt{(C_m'^2 - C_G'^2)} = \frac{C_2}{1 + \rho C_2}$$

8 Principal directions of curvature are parallel since the coefficients of the characteristic equation for \mathbf{x}'

$$(F'g' - G'f')\mu^2 + (E'g' - G'e')\mu + (E'f' - F'e') = 0$$

are proportional to those of \mathbf{x} , by means of a factor equal to the area factor.

APPENDIX 2

1 Statistical characteristics of the joint probability distribution of f , f_u and f_{uu} which is assumed to be Gaussian

The joint probability distribution is wholly characterized by

$$E(f), E(f_u), E(f_{uu}), E(f^2), E(f_u^2), E(f_{uu}^2), E(ff_u), \\ E(ff_{uu}), E(f_u f_{uu})$$

If f is a stationary, twice-derivable Gaussian process with zero mean and variance σ^2 , $E(f) = 0$, $E(f^2) = \sigma^2 = R(0)$, the stationarity property implies that $E(f_u) = 0$, $E(f_{uu}) = 0$. Derivability implies

$$E(ff_u) = \left(\frac{dR}{du} \right)_{u=0} = 0,$$

$$E(f_u^2) = - \left(\frac{d^2R}{du^2} \right)_{u=0} = \sigma_u^2,$$

$$E(ff_{uu}) = \left(\frac{d^2R}{du^2} \right)_{u=0} = -\sigma_u^2,$$

$$E(f_u f_{uu}) = - \left(\frac{d^3R}{du^3} \right)_{u=0} = 0,$$

$$E(f_{uu} f_{uu}) = \left(\frac{d^4R}{du^4} \right)_{u=0} = \sigma_{uu}^2$$

All the parameters only depend on the shape of the autocorrelation function for f , near zero.

2 Distribution function for the process f' obtained by means of transformation defined by Equation (1) applied to f

$$P_{f'}(\mu) = \text{prob}(f' \leq \mu) \\ = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_D du' \\ = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T t(f, f_u, \mu) \left[1 - \rho \frac{f_{uu}}{(1+f_u^2)^{3/2}} \right] du \\ = \int_{\eta=-\infty}^{\infty} \int_{\xi=-\infty}^{\infty} \int_{\zeta=-\infty}^{\infty} t(\xi, \eta, \mu) \left[1 - \frac{\rho \xi}{(1+\eta^2)^{3/2}} \right] \\ \times p_{ffufuu}(\xi, \eta, \zeta) d\xi d\eta d\zeta \\ = \int_{\xi=-\infty}^{\infty} \int_{\zeta=-\infty}^{\infty} \int_{\eta=-\infty}^{\infty} \frac{\mu + \rho - \rho/\sqrt{1+\eta^2}}{\xi} \\ \times \left[1 - \frac{\rho \xi}{(1+\eta^2)^{3/2}} \right] p_{ffufuu}(\xi, \eta, \zeta) d\xi d\eta d\zeta$$

where D is the set of all points of the $(-T, T)$ interval for which $f' \leq \mu$ and

$$t(f, f_u, \mu) = 1$$

if

$$f - \rho + \frac{\rho}{\sqrt{1+f_u^2}} \leq \mu$$

and zero otherwise.

For a Gaussian process the formula simplifies to

$$p_{f'}(\mu) = \frac{1}{\sigma_u} \int_{-\infty}^{\infty} g\left(\frac{\eta}{\sigma_u}\right) \left\{ G\left[\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma\sqrt{1+\eta^2}}\right] \right. \\ \left. - \frac{\rho\sigma_u^2}{\sigma(1+\eta^2)^{3/2}} g\left[\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma\sqrt{1+\eta^2}}\right] \right\} d\eta$$

3 Probability function, mean value, Ra and Rq parameters

$$p_{f'}(\mu) = \frac{dp_{f'}}{d\mu}(\mu) = \frac{1}{\sigma\sigma_u} \int_{-\infty}^{\infty} g\left(\frac{\eta}{\sigma_u}\right) \\ \times g\left(\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma\sqrt{1+\eta^2}}\right) \\ \times \left[1 + \frac{\rho\sigma_u^2}{\sigma^2(1+\eta^2)^{3/2}} \left(\mu + \rho - \frac{\rho}{\sqrt{1+\eta^2}} \right) \right] d\eta$$

$$m' = \int_{-\infty}^{\infty} \mu p_{f'}(\mu) d\mu$$

$$= \frac{\rho}{\sigma_u} \int_{-\infty}^{\infty} g\left(\frac{\eta}{\sigma_u}\right) \frac{1}{\sqrt{1+\eta^2}} \left(1 + \frac{\sigma_u^2}{1+\eta^2} \right) d\eta - \rho$$

$$R'a = \int_{-\infty}^{\infty} |\mu - m'| p_{f'}(\mu) d\mu$$

$$= m' + \frac{2\rho}{\sigma_u} \int_{-\infty}^{\infty} g\left(\frac{\eta}{\sigma_u}\right) \left\{ \frac{\sigma}{\rho} g\left[\frac{\rho}{\sigma} + \frac{m'}{\sigma} - \frac{\rho}{\sigma\sqrt{1+\eta^2}}\right] \right.$$

$$\left. + G\left[\frac{\rho}{\sigma} + \frac{m'}{\sigma} - \frac{\rho}{\sigma\sqrt{1+\eta^2}}\right] \right.$$

$$\left. \times \left[1 + \frac{m'}{\rho} - \frac{1}{\sqrt{1+\eta^2}} \left(1 + \frac{\sigma_u^2}{1+\eta^2} \right) \right] \right\} d\eta$$

$$R'q^2 = \int_{-\infty}^{\infty} (\mu - m')^2 p_{f'}(\mu) d\mu$$

$$= \sigma^2 - (\rho + m')^2$$

$$+ \frac{\rho^2}{\sigma_u} \int_{-\infty}^{\infty} \frac{g(\eta/\sigma_u)}{1+\eta^2} \left(1 + \frac{2\sigma_u^2}{1+\eta^2} \right) d\eta$$

This parameter, as well as the mean value m' , can be directly found as

$$R'q^2 = E \left\{ \left[1 - \frac{\rho f_{uu}}{(1+f_u^2)^{3/2}} \right] \left[f - \rho + \frac{\rho}{\sqrt{1+f_u^2}} \right]^2 \right\} - m'^2 \\ = \sigma^2 - (m' + \rho)^2$$

$$+ \rho^2 E \left[\frac{1}{1+f_u^2} \left(1 + \frac{2\sigma_u^2}{1+f_u^2} \right) \right]$$

$$m' = E \left\{ \left[1 - \frac{\rho f_u}{(1+f_u^2)^{3/2}} \right] \left[f - \rho + \frac{\rho}{\sqrt{1+f_u^2}} \right] \right\}$$

$$= \rho E \left[\frac{1}{\sqrt{1+f_u^2}} \left(1 + \frac{\sigma_u^2}{1+f_u^2} \right) \right] - \rho$$

APPENDIX 3

1 Statistical characteristics of the joint probability distribution function of $f, f_u, f_v, f_{uu}, f_{vv}, f_{uv}$ assumed to be Gaussian

The joint probability distribution function is wholly characterized by six mean values, six variances and fifteen covariances. If f is a homogeneous, twice-derivable random field with zero mean and variance σ^2 , the mean values will all be zero by the homogeneity hypothesis. Derivability implies independence between the pairs

$$f, f_u; f, f_v; f, f_{uu}; f_u, f_v; f_u, f_{uu}; f_u, f_{vv}; f_u, f_{uv}; f_v, f_{uu}; f_v, f_{vv}; f_v, f_{uv}; f_{uu}, f_{uv}; f_{vv}, f_{uv}$$

and

$$E(f_u^2) = E(f_v^2) = \sigma_u^2 = -\left(\frac{\partial^2 R}{\partial u^2}\right)_{u=0}$$

$$E(f_{uu}^2) = E(f_{vv}^2) = \sigma_{uu}^2 = \left(\frac{\partial^4 R}{\partial u^4}\right)_{u=0}$$

$$E(ff_{uu}) = E(ff_{vv}) = -\sigma_u^2$$

$$E(f_{uu}f_{vv}) = \sigma_{uu}^2 = E(f_{uv}^2) = \sigma_{uv}^2 = \frac{1}{3}\sigma_{uu}^2$$

Thus all the parameters only depend on the shape of the autocorrelation function near the origin.

2 Distribution function for the process f' obtained by means of transformation defined by Equation (1) applied to f

$$P_{f'}(\mu) = \text{prob}(f' \leq \mu) = \lim_{T \rightarrow \infty} \frac{1}{4T^2} \int_D dA'$$

$$= \lim_{T \rightarrow \infty} \frac{1}{4T^2} \int_D t(f, f_u, f_v, \mu)$$

$$\times \lambda(f, f_u, f_v, f_{uu}, f_{vv}, f_{uv}) dA$$

$$= \int_{\nu=-\infty}^{\infty} \int_{\eta=-\infty}^{\infty} \int_{\xi=-\infty}^{\infty} \int_{\psi=-\infty}^{\infty} \int_{\varphi=-\infty}^{\infty} \int_{\zeta=-\infty}^{\infty}$$

$$\times t(\xi, \eta, \nu, \mu) \lambda(\xi, \eta, \nu, \zeta, \varphi, \psi)$$

$$\times P_{ff_{uu}f_{vv}f_{uv}}(\xi, \eta, \nu, \zeta, \varphi, \psi)$$

$$\times d\xi d\eta d\nu d\zeta d\varphi d\psi$$

$$= \int_{\nu=-\infty}^{\infty} \int_{\eta=-\infty}^{\infty} \int_{\psi=-\infty}^{\infty} \int_{\varphi=-\infty}^{\infty} \int_{\zeta=-\infty}^{\infty} \int_{\xi=-\infty}^{\frac{\mu + \rho - \rho / \sqrt{(1 + \eta^2 + \nu^2)}}{1 + \eta^2 + \nu^2}}$$

$$\times \lambda(\xi, \eta, \nu, \zeta, \varphi, \psi)$$

$$\times P_{ff_{uu}f_{vv}f_{uv}}(\xi, \eta, \nu, \zeta, \varphi, \psi)$$

$$\times d\xi d\eta d\nu d\zeta d\varphi d\psi$$

with D the set of all points of a square $2T \times 2T$ for which

$$f' \leq \mu$$

$$t(f, f_u, f_v, \mu) = 1 \text{ for } f - \rho + \rho / (\sqrt{(1 + f_u^2 + f_v^2)}) \leq \mu$$

$$= 0 \text{ otherwise}$$

$$\lambda(f, f_u, f_v, f_{uu}, f_{vv}, f_{uv}) = 1 - 2\rho C_m + \rho^2 C_G^2$$

the area factor defined in Equation (5).

In the Gaussian case considered, with $w^2 = \eta^2 + \nu^2$ the formula simplifies to

$$p_{f'}(\mu) = \frac{\sqrt{(2\pi)}}{\sigma_u^2} \int_0^{\infty} wg \left(\frac{w}{\sigma_u}\right) \left[G \left(\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma \sqrt{(1 + w^2)}} \right) \right.$$

$$- \frac{\rho \sigma_u^2 (2 + w^2)}{\sigma (1 + w^2)^{3/2}} g \left(\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma \sqrt{(1 + w^2)}} \right)$$

$$- \frac{\rho^2 \sigma_u^4 ((\mu/\sigma) + (\rho/\sigma) - \rho/\sigma \sqrt{(1 + w^2)})}{\sigma^2 (1 + w^2)^2}$$

$$\left. \times g \left(\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma \sqrt{(1 + w^2)}} \right) \right] dw$$

3 Probability function, mean value, Ra and Rq parameters

$$p_{f'}(\mu) = \frac{dp_{f'}}{d\mu}(\mu)$$

$$= \frac{\sqrt{(2\pi)}}{\sigma \sigma_u^2} \int_0^{\infty} wg \left(\frac{w}{\sigma_u}\right) g \left(\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma \sqrt{(1 + w^2)}} \right)$$

$$\times \left[1 - \frac{\rho^2 \sigma_u^4}{\sigma^2 (1 + w^2)^2} \right.$$

$$+ \frac{\rho \sigma_u^2 (2 + w^2)}{\sigma (1 + w^2)^{3/2}} \left(\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma \sqrt{(1 + w^2)}} \right)$$

$$\left. + \frac{\rho^2 \sigma_u^4}{\sigma^2 (1 + w^2)^2} \left(\frac{\mu}{\sigma} + \frac{\rho}{\sigma} - \frac{\rho}{\sigma \sqrt{(1 + w^2)}} \right)^2 \right] dw$$

$$m' = \int_{-\infty}^{\infty} \mu p_{f'}(\mu) d\mu$$

$$= \frac{\rho \sqrt{(2\pi)}}{\sigma_u^2} \int_0^{\infty} \frac{w}{\sqrt{(1 + w^2)}} \left(1 + \frac{2 + w^2}{1 + w^2} \sigma_u^2 \right)$$

$$\times g \left(\frac{w}{\sigma_u} \right) dw - \rho$$

$$R'a = \int_{-\infty}^{\infty} |\mu - m'| p_{f'}(\mu) d\mu$$

$$= m' + \frac{2\rho}{\sigma_u^2} \sqrt{(2\pi)} \int_0^{\infty} wg \left(\frac{w}{\sigma_u}\right)$$

$$\times \left\{ \frac{\sigma}{\rho} g \left[\frac{\rho}{\sigma} + \frac{m'}{\sigma} - \frac{\rho}{\sigma \sqrt{(1 + w^2)}} \right] \right.$$

$$+ G \left[\frac{\rho}{\sigma} + \frac{m'}{\sigma} - \frac{\rho}{\sigma \sqrt{(1 + w^2)}} \right]$$

$$\left. \times \left[1 + \frac{m'}{\rho} - \frac{1}{\sqrt{(1 + w^2)}} \left(1 + \frac{\sigma_u^2}{1 + w^2} \right) \right] \right\} dw$$

$$R'q^2 = \int_{-\infty}^{\infty} (\mu - m')^2 p_{f'}(\mu) d\mu$$

$$= \sigma^2 - (\rho + m')^2 + \frac{\rho^2}{\sigma^2} \sqrt{(2\pi)} \int_0^{\infty} \frac{w}{1 + w^2}$$

$$\times \left(1 + 2 \frac{2 + w^2 + \sigma_u^2}{1 + w^2} \sigma_u^2 \right) g \left(\frac{w}{\sigma_u} \right) dw$$

$R'q$, as well as the mean value m' , can be directly found as

$$\begin{aligned} R'q^2 &= E \left[(1 - 2\rho C_m + \rho^2 C_G^2) \right. \\ &\quad \left. \times \left(f - \rho + \frac{\rho}{\sqrt{(1+f_u^2+f_v^2)}} \right)^2 \right] - m'^2 \\ &= \sigma^2 - (\rho + m')^2 + \rho^2 \\ &\quad \times E \left[\frac{1}{1+f_u^2+f_v^2} \left(1 + 2 \frac{2+f_u^2+f_v^2+\sigma_u^2}{1+f_u^2+f_v^2} \sigma_u^2 \right) \right] \\ m' &= E \left[(1 - 2\rho C_m + \rho^2 C_G^2) \right. \\ &\quad \left. \times \left(f - \rho + \frac{\rho}{\sqrt{(1+f_u^2+f_v^2)}} \right) \right] \\ &= \rho E \left[\frac{1}{\sqrt{(1+f_u^2+f_v^2)}} \left(1 + \frac{2+f_u^2+f_v^2}{1+f_u^2+f_v^2} \sigma_u^2 \right) \right] - \rho \end{aligned}$$

APPENDIX 4 LATERAL DEFLECTIONS OF THE STYLUS

1 Equilibrium condition

Assuming $T = -k\delta_H$, $F = \text{constant}$, and friction negligible (figure 7), the equilibrium condition is

$$\begin{aligned} F \sin \theta - k\delta_H \cos \theta = 0 &\Rightarrow \delta_H = \frac{F}{k} \tan \theta_e \\ &= \frac{F}{k} \left[\frac{\partial f'(x_1, x_2)}{\partial x_2} \right]_{x_1, x_2 = \delta_H} \end{aligned}$$

2 Stability condition

Maximum negative curvatures found in the secondary effective surface correspond to the stylus traversing an edge (figure 8). In this case the potential energy associated to F and to the restoring elastic force T is defined by the expression

$$U = F\rho \cos \theta + \frac{1}{2}k\rho^2(\sin \theta - \sin \theta_0)^2$$

$d^2U/d\theta^2$ at the equilibrium positions $-dU/d\theta = 0$ — is given by

$$\left(\frac{d^2U}{d\theta^2} \right)_{\partial U/\partial \theta = 0} = -F\rho \frac{1}{\cos \theta_e} + k\rho^2 \cos^2 \theta_e$$

Stability requires $d^2U/d\theta^2 \geq 0$, and consequently

$$F/k \leq \rho \cos^3 \theta_e$$

3 Relationship between δ_H and δ_v

For a given δ_H , maximum δ_v corresponds to the stylus contacting a plane of maximum slope. If $\alpha = 90^\circ$, $\delta_{\max} = \delta_H$. Minimum δ_v corresponds to the stylus traversing an edge. In this case

$$\delta_H = r(\sin \theta - \sin \theta_0) = \frac{F}{k} \tan \theta$$

$$\delta_{v\min} = r(\cos \theta - \cos \theta_0)$$

and eliminating θ and θ_0

$$\begin{aligned} \delta_{v\min} &= r \left\{ \cos \arctan \frac{\delta_H}{F/k} \right. \\ &\quad \left. - \cos \left[\arcsin \left(\sin \arctan \frac{\delta_H}{F/k} - \frac{\delta_H}{r} \right) \right] \right\} \end{aligned}$$

where $r = \rho \cos(\alpha/2)/\cos \theta_e$.

REFERENCES

1. ISO R 468. *Surface roughness*. 1966.
2. ANSI B46.1 *Surface texture*. 1971.
3. BS 1134. *Method for the assessment of surface texture*. 1972.
4. NF E 05. *Etat des surfaces des produits*. 1972.
5. MIT Report no. 15. Friction and surface finish. *Proc. Conf. held at MIT in 1940*, MIT Press, 1969.
6. E. Picciolato. *Tolerancias de Fabricación*. Instituto Nacional de Racionalización (Spain), 1964.
7. J. Agullo and S. Cardona. *Valoración de la Rugosidad Superficial*. ETSIB Centro de Estudios e Investigación de Máquinas Herramienta, 1973.
8. Jiro Nara. Two-dimensional representation of surface roughness. *CIRP Ann.* (1969), 485-93.
9. A. Ber and S. Braun. Spectral analysis of surface finish. *CIRP Ann.* (1968), 53-9.
10. M. Kubo and J. Peklenik. An analysis of micro-geometrical isotropy for random surface structures. *CIRP Ann.* (1968), 235-42.
11. J. Peklenik and M. Kubo. A basic study of a three-dimensional assessment of the surface generated in a manufacturing process. *CIRP Ann.* (1968) 257-65.
12. P. Ranganath Nayak. Some aspects of surface roughness measurement. *Wear*, 26 (1973), 165-74.
13. A. Papoulis. *Probability, random variables and stochastic processes*. McGraw-Hill, 1965.
14. E. Wong. Homogeneous Gauss-Markov random fields. *Ann. Math. Stat.* (1969) 40, 1625-34.
15. B. I. Shkurskiy. Mathematical simulation of two-dimensional random fields. *Engineering Cybernetics* (1969) 6.

EDM AND ECM

CONDITIONING A DIAMOND-GRIT CUP-WHEEL FOR ELECTROCHEMICAL GRINDING

by

M. M. SFANTSIKOPOULOS* and C. F. NOBLE*

SUMMARY

The paper discusses truing, dressing and cleaning of a diamond-grit cup-wheel by mechanical and electrochemical means. Particular emphasis is given to subsequent use for electrochemical grinding. Electrochemical truing is shown to offer no advantage over conventional methods except for fully retracting the metallic bond from the surface of the diamond grits. Similarly electrochemical dressing is shown to remove conductive products from the inter-grit spaces. Particularly advantageous wheel-wear characteristics are shown to exist during subsequent E.C.G. use. Invaluable features are further exhibited by in-process electrochemical cleaning, provided it is carried out prior to passive conditions arising.

INTRODUCTION

Available grinding wheels for Electrochemical Grinding (E.C.G.) fall into three main categories: Diamond metal-bonded wheels; formable conventional-grit wheels with a copper powder-resin, or graphite, bond; and, finally, non-grit graphite wheels. Of these the application of wheels belonging to the third group is rather limited, at least at present, whereas aluminium oxide grinding wheels are used in electrochemical form grinding. Almost exclusively diamond wheels are employed either for face (plunge) grinding operations—the most frequently met industrial E.C.G. application today—or for surface grinding (peripheral or of the vertical spindle type). It is the object of the present paper to discuss, analyse, and experimentally assess some aspects of the ‘conditioning’ methods available for these wheels with respect to the specific requirements posed by the features of the E.C.G. process. The term ‘wheel conditioning’ is used hereafter to imply all three operations, truing, dressing and, especially E.C.G., cleaning. By ‘wheel cleaning’ is meant the removal of any passive films and/or mechanical and/or electrochemical debris from its active surface. It is a particularly important operation for process prestige and efficiency.

Conventional methods

The main field of application for diamond wheels remained rather limited until they were introduced for grinding steels and advanced ‘space-age’ materials. Unfortunately there is very limited detailed information on the conditioning procedures for these wheels,

and only very limited research work appears to have been conducted.

Because single or multi-point diamond truing—as applied to the usual abrasive wheels—would result in rather severe damage to both the truing tool and the wheel, the pattern followed is that of abrading the wheel bond to such a degree that the diamond grits fall away. The mechanics of the method is based on the diamond grits cutting through the bond and grit of the ‘truing wheel’ or the ‘truing block’ so producing abrasive debris at the interface of the two units which, in turn, abrades the bond of the diamond wheel. The truing wheel may be driven by contact with the diamond wheel or it may be driven independently.

Selby¹ classifies truing methods as: (a) a mounted silicon carbide block; (b) a brake-controlled silicon carbide truing device (i.e. truing wheel); (c) a metal bond, diamond-grit, truing block. The latter method has only recently been developed by De Beers and the trued wheel possesses a highly polished surface. A silicon carbide block subsequently serves to ‘open’ the structure. The final ‘wheel dressing’ is left to a ‘self-conditioning’ procedure. Davies² conducted a series of conditioning tests with resinoid bond diamond wheels for surface grinding of steel. He made use of the brake-controlled method (b) for wheel truing and tested four dressing methods with the aim of producing an ‘open’ wheel with a low initial wear: (i) a braked truing wheel as for truing; (ii) an aluminium oxide dressing stick held by hand or by the machine; (iii) carbide conditioned, i.e. grinding a carbide block and relying, therefore, on a self-dressing process; (iv) a rotary wire brush. It may

* University of Manchester Institute of Science and Technology

be concluded that the truing wheel or the carbide block could be applied with satisfactory results and that, by suitably choosing the truing-wheel speed, the wheel surface structure could be controlled. A further method, referred to by Kobayashi^{2,3}, is the use of abrasive particles suspended in a liquid between dresser and wheel together with an 'in-feed' towards the wheel.

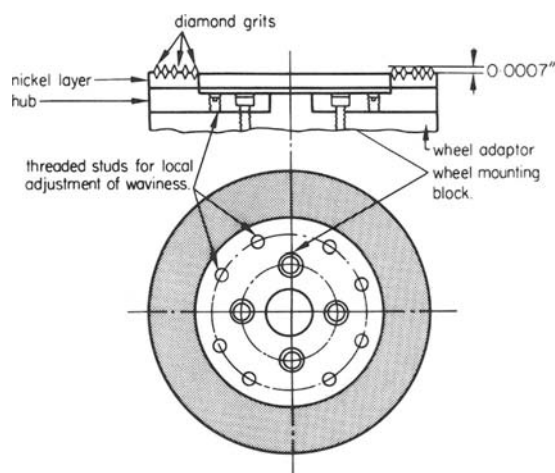


Figure 1. Metzger & Keeleric E.C.G. wheel (reference 3).

The importance of conditioning an E.C.G. wheel made itself apparent during the first applications of the process. Metzger and Keeleric³ described in 1952 the manufacture of an electroplated diamond wheel with a 'preset grit projection' of 0.0007 in and reported a patented 'truing method' for achieving the desirable E.C.G.-gap constancy (figure 1). Fishlock⁴ also reports attempts to produce E.C.G. wheels with an accuracy of 0.0001 in by crush-forming which proved, nevertheless, impractical. It may be said that, in general, the preparation and maintenance of E.C.G. wheels has followed more or less the same patterns as those of conventional grinding wheels with the exception of the 'electrochemical methods' which are discussed hereafter.

Electrochemical methods

Electrochemical wheel conditioning, first referred to (1953) by Keeleric and Williams⁵, has been faced with much controversy. It involves an inversed polarity (wheel: the positive electrode; dresser: the negative electrode) and electrochemical machining of the conductive wheel bond. Some mechanical action is also encountered. The main aim is to establish the 'grit exposure from the bond' and as a convenient method for avoiding alternative mechanical operations.

Fishlock⁴ mentions (1960) a special diamond dressing-shoe operating electrochemically with a claimed accuracy of 0.0001 in, while Unterbusch⁶ reported (1961) an electrochemical hand-dresser for cleaning of passive films from an E.C.G. wheel face. On a different basis, Brandi⁷ recommended 'wheel deplating' for unused wheels in order to reveal the diamond grits at their 'correct' height from the bond, but cleaning had to be effected afterwards by means of conventional abrasive sticks. Nevertheless he points out that a deplated wheel-matrix disposes an irregular

and porous surface having a greater capacity to carry electrolyte solution.

Another approach to the problem is quoted by Fishlock⁸ (1961): mechanical dressing with silicon carbide and then electrochemical dressing with hardened high-speed steel. The method was claimed to hold a high and constant current and to reduce the frequency of re-dressing. A similar procedure is also suggested by the Thompson Grinder Company⁹. According to Balashov¹⁰ Russian built E.C.G. machine tools are equipped with a polarity switch for rapid wheel dressing. A more recent Russian work (Pronin, et al.)¹¹ is rather indecisive about the effectivity of electrochemical dressing, but the present authors interpret these results as favouring electrochemical wheel treatment. Different revisions of the method are today in use though in every case different reasoning is quoted to justify their employment (e.g. reference 12).

Chalkley¹³ has rejected that there is any advantage to be had from electrochemical wheel conditioning, commenting that it could have the opposite effect by destroying the geometrical equilibrium existing between bond and diamond grits. Instead he suggests the use of aluminium oxide abrasive sticks for cleaning the working face. This type of wheel cleaning, however, has not proved satisfactory—at least for grinding tungsten carbide—by the present authors¹⁴, where the non-conductive wheel layer, once formed, remained practically untouched. Nikitin and Stepanov¹⁵ also stress the inadequacy of dressing-sticks or dressing-wheels for exposing the diamond grits from the wheel bond, even by repeated action. Cutting power measurements (conventional grinding) indicated that considerable power was absorbed in overcoming friction between bond and workpiece. Under E.C.G. conditions this would result in short-circuiting and arcing. Their observations are in broad agreement with the present authors' problems experienced during earlier E.C.G. studies¹⁴. Nikitin and Stepanov suggest, therefore, a grit protrusion height of 30-40 per cent of the grit mean diameter. To reveal the diamond grits they etched the wheel's diamond-bearing layer with (undiluted) nitric acid and found that for both grit-exposure methods (i.e. conventional and etching) the specific consumption of diamond varied within the same limits.

A simultaneous electrochemical grinding and electrochemical dressing arrangement was tried by Büttner and Lindenbeck¹⁶ using independent 'grinding' and 'dressing' d.c. power supplies. A similar procedure was applied by Keeleric and Williams⁵ but with successive stages of electrochemical grinding and dressing using the same d.c. power unit. They were nevertheless particularly interested in testing metal-bonded diamond wheels used for conventional grinding of steel. The results obtained were considered encouraging for further development work.

ASSESSMENT OF CONDITIONING METHODS FOR E.C.G. WHEELS

The reviewed 'conventional' and 'electrochemical' wheel conditioning methods in conjunction with the

geometric and dynamic considerations of E.C.G. (i.e. rotational trueness of the wheel-bond surface, grit-protrusion height, average electrochemical gap, table traverse feed rate, depth of cut, forced vibrations, geometric accuracy of the machine tool (e.g. reference 17) allow the following conclusions to be drawn.

Conventional truing and dressing techniques for metal-bonded diamond wheels, though they may prove satisfactory in some respects for conventional mechanical grinding cannot guarantee the geometry of the bond-grit configuration required for E.C.G. This is because of the principle of their mechanics. In conventional grinding part of the contact between wheel and workpiece always represents a direct contact between the wheel bond and the surface being ground. Usually the 'stable secondary stage' of wheel wear is reached, as shown later in the paper, after more or less permanent damage to the wheel is caused by arcing and by considerable loss of wheel material. The problem of wheel cleaning in particular has not found any effective and economic mechanical solution. It has been pointed out in a previous paper¹⁴, and is clearly demonstrated by the experimental results shown in figure 8, that wheel layer formation is detrimental to process efficiency and prestige.

Electrochemical wheel-conditioning techniques have found only very limited use. This is possibly due to difficulties arising in its application; especially to incorrect operating resulting in considerable diamond loss. Another reason, equally significant, is that, as already stated, the main E.C.G. method used in industry is still that of electrochemical face grinding. When surface grinding, however, the machining conditions are different; the requirement for accuracy and economics is much higher and hence the question of grinding-wheel conditioning all the more important.

Truing an E.C.G. wheel

Running accuracy of the wheel should refer to the bond surface rather than to the grit envelope since it has been shown¹⁷ that reference to the grit envelope is not usually accurate enough and leads to faulty conclusions. After a conventional mechanical truing (e.g. by means of a silicon carbide wheel) the wheel-bond surface can be trued under practical conditions, by an 'electrochemical truing' operation.

If ΔB_{\max} is the wheel-bond waviness and if the electrochemical truer is being fed towards the wheel face in separate 'steps', then it can be shown that above a certain truing time, the remaining wheel bond waviness $\Delta B'_{\max}$ per step is approximately given by

$$\Delta B'_{\max} \sim \epsilon_1 \Delta B_{\max} \quad (1)$$

where $\epsilon_1 \text{ min} \sim 0.4$. It can be also shown that for a particular wheel/truer geometry the electrochemical truing time per step depends on the square of the wheel waviness. A continuous up-feed is impracticable because it is likely to lead to electrolyte starvation at high spots and because the grits will then mechanically machine the truing surface. If an electrochemical truing operation is to be effected it must take place in successive steps and only for reasonably low values of waviness. Figure 2 records bond waves of an E.C.G. metal bonded diamond wheel (Diamond grit size 120, concentration 50, 8 in o.d. \times 1 in rim width \times $\frac{1}{16}$ in diamond impregnation.) which correspond to three successive electrochemical truing steps I, II, III. A mechanical truing preceded by means of a brake-control device using a 37C60-MVK silicon carbide wheel. The aqueous electrolyte solution was a 10 per cent concentration of sodium nitrite and sodium nitrate¹⁴. The records were taken and are superimposed according to the

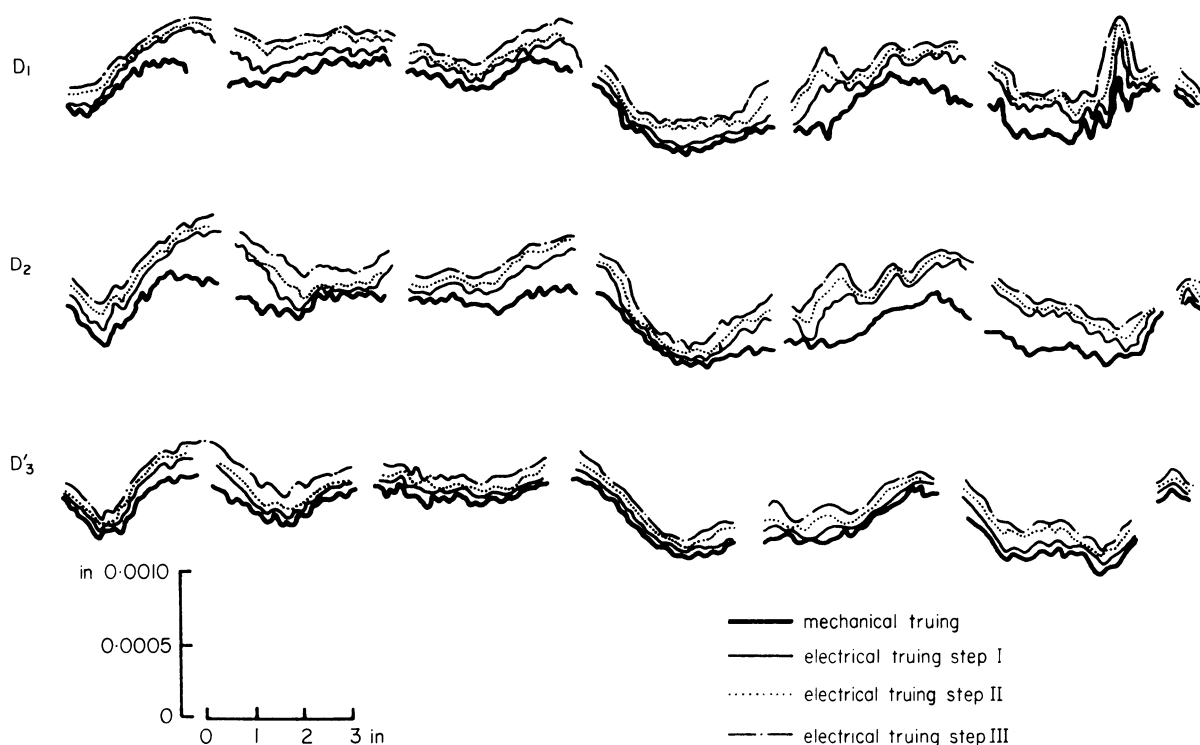


Figure 2. Bond waves following mechanical and electrochemical truing sequence.

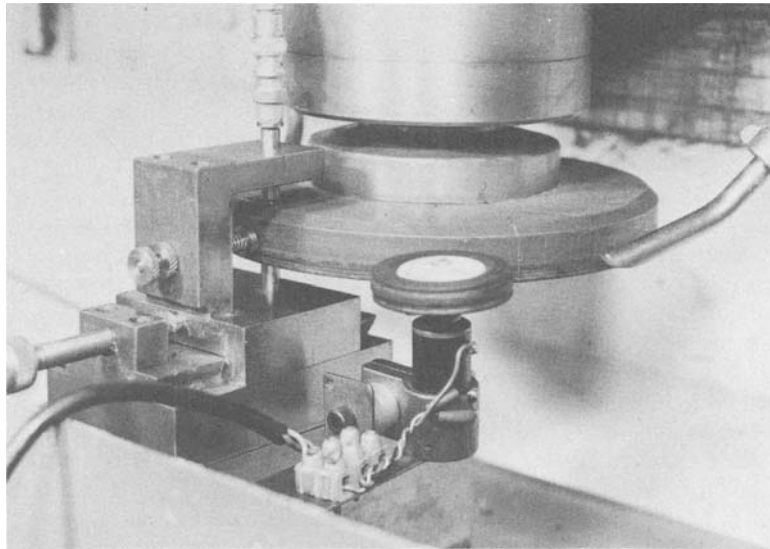


Figure 3. Rig for capacitive transducer measurements of wheel wear.

observed wheel 'wear', by means of a specially designed and calibrated experimental rig employing capacitance (distance) transducers. Part of it is shown in figure 3 fixed on the table of a vertical spindle, Abwood surface grinding machine, model SG4HE. The material for the truer was tungsten carbide I.S.O. P10 block $\frac{1}{2}$ in \times $1\frac{1}{2}$ in². Figure 4 diagrammatically represents the truing-current density and (linear) wheel-wear variation per step. The latter has been evaluated as the average of the graphical integration of the wheel-wear area obtained by the superimposed records of the wheel-bond waves. Characteristic machining current and spindle load u.v. records for each case are quoted in figure 5. It can be seen that the largest wheel-bond removal occurred with the

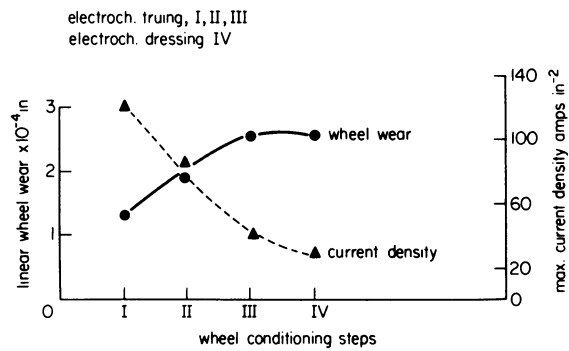


Figure 4. Wheel wear and current density during electrochemical conditioning.

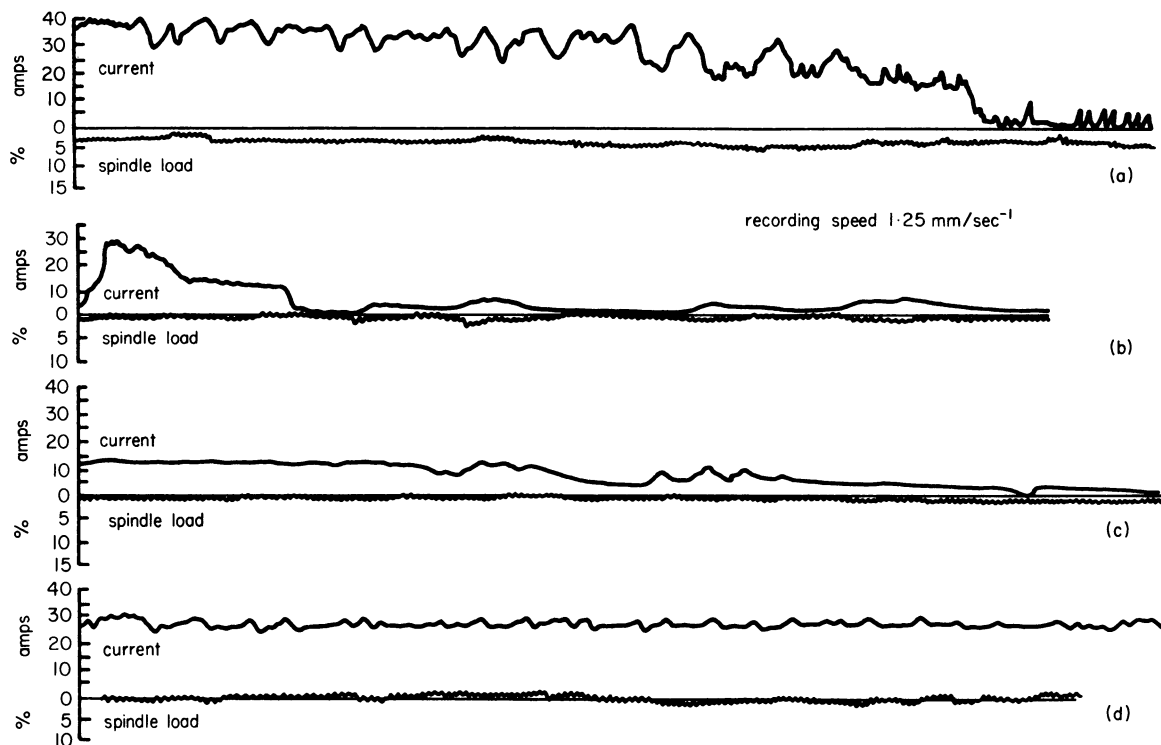


Figure 5. U.V. records of current and spindle load, (a) (b) (c) during truing and (d) during cleaning.

first truing step. Apart from the apparent material removal from the 'higher' parts of the bond surface rather than from the 'lower' ones, local electrochemical attack can also be noticed. This is probably caused by inhomogeneities of the wheel body on the wheel assembly. Nevertheless the record of figure 5(a) demonstrates a fluctuating, relatively high, current density for a long time and suggests that the apparent flatness obtained through the mechanical truing is due to a 'polishing action' which fills the gaps with bond and grit debris. Electrochemical dissolution and limited spark erosion during this first step removed almost all of the loose bond-grit material. It is, in effect, a 'wheel de-burring' operation. The average gap increased and the machining current-density gradually dropped during subsequent steps, as might be expected. The way that the current was developed, however, during these steps implies formation of an anodic film which fortunately restricts heavy electrochemical attack. From this point of view, the produced grit protrusion height is discussed later in the paper.

The total (linear) wheel wear just exceeds 0.00025 in. In general, however, there is no remarkable improvement over the initial wheel-bond waviness. In fact it has been made slightly worse. Compared with theoretical predictions for the method (Equation (1) and similar) the practical problems of inhomogeneities of the electrical circuit, the electrolyte solution flow, and distribution conditions in the inter-electrode space, result in a poorer tool, from the waviness point of view, than that from a conventional method. It is, however, very necessary for removing all the loose and poorly held bond-grit debris which fills the space between the diamond grits and, even more important, for removing the conductive 'bond peaks' left by the mechanical action and which protrude to the same height on the grits. If these remain considerable initial wheel wear must be expected, due to high current arcing, short-circuitings and sparking, and will seriously affect wheel performance.

Dressing the E.C.G. wheel

Theoretical, as well as experimental evidence, has shown that for the range of diamond grit size (i.e. 85 . . . 170) applied in E.C.G., the grit protrusion coefficient which corresponds to the average electrochemical gap required for continuous and efficient electrochemical action is given approximately by

$$\lambda \sim 2RY_0 \quad (2)$$

where λ is the grit protrusion coefficient, R the grit size and Y_0 the average electrochemical gap. For gaps varying between 0.0006 in and 0.0012 in it is concluded that λ lies well below 50 per cent, with an average of about 25 per cent, for most E.C.G. cases. If, now, an 'electrochemical dressing' is attempted for increasing the average grit protrusion coefficient from λ to λ' , it can be shown that the required time, assuming a negligible wheel-bond waviness, is given by the expression

$$T \sim K \frac{\lambda'^2 - \lambda^2}{R^2} \quad (3)$$



Figure 6. 105x magnification of active surface of wheel after electrochemical dressing.

where K is a constant of the dressing conditions, including geometry. Equation (3) stresses the extreme sensitivity of any electrochemical dressing operation, particularly when λ and λ' are close to each other.

Measurements of the average grit-protrusion height for the electrochemical truing case of figure 2 showed that, in fact, the process should be extended in order to achieve a 25-30 per cent protrusion. Grit projection measurements (as well as those for wheel wear and process records) were made without removing the wheel from the machine spindle, again using capacitance transducers and taking the average of a large number of values. Results obtained with this method were found to be the same as those taken by employing a short focus microscope and/or wheel surface replicas, but such methods require removal of the wheel, and this must be avoided because it is extremely difficult to exactly remount to the spindle. The in-process method described has the added advantage of distinguishing between 'bond peaks' or conductive areas and the revealed diamond grits. The relative frequency with which the two phenomena are revealed can then be used as a measure of the condition of the wheel's working surface.

Electrochemical dressing was conducted using EN8 steel as dresser material, with a maximum current density of 27 A/in² for 10 minutes. For the central area of the diamond layer of the wheel a grit-protrusion height of 0.0014 in was produced with almost immeasurable wheel wear (figure 4). Figure 6 is a magnification of the wheel's active surface with the diamond grits protruding from it. A final bond ratio¹⁷ of 0.4 was obtained.

Average grit projection after mechanical truing varied between 0.0012 and 0.0014 in. Attempts to remove from the wheel the conductive bond peaks and loose or semi-held bond-grit material using aluminium oxide WA12OLV abrasive sticks resulted in some improvement (linear wheel wear approximately 0.0001 in and increase of grit protrusion of the order of 0.0002 in). However, the wheel surface still possessed a very high density of conductive spots.

E.C.G.-wheel performance

Figure 7 illustrates (linear) wheel-wear diagrams following actual electrochemical grinding of tungsten carbide I.S.O. K20 $\frac{3}{4} \times \frac{3}{4} \times \frac{5}{16}$ in³ workpieces. The grinding conditions were 5 V applied potential, 0.0002 in up-feed per table stroke and 1 ft/min table

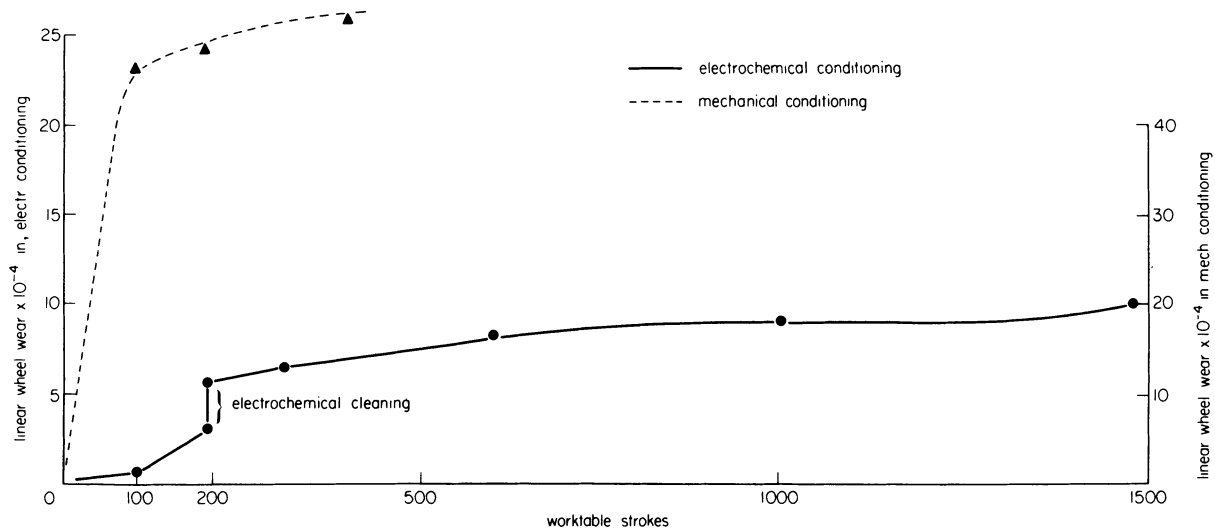


Figure 7. Linear wheel wear during E.C.G. after mechanical and after electrochemical conditioning (note scale difference).

feed rate. Electrolyte flow rate was 6.5 litres/min and the wheel speed 3000 rev/min.

Low values of voltage, up-feed and table feed-rate were selected^{14,17} in order to avoid poor machining conditions through increased mechanical removal, sparking, arcing, and strained electrolyte supply in the machining gap.

As can be seen the initial wheel wear following mechanical conditioning—mainly due to high arcing and sparking as previously discussed—impressively surpasses that following electrochemical conditioning. Even taking into account 'wear' during electrochemical conditioning (figure 4), the observed loss represents approximately 23 times increase or about 7 per cent of the initial diamond-layer thickness. Although after successive mechanical removal of arc-damage from the wheel's working surface by means of an abrasive stick a lower wear rate is evident, nevertheless, it is not always possible to clean the wheel effectively over its whole surface, particularly when heavy arcing has occurred. As a consequence further arcing and sparking is initiated and an overall

reconditioning is needed. It is, of course, possible to 'condition' the wheel by applying a sequence of electrochemical grinding and frequent interruptions for mechanical removal of the arc-damage until a final stable situation is reached. This is, however, a more tedious final conditioning procedure than the electrochemical one and involves the danger of permanently damaging the wheel, in addition to costing much more in terms of useful overall wheel life.

Cleaning the E.C.G. wheel

The variation of current density and spindle load with number of worktable strokes is shown in figure 8. The wheel was conditioned electrochemically, and figure 8 corresponds to figure 7 as far as the electrochemical conditioning is concerned.

A rapid drop in current density coupled with a rapid increase of spindle load is demonstrated after the first hundred strokes. The wheel wear-rate also increases as a result of the considerable mechanical cutting action then taking place. On the working face

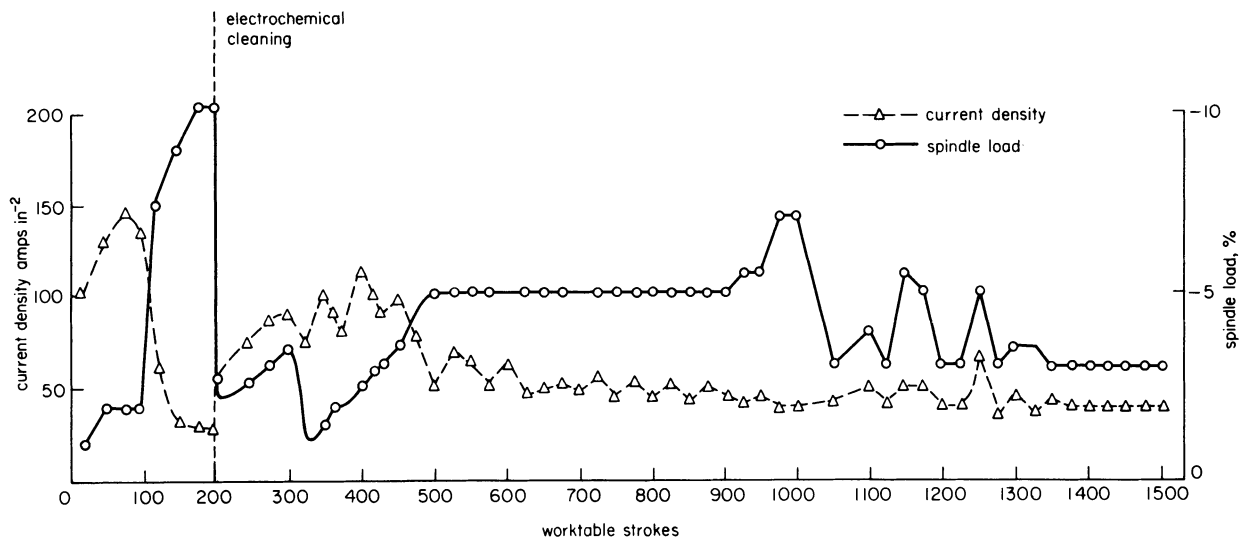


Figure 8. Current density and wheel load during E.C.G. after electrochemical conditioning.

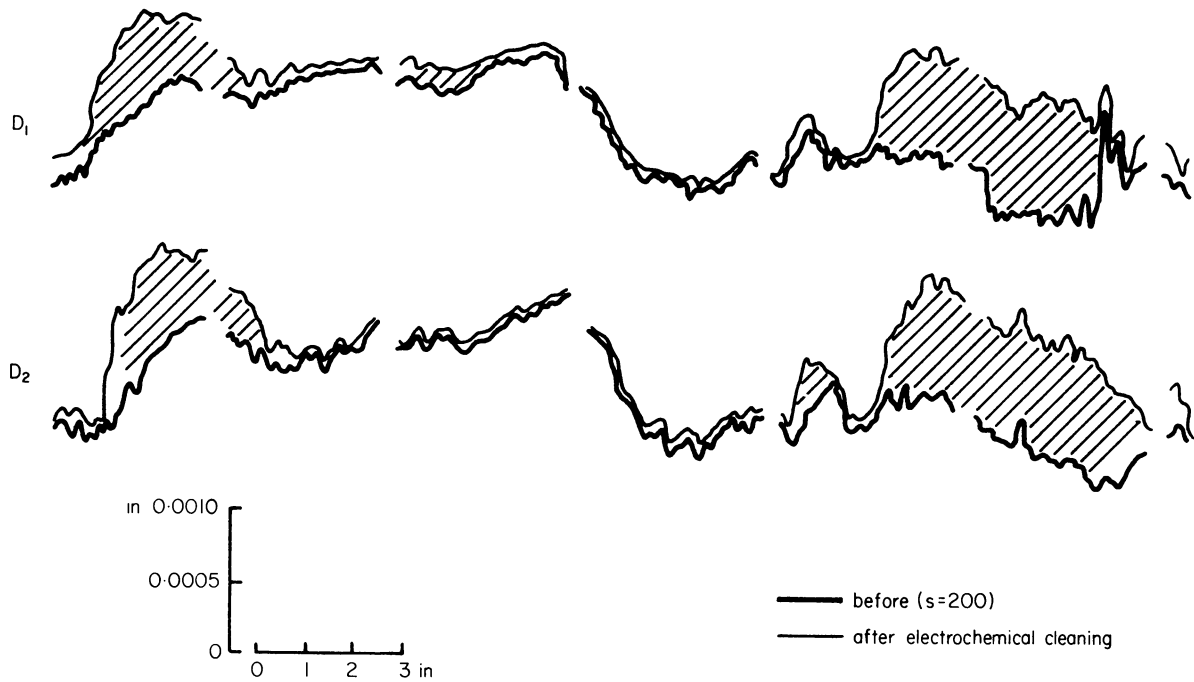


Figure 9. Bond waves after electrochemical conditioning followed by 200 E.C.G. strokes, and immediately after electrochemical cleaning.

of the wheel a highly resistive bond layer had formed. This layer is very characteristic of E.C.G., especially when grinding tungsten carbide, and so is thought to consist of a thin film of cobalt and tungsten-oxides. However, experimental evidence has also shown that the type of wheel bond plays a decisive role in its formation^{14,18,19,20,21,22}.

Results obtained after electrochemical cleaning of the wheel (based on the same principles as those after dressing) are incorporated in figures 7 and 8. Cleaning current and related spindle load are included in figure 5. 'Cleaning-wear' diagrams in figure 9 confirm the effect of wheel layer in restricting current during grinding and demonstrate the nature and magnitude of electrochemical attack—in fact preferential removal occurs over almost half the wheel surface. Three comments may be made here: (1) The layer does not uniformly cover the whole of the wheel's working face but its build-up and severity depends on local conditions. (2) The uncleaned parts coincide with the valleys of the bond waves where electrolyte supply condition is less strained and hence where electrochemical activity during grinding is most intense. (3) A thorough wheel cleaning was not achieved even after 20 minutes operation, but further electrochemical action is not thought to be reasonable as it would rather amplify the wheel waviness. It is to be noted that no significant difference in the grit-protrusion height was produced.

Figures 7 and 8 also incorporate grinding results after wheel cleaning, to complete the picture of a typical sequence in grinding practice, i.e. grinding-cleaning-grinding. Figure 7 also includes the average cleaning wear. As can be seen the current density never again reached the level of the 'pre-layer' values, for reasons explained above. In fact, as the number of strokes increased, the mode of current reduction is a gradually damped oscillation, as the wheel layer

begins to build up again. It is followed by a gradual rise in mechanical action. Two additional remarks on figure 8 are of interest: (4) the gradual reduction in current density to below 50 A/in² after about 1000 strokes leads to an increase in mechanical action and subsequently to current and spindle load fluctuations. The phenomenon suggests partial breaking-up of the bond-layer due to increased mechanical forces and wear but not sufficiently to restore high current or to assure a 'self-cleaning' procedure. After a few hundred further strokes the current density returned to its previous low level. (5) The wheel layer effect developed most rapidly and with greater severity in the case of the freshly conditioned wheel (during the first 200 strokes) and was associated with maximum current densities. It is apparent that after 450-550 strokes wheel reconditioning is needed.

CONCLUSIONS

Following a mechanical truing operation the application of 'electrochemical wheel conditioning' to remove mechanical debris, conductive bond 'peaks' or plateaus, and to expose the diamond grits is shown to be quite reasonable. It will significantly extend the useful life of the wheel and assure stable electrochemical grinding. It must be carried out in steps and with extreme care, however, so that it will not destroy the wheel-bond geometry or over-expose the diamond grits.

Formation of a resistive layer on the working surface of the wheel is detrimental to process prestige but 'electrochemical cleaning' of the grinding wheel can considerably improve process efficiency provided it is applied during early stages of the operation. An automatic grinding/electrochemical cleaning/grinding cycle will appear to be a sensible development, at least when grinding tungsten carbide.

ACKNOWLEDGMENTS

The authors thank Abwood Machine Tools Limited for the use of their electrochemical grinding machine, Hoy Carbides Limited for tungsten carbide workpieces and U.M.I.S.T. for instrumentation and equipment used during the study in the Royce Laboratory. Thanks are also due to the National Technical University of Athens for the S. Niarchos scholarship held by the first-named author.

REFERENCES

1. J. S. Selby. The pre-grinding preparation of diamond abrasive grinding wheels, *Carbide and Steel Grinding for the Seventies*, De Beers Industrial Diamond Division, London, 1970.
2. C. Davis. The grinding of an annealed steel with diamond abrasive wheels, Ph.D. Thesis, U.M.I.S.T. 1970.
3. L. H. Metzger and G. Keeleric. Electrolytic grinding, *American Machinist* (10 Nov., 1952) pp. 154-8.
4. D. Fishlock. What electrochemical machining can do, *Metalworking Production*, (7 Dec. 1960) pp. 94-8.
5. G. F. Keeleric and L. A. Williams. Verfahren zum Abrichten eines Schleifwerkzeuges, Deutsches Patentamt Auslegeschrift 1113388/830929Ib/67a 31 Aug. 1961 (U.S.A. Patent No. 355932, 19 May 1953).
6. W. Unterbusch. Wirtschaftliches Arbeiten mit Elektrolytischen Schleifmaschinen in der Hartmetallaufbereitung, *Werkstatt und Betrieb*, (1962) 95, no. 8, pp. 539-42.
7. R. R. Brandi. Improving diamond wheel performance in electrolytic machining, *The Tool Manufacturing Engineer* (Jan. 1961) pp. 80-2.
8. D. Fishlock. Electrolytic grinding scores five ways, *Metalworking Production* (30 Aug. 1961) pp. 74-6.
9. J. C. Wilson. *Electrolytic Surface Grinding*, The Thompson Grinder Company, Springfield, Ohio.
10. Yu. A. Balashov. Electrolytic grinding, *Machines and Tooling* (1966) 38, no. 3, pp. 31-4.
11. E. K. Pronin, et al. Wheel dressing methods for electrolytic grinding, *Russian Engineering Journal* (1969), vol. XLIV, no. 7.
12. T. D. Kelso. E.C.G. helps Buick grind down costs, *Manufacturing Eng. Management* (Aug. 1970) pp. 23-4.
13. J. R. Chalkley. Electrolytic tool grinding, Lecture given at the Coventry Technical College, October 1967 (issued by Impregnated Diamond Products Limited).
14. M. M. Sfantsikopoulos and C. F. Noble. Vertical spindle electrochemical surface grinding, *Proc. 12th Internat. M.T.D.R. Conf.* 1971.
15. A. P. Nikitin and G. I. Stepanov. Etching to expose the grains of diamond wheels, *Machines and tooling* (1969) 40, no. 8, pp. 46-7.
16. A. Büttner and D. A. Lindenbeck. The Electrolytic dressing of diamond wheels for use in steel grinding, *Industrial Diamond Review* (1969) 29, no. 348, pp. 450-54.
17. M. M. Sfantsikopoulos and C. F. Noble. Dynamic and geometric aspects of vertical spindle E.C.G., *Proc. 13th Internat. M.T.D.R. Conf.* 1972.
18. R. R. Brandi. Improving diamond wheel performance in electrolytic machining, *The Tool and Manufacturing Engineer* (Jan. 1961) pp. 80-82.
19. M. M. Sfantsikopoulos. Electrochemical surface grinding. Ph.D. Thesis, UMIST, Jan. 1973.
20. H. Reinhart and W. Grunwald. Zum Elektrolytischen Abtrag von Hartmetall mit diamantschleifscheiben, *Werkstatt und Betrieb* (1962) 95, no. 4, pp. 212-18.
21. M. G. Poore. Electrolytic grinding of tungsten carbide, Part II, Thesis, The College of Aeronautics, Cranfield, 1966.
22. V. N. Gavrilov and Yu. B. Burochkin. Effect of wheel bond on electrochemical diamond grinding, *Machines and Tooling* (1969) 40, no. 2, pp. 43-4.
23. A. Kobayashi. Recent development in diamond wheel grinding, *Annals of the C.I.R.P.* (1970) vol. XVIII, pp. 31-47.

ELECTRO-DISCHARGE MACHINED SURFACES

by

J. R. CROOKALL* and B. C. KHOR*

SUMMARY

Existing published work on the nature of electro-discharge machined surfaces is reviewed and discussed, including deformation and damage, and metallurgical structure and physico-chemical changes.

Using both optical microscopy and scanning-electron microscopy, the results of an investigation of surface effects and characteristics of both pulse-generator and relaxation-circuit EDM machines on HW5 tool steel, EN1 steel, and copper workpiece materials are presented. The very complex nature of these surfaces is demonstrated. They possess characteristics peculiar to EDM on account, not only of the severity and rapidity of local heating and cooling which occurs, but also due to the discrete, random attack of the discharges. In the light of the evidence presented, conventional roughness measurements by stylus instruments provide a very limited indication of the effective surface characteristics.

INTRODUCTION

Electro-discharge machining is carried out within a dielectric fluid, with a small gap, typically of 0.08 mm, in which the rapid succession of discharges occur. Thus to all intents and purposes, no force is exerted by one electrode upon the other, and material is removed by melting, and possibly evaporating, the discrete areas at which the discharges occur. The material is then expelled into the surrounding fluid, following the collapse of pressure on cessation of the discharges. Although material removal at each discrete location is on a microscopic scale, it is nonetheless a violent process, which, in the course of time, attacks each portion of the machined surface. It is capable of eroding any metallic (electrically conducting) material, irrespective of its mechanical strength, and, as might be expected from the nature of the process, metal removal ability is related fairly consistently to workpiece thermal properties.

The resulting applications are frequently to the machining of complex shapes, thin slots and holes (e.g. in turbine blades) and particularly in materials difficult to machine by conventional methods, including the 'high strength temperature resistant' materials. This implies that a relatively high proportion of applications are in situations requiring high resistance to stress, fatigue, creep and possibly stress-corrosion.

Surface integrity is a general term covering various ways in which the state of the surface, or near-surface region exhibits properties inferior to, or at least

different from, the bulk parent material. This inevitably involves consideration of the metallurgical, textural, mechanical (e.g. hardness), and (in EDM particularly) compositional properties. However, the techniques of investigation, used in the evaluation of each of these aspects, are widely different. Hence in other than a really comprehensive investigation into all relative factors, surface integrity in a given situation is inevitably defined in terms of the particular technique or techniques used. Even when many such aspects are known about the quality of a surface, their interpretation for functional purposes when present simultaneously, is difficult.

In a previous paper¹ the residual stresses, due to three types of EDM machine, were investigated. A significant finding was that although the *distribution* of residual stress with depth into the subsurface region varied in certain respects from one machine to another, and with machining conditions on a particular machine, the *maximum value* of residual stress, which normally occurred at or near the surface, was to a first approximation less influenced in this way. High tensile residual stresses were invariably found near the machined surface, which approached the UTS of the material. The (reviewed) work of others indicated, however, that the reductions in fatigue life accompanying electro-discharge machining was due more to the textural and other properties of the surface than to residual stresses alone.

The present paper is concerned primarily with surface effects revealed by optical and scanning electron microscopy. Some observations are made on

* Imperial College of Science and Technology

the validity of surface texture measurements revealed by stylus-type instruments. The compositional and physical characteristics of electro-discharge machined surfaces are reviewed.

THE ELECTRO-DISCHARGE MACHINED SURFACE

Discharges and their effects

During an individual discharge three distinguishable stages may be observed

- (1) An *ignition* phase consisting of ionization and arc formation at some localized area
- (2) The main discharge—an electron *avalanche* which strikes the anode. Very high current densities are reached due partly to the narrow zone of nevertheless low resistance following stage 1, and also maintained by hydraulic restriction of the dielectric fluid and the magnetic *pinch* effect. Later, ions strike the cathode which is heated less rapidly than the anode
- (3) Local melting and even evaporation of the electrode metal occurs at the site of the discharge, followed by its expulsion. Expulsion appears to occur mainly after the electrical discharge has ceased, and is frequently accompanied by the evolution of gas. With increasing discharge duration, current density decreases, and the discharge tends to assume the characteristics of an arc. Electro-discharge machined surfaces have a characteristically *matt* appearance comprising microscopic craters (diameter to depth ratio² from 5:1 up to 50:1) associated with the discrete discharges

Naturally, studies of the craters formed are reported in the work of quite a number of contributors to the EDM field. Observations and sometimes empirical expressions for the volume or diameter and depth of craters, related to electrical parameters, are to be found. Some of these date back to earlier work in the 1950's when relaxation-circuit generators predominated. Such relationships, although undoubtedly true within the range of conditions adopted by the experimenters, can be seen in retrospect to be insufficiently general.

However the importance of this aspect is mainly twofold, namely that crater volume itself (metal removed per discharge) is directly related to the surface roughness produced, and also, when it is multiplied by the effective discharge frequency, yields (in theory at least) the metal removal rate. Nevertheless the actual situation is far from being as straightforward as this because, without specialized equipment for counting discharges, the *effective* pulse frequency is not known. Furthermore other than with a current-pulse timed generator (isopulse) many active pulses will be only fractionally effective, and metal removed per pulse is not directly proportional to the pulse charge or energy. These aspects are discussed in some detail elsewhere³.

It should also be noted that the experimental expedient of dividing metal removed over an appreciable period of machining (e.g. 15 minutes) by the actual or estimated number of effective pulses, to

yield a figure for metal removed per pulse, inevitably yields only a time-averaged result, and *ipso facto* begs the question of the true value and its characteristic dependencies.

Progress in understanding the basic underlying principles has resulted from careful quantitative studies of the thermal basis of metal removal by Zolotych⁴, and recently by Van Dijk⁵, using an analytical heat-flow model in which growth of the plasma channel with time during the discharge was an assumed condition. Such work has demonstrated not only a general basis for the explanation of removal characteristics related to power and duration aspects of the discharge and the physical conditions involved, but also good quantitative agreement with observations over a range of conditions and circumstances. For example, the basic differences in effect between the relaxation and the pulse generator types of supply are explicable in such terms, including *preferred polarity* effects and electrode physical properties. A summarizing account of this and other relevant and related factors is to be published shortly⁶.

Such investigations, however, must be classified by implication as being upon the *single discharge* basis. Between this and the relatively unsophisticated *time-averaged* approach lies an important area of investigation in relation to the mutual interactions and successive dependencies within a train of discharges, which constitutes the practical domain of electro-discharge machining. Here stochastic as well as deterministic processes are evident⁷. It is not surprising therefore that erosion rate calculations, based upon the repetition of deterministically derived, or measured, results of single discharges, can be of only limited validity. In practice also it is known that craters are partially superimposed, and may become contaminated with transported material from the tool electrode, and products from the dielectric fluid. However it is probable that an early conclusion due to Kravetz⁸, showing a reduction in surface roughness by one half requires about an eightfold reduction in machining rate, is nevertheless still reasonably representative, at least as a general approximation.

Cole, Bucklow et al. have used a scanning-electron microscope to examine the appearance of the sparked surface of aluminium, germanium, iron, and tungsten⁹ and aluminium and stainless steel². They reported that the surfaces appeared similar for all the materials, as though a pool of molten metal had been ejected from a series of overlapping craters. Depending on the input energy and material, crater sizes were about 10 μm to 15 μm deep. Other features included *pock marks*, *chimneys* and re-solidified metal droplets.

Surface deformation and damage

Fusion and plastic deformation are always associated with the craters due to electro-discharge machining¹⁰. This is attributed to the molten liquid solidifying epitaxially when quenched by the dielectric. The high thermal contraction produces high stresses which in turn cause severe slip, twinning and cleavage, depending on the crystal structure. Another probable reason is that shock waves are produced which propagate from the discharge region and which

deform the metal¹⁰. The material is also known to work harden.

These changes accentuate the grain-boundary weakness in the epitaxial layer, and grain boundary cracking is often found, even in ductile materials. 70/30 brass¹¹ and some hardened martensitic steels¹² were found to be particularly prone to deep grain-boundary cracking. Cleavage cracks were found in chromium and tungsten¹³ at temperatures below their ductile/brittle transition. Barash¹⁴ calculated the thermal stresses on tool steel due to a single spark, and found that this stress was very near to the UTS value, and could cause rupture of the material. He deduced the presence of permanent damage such as microcracks, and estimated the depth to which these extend to be between 2.5 and 5 times the CLA roughness value of the surface¹⁵. This compared with a factor of 10 times for a previous experiment, but it was noted from analytical considerations¹² that the thickness of the damaged zone would be dependent upon the parameters of the discharge, and a higher energy would give a thicker zone. Aleksandrov¹⁶ has also reported that the longer the pulse duration, the deeper the damage. By layer removal and fatigue testing, he showed that the damaged layer with a pulse of 1 050 μsec duration was 0.2 mm, but at 130 μsec it was only 0.04 mm.

Using the broadening of X-ray diffraction lines as an indication of deformation, by successive removal of layers, Wilms and Wade¹⁷ reported that the depth of damage below the bottom of the craters was 0.005 in (127 μm) for chromium, 0.008 in (203 μm) for aluminium, and 0.006 in (152 μm) for iron and for antimony. Microscopic examination also showed the four metals to be plastically deformed around the craters. This was most marked for aluminium and chromium, on which slip lines were observed some 0.050 in (1 270 μm) from the edge of the hole: it was 0.010 in (254 μm) in iron. Antimony, which also showed cleavage, appeared to have more localized damage. Cracks were observed in chromium and antimony (brittle materials) but not in iron and aluminium (ductile materials).

In addition to the above manifestations of surface deformation, the extent of which is known as the *macro-strained zone*, there is also a *microstrained zone*. No slip, twinning or cracking occurs, but extra dislocations are induced by EDM. While knowledge of this may not be particularly relevant to many practical applications, it is usually essential, when preparing specimens for research purposes, to know the degree and depth of damage which has occurred. This is done by dislocation etching; however the technique is confined to single crystals in a relatively undamaged condition, because dislocation densities greater than 10^6 mm^{-2} cannot be resolved. Bucklow and Cole² have reviewed the existing work on this aspect. The results demonstrate that the dislocations penetrate much further in copper for example, than in silicon iron, that the background dislocation density has an appreciable influence on the 'depth of detectable damage' (i.e. the depth to which any changes whatsoever could be detected by the most sensitive means), and that the form of the tool used may also be important.

However, it must be pointed out that the sensitivity of the techniques are different. The dislocation etching technique, where applicable, would give the highest sensitivity. Brown² and Woodhouse² demonstrated that the X-ray peak intensity was almost as sensitive as dislocation etching, but confirmed Michelle's² observation that X-ray diffraction line broadening was not very sensitive. It gave depths underestimated by a factor of about 10-20, when compared with the X-ray intensity method. The sensitive as dislocation etching, but confirmed on the microscopic manifestations of the damage.

Structural and physico-chemical changes

Unusual phase changes in electro-discharge machined surfaces can be expected in view of the high temperatures attained. The molten surface of a spark crater is suddenly charged with alloying material, only to be rapidly quenched by conduction through the crater base, and possibly the dielectric.

The presence of a *white layer* has been reported by many workers in the past, but not fully investigated. Rudiger and Winkelmann¹⁸ reported a non-etchable *white covering layer* which was far harder than the basic grain structure. Barash and Sri-Ram¹⁵ observed its formation under all machining conditions, even when using water as a dielectric. Walton and Wendell¹⁹ also observed such a layer during routine examination of ED-machined tools. They found the layer on heat-treated materials, subsequently machined by electro-discharge machining, but not on tools which were hardened after machining. This is because the layer had been masked by the heat-treatment.

Lloyd and Warren²⁰ have reviewed some work on the phase changes in spark machined surfaces on pure iron, steels, titanium and copper. In those cases where the dielectric cracked to release pyrolysis products of carbon and other unsaturated hydrocarbons, e.g. paraffin, new phases were formed. An austenite/carbide surface formed on pure iron, and a titanium carbide surface on the titanium. Higher alloy steels and high-tungsten tool steels showed extensive phase changes and indicated a carbon-saturated surface²¹, together with a sharp gradient in the underlying material of carbon and other alloying elements introduced via the cutting electrode. This, together with the thermal cycle, gave rise to marked structural changes in the subcutaneous zone, in a manner analogous to welding processes.

The use of water as a dielectric produced oxides at the surface. In the case of titanium, a very hard and cleaved surface of almost pure hexagonal Ti_2O was formed, while mixed oxides formed on ferrous materials.

Severe discharge conditions under a paraffin dielectric (e.g. when using graphite tools) gave a fused layer of dendritic austenite and an austenite-cementite eutectic, i.e. the ledeburite structure of a hypoeutectic white cast iron, which has a hardness of 1200-1300 kg/mm^2 . Less severe discharge conditions (e.g. copper electrodes) gave a completely austenitic surface with an intermediate austenite-martensite matrix above the base material. This structure, even in pure iron, was attributed to the introduction of

carbon via the cracked hydrocarbon dielectric, and indicated heating to above its critical temperature. Where the tool electrode is compatible, alloying at the surface was found. Copper, for instance, alloyed readily with high nickel alloys to form a cupro-nickel surface, and copper diffusion was present in ferrous metals. It was considered that this alloying effect could be used, for example, to enhance surface hardness, or to reduce residual tensile stresses by a suitable source of alloying element. Johnson and Radon²² have used an electron probe micro-analyser to obtain plots of copper across sections of the surface machined with a copper-beryllium electrode and found 90 per cent of the total transferred copper located in a layer about 10 μm thick adjacent to the eroded surface, but lower concentrations were present up to about 17 μm from the eroded surface.

Opitz²³ reported the presence of a ledeburite structure which changed from the outside inwards from a hyper- to a hypo-eutectoid concentration (i.e. the border zone) when he machined a hot-forging die material and a 0.5 per cent C steel at high energies in kerosene with a copper electrode. Copper inclusions were also observed. He also suggested carburization was from the dielectric, and found with pure iron, carburization of the border zone with 0.03 per cent to 0.28 per cent C. He distinguished and categorized the structure formation into a series of zones, which he called the *border zone*, *intermediate zone*, *conversion zone*, and the basic grain structure. The border zone was characterized by the change of composition of the original workpiece material. At lower electrical power inputs, the *border zone* was a non-etchable *white layer*. The *intermediate zone* appeared only after a rough machining operation, was only a few μm wide, and had an austenitic grain structure. The *conversion zone* was characterized by a change of structure from the basic grain structure.

The depth of these zones depended upon the characteristics of the pulse used. It was suggested that this indicated that it is not the energy of the individual pulses, but the converted power in the working gap that determines the depth of the surface changes. He also showed that the width of the zones depended on the shape and work area of the electrodes. Small deep holes or sharp angles encourage arcing due to difficult swarf removal, causing a wider modified zone.

Not unnaturally, reports on phase changes have been mainly concerned with the carburization of ferrous materials. However, non-ferrous alloys do exhibit phase changes also. Kawai et al.²⁴ reported 'abnormal structures' (i.e. structures which differ considerably from the equilibrium or chill-cast state) on the surfaces of Ag-Cu alloys and cemented carbides. It was considered that they were formed by a similar mechanism to the martensite transformation in ferrous materials, i.e. the suppression of the eutectic reaction. On eroded surfaces of Ag-Cu alloys which showed typical eutectic structures were formed homogeneous solid solution, which reverted to the equilibrium state of two phases with marked precipitation hardening on heating above 200°C. The cemented carbide alloy surface showed 'abnormal structures' of γ and ϵ above an α phase matrix, and

these also reverted to the normal phase on tempering from about 650° to 850°C. This was taken to suggest that the origin of these 'abnormal structures' was not a change in composition, but simply the change in the crystal structures.

Other effects

Crystal lattice defects have also been reported. Reference was made by Lloyd and Warren²⁰ to the work of Palatnik²⁵ in a study of Cu, Ag and Au in which a reduction in the lattice parameters occurred. This was attributed to a vacancy concentration of 0.4 per cent, which is very high (over a hundred times greater than is found with normal quenching) further indicating the extraordinarily rapid quenching of the fused layer.

EXPERIMENTAL TECHNIQUES AND MATERIALS

In the present investigation, the structural and surface effects of electro-discharge machining were investigated by optical microscopy and scanning-electron microscopy. In the light of this, the evidence of stylus-type surface roughness measurements were examined.

Optical microscopy (normal illumination)

Normal illumination optical microscopy is applicable where the reflecting surfaces are sufficiently flat to keep the full field of view in focus. However, its application is limited to a resolution of about 2500 Å in green light, and a depth of field of 1 μm at a resolution of about 5000 Å. Although this method is suitable for examination of structures in subsurface layers of the specimen, it is difficult for surface examination of EDM specimens, due to its highly irregular nature and particularly the relatively high slopes which much of the surface presents to the illumination direction.

However, optical metallography was used to examine the structures beneath the eroded surface, by observation of transverse sections of specimens. Plating of the specimens before metallographic preparation is recommended²⁶ to protect the surface structures. Copper plating was utilized for the copper specimens using an acidic copper sulphate bath (average concentration 0.16 kg per litre $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0.061 kg per litre H_2SO_4). The anode was a pure copper plate and a current density of 50–65 amps per sq metre²⁷ was employed. To copper plate steels, it would have been necessary to use a cyanide bath as the acidic bath gave a non-adherent deposit, but, because of its toxic nature, copper plating on steels was not pursued. Iron plating was therefore used on EN1 and HW5 steels using ferrous chloride (0.25 kg per litre) and calcium chloride (0.28 kg per litre) at around 80°C, which was kept slightly acidic by periodic additions of HCl. The anode was of pure iron and a current density of 200 amps per sq metre²⁷ was used.

After plating, the specimens were mounted, polished and etched in 5 per cent nital for the steels, and an ammoniacal solution of ferric chloride for copper. The specimens were examined on a Vickers projection microscope.

Scanning-electron microscopy

The scanning-electron microscope has a resolution, depending on the diameter of the electron probe and diffusion of the electrons in the specimen, typically of about 300 Å at a probe voltage of 10 kV to 20 kV. Depth of field can be of the order of 1 mm, and it is thus ideal for the examination of EDM surfaces.

An electron probe is scanned across the surface in a two-dimensional raster, and the back-scattered electrons are collected with a scintillation counter. This signal is fed to a cathode ray tube, scanned synchronously with the electron probe, and a point-by-point image of the specimen is obtained. To improve contrast and give a three-dimensional appearance to the image, the specimen is inclined at an angle to the incident beam.

The Stereoscan* scanning-electron microscope was used to examine the surfaces of EDM specimens prepared from strips about 3 mm in width cut from the machined workpieces. These were ultrasonically washed and mounted for examination in the Stereoscan at an illumination angle of 40°.

Surface-roughness measurement

Surface profiles of the ED-machined specimens were obtained with a Talysurf II† instrument. However, the roughness of surfaces produced during these tests was generally outside the normal range of the instrument. Thus a 'side-acting gauge' head was used in place of the standard head to obtain reduced sensitivity. This meant that the chart recorder could be used and calibrated for use, but the integrating meter would no longer give reliable readings on centre-line average roughness. 'Average peak-to-valley roughness' was thus taken from the chart recordings in these tests.

Materials and machines

Three different workpiece materials were used, namely HW5 tool steel, EN1 steel and commercially pure copper. The tool steel was selected as being representative of a material frequently machined by EDM, and upon which surface effects could be of significance. EN1 steel and copper were used because of their contrasting properties; the latter representing a non-transforming material.

The composition (per cent by weight) of the steels were as follows:

Darwin HW5 tool steel

C = 0.3; W = 1.0; Cr = 5.0; Mo = 2.0; V = 0.25

EN1 steel

C = 0.07-0.15; Si = 0.10 maximum; Mn = 0.80-1.20; S = 0.20-1.30; P = 0.07 maximum.

Two electro-discharge machines were used

- (1) Agietron‡ with Agiepuls generator
- (2) Servomet SMD§ with relaxation-circuit generator

The results are distinguished by the quoted pulse energy and duration, or circuit capacitance and resistance, respectively.

For a relaxation machine, the value of capacitance used, together with an experimental figure for the mean breakdown voltage, would indicate a value of average energy per discharge. As has been stated already, however, considerable differences exist between such pulse characteristics and their effects, and those from a rectangular-pulse generator, albeit stemming from several-order-of-magnitude differences in electrical, time, and energy density parameters particularly, rather than from any basic difference in 'kind'. Thus any simple measure of equivalence based upon discharge energy can be taken only as a rough first approximation. Some discussion of principles, upon which a quantitative measure of equivalence could be based, has been made³, but this is not a subject upon which it is possible to be conclusive at the present time. Nevertheless energy values calculated according to the above simple criterion will be quoted where relevant. (Brass tool electrodes and commercial paraffin dielectric were used for all tests.)

EXPERIMENTAL RESULTS AND DISCUSSION

Optical microscopy

HW5 tool steel

The photomicrographs in figure 1 are typical transverse sections through the electro-discharge machined surface of the HW5 tool steel. The section in figure 1(a) shows the protective iron plating, a sharply delineated white border zone and the matrix: micro-hardness indentations are also evident. The surface of

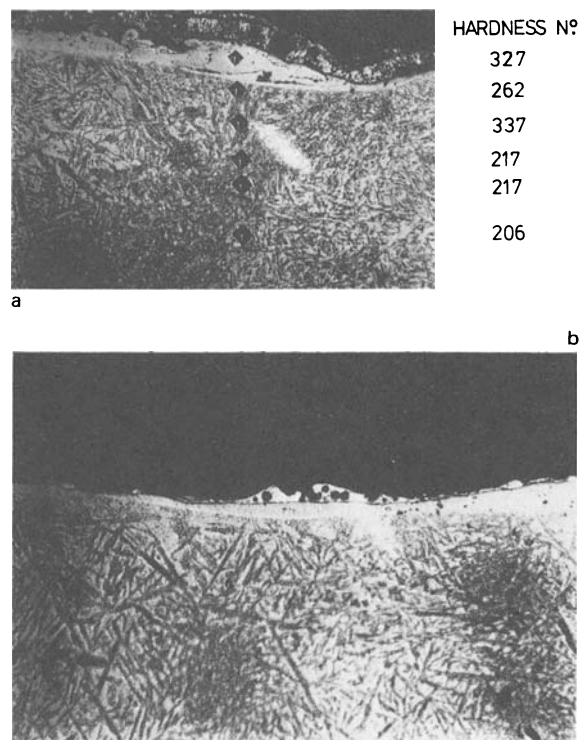


Figure 1. Photomicrographs of HW5 tool steel, x 500 (C = 1.0 μF; R = 108 Ω). (a) Iron plated specimen showing microstructural changes and micro-hardness values. (b) Unplated specimen showing tool-material (brass) transfer.

* Cambridge Instrument Company Ltd.

† Rank Taylor Hobson, Ltd.

‡ Agie Ltd, Switzerland.

§ Metals Research Ltd.

the white layer is irregular, showing the craters, and also signs of splashing during the ejection and resolidification. Porosity, evidently caused by entrapped gases, is also evident, and in a few cases pores, which are comparatively closely spaced and sometimes 'in line', have weakened the material sufficiently to cause cracking.

The *heat-affected zone* is not very evident in figure 1(a), but is clearer in figure 1(b). This zone is below the white layer and exhibits a much finer structure of martensite than the matrix, which has a tempered martensitic structure. Evidence of tool material transfer is also clear: material from the brass electrode is also ejected during the discharge cycle and some of this has resolidified on the workpiece as a separate layer above the *white layer*. This clearly suggests that the cathode heating occurs later than that at the anode. Entrapped gas, causing porosity, is evident in both the transferred tool material and the resolidified white layer. The thickness of the latter is not uniform and can be seen to vary from about 2 μm to 20 μm .

Vickers microhardness readings taken across the EDM-affected zone are also shown in figure 1(a), indicating the substantially increased hardness in the resolidified layer. The total depth affected by observable hardness variation was about 70 μm , whilst the white layer was about 5 μm to 30 μm in thickness.

EN1 steel

The transverse section of EN1 steel shown in figure 2 shows ejected metal in globular form, including a hollow globule (of about 30 μm diameter). This provides further evidence confirming Krasnyuk's²⁸

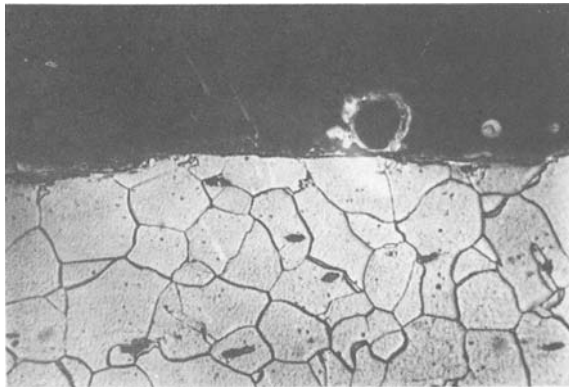


Figure 2. Photomicrograph of EN1 steel showing hollow globule of resolidified metal, $\times 500$ ($C = 1.0 \mu\text{F}$; $R = 108 \Omega$).

observation that most materials are removed as molten droplets which freeze into hollow spheres. No *white layer* was observed with this material, though a very thin layer of transformed material is just detectable.

Copper

The effects due to sparking on copper can be seen in figure 3, for which the specimen was copper plated; nevertheless the interface can be discerned quite clearly. In the near-surface region (10 μm), there is evidence of a much finer grain structure. This probably demarks the melted zone, which has resolidified



Figure 3. Photomicrograph of copper (copper plated), $\times 625$ ($C = 1.0 \mu\text{F}$; $R = 108 \Omega$).

rapidly due to the high thermal conductivity of the material. Also there is evidence of severe deformation, the slip lines being quite clear; however, neither porosity nor cracks were detected.

Brass tool-electrode

The structure of the brass tool-electrode after EDM-ing a brass workpiece is shown in figure 4. The specimen has been copper plated, and it indicates that

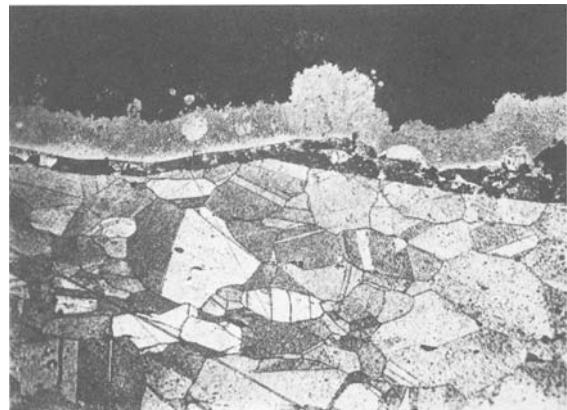


Figure 4. Photomicrograph of brass tool electrode (copper plated) after machining a brass workpiece, $\times 500$.

the tool electrode is also subjected to changes in structure, as is evidenced by the dark etched band. Small cracks can be detected within this band. In the results so far, the conditions used represent approximately the semi-finishing region of EDM.

Scanning-electron microscopy

HW5 tool steel

Figure 5 shows scanning-electron photomicrographs of the tool steel HW5 surface after EDM-ing at a pulse energy of 0.5 Joule, for pulse durations of 250 μsec and 500 μsec . Such conditions are typical of *rough* machining where fast metal removal and low electrode wear are sought with pulse-generator machines. It can be seen that at a magnification of about 200 \times , the surface is highly irregular, and that surface melting has unquestionably taken place. With

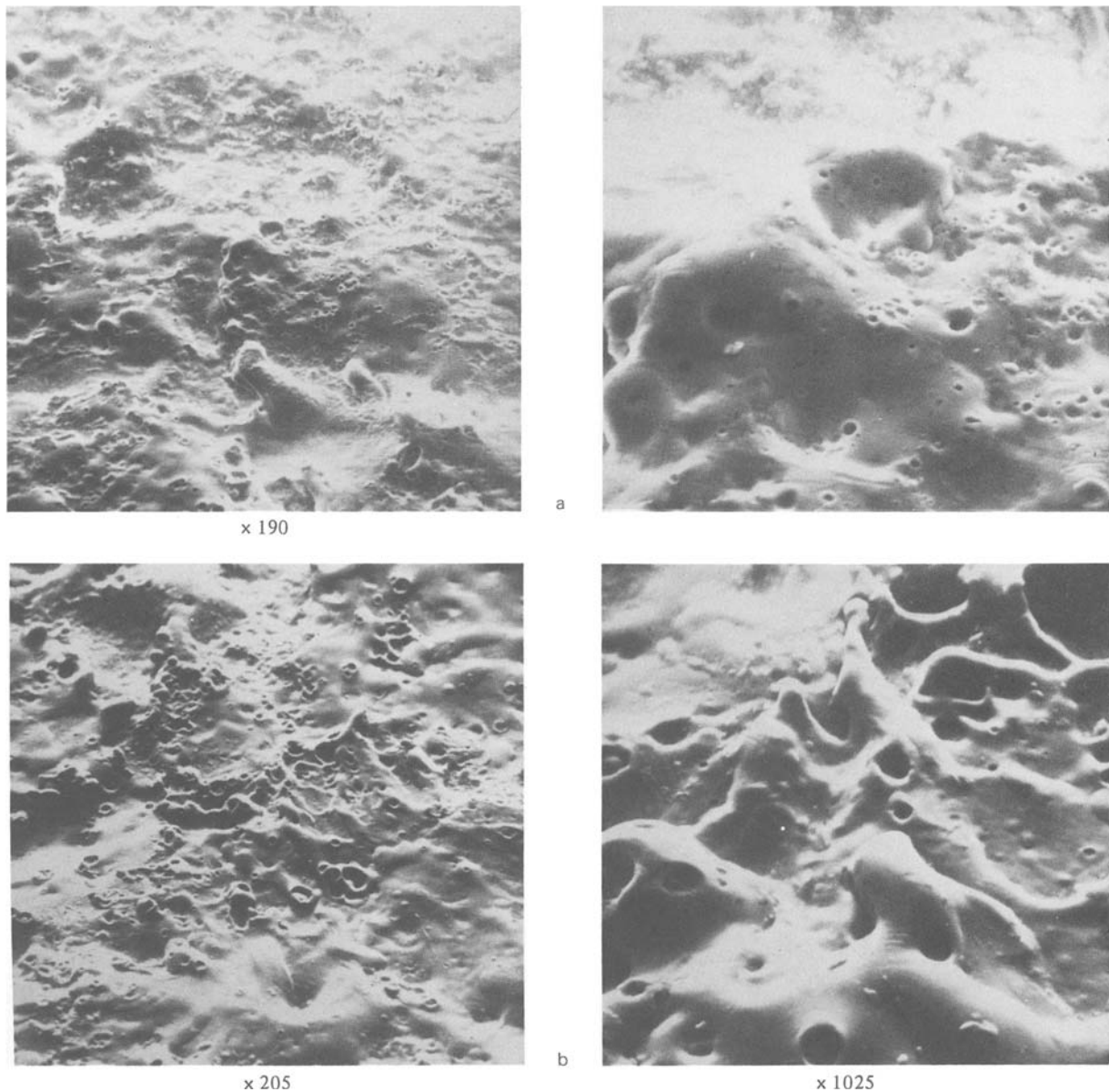


Figure 5. Scanning-electron micrographs of HW5 tool steel machined at 0.5 joule. (a) Pulse duration 250 μsec . (b) Pulse duration 500 μsec .

the longer pulse duration the *pock marks* (as observed by Bucklow and Cole²) are larger, and, as is clearly evident at the higher magnification, the surface shows greater *turbulence*.

One noteworthy feature is that the nature of the surface revealed by the scanning-electron microscope is hardly evident from the transverse sections of Figure 1. However, acknowledging the difference in magnification, it appears that the pock marks of Figure 5 are in fact bubbles of entrapped gas which have reached the surface, and indeed that they are very numerous. Further, the effect of the longer pulse duration is to allow the gas bubbles to grow larger, causing the larger pock marks at the surface and greater *turbulence* beneath it.

Evidently the surface is frozen, virtually instantaneously, when the discharge ceases. However, the shapes of the pock marks, and particularly their rims, are indicative of their sudden and simultaneous rupture, coinciding with the sharp decrease in pressure as

the discharge is cut off. Also the tendency towards greater shape irregularity in the larger pock marks is indicative of the greater sudden volumetric expansion, associated with the larger bubbles. Due mainly to heat conduction losses into the body of the electrodes, solidification follows almost immediately upon cessation electrically of the discharges, causing rapid *freezing* of the surface. Some overlapping of resolidified layers is evident in figure 5(b), presumably due to the effect of partly superimposed adjacent discharges.

The surface of the tool steel machined with the relaxation-circuit generator is shown in figure 6 for two values of capacitance. The current pulse from such a generator is generally something like two orders of magnitude greater in amplitude than from a pulse generator (consequently current density is very high), but is of very brief duration²⁹. The relatively low energy associated with discharges from a 0.25 μF capacitor (of the order of 0.005 joule) is virtually

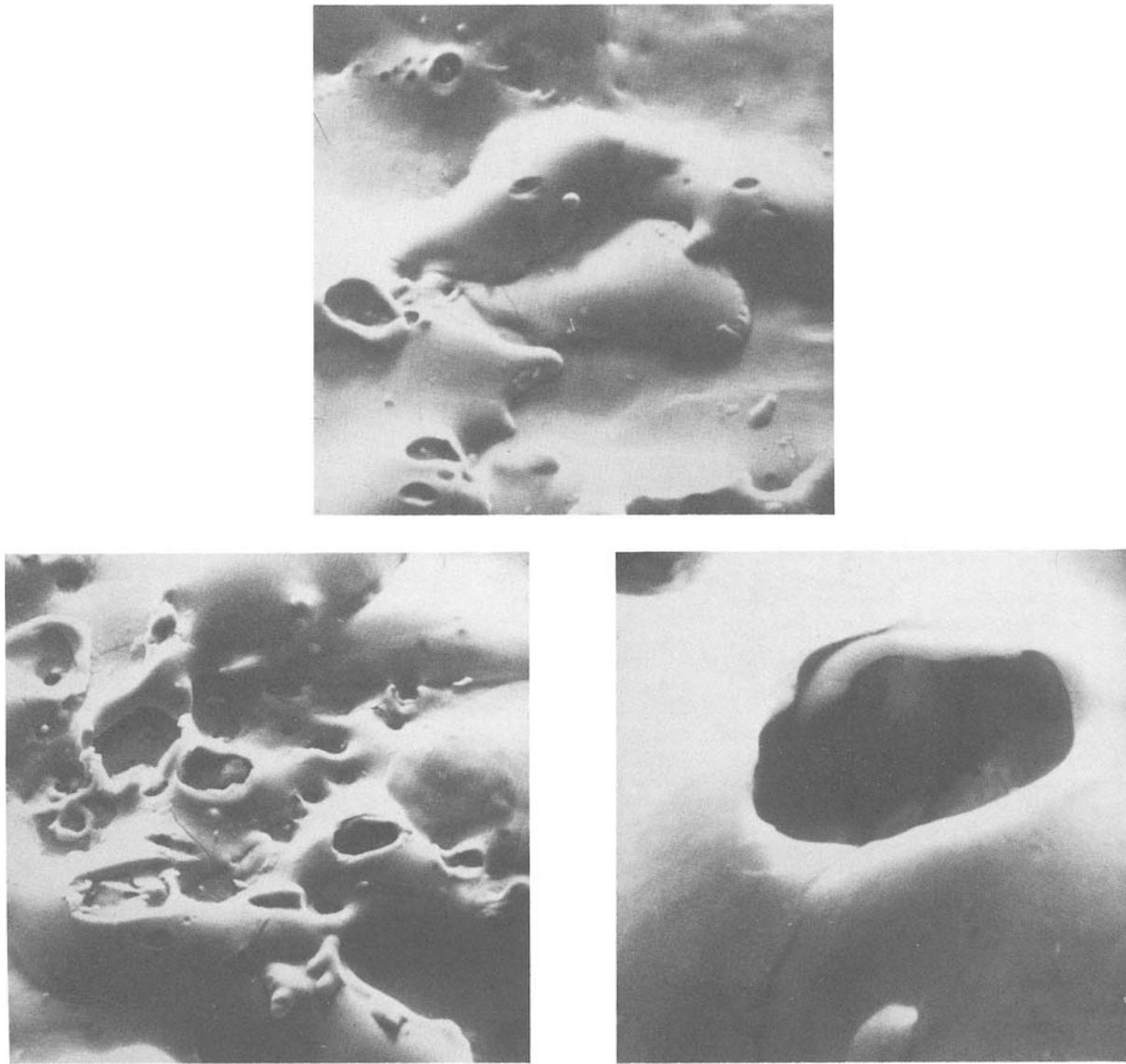


Figure 6. Scanning-electron photomicrographs of HW5 tool steel machined with an RC generator. (a) $C = 0.25 \mu\text{F}$; $R = 235 \Omega$; $\times 900$. (b) $C = 1.0 \mu\text{F}$; $R = 108 \Omega$; $\times 975$. (c) $C = 1.0 \mu\text{F}$; $R = 108 \Omega$; $\times 4900$.

within the *finishing* region of operation of the machine. Figure 6(a) shows comparatively few pock marks, due presumably to both low energy and the lack of time for their formation.

However, there is evidence of cracking, which can be seen in one region of figure 6(a) to follow and link up the general line of surface pock marks, whereas elsewhere cracking occurs without surface pock marks, although the existence of subcutaneous pores cannot be ruled out. This is evident also in the transverse microsections of figure 1, in which cracks can be seen to follow lines of pores, and thus penetrate the re-solidified layer.

It is known from residual stress measurements, made with the same specimens and machining conditions and reported elsewhere¹, that high tensile residual stresses occur near the surface even with low-energy discharges. It is thus very likely that the immediate subsurface layers are in a state of incipient crack formation and propagation, due to the biaxial tensile stresses. Hence cracks, approximately normal to the surface, form fairly readily in the presence of

localized porosity, which is itself a feature of the sub-surface region of this material.

At the higher capacitance of $1.0 \mu\text{F}$, figure 6(b), the extent of gas evolution from the surface is most marked, and a great deal of overlapping and cracking is evident: however this condition would not be used for finishing purposes. The turbulent interior of a burst surface globule (through which runs a crack) can be seen in the highly magnified photomicrograph of figure 6(c). Roughly spherical debris particles can be seen still attached within the globule. Assuming that the globule could have become detached from the surface under certain conditions, the size range of debris particles commonly found in EDM is illustrated thereby; furthermore the debris particles of figure 2 bear comparison with the evidence of figure 6. In spite of the short duration of capacitor discharge at the moderate discharge energy of approximately 0.02 joule, the very high electrode surface temperatures produced³⁰ inevitably cause a high rate of gas and electrode metal vapour evolution, leading to the 'grotesque' appearance of figure 6.

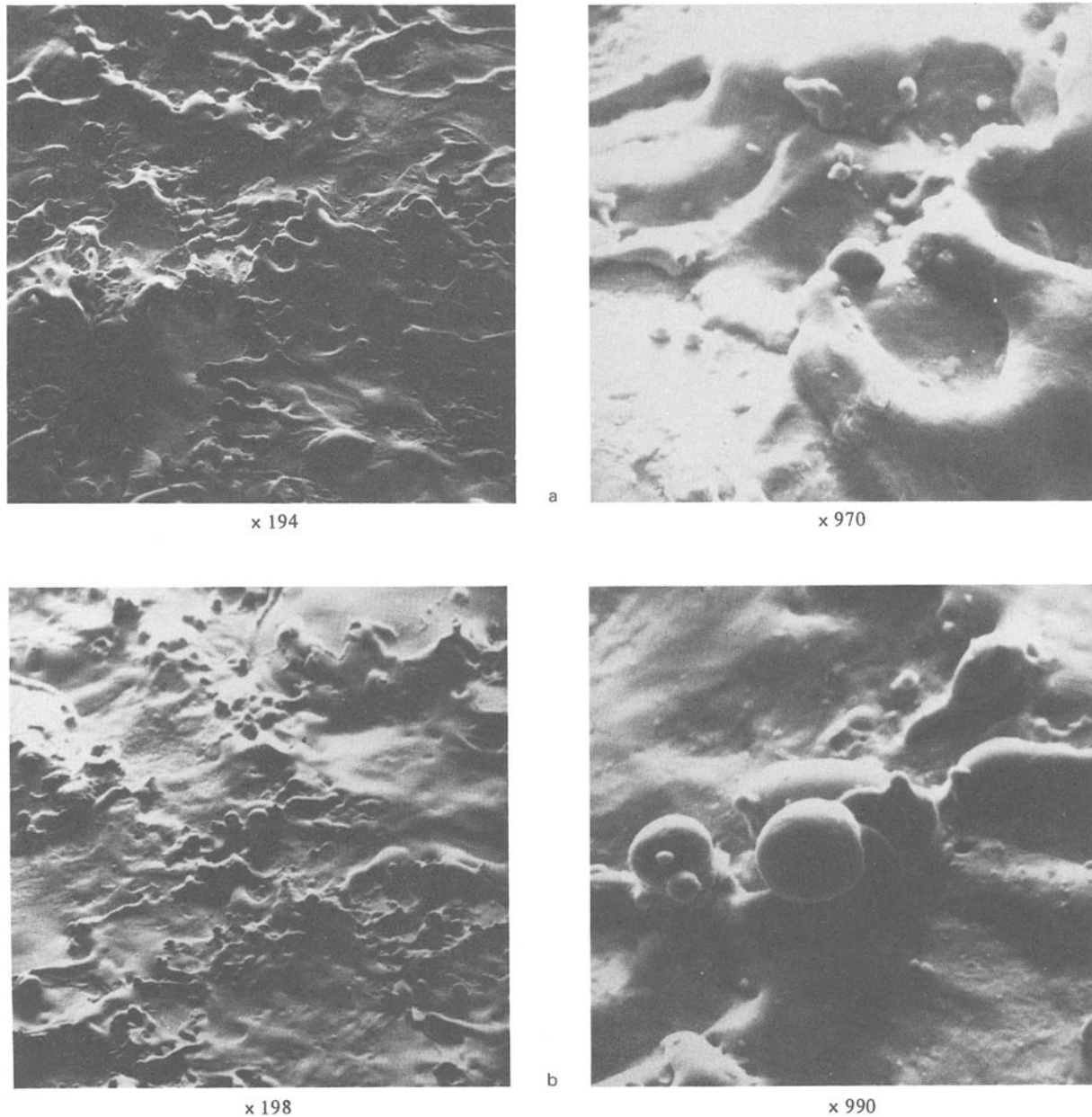


Figure 7. Scanning-electron photomicrographs of EN1 steel machined with an RC generator. (a) $C = 30 \mu\text{F}$; $R = 27 \Omega$. (b) $C = 6 \mu\text{F}$; $R = 43 \Omega$.

EN1 steel

Figure 7 shows the surface of 0.1 per cent C EN1 Steel machined with the relaxation-circuit generator at the higher capacitance values of $30 \mu\text{F}$ and $6 \mu\text{F}$ (approximately 0.1 and 0.2 joules, respectively). Evidence of gas evolution is less marked, but the characteristic 'lunar' surface is similar to that for the tool steel. Spherical debris particles of about $20 \mu\text{m}$ in diameter can be seen loosely attached to the surface.

Copper

The surfaces of copper workpieces machined with 0.5 joule pulse with pulse durations of $250 \mu\text{sec}$ and $500 \mu\text{sec}$ are shown in figure 8. These conditions compare with those used on the tool steel and shown in figure 5. On copper the effect of the longer pulse is similar in producing a more *turbulent* surface, but the copper surface shows markedly less irregularity than

the tool steel (figure 5). Undoubtedly the reason lies in the smaller volume of material melted by discharges on copper due to the high thermal conductivity, and hence rate of heat loss to the substrate (thermal conductivity of copper is approximately eight times that of steel). It is of course for this reason particularly that copper is frequently employed as a tool-electrode material.

This is supported by evidence of the residual stresses in the copper, reported earlier¹, which were confined to a very narrow region near the surface. Stresses virtually at the surface were high, but fell away very rapidly with depth, indicating the shallow penetration of the higher temperatures.

Surface finish

The finish or roughness of a surface is frequently considered to be one of its most important qualities. However, the investigations of the near-surface struc-

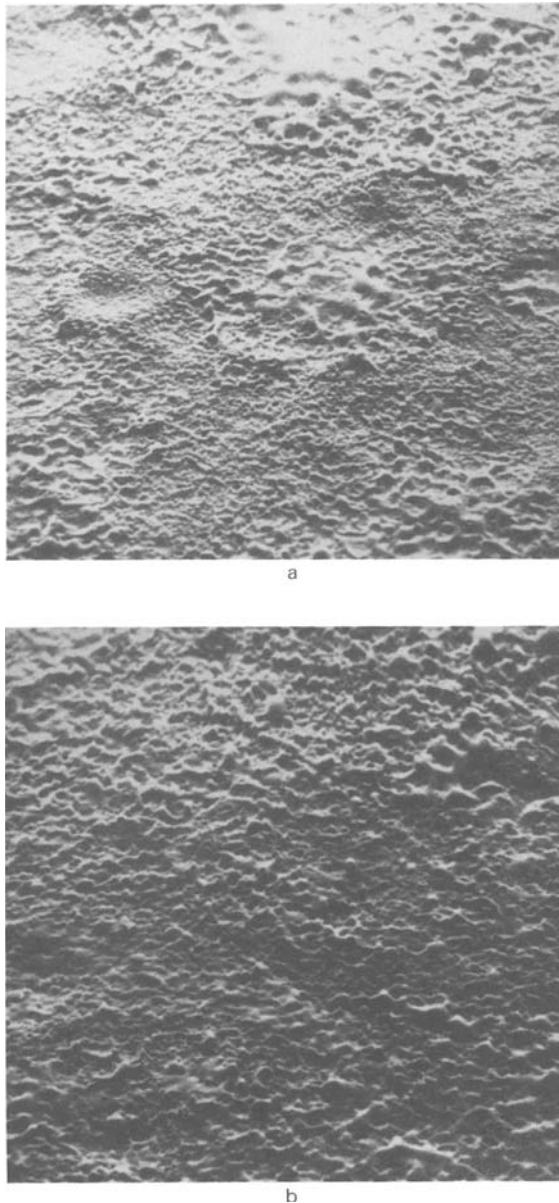


Figure 8. Scanning-electron photomicrographs of copper machined at 0.5 joule, $\times 190$. (a) Pulse duration 250 μsec . (b) Pulse duration 500 μsec .

ture, and of the surface itself with the scanning-electron microscope, have indicated that topographical measurements with stylus instruments can be expected to convey only very limited information on the true nature and 'integrity' of surface produced by electro-discharge machining. Surface roughness measurements were in fact taken, but these served mainly to emphasize such limitations. Only a linear trace is so obtained, and it is apposite to observe also that little evidence is obtained in this case from the transverse microsections observed optically in figures 1 to 3, compared with the almost 'three-dimensional' images obtained with the Stereoscan (figures 5 to 8). Furthermore it is abundantly evident that a standard stylus (and incidentally any stylus which would represent a practical proposition) would be incapable of detecting the finer detail in the topographical structure, such as cracks and re-entrant or steeply-sloped features. Thus stylus measurements must be regarded as providing limited information, even topo-

graphically, in the case of EDM surfaces, although a general indication of average amplitude is obtained which, for practical purposes, is of some value, and the technique is undeniably convenient.

It is unwise therefore, to attempt to read too much into stylus measurements of electro-discharge machined surfaces. Nevertheless surface roughness measurements are frequently made and quoted, and undeniably, some way of specifying simply and conveniently the required final state of the surface, in terms of roughness, is necessary in practice. Some guidance on achieving this aim is available in the German Standard VDI 3400³¹. This is either by 'categorization' using commercially available tactile comparison specimens, or by direct measurement with a stylus instrument. Furthermore a standardized comparative scale for these two methods has been adopted. Nevertheless, although the difficulties inherent in obtaining a meaningful numerical assessment of surface roughness are acknowledged in the Standard, the particular reasons which are quoted in it are quite general, and not a specific statement representative of the particular problems posed by the EDM process.

Quite evidently, the true nature of electro-discharge machined surfaces is complex, and the quantification of such a variety of factors in any single, easily expressible form, is inevitably more remote for the foreseeable future than it is even in the case of surface roughness alone.

CONCLUSION

- (1) The basic nature of electro-discharge machined surfaces has been examined with respect to typical electrode materials, and using two different types of generator. The object has been to elucidate their characteristics rather than to provide data on specific regimes of machining. EDM surfaces are associated with fusion, thermally affected zones, hardness changes, plastic deformation, and residual stresses.
- (2) The surfaces produced are usually very complex in nature, and possess many characteristics peculiar to EDM. This is due not only to the severity and rapidity of local heating and cooling which occurs, but also to the discrete and mainly random attack of the discharges, which to some extent tends to obliterate the evidence of individual discharges.
- (3) The results of the microscopic examinations support the conclusions of the fundamental physical research essentially on the *single discharge* basis, and particularly recent thermal analyses of crater formation, which show that liberation of dissolved gases or vapour is associated with metal removal during the *ejection* stage.
- (4) Conventional topographical surface finish measurements, made with stylus instruments, are of limited use in describing electro-discharge machined surfaces. Features such as cracks, re-entrant craters and 'pock marks' cannot be effectively detected by stylus instruments, but

these may be significant in applications requiring high surface integrity.

- (5) Different EDM generators can vary in their effective electrical characteristics, which in turn may influence the nature of the surface effects produced. Differences are likely to occur in the surfaces produced by pulse generators and those of the relaxation type; the latter being sometimes used for finishing purposes, and in applications such as fine-hole drilling and wire-electrode profiling.
- (6) In the large majority of *general* applications of EDM, however, surface characteristics other than roughness are of no great practical importance. Sub-surface effects are confined to a relatively shallow zone, which, together with roughness amplitude, diminishes as discharge energy is reduced. In applications where this is unacceptable (mainly in the aero-space industry) the surface layer may be subsequently removed, e.g. by chemical milling, or its properties may be enhanced by shot peening. Nevertheless in such cases prior knowledge of the surface characteristics is desirable.

Various observational techniques can be employed, and evidently no short-cut or 'overall' method can safely be relied upon for the *identification* of such effects. Nevertheless, investigation of the *combined resultant effect* can be carried out by fatigue testing, for example, but upon the presupposition that the in-service stress history, particularly in respect of the 'fibre stresses', is known beforehand.

REFERENCES

- J. R. Crookall and B. C. Khor. Residual stresses and surface effects in electro-discharge machining. *Proc. 13th M.T.D.R. Conf. Birmingham* (Sept. 1972) 331-38.
- I. A. Bucklow and M. Cole. Spark machining. *Met. Reviews* (June 1969) 3, No. 6, 103-13.
- J. R. Crookall. A basic analysis of pulse trains in electro-discharge machining. *Int. J. Mach. Tool Design and Research* (1973) 13, 199-213.
- B. N. Zolotykh. Theorie zum Phänomen der funkenerosiven bearbeitung. *Fertigung* (1971) 6, 185-92.
- Frans Van Dijk. Physico-mathematical analysis of the electro discharge machining process. *Doctoral Thesis, Katholieke Universiteit Te Leuven* (1973).
- F. Van Dijk, J. R. Crookall, C. J. Heuvelman and R. Snoeys. Some results of physical research in EDM. Paper to be presented at International Symposium for Electromachining, Bratislava (Sept. 1974).
- R. F. Cooke and J. R. Crookall. An investigation of some statistical aspects of electro-discharge machining. *Int. J. Mach. Tool Design and Research* (1973) 13, No. 4, 271-86.
- A. T. Kravetz. *Machines and Tooling* (1959) 6, 22.
- M. Cole, I. A. Bucklow and C. W. B. Grigson. Technique for the rapid, accurate and strain-free machining of single crystals. *Brit. J. App. Physics* (1961) 12, 296.
- L. S. Palatnik and A. A. Levchenko. *Kristallografiya* (1958) 3, 613.
- L. E. Samuels. *J. Inst. Metals* (1963) 91, 191-92.
- I. Z. Mogilevski and Y. L. Linetskii. Investigation into the physical and chemical changes into the surface layers of steels and alloys after electro-erosion machining, Collection: Problems of Electrical Machining *Ac. Sc. USSR* (1960) 98 (Russian).
- V. P. Aleksandrov and B. N. Zolotykh. On the selection of the optimum conditions in machining nickel-based heat-resistant alloys by the electro-spark method. *Izvestiya AN SSSR, OTN* (1958) 6.
- M. M. Barash. Effect of EDM on the surface properties of tool and die steels. *Metals Eng. Quarterly, ASM* (Nov. 1965) 48-51.
- M. M. Barash and M. G. Sri-Ram. Some properties of the spark-machined heat-treated steels. *Proc. 3rd Int. M.T.D.R. Conf., Univ. of Birmingham* (Sept. 1962) 85-91.
- V. P. Aleksandrov. Residual stresses and the long term and fatigue strengths of heat-resistant materials after electro-spark machining, in book: B. R. Lazarenko (Ed.) *Electrospark Machining of Metals* 3, Consultants Bureau, N.Y.
- G. R. Wilms and J. B. Wade. Some observations with the electric spark machining processes. *Metallurgia* (1956) 54, 263.
- Rudiger and Winkelman, *Metall.* (1958) 5.
- L. R. Walton and C. R. Wendell. EDM metallurgical aspects as related to tool steels, *ASTME, 33rd Annual Meeting, Collected papers* (1965) 65, Book 1, paper 705.
- H. K. Lloyd and R. H. Warren. Metallurgy of Spark Machined Surfaces. *J.I.S.I.* (March 1965) 238-47.
- G. F. Kosolapov and Y. D. Tyapkin. *Metalloved Term Obra. Mat.* (1955) 41, 226-32.
- F. A. Johnson and J. C. Radon. Spark eroded notches in fracture toughness testing of high strength steels. *Fracture Group Research Report* (Jan. 1970) FG 24, Imperial College, London.
- H. Opitz. Metallurgical aspects and surface finish. *Spark Machining Symposium, Metal Treatment and Drop Forging* (June 1960) 237-50.
- Kawai Eiichiro and Mutsuzaki Kensuke. Abnormal structures of the surface of alloys eroded by electrical discharge machining. *Bul. Japan Soc. of Precision Eng.* (1968) 2, No. 4, 8, 296-302.
- L. S. Palatnik et al. *Soviet Physics Crystallography* (1962) 6, 472-75 (Translation from Russian).
- Photomicrography with the Vickers Projection Microscope*. Cooke Troughton and Simms Ltd., York, England.
- Modern Electroplating*. Edited by A. G. Gray. Chapman and Hall Ltd., London.
- B. A. Krasnyuk *Electrospark Machining of Metals* (1963) 3, 110 (Consultants Bureau, N.Y., 1965).
- J. R. Crookall and C. J. Heuvelman. Electro-discharge machining—the state of the art. *Annals of CIRP* (1971) 20/2 *Keynote papers* 113-20.

30. F. Van Dijk and R. Snoeys. Thermo-mathematical analysis of electro-discharge machining operations. *Proc. CNTN Conf., Timisoara, Roumania* (Oct. 1971).
31. *Electroerosive machining: definitions, process, application*. German Standard VDI 3400 Verein Deutscher Ingenieure—Handbuch Betriebs-technik, Register Nr. 2/6 (Sept. 1970).

SOME EFFECTS OF DEBRIS CONCENTRATION ON EROSION AND ELECTRODE WEAR IN ELECTRO-DISCHARGE MACHINING

by

J. R. CROOKALL* and P. W. LEE*

SUMMARY

Electro-discharge machining has been studied with controlled 'pollution' of the dielectric up to very high levels by unfiltered recirculation. The effects of debris concentration on erosion of the electrodes, and the tool/work erosion ratio, varied, depending upon the materials and conditions. In some cases, however, the effect was negligible.

Local distribution of erosion (shape degeneration) differed markedly among the cases investigated. One of these cases involved the use of distilled water as a machining fluid. The underlying reasons for this behaviour are considered to be thermal in origin.

INTRODUCTION

Little previous work, concerned directly with the effect of debris concentration (such as erosion ratio, metal removal rate and surface roughness) within the gap upon machining characteristics in electro-discharge machining, appears to have been published. However, previous work^{1,2} on the electrode shape degeneration, due to wear, has shown that the erosion of corners is different in degree to that found in plane regions of the gap. Thus the normal test case for wear involving a *plane orthogonal gap* is only indirectly relevant to the erosion of corners or regions of high curvature. The reasons are two-fold:

- (a) a *geometrical* effect—termed 'relative duty'—whereby an external corner of the tool, being of smaller radius (and hence surface area), suffers a greater rate of recession of the surface at the corner itself than at plane regions of the surface, other factors equal.
- (b) the *local erosion ratio* effect—whereby the relative erosion (i.e. volume loss) is higher locally for an external corner than for plane regions of the gap, or indeed from the simple standard 'orthogonal' test.

For these two reasons, electrode corners degenerate to some extent even with very low erosion ratios, measured in the normal way.

If the erosion ratio is significantly influenced by debris concentration within the gap, then it is possible that the restriction on flow through a corner causes, by this mechanism, an increase in erosion

ratio at the corner itself. It is of interest to know also whether the build up of debris within the gap itself affects the metal removal rate and erosion ratio.

The experimental problems, in investigating such effects, are primarily the control of the debris concentration, for test purposes, without influencing other parameters of the process other than as a natural result of deliberately changing the gap conditions, and also the determination of the level of debris 'pollution' of the dielectric achieved.

OTHER RELEVANT WORK

Naturally, the changes to the pure (or at least filtered) dielectric, brought about by discharging within the gap, are directly relevant. Vajta et al.³ observed the changes in density, viscosity, flash point and refractive index of kerosene, after use as a dielectric fluid. Some of the kerosene was decomposed into carbon, gas, and a mixture of liquid hydrocarbons; the proportions were dependent upon the discharge energy level. The specific gas evolution and carbon formation (i.e. regarding unit weight loss of workpiece) decreased as the discharge energy level was raised.

Hockenbury⁴ observed the movement of gas and solid particles during the discharge cycle. He claimed that the dispersed and metallic particles participated in the mechanism of breakdown of the gap. Partial bridging of the gap by conducting particles aided the breaking down, which therefore sometimes took place at points other than at the closest approach of

* Imperial College of Science and Technology

the electrodes themselves. Furthermore the high electrical field, formed at protruding irregularities on the electrodes, tended to align debris particles at these points. He also recorded a generally observed fact that some contamination of the gap leads to more regular breakdown characteristics. This results in greater stability of EDM operation, a larger gap and sometimes a decreased machining rate.

De Bruyn⁵ has also demonstrated that the gap at the entry point of the dielectric fluid is less than that at the exit, due to contamination with debris. He found that this condition could be changed by employing alternating periods of discharging and flushing.

Mironoff⁶ reported a decrease in metal removal rate with increasing debris concentration, and also that temperature affects erosion rate with some dielectric fluids, whereas petroleum was not affected.

Kurafuji and Suda⁷ investigated the effects of adding conducting powders (principally of iron, copper or graphite) to a kerosene dielectric. It is noted from their results that the charging resistance (a relaxation-circuit generator was used) varied quite widely from one test to another. This will of course by itself influence the results obtained. Nevertheless they quote an increase in erosion rate due to the added powders; the effect being greater for powders of fine size and low density. An increase in gap size, higher discharge frequency and more stable machining were observed.

In terms of metal removal rate the evidence appears to be somewhat contradictory, although conclusions on gap and stability are widely corroborated. It must be acknowledged that contradictory evidence is not uncommon in the results of EDM research. This is not surprising as generators, servo-systems, materials and machining conditions vary widely between experiments, and yet, they are known to influence results to an extent which makes direct comparison somewhat hazardous. In particular, the introduction of 'alien' powders can be quite different from the effects of natural debris comprising solid, gaseous and soluble contaminants which are actively formed.

During machining the debris which is formed must be removed from the gap. This usually occurs partly through the 'explosive' agitation of the discharges, and partly through induced dielectric flow. Indeed with complex or deep cavities, debris clearance becomes a major problem, and various subterfuges such as electrode vibration and periodic lift are used. These inevitably reduce the time during which correct machining is taking place, and can cause an increase in electrode wear. Obviously, part of the debris comprises the material removed from the electrodes, but the remainder originates from the products of breakdown of the dielectric. Hydrocarbon dielectrics all break down in an approximately similar way, but distilled water is quite different in many ways which affect its behaviour as an EDM fluid. Some of these are turned to advantage for fine work with EDM⁸. Carbon is, of course, totally absent in distilled water, and, in addition to resolidified electrode metal, some metallic oxides and hydrogen and oxygen may be produced; overall, less solid material is produced.

CONTROLLING AND MEASURING DEBRIS CONCENTRATION

For experimental purposes it is necessary to vary the concentration of debris in the gap, keeping other factors constant. Also it is implicit here that natural EDM debris should be used, rather than 'alien' substances. Debris may be added via the dielectric feed, and for various reasons the method adopted was that of recirculating the dielectric continuously without filtration. Thus debris concentration level increases continuously throughout the period of the test. Hence by monitoring this level at regular intervals, together with the erosion of work and tool electrodes, and machining conditions in terms of recorded electrical behaviour, one test would serve to provide results over a wide range of debris concentration.

Evidently only the solid debris could be conveniently recirculated in this way, and even so the differential rates of settling of constituents of different densities inevitably caused some 'selection' of the material which was effectively reintroduced into the gap. Such effects could be reduced by continuous agitation to retain suspension of as much solid matter as possible. Such problems are likely with any method of artificially introducing debris, and of the various schemes considered, the adopted method appeared to provide the most practical combination of circumstances.

In the measurement of debris concentration a number of possible methods were considered. In view of the importance of conducting particles in influencing gap conditions, methods of estimating the concentration of these were considered. These included the possibility of estimating separately the metallic and carbon contents. Dissolution of the metal particles in acid, and determination of the residual carbon by weighing, suffered from inaccuracy, due particularly to the low density of carbon. Methods involving separation by combustion were deemed to be no better for the same reasons. Indeed a volumetric determination appears to be more practical than weighing methods.

The technique of *nephelometry* was tried. It estimates solid material in suspension from the scattering of a light beam passed through it. For EDM debris, this method suffered from the relatively rapid settling of larger particles, and also in that the suspension became largely opaque at moderate levels of concentration.

The possibility of separating carbon and metal particles was abandoned, and the method adopted for determination was by centrifuging in a specially shaped tube (fig. 1(a)), and estimating the proportion of solid debris by volume. To regularize the packing density, a standard duration for centrifuging was used.

APPARATUS

The electro-discharge machine was a Servomet* employing a relaxation-circuit generator and electro-mechanical servo. However a special tank was used

* Metals Research Ltd.

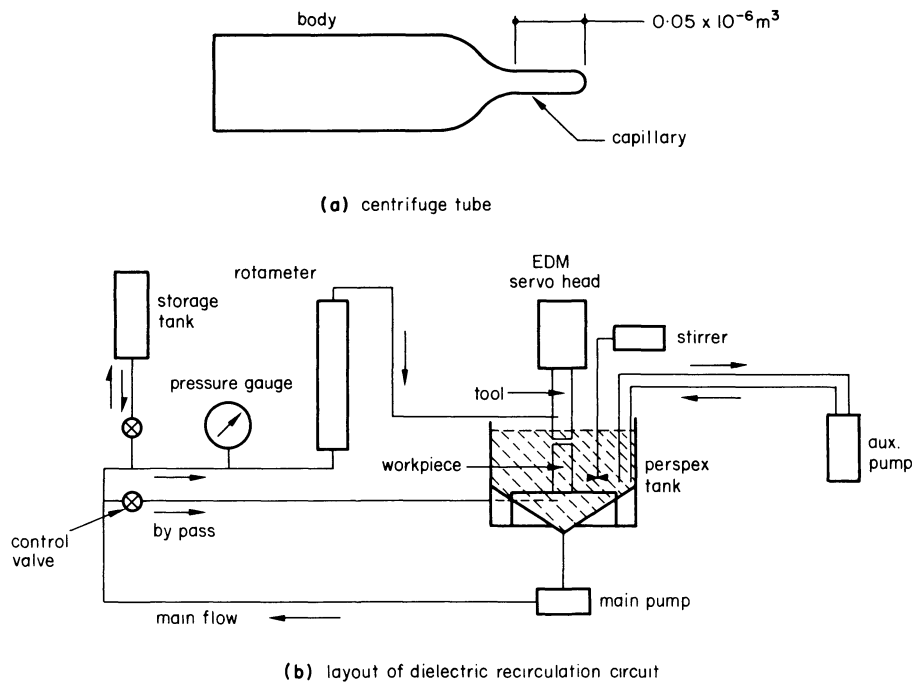


Figure 1. Diagrammatic representation of apparatus. (a) Centrifuge Tube. (b) Layout of dielectric recirculation circuit.

(figure 1(b)) which allowed the debris, formed during machining, to be collected and recirculated. The dielectric in this tank was continuously agitated, and the pressure and volume flow rate of the 'polluted' fluid fed to the tool electrode was measured and controlled. The electrodes, dielectric fluid and energy settings of the machine used during the tests are given in Table 1.

TABLE 1 Conditions and materials used in the tests

Figure number	Electrode Materials		Dielectric Fluid	Parameters of RC circuit	
	Workpiece	Tool		C (μF)	R (ohms)
3	Mild Steel	Brass	Kerosene	30	27
4	Die Steel*	Brass	Kerosene	30	27
5	Mild Steel	Brass	Kerosene	6	43
6	Mild Steel	Brass	Distilled Water	30	27

* Darwin HW5 die steel, of the following nominal composition: C = 0.3%, W = 1.0%, Cr = 5.0%, Mo = 2.0%, V = 0.25%.

Cylindrical electrodes, nominally 25 mm in diameter, were used. The brass tool electrodes had a central flushing hole of 1.6 mm diameter.

RESULTS AND DISCUSSION

The steady increase in debris concentration level with elapsed machining time is shown in figure 2 for the three tests made in kerosene dielectric. The lower of the three curves indicates the lower rate of debris production at the smaller energy level, as expected.

For the same energy level, the other two curves indicate a small difference in the rate of production of debris for the two different steel workpieces employed.

The test with the mild steel workpiece (brass tool electrodes were used throughout) was conducted over a machining period of three hours. Tool and work erosion rates, and also the tool/work erosion ratio are shown plotted in figure 3, against debris concentration level. It is evident that the effect of debris concentration upon the erosion characteristics is very slight indeed, particularly when it is remembered that the concentration reached in this experiment must be considerably greater than would normally be encountered in practice, with the possible exception of shapes which were particularly difficult to flush. Short-circuit conditions which were observed to occur periodically at the start of machining, became less frequent as the debris concentration level

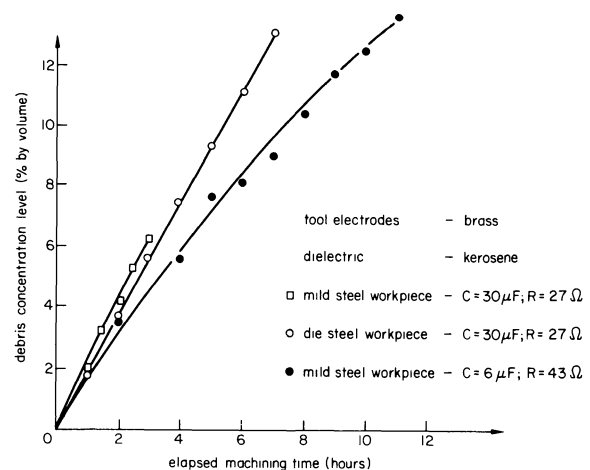


Figure 2. Growth of debris concentration with elapsed machining time.

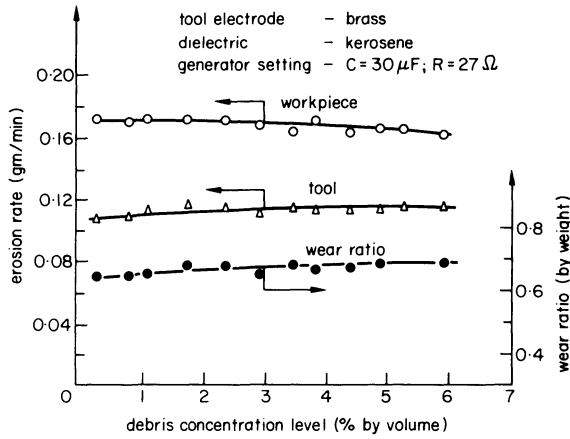


Figure 3. Effect of debris concentration on erosion and wear for mild steel.

increased. The breakdown voltage, measured from the recorded electrical waveform, decreased fairly steadily with the increase in debris concentration level. This is indicative of the effect of conducting particles in aiding the breakdown of the gap.

The results for the die steel (figure 4) show a definite difference. A slight though definite decrease in workpiece erosion rate is evident, together with a corresponding increase in tool erosion. Thus the wear ratio increases from about 0.6 to 0.95 over the seven-hour period of machining. The stability of machining, and the observed breakdown voltage characteristics, were fairly similar to those of the previous test. Thus it appears that a relatively small change in composition of the workpiece material can affect the erosion behaviour of both electrodes, and that erosion ratio may or may not be dependent upon debris concentration level.

Figure 5 shows that for the mild steel workpiece at a lower discharge energy setting, apart from the obvious fact that the erosion rates are lower, erosion ratio is again affected by debris concentration level, and also that the erosion ratio at all levels is higher (a well-known result). Short-circuit conditions were less frequent and virtually absent towards the end of the test (11 hours of machining). Again, breakdown voltage decreased over this period, and the discharge repetition frequency increased.

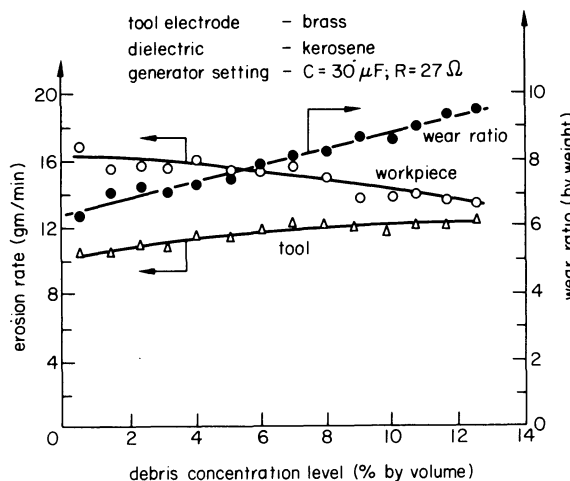


Figure 4. Effect of debris concentration on erosion and wear for die steel.

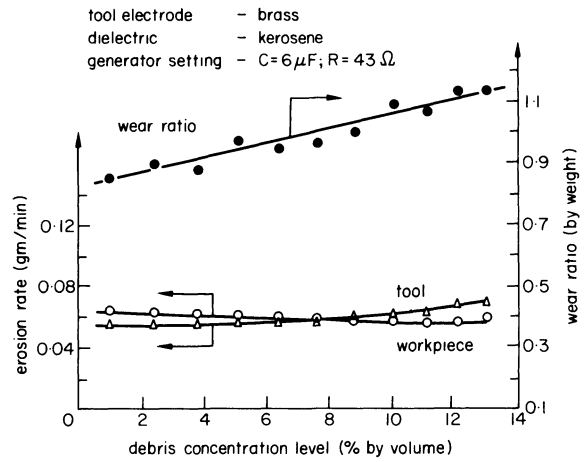


Figure 5. Effect of debris concentration on erosion and wear for mild steel.

One fact, noted from all of the results presented so far, is that the erosion ratio is very sensitive to the trends of erosion of the individual electrodes. Thus, if the erosion ratio is left out of consideration, the difference in trends for the electrodes individually could be deemed to be less remarkable.

It may appear somewhat surprising that less frequent shorting and more stable machining conditions were found to occur at the very high debris concentration levels. This is undoubtedly due to the fairly small gap area (490 mm²), and the straightforward flushing of the axisymmetrical, plane-orthogonal gap used. Hence average debris concentration built up *slowly* over a period of some hours. Conversely, instability and shorting is likely to occur with cases which are more difficult to flush due to the effect of *rapid* changes in local debris concentration.

The result for distilled water as a machining fluid is shown in figure 6, plotted against elapsed machining time. Comparison with figure 3, for which the electrical conditions of machining are the same, indicates some similarity, but with the obvious difference that the erosion ratio is approximately 30

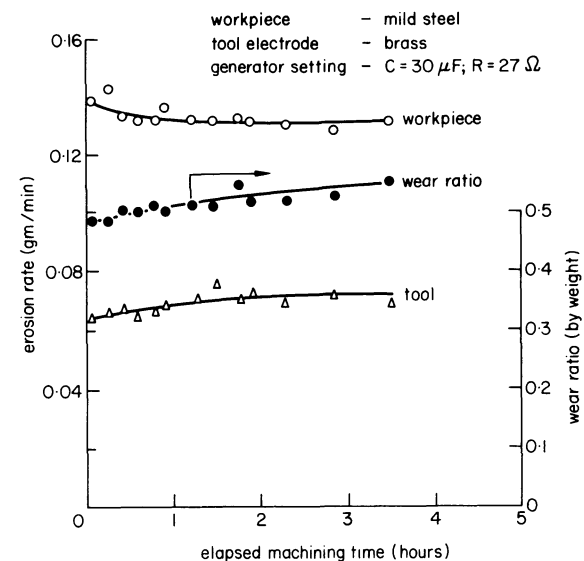


Figure 6. Effect of machining time on erosion and wear with distilled water.

per cent lower. The specific heat of water is about 2.1 times that of kerosene, and is thus considered to be a factor in reducing the rate of erosion on both electrodes, but particularly on the tool.

The heat-withdrawing effect of the fluid is more dramatically illustrated in terms of the *distribution* of erosion material. The final shapes of the tool electrodes used with mild steel as a workpiece material are shown in figure 7. Further work is necessary before a conclusive statement can be made on the reasons for generation of these particular shapes. However, comparing shapes achieved under different energy levels in kerosene (figures 7(a) and (b)) it is immediately evident that shape distortion is greater at higher energy, in spite of the fact that mean erosion ratio is lower (the actual amount of erosion is very similar).

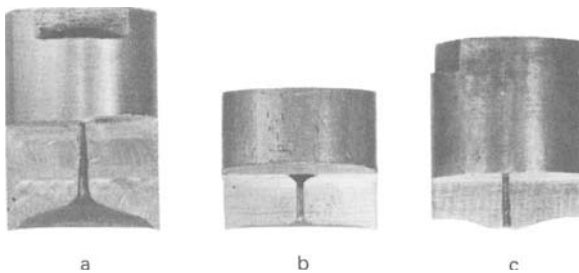


Figure 7. Sectioned brass electrodes showing shape degeneration under different conditions. (a) High energy with kerosene dielectric, $C = 30 \mu\text{F}$; $R = 27\Omega$. (b) Lower energy with kerosene dielectric, $C = 6 \mu\text{F}$; $R = 43\Omega$. (c) High energy with distilled water, $C = 30 \mu\text{F}$; $R = 27\Omega$.

The rate of production of debris must be higher at the higher energy level, and for this condition (figure 3) erosion ratio is virtually unaffected by debris concentration. Hence any variation radially in debris concentration within the gap is unlikely to be the major cause of the shape generated in figure 7(a). It is concluded therefore that the local variation in relative erosion has arisen because of variations due to electrode geometry in the local thermal capacity, at different points on the eroding surface of the electrode.

The lower instantaneous heat release of the lower energy condition (figure 7(b)) would imply that relative erosion characteristics would be less sensitive to local thermal capacity, and it is evident from the final electrode shape that a lower local variation in erosion has resulted. Turning to a case of distilled water (figure 7(c)) it is evident that the rapid quenching ability has modified the shape quite considerably. Indeed the thermally rather exposed region around the dielectric flushing hole, which has been eroded so rapidly in figure 7(a), shows less wear than elsewhere with distilled water. This is due to the supply of cold dielectric with a high specific heat at this point. Indeed the region at the exit from the gap illustrates what is presumed to represent the effect of a large body of dielectric fluid in drawing heat quickly from peripheral regions of the electrode. In the case of distilled water, all regions are less influenced by local thermal isolation as the heat capacity of the fluid actually within the gap is much higher anyway.

Returning to the increased erosion ratio which was found to occur at corners², it now seems most probable that the major contributory factor is the thermal isolation of each point on the electrode surface which is in erosive engagement. In other words when a discharge strikes at any general local point, the solid angle subtended at that point into which heat can be conducted into the body of the electrode determines the penetration of the melting isotherm, and hence the amount of discrete erosion at that point (for either electrode). However this condition is modified by the heat capacity of the dielectric locally. If large it can take an increased proportion of the heat generated through the medium of the fluid. The result is also a reduction in penetration of isotherms into both electrodes (depending upon their relative thermal conductivity), the cumulative effect of which is to alter the erosion rates, and hence the wear ratio.

CONCLUSIONS

- (1) The effect of debris concentration upon erosion characteristics in EDM has been observed experimentally by varying the general level of solid debris within the gap by recirculation of unfiltered dielectric. Erosion measurements are made at regular intervals as debris concentration rises.
- (2) It is considered that the debris concentration levels reached during the tests, which were of some hours duration, are considerably higher than occurs locally, in even 'difficult' practical cases. Bearing in mind that the 'plane orthogonal gap' is more easily cleared by direct servo action than 'convoluted' gaps, machining stability improved with increasing debris concentration. Breakdown voltage decreased, and the effect on workpiece removal rate varied from insignificant, to a progressive slight decrease. Indeed this could be attributed to the changing discharge frequency/energy conditions (an RC circuit was used) rather than directly to the presence of debris itself.
- (3) The effect of debris concentration on tool/work erosion ratio varied from negligible to a progressive increase of up to 50 per cent for a debris concentration level of about 13 per cent by volume (a very high level). Die steel showed this trend, whereas for mild steel the effect was negligible, hence electrode materials are significant in this respect.
- (4) Distilled water, as a dielectric, produced lower erosion rates and ratios. Shapes generated on the electrodes in each case were different.
- (5) All the evidence points to the importance of the 'thermal capacity' which lies behind all points on an electrode surface which are in erosive engagement. The 'thermal capacity' of the dielectric fluid, in altering the proportion of heat transmitted through the electrode, can thus influence erosion and erosion ratio both locally near feed holes and free surfaces, and generally as an average effect.

REFERENCES

1. J. R. Crookall and R. J. Fereday. An experimental determination of the degeneration of tool electrode shape in electro-discharge machining. *Part 1, Microtecnic, No. 2* (March 1973) **XXVII**, 97-99, *Part 2, Microtecnic, No. 4* (May 1973) **XXVII**, 197-200.
2. J. R. Crookall and A. J. R. Moncrieff. A theory and evaluation of tool electrode shape degeneration in electro-discharge machining. *I. Mech. E. Proc.* (1973) **187**, 6-73.
3. L. Vajta, P. Siklos and L. Nagy. Study on the decomposition of kerosene used as a dielectric medium in spark machining. *Periodica Polytechnica, Budapest, Hungary.* (1967) **11**, 3-4.
4. T. O. Hockenberry. The role of the dielectric fluid in electrical discharge machining I. *Society of Automobile Engineering—Paper 680635* (7-11 October 1968).
5. H. E. De Bruyn. Some aspects of the influence of gap flushing on the accuracy in finishing by spark erosion. *Annals of C.I.R.P.* **XVIII**, 147-151.
6. N. Mironoff. Electro-erosive metal working, its physical fundamentals and industrial applications. *Microtecnic* (1965) **19**.
7. H. Kurafuji and K. Suda. Study on electrical discharge machining I. *Journal of the Faculty of Engineering, University of Tokyo* No. 1 (1965) **XXVII** 1-18.
8. J. R. Crookall and C. J. Heuvelman. Electro-discharge machining—the state of the art. *Annals of C.I.R.P.* (1971) **20/2**.

INVESTIGATION OF THE MODULATION OF INTERFACE WAVES IN EXPLOSIVE WELDING

by

S. R. REID* and D. J. LANGDALE*

SUMMARY

A characteristic feature of explosive welding is the wave-like geometry of the interface between the welded components. In this paper the results of an examination of the variation in the amplitude and wavelength of the waves along the weld interface are presented for a number of steel-to-steel welds produced using Metabel explosive. It is concluded that the variation possesses a certain degree of regularity. The modulation of the waves is discussed and it is concluded that this phenomenon warrants further study.

NOTATION

F	boundary of region in which fluid-like behaviour occurs
S	stagnation point (see figure 3)
$U_d, V_f, V_j, V_p, V_s, V_w$	velocities relating to the kinematics of the flyer plate collapse defined in the Appendix
a	amplitude of interface waves
\bar{a}	mean amplitude of interface waves
c_0	elastic wave speed
f	length of region within which fluid-like behaviour occurs
t_b	thickness of rubber buffer
t_e	thickness of explosive
t_f	thickness of flyer plate
α	set-up angle
β	impact angle
δ	dynamic bend angle
λ	wavelength of interface waves
$\bar{\lambda}$	mean wavelength of interface waves

INTRODUCTION

Explosive welding is now accepted as a reliable method of producing high quality welds, particularly in cladding processes¹ and in the welding of tubes². Since its accidental discovery³ during the development of certain explosive-forming processes some thirty years ago, it has attracted the interest of a large number of researchers and many aspects of the process have been examined. Details of the process and some of its practical applications can be found in the review articles by Crossland and Williams³ and Ezra⁴.

The strength of explosively produced welds stems from the nature of the interface between the two welded components. This interface frequently has a wave-like appearance (see figure 1) the mean dimensions of which change according to the metals joined,

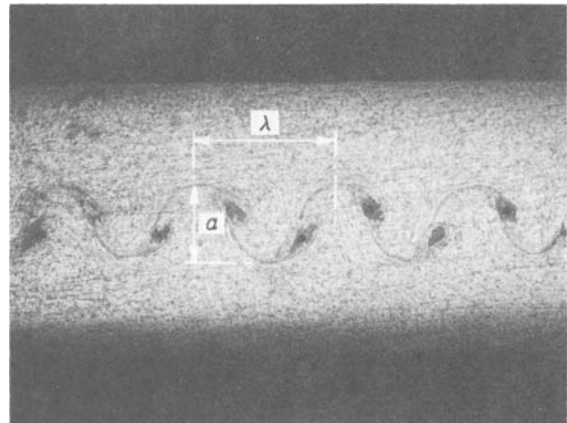


Figure 1. Photomicrograph of interface waves between two explosively welded steel plates showing definitions of amplitude and wavelength.

the explosive used, the thickness of the components and the geometry of the configuration in which the components are set prior to welding. Details of the inclined flyer plate set-up, which was used throughout the series of tests described later, are shown in figure 2.

Along the interface there is evidence of high pressures having been produced and often regions

* Department of Mechanical Engineering, University of Manchester Institute of Science and Technology

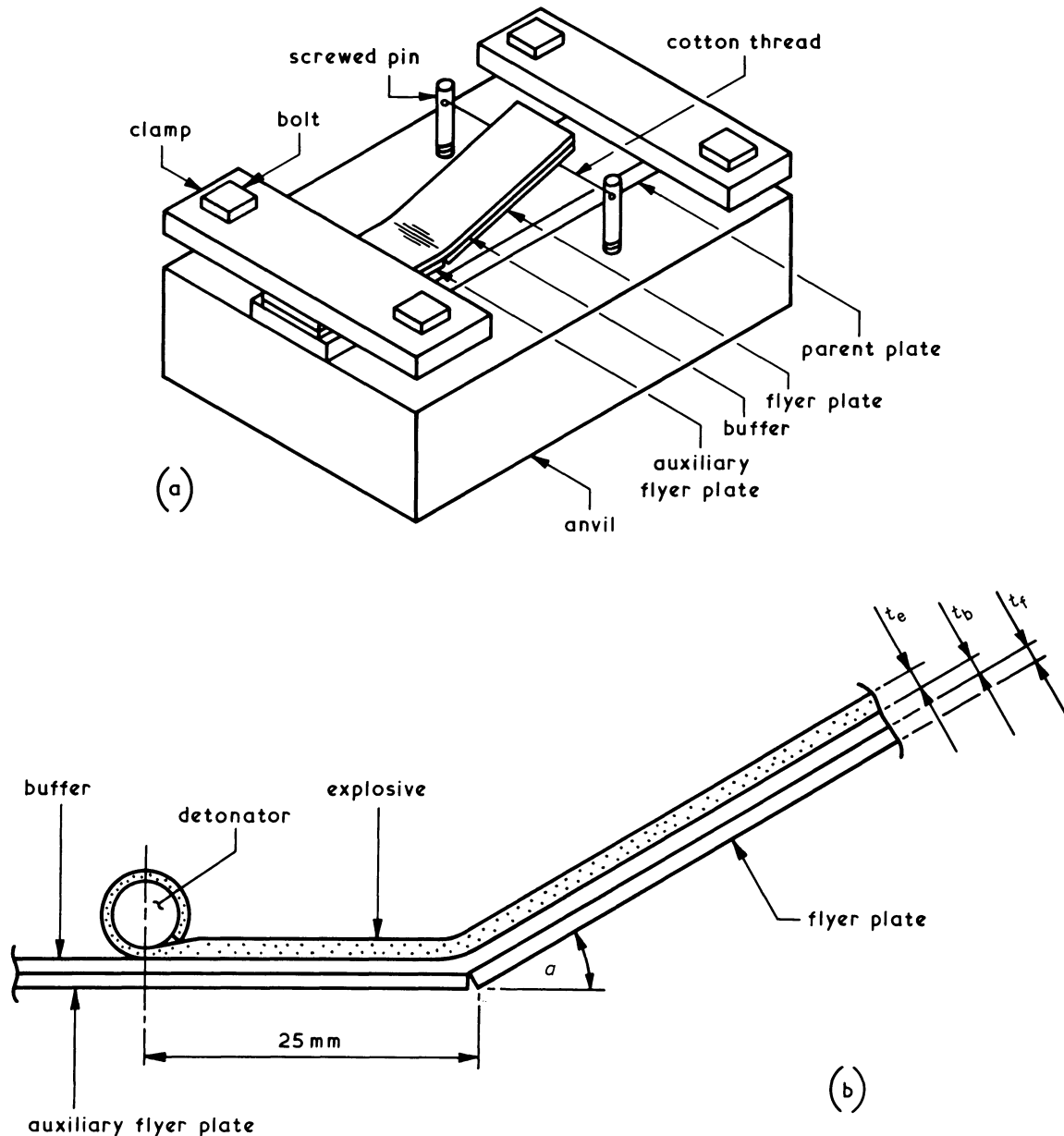


Figure 2. (a) Inclined flyer plate arrangement. (b) Details of flyer plate/rubber buffer/explosive combination.

exist within which fusion and mixing of the components has occurred. Whilst the appearance of these features may encourage the view that an explosive weld is a form of either pressure weld or fusion weld, such a conclusion would be false. It is possible with explosive welding to join metals having very different melting points (for example lead and steel) which cannot therefore be welded by the usual heat welding methods. Similarly, metals with yield stresses sufficiently different to prevent their being pressure welded can often be joined explosively. Clearly the mechanisms responsible for creating an explosive weld are quite different from those involved in the more conventional welding processes.

In order to control the quality of explosive welds it is necessary to be able to control the nature of the interface waves. A certain amount of waviness is desirable as it provides an increased contact area between the components as well as providing a degree of mechanical interlocking. However, if one is not

careful the interface can become a continuous fusion layer between the two components and will have poor fatigue properties.

In view of the variation in the quality of the weld with the interface geometry, several attempts have been made to construct theories⁵ explaining interface wave formation. To be of use in practice these theories should ultimately enable the set-up parameters to be determined so as to give a selected degree of waviness at the interface. In order to test them and to gain a better understanding of other features of the interface waves, the authors have conducted several series of tests in which steel-to-steel welds were produced using Metabel explosive⁶. After welding the components, the weld interfaces were examined under a low-power travelling microscope in order to measure the dimensions of the waves. Apart from confirming many of the results which have appeared in the literature regarding the variation in the size of the waves as the set-up parameters (see

figure 2(b)) are varied, the tests also revealed another interesting feature of the waves which has not been discussed in detail before in the literature. The size of the waves and in particular their amplitude varies significantly along the interface. On the basis of the results presented it is concluded that this variation is not entirely random. In order to infer that the amplitude is affected by some regular phenomenon, the interface waves are said to be modulated.

After briefly describing the experimental procedure and giving representative examples of the results obtained, they are discussed and a qualitative explanation for the appearance of the waves is given on the basis of a mechanism for interface wave formation recently described in the literature⁵.

EXPERIMENTS: APPARATUS AND PROCEDURE, RANGE OF TESTS AND THEIR PURPOSE

Experimental procedure and flyer plate kinematics

The set-up configuration used for all the tests was that shown in figure 2(a)—the inclined flyer plate arrangement. Such an arrangement is essential in order to produce welds using Metabel because of its high detonation velocity³. When the explosive is detonated the flyer plate collapses at high speed on to the parent plate. When the conditions are such that the components become welded, it can be shown⁷ that the pressure produced in the vicinity of the contact region considerably exceeds the yield stresses of the two plates. It is usually assumed that the components behave essentially like impinging fluid jets^{3,7} and that the flow has the appearance shown in figure 3 at an instant part way through the welding process.

From figure 3 it can be seen that the flyer plate is assumed to be bent by the explosive pressure through an angle δ , the dynamic bend angle, so that the actual angle of impact is not α , the set-up angle, but $\beta = \alpha + \delta$. In the Appendix the velocities indicated in figure 3 are defined and it is shown how they depend upon U_D , the detonation speed of the explosive, and upon α and δ . The angle δ varies according to the thicknesses of the flyer plate, the rubber buffer and

the explosive layer, whereas U_D is a property of the explosive alone. These two parameters must be measured in tests separate from those described in this paper. The authors used the pin contactor method^{6,8,9} to measure V_w for two different values of α . The method of calculating U_D and δ from these two values is indicated in the Appendix.

After the two plates had been welded they were sectioned along their longitudinal centre lines and the section was polished. The specimens were then etched with 2 per cent Nital, washed and dried, after which the profile of the interface between the two plates could be clearly seen even with the naked eye. The wavelengths and amplitudes of the waves (defined in figure 1) were then measured using a travelling microscope.

Purpose and range of tests

In order to prevent surface damage to the flyer plate by the explosive and also to vary the flyer plate velocity, a rubber buffer is inserted between the flyer plate and the sheet of explosive. Apart from the collection of data on the interface waves, one of the other aims of the investigation described in reference 6 was to quantify the effects of varying the thickness of the buffer, t_b . In the three series of tests discussed in this paper the buffer thickness was the only dimension that varied from one series to the next. The details of the components used in the tests are as follows:

- (1) The width and length of the flyer plate/buffer/explosive combination (see figure 2(b)) were 25.4 mm and 150 mm respectively.
- (2) The thicknesses of the components were:
flyer plate, $t_f = 1.25$ mm
buffer, $t_b = 1.59$ mm (series A), 3.18 mm (series B), 4.77 mm (series C)
Metabel explosive, $t_e = 4.75$ mm
- (3) Mild steel flyer plates and parent plates (thickness 12.7 mm) were used throughout. The set-up angle, α , was 7° throughout, this being chosen after performing a set of preliminary tests which determined the ranges of

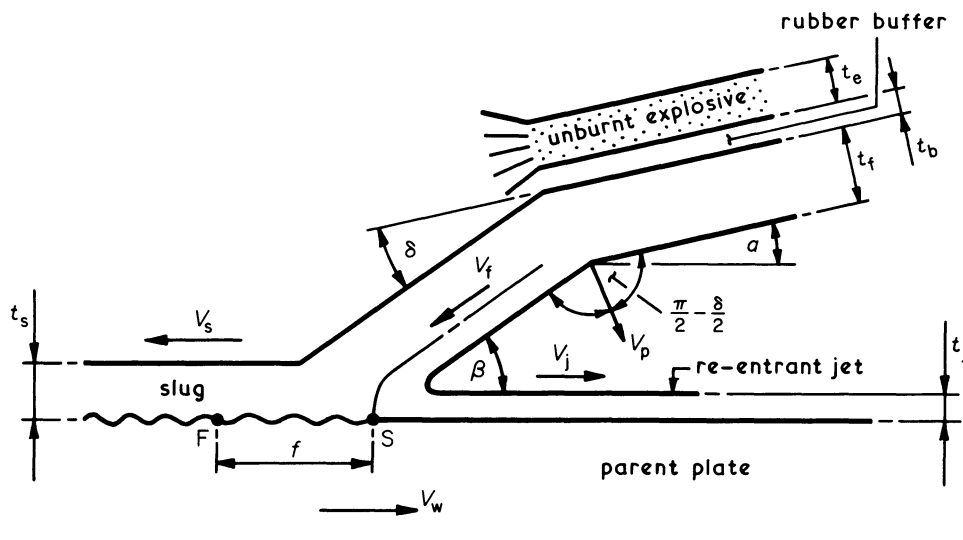


Figure 3. Assumed flow configuration during cladding process.

set-up angles within which waves were produced for each value of t_b .

- (4) The explosive was detonated with an ICI no. 6 detonator.
- (5) In each series (that is, for each value of t_b) ten identical tests were performed.

The reason for confining attention to such a narrow range of variables and for repeating the tests a considerable number of times was that it was hoped to produce data which would be sufficiently reliable to enable an appraisal to be made of the detailed behaviour of the interface waves. There is always some uncertainty about the details of the results of a particular test because of the nature of the explosive loading and because of the possibility of slight irregularities in both the dimensions and the material consistency of the components. For example, the interface geometry at the detonator end of the specimen may be affected by transient phenomena associated with the establishment of the flyer plate collapse mechanism or the development of the detonation wave propagating through the explosive. Effects such as these might be expected to vary even for nominally identical set-up conditions. It was hoped to minimize the effect of the development of the detonation wave by detonating the explosive away from the end of the flyer plate by using an auxiliary flyer plate, as shown in figure 2(b).

The performance and analysis of a number of repeated tests should enable irregularities such as these to be eliminated and conversely it should allow features which are truly characteristic of the wave-making process to be made more apparent.

RESULTS

Typical data from weld interface examination

When the weld interfaces were scanned it became clear that the amplitude of the waves was not constant. In all cases the waves appeared at a distance of approximately 7 mm from the end of the specimen. The amplitude and wavelength then increased along the interface in the direction of welding until it reached a maximum value at a point which varied from test to test. However, instead of remaining constant over the remainder of the interface, in all cases it was noted that the amplitude appeared to pass through a succession of maxima and minima. It was decided therefore to measure the amplitude and wavelength of the waves at several points along the interface, particularly near to the observed maxima and minima. These were then plotted in graphs of which figures 4(a) and (b) are representative.

Mean amplitude and wavelengths: checks with previous results

From the complete set of graphs, the mean amplitude and wavelength for each series of tests were found. Both amplitude and wavelength varied along the weld interface and in the same kind of way. In terms of percentage variation the wavelength varied over a much narrower band than did the amplitude. However, in both cases a sufficient number of tests had

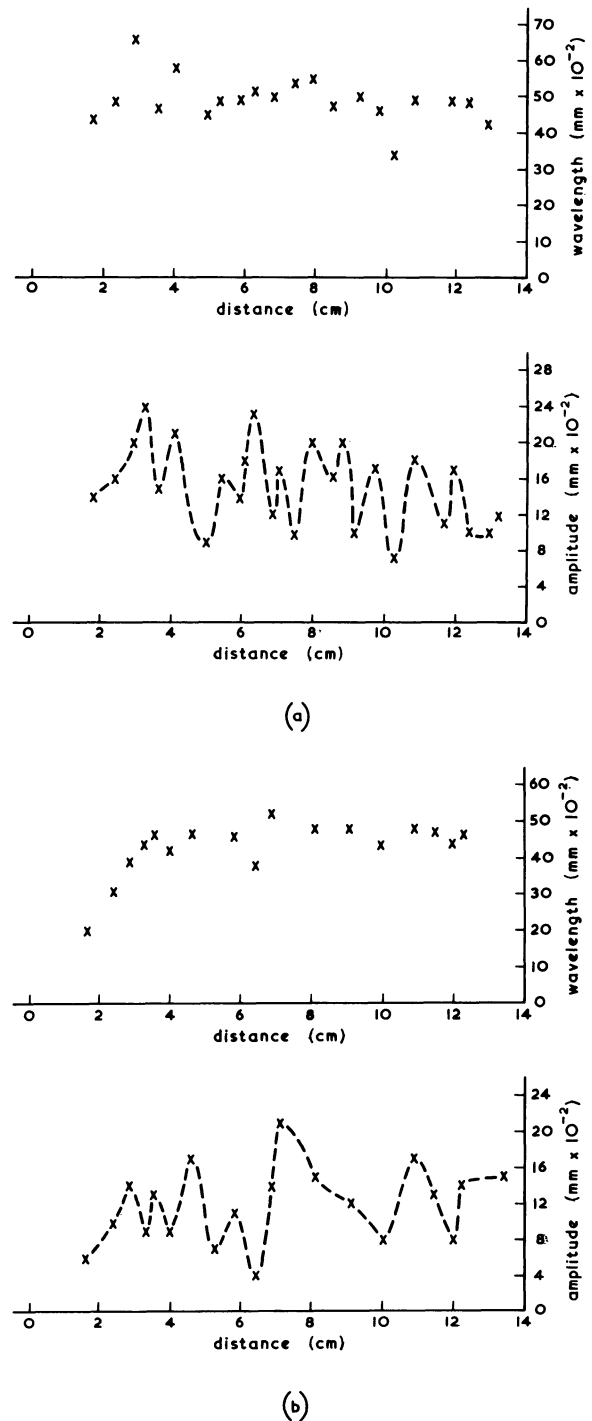


Figure 4. Typical wavelength and amplitude variations along the interface (both taken from series B).

been performed to enable representative mean values of the two parameters to be estimated for each series.

The mean amplitudes, \bar{a} , and wavelengths, $\bar{\lambda}$, are given in Table 1, together with the dynamic bend angles found from separate tests. Also shown is the ratio $\bar{\lambda}/\beta^2$. Several of the theories which have been proposed to explain the mechanism of interface wave formation indicate that

$$\bar{\lambda} \propto t_f \beta^2$$

for small angles of impact⁵. This has been confirmed experimentally by several authors (see for example Deribas *et al.*¹⁰) and Table 1 indicates that

the three series of tests under discussion also satisfy this relationship reasonably well.

TABLE 1 Dynamic bend angle, impact angle, mean amplitude and mean wavelength for each series of tests.

Series	δ	β	$\bar{\lambda}$ (mm)	\bar{a} (mm)	$\bar{\lambda}/\beta^2$	$\bar{a}/\bar{\lambda}$
A	7°	14°	0.58	0.19	9.7	0.33
B	5°30'	12°30'	0.48	0.16	10.0	0.33
C	4°30'	11°30'	0.40	0.14	9.9	0.35

The ratio of the mean amplitude to mean wavelength is also shown in Table 1. In spite of the large variations in the amplitude this ratio is also remarkably constant at about 0.33. This feature of the results has also been noted before⁵ and will be discussed later.

Modulation of the waves

The distances between successive maxima and minima were measured from the graphs giving the amplitude data for each test. These distances will be referred to as the modulated wavelength. The statistics of the number of appearances of particular values of the modulated wavelength for the combined results of all three tests are shown in figure 5. The major feature of the results is that almost a third of the distances measured lie in the range 1.1 cm to 1.4 cm. There is a number of modulated wavelengths less than 1.1 cm but most of these lie in a band between 0.55 cm and 0.7 cm, the limits of which are half those of the most

dominant range of modulated wavelengths. Following the terminology used in acoustics, this seems to imply that the wavelengths in the most dominant group, group II in figure 5, are *possibly* the first harmonics of the fundamental band of wavelengths in group I. Continuing this process, bands corresponding to higher harmonics can be constructed and these are shown as III and IV in figure 5. Beyond group IV the ranges overlap and are not shown. It can be seen that a large proportion of the modulated wavelengths less than 2.8 cm fall within groups I to IV. This behaviour led the authors to conclude that the amplitude variation was not completely random but that some regular phenomenon was influencing the waves as they were being formed.

Taking the results of the three series of tests separately does not reveal any significant difference in the bandwidths and their limits. The only feature which could be noted was that series A and C showed a slight preference for group IV, although still exhibiting large proportions of results in groups II and III.

DISCUSSION

Amplitude variation

In reference 5 the question of interface wave formation in explosive welding is discussed in some depth and references are given to many of the papers which have considered the theoretical aspects of the problem, as well as those which contain experimental data on wave dimensions. Most of them follow Deribas et

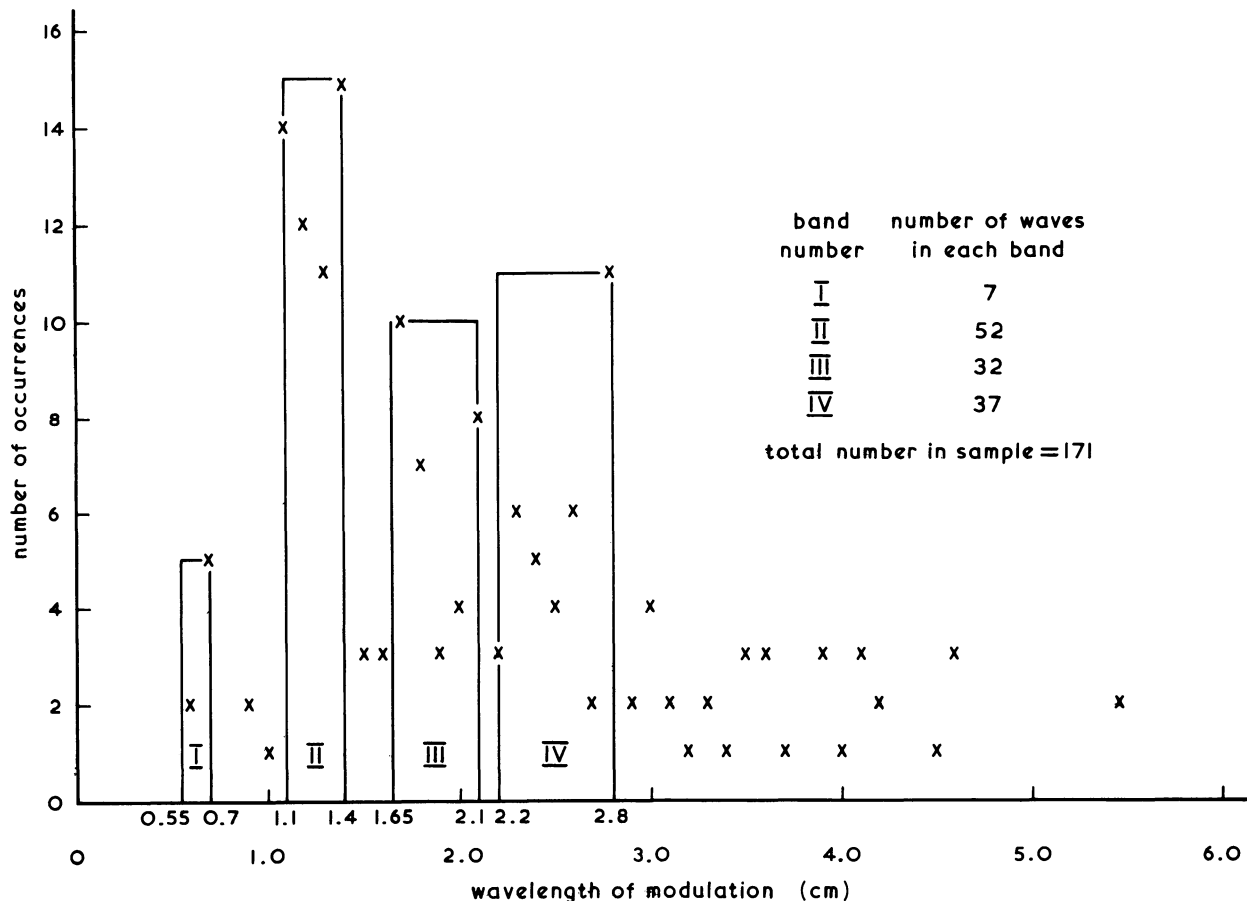


Figure 5. Distribution of modulated wavelengths for all the tests.

al.¹¹ and 'select the wavelength as the basic experimental parameter for the quantitative characterization of the wave dimensions'. In terms of mean wave dimensions this is reasonable in view of the relative constancy of the mean amplitude to mean wavelength ratio which has been demonstrated above and confirmed by several authors, including Kowalick and Hay¹². However, whenever this ratio is referred to in the literature there is always mention of the amount of scatter in its value, and it is argued in reference 5 that this is primarily due to scatter in the amplitude. This view is supported by the results of the tests performed by the authors as well as those of Klein¹³, which also show a much wider variation in the amplitude data than in the corresponding wavelengths.

If one is attempting to construct a theoretical model for explosive welding, it is important to know whether or not such scatter is random, or whether it is another inherent characteristic of the mechanism which is responsible for the production of the waves. If the latter is the case, then any theory which is proposed should be capable of explaining the amplitude variation.

Qualitative explanation of modulation

The conclusion reached in reference 5 is that, of all the different mechanisms that have been proposed, the so called vortex-shedding mechanism¹² best fits the experimental data. According to this model the confluence of the flyer plate and parent plates in the region behind the stagnation point S is analogous to the flow of a fluid round an obstacle, such as a cylinder, in hydrodynamics. This flow is characterized by the alternate shedding of vortices from either side of the cylinder, forming a Karman vortex street, provided the Reynolds number of the flow lies in a certain range. In reference 5 it is argued that the same sort of mechanism that produces the vortex street in fluids produces interface waves in explosive welding and a discussion is given of the various ways in which different vortex shedding theories could be applied to explosive welding.

The situation in explosive welding is of course much more complex than in the hydrodynamic analogue since the region over which the materials exhibit fluid-like behaviour is finite. The interface waves which one sees and measures are 'frozen' into the components after leaving the fluid region which moves along behind the stagnation point with a velocity V_w .

In a real (viscous) fluid the vortex street grows in lateral extent downstream of the obstacle, due to the annihilation of vorticity by viscous diffusion. However it can be shown theoretically that the mean longitudinal separation of the vortices (analogous to the wavelength of the interface waves) remains constant, and this is supported by experimental observation. According to reference 5, therefore, the interface waves are analogous to parts of a vortex street which have been frozen at a certain stage in their development. Suppose for simplicity that in figure 3 the point F, which is a distance f behind the stagnation point, represents the limit of the region within which the pressure is sufficiently high for the

metals to behave in a fluid-like manner. As the stagnation point sweeps across the parent plate the waves are formed near to S, develop within distance f and are frozen as they emerge beyond F; they are frozen due to the drop in pressure below some critical level. Provided f remains constant, the waves which emerge have all developed over the same length and therefore will all have the same amplitude and wavelength, that is the interface waves will be uniform. From the experimental results which have been described this apparently is not the case. The length f must vary so that the waves are frozen at different stages in their development and so give the appearance of being modulated along the interface. Furthermore, from the results of the tests it would appear that the changes in f are produced by some phenomenon which has periodic characteristics.

A precise model for vortex shedding has still not been found, even for the hydrodynamic situation; however it is clear that in the vicinity of the cylinder the fluid is subjected to periodic disturbances transverse to the flow direction, the frequency of which is equal to the vortex shedding frequency. This motion has been modelled by a swinging wake⁵. As far as explosive welding is concerned these periodic disturbances generate two series of pulses which are alternately compressive and tensile, and which propagate into the flyer plate and the parent plate. These pulses expand cylindrically and are reflected as tensile and compressive waves respectively from the free surface of the flyer plate, and possibly from the interface between the parent plate and anvil. The front of the tensile wave produced by the reflection of a compressive pulse from the free surface of the flyer plate is shown in figure 6. These waves

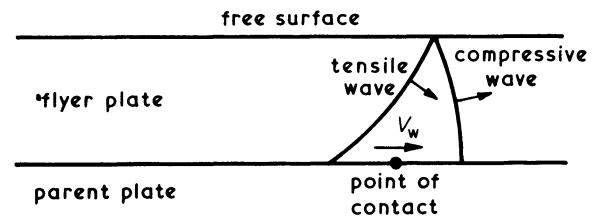


Figure 6. Compressive wave and reflected tensile wave propagating in flyer plate (following reference 12).

propagate with a speed of the order of the elastic wave speed which exceeds V_w , as indicated in the Appendix, and therefore they catch up with the stagnation point, modifying the pressure in the region behind the stagnation point and hence varying the freezing length, f . A detailed analysis of the effects of the reflected waves in the welding configuration used in the tests is almost impossible because of the unknown interface conditions between the parent plate and the anvil. However, a crude estimate will now be made of the order of magnitude of the modulated wavelength by considering the time interval over which particular tensile waves catch up with the stagnation point.

Modulated wavelength: previous treatment and order of magnitude

The only other reference to the possibility that interface waves are modulated which the authors have

found is that given briefly in a paper by Godunov et al.¹⁴. They report that in some tests which they performed, the interface waves also exhibited very long, shallow, modulating waves whose wavelength exceeded many times the wavelength of the individual interface waves. Unfortunately no details are given of any dimensions, set-up conditions, etc.

In reference 14 the stagnation point, S, is treated as a travelling *uniform* pressure source, unlike the model just described which implies that the pressure source is pulsating with a frequency analogous to the vortex shedding frequency. (A constant pressure source is incapable of producing a periodically varying pressure field in the vicinity of the stagnation point. The reason for this is that the pressure field resulting from the reflection of the pressure waves generated at the stagnation point is independent of time when referred to a set of co-ordinates moving with the stagnation point, provided that the pressure near to the stagnation point is constant.) Notwithstanding this, it is shown in reference 14 that if a compressive wave is emitted from the stagnation point at a certain time, then the tensile wave produced by reflection of the compressive wave from the free surface of the flyer plate catches the stagnation point after a time τ given by

$$\tau = \frac{2t_f}{\sqrt{(c_0^2 - V_w^2)}} \quad (1)$$

where c_0 is the speed of the compressive and tensile waves in the flyer plate. (In general, when considering solids, one should also include a shear wave as part of the reflected system of waves. This is ignored by Godunov et al. since they treated the materials as essentially inviscid fluids. However, the shear wave would always be behind the tensile wave because of its lower propagation speed and would presumably only produce secondary effects.) During the time interval τ the stagnation point has travelled a distance

$$\frac{2V_w t_f}{\sqrt{(c_0^2 - V_w^2)}} \quad (2)$$

According to Godunov this gives a measure for the modulated wavelength. Using $t_f = 1.25$ mm, $V_w = 2500$ m/s and $c_0 = 5000$ m/s, which are applicable to the tests performed, the modulated wavelength given by (2) is only 1.44 mm, which is an order of magnitude too small for the dominant wavelength shown in figure 5. However, basing equation (1) on the parent plate thickness of 12.7 mm produces a modulated wavelength which is commensurate with this dominant wavelength. There may therefore be something of value in an approach of the type described in reference 14.

Future investigation of modulation

Although the tests described indicate that the interface waves are subjected to modulation, it is clear that in order to examine the effect more closely, the experimental arrangement requires modifying. Several things could be tried, including:

- (1) varying the thickness of the parent plate,
- (2) changing the method of support of the parent

plate to remove any uncertainty about the boundary conditions at its lower surface,

- (3) producing welds between two colliding flyer plates with identical explosive loading using the symmetrical set-up⁵. According to the previous section this should produce modulations with much smaller wavelengths, since both the length dimensions involved are small.

CONCLUSION

From a systematic analysis of the results of three series of tests it would appear that the variation in the amplitude (and wavelength to a lesser degree) of the interface waves is not random. Assuming that the waves are formed by a mechanism which is analogous to vortex shedding in hydrodynamics, it is proposed that the modulation results from the modification of the length of the region within which the waves develop by stress waves. These stress waves result from the reflection of a series of pulses which are generated by the oscillatory process which produces the interface waves.

ACKNOWLEDGMENTS

The authors would like to thank Professor W. Johnson for his comments on the draft of this paper and Mr N. H. S. Sherif for the use of the photomicrograph shown in figure 1.

REFERENCES

1. B. Crossland, J. D. Williams and V. Shribman. Explosive cladding of large plates. *Select Conference on Explosive Welding*, Paper 12, London (Institute of Welding) (1968) p. 15.
2. T. Z. Blazynski and A. R. Dara. The use of explosives for implosive welding of duplex cylinders. *Proc. 11th Int. MTDR Conf.* Pergamon Press, 1968.
3. B. Crossland and J. D. Williams (1970). Explosive welding. *Metallurgical Revs., Review 144, Metal and Materials* (1970) 7(7), 79.
4. A. A. Ezra. *Principles and Practice of Explosive Metal Working*, Vol. 1. Industrial Newspapers Ltd., 1973.
5. S. R. Reid, A discussion of the mechanism of interface wave generation in explosive welding, *Int. J. Mech. Sci.* (1974) 16, 399.
6. D. J. Langdale, M.Sc. Dissertation, University of Manchester Institute of Science and Technology, 1973.
7. J. N. Hunt. Wave formation in explosive welding. *Phil. Mag.* (1968) 17, 669.
8. F. W. Travis and W. Johnson. Explosive welding of stellite to stainless steel. *Proc. 8th Int. MTDR Conf.*, Pergamon Press, 1967.
9. V. Shribman and B. Crossland. An experimental investigation of the velocity of the flyer plate in explosive welding. *Proc. 2nd Int. Conf. of the Centre for High Energy Forming*, vol. 2, Denver, Colorado, 1969.
10. A. A. Deribas, V. M. Kudinov, F. I. Matveenkov and V. A. Simonov. Simulation of the process of wave formation in explosive welding. *Combustion, Explosion and Shock Waves* (1968) 4 (1), 100.

11. A. A. Deribas, V. M. Kudinov and F. I. Matveenkov. Effect of initial parameters on the process of wave formation in explosive welding. *Combustion, Explosion and Shock Waves* (1967) 3 (4), 561.
12. J. F. Kowalick and D. R. Hay (1971). A mechanism of explosive bonding, *Metallurgical Transactions* (1971), 2, 1953.
13. W. Klein. Gefügeänderungen beim Explosivplattieren von Stahl mit Stahl und einigen Nichteisenmetallen. *Z.f. Metallkunde* (1965) 56, 261.
14. S. K. Godunov, A. A. Deribas, A. V. Zabrodin and N. S. Kozin. Hydrodynamic effects in colliding solids, *J. Computational Physics* (1970) 5, 517.
15. G. Birkhoff, D. P. McDougall, E. M. Pugh and G. I. Taylor. Explosives with lined cavities. *J. Appl. Physics* (1948) 563.

APPENDIX

Flyer plate collapse kinematics

The parameters U_d , V_p , V_w , V_f , V_s and V_j shown in figure 3 are defined as follows:

- U_d detonation speed of explosive
- V_p flyer plate velocity
- V_w velocity with which stagnation point sweeps across the parent plate
- V_f velocity of flyer plate relative to stagnation point (that is, velocity of flyer plate material in a reference frame moving with the stagnation point)
- V_s slug velocity
- V_j velocity of re-entrant jet

The equations relating these velocities can be found in several places (see for example reference 15) and therefore they will simply be stated.

Since V_p is the vector sum of V_w and V_f , the velocity triangle implies that

$$\frac{V_p}{\sin \beta} = \frac{V_w}{\cos(\delta/2)} = \frac{V_f}{\cos((\alpha + \beta)/2)} \quad (\text{A1})$$

Adapting the geometrical argument given in the above reference to the case in which the detonation wave travels along the flyer plate, we have

$$V_p = 2U_d \sin \frac{\delta}{2} \quad (\text{A2})$$

If required, $V_s = V_f - V_w$ and $V_j = V_f + V_w$. Since V_p depends only on U_d and δ , all the velocities in (A1) depend only on U_d , δ and α .

Stagnation point velocity, V_w , and determination of U_d and δ

From (A1) and (A2),

$$V_w = \frac{V_p \cos(\delta/2)}{\sin \beta} = \frac{U_d \sin \delta}{\sin(\alpha + \delta)} \quad (\text{A3})$$

since $\beta = \alpha + \delta$. It is often stated that one of the limiting conditions for welding to occur is that the velocity of the stagnation point must be less than the 'sound' speeds of the two components³. As mentioned in reference 5 it is not often made clear as to what precisely is meant by the sound speed. In the calculation of the modulated wavelength in the discussion above the elastic wave speed, c_0 , has been taken. Whether this is suitable is a matter for future discussion.

The relationship between V_w , U_d and δ given by equation (A3) is the basis of the pin-contact method used to find U_d and δ . In this technique pins are set into the parent plate at regular intervals⁸ and the interval of time between the flyer plate contacting adjacent pins is measured electronically. The velocity with which the line contact sweeps over the parent plate, V_w , can therefore be easily calculated. If the flyer plate is initially set parallel to the parent plate ($\alpha = 0$) then, from equation (A3), $V_w = U_d$. Thus the detonation speed can be determined. Measuring V_w for a non-zero value of α and using the previously measured value of U_d , one can then calculate δ again using equation (A3).

IMPLOSIVE WELDING OF COMPOSITE METAL-FOIL CYLINDERS

by

H. EL-SOBKY* and T. Z. BLAZYNSKI†

SUMMARY

General problems involved in the use of explosives for the welding of multi-layer metal-foil cylinders are discussed. The quality of the welds, the question of the interfoil radial clearance and the weldability of some metal combinations are examined.

INTRODUCTION

Implosive and explosive welding of mono- and bi-metallic solid duplex cylinders has now become well established and has been investigated and described by Blazynski and Dara¹⁻³. Specialized industrial applications, however, demand the production of metallic composite cylinders containing either metallic or non-metallic filament reinforcements. The technique of welding adopted in the case of solid cylinders is not satisfactory when multi-layered laminates are required and, consequently, cylinders made up from layers of metal foil have to be made.

Multi-layered, metal-mesh reinforced plates have been successfully, explosively, bonded by, among others, Jarvis and Slate⁴, Wylie⁵ et al. and Reece⁴. In most of these experiments a single layer reinforcement would be sandwiched between two, relatively thick, plates. Problems of misalignment would often arise and basic structural inhomogeneity would be introduced. Initial tests carried out by El-Sobky and Blazynski⁷ on metal foil cylinders, reinforced with metal mesh, welded together implosively, indicated clearly that the problems associated with the fractional foil-surface welding, extended in depth to a number of layers, required further detailed attention. The quality of the weld obtained between the successive layers is of particular importance since, as experience shows⁷, the relationship between the thickness of the foil, thickness of the reinforcement and filament diameter can be critical.

The present paper is concerned with the implosive welding of multi-layered metal-foil cylinders in their basic form, i.e. excluding filament reinforcement. The following parameters have been investigated:

- (i) Quality of the weld along the multiple surfaces of interfacial contact of the layers across the

thickness, as related to energy transmitted from the charge.

- (ii) The effect of the very small 'stand-off' or interlayer clearance
- and
- (iii) The weldability of the various metallic combinations.

THEORETICAL CONSIDERATIONS

With reference to figure 1, which shows the welding arrangement prior to detonation, it is clear that as the detonation front travels downwards pressure waves are propagated obliquely through the multi-layered cylinder. These waves will be reflected from the core of the assembly and will set up a system of stresses within the cylinder. The reflected wave intensity is dependent on the difference in the acoustic impedances of the metals involved. Acoustic impedance is defined as ρc , where ρ is the density and c is the velocity of propagation of the wave. If the particle velocity at a point within a longitudinal wave is u , the stress in the xx direction is given by

$$\sigma_{xx} = \rho c_1 u \quad (1)$$

and the transverse stress

$$\tau = \rho c_2 v \quad (2)$$

where c_1 , and c_2 are the velocities of wave propagation in two directions, and v is the particle velocity in the transverse direction.

In the usual notation, the two velocities are given by

$$c_1 = \left| \frac{E(1-\nu)}{\rho(1+\nu)(1-2\nu)} \right|^{1/2} \quad (3)$$

* Department of Mechanical Engineering, Al-Azhar University, Cairo (on study leave at University of Leeds)

† Department of Mechanical Engineering, University of Leeds

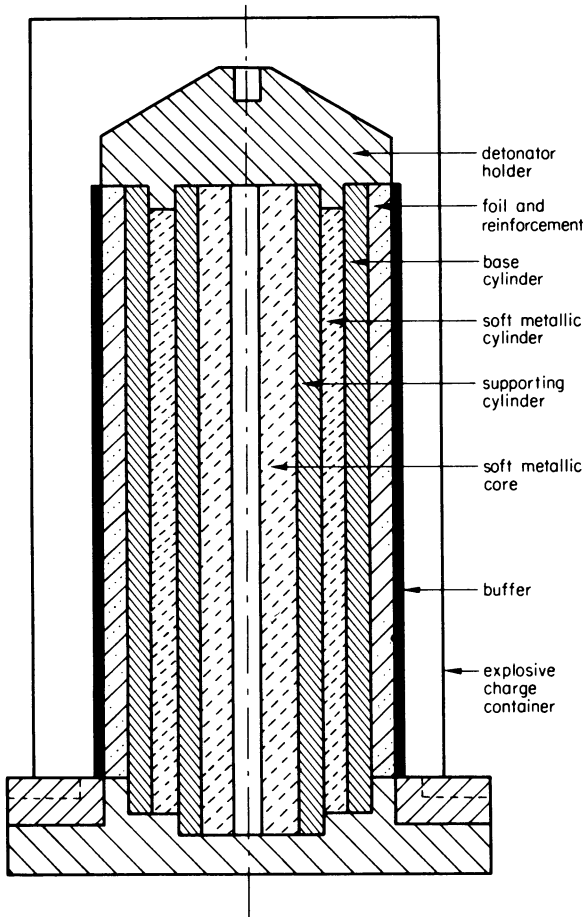


Figure 1. Diagrammatic arrangement of the welding system.

and

$$c_2 = \left| \frac{E}{2\rho(1+\nu)} \right|^{1/2} \quad (4)$$

Transmission and reflection of waves through a number of layers of foil is shown, diagrammatically, in figure 2. When reflection occurs, the relation

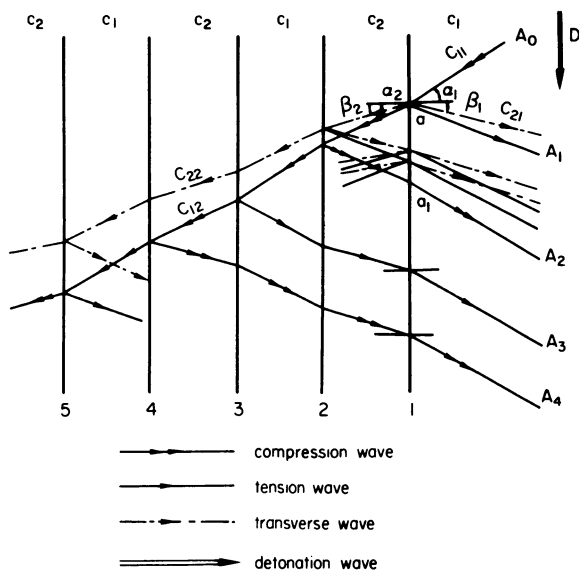


Figure 2. Successive reflections and refractions of a longitudinal compression wave in bi-metallic multi-layer welding systems.

between the stress produced by the incident wave (σ_i) and that produced by the reflected one (σ_r) is given by

$$\sigma_r = R\sigma_i \quad (5)$$

For the transverse, shear, stress the expression is

$$\tau_r = [(R + 1) \cot 2\beta] \sigma_i \quad (6)$$

The coefficient of reflection R is

$$R = \frac{\tan \beta \tan^2 2\beta - \tan \alpha}{\tan \beta \tan^2 2\beta + \tan \alpha}$$

Where the relationship between α and β is given by

$$\frac{\sin \alpha}{c_1} = \frac{\sin \beta}{c_2}$$

Equation 5 can also be expressed as

$$\sigma_r = \left[\frac{\rho_2 c_{12} - \rho_1 c_{11}}{\rho_2 c_{12} + \rho_1 c_{11}} \right] \sigma_i \quad (5a)$$

The equation indicates that the reflected wave depends on the interplay of acoustic impedances of the two metals, and that the coefficient of reflection is a function of the angle of incidence. It follows from Equation (5a) and figure 2 that the reflected stress wave will change signs alternately. Assuming that $\rho_1 c_{11} > \rho_2 c_{12}$ and considering, for simplicity, the longitudinal compression wave only and retaining general intensity A_0 , it is found that A_1 and A_2 will be tensile and compressive respectively. It is also clear that the angle of incidence constitutes the only factor that governs displacements between the points of intersection with the surfaces.

In the investigated system, the initial incident compression wave is caused by the pressure generated by the explosive at the point of collision. Assuming parallel layers, the point of collision can be said to be moving with a velocity $V_c = D$ (velocity of the detonation front). The distance between the points of incidence and reflection is given by

$$aa_1 = 2T \tan \alpha \quad (7)$$

Where T is the thickness of foil.

The velocity of the point of reflection is $c_1 \sin \alpha$.

If the collision point and the incident wave (A_0) are travelling at the same speed, i.e. $D = C_1 \sin \alpha$, then interference will occur between A'_0 , the tensile reflection of wave A_0 , and the compressive reflection of wave A_1 . The same situation, although in reverse order, will obtain at the next interface if $\rho_2 c_{12} > \rho_1 c_{11}$. On the other hand, if $D > c_1 \sin \alpha$, a situation may arise in which a reflected wave, striking a newly welded surface, behind the point of collision, will give rise to a very high stress which, in turn, may cause separation of surfaces. Clearly, therefore, the required condition is that

$$D < C \sin \alpha \quad (8)$$

The velocity of detonation depends on the density and thickness of the charge, but it is more sensitive to the changes in density. In fact, it increases with an increase in density and/or in thickness [$D = D(\rho^3)$, $D = D(t)$]. It is therefore preferable to obtain the

required energy levels by increasing the charge thickness rather than its density.

EXPERIMENTAL DETAILS

Figure 1 shows the implosive welding arrangement. The mechanism of welding is fully described in references 1 to 3. The bore of the multi-layered cylinder is supported by a soft metallic core stiffened by means of a steel cylinder. The soft metallic cylinder, contained between the base and supporting tubes, is capable of absorbing excess energy and of allowing the base tube to deform radially. Three materials were used as base tubes, viz. Al, Cu and mild steel. Cylinder compositions and other data are given in Table 1. The relevant properties of the alloys used are listed in Table 2.

TABLE 2 Material Properties

Material	Velocity of propagation		Density (g/cm ³)
	C ₁ (m/s)	C ₂ (m/s)	
Al	6320	3100	2.7
Brass	4300	2100	8.3
Cu	4560	2250	8.9
M.S.	5940	3220	7.8

Typical, welded specimens, are shown in figure 3.

The I.C.I. Trimonite No. 1 powder explosive was used throughout. The explosive has an energy release level of about 1260 Kcal/g and a density of 1.1 g/cm³. On average, a 16 mm layer of explosive was used, giving a detonation velocity of about 3 km/s.

TABLE 1 Materials and sizes of welded cylinders

Test no.	Base tube			Foil 1		Foil 2		Total O.D. before welding (mm)
	Material	O.D. (mm)	I.D. (mm)	Material	t (mm)	Material	t (mm)	
1	Cu	58.4	50.8	Cu	0.1016	Brass	0.1016	61.0
2	Cu	58.4	50.8	Cu	0.1016	Al	0.1016	61.0
3	M.S.	58.4	50.8	M.S.	0.0762	Al	0.1016	61.0
4	M.S.	58.4	50.8	M.S.	0.0762	Cu	0.1016	61.0
5	M.S.	58.4	50.8	M.S.	0.0254	Brass	0.1016	60.0
6	M.S.	58.4	50.8	M.S.	0.0762	M.S.	0.0762	60.4
7	M.S.	58.4	50.8	Brass	0.1524	Al	0.1016	62.2
8	M.S.	58.4	50.8	Al	0.1016	Al	0.1016	61.0
9	Al	58.4	50.8	Al	0.0127	Al	0.0127	60.0

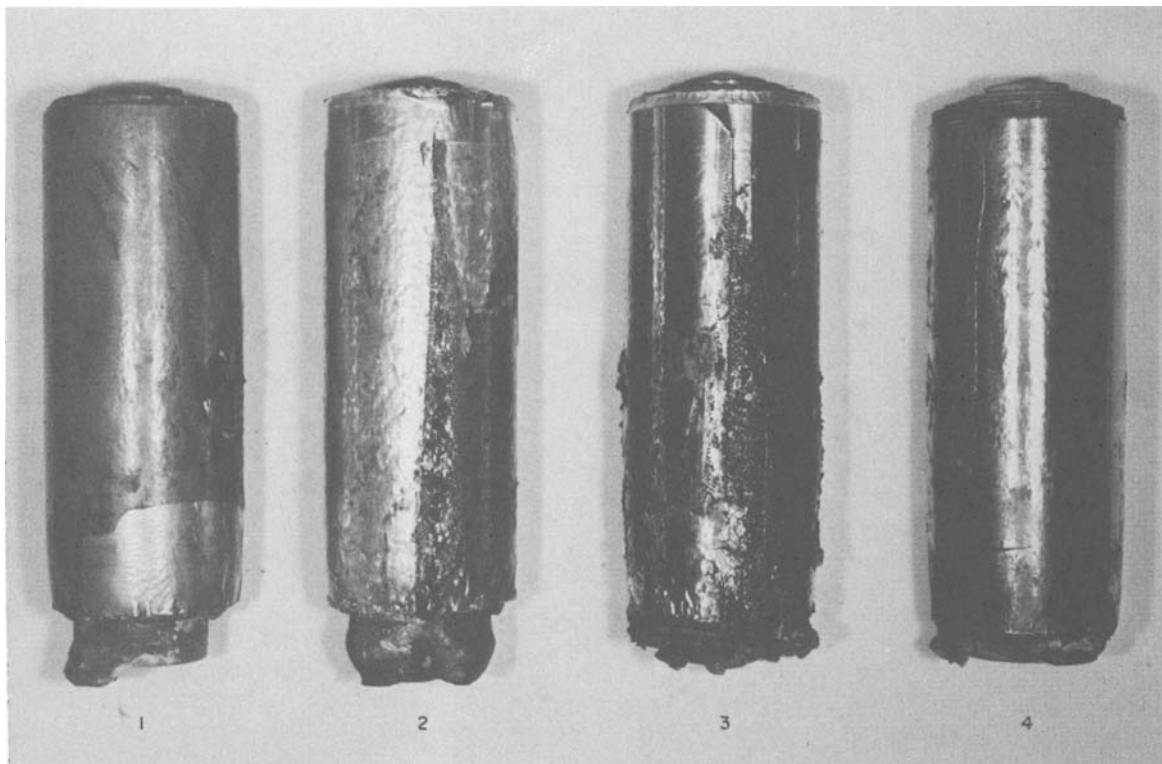


Figure 3. Typical foil cylinders after welding. 1-Cu-brass, 2-Cu-Al, 3-Al-Steel, 4-Cu-Steel.

DISCUSSION OF RESULTS

Geometry of the Composite Cylinder

The experimental welding arrangement, shown in figure 1, proved satisfactory and made it possible to achieve the welding of uniform multi-layered cylinders (figure 3). The inner steel tube provides a rigid support, but, of course, it also introduces an additional wave-reflecting surface. It was found that in those cases in which the acoustic impedance of the base tube was much higher than that of the welded foil, separation would occur in the vicinity of the base tube interface. At the same time, however, composite layers of the foil would be welded together. Examination of the base interface would usually reveal the presence of minute fragments of welded layers of foil, thus indicating that the original weld had been broken as a result of wave interference created in this area.

At the instant of separation, the sharp edge of the innermost foil would tend to spring back and to act

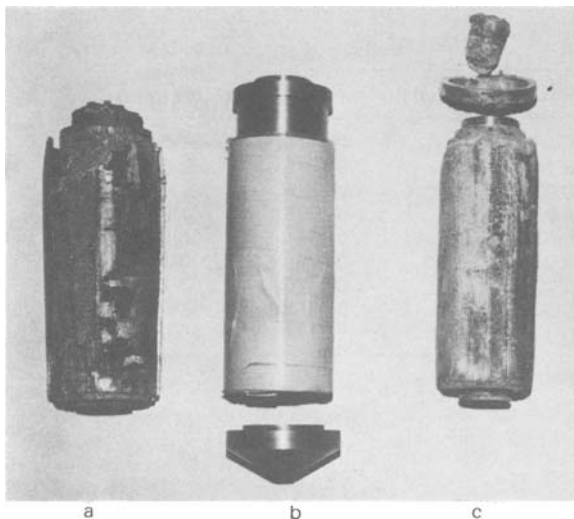


Figure 4. Before and after welding: (a) split due to separation at the base tube, (b) assembly before explosion, (c) successfully welded cylinder.

as a guillotine producing a sharp cut along the cylinder (figure 4).

Quality of the weld

Typical results obtained with a variety of mono- and bi-metallic combinations are shown in figures 5 to 8.

On the whole, welds in mono-metallic foil cylinders were of satisfactory quality, but they would tend to deteriorate radially inwards in relation to the quality of the outer layers. In some cases, particularly with aluminium, complete separation in the inner layers would occur. Figure 5 shows an Al-Al combination in which separation, in the form of a standing wave, is present.

Successful welding of bi-metallic combinations depended on the relative values of acoustic impedances of the metals involved. In this respect, the acoustic impedance of aluminium would usually substantially differ from that of any other metal, and, consequently, bi-metallic combinations involving aluminium would introduce certain difficulties. In particular, the Al-Cu composites were characterized by a large number of separations and satisfactory welds were extremely difficult to obtain.

The welds were very good when brass-copper foils were welded on a copper base tube. Figure 6, for instance, shows that the 'wavy' type weld occurs on alternate surfaces, with 'line' metal to metal weld on the remaining interfaces. The wave amplitude increases radially towards the base tube whereas, at the same time, the frequency of the wave is reduced.

With reference to figure 2, it can be noted that the 'wavy' interface occurs in those sections in which pressure waves are transmitted from the medium of a high acoustic impedance to the medium of a lower one, i.e. in the sections in which Equation (5a) predicts a change in the nature of the reflected longitudinal wave.

The fact that the frequency of the interfacial waves increases towards the outer surfaces, where the number of reflected waves and hence the interference effects are greater (fig. 2), indicates that there exists a

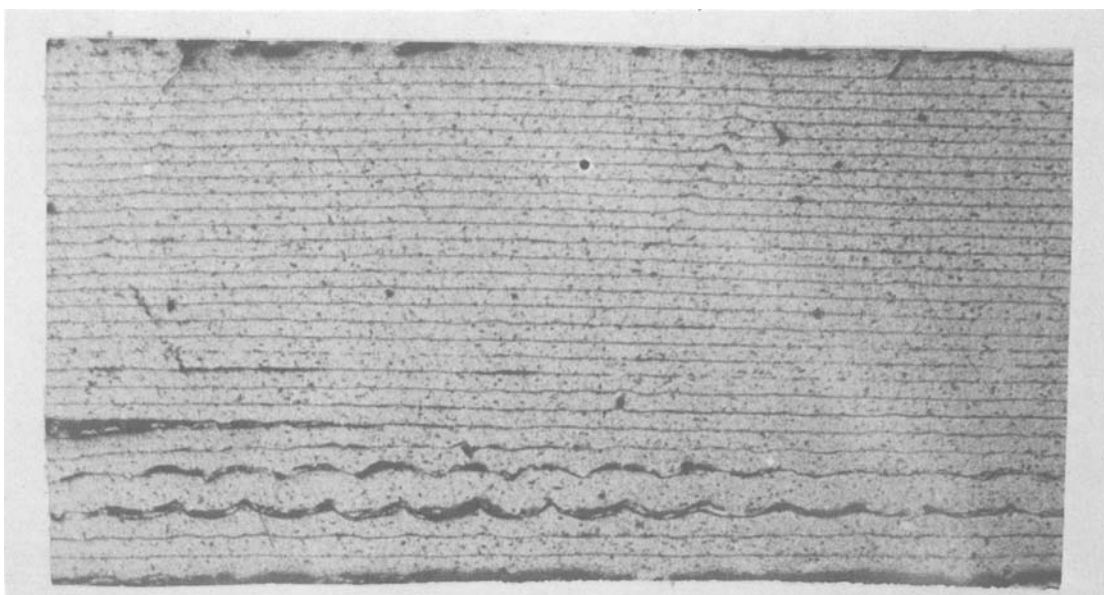


Figure 5. Al-Al foil weld. ($\times 150$).

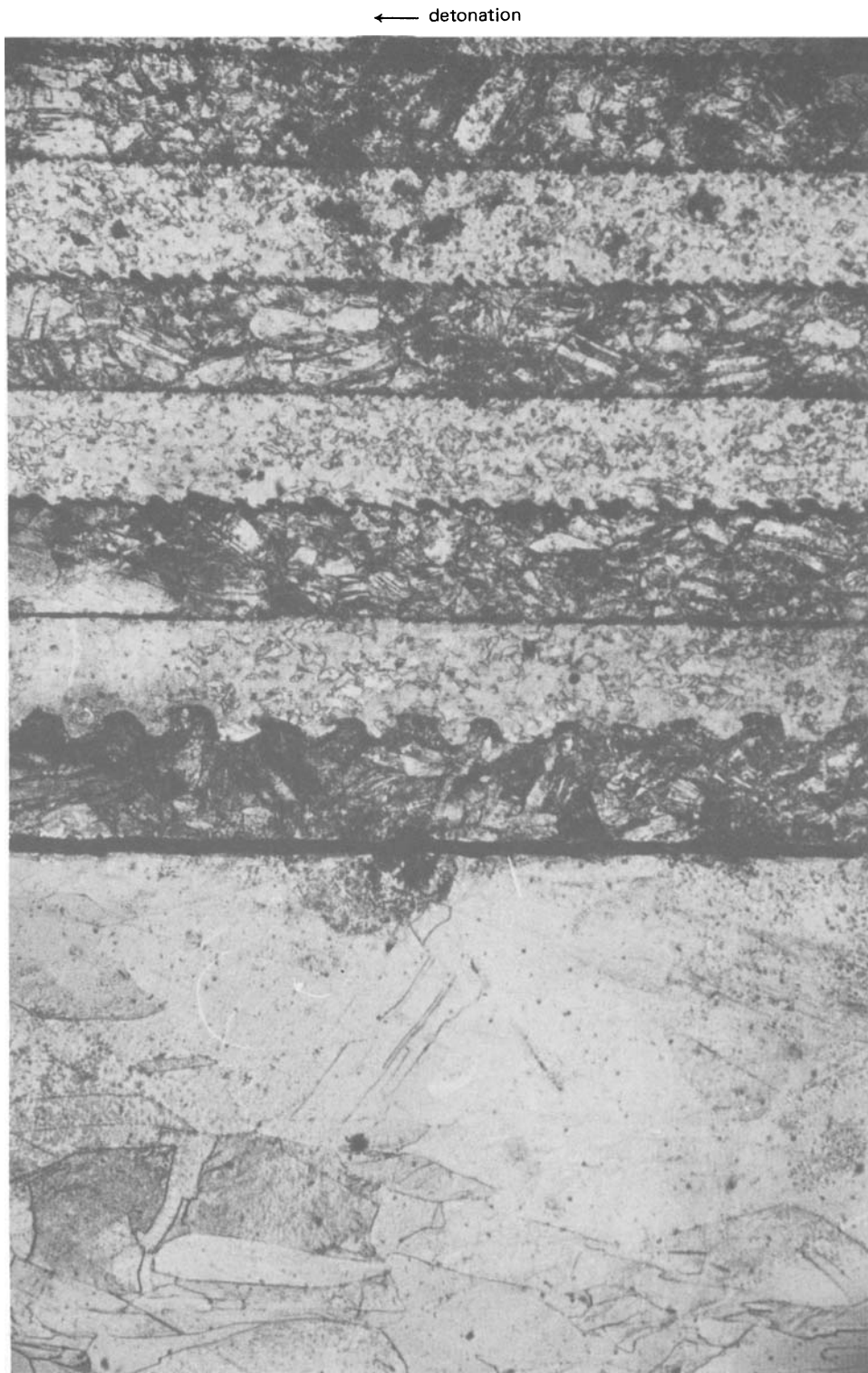


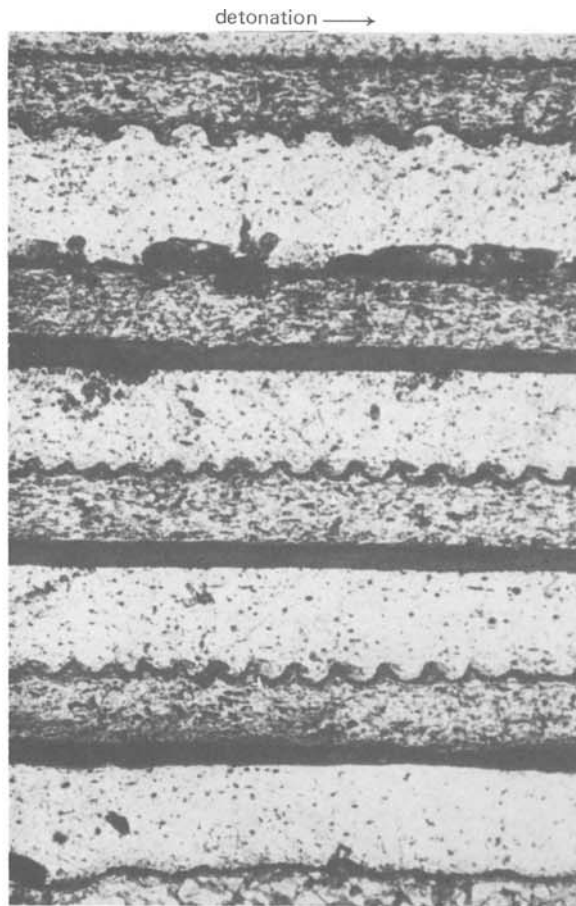
Figure 6. Cu-brass bi-metallic foil cylinder ($\times 160$).

relationship between the phenomenon of the surface wave formation and the nature and type of the pressure/stress wave, including, of course, the effect of the transverse waves.

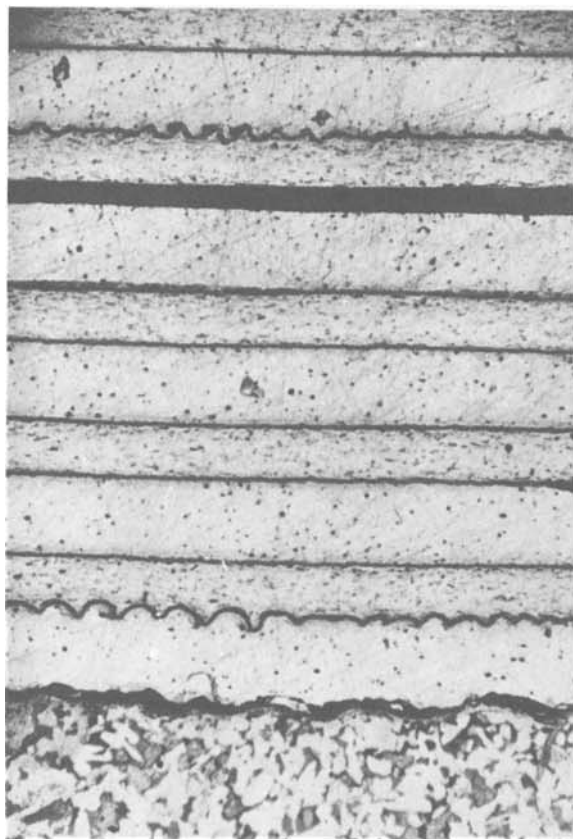
When welding steel-copper foil composites, the same arrangement of 'wavy' and non-wavy welds was observed. Figure 7a shows a situation in which separation between welded layers occurred (black areas).

Figure 7b deals basically with the same composition, but with a change in the level of supplied energy. Non-wavy welds are predominant, the wavy ones being associated with the layers in which separation has occurred.

Figure 8 shows a steel-aluminium composite with no 'wavy' welds, but with some traces of intermetallic compound produced as a result of an increase in the explosive charge.



a



b

Figure 7. (a) Cu-Steel bi-metallic foil cylinder showing multiple welds and separation ($\times 280$). (b) Cu-Steel bi-metallic foil cylinder ($\times 190$).



Figure 8. Al-Steel bi-metallic foil cylinder ($\times 150$).

CONCLUSIONS

The success of welding multi-layered foil cylinders is associated with the relative acoustic impedances of the metals involved. Satisfactory welds can be obtained with many engineering alloys. However, some difficulty is experienced when using aluminium in bi-metallic composites.

In multi-layered combinations the amplitude of 'wavy' welds increases radially whilst the frequency is reduced. Theoretical considerations discussed earlier provide a guide to the potential pattern of behaviour.

The extremely small annular clearances present are sufficient to produce the required amount of jetting.

ACKNOWLEDGMENTS

The authors wish to thank the S.R.C. for the financial support, and the University of Leeds for a scholarship to H. El-Sobky.

REFERENCES

1. T. Z. Blazynski and A. R. Dara (1971). The use of explosives for implosive welding of duplex cylinders. *Proc. 11th Int. MTDR Conference*, p. 940, Pergamon Press.
2. T. Z. Blazynski and A. R. Dara (1971). A comparison between the implosive and explosive systems for welding of duplex cylinders.

- Proc. 3rd Int. Conference of the Center for High-Energy Forming*. University of Denver, Colorado.
3. T. Z. Blazynski and A. R. Dara (1973). Determination of residual stress levels and patterns in implosively welded mono- and bi-metallic compound cylinders. *Proc. 4th Int. Conference of the Center for High-Energy Forming*, University of Denver, Colorado.
 4. P. M. B. Slate and C. V. Jarvis (1972). *J. Inst. Met.* **100**, 21.
 5. H. K. Wylie, J. D. Williams and B. Crossland (1971). Explosive fabrication of fibre reinforced aluminium. *Proc. 3rd Int. Conference of the Center for High-Energy Forming*. University of Denver, Colorado.
 6. O. Y. Reece (1971). Molybdenum wire reinforced columbium composites. *Proc. 3rd Int. Conference of the Center for High-Energy Forming*. University of Denver, Colorado.
 7. H. El-Sobky and T. Z. Blazynski (1973). A preliminary investigation of the implosive welding of filament reinforced metal foil composite cylinders. *Proc. Int. Conference on the use of high energy-rate methods for forming welding and compaction*. University of Leeds.

THE COLD DRAWING OF IMPLOSIVELY WELDED BI-METALLIC TUBES

by

S. TOWNLEY* and T. Z. BLAZYNSKI*

SUMMARY

Bi-metallic, implosively welded, tubes were cold-drawn to various sizes to provide experimental data for establishing a mathematical model of the operation. A theoretical approach to the mechanics of the process, proposed in the paper, offers an insight into the potential modes of failure, and provides a useful guide to the magnitude of drawing loads required at higher strains.

NOTATION

f	The volume fraction of outer material in the combination
k	The shear yield stress of a material
k_1	The shear yield stress of the inner material
k_2	The shear yield stress of the outer material
k_*	The effective shear yield stress of the combination
m_0	The friction factor at the plug
m_α	The friction factor at the die
r_0	The radial position at the die inlet
r_1	The radial position at the die outlet
α	The die semi-angle
β	The angular position of the interface
σ_r	The radial stress
σ_θ	The stress in direction θ
σ_m	The mean drawing stress
$\tau_{r\theta}$	The shear stress associated with the r, θ directions

Suffixes 'm' and 'c' refer to the matrix and cladding respectively.

INTRODUCTION

A technique involving the use of high explosives for implosive welding of solid, seamless, tubular components, to form integral bimetallic duplex cylinders was developed by Blazynski and Dara^{1,2}. The ease of manufacture of large diameter, short length bi-metallic cylinders, afforded by this relatively simple operation, makes it economically attractive to examine the problems associated with a further, conventional, processing of such cylinders; the ultimate objective being the forming of long tubular components. The intention is to provide a relatively cheap, but structurally strong basic component, with a thin sheath of a more expensive metal, possessing,

say, anti-corrosive properties. The application of such compound cylinders ranges from the fields of chemical engineering and atomic power plant, where the transport of high temperature/high pressure fluids is of interest, to the situations in which efficient heat transfer systems, that combine strength with inexpensive means of surface protection, are required.

Cold drawing, involving a reduction in both diameter and wall thickness of an integrally welded combination, constitutes a possible solution to the conventional processing of a short, large diameter, cylinder. The potential disparity in the mechanical properties of the alloys forming a duplex cylinder introduces, on the one hand, an additional difficulty in assessing the practical problems that may arise in drawing, and, on the other hand, in developing a satisfactory analytical model of the operation that would enable a prediction of the likely material behaviour to be made.

Some mathematical analyses of the phenomena associated with the processing of bi-metallic, generally non-welded, combinations have been made in the past by a number of investigators; mostly in connection with the sandwich rolling of strip.

Arnold and Whitton³ analysed sandwich rolling, as a means of deforming hard metals by cladding them with softer ones. They suggested that because of the interfacial frictional stresses set up during the operation, tensile and compressive stresses were established in the hard and soft layers respectively, acting along the axis of the strip. The presence of such stresses lowered the roll force required to deform the hard material. The following concept of an *equivalent* or *modified* yield stress for the sandwich was introduced

$$k_* = (1 - f)k_m + fk_c$$

* Department of Mechanical Engineering, University of Leeds

Atkins and Weinstein⁴ used the idea of an equivalent yield stress, and, applying the same force-equilibrium principles, analysed the drawing and extrusion processes.

The mechanism of simultaneous yielding, applied by all these workers, is illustrated in figure 1(a) by considering the frictionless plane strain compression of a sandwich. Figure 1(b) shows Mohr's circle at the onset of yield in the matrix, and figure 1(c) the onset of yield in the cladding. Figure 1(d) represents a sandwich situation in which the pressure is constant throughout, $p_m = p_c$. The stresses at right angles to the pressure stresses, i.e. σ_c and σ_m , have such values

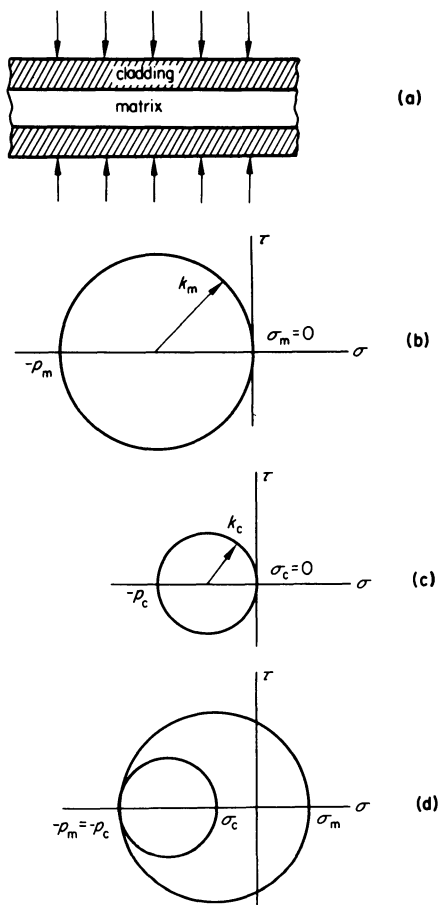


Figure 1. (a) Frictionless plane strain compression of a sandwich. (b) Mohr's circle for compression of matrix alone. (c) Mohr's circle for compression of cladding alone. (d) Mohr's circle for compression of the combination.

as to satisfy the boundary condition of no net horizontal force. The material interfaces must be able to withstand shear stresses in order to support σ_c and σ_m and to prevent relative motion between the layers; otherwise homogeneous deformation is not possible and the model breaks down.

Whitlock and Alexander⁵ investigated the extrusion of bi-metallic strip from separate containers through a single die. Tin was successfully extruded with lead, at room temperature, and copper with silver at 750°C. Satisfactory pressure welding was achieved, provided that the reduction in the die exceeded 60 per cent. Upper bound methods were employed to estimate the extrusion pressures required.

Recently Osakada⁶ et al., using upper bound techniques, have analysed hydrostatic extrusion of bi-metallic rod, in an attempt to determine the various possible modes of deformation. Rods having copper cores and aluminium sleeves were used, the volume fraction of the copper being 0.16. It was found that uniform deformation occurred at low extrusion ratios, and that the copper would fail, in tension, at high ratios. The critical extrusion ratio increased as the die angle decreased. High interfacial friction was found to promote uniform or homogeneous deformation.

It is clear, even from this brief review of some of the existing analytical approaches, that the shear stress distribution, and the interfacial shear stress in particular, play an important part in determining the mode of deformation. Potential failure of the component is, by implication, closely related to the physical conditions generated by these stresses. In the circumstances, it was felt that an analytical approach to the problem of cold drawing of welded cylinders should be based on a model that fully incorporates the effect of shear. Shield's⁷ approach forms such a basis in spite of the fact, that, by virtue of dealing with flow in an infinite channel, it places limitations on the systems in which end effects may play some role. Numerical inaccuracies associated with these effects will be pronounced only in the case of low total strains imposed on the processed material, and will be insignificant at large deformations.

Using cold drawing of steel-copper and steel-brass, implosively welded tubes as examples, the present paper describes the development of a mathematical model of the mechanism of the process, based on a force equilibrium approach incorporating however, a shear stress system.

In the interest of brevity, the derivation of the appropriate expressions is limited, in the next section, to the indication of the basic steps only.

THEORETICAL CONSIDERATIONS

The method of analysis is based on the approach suggested by Shield⁷, and is applied to the drawing of bi-metallic implosively welded tubes. The shear stress distribution is assumed to be constant as the tube is drawn through the die, and end effects are neglected. Strain hardening is accounted for by using an average yield stress for the strain produced. The tube drawing

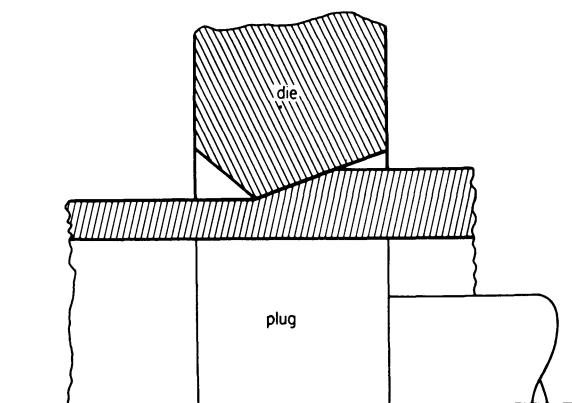


Figure 2. Close-pass drawing with fixed plug.

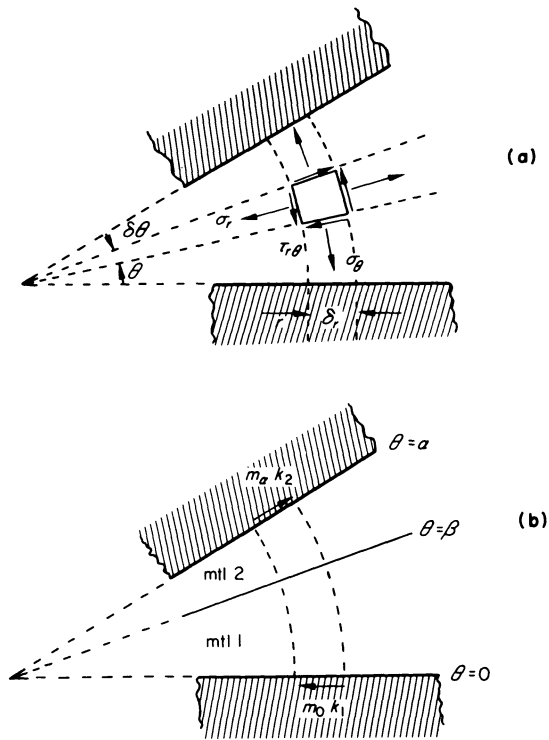


Figure 3. (a) Stresses acting on an element. (b) Interface and frictional stresses acting on the walls.

is considered to be a plane strain process since it is a *close-pass* drawing in which the mean tube diameter changes very little (figure 2).

The flow of a plastic-rigid, bimetallic material, through a rigid triangular wedge-shaped channel, is considered. Von Mises yield criterion is used, and the theory is applied to close-pass bi-metallic tube drawing. With reference to figure 3, consider the following conditions.

Equilibrium in the r direction

$$\sigma_r + r \frac{\partial \sigma_r}{\partial r} + \frac{\partial \tau_{r\theta}}{\partial \theta} - \sigma_\theta = 0 \quad (1)$$

Equilibrium in the θ direction

$$\frac{\partial \sigma_\theta}{\partial \theta} + 2\tau_{r\theta} + r \frac{\partial \tau_{r\theta}}{\partial r} = 0 \quad (2)$$

Yield criterion of Von Mises

For plane strain $\frac{1}{2}(\sigma_r - \sigma_\theta)^2 + \tau_{r\theta}^2 = k^2$

Making the assumption that $\tau_{r\theta}$ is a function of θ only

$$\tau_{r\theta} = k\tau(\theta) \quad \text{i.e.} \quad \frac{\partial \tau_{r\theta}}{\partial r} = 0$$

The yield equation becomes:

$$\sigma_r - \sigma_\theta = 2k(1 - \tau^2)^{1/2} = G \quad (3)$$

Where G is a function of θ only, $\tau = \tau(\theta)$

Equation (2) may now be solved for σ_θ

$$\sigma_\theta = kg(r) - \int 2k\tau d\theta = \sigma_r - G \quad (4)$$

Using Equation (4) in Equation (1)

$$G + rkg'(r) + k\tau'(\theta) = 0$$

As $G + k\tau'(\theta)$ is a function of θ only, and $rkg'(r)$ is a function of r only

$$2(1 - \tau^2)^{1/2} + \tau'(\theta) = c \quad (5)$$

$$g(r) = -\ln r + A \quad (6)$$

where c and A are constants. Equation (5) may be solved for θ in terms of τ . Two sets of equations are formulated to describe the two components of the composite tube. Consideration of the interfacial boundary condition, i.e. that shear stresses must be equal on the boundary, results in $k_1 c_1 = k_2 c_2$. With this knowledge, the values of c_1 and c_2 are determined by equating the shear stresses at the boundary.

The remaining constants A_1 and A_2 are found by using the boundary conditions at entry

$$\int_0^\alpha \sigma_r r_0 d\theta \cos \theta = 0$$

The mean drawing stress may be shown to be

$$\sigma_m = k_1 c_1 \ln \left(\frac{r_0}{r_1} \right)$$

For the *frictionless drawing case* ($m_0 = m_\alpha = 0$), it may be shown that

$$\sigma_m = 2k_* \ln \left(\frac{r_0}{r_1} \right)$$

where $k_* = \{(1-f)k_1 + fk_2\}$

A digital computer was used to perform the iterations and numerical integration involved in this work.

EXPERIMENTAL DETAILS

The bi-metallic tubes consisted of steel/brass and steel/copper combinations formed by the implosive technique, developed by Blazynski and Dara^{1,2}. The materials were supplied as

- (i) Steel—hot drawn tube
- (ii) 70/30 Brass—cold drawn tube
- (iii) Commercially pure copper—cold drawn tube

The brass and copper tubes were annealed at 650°C for 45 min in an air circulated furnace. The tubes were then machined to the required size for the implosive welding operation. After successful welding, the tube bores were machined to various sizes to provide a range of reduction.

The die used had a semi-angle of 7.5°. A range of plugs was used to match the range of tube bores in order that close-pass drawing might be achieved (Table 1 gives tube and plug sizes). At first a soap based lubricant was used, but later ROCOL RTD, an extreme pressure, cutting lubricant, was successfully employed.

The friction drag on the plug was measured by means of a strain-gauge load cell placed at the end of the plug-bar. For steel, with ROCOL, a friction factor of 0.026 was calculated. For copper, the friction measurements were unsatisfactory and erratic.

TABLE 1 Tube dimensions and materials

No	Before drawing (mm)			After drawing (mm)			f	Strain e
	Tube od	Interface dia	Tube id	Tube od	Interface dia	Tube id		
1	56.1	51.5	47.0	53.8	50.3	46.87	0.505	0.314
2	56.1	50.35	47.0	53.8	49.42	46.87	0.632	0.314
3	56.1	51.5	47.0	53.8	50.12	46.5	0.505	0.255
4	56.1	50.35	47.0	53.8	49.19	46.5	0.632	0.255
5	55.0	50.4	45.0	53.8	49.53	44.5	0.46	0.082

Accordingly a friction factor of 0.06 was used, as suggested by Avitzur⁸. The same friction factor was assumed for brass.

The tubes were drawn on a 50 ton single ram-type Hydraulic Drawbench at a speed of 5 cm/s. The drawing load was measured by sensing ram pressure via a Kistler pressure transducer. Prior to drawing, the stress-strain curves for the three materials were established (figure 4). The Watts and Ford⁹ plane strain compression test was used and suitable test pieces, in the form of flat strips, were machined out of the tube walls.

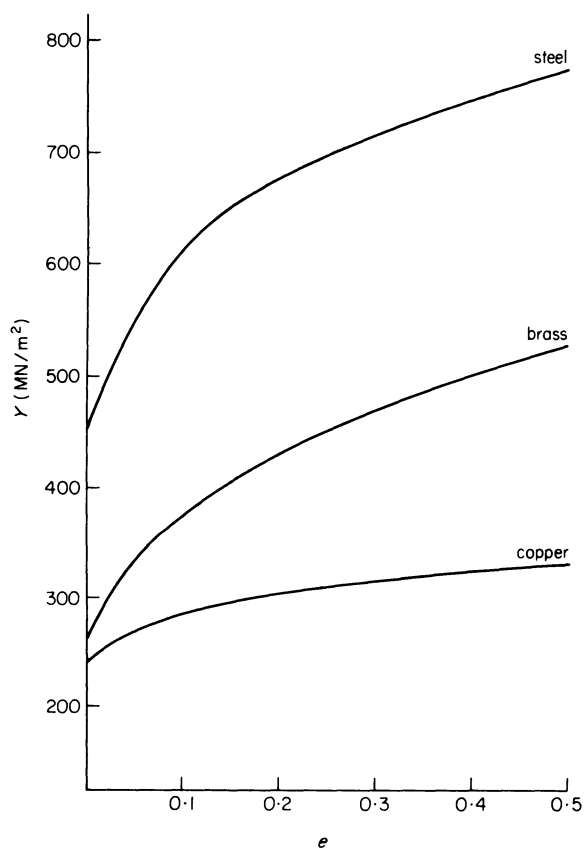


Figure 4. Stress-strain curves for the alloys employed.

DISCUSSION OF RESULTS

Some indication of the practical problems encountered in the drawing of duplex cylinders is afforded by figure 5. Figure 5(a), for instance, shows that, immediately after welding, the compound cylinder is not of a completely uniform diameter

along the whole of its length, and that, therefore, some non-uniformity in processing has to be expected. On the other hand, however, a reduction in the diameter at the lower end of the tube can be utilized in swaging the workpiece in preparation for the actual drawing.

Pronounced differences, in mechanical properties of the constituent materials, may result in differential drawing shown in figure 5(b). Although, at least within the range of employed materials and deformations, no actual failure of a component was registered, it is clear that conditions for such a failure can easily arise. They may be related to the quality of the weld (figure 5(c)) and to the pattern of flow of the respective materials in the die.

The relative behaviour of the different alloys depends, first of all, on their mechanical properties and, secondly, on the operative volume fraction f . It is for this reason that the summary of the preliminary results, illustrated by figures 6 to 8, is based, primarily, on the respective effects of friction, as defined by the factor m , and the material content, as reflected by f .

If the drawing stress is plotted against total strain for all the material combinations, irrespective of the values of f and m (figure 6), it becomes clear that, as in a conventional single tube drawing, the level of energy involved increases with strain. The scatter of experimental points, shown in the figure is caused by both the variation in the respective volume fractions employed, and by the actual differences in the material properties. It is, however, clear that, within the range of experimental conditions used, the variations in the properties of respective material combinations are less pronounced at higher strain levels than at lower ones. In fact, the effect of neglecting f and the mechanical properties, for a given strain, introduces a scatter lying within a range of 34 per cent to 17 per cent for strains ranging from 0.1 to 0.3.

Figures 7(a) to 7(c) show the degree of correlation between the theoretical curves plotted on the basis of the developed analysis, and the experimental points obtained in the course of drawing. All the 0.082 strain points, in these figures, relate to a volume fraction $f = 0.46$. In view of what has already been said about the limitations of the Shield's type of analysis, it is not surprising that the theoretical curves underestimate the magnitude of the mean drawing stress at these low strain levels. End effects at the entry to and exit from the die form a substantial portion of the drawing stress. Bulging may well

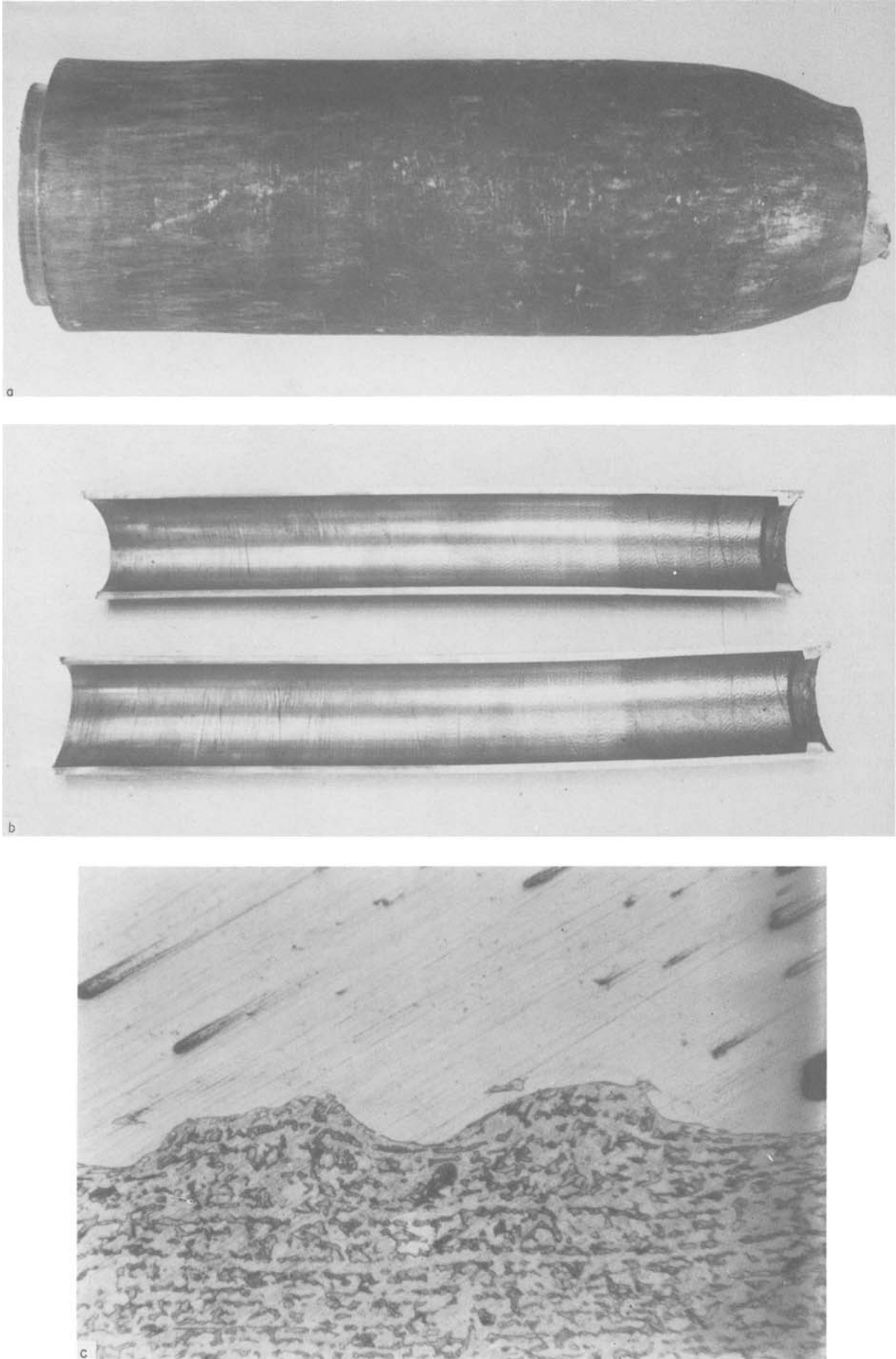


Figure 5. (a) Implosively welded bi-metallic tube prior to drawing. (b) Sectioned, drawn, tube showing differential end drawing effects. (c) Typical interfacial weld (copper-MS).

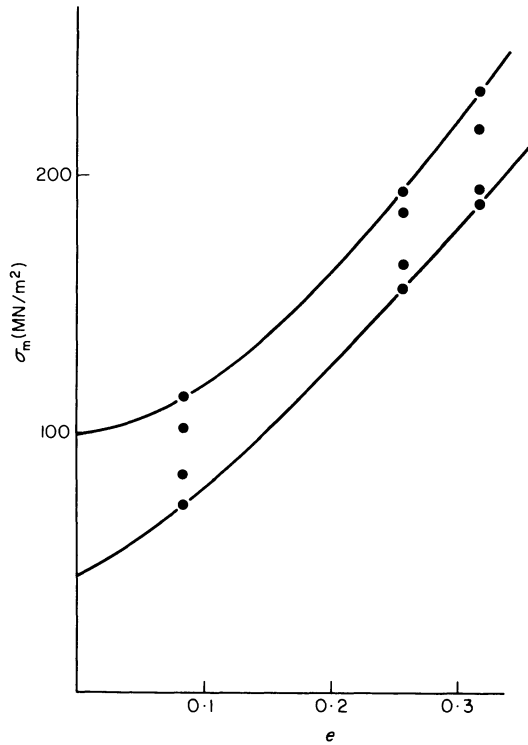


Figure 6. Drawing stress-total strain curves. Envelope of experimental results for all drawing combinations.

intrude at low strains, and this too will increase the actual stress required to effectively deform the tube.

The higher strains of 0.255 and 0.314 are for volume fractions of 0.505 and 0.632 respectively. It can be seen that at these higher strains there exists a much better correlation between the theoretical and experimental results since the end-effects form only a small proportion of the mean drawing stress. The scatter is of the order of ± 12 per cent. Since, after all,

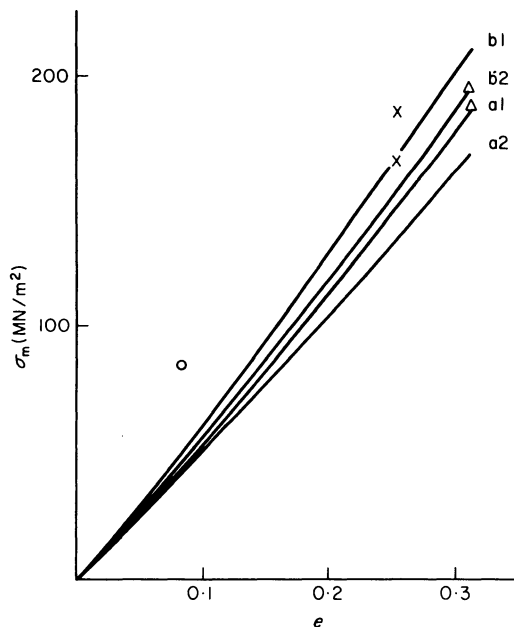


Figure 7. (a) Mean drawing stress against strain (copper on steel)

'a' $m_0 = 0.026, m_\alpha = 0.06$
 'b' $m_0 = 0.06, m_\alpha = 0.06$
 '1' $f = 0.46, '2' f = 0.632$

Experimental points 'O' $f = 0.46$, 'x' $f = 0.505$, ' Δ ' $f = 0.632$.

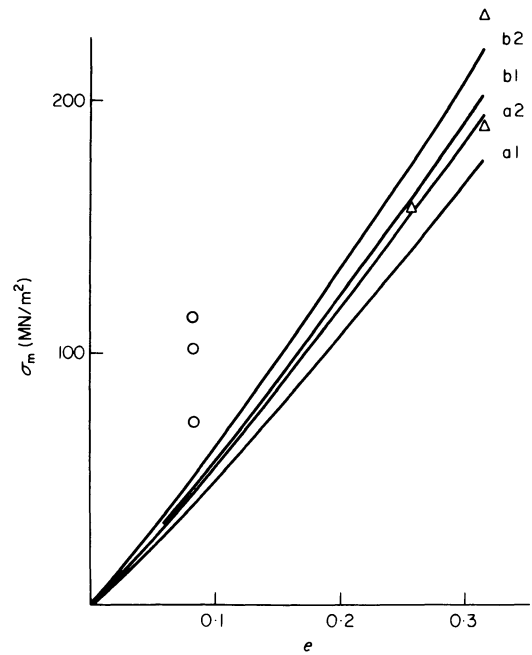


Figure 7. (b) Mean drawing stress against strain (steel on copper)

'a' $m_0 = 0.06, m_\alpha = 0.026$
 'b' $m_0 = 0.06, m_\alpha = 0.06$
 '1' $f = 0.46, '2' f = 0.632$

Experimental points 'O' $f = 0.46$, ' Δ ' $f = 0.632$

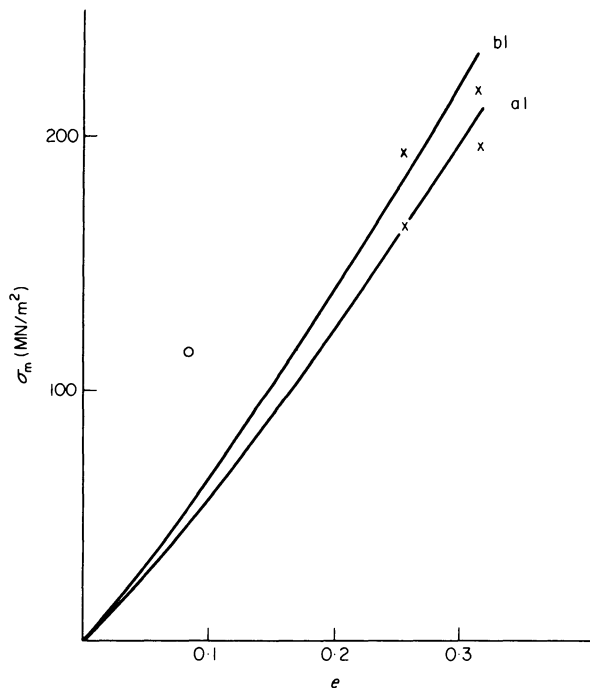


Figure 7. (c) Mean drawing stress against strain (brass on steel)

'a' $m_0 = 0.026, m_\alpha = 0.06$
 'b' $m_0 = 0.06, m_\alpha = 0.06$

it is at the level of higher strains that the process is of industrial interest, it is clear that the proposed theory, in spite of its limitations, does provide a useful indication of the required drawing loads.

In general, two major causes of potential defects are tensile failure of the tube material at the exit

from the die, and the interfacial shear stress in excess of the yield stress in shear of the weaker material. It is interesting to note that in the processing of an unwelded rod-sleeve combination, carried out by Osakada⁶, it was in fact the low interfacial shear strength that contributed to the tensile failure of the material.

In the case of compound cylinders, tensile failure can be expected only at relatively high strains. Failure in shear, on the interface, is less likely, but could be expected for high values of the ratio k_2/k_1 (figure 8), and for a high volume fraction f of about 0.6.

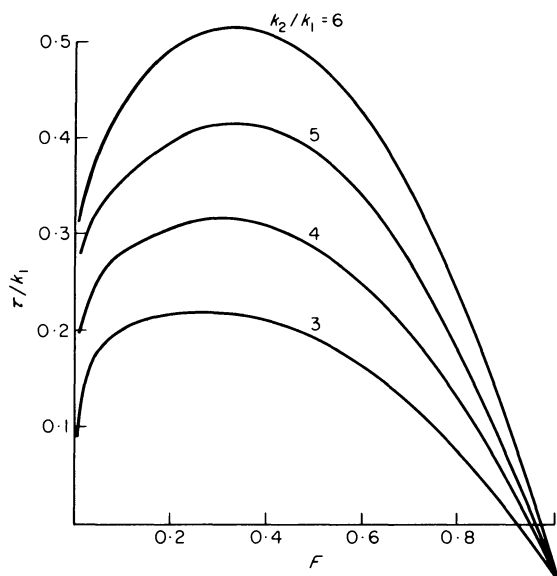


Figure 8. Variation of τ/k_1 with volume fraction.

Neither of these modes of failure was encountered in the present investigation, but the conditions leading to either can be predicted from the proposed theory. The theoretical plot in figure 8 indicates therefore that the variation in interfacial shear stress with f , the volume fraction, of the cladding is clearly defined, and that high interfacial shear stress is promoted by large k_2/k_1 ratios, large values of the frictional factor m_α and low values of m_0 . Homogeneity of deformation is improved by keeping the ratio k_2/k_1 as near unity as possible, m_α as small as possible, and, in theory, by making m_0 large. In practice however this would result in a poor surface finish in the bore of the tube, and, possibly, in a pick-up on the plug.

The most important parameter is nevertheless the ratio k_2/k_1 , whereas m_0 and m_α are of secondary importance. Generally, if the ratio has to be unavoidably large then the interfacial shear strength is likely to be important in deciding whether or not homogeneous deformation is possible, but, since the present work involves implosively welded tubes, the interfacial shear strength is already automatically maximized, and therefore good chances are created for homogeneity to occur. It should be borne in

mind, however, that the weaker the interfacial bond, the more likely it is that homogeneity of deformation will be lost.

CONCLUSIONS

The results of the preliminary investigation of the theoretical aspects of cold-drawing, previously implosively bonded bi-metallic cylinders, indicate that the proposed model of the mechanics of the operation provides a useful insight into the pattern of material behaviour. Reasonably good correlation for calculated and measured drawing loads exists at higher strains which are of greater industrial interest.

In the examined range of materials and sizes no failure in the weld, during drawing, has been observed.

ACKNOWLEDGMENTS

The work was carried out in the Department of Mechanical Engineering, University of Leeds.

The authors wish to thank the Science Research Council for the grant, enabling the purchase of the draw-bench, and for the financial support given to S. Townley.

REFERENCES

1. T. Z. Blazynski and A. R. Dara. A comparison between the implosive and explosive systems for welding of duplex cylinders. *Proc. 3rd International Conference of the Centre for High-Energy Forming*, University of Denver, Colorado (1971).
2. A. R. Dara and T. Z. Blazynski. Explosive welding systems for the manufacture of compound cylinders. *Proc. International Conference on 'The use of high energy-rate methods for forming, welding and compaction'*. University of Leeds (1973).
3. R. R. Arnold and P. W. Whitton. Stress and deformation studies for sandwich rolling hard metals. *Proc. Instn. Mech. Engrs.* (1959) 173, 241.
4. A. G. Atkins and A. S. Weinstein. The deformation of sandwich materials. *Int. J. Mech. Sci.* (1970) 12, 641.
5. B. C. Whitlock and J. M. Alexander. Extrusion of bi-metallic strip from separate containers. *Proc. Instn. Mech. Engrs.* (1969-70) 184, part 1.
6. K. Osakada, M. Limb and P. B. Mellor. Hydrostatic extrusion of composite rods with hard cores. *Int. J. Mech. Sci.* (1973) 15, 291.
7. R. T. Shield. Plastic flow in a converging conical channel. *J. Mech. Phys. Solids* (1955) 3, 246.
8. B. Avitzur *Metal Forming: Processes and Analysis*. McGraw-Hill Book Co. (1968).
9. A. B. Watts and H. Ford. On the basic yield stress curve for a metal. *Proc. Instn. Mech. Engrs.* (1955) 169, 1141.

HOT FORMING

TRANSVERSE DEFORMATION IN SECTION ROLLING AND FORGING

by

G. W. ROWE*, I. M. DESAI† AND H. S. SHIN*

SUMMARY

This paper presents an apparently new approach to a theoretically-based method of designing roll profiles, and roll passes, for production of bars with specified transverse sections. Simple examples are given for forged and rolled Tee sections.

Circular-mesh grids have been etched on polished cross-sections before rolling so that subsequent deformation can be measured. Slip-line fields, deduced from such measurements, have been compared with fields for plane-strain forging, using identical geometric shapes. They have also been compared with maximum shear-stress trajectories obtained from photo-elastic measurements.

It is proposed that photo-elastic experiments can be combined with plasticity theory to provide a basis for tool design for rolling and for forging.

INTRODUCTION

Designing roll profiles, to produce bars of given cross-sectional shape, is still a largely empirical art. Computers may be used for tool design when the over-riding consideration is the stress in the tool, but in most shape-rolling this is of less importance than the flow of metal and the homogeneity of the product. It is not usually possible to produce the final desired shape and uniformity in a single pass, but the selection of optimal intermediate passes requires considerable skill and experience. There are no fundamental guide-lines, and empirical trials are very expensive.

The purpose of this paper is to suggest a simple and inexpensive technique, combining photo-elastic experiments and plasticity theory, to develop shear patterns that can be used as an aid to the prediction of flow in forging or in rolling. Two simple examples are given in which the results can be checked by reference to known solutions. These relate to the compression of a rectangular bar, and to the formation of Tee shapes by cold rolling. These, in fact, formed part of a more general study¹ leading to a rolling process for production of complete surgical needles², and to a forging sequence for turbine rotor discs³.

EXPERIMENTAL TECHNIQUES

Assessment of plastic deformation

All plastic flow experiments were conducted with commercially pure aluminium that had been cast as

75 mm (3 in) diameter billets and extruded to 25 mm (1 in) bar at room temperature, using good lubrication. The starting material was thus reasonably uniform and fully strain-hardened.

To examine transverse flow, each bar specimen was cut into halves, and the cut faces were carefully lapped flat and accurately normal to the axis. They were then polished with diamond paste to a Roughness Average (Ra) value of less than 0.02 μm (1 μin), taking care not to round off the edges. After degreasing, one surface was coated with a thin uniform film of a commercial photo-resist polymer (Kodak KMER) by a standardized spinning procedure. It was then prebaked and exposed to ultraviolet light through a master negative of a circular mesh grid. The image was developed in accordance with the manufacturer's instructions, followed by washing and final baking to give strength to the film. The specimen surface was then swabbed lightly with cotton wool moistened with 10 per cent NaOH solution to etch the grid pattern.

The gridded face and the matching polished face of the other half-specimen were carefully re-located, and the whole specimen was firmly clamped in a narrow jig that maintained close contact at the interface during subsequent deformation. The transverse strain pattern produced by rolling or forging was then studied by recording the orientation and aspect ratio of each ellipse formed from an original circle. The characteristic mesh diameter was 0.5 mm (20 thou) with 0.05 mm line thickness, and measurements were made at a magnification of 50 \times , using a

* Department of Mechanical Engineering, University of Birmingham

† Department of Industrial Metallurgy, University of Birmingham

Zeiss MP320 measuring projector with a two-co-ordinate stage and 0.001 mm digital read-out, fed to a punched tape.

Photo-elastic measurements

The specimens for photo-elastic compression studies were prepared from commercial stress-free epoxy resin sheet 6 mm (1/4 in) thick. Discs of this material were cut to precisely scaled dimensions and the two flat faces were carefully lapped parallel to within 0.2×10^{-3} in over a 1 in face length. A jig was used to keep the faces normal to the sides. If a greater tolerance in parallelism was allowed, the pattern of stress became asymmetric.

The prepared specimens were loaded against accurate platens with interposed foils of polytetrafluoroethylene. The assembly was located on a large photo-elastic bench, and the strain-birefringence pattern was observed at 12 times magnification. Quarter-wave plates were used to provide circularly-polarized light for study of the isochromatic fringes (of equal stress increment). For most of the work, however, the main interest was in the isoclinics (lines of equal maximum-stress inclination). These extinction lines were traced after each 10° synchronous rotation of the polarizer and analyser, and also, where necessary, after smaller increments. Thus a complete field of principal stress directions was built up, in a conventional way. The pattern of maximum shear stress could be derived from this, as described below.

EXPERIMENTAL PROCEDURE AND RESULTS

Simple compression of rectangular aluminium bar

The first experiment is designed to check the validity of interpretation of the deformed circular grid pattern, using a compression of a rectangular block, under sticking friction conditions, for which the slip-line field solution is well known. An aluminium block was machined carefully to rectangular form, and cut into halves which were lapped and polished on the transverse interface, as described above. A circular-mesh grid was then etched on one surface, and the reassembled block was compressed between flat degreased platens, with 5 per cent height reduction to give a final width:height ratio of 2.4:1. Figure 1 shows the deformed grid after this com-

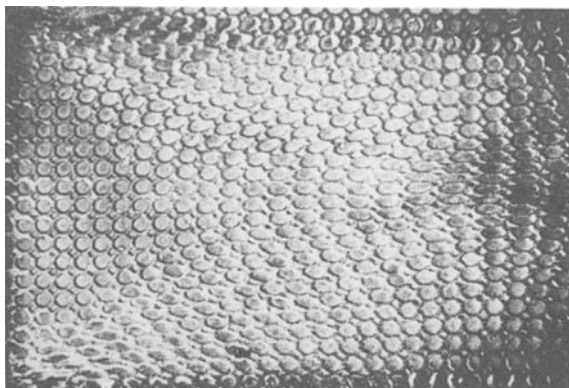


Figure 1. The deformed circular-grid pattern on an aluminium specimen compressed between flat, dry platens by 5 per cent reduction in height under plane-strain conditions to a width:height ratio of 2.4:1.

pression. The elliptical form of some of the meshes is clearly visible.

The major axis of each ellipse was recorded, using the projection microscope, and short lines were drawn at 45° to this axis at the centre of each ellipse, to specify the local directions of maximum shear strain. The result is shown in figure 2. Since the

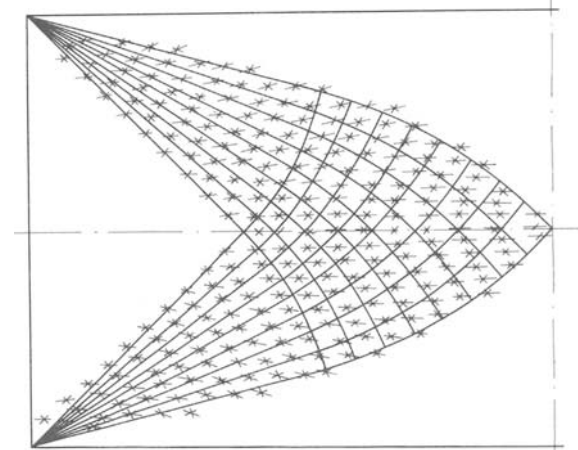


Figure 2. The maximum shear-strain directions deduced from figure 1, and shown as short crosses superimposed on the principal strain directions. The full lines represent the theoretical slip-line field for sticking friction.

aluminium was homogeneous and isotropic, these lines will be parallel to the directions of maximum shear-strain rate and maximum shear-stress. They are thus directly comparable with the slip-lines. The slip-line field for sticking friction, with this 2.4:1 ratio of width to height, is superimposed on the experimental shear direction in figure 2.

There is very close agreement between the experimental and theoretical directions over most of the specimen. Also the rigid zone boundaries, beyond which ellipticity was not detectable, were clearly related. In the vicinity of the singularities of the field, at the specimen corners, the agreement was less precise. This may be due to the idealization of a point centre to the fan shaped region, but is probably attributable to the finite deformation (5 per cent) necessary to develop discernable ellipses. Further experiments with larger deformation increments¹ show that as the deformation is increased the shear pattern becomes less well defined, as would be expected.

These results establish confidence in the technique, and also indicate the errors to which it may be susceptible. If the specimen is annealed, instead of being fully work-hardened to match the slip-line field assumption, the field will spread over a wider area, as has been shown elsewhere^{4,13}, but is of a basically similar form.

Rolling of tee section from round stock

The second example is of a more complex shape produced by rolling round aluminium bar between two rolls; one of which is cylindrical, and the other of which has a deep parallel-sided circumferential groove. In rolling, the specimen is elongated, so that it is not possible to maintain two flat transverse

surfaces of a split billet in close contact. The etched-grid technique cannot be directly employed, so it is necessary to restrain the specimen in some way. Various methods are possible, but the most satisfactory found in this work was to insert the cylindrical billet, after sectioning and reassembly, in a closely-fitting rectangular hole in a steel plate. The thickness of the plate is chosen so that it will pass easily between the rolls but will support an adequate meridian zone at the ends of the specimen.

In this way, it is possible to roll with no significant longitudinal extension. The results are therefore approximately comparable with plane-strain compression fields. The general pattern of shear-stress trajectories is however similar, whether the sample is restrained longitudinally or not. This is known, for example, in the analogy between the slip-line fields for strip drawing and for forging⁵. It was illustrated in the present work by examining the macrostructure of cross-sections rolled with and without constraint. Figure 3 shows the deformed regions with 10 per cent

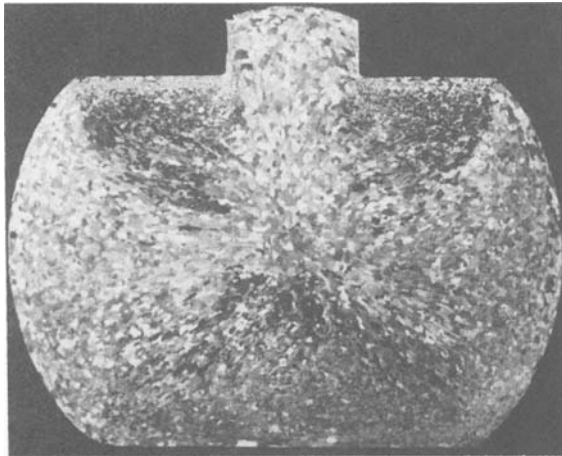


Figure 3. A macro-etched section of a round bar reduced in height by 10 per cent during rolling between a smooth roll and one carrying a deep circumferential groove. The dark regions show the heavy deformation zones, which may be compared with figure 4.

reduction in total height. These may be compared with the slip-line field areas in figure 4.

Slip-line fields deduced from circular-mesh grids

In figure 2, the maximum shear-stress directions were compared with a known slip-line field, but in the present example the slip-line field is not known beforehand, even for two dimensional forging. To deduce a slip-line field from the orientations of the ellipses, as measured, it is necessary to establish an objective procedure, uninfluenced by prior knowledge of likely fields. This can be done by dividing the whole diagram into suitable small squares; selecting any one, well within the specimen, and starting the field by drawing a line right across that square, parallel to the experimentally-determined maximum shear-strain direction. Then, starting from the point of intersection of this line with the square-mesh boundary, another line is drawn right across the adjacent square, parallel to the measured shear-strain direction nearest to that square, or preferably to an

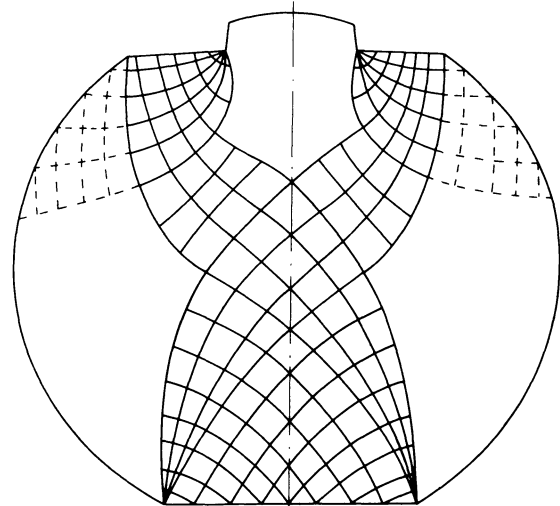


Figure 4. The slip-line field deduced from a deformed circular-mesh grid on a transverse section of an aluminium bar rolled with 7.5 per cent reduction in height between a smooth lubricated roll and a roll carrying a deep circumferential groove.

interpolated direction. In this way a complete network of shear-strain directions is built up throughout the deformed region of the specimen.

Regions in which the circles remain circular are taken to be the dead or rigid zones, although the precision with which they can be determined depends on the sensitivity with which ellipticity can be detected. Attempts are being made to define the plastically-deformed regions more accurately by use of an etch-pit technique to reveal dislocations⁶.

Figure 4 shows the slip-line field deduced from an etched grid on a transverse section of a bar rolled with longitudinal constraint, with a total height reduction of 7.5 per cent. The corner fans have been extrapolated in the vicinity of the singularities because, as explained earlier, the technique is unreliable immediately adjacent to a corner.

It will be seen that the part of the field reaching from the cylindrical roll to the centre closely resembles the accepted field for forging, with height approximately twice the platen breadth, under good lubrication conditions. The field is not, however, quite correct, because the lines do not all intersect orthogonally. The differences from 90° intersection are fairly small and may be attributable to the finite (7.5 per cent overall) deformation imparted.

The upper part of the field has two components. One is a continuation of the forging-type field reaching from the flat surface across the specimen to an opposite shoulder. The other is an indentation-type field immediately below the shoulder and seen more clearly in figure 5, where it is completely separated. The latter figure is taken from a specimen rolled with only 5 per cent total height reduction.

It is well known that in forging⁷ or in strip drawing⁸ and wire drawing⁹ there is a transition from a deformation pattern right across the specimen to a bulge formation when the height greatly exceeds the platen breadth. (There is also an analogy with zero-rake cutting). Figure 5 appears to be incomplete because the lower plastic zone is completely constrained by rigid material. Presumably it does connect

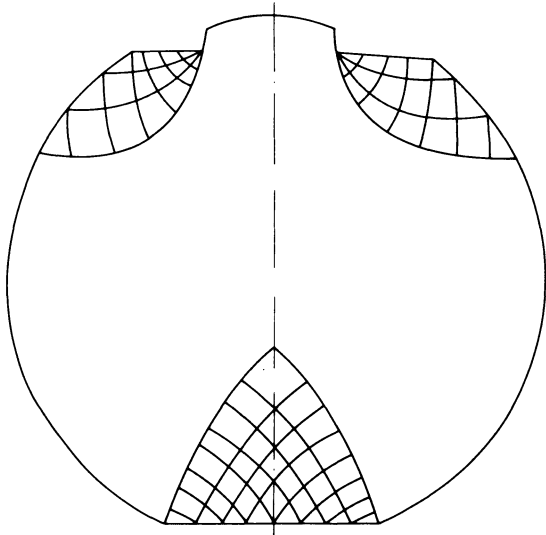


Figure 5. A slip-line field (probably incomplete) deduced from an experiment identical to that related to figure 4 but with 5 per cent reduction in height.

with the upper zones, but through a region where the shear strain is too small to detect.

Comparison may be made between figure 4 and figure 3. The general shape of the dark deformed zones in figure 3 resembles the slip-line field area of figure 4 although the bulge regions do not appear. This is probably because of the heavier reduction (20 per cent) in the macro-section. Slip-line field theory for forging predicts this bulging only with light reductions and large height:width ratios.

Considering figure 4 in more detail, it can be seen that the lines of one family do not always keep to the same angular separation where they are cut by lines of the other family. This is contrary to Hencky's first theorem and indicates some error, possibly due to the finite deformation. Despite this minor inaccuracy, there is clear evidence of an important difference from the commonly-accepted slip-line fields that assume constant coefficient of friction at a tool face.

The angle θ , at which the slip-lines meet a tool-face boundary, is determined by the ratio of the friction stress (μq) to the shear yield stress k according to the equation $\cos 2\theta = (\mu q)/k$, in which the coefficient of friction μ is assumed to be a constant of proportionality. In figures 4 and 5, the angle θ clearly changes steadily with distance from the centre. On the plane of symmetry, the angle is 45° , in accordance with theory, to balance the shear stresses exactly. The pressure distribution q can be calculated from the slip-lines in figure 4, starting from the boundary adjacent to the rigid zone. The hydrostatic pressure is p , with a shear yield stress k , and using the Hencky equations $p \pm 2k\phi = \text{constant}$. In this way it is determined that the coefficient of friction increases from zero at the centre to 0.054, 0.107, 0.162 at each intersection, moving outwards along the face, and to $\mu = 0.216$ close to the outer corner. The surface of this cylindrical roll had been polished.

The grooved roll, in contrast, was machined. It can be seen that the slip-lines meet this roll surface at a much steeper angle, close to the theoretical value of 90° for sticking friction. Nevertheless, there is again a decrease in friction towards the central zone.

Figures 4 and 5 show that there is a rigid region beneath the stem of the Tee, and that this diminishes as the overall percentage reduction in height per pass is increased. Because of this rigid zone, the slot does not readily fill. Indeed, a later and more detailed study¹⁰ shows that the strain, in the slip-line field surrounding this rigid zone, is small, and that the effectively dead metal below the slot is significantly greater in area though of similar general shape to that shown in figure 4.

Maximum shear-stress directions shown by photo-elastic study

As has been mentioned above, the plastic deformation of a grid requires a finite overall deformation that may in some regions, especially near the corners, impair the accuracy of the deduced field, and may lead to inconsistencies such as the requirement for mutual orthogonality of the maximum shear-stress directions.

Following a chance observation that a photo-elastic analysis of stress directions, in cross-forging of a cylinder¹¹, resembled the pattern of slip-lines in the lower half of figure 4, a photo-elastic study was made of a compressed disc cut to the geometric form of figure 4 (corresponding to 7.5 per cent deformation from round). The specimen was prepared carefully with flat polished surfaces, as described above, from a sheet of Araldite CT 200. The lower platen was lubricated with a thin film of polytetrafluoroethylene, to simulate the polished roll.

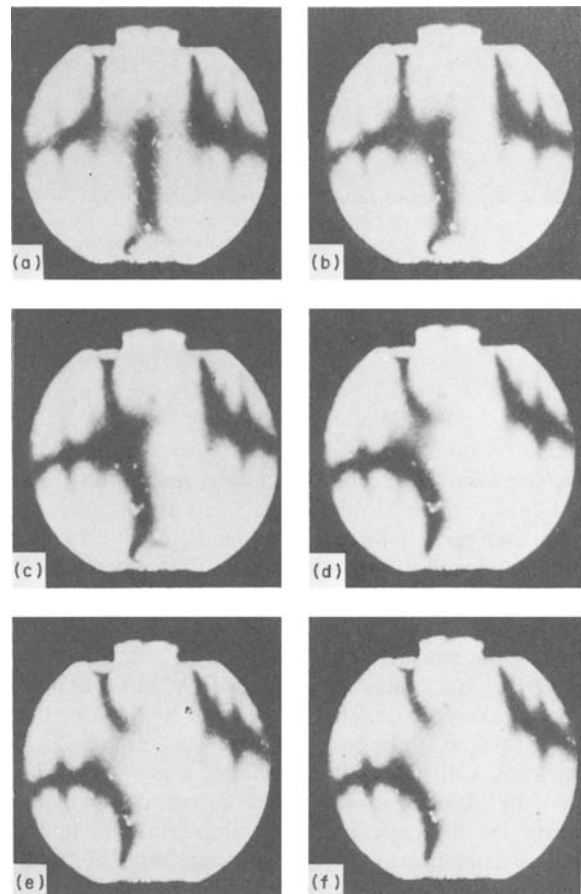


Figure 6. A set of photographs showing the isoclinics at 2° intervals of synchronous rotation of the polarizer and analyser in the photo-elastic experiment.

Using the photo-elastic bench as described above, isoclinics were traced out, under a load equivalent to a nominal platen pressure of 600 lbf/in² (4.1 kN/mm²). A typical set of results is shown in figure 6. As the isoclinics are never very well defined, the results were repeated several times until a symmetrical, and correctly-spaced, pattern was obtained, as shown in figure 7. This field then represented the loci of equal inclination of the principal-stress directions to the reference centre-line. Thus along the 20° line, for example, the principal stresses are always at 20° (and 110°) to the vertical. The maximum shear-stress directions are therefore at 65° and 155°. Short lines with these directions are drawn at intervals on the isoclinics of figure 7. The 0° isoclinics are the

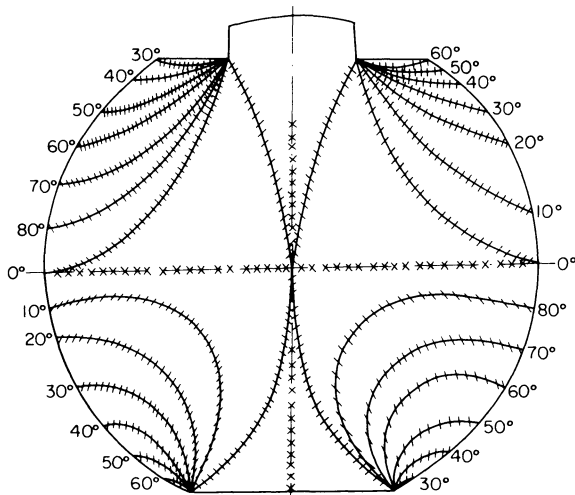


Figure 7. The set of isoclinics at 10° intervals, as drawn from several experiments of the type recorded in figure 6, with maximum shear-stress directions superimposed. The geometric form of the specimen corresponds exactly to the aluminium rolled with 7.5 per cent reduction as in figure 4.

horizontal and vertical axes, so the maximum shear-stress directions are, in these instances, at 45° to the isoclinics.

To construct a pseudo-slip-line field, namely a field of maximum elastic shear-stress directions, a

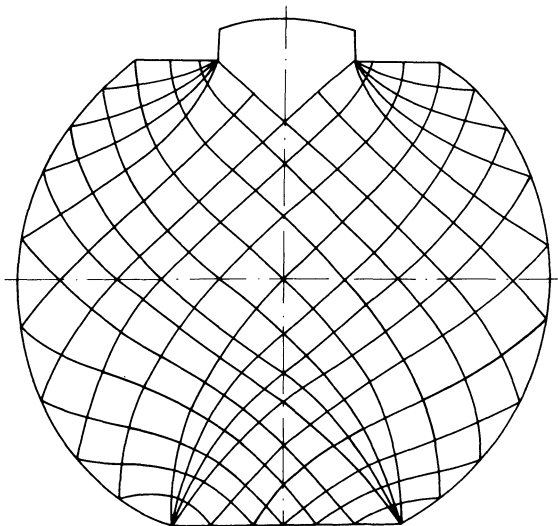


Figure 8. The directions of maximum elastic shear-stress deduced from the isoclinics in figure 7.

procedure analogous to that described above is used. The field is covered with a mesh of small squares, and each shear line is started at an interior square, drawing a line across the square at the inclination shown by figure 7. A line is drawn across the next square from the point of intersection, parallel to the nearest indicated line, so that a continuous network of shear lines is built up. Figure 8 shows the result.

The general similarity between figures 4 and 8 is striking, especially when it is recognized that the dead zones of figure 4 will in fact be elastically stressed.

Considering figure 8 in detail, it can be seen that all the lines meet the free surfaces at, or very near, 45°, as theory requires. Most of the lines intersect at 90°, again in accordance with theory. But some intersections are far from orthogonal, and the lines do not always maintain constant angular separation. Indeed, these apparent errors emphasize the similarity between the elastic and plastic fields (figures 4 and 8). A further similarity of great interest is the steady increase in tangential stress component with distance from the centre outwards along the flat platen.

DISCUSSION

This paper can be considered as a progress report on part of a current wider investigation concerned with prediction of deformation patterns in rolling and forging. The major practical objective is to find a rational procedure for designing tool profiles and passes, or at least aiding their design, when the dominant criterion is homogeneity of deformation¹².

The use of plastically-deformed models is clearly less expensive and time-consuming than full-scale trials, and this method has often been used to study metal flow. The simple compression with sticking friction, illustrated in figures 1 and 2, shows that a close correlation with theory can be obtained. This means of course that if the theoretical slip-line field can be drawn, the real deformation of the non-hardening model can be predicted. It is relatively easy to draw such fields for a variety of simple geometric shapes, and for sequential passes. Similarly good correlation has been found by many workers, using small square grids and other indicators¹³. The circular grids¹⁴ are practically convenient because the major axis is easily located and measured.

There are still problems to be faced in hot forming when there are large temperature gradients or strain-rate gradients that affect flow stress distribution, but a basic knowledge of the flow pattern for uniform flow stress can be a useful guide, considerably reducing the amount of empirical trial necessary. Approximate allowance for the effect of strain hardening can often be made for cold-working applications.

The use of gridded aluminium models itself can, however, involve appreciable costs for the manufacture of tooling, especially for more elaborate studies of optimum pass shapes and schedules. The present work suggests that photo-elastic models can be used to obtain preliminary information.

The two figures 4 and 8 imply that there is a close similarity between the slip-line field, as experimentally determined for a shape not previously

examined, and the elastic shear-stress directions for an identical shape with closely-matched boundary conditions. Examination of the literature shows that a related observation was made by Coker¹⁵, apparently with little comment, as long ago as 1926.

Since we are not here concerned with significantly anisotropic materials, the directions of maximum shear-strain, maximum shear-strain rate and maximum shear-stress will coincide in the plastically-deforming metal, and the above results imply that the directions of maximum shear-stress do not change as a non-hardening metal passes from elastic to plastic deformation. Further plastic deformation will of course lead to changes in these directions, but these are predictable from classical slip-line field theory, utilizing the slip-line fields and hodographs for sequential small deformation increments. The local strain can also be deduced, and, consequently, the strain distribution. If approximate slip-line fields for annealed metals are used^{4,9}, the strain distribution can be used as a guide to the homogeneity of the product in terms of hardness, ductility, or grain size on subsequent heat treatment¹⁶.

The effects of varying the die shape, for example with convex or concave intermediate profiles, and the sequence of passes, can then be examined without the necessity for making expensive experimental dies and practical trials. One example of the prediction of a forging schedule from slip-line field has been given¹⁷ for *butterpat* forging in which the objective is to obtain extensive uniform deformation throughout a cube of metal. The theory showed that improved homogeneity could be obtained if, in a sequence of compressions, the top and bottom of the billet were alternately lubricated, the opposite surface remaining dry.

CONCLUSION

It is proposed that a combination of experimental photo-elastic stress-trajectory determinations with plasticity theory can be used as a basis for rational tool design for maximum homogeneity in forging and rolling. This should considerably reduce the total number of empirical tests necessary for a new design, though some final modifications to match real conditions are likely to be needed.

An interesting subsidiary conclusion from this work is that the coefficient of friction is not constant along the tool face even when a solid lubricant is used. Previous observations of varying coefficient of friction in forging have been explained in terms of trapping of viscous lubricant^{18,19}.

ACKNOWLEDGMENTS

The authors would like to thank Professors S. A. Tobias and D. V. Wilson for provision of facilities in the Departments of Mechanical Engineering and Industrial Metallurgy, respectively. They would also like to thank Mr E. Amini, Dr H. Kaliszer and Mr T. F. Li for useful discussions, Mr Amini for the loan of a photo-elastic bench, Dr Kaliszer for the loan of a digitized metrological microscope, Mr A. Burgess and Mr E. Cottam for help in photography.

REFERENCES

1. I. M. Desai. *PhD Dissertation, Birmingham University*, 1974.
2. I. M. Desai and G. W. Rowe. *Brit. Pat. Appln.* 1973.
3. H. S. Shin. *MSc Thesis, Birmingham University*, 1972.
4. J. A. Newnham and G. W. Rowe. An analysis of compound flow of metal in a simple extrusion-forging process. *J. Inst. Metals* (1973) **101**, 1-9.
5. A. P. Green. Plane strain theories of drawing. *Proc. Inst. Metals* (1960) **174**, 847-64.
6. K. Sato and T. F. Li (1974) Unpublished.
7. A. P. Green. A theoretical investigation of the compression of a ductile material between smooth flat dies. *Phil. Mag.* (1956) **42**, 900-18.
8. R. Hill. *The Mathematical Theory of Plasticity*. Oxford University Press, 1950.
9. R. W. Johnson and G. W. Rowe. Redundant work in drawing cylindrical stock. *J. Inst. Metals* (1968) **96**, 97-105.
10. T. F. Li (1974) Unpublished.
11. Goku *et al. Nippon Kinzoku Gakkai-Si* (1958) **22**, 572.
12. D. L. Figgis. *MSc Thesis, Birmingham University*, 1966.
13. L. E. Farmer and P. L. B. Oxley. A slip-line field for plane strain extrusion of a strain hardening material. *J. Mech. Phys. Solids* (1971) **19**, 369-81.
14. P. V. Vaidyanathan and T. Z. Blazynski. A theoretical method of efficient extrusion die design. *J. Inst. Metals* (1973) **101**, 79-84.
15. E. G. Coker. *Elasticity and Plasticity. Proc. Inst. Mech. Engrs.* pp. 897.
16. H. S. Shin (1974) Unpublished.
17. G. W. Rowe. Recent studies of cold extrusion. *Mecanique, materiaux, electricite* (1972) **26**, 14-19.
18. L. H. Butler. The effect of interposed lubricants on the surface deformation of metals during plastic working. *J. Inst. Metals* (1959) **88**, 337-43.
19. P. F. Thomason. On the plane-strain compression of a rigid plastic material, with an entrapped viscous lubricant on the platen face. *Int. J. Mech. Sci.* (1972) **14**, 279-84.

A STUDY ON PLASTIC WORKING OF ALLOYS IN THEIR MASHY STATE

by

S. FUKUOKA* and M. KIUCHI*

SUMMARY

It is well known that flow stress of metal decreases as its temperature rises. This property is applied to various hot plastic workings. Some working processes are now being developed which include metal deformation in the temperature range over solidus line, such as extrusion casting, ACURAD process and continuous casting at high speed. In order to study these processes, it is very important to get the basic data (little of which is known) on flow stress and flow behaviour of metal in the temperature range over its solidus line.

Bearing this in mind, the authors made compression tests on the alloys of Pb-Sn and Al-Cu in their mashy state (that is in the temperature range over solidus line). The results of the compression tests were as follows

- (1) the relationship between flow stress and the solid fraction was clarified
- (2) according to the decrease of solid fraction from 100 per cent to 75 per cent, the flow stress decreases to $\frac{1}{4} - \frac{1}{5}$
- (3) the alloys in their mashy state fracture in the strain range of 3-10 per cent in uni-axial deformation

From the extrusion tests, the following results were obtained.

- (1) bars were successfully extruded in the reduction of 75-99.94 per cent
- (2) pipes were successfully extruded in the reduction of 74.7-93.4 per cent
- (3) by selecting an optimum temperature, sound products can be obtained at small extrusion force and at high speed.

INTRODUCTION

Various metal products are manufactured through plastic working processes including hot plastic workings which utilize the property of metal that the flow stress decreases as its temperature rises. The productivity of hot plastic workings, however, has been thought to be restricted by the decrease in tool strength with the rise in temperature.

Some working processes are now being developed in which high pressure is loaded on metals in the temperature range over solidus line; but in essence they are casting processes.

Compared with these, the working process that the authors are now developing is essentially different from casting processes. It is the working process of metals, especially of alloys using characteristics of their mashy state.

This working process in its mashy state has the following merits and possibilities:

- (1) working force decreases remarkably
- (2) higher working rate (such as high reduction in extrusion and rolling) is attained

- (3) thin products can be easily made from comparatively big billets or ingots
- (4) direct rolling, extrusion or the other working process of the molten metal through mashy state, is possible with some cooling devices
- (5) products of high quality can be obtained because of the high hydraulic pressure

For the purpose of studying deformation behaviours of metals or alloys, as well as exploring the possibility of new plastic working processes of metal in the mashy state, it is very important to get the basic data on flow stress and flow behaviour of metal in the temperature range over its solidus line. Very little of this data is known. This data is needed in order to understand the above-mentioned casting processes.

From this viewpoint, the authors firstly made compression tests using a cam-plastometer on the alloys of Pb-Sn and Al-Cu in their mashy state (in the temperature range between solidus and liquidus line) as a *fundamental research*, and secondly made extrusion tests in the same state as an *application research*. Some data about the flow stress and the

* Institute of Industrial Science, University of Tokyo

flow behaviour of alloys in their masy state and about the extrusion process in that state are reported.

COMPRESSION TESTS: EXPERIMENTAL APPARATUS AND METHOD

Under the tests, strain rate was fixed at about 0.2/sec. Test pieces were made from Pb, Sn, Al of 99.9 per cent purity and Cu of 99.99 per cent purity. They were melted into alloys in a crucible at about 450°C for Pb-Sn and at about 900°C for Al-Cu. Each alloy was cast in a metal mould and finished into a prescribed shape under a lathe. Two kinds of test pieces were used and their dimensions were 20 mm in diameter, 18 mm in height (only for Pb-Sn), and 26 mm in diameter, 23.5 mm in height (for both alloys). Each end surface was finished as smooth as possible, and the dimension error was within 0.05 mm for both diameter and height. The components of Pb-Sn alloy were Pb : 80.8 per cent and Sn : 19.2 per cent, those of Al-Cu alloy were Al : 94.3 per cent and Cu : 5.7 per cent.

A sub-press was used in order to keep the test piece at a constant high temperature during the test. It was made from SUS42 of JIS (a kind of stainless steel) and its shape and dimensions are shown in figure 1.

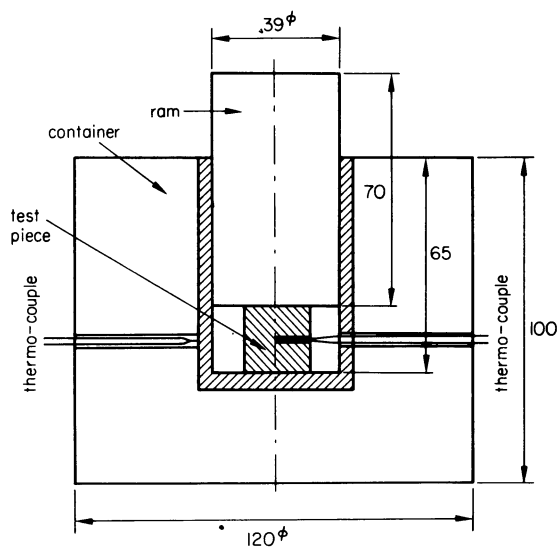


Figure 1. Shape and dimensions of the sub-press.

The procedure of test was as follows. The test piece was set in the centre of the sub-press and heated in the electric furnace. After both the temperatures of the sub-press, and of the test piece, reached the prescribed value, these were kept constant for about 20 minutes. Then the sub-press, containing the test piece, was set in the cam-plastometer and subjected to the compression test. Throughout this process, the temperature of the test piece was measured with a thermo-couple of chromel-alumel which was inserted through a hole of 1 mm diameter, and which reached its centre. The electric furnace was controlled automatically according to the measured temperature of the sub-press. The temperature error of the test piece was assumed within 1°C. During the process, compression force was measured with a load cell and displacement with a differential transformer.

RESULTS AND CONSIDERATIONS

The relationship between temperature and solid fraction

The solid fraction of test piece in the masy state can be obtained from its temperature by using the phase diagram of the alloy. In the phase diagram (figure 2),

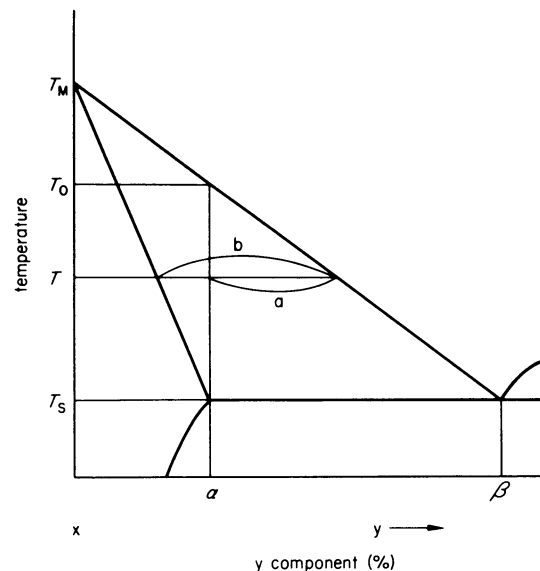


Figure 2. A phase diagram of binary alloy.

the solid fraction ϕ of X-Y alloy containing metal Y of α per cent can generally be written as

$$\phi = 100(T_0 - T)(T_M - T_S)/(T_0 - T_S)(T_M - T) \quad (1)$$

where

$$T_0 = (1 - \alpha/\beta)T_M + (\alpha/\beta)T \quad (2)$$

From Equations (1) and (2), the solid fraction ϕ (per cent) at the temperature T is obtained.

The relationship between stress and strain

When it is assumed that the volume of the test piece is kept constant and it deforms into the shape of column, the mean compression stress (σ) and the mean logarithmic strain ($\bar{\epsilon}$) is calculated as follows

$$\sigma = 4P/(\pi D^2) = (4/\pi)(P/D_0^2)(H/H_0) \quad (3)$$

$$\bar{\epsilon} = \ln(H/H_0) \quad (4)$$

Results on Pb-Sn alloy

Some typical examples of compression stress-strain curves are shown in figure 3, where strain rate is about 0.2/sec. As shown with Pb-Sn alloy, test pieces fracture at the strain range of 3-5 per cent and flow stress decreases suddenly. This tendency does not depend on the solid fraction.

Figure 4(a) shows the relationship between solid fraction and flow stress corresponding to 4 per cent strain. The flow stress of the alloy with 100 per cent solid fraction, that is on the solidus line of the phase diagram, is about a half of that at room temperature. (It is 4.5 kg/mm at 4 per cent strain). Moreover, the

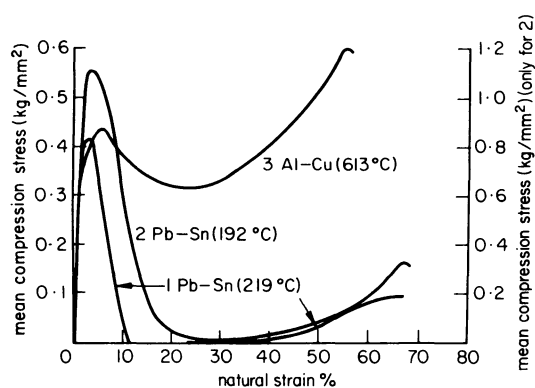


Figure 3. Typical stress-strain curves in uni-axial compression.

flow stress decreases as much as $\frac{1}{5}$ of that of 100 per cent solid fraction alloy, according to the decrease in the solid fraction from 100 per cent to 80 per cent.

Results on Al-Cu alloy

Figure 4(b) shows the relationship between flow stress corresponding to 4 per cent strain and solid fraction. From this result it is seen that the flow stress decreases to about $\frac{1}{4}$ for the change in solid fraction from 100 per cent to 75 per cent, and that it decreases no more for any further decrease in solid fraction. Figure 4(c) shows the change in flow stress for wider range of temperature, where the flow stress at melting point (that is 548°C on the solidus line) corresponds to that of 100 per cent solid fraction in figure 4(b).

Comparison between Pb-Sn and Al-Cu

With regard to the flow stress of these two alloys in their mushy state, the following points may be made

- (1) according to the decrease of solid fraction from 100 per cent to 75-80 per cent in the temperature range over the solidus line, the flow stress decreases to $\frac{1}{4}$ - $\frac{1}{5}$ of that on the solidus line
- (2) over the solidus line, both alloys are apt to fracture at the stage of comparatively small strain in these uni-axial deformation processes, but this tendency is sharper on Pb-Sn alloy.

EXTRUSION TESTS: EXPERIMENTAL APPARATUS AND METHODS

Test pieces

Test pieces were made as follows. Pb and Sn of 99.9 per cent purity, or Al of 99.9 per cent and Cu of 99.99 per cent purity were melted into alloy of Pb-19.2 per cent Sn or Al-5.7 per cent Cu respectively in a crucible, and billets of 41 mm in diameter were cast in metal moulds. These billets were finished into test pieces of 40 mm diameter under a lathe. As shown in figure 5, the lateral length of the column part of the test piece was 40 mm (for Pb-Sn), and a nose was annexed corresponding to 60° and 120° die. At its centre, a hole of 2.5 mm diameter, which reaches the longitudinal centre, was cut to measure the temperature with a thermo-couple of 0.32 mm thickness.

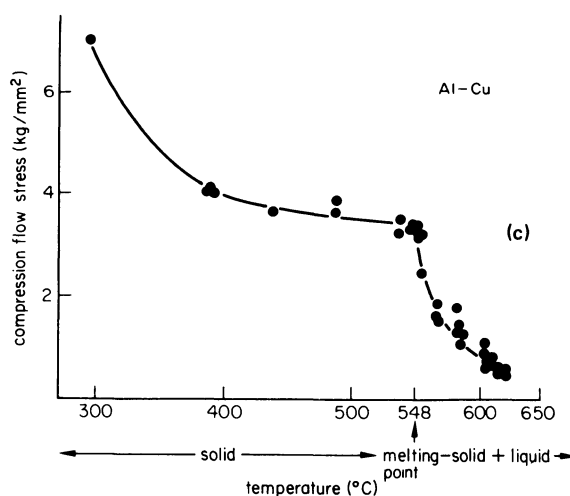
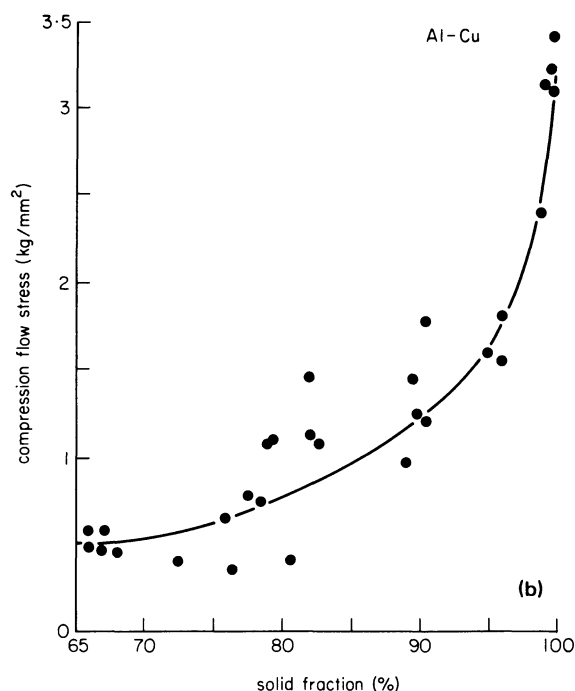
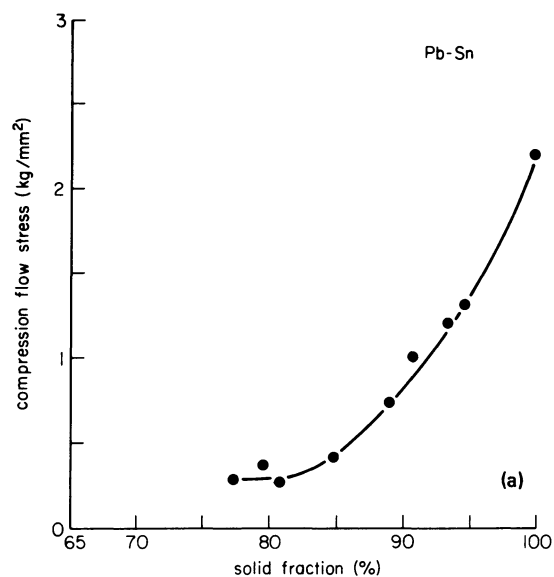


Figure 4. The relationship between solid fraction and flow stress corresponding to 4 per cent strain (a) Pb-Sn (b) Al-Cu (c) The change of flow stress of Al-Cu alloy for wider temperature range.

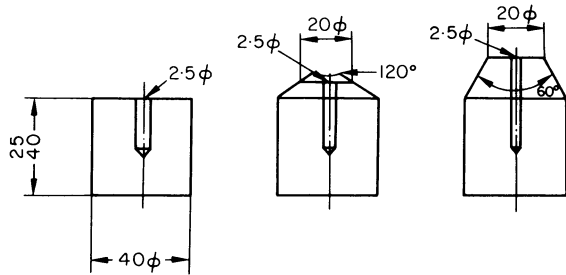


Figure 5. Shape and dimensions of the specimens.

Extrusion apparatus

The extrusion apparatus was a kind of compression testing machine of horizontal type annexed to an electric furnace. It had sub-jacks for keeping the die and the container tight. Maximum working force of the main ram was 100 tons and maximum tightening force of sub-jacks was 20 tons. Maximum speed of the main ram was 70 mm/min. The capacity of the annexed electric furnace was 5 kW for 100 V; its highest temperature being 700°C. The heated parts of the apparatus, except the container, were cooled by means of a water cooling pipe, wound round them. Thus the influence of heat on the load cell was excluded.

Container for extrusion

The dimensions of the container were 120 mm outer diameter and 40 mm inner diameter. It was made of SKD61 of JIS (a kind of steel for hot press working), and finished ground after being hardened on fitting surfaces. Its effective cylinder length was 180 mm.

Dies and punches

The diameter of the dies' holes (orifices) were 20 mm, 10 mm, 1 mm, and the dies' angles were 60°, 120° and 180°. The length of their bearing part was 6-6.5 mm. The punch was 40 mm in diameter and 70 mm in length. The clearance between punch and container was 0.05-0.12 mm. Dies and punches were made of SKD61 and finished ground after being hardened.

Measuring instruments

A load cell was used to measure the applied load and a slide register was used to measure the ram stroke. Temperature was measured with thermo-couples (0.32 mm thick chromel-alumel) inserted in the die, the container and the specimen.

Experimental method

After the temperature of the specimen was raised and kept at a prescribed value for a given length of time by the control equipment of the coupled electric furnace, a tightening force between the container and the die was applied with sub-jacks, and then an extrusion force was unloaded at a prescribed speed by the main jack. The load and the displacement was recorded in a two-pens recorder. Figure 6 shows the experimental apparatus.

EXPERIMENTAL RESULTS AND CONSIDERATIONS

Pressure-stroke curves

Some examples of extrusion pressure-stroke curves are shown in figure 7. In each case a peak force appears at the first stage, and later the short extrusion pressure becomes constant. The larger the die angle, the sharper the peak. This tendency does not depend on the temperature or the fraction of solid.

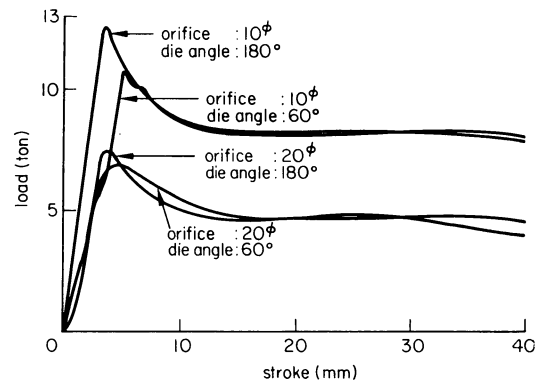


Figure 7. Pressure-stroke curves in extrusion.

Kudo et al.¹ consider this peak to be caused by the formation of dead metal, but, as it is difficult to assume the formation of dead metal for a 60° die angle, other reasons, such as the increase of flow stress accompanying the temperature drop, could be considered. However this needs clarification.

The relationship between extrusion pressure and solid fraction

Figure 8(a) and (b) show the relationship between extrusion pressure and solid fraction, where the extrusion pressure p_m means the constant value after the

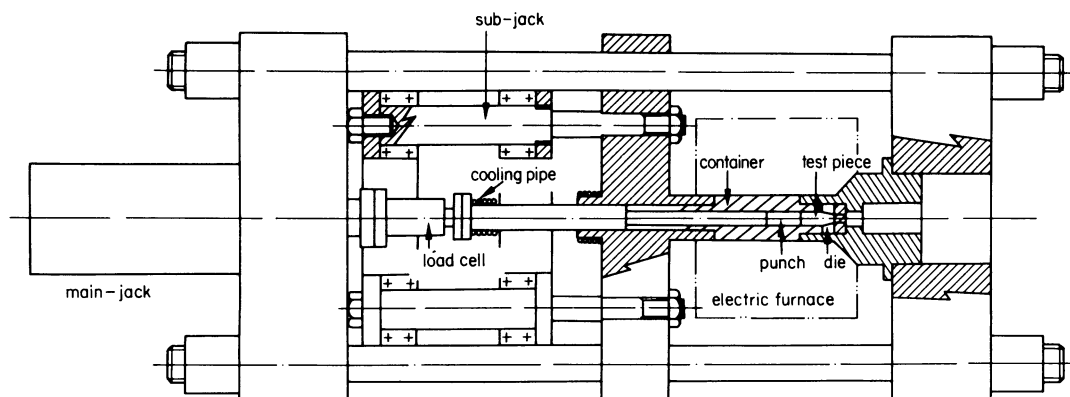


Figure 6. General view of experimental apparatus.

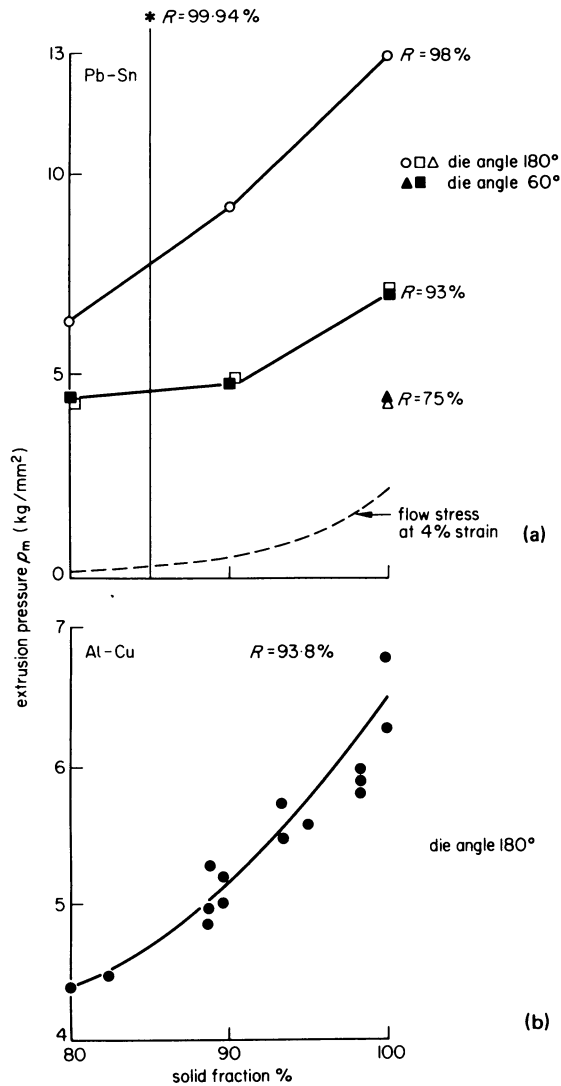


Figure 8. The relationship between extrusion pressure and solid fraction (a) Pb-Sn (b) Al-Cu.

peak appeared in figure 7. Figure 8(a) shows the results on Pb-Sn alloy for three degrees of reduction and figure 8(b) shows the result on Al-Cu alloy for the reduction of 93.8 per cent. The broken line in figure 8(a) shows the flow stress of Pb-Sn alloy at the uni-axial strain of 4 per cent. From these curves, we can state the following facts.

- (1) In the case of Pb-Sn, according to the change of solid fraction from 100 per cent to 80 per cent, extrusion pressure decreases about 30-40 per cent of that of 100 per cent solid fraction, where the reduction is 93.8 per cent (that is the reduction when a billet of 40 mm in diameter is extruded into a bar of 10 mm in diameter), and the decrease reaches about 50 per cent for the reduction of 98 per cent. The decreasing rates of both cases, however, are smaller than that of the flow stress known from the uni-axial compression tests.
- (2) The decreasing rate of extrusion pressure of Al-Cu alloy for the change of solid fraction from 100 per cent to 80 per cent, is about 40 per cent. This is also smaller than that flow stress, known from the uni-axial compression tests.

- (3) These facts might be caused by the difference in flow behaviour of the fluid, held in the specimens between extrusion and uni-axial compression.

The relationship between extrusion pressure and extrusion ratio (the ratio of billet's to product's cross-sectional area)

Figure 9 shows the relationship between normalized extrusion pressure (the ratio of extrusion pressure to flow stress at 4 per cent uni-axial strain) and extrusion ratio. The broken lines in the figure indicate the values theoretically calculated by Kudo et al.¹, using the *slip-line-field* method. The uppermost line corresponds to the value when both the surface of the

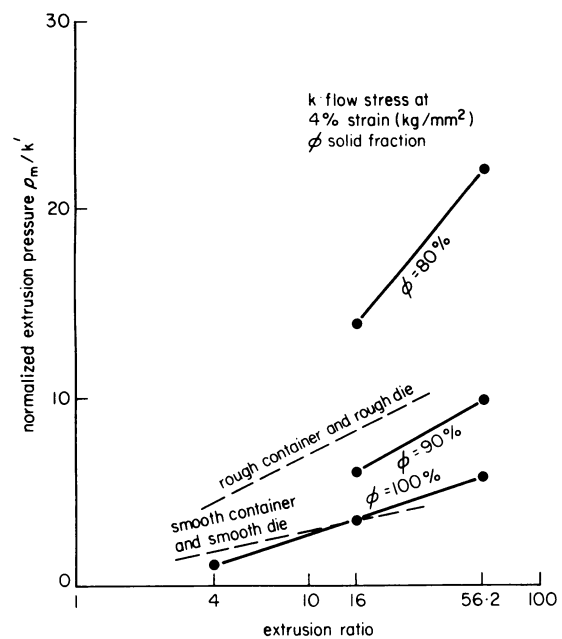


Figure 9. The relationship between dimensionless extrusion pressure and extrusion ratio.

container and of the die are rough, while the lower line corresponds to the value when these are smooth. Experimental lines for 90 per cent and 100 per cent of solid fraction are between two theoretical lines. This suggests that the flow behaviour of metal, in its mashy state with 90-100 per cent solid fraction, does not differ in essence from that in the temperature range for hot plastic workings. Thus, the experimental results agree approximately with those obtained by the *slip-line-field* method. The experimental result for 80 per cent of solid fraction is far apart from the analytical ones. This suggests that the fluid part of the specimen, held among crystallites, flows in a different way for extrusion from uni-axial compression, and that for low solid fraction, as we have here, the flow stress values obtained from uni-axial compression tests may not be applicable to analysis of the extrusion.

Effect of die angle

The relationship between extrusion pressure and die angle is shown in figures 10(a), (b) and (c). Figure

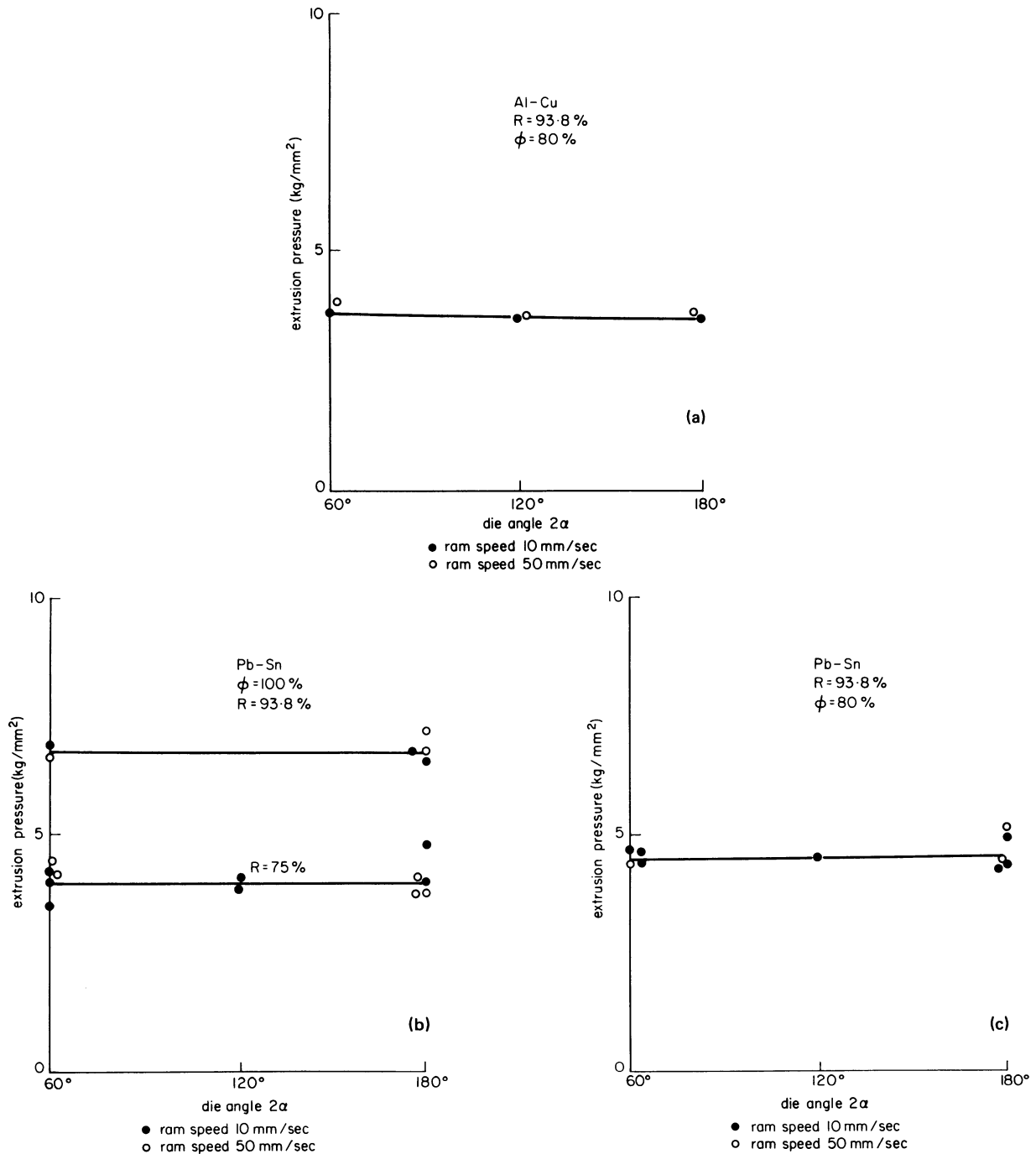


Figure 10. The relationship between extrusion pressure and die angle (a) 100 per cent solid fraction for Pb-Sn (b) 80 per cent solid fraction for Pb-Sn (c) 80 per cent solid fraction for Al-Cu.

10(a) corresponds to the extrusion in 100 per cent of solid fraction, figure 10(b) to 80 per cent for Pb-Sn alloy and figure 10(c) to 80 per cent for Al-Cu alloy. It can be seen from these figures that the extrusion force is not affected by die angle for any amount of solid fraction, and that both friction pressure and shearing work are small.

Effect of extrusion speed

In the speed range (ram speed is 10-50 mm/min) of these experiments, the effect of the extrusion speed cannot be recognized (figures 10(a), (b) and (c): black points correspond to 10 mm/min and white points to 50 mm/min.)

Properties of the products

Products of 10 mm diameter (the reduction being 93.8 per cent) of the solid fraction under 80 per cent had side cracks for both alloys of Pb-Sn and Al-Cu, while, for the solid fraction of 90-100 per cent, products of Pb-Sn had side cracks, but those of Al-Cu did not always have them. For 100 per cent solid fraction all the products of Pb-Sn were good, but some Al-Cu products had side cracks. The products of Al-Cu alloy were put to tension tests, and the results were as follows

- (1) the tensile strength of the sound products extruded was 16-20 kg/mm², and it did not depend on the solid fraction

- (2) the tensile strength of the extruded products was higher than that of the billet as cast by 50 per cent and over
- (3) while the billet fractured in a brittle way, the sound products fractured after a certain elongation had occurred.

Further application of extrusion of mashy alloy

As stated above, the flow stress of alloys in their mashy state, and the characteristics of extrusion of mashy alloys, were grasped to some extent. The authors then tried to apply this working process to more difficult production.

Firstly the upper limit of reduction (or extrusion ratio) was investigated, and the authors succeeded in extrusion of 99.94 per cent reduction (or 1600 in extrusion ratio) for Pb-Sn alloy at solid fraction of 85 per cent. The product was of a 1 mm diameter (its billet was 40 mm in diameter) and it was sound in mechanical properties. By using some cooling devices and selecting optimum temperature, it may be possible to carry out extrusion in higher reduction.

Secondly extrusions of Al-Cu tubes were attempted. In these cases, a billet with a hole whose diameter agreed with that of the tube produced, and a punch with a mandrel of the same diameter, were used. The outer diameter of the tubes was 20 mm and their thicknesses were 2, 1, and 0.5 mm. They were extruded in the solid fraction of 90-95 per cent, at a speed of 10 cm/min-70 cm/min, with an extrusion force of 5-13 kg/mm². They were sound, but a longitudinal crack appeared when the working conditions were not suitable. In these experiments, extrusion of tubes of a thickness less than 0.5 mm was not successful, but with some proper working devices it should become possible.

CONCLUSIONS

From uni-axial compression tests and extrusion tests on Pb-Sn and Al-Cu alloys in their mashy state, the authors made the following conclusions

- (1) the flow stress of alloys at the temperature on the solidus line of the phase diagram is about half that at room temperature
- (2) according to the decrease of solid fraction from 100 per cent to 75-80 per cent, the flow stress decreases from 1 to $\frac{1}{4}$ - $\frac{1}{5}$
- (3) the alloys in their mashy state fracture in the strain range of 3-10 per cent in uni-axial deformation
- (4) the extrusion force of the alloys in their mashy state decreases with the decrease of solid fraction, but the decreasing rate is not as great as that of flow stress in uni-axial deformation
- (5) the extrusion force of mashy alloys does not depend on die angle
- (6) the maximum extrusion ratio attained in these tests was 1600 for bar extrusion
- (7) side cracks were apt to grow in bar products and longitudinal cracks were observed on a few tubes extruded in this process.

ACKNOWLEDGMENT

This study was made in the Kiuchi laboratory of the Institute of Industrial Science, University of Tokyo. The authors would like to express their gratitude to Professor H. Suzuki, the president of the Institute, for his invaluable advice.

REFERENCES

1. H. Kudo et al. *Journal of the Japan Society of Mechanical Engineers* (1964) **67**, 392.

STRESS STRAIN CURVES FROM THE RING TEST

by

A. N. BRAMLEY* and N. A. ABDUL†

SUMMARY

A new procedure for the establishment of stress strain curves from the ring test is described. The underlying upper bound theory, due to Avitzur, has been used to produce a very practical calibration chart for easy reduction of experimental data. The validity of the technique has been assessed by conducting ring tests and tensile/compression tests on cold aluminium and hot steel. It has also been shown that varying friction conditions and continuous or series testing methods do not invalidate the technique.

INTRODUCTION

Although analyses of hot working processes are in general unable to accommodate strain hardening and rely on the use of a mean effective yield stress, there is a sustained interest in the determination of accurate stress-strain data under hot working conditions. A variety of methods, ranging from the simple tensile test to the relative sophistication of cam plastometer studies, have been used to produce a tremendous amount of data, covering a wide range of materials and conditions. However, when considering new materials, specific temperature and strain rates, there is a strong case for obtaining the data oneself. There is therefore a demand for a simple direct method which involves the minimum of sophisticated equipment and procedures.

The latest advance in the use of the widely known ring test goes some way towards fulfilling this demand, and enables stress-strain curves, relevant to any temperature and strain rate configuration, to be obtained over a realistic range of strain. It is anticipated that in due course general methods of analysis for metalforming processes will be able to accommodate strain hardening and will therefore be able to make good use of the ring test.

THE THEORY OF THE RING TEST

Considerable use has been made in recent years of the ring compression test to determine coefficients of friction relevant to metalforming operations. When compressed in its axial direction, a flat ring will exhibit inward and outward radial flow in a way which is dependent solely upon the prevailing frictional conditions on the tool-workpiece interface.

Analytical prediction of the relationship between the geometric changes of the ring and the prevailing frictional condition, and also of the reduced forging pressure, have been made by Avitzur¹. The possibility of using the test to predict yield stress-strain relationships emanates from this, and was proposed by Saul, Male and De Pierre² who used Avitzur's theoretical equations to produce tabulated data which related the average forging pressure and geometry change to the yield stress of the material. The built-in measure of the friction provided by the ring enables the test to be used as a method of determining strain hardening characteristics. The validity of the technique was confirmed in part by correlation with results from the Polakowski re-machining technique.

The purpose of this paper is to show that a more realistic experimental validation of the technique can be made, and that the calibration data can be presented in such a way that the test can be used directly, without the aid of computational facilities, in a very simple way. The full details of the theoretical analysis are not therefore included, and the presentation confines itself to re-stating the assumptions made and the final working equations.

The lower upper-bound solution presented by Avitzur assumes that the deformation is uniform throughout the thickness of the specimen. This implies that the frictional shear stress, generated at the die-workpiece interface, will be uniformly distributed throughout the thickness; an assumption which will incur increasingly greater errors as the aspect ratio of the rings under consideration increases. A more recent paper by Avitzur³ does however present a theoretical analysis of the ring which accommodates barrelling, and presumably this improvement could be

* Department of Mechanical Engineering, Leeds University

† Department of Mechanical Engineering, Lagos University, Nigeria

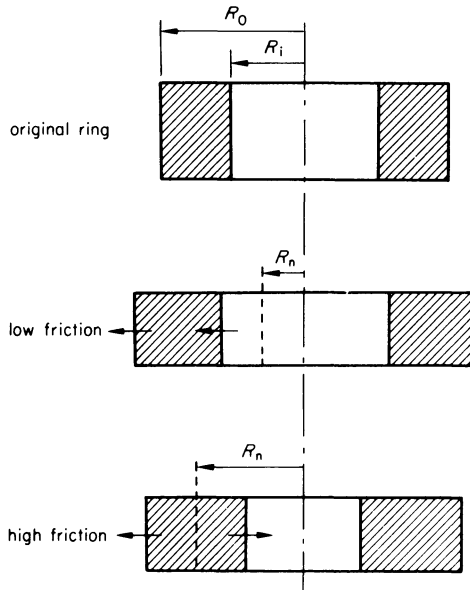


Figure 1. Ring deformation modes

incorporated into the technique. The frictional condition at the interface is described by a friction factor m , assumed constant across the interface, and defined by

$$m = \sqrt{3\tau_f/\sigma_0}$$

where τ_f is the frictional shear stress and σ_0 is the yield stress of the ring material.

There are two distinct modes of deformation defined by the position of the neutral radius R_n (figure 1), which separates radially inward and outward flow of material. In the case of wholly radially outward flow, both the inside and outside diameters of the ring increase, and the neutral radius is imagined to lie within the hole of the ring. Theoretically, this mode will occur when

$$\frac{mR_0}{T} \leq \frac{1}{2(1 - R_i/R_0)} \ln \left[\frac{3(R_0/R_i)^2}{1 + (1 + 3(R_0/R_i)^4)^{1/2}} \right] \quad (1)$$

and for this condition of $R_0 > R_i > R_n$, the neutral radius R_n , after being used as a minimising parameter in the upper bound analysis, is given by

$$\left(\frac{R_n}{R_0}\right)^2 = \frac{\sqrt{3}}{2} \cdot \frac{1 - (R_i/R_0)^4 x^2}{[(x(x-1)1 - (R_i/R_0)^4 x)]^{1/2}} \quad (2)$$

where

$$x = \left[\frac{R_0}{R_i} \exp \left\langle -\frac{mR_0}{T} \left(1 - \frac{R_i}{R_0}\right) \right\rangle \right]^2$$

and the forging pressure p is given by

$$\begin{aligned} \frac{p}{\sigma_0} = & \frac{1}{1 - (R_i/R_0)^2} \left[\left(1 + \frac{1}{3} \left(\frac{R_n}{R_0}\right)^4\right)^{1/2} \right. \\ & - \left. \left(\left(\frac{R_i}{R_0}\right)^4 + \frac{1}{3} \left(\frac{R_n}{R_0}\right)^4 \right)^{1/2} \right] \\ & + \frac{2}{3\sqrt{3}} \frac{mR_0}{T} \left\langle 1 - \left(\frac{R_i}{R_0}\right)^3 \right\rangle \end{aligned} \quad (3)$$

For the case where the neutral radius lies within the material of the ring, the flow is such that the outer diameter increases whilst the inner diameter decreases. This mode occurs when

$$\frac{mR_0}{T} \geq \frac{1}{2(1 - R_i/R_0)} \ln \left[\frac{3(R_0/R_i)^2}{1 + (1 + 3(R_0/R_i)^4)^{1/2}} \right] \quad (4)$$

As before R_n is used as a minimizing parameter and is given by

$$\begin{aligned} 2 \left(\frac{mR_0}{T} \right) \left(1 + \frac{R_i}{R_0} - \frac{2R_n}{R_0} \right) \\ + \ln \left[\frac{\left(\frac{R_i}{R_0}\right)^2 \left(\frac{R_n}{R_0} \right)^2 + \left(3 + \left(\frac{R_n}{R_0}\right)^4 \right)^{1/2}}{\left(\frac{R_n}{R_0}\right)^2 + \left(3 \left(\frac{R_i}{R_0}\right)^4 + \left(\frac{R_n}{R_0}\right)^4 \right)^{1/2}} \right] = 0 \end{aligned} \quad (5)$$

and the forging pressure p in this case is given by

$$\begin{aligned} \frac{p}{\sigma_0} = & \frac{1}{1 - (R_i/R_0)^2} \left[\left(1 + \frac{1}{3} \left(\frac{R_n}{R_0}\right)^4\right)^{1/2} \right. \\ & - \left. \left(\left(\frac{R_i}{R_0}\right)^4 + \frac{1}{3} \left(\frac{R_n}{R_0}\right)^4 \right)^{1/2} \right] \\ & + \frac{2}{3\sqrt{3}} \frac{mR_0}{T} \left\langle 1 + \left(\frac{R_i}{R_0}\right)^3 - 2 \left(\frac{R_n}{R_0}\right)^3 \right\rangle \end{aligned} \quad (6)$$

Inspection of these equations reveals that p/σ_0 is independent of ring thickness T , and dependent on the various radius ratios, R_i/R_0 , R_n/R_0 , etc. In principle mR_0/T from Equations (2) and (5) could be substituted into Equations (3) and (6) respectively to give p/σ_0 directly as a function of the ring geometry. Unfortunately Equation (2) transforms into a quartic for which it has not been possible to obtain an analytical solution. A further possibility is that of substituting experimental data into these equations and solving them numerically, but this would involve the use of computational facilities to reduce every set of data. The most appropriate procedure is to solve the equations incrementally in the following way:

Consider a ring with a certain ratio of initial dimensions $R_0:R_i:T$ and an arbitrary friction factor m where $1 > m > 0$. Check Equations (5) and (6) to establish whether $R_n > R_i$ or $R_n < R_i$. Solve appropriately Equations (2) or (5) for R_n , and, subsequently, Equations (3) and (6) for p/σ_0 . Allow a small increment of deformation (the smaller the increment, the less the cumulative errors). In this investigation a 4 per cent decrement in T was used for each step. Assuming R_n remains constant for the step, calculate the new geometry R'_i , R'_0 and T' at the end of the step by using volume constancy conditions for the ring. This gives:

$$(R_i')^2 = \frac{(R_i^2 - R_n^2)T - T'R_n^2}{T'}$$

and

$$(R_o')^2 = \frac{(R_o^2 - R_n^2)T + T'R_n^2}{T'}$$

The procedure is continued until the necessary range of deformation is covered, p/σ_0 being derived at each

step. The whole sequence is then repeated for a different value of m .

The results plotted in figure 2 are for a ring with an initial radius ratio R_o/R_i of 2 as a variation of internal radius change and thickness change against p/σ_0 . As anticipated p/σ_0 is independent of the initial ring thickness, and, also, the friction factor m does not appear directly in figure 2. It is interesting to observe that it is not a pre-requisite that the friction

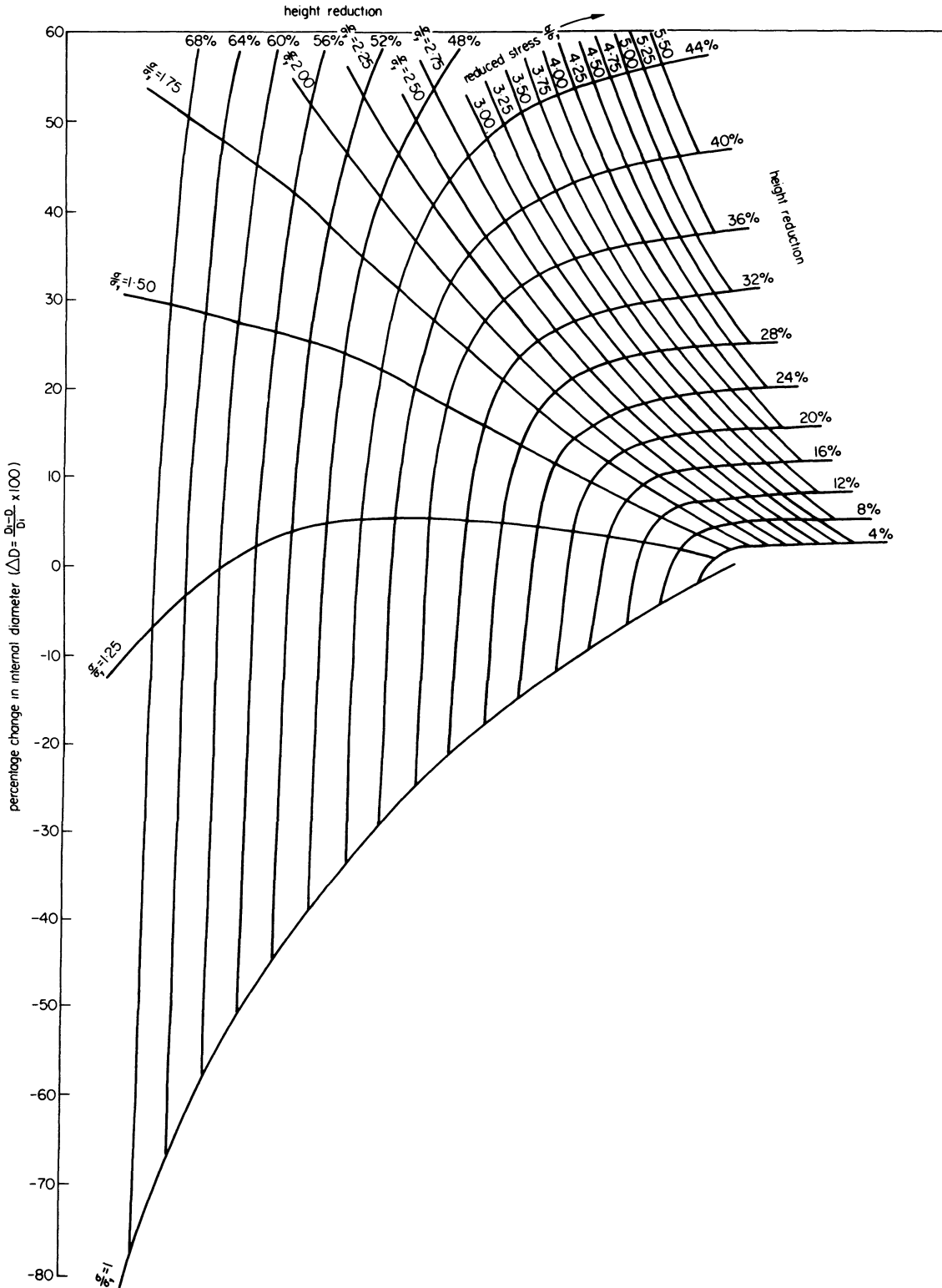


Figure 2. Ring test calibration chart.

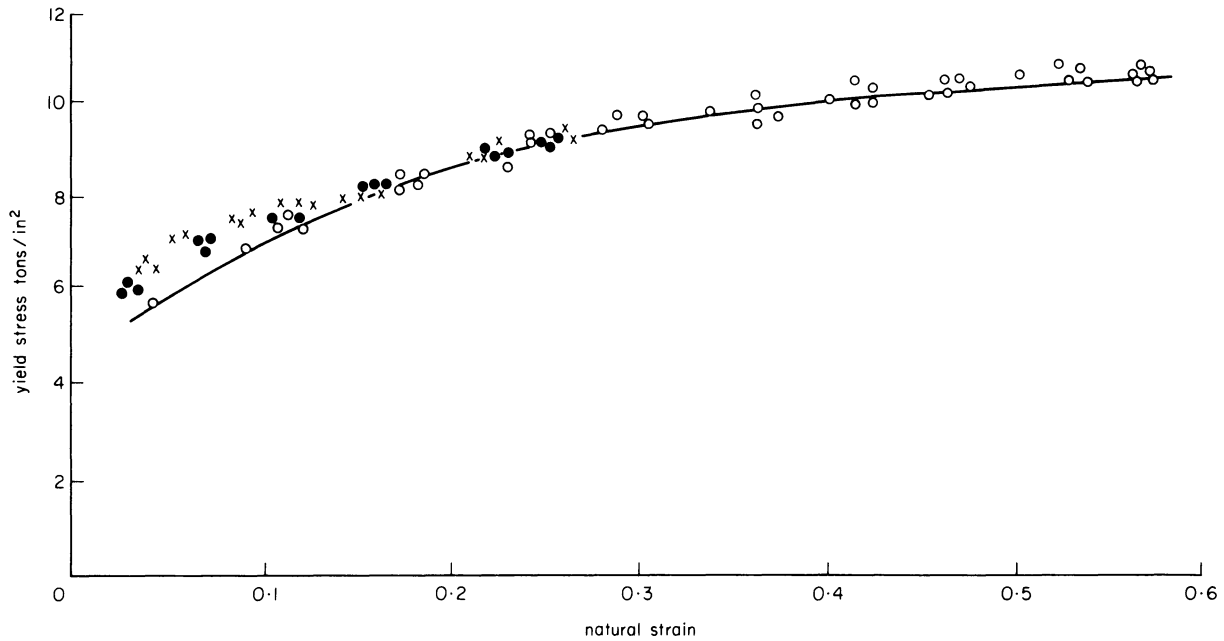


Figure 3. Cold aluminium stress-strain curves, \circ —ring tests, \times —tensile tests, \bullet —compression tests.

factor should remain constant during the whole deformation, but only during the increment. A built-in measure of friction is provided by the pressure p and the change in geometry. Similarly σ_0 may vary between successive increments to accommodate strain hardening. All three variables in figure 2 are continuously scaled to give a high read-out accuracy; the procedure for its use being as follows:

At an arbitrary number of stages during the deformation, experimentally determined values of percentage change in internal diameter and height reduction are fed into the graph to establish corresponding values of the reduced stress p/σ_0 . A knowledge of p , from the applied load and current cross

sectional area of the ring, enables σ_0 to be determined at each stage. σ_0 , plotted against the height reduction (expressed as a strain), then yields the relevant strain hardening characteristic.

EXPERIMENTAL WORK

Initial experimental confirmation of the technique was achieved via carefully controlled incremental tests on a slow speed testing machine. Figure 3 shows the spread of results obtained for commercially pure aluminium rings of external diameter 1.5 in, internal diameter 0.75 in and height 0.5 in, tested at room temperature. Under these testing conditions, it is possible to perform accurate tensile and compression

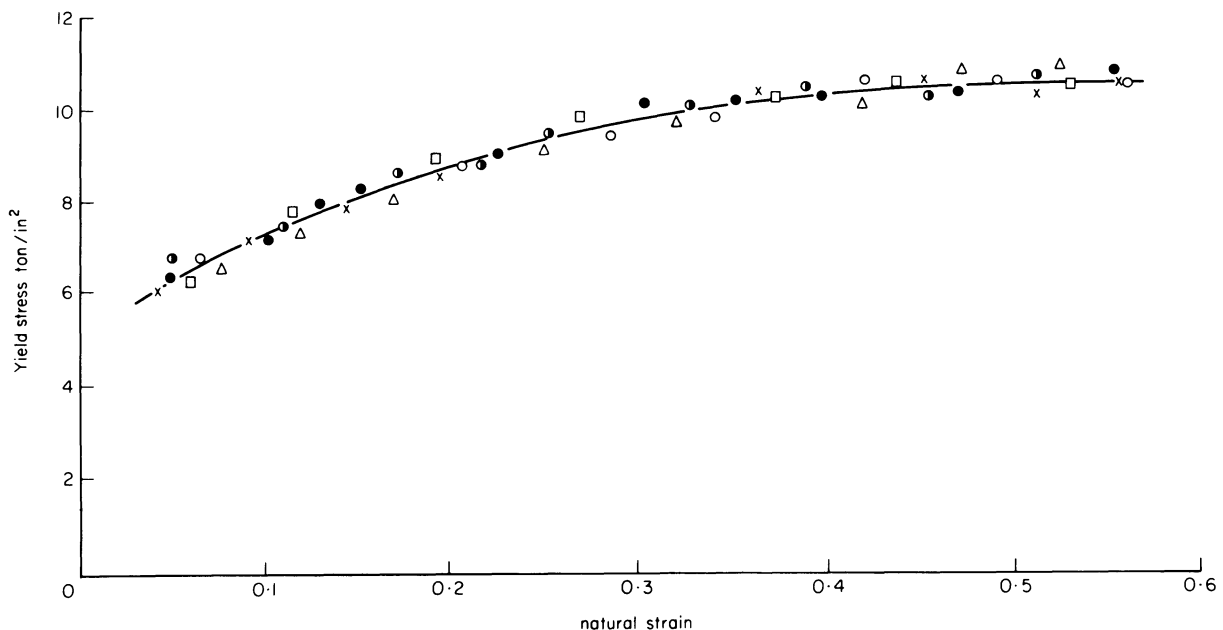


Figure 4. Stress-strain curves derived from aluminium rings of different initial heights; \circ \times \triangle \square —6:3:2 rings, \bullet —6:3:0.5 rings.

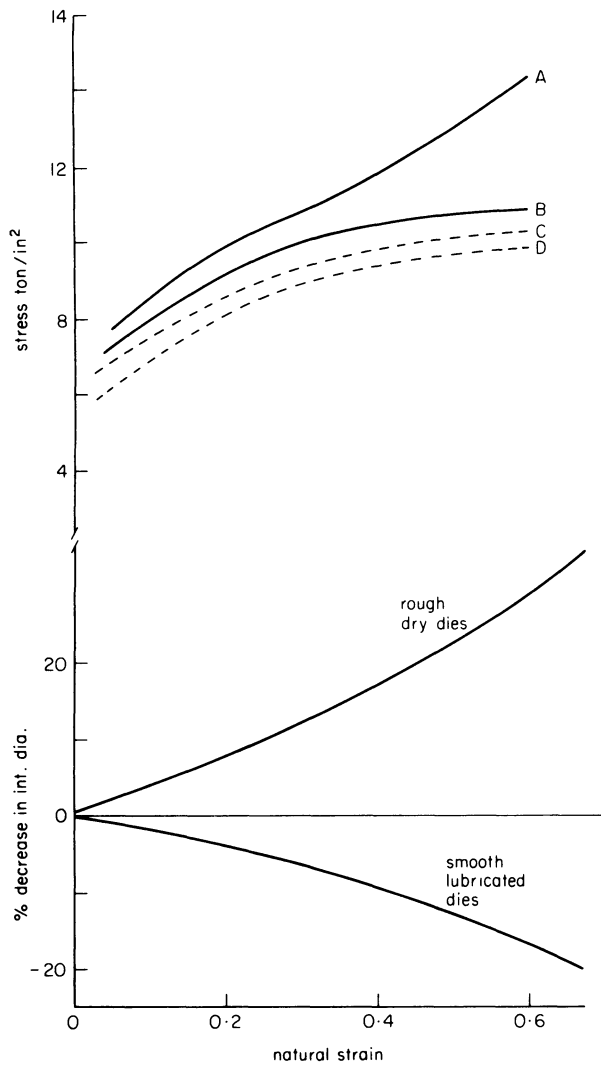


Figure 5. Stress-strain curves derived from aluminium rings under different frictional conditions; *A*—Actual flow stress for high friction, *B*—Actual flow stress for low friction, *C*—Derived stress-strain curve from high friction test, *D*—Derived stress-strain curve from low friction test.

tests, for comparison purposes, over part of the range of the ring test. Such results are included in figure 3, and suggest that there may be an overestimation of the yield stress at strains less than 0.1, but that from strains of 0.1 to 0.25, the limit of uniform strain for the tensile and compression tests, the ring test gives very accurate results. According to the analysis, the derived stress-strain curve for the ring test will be independent of the height of the specimen for all values of R_o/R_i . To test this feature, some rings of 0.125 in height were compressed, and the derived results are plotted in figure 4 from which it is clear that such a dramatic change in the height of the ring has no effect on the derived stress-strain curve. This confirms the validity of the theoretical assumption.

A further important feature of the ring test theory is its ability to accommodate different frictional conditions from one test to the next, and also friction conditions which may vary during the test. The latter condition is difficult to impose with any degree of certainty, and, in fact, is probably a situation which obtains during the ring test in any case. The former condition, however, can be imposed by testing with rough dry dies, and then with smooth lubricated dies. The effect of doing this is shown in figure 5 on a different batch of aluminium. The widely different geometrical changes which occur are shown on the graph, together with the actual forming stress curves and the derived yield stress curves for the two frictional conditions. The difference between the derived curves is of the order of 0.25 tons/m².

Tests on En3B steel at 1200°C were conducted on a Rhodes 75 ton crank press, and, because of the impracticality of using an incremental compression technique with continual reheating, a series-compression technique was adopted. This involved the compressing of a series of rings, of equal initial size, by different amounts extending over the required range of deformation. To confirm the validity of this approach, a special displacement transducer which

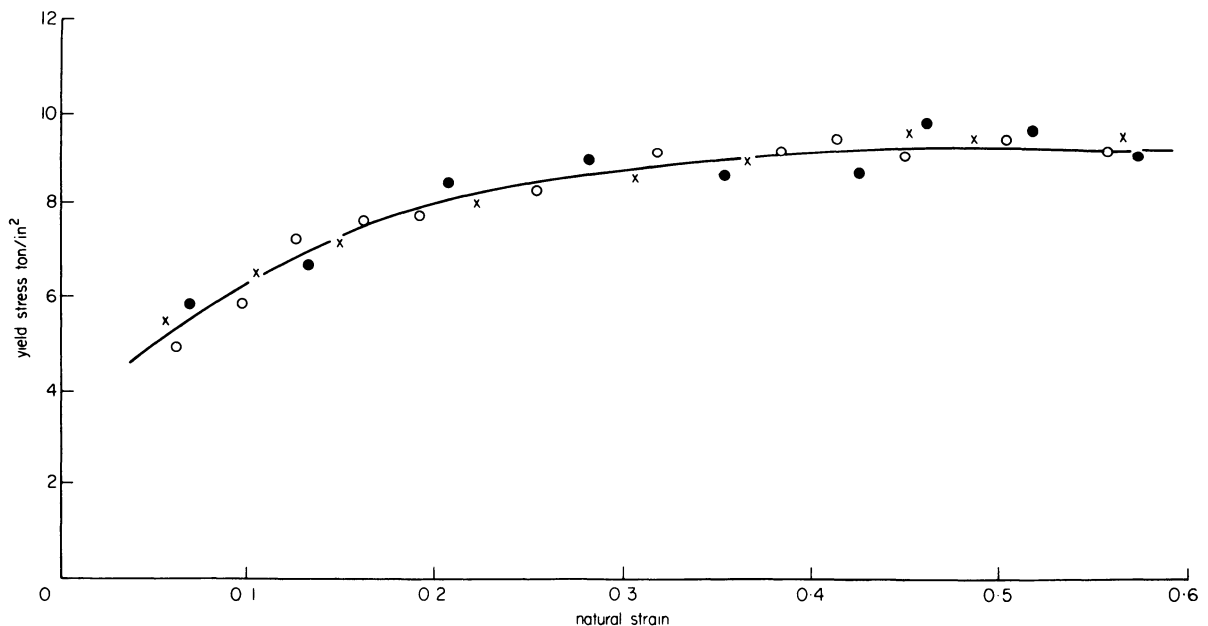


Figure 6. Stress-strain curves derived from En3B steel rings at 1200°C; X—low friction, continuous test, ●—high friction, continuous test, ○—series compression tests.

would record changes in the external diameter* of the ring during the course of compression was constructed. These readings were converted to values of internal diameter change by using the corresponding thickness change readings obtained from a conventional vertical displacement transducer in the die set. The results of both the series-compression and the continuous tests are shown in figure 6.

CONCLUSIONS

The ring test analysis made by Avitzur can be exploited to develop a simple procedure for the experimental determination of stress-strain curves. It can be used under conditions of strain hardening, variable frictional level and series incremental testing. Rapid reduction of the experimental data into a stress-strain curve is effected with a prepared calibration chart. A chart for rings having an initial radius ratio $R_0/R_i = 2$ is included in this paper, and large scale copies of this chart are available on request from the Mechanical Engineering Department at Leeds University.

It is anticipated that in cases where the stress-strain curve is required for application to a particular operation such as forging, the prevailing strain-rate and chilling conditions can be reproduced in the ring

* The measurement of internal diameter changes was originally claimed by Male and Cockcroft⁴ to be more sensitive than outside diameter changes. This is not always the case: for example when the internal diameter sustains only slight variation due to the particular frictional conditions, the outside diameter is changing by substantial amounts and is correspondingly a more sensitive parameter for measurement.

test by using a similar aspect ratio ring specimen and the same type of forging equipment.

Finally, a by-product of the technique is that one can use the ring test data to assess frictional conditions in the traditional way which was originally developed by Male and Cockcroft⁴.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support of the following:

Science Research Council, for a research grant.

Government of Nigeria, for financial support for N. A. Abdul.

Professor B. N. Cole, for the provision of research facilities.

REFERENCES

1. B. Avitzur. Forging of hollow discs. *Israel J. Tech.* (1964) 2, 295.
2. G. Saul, A. T. Male and V. De Pierre. A new method for the determination of material flow stress values under metalworking conditions. *USA AFML Tech. Report* (1970) TR-70-19.
3. B. Avitzur. Bulge in hollow die forging. *Departmental Report Lehigh University* (1969).
4. A. T. Male and M. G. Cockcroft. A method for the determination of the coefficient of friction of metals under conditions of bulk plastic deformation. *J. I. Metals* (1964-65) 93, 38.

AN ELEMENTAL UPPER-BOUND TECHNIQUE FOR GENERAL USE IN FORGING ANALYSIS

by

R. P. McDERMOTT* and A. N. BRAMLEY*

SUMMARY

This paper outlines a general method of analysis developed from the procedure described by Kudo in 1960 in which the forging cross section is divided into a number of characteristic elemental regions. As proposed by Kudo, the procedure was essentially manual and only rectangular sectioned elements were considered; indeed it does not appear to have enjoyed any significant industrial application.

The procedure developed, extends this basic approach to include conditions of axial symmetry. Equations have been established for triangular, convex and concave shaped regions and, with the appropriate calculations programmed into a computer, it is a relatively simple procedure to input data describing the flow stress, friction, geometry, etc. for each region and obtain an output of forging load prediction. Also embodied in the technique is the facility for predicting optimum flash geometry which is of major significance in an industry conscious of the need to conserve material. Preliminary experimental work shows promising results.

NOTATION

\dot{E}_i	the rate of internal energy dissipation
\dot{Y}	yield strength of the forged material
$\dot{\epsilon}_R$	strain rate in radial R direction
$\dot{\epsilon}_Z$	strain rate in axial Z direction
$\dot{\epsilon}_\theta$	strain rate in circumferential θ direction
$\dot{\gamma}_{RZ}$	shear strain rate in $R-Z$ plane
F	coefficient of friction on region boundaries
S	rate of relative slip on boundaries

INTRODUCTION

There are basically two types of approach open to the potential analyst when considering even the most simple of axisymmetric forgings. The choice, however, is dependent more upon the environment in which the analysis is made rather than any other consideration. The forging industry depends largely on methods of analysis, similar to those described by Seibel¹ and summarized by Lange² and Altan and Fiorentino³, which use plan areas, flow stress and shape factors to arrive at a forging load prediction; based primarily on experience and guesswork rather than formulation. This has the advantages of speed and simplicity which are necessary prerequisites of production situations, but which can also result in gross overestimates of plant requirement and excessive flash production. Academic establishments on the other hand prefer more elegant/exact approaches such as the slip line field technique⁴ and upper bound

methods⁵ which, by considering far more parameters, achieve answers at least an order of accuracy better, whilst at the same time suffering from a high degree of complexity.

In 1960 for example, Kudo⁶ produced a complex academic method, based on the upper bound technique (described later), for analysing axisymmetric shapes, with the restriction that the overall forging must be capable of being split into "elemental rings" of rectangular section only. The energy for each such ring element was based solely on the flow in and on the boundaries surrounding that region (figure 1). Despite these restrictions, it produced maximum load

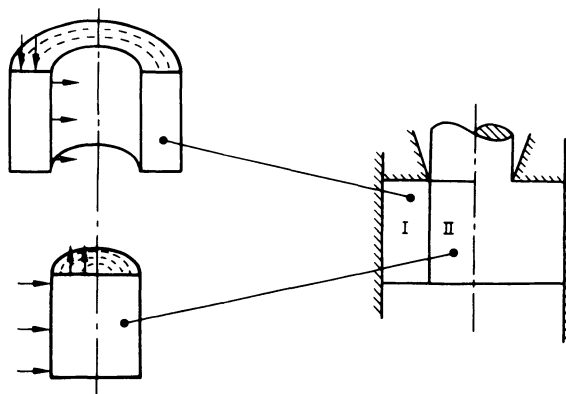


Figure 1. Analysis of simple forging/extrusion as envisaged by Kudo.

* Department of Mechanical Engineering, University of Leeds.

estimates within +(10 per cent → 15 per cent), an unprecedented accuracy as far as the forging industry was concerned, but of little use because of its academic tedium. Later work by Kobayashi et al⁷ led to further marginal improvements in accuracy using similar methods with the same rectangular shape and academic restrictions.

ELEMENTAL UPPER BOUND TECHNIQUE

This original concept by Kudo⁶, involving the breakdown of the forging into elemental regions, is basically easy to understand. By updating this nomenclature, therefore, a similar method of analysis, now free from the previous shape and academic limitations, is proposed. In this new approach eight elemental rings each capable of being linked to all others are considered. They include not only rectangular but also triangular and circular cross sectioned elements which enable forgings with significant draft and radii to be analysed.

These eight regions are divided so that two modes of deformation, both concerned with an element having an outer radius of unity, an inner radius B , with height to outer radius ratio of A are considered. In figure 2(a), (b), (c) and (d), each of the four regions is deformed in such a manner that its top surface descends vertically as a result of an external force of unit velocity which causes the inner side surface to move inwards as a straight line. The surroundings to these regions may be considered as either rigid tools or as rigid parts of surrounding workpiece; no flow being allowed across the remaining faces in contact with either tool or workpiece. The second type of deformation occurring in the remaining four regions is that shown in figure 3 (a), (b), (c) and (d). This differs from the first mode only in that the inner surface does not move whilst the outer cylindrical surface moves radially outwards, remaining cylindrical.

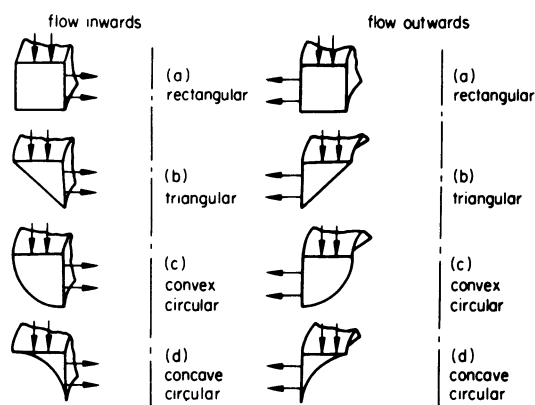


Figure 2 and 3. Diagrams showing the eight basic 'elemental rings'.

For each of these eight regions shown, a general admissible velocity field (i.e. one which is a distribution of particle velocities, kinematically compatible within itself, and with externally applied forces), was considered as the example shown in Appendix A.

Similar methods were used to obtain the velocity fields for the following regions:

Figure 2(a)—Rectangle flow inwards

$$\dot{u} = -\frac{(1-R^2)}{2AR}$$

$$\dot{w} = -\frac{Z}{A}$$

also

$$\dot{u} = -\frac{(1-B)(1-R)}{2AR}$$

$$\dot{w} = \frac{(1-B)Z}{2AR}$$

Figure 3(a)—Rectangle flow outwards

$$\dot{u} = \frac{R^2 - B^2}{2AR}$$

$$\dot{w} = -\frac{Z}{A}$$

also

$$\dot{u} = \frac{(1-B)(B+R)}{2AR}$$

$$\dot{w} = \frac{(1-B)Z}{2AR}$$

Figure 2(b)—Triangle flow outwards

$$\dot{u} = -\frac{\cot \alpha}{2} \left(1 + \frac{1}{R}\right)$$

$$\dot{w} = \frac{\cot \alpha}{2} \frac{Z}{R} + 1$$

Figure 3(b)—Triangle flow outwards

$$\dot{u} = \frac{\cot \alpha}{2} \left(1 + \frac{B}{R}\right)$$

$$\dot{w} = -\frac{\cot \alpha}{2R} Z + 1$$

Figure 2(c)—Convex circular flow inwards

$$\dot{u} = \frac{(R^2 - 1)}{2R(1 - 2B + 2BR - R^2)^{1/2}}$$

$$\dot{w} = -\frac{Z(-R^3 + 3BR^2 - 4BR + R + B)}{2R(1 - 2B + 2BR - R^2)^{3/2}} + 1$$

Figure 3(c)—Convex circular flow outwards

$$\dot{u} = \frac{(R^2 - B^2)}{2R(B^2 - 2B + 2R - R^2)^{1/2}}$$

$$\dot{w} = -\frac{Z(-R^3 + 3R^2 + B^2R - 4BR + B^2)}{2R(B^2 - 2B + 2R - R^2)^{3/2}} + 1$$

This type of approach, however, was found inapplicable to regions shown in figures 2(d) and 3(d)

(i.e. concave circular flow inwards and outwards respectively), because of the contradictory conditions obtained at the tangential intersections of the two straight line boundaries with the curve. Therefore, as a means to overcoming this, an approximate solution was initiated consisting of four triangles which linked together to give a 10 per cent overestimate in material volume but a 5 per cent underestimate in outlet velocity (complex velocity components not shown here).

Once the velocity components for any of the eight basic regions are known, by definition

$$\begin{aligned} \dot{\epsilon} &= \frac{\partial \dot{u}}{\partial R} \\ \dot{\epsilon}_Z &= \frac{\partial \dot{w}}{\partial Z} \\ \dot{\epsilon}_\theta &= -(\dot{\epsilon}_R + \dot{\epsilon}_Z) \\ \dot{\gamma}_{RZ} &= \left(\frac{\partial \dot{u}}{\partial Z} + \frac{\partial \dot{w}}{\partial R} \right) \end{aligned}$$

We can obtain all the strain rates, and so the rate of internal energy dissipation can be calculated for that particular field using Hill⁸.

$$\dot{E} = \sqrt{\frac{2}{3}} Y \int_v (\dot{\epsilon}_R^2 + \dot{\epsilon}_\theta^2 + \dot{\epsilon}_Z^2 + \frac{1}{2} \dot{\gamma}_{RZ}^2) dV + \dot{Y} \int_s F \dot{S} ds \quad (10)$$

The first integral being carried out throughout the entire volume of the workpiece which is deforming continuously, and the second over all surfaces of velocity discontinuity on workpiece/tool and inter-region boundaries.

It should be noted that for both modes of deformation the quantity \dot{E} remains unchanged if the same unit region is compressed radially and horizontally on its cylindrical surface. This results in its top surface being raised in contact with a rigid external body which moves freely in the vertical direction. The value of \dot{E} for any particular region can now be expressed non-dimensionally by division with the product of yield stress, surface area of pressing, and pressing speed of that unit region. This gives what might now be called the *coefficient of internal energy dissipation*. Thus, when considering the total rate of energy dissipation for a certain velocity field in some given complex forging, consisting of many such unit regions, it could be represented by:

$$\dot{E}T = Y \sum_{i=1}^{i=n} e_i A_i V_i \quad (11)$$

- e_i being the coefficient of internal energy dissipation for region i
- A_i the pressing area for region i
- V_i the pressing speed for region i
- n being the total number of unit regions comprising the total forging

Now an upper bound for the mean extrusion pressure for example over the whole sectional area of the billet

could be expressed by

$$\frac{P}{Y} = \frac{\dot{E}T}{\pi \cdot D_T^2/4 \cdot V_T} \quad (12)$$

$\pi \cdot D_T^2/4$ being the total pressing area of the forging and V_T the closure speed. At this point if solutions to Equations (10) and (11) involved simple addition of the terms as shown, this method of analysis would without doubt find pride of place with the forging industry. However, in determining a general admissible velocity system at any point inside and along the boundaries of each region, the mathematics becomes very involved. Furthermore, to determine the strain rate terms in Equation (10) and to perform the volume integral with them becomes impossible by direct mathematical methods. It is for this reason that a system was evolved using the I.C.L. 1906A computer to again simplify the analysis.

In order to determine the load required to deform any of the eight regions, as a direct reading in tons, all that is required is input of the following data:

- (1) Type of region
- (2) Yield stress of material being deformed
- (3) Pressing speed at which it is being deformed
- (4) Internal radius
- (5) External radius
- (6) Height
- (7) } Friction factors
- (8) }
- (9) }
- (10) Additional friction factor if four sided region (figures 2(a) or (3))

Hence analysis of a complex forging can now be completed by subdivision into elemental shapes, input of data for each, and summation to find the total load. It should be stressed, however, that this method of analysis is such that it considers the forging only at the end of the process when all the space in the die cavity is filled, that is, maximum load. As such, it is only suitable for forgings of the press type as drop hammers require energy calculations for the full process.

In figure 4 is shown a hypothetical axisymmetric forging for which a typical region breakdown is given. It is evident that all the composite regions of this forging link onto one another to produce an

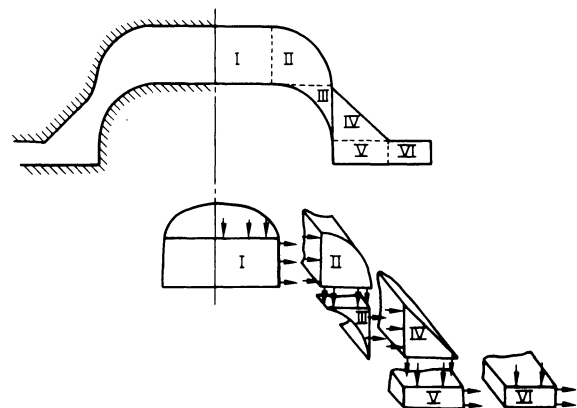


Figure 4. Diagram showing hypothetical forging.

apparently smooth transition of flow on an inter-elemental scale. The question must now arise as to how the analysis can take place if such a transition cannot be achieved.

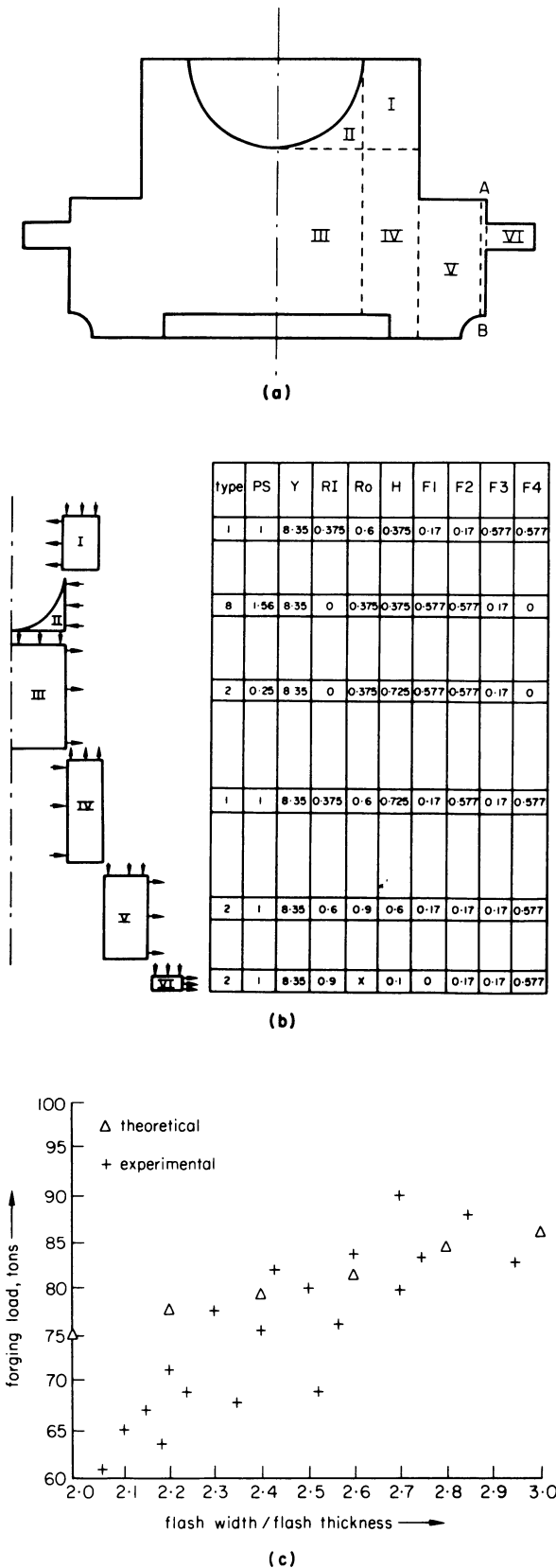


Figure 5. (a) - Diagram showing actual forging. (b) - Its regional breakdown and associated parameters. (c) - Compares the theoretical results so obtained with experimental.

Figure 5(a) illustrates such a forging (after Balogun⁹). Here there is a cyclic type flow existing in regions I → IV and, in order to include regions V and VI, we must consider the material to be filling an infinitesimal elemental gap, and so maintaining the continuity of analysis. Such a gap would be AB, where at the moment of *die filling* region V would be flowing outwards to close this gap AB. The analysis was completed in this way to produce the elements shown and their associated parameters in figure 5(b). The flash thickness was kept constant whilst the flash width ($X - 0.9$ inches) varied and the theoretical results are compared with those obtained experimentally in figure 5(c). Although this approach varies

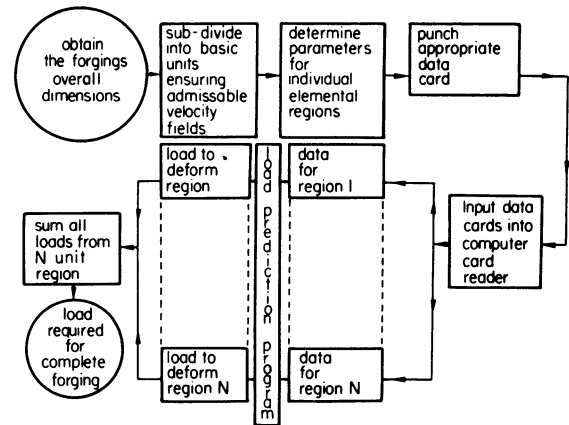


Figure 6. Flow diagram illustrating analysis procedure.

slightly to that in figure 4 the same method of solution is relevant, i.e. if the appropriate data is determined for each composite region and the analysis is completed as illustrated on the flow diagram shown in figure 6, all that would be required would be the punching of six cards (one for each region), inputting these to a computer card reader and simply obtaining the predicted load from the output. However, computer time is expensive, and an industry already committed to its existing plant size might argue that this cost is unacceptable. Nevertheless, there is a further use for this method that almost certainly justifies its adoption.

A recent paper¹⁰ published by members of the British Drop Forging Research Association illustrated the relative importance of material production and die cost in forging (figure 7). It is interesting to note that on average 50 per cent of the total cost of forgings is made up by material costs. Thus, a drop of 10 per cent in material costs would result in a 5 per cent decrease in forging costs. In closed die forging the yield from material can vary between 50-90 per cent¹¹ whilst the average yield may be about 70 per cent. Thus, neglecting the revenue from scrap metal, an estimated 15 per cent of the total cost in producing forgings without any losses in flash, scale, sawing, etc. However, by improving the existing flash designs it might be possible to increase the overall material yield by a few per cent. This method of analysis fortunately can do just that, that is it predicts the dimensions of a flash that is *just*

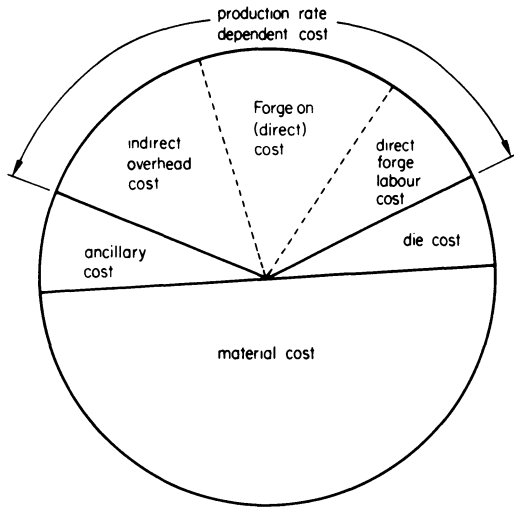


Figure 7. Distribution of forging costs.

sufficient to ensure complete die filling and so it optimises flash wastage and also increases die life as a result of the reduced peak die load.

The procedure is based on the fact that at the point of minimum die separation the biggest contributing factor to the total load is that from the flash itself¹². Consider, therefore, a die closing incrementally: for each step of the closure process, again making the assumption that all regions inside the die cavity are fully filled, certain parameters (e.g. height, A) of each of the elements making up the whole forging, would change as would therefore the load at this point. At some stage in this closure process, there comes a point when the load contributions from the flash would be greater than that from all the other regions combined. Here it could be said that it required less energy to achieve the formation of the complete shape inside the die than it did for the flash alone, and so material being forged finds it easier to completely fill the die cavity than it does to expand through another increment in flash growth. At the point of balance, i.e. when the load to deform the

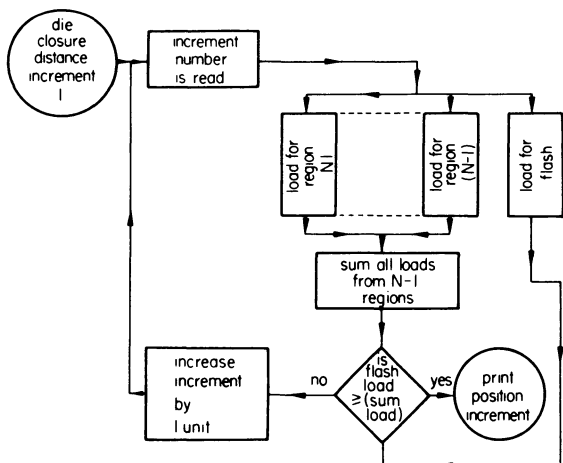


Figure 8. Flow diagram illustrating flash prediction procedure.

flash equals the load to deform the rest of the forging, is that point when hypothesis becomes fact

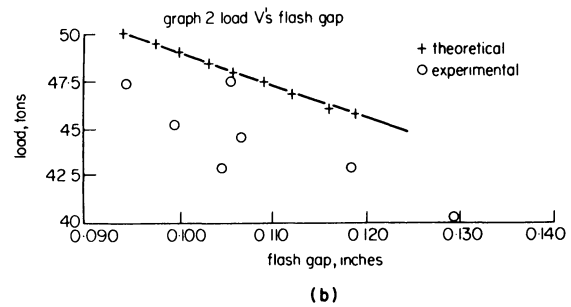
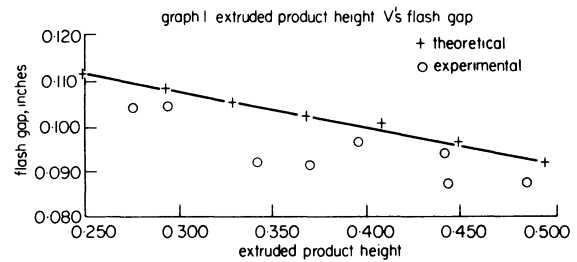
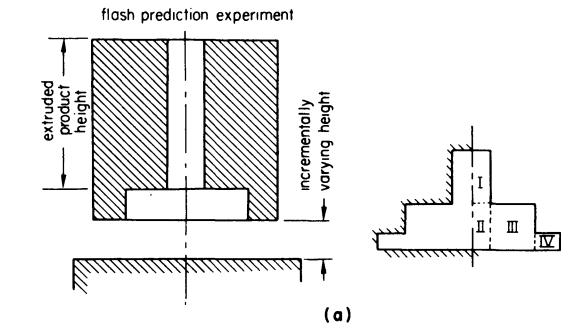


Figure 9. (a) – Die used in flash prediction experiment. (b) – Compared theoretical prediction with experimental results.

and the die cavity is just completely filled. By noting the die separation here, we can determine the complete flash geometry using volume constancy, and so achieve the optimum flash size and minimum material wastage—the few percent potential saving to industry mentioned previously.

Again to facilitate direct access by industry to this method it has been packaged with the use of the 1906A, and requires only the die separation for each set of element parameters to produce the required flash geometry, as shown in the flow chart in figure 8. A copy of both this and the load determination computer programme are available from the Drop Forging Research Association, Sheffield.

Figure 9 illustrates a simple die used experimentally to determine the flash geometry required to achieve a certain height of extruded product. This is a difficult example since material can flow in either flash or extruded product. The results, however (figure 9(b)), show a high degree of correlation between the theoretical predictions and experimental fact.

CONCLUSION

Finally, it should be stressed that provided a suitable velocity field can be established for the forging as a whole (i.e. one that may deviate from reality but must not be impossible), there appears to be no limit to the complexity of forgings open to analysis by this method; not only in terms of load prediction but also for optimum flash geometry determination.

ACKNOWLEDGMENTS

The authors wish to acknowledge the Drop Forging Research Association and Guest, Keen, Nettlefold Forgings Limited, Bromsgrove for assistance in providing background information for this work; the Science Research Council for the research grant to cover the work and the research studentship for R. P. McDermott; and Professor B. N. Cole for providing research facilities.

REFERENCES

1. E. Siebel. *Werkstattstechnik und Maschinenbau* (1950) **40**(11) 373-80.
2. K. Lange. *Theory and basic principles of drop forging. Metal Treatment and drop forg.* (May 1965).
3. Altan and Fiorentino. Predictions of loads and stresses in closed die forging. *Journal of Eng. for Industry* (May 1971), 477-84.
4. W. Johnson, R. Sowerby and J. B. Haddow. *Plane Strain Slip Line Fields*, Arnold (1970).
5. W. Prager and P. G. Hodge. *Theory of Perfectly Plastic Solids*, Chapman and Hall (1951).
6. H. Kudo. Some analytical and experimental studies of axi-symmetric cold forging and extrusion I. *Int. J. Mech. Sci.* (1960) **2**, 102-27.
7. S. Kobayashi. Upper bound solutions of axi-symmetric forming problems. *Trans. ASME* (May 1964) 122.
8. R. Hill. *The Mathematical Theory of Plasticity*, Clarendon Press, Oxford (1950).
9. S. A. Balogun. *Die load and stress in press forging*, PhD. thesis, Aston University (1971).
10. T. Alton and H. J. Henning. Closed-die forging of round shapes. *Metallurgia and Metal Forming* (March, 1972) p. 83.
11. A. C. Hobdel and A. Thomas. Approaches to cheaper forgings. *Metal Forming* (Jan, 1969) p. 17.
12. K. Vieregge. *A contribution to flash design in closed die forging*. Doctoral dissertation, Technical University, Hanover (1969).

APPENDIX A EXAMPLE SHOWING THE GENERATION OF AN ADMISSIBLE VELOCITY FIELD FOR A TRIANGULAR REGION FLOW INWARDS

In figure 10(b) let it be assumed that \dot{u} be independent of Z (in order to facilitate calculation of internal rate of energy dissipation and to allow connection with neighbouring deforming regions) and that the inlet and outlet velocities from any region boundary are constant along that boundary.

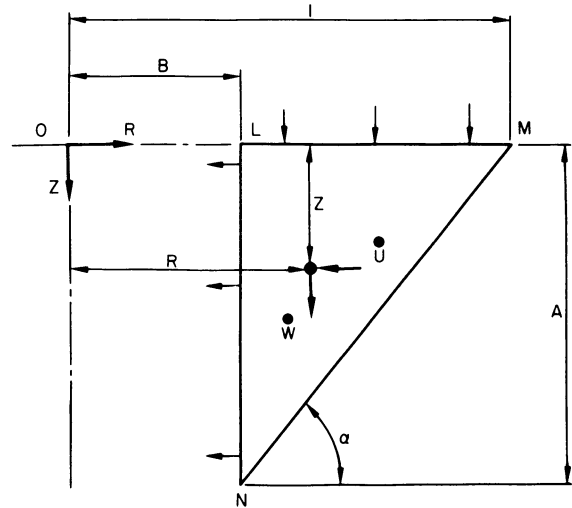


Figure 10. Unit triangular region flow inwards.

By considering volume constancy

$$\dot{\epsilon}_R + \dot{\epsilon}_\theta + \dot{\epsilon}_Z = \frac{\partial \dot{u}}{\partial R} + \frac{u}{R} + \frac{\partial \dot{v}}{\partial Z} = 0$$

and since \dot{u} contains only R as a variable then \dot{w} must contain both R and Z , but since R and Z are independent variables the above equation dissolves into

$$\frac{d\dot{u}}{dR} + \frac{\dot{u}}{R} = f(R) \quad (1)$$

and

$$\frac{\partial \dot{v}}{\partial Z} = -f(R) \quad (2)$$

Where $f(R)$ is an arbitrary function of R Equation (2) leads to

$$\dot{v} = -Zf(R) + g(R)$$

and by using the boundary conditions that

$$\dot{v} = 1 \quad \text{at} \quad Z = 0$$

$g(R)$ is reduced to 1, so that

$$\dot{v} = -Zf(R) + 1 \quad (3)$$

On the other hand, the general solution of Equation (1) is

$$\dot{u} e^{\int dR/R} - \int f(R) e^{\int dR/R} dR = C$$

Since $e^{\int dR/R} = e^{\ln R} = R$,

$$\dot{u} = \frac{\int f(R)R dR + C}{R} \quad (4)$$

Now along MN , the resultant velocity vector must be parallel to MN from continuity requirement, provided that the region below MN is at rest. This requires that

$$\frac{\dot{v}}{\dot{u}} = -\tan \alpha \quad \text{along} \quad \frac{Z}{1-R} = \tan \alpha \quad (5)$$

Substituting Equations (3) and (4) into Equation (5) we have

$$R(1-R)(\tan \alpha)f(R) - R = \tan \alpha \left(\int f(R)R dR + C \right) \quad (6)$$

Differentiation of this with respect to R leads to

$$d \frac{f(R)}{dR} + \frac{1-3R}{R(1-R)} f(R) = \frac{\cot \alpha}{R(1-R)} \quad (7)$$

The general solution of this is

$$f(R) = -\frac{\cot \alpha}{2R} + \frac{C_1}{R(1-R)^2} \quad (8)$$

Substituting equation (8) into equation (6), we have

$$C_1 = -\frac{\cot \alpha}{2}$$

hence, from equation (4)

$$\dot{u} = -\frac{\cot \alpha}{2} \left(1 + \frac{1}{R}\right) + \frac{C_1}{R(1-R)}$$

The boundary condition that $\dot{u} = -\cot \alpha$ at $R = 1$ requires that $V_1 = 0$. This is also required to avoid infinite \dot{u} at $R = 1$. Finally we obtain

$$\dot{u} = -\frac{\cot \alpha}{2} \left(1 + \frac{1}{R}\right) \quad (9)$$

$$\dot{v} = \frac{\cot \alpha Z}{2R} + 1$$

APPLICATION OF A MODULAR APPROACH TO ESTIMATE LOAD AND ENERGY IN CLOSED DIE FORGING

by

S. K. BISWAS* and B. W. ROOKS*

SUMMARY

This paper presents a modular simulation technique which predicts the forging characteristics of given components when deformed on work-restricted forging machines. The technique considers each component to be made up of simple modules, and the whole forging operation is simulated approximately by linking the deformation of each module in a sequential manner. The slab method is used to calculate the stress distribution and the load at the completion of each module and hence the energy to deform that module. The total energy required to deform the whole component is obtained by summing the deformation energies of the individual modules.

The theoretical predictions of the modular technique are compared with results from experiments on a Petro-Forge high-speed hammer. In these tests two component shapes are forged in medium carbon steel (EN8), each with two types of flash opening and under three conditions of forging temperature.

As there is, in general, a good agreement between the theoretical and the experimental predictions, it is concluded that the technique could form the basis of a method of direct use to industry. The application of the technique under industrial conditions is discussed, and whilst application to axisymmetric components of the type tested is relatively easy it is acknowledged that there are problems to be overcome, particularly with more complex components. Also the necessary technical information to implement the technique is discussed and it is concluded that more precise data with respect to die friction under production forging conditions is required.

INTRODUCTION

Drop forging companies are very much craft-based, relying on the experience and skills of their personnel. A company is able to attract profitable business by utilizing this experience to predict technical requirements such as forging weight, number of stages and strokes, die size, die life, size of hammer or press and so on. Some of these factors could be predicted by analytical methods, for example techniques are available for estimating die loads and deformation energies in forging. Die loads affect die design and selection of the press or hammer, whilst deformation energy determines the number of stages and strokes and therefore the production rate to produce a component.

Two methods available for the prediction of forging load are the slip-line field and upper bound theories. These approaches have been applied to steady-state axisymmetric deformation problems such as extrusion and upsetting. However deformation energy is calculated from the area bounded by the load-deformation characteristic and therefore requires knowledge of the load history. Hence, to determine the energy requires an estimate of the most

suitable velocity field to satisfy the upper bound condition at each stage of deformation. Such an approach demands great expertise and considerable experience on the part of the designer, as well as very large computing facilities, computer programming knowledge and above all the time in which to carry out the exercise. These requirements are not normally available to the small and medium-size forging company. To attempt a general solution to the upper bound technique which can be readily adapted to a wide range of components is impossible at the present time, and therefore the technique is still limited to academic-type exercises.

The other approach to determine forging characteristics is the slab method, which in its original form was applicable only to steady-state problems. Kobayashi¹ used this method to predict the load at the last stage of forging a complex axisymmetric part. The dynamic problem has been attempted, primarily by Altan² who used a computer simulation technique that utilized the slab method to establish the characteristics of a complex component by estimating the stress distributions in assumed flow models or zones at incremental steps of the deformation stroke. The

* Department of Mechanical Engineering, University of Birmingham

authors³ have modified Altan's method by introducing a zonal flow stress dependent on the changing strain rate and temperature in each zone, and have extended it to apply to work-restricted forging machines. Although this technique has found good experimental agreement³, its application, as with the slip-line field method, still has practical limitations. A designer would need to assume the existence of certain flow models to describe the deformation of a particular component and would have to determine the optimum geometry at each incremental stage of deformation. Thus, again specialized skills and large computer facilities are necessary. Whilst the simulation technique could be used for one component at a time it does not lend itself to a general application and therefore the computer program for each component has to be developed separately. It is not even generally applicable to computer-aided design or production planning such as preform specification unless a large computer with a memory capable of storing the individual program of each component shape likely to pass through a company's order books can be dedicated to this one task.

This paper presents a technique which is a simplification of the slab simulation technique and is aimed at overcoming some of the problems of application just discussed. With the technique developed a general approach can be made to estimate load and energy and a wide range of components can be analysed from one basic programme. Each component is broken down into modules of deformation, for example upsetting, extrusion and flash formation, all of which occur to a greater or lesser extent in the commoner types of axisymmetric forgings. A computer sub-routine has been developed for each of these modules and gives the peak load and the total energy for each module of deformation. The function of the basic programme is to 'string' these modules together in a sequential manner and to calculate the load-deformation characteristic and hence the total energy required to produce the component. The paper also includes a verification of this technique in which the theoretical characteristics are compared with experimental results obtained from forging the four components shown in figure 1 on a Petro-Forge high-speed hammer under various conditions of strain rate and temperature.

THE TECHNIQUE

In the step-by-step computer simulation technique^{2,3}, the forging at any stage of deformation is assumed to consist of zones of characteristic metal flow. On the introduction of an increment of deformation the stresses in all the zones change simultaneously, giving rise to a change in the stress distributions which determines the energy required to cause that increment. The technique developed in this paper differs in that the deformation is simulated by dividing the die cavity into separate modules which generally deform sequentially rather than simultaneously. Each of these modules is identifiable as a cavity or region of metal flow. The filling of each of these modules occurs in sequence with the exception of one, the central flange or body of the component,

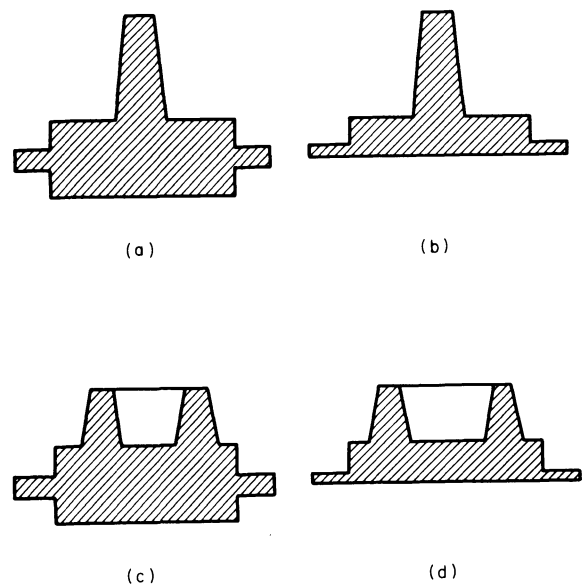


Figure 1. Forged Components.

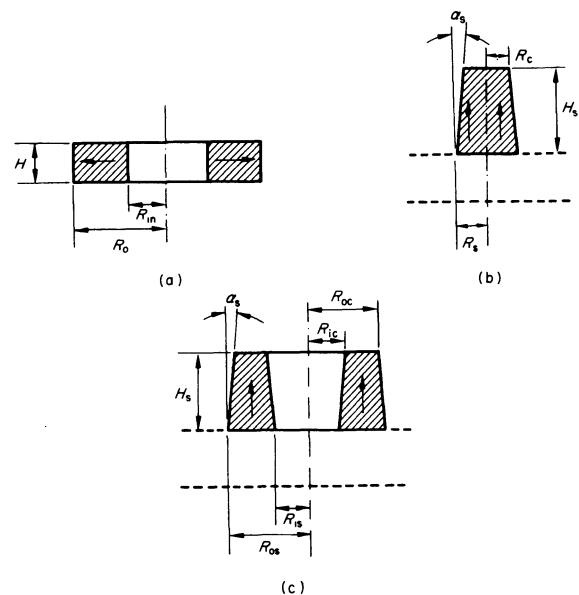


Figure 2. Deformation modules showing flow: (a) Upsetting—module 1; (b) central extrusion—module 2; (c) annular extrusion—module 3.

which loses metal to the modules. Thus the completion of each module is accompanied by a reduction in the volume of the central flange by the amount required to fill that module. The total load required to deform the central flange and the filled module is estimated and represents the 'instantaneous' load corresponding to the current flange height. It is assumed that the change in load from the completion of one module to the next is linear. Thus the energy required to deform a module is the product of the mean load and the change in height of the flange.

The majority of axisymmetric components used in practice can be constructed by combining any two or three of the modules shown in figure 2. The equations for calculating the load (by the slab method^{2,4}) required to deform these modules are

given in Appendix 1. Thus for estimating the approximate load-deformation characteristics of an axisymmetric component, the designer is required:

- (1) to choose the necessary combinations of modules to construct the component, and
- (2) to select the sequence in which these modules are going to be deformed.

APPLICATION OF THE TECHNIQUE

An example is given of the application of this technique to the component shown in figure 1(c).

The cavity of the die used for forging this component can be considered as made up of three constituent cavities or modules. These are, in the sequential order of filling, the central flange cavity (module 1, figure 2(a)), the extrusion cavity (module 3, figure 2(c)) and the flash cavity (module 1, figure 2(a)). The complete sequence is shown in figure 3, commencing with the billet (figure 3(a)) and finishing with the completely formed component (figure 3(d)),

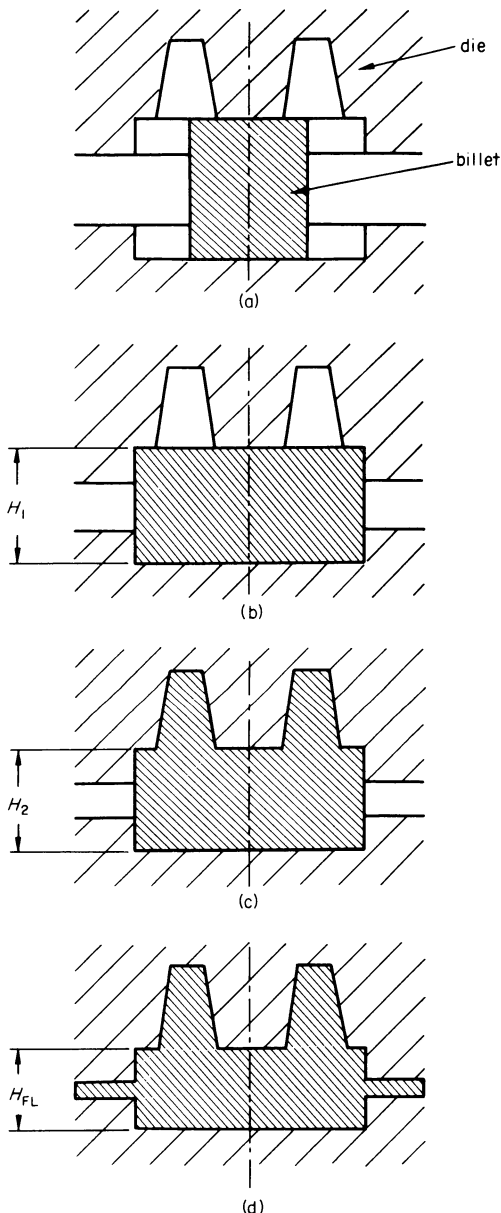


Figure 3. Sequence of modular filling of the die.

whilst the corresponding load-deformation characteristic for this above sequence is given in figure 4.

Stage 1 of the sequence is completed with the filling of the central flange with a flange height H_1 (figure 3(b)) and load P_1 (figure 4); the latter is calculated using Equation (A1) of Appendix 1. The energy absorbed in stage 1 (E_1) is given by the area of triangle OAB in figure 4. Similarly completion of stage 2 occurs with the filling of the extrusion cavity (figure 3(c)) when the flange height is equal to H_2 .

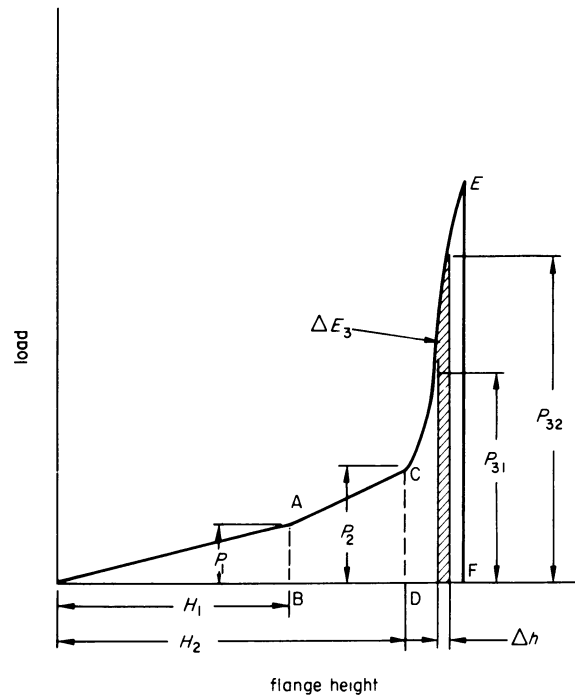


Figure 4. A typical load-deformation curve obtained by the application of the modular technique.

The load to extrude metal into the extrusion cavities (that is module 3) is P_{21} , which is calculated by proper substitution into Equation (A3) of Appendix 1, whilst the load to reduce the flange height from H_1 to H_2 (that is module 1) is P_{22} . Thus, the total load at the completion of stage 2 is P_2 , where

$$P_2 = P_{21} + P_{22}$$

The energy absorbed during stage 2 (E_2) is given by the area ABCD in figure 4. Thus the total energy absorbed at the end of stage 2 is equal to $E_1 + E_2$.

The final stage of the deformation sequence reduces the flange height from H_2 to the final flange height H_{FL} during which the excess metal is forced through the flash gap. As the actual load-deformation characteristic during this stage is comparatively steep, a large error in energy estimation can be expected if a linear relationship based on the final peak load was assumed to represent this last stage. Therefore this last stage of deformation (stage 3) is simulated to proceed in small steps. A small step of deformation is introduced by reducing the height of the flange and the flash by an amount Δh . The outer radius of the material extruded through the flash is calculated from volume constancy considerations. The load equation for module 1 is applied to estimate the load P_{F1} in the flange and that in the flash, P_{F2} .

The total load after a die descent Δh , as indicated in figure 4, is given by

$$P_{32} = P_{F1} + P_{F2}$$

The energy absorbed during a die descent Δh is shown by the shaded element in figure 4 and is equal to

$$E_3 = \frac{1}{2} \Delta h (P_{32} - P_{31})$$

Referring to figure 4 the total energy absorption during stage 3 is given by the area FOCE.

FLOW STRESS ESTIMATION

The main feature of hot-working is that large strains are applied to materials at high rates of strain and at temperatures above half the melting point temperature of the materials. Hot-working is similar to steady-state creep in that it is temperature and strain-rate sensitive, and equations similar to the ones used to describe stress-strain relationship in creep have been found⁵ to fit hot-working data reasonably well. The flow stress in a module is estimated by substituting the values of the average modular strain rate and modular temperature into one such equation (see Equation (A4) of Appendix 2) expressed in terms of strain rate and temperature-independent material constants. These constants are determined by applying a statistical method^{4,6} to flow stress data⁷.

In reality the strain rate varies within the forging and also throughout the deformation process. Here, for simplicity, the average value of strain rate is assumed to prevail uniformly within a module and throughout deformation, although it may differ for each module. This average value is estimated by proper substitution into Equations (A5), (A6) and (A7) of Appendix 2.

In determining the temperature of a module the speed of forging and the ratio of surface area to volume must be considered. If the ratio of surface area to volume of a module is high and the forging velocity is low (as in a hydraulic press) the temperature drop due to conduction can be considerable and must be taken into account. However, if the process is a high-speed one, as in the present case, the temperature drop due to conduction even after full extrusion of the flash can be very small. Thus under high-speed conditions the only temperature change calculated for each module is that using Equations (A8), Appendix 2.

ESTIMATION OF THE LOAD AND ENERGY CHARACTERISTICS FOR WORK-RESTRICTED FORGING

The modular technique estimates the value of deformation energy required to produce a certain billet reduction. In practice a certain portion of the energy delivered by the platen at impact is used in causing vibration of the machine frame and in straining various machine members. This amount of redundant energy (which though delivered is not used in deforming the workpiece) is a characteristic of the machine structure and a function of the peak load. In

the case of the Petro-Forge used in this investigation the empirical relation for redundant energy derived by Chaherli⁸ is given by

$$E_{ST} = F^2 \times 3.76 \times 10^{-5}$$

where F is the peak load in tons.

The total forging energy required to effect a certain reduction as calculated by the modular technique thus consists of the redundant energy E_{ST} and the deformation energy estimated at an assumed impact velocity. Hence this assumed velocity is related to the estimated forging energy by its influence on the flow stress of the workpiece material through strain-rate effects. The energy characteristics of an actual hammer operation, on the other hand, are related to the impact velocity V by the fixed relationship

$$E = \frac{1}{2} m \left(\frac{M}{M+m} \right) V^2 = M' V^2 \text{ (say)} \quad (1)$$

where m is the moving mass of the upper platen and M is the moving mass of the lower platen (counter blow machine).

The theoretically devised energy-velocity relationship is not necessarily the same as the actual relationship and in that case the estimated characteristics are not applicable to a work restricted system. To overcome this problem a method was developed to estimate the load and energy characteristics, compatible to a work-restricted system by constructing theoretical energy-velocity characteristics at a constant reduction. These characteristics were found to be linear for the components forged under the present investigation. This implies that there is one value of compatible velocity for a given reduction and the required energy-reduction characteristic consists of points each of which represent a different compatible impact velocity.

If the equation of an estimated energy-velocity characteristic for a reduction x_i is given by

$$E = A_i + B_i V \quad (2)$$

where A_i and B_i are constants corresponding to reduction x_i , substituting Equation (1) into Equation (2) yields

$$M' V^2 = A_i + B_i V \quad (3)$$

This equation was found to have only one positive root which is the desired compatible value of V_i . Substituting the compatible V_i into equation (1) gives the compatible energy requirement to effect reduction x_i . The corresponding compatible load is estimated by substituting the velocity V_i into the relevant constant reduction load-velocity equation.

The above procedure has been incorporated into the computer program implementing the modular technique so that given the moving weights of the upper and lower platens, the outputs from the program gives the compatible characteristics.

EXPERIMENTAL DATA

All the tests were carried out on a Mk IIA Petro-Forge high-speed hammer⁹, which has a nominal output rating of 10 000 ft lbf given by a platen

impact speed of 50 ft/s. Lower energies are obtained at correspondingly lower impact speeds. In this investigation the impact speeds covered the range of 15–45 ft/s (1050–9450 ft lbf).

Instrumentation fitted to the forging machine included a strain-gauge load cell¹⁰ placed under the bottom die to measure peak die load, a capacitance-type displacement transducer to measure upper platen position and a magnetic induction-type transducer to measure the platen velocity.

The four component designs forged in this investigation are shown in figure 1 and the corresponding billet sizes are given in Table 1.

TABLE 1

Component	Height H_0 (in.)	Diameter D_0 (in.)
A and C	1.5	1.0
B and D	1.125	1.0

Both sets of billets were from 'bright' EN8 bar, those of component B and D being cropped to length whilst those of component A and C were sawn.

The tests were carried out at three nominal forging temperatures of 1000, 1100 and 1200°C, the billets being heated in an electric muffle furnace.

No lubrication was applied either to the billets or to the dies and hence for estimation of the die-billet interface friction, metal-to-metal contact was assumed at this interface.

For each test the peak load, peak displacement and peak velocity were stored in and recorded with the aid of Kistler peak volt meters. The peak voltage output from the velocity transducer was computed to yield the kinetic energy at impact.

EXPERIMENTAL AND THEORETICAL RESULTS

A complete set of theoretical curves and superimposed experimental points are given in figures 5–8, each at three forging temperatures of 1000, 1100 and 1200°C. In figure 5 the load–flange reduction characteristic $(H_0 - H_{FL})/H_0$ for components A and B are given, whilst figure 6 shows the deformation energy–flange reduction characteristics for the same two components. Figures 7 and 8 show deformation energy–flange reduction characteristics respectively for components C and D. The theoretical curves in every case are seen to be in three distinct stages which correspond to the filling of the three module cavities exemplified respectively by figure 2. Generally, the agreement between theory and experiment is seen to vary from fair to good.

For component A, the load (figure 5) is significantly underestimated during the second stage and to a lesser extent in the third stage. The load estimations for component B (figure 5) are seen to be in close agreement with the experimental points particularly in the important final stage.

A comparison of the theoretical and experimental energy characteristics for components A and B (figure 6) is less easy because of the increased scatter of the experimental points. For component A the energy is underestimated during stage 2 but improves as de-

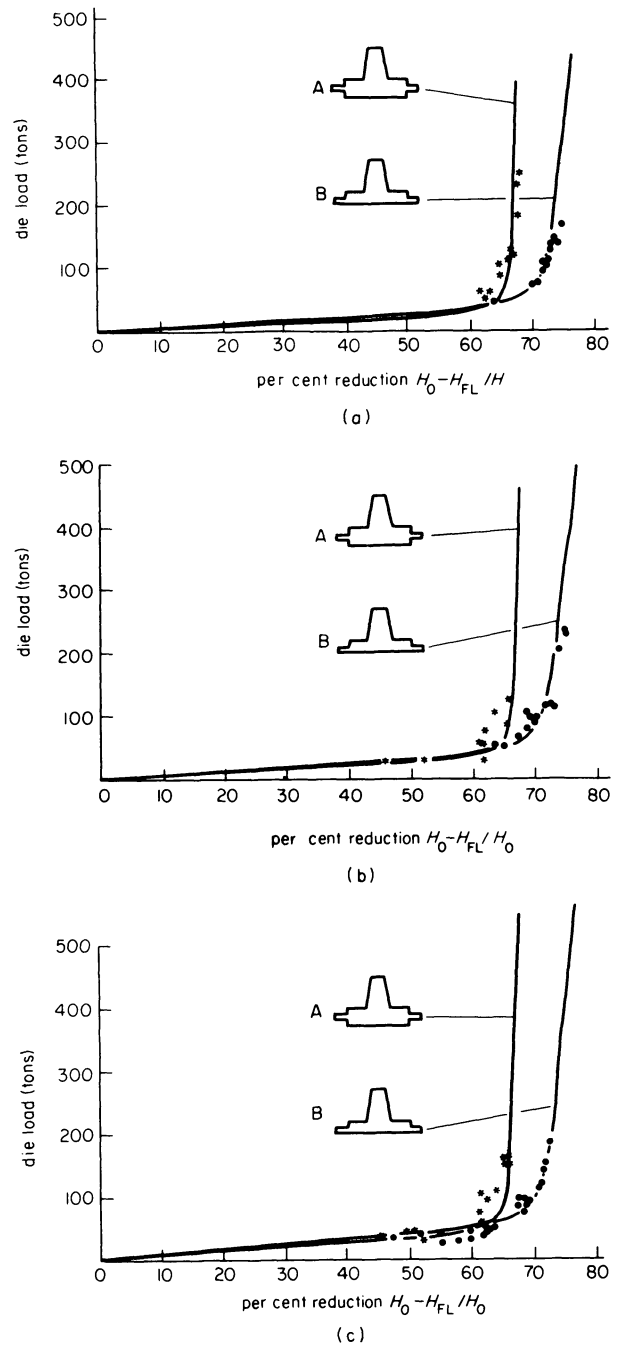


Figure 5. Variation of load with percentage reduction for components A and B at three forging temperatures (a) 1200°C, (b) 1100°C, (c) 1000°C.

formation increases whereas for component B energy tends to be overestimated during stage 2, but also agreement improves during the final stage.

For components C and D the load characteristics (figure 7) show a good agreement between theory and practice, similar to those for components A and B.

The theoretical energy characteristic for component C (figure 8) tends to under-estimate the experimental values at all reductions, whereas that for component D only under-estimates at the higher reductions, good correlation occurring at low reductions. This latter observation contrasts with that for component B (figure 6) which shows a theoretical over-estimation at low reductions but improves at high reductions.

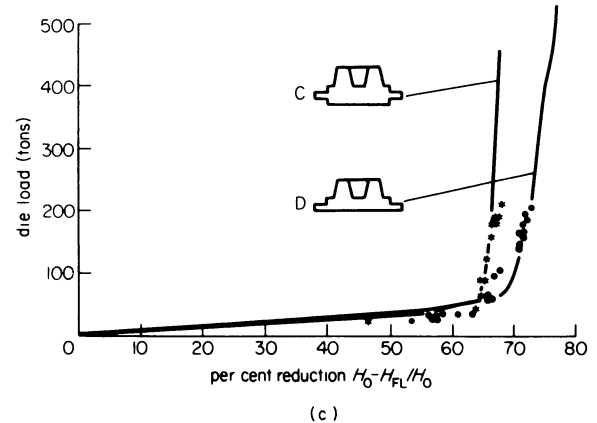
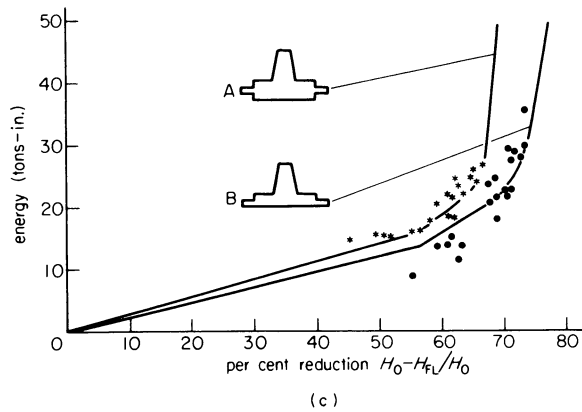
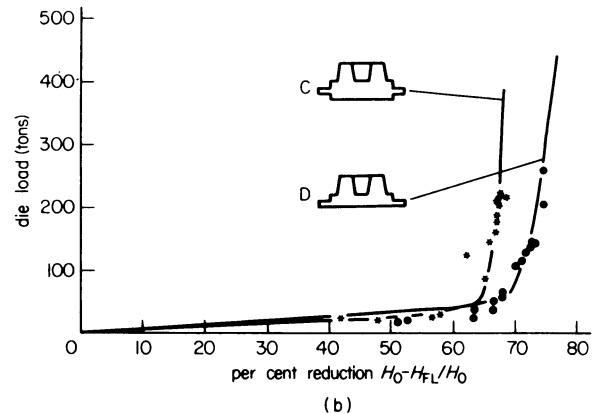
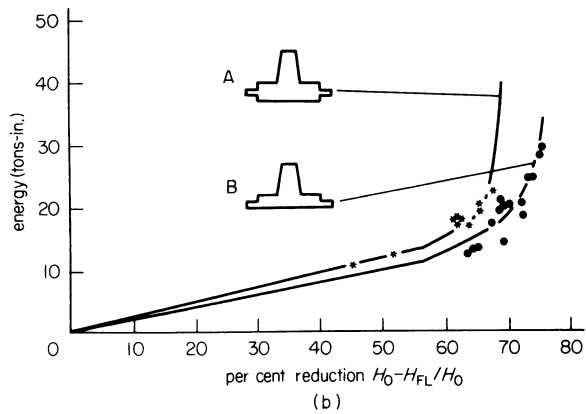
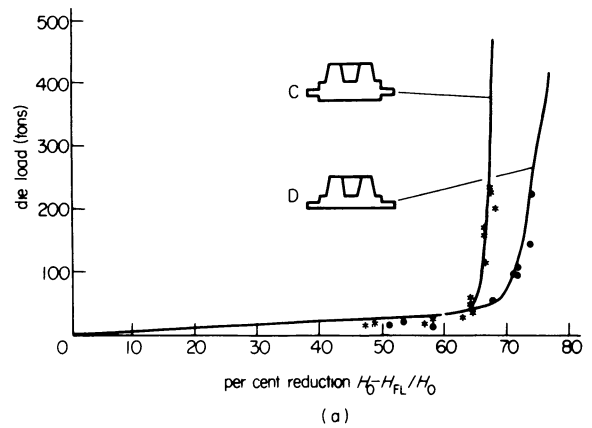
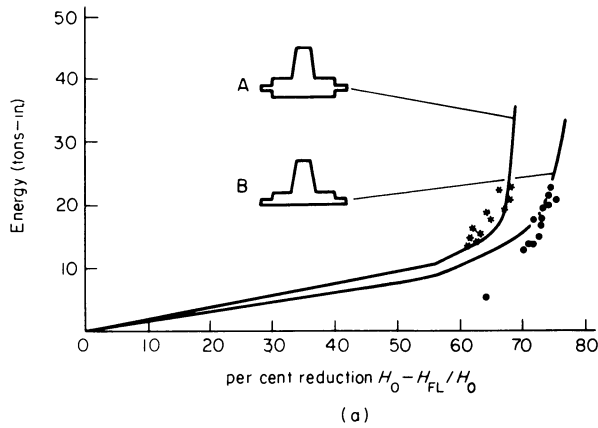


Figure 6. Variation of energy with percentage reduction for components A and B at three forging temperatures (a) 1200°C, (b) 1100°C, (c) 1000°C

Figure 7. Variation of load with percentage reduction for components C and D at three forging temperatures (a) 1200°C, (b) 1100°C, (c) 1000°C.

DISCUSSION OF RESULTS

From plasticity theory considerations the simplified model used in the modular technique will inevitably lead to errors. Thus, as the results have shown, discrepancies do occur. However they are explainable and in any case do not detract from the suitability of applying the technique in a practical situation.

Initially it is relevant to discuss the comparison between the theoretical and experimental characteristics so as to highlight the limitations of the modular technique.

The greatest error in prediction of the transient values occurs in stage 2. This is particularly true of the load characteristics (figures 5(a), 5(b), 7(a) and 7(b)). This is understandable if it is considered that the modular technique assumes stage 2 to be entirely

the filling of the rib cavity, whereas in reality there is simultaneous formation of the flash, upsetting of the central flange and extrusion into the rib cavity³. Thus the theoretical prediction takes an average line in stage 2, over-estimating in the first part and under-estimating in the second part.

The transient errors in the energy prediction are less pronounced in stage 2 due to the fact that energy is a cumulative parameter so that variations are damped out to some extent. However, as the errors are also cumulative they result in an energy prediction less accurate than the load prediction during the final stages of operation, particularly for the double-rib components (C and D), as is seen in figures 8(a) and 8(b).

Whilst the agreement between the theoretical and

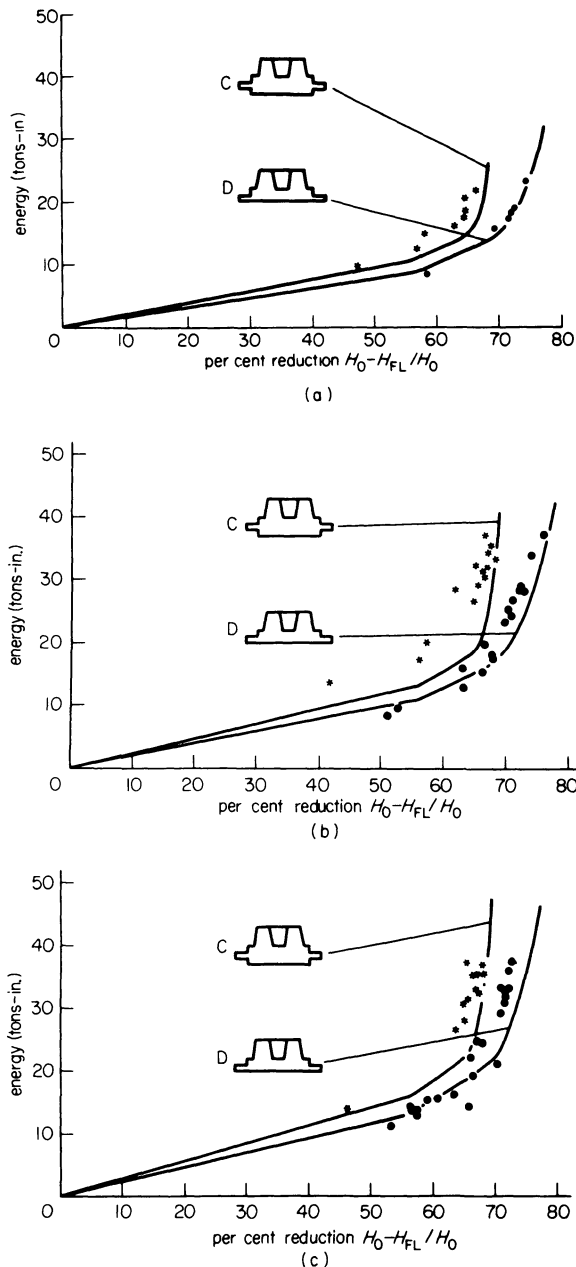


Figure 8. Variation of energy with percentage reduction for components C and D at three forging temperatures (a) 1200°C, (b) 1100°C, (c) 1000°C.

experimental load characteristics is extremely good over the last stage of deformation, some over-estimation does occur at high reductions, as can be seen in figures 5(b), 5(c), 7(b) and 7(c). This is due to the assumed model in the final stage which could be improved by adopting either the disc model of Kobayashi¹ or the lens-shaped model of Altan² to describe the shear surfaces that exist within the flange cavity whilst the flash is being extruded. This would increase the complexity of the technique and at this stage is not considered justifiable. Thus, the comparison between the theoretical prediction and the experimental results can be summarized as follows:

- (1) During the final stages of all four components the load prediction is within 10 per cent of the experimental results, which is considered to be a very close agreement.

- (2) For the single-rib components (A and B) the total energy required to achieve a high degree of filling is predicted within 15 per cent; this figure is reduced to within 6 per cent as the forging temperature is increased, due to a decrease in the experimental error.
- (3) For the double-rib components (C and D) the total energy required to achieve a high degree of filling is predicted within 30 per cent at a forging temperature of 1000°C but within only 5 per cent at 1200°C.

PROBLEMS OF IMPLEMENTATION

For industrial application purposes the above accuracies of prediction are very good, and if they could be achieved under industrial conditions it is certain that the industry would be satisfied. However it is acknowledged that the experimental results were achieved under laboratory conditions and therefore will be considerably more accurate than would be achieved under production conditions. Further, the components are not particularly complex and are familiar to the investigators, so that the choice of modules and their sequence was a relatively simple problem. More complex and less familiar components could reduce the accuracy of prediction. However, the results are considered to indicate that relatively accurate predictions could be made under industrial conditions, and the co-operation of any industrial company in applying this technique to real components would be welcomed.

The application of the technique breaks itself down into two problems. The first is the choice of modules to achieve the desired level of deformation. These can be selected with experience and then the technique presented in this paper can be utilized to predict levels of load and energy required; computer sub-routines requiring a relatively small computer are now available for this purpose.

The second problem is the provision of the data which is required to implement these computer programs. The following are required:

- (1) the initial billet or preform and final forging dimensions—these should be generally available from production drawings,
- (2) material structural constants—provided the material specification can be given, the constants relating the materials flow stress to forging temperature and velocity are available and in a computer-aided system could be provided 'on file',
- (3) billet temperature;
- (4) impact velocity;
- (5) friction coefficient.

Whilst (1) and (4) are relatively easy to obtain, the friction coefficient value is likely to prove the most difficult. If no lubricant, coolant, sawdust or the like is applied to the die, that is dry friction conditions, a reasonable estimate of the coefficient can be made. Otherwise the amount of information relating the friction coefficient to the type of lubricant, the quantity, die temperature, die geometry and so on is minimal. At present this is a problem that is not

easily overcome and can only be finally resolved by further experiment both in the laboratory and by collection of data from industry, the latter requiring some basic instrumentation of a production forging machine.

This paper, therefore, presents a technique which is merely a starting point and can only be advanced if the industry so desires it. Perhaps the discussion on this paper will reveal how further work should be directed.

ACKNOWLEDGMENTS

The work presented in this paper was part of the programme to develop an automatic multi-station forging unit sponsored by the Wolfson Foundation and carried out under the general supervision of Professor S. A. Tobias.

The authors acknowledge the financial support given by the Wolfson Foundation and are grateful to Professor Tobias for his friendly help and guidance.

APPENDIX 1 MODULAR LOAD EQUATIONS

(a) Module 1 load Equation ² (figure 2(a))

$$P_1 = 2\pi \left[\frac{A_1}{3} (R_0^3 - R_{in}^3) + \frac{B_1}{2} (R_0^2 - R_{in}^2) \right] \quad (A1)$$

where

$$A_1 = \frac{-2\bar{\sigma}_1 f_1}{H}$$

$$B_1 = -A_1 R_0 + \bar{\sigma}_1$$

(b) Module 2 load Equation ² (figure 2(b))

$$P_2 = \pi R_s^2 K_1 \ln \left(\frac{R_s}{R_s - H_s \tan \alpha_s} \right) \quad (A2)$$

where

$$K_1 = \frac{2\bar{\sigma}_2 [f_2(1 + \tan^2 \alpha) + \tan \alpha]}{\tan \alpha}$$

(c) Module 3 load Equation ⁴ (figure 2(c))

$$P_3 = \pi (R_{os}^2 - R_{is}^2) \frac{K_3}{K_4} \ln \frac{K_2}{K_2 + K_3 H_s} \quad (A3)$$

where

$$K_2 = R_{os}^2 - R_{is}^2$$

$$K_3 = -2 \tan \alpha (R_{os} + R_{is})$$

$$K_4 = -2(R_{os} + R_{is})\bar{\sigma}_3(f_3 + \tan \alpha + f_3 \tan^2 \alpha)$$

APPENDIX 2 FLOW STRESS CALCULATIONS

The general flow stress equation for hot working^{3,5} is

$$\dot{\epsilon}_i = A_2 (\sinh A_1 \bar{\sigma}_i)^{A_3} \exp \left(\frac{-Q}{RT} \right) \quad (A4)$$

where *i* denotes a module.

The values of the strain rate for each module ($\dot{\epsilon}_i$) is determined as follows.

(1) *Module 1*. The mean strain rate (referring to figure 2(a)) is

$$\dot{\epsilon}_1 = \frac{V_0}{H_0} \quad (A5)$$

(2) *Module 2*. The mean strain rate⁴ (referring to figure 2(b)) is

$$\dot{\epsilon}_2 = \frac{6V_0 D_s^2 \tan \alpha \ln(D_c/D_s)}{(D_s^3 - D_c^3)} \quad (A6)$$

where

$$D_s = 2R_s$$

$$D_c = (D_s - 2H_s \tan \alpha_s)$$

(3) *Module 3*. The mean strain rate⁴ (referring to figure 2(c)) is

$$\dot{\epsilon}_3 = \ln \frac{R_{oc} - R_{ic}}{R_{os} - R_{is}} \frac{V_0 \tan \alpha (R_{oc} - R_{ic})}{(R_{oc} - R_{is})(R_{ic} - R_{is})} \quad (A7)$$

The value of the module temperature will be dependent on the initial billet temperature and the temperature rise (or loss) during forging. At high strain rates such as are used in the present investigation, there is insufficient time for the heat to be transferred to the surrounding media, and therefore adiabatic conditions apply. The adiabatic temperature rise at the workpiece cavity *T_A* is given by⁵

$$\Delta T_A = \frac{\Delta E/v}{C_p} \cdot \frac{1}{J} \quad (A8)$$

NOTATION

$\bar{\sigma}_i$	Modular flow stress
f_i	Modular die-material friction factor
$\dot{\epsilon}_i$	Modular mean strain rate
A_1, A_2, A_3, Q	Structural material constants
<i>T</i>	Workpiece temperature in K
<i>R</i>	Boltzmann's constant
ΔT_A	Temperature rise due to adiabatic heating
ΔE	Incremental energy input
ρ	Density of workpiece material
C_p	Specific heat of workpiece material
<i>v</i>	Volume of workpiece
<i>J</i>	Mechanical equivalent of heat

REFERENCES

1. S. Kobayashi, V. McDonald and E. G. Thomsen. Some aspects of press forging. *Intern. J. Mech. Sci.* (1960) 1, 282.
2. T. Altan, H. J. Henning and R. J. Fiorentino. The use of analytical method in predicting load and stresses in closed die forging. *A study of the mechanics of closed die forging. Final report.* Battelle Mem. Inst., 1969, chapters 3 and 4.
3. S. K. Biswas and B. W. Rooks. Application of a computer simulation technique to estimate load and energy in axisymmetric closed die forging. *Proc. 13th Int. MTDR Conf.*, Macmillan, 1973.
4. S. K. Biswas. Identification of the hot forging process. Ph.D. Thesis, Univ. of Birmingham, 1973.

5. S. K. Samanta. The dynamic compression of steels at elevated temperatures. *Proc. 11th Int. MTDR Conf.*, Pergamon, 1971.
6. F. Garfalo. An empirical relation defining the stress dependence of minimum creep rate in metals. *Trans. Met. Soc. AIME* (1963) p. 351.
7. H. Suzuki *et al.* Studies on the flow stress of metals and alloys. *Report of Inst. of Ind. Sci.*, The Univ. of Tokyo, 1968, p. 117.
8. A. Chaherli. Ph.D. Thesis, University of Birmingham, 1971.
9. L. T. Chan. Ph.D. Thesis, University of Birmingham, 1967.
10. S. C. Jain and E. Amini. Development of a short load cell for metal forming application. *Proc. 9th Int. MTDR Conf.*, Pergamon, 1969.

THE ELEVATED TEMPERATURE EXTRUSION OF FLUTED SECTIONS IN HIGH-SPEED STEELS

by

C. E. N. STURGESS* and T. A. DEAN*

SUMMARY

Data are presented for the extrusion of fluted products of 0° , 15° and 30° spiral angles in 18 : 4 : 1 (T1) and 6 : 5 : 2 (M2) high-speed steels. The tests were performed at reductions of 1.65 : 1 and 2.6 : 1 over the temperature range 800–1100°C, at low (0.4 m/s) and high (8 m/s) velocities using Copaslip as a lubricant.

At a given temperature and reduction larger helix angles require higher extrusion pressures. At a given temperature and helix angle the pressure differences for the various helix angles are roughly the same for the two reductions in area.

The optimum conditions to produce sound, useful products were the warm (800°C) extrusion of 6 : 5 : 2 (M2) high-speed steel at high velocities.

NOTATION

A, B, C	Constants
A_1	Area at point 1
A_2	Area at point 2
A_m	Mean area
a	Side length of idealized square product
D	Diameter of container
$k = Y/2$	Shear resistance to deformation
L	Length of container
L_h	Length of helical portion of die
m	Frictional shear factor
P	Basic extrusion pressure
P_E	Extrusion pressure with container friction
P_T	Torsional work done/unit volume
r	Radius of container
r_c	Circumscribed radius of square product
R	Reduction in Area
T_1	Plastic Torque of section at 1
T_2	Plastic Torque of section at 2
T_m	Mean plastic torque
V_m	Mean volume of helical die zone
Y	Flow strength of material
θ	Product rotation
β	Product spiral angle

INTRODUCTION

High-speed steels are expensive materials, current prices being approximately £2/kg. Although largely superseded for turning tools by sintered carbides, high-speed steels still form the mainstay of the milling

cutter and drill industries. The presently most popular grades of high-speed steel are 6 : 5 : 2 (M2) and 18 : 4 : 1 (T1).

The conventional manufacture of fluted milling cutters by machining results in 50 per cent material wastage, accounting for around one-sixth of the cost of the final product¹. In this environment it appears useful to investigate the use of forming techniques as material savings are one of the major attributes of such processes.

To date although drills, which have tapered webs, are produced by rolling, the forming of prismatic multi-flute shapes has received little attention.

The Allegheny Ludlum Corporation² have been extruding straight-flute, somewhat idealized, cutting-tool shapes in M2 since 1952, the weight saving is reckoned to be around 35 per cent.

In 1959 Sukolski and Hoyle³ extruded T1 into round and flat bars with reductions up to 54 : 1 at temperatures of 1250–1280°C. The products were metallurgically comparable to rolled or forged sections.

Pugh⁴ has hydrostatically extruded spiral drill sections in k9 tool steel, but at present this technique is not considered commercially viable.

The only attempt to conventionally extrude helical milling cutters was conducted by General Dynamics (Fort Worth Texas) in 1962⁵, when slow, spiral two-flute cutters were produced on a Dynapak machine. It was claimed that the production costs could be cut from \$20 to \$7 per piece by this technique.

* Department of Mechanical Engineering, University of Birmingham

The foregoing indicates that considerable economic benefits would be achieved if milling cutter blanks could be formed. This paper describes part of an investigation designed to assess the overall feasibility of producing helically fluted cutter blanks from high-speed steels, by elevated temperature extrusion forging.

EXPERIMENTAL EQUIPMENT

Eccentric press

A Wilkins and Mitchell, single-point, 2000 kN capacity, 200-mm stroke eccentric press, capable of 45 strokes per minute, was used for the low-speed extrusions. The ram speed at impact was approximately 0.4 m/s.

Petro-Forge

High velocity (10 m/s) extrusions were performed on a Petro-Forge⁶ combustion-actuated high speed hammer.

The Petro-Forge employed for these investigations was a Mk 11B characterized by:

weight of moving platen	225 kg
working velocity	3–15 m/s
energy per blow	0.75–12.5 kJ

Extrusion tooling

Optimized duplexed assemblies were used for both container and dies⁷ (figure 1). The component cylinders of the duplexes were produced with matching 2° including tapers on the interface diameters, the interference fits were effected by pressing.

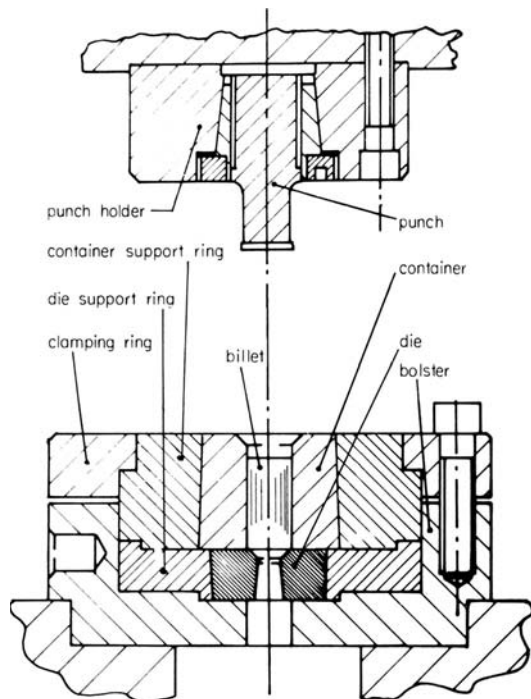


Figure 1. Extrusion tooling.

The tool material was a 5 per cent chromium hot-working die steel, heat treated to 48–50 Rockwell C. This material is sensitive to thermal shock, hence the tools were preheated before use by an electrical resistance bore heater.

The design sizes of the tools were:

Container

Cylinder radii	12.7, 32.5, 83 mm
Interference	0.3 mm (0.5%)
Maximum pressure capability	1250 MN/m ²

Dies

The extrusion dies were of 90° included-angle conical profile through which the forms shown in figure 2 were spark eroded.

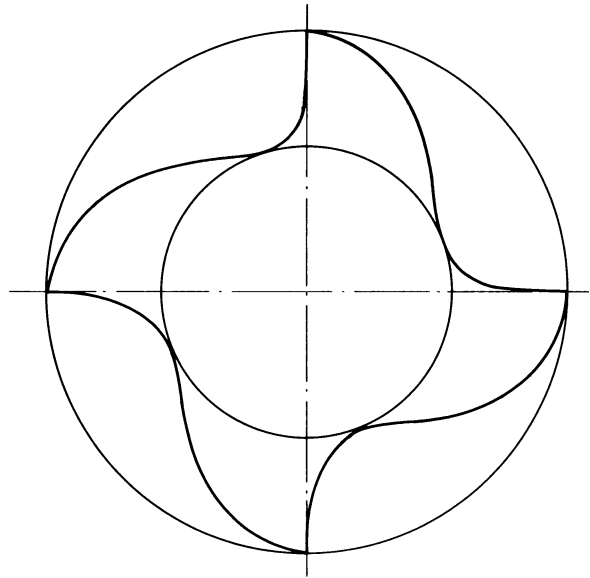


Figure 2. Fluted sections.

To enable the helical dies to be produced a rotating member and a former were interposed between the spindle and the electrode, figure 3, the electrode then describing a helix through the die. The design sizes of the dies were as given in Table 1.

The dies were relieved by spark eroding from the reverse side with oversized electrodes to give minimum lands of 1.5 mm (figure 4).

Extrusion punch

The extrusion punch was manufactured from a maraging steel, solution treated at 820°C and aged for 3 h at 500°C, to provide high strength (1850 MPa) with good ductility.

High-speed steel stock

The materials were obtained as 25.4-mm diameter, centreless ground bar in the fully annealed condition, the compositions and hardnesses were as given in Table 2. The billets for the trials were 25.4-mm diameter x 50 mm long.

Billet heating

The billets were heated in an electric muffle furnace for periods adequate to ensure through heating⁸.

Lubricants

The lubricant employed for the tests was Copaslip, a suspension of copper and lead particles in a bentone grease base, applied to the dies by brush immediately prior to extrusion.

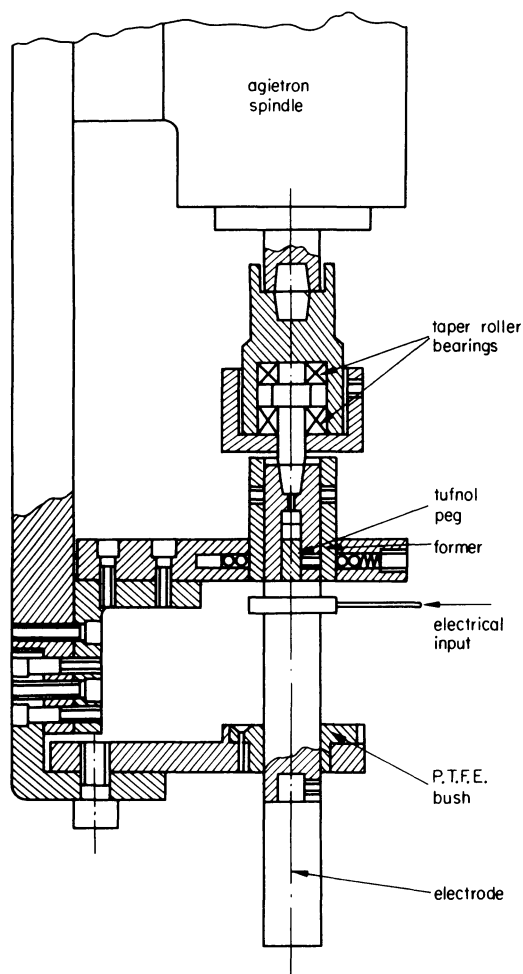


Figure 3. Adaptor for spark-eroding helical dies.

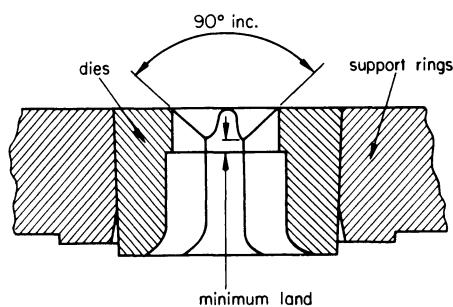


Figure 4. Extrusion dies.

Instrumentation

Load cell

Four Hawker Siddeley KW8/250 Ω strain gauges were positioned on the punch and arranged in a full bridge to eliminate the effects of bending stresses.

The load cell was used in conjunction with a 50-kHz carrier frequency impedance bridge amplifier. The amplifier possessed a 'flat' frequency response up to 5 kHz which was considered adequate to avoid attenuation of the output.

Displacement transducer

The movement of the punch during extrusion was monitored by the use of a coaxial capacitance displacement transducer, of the type described by Vinycombe and Martyr⁹, comprising a steel rod surrounded by concentric cylinders of Tufnol (dielectric) and brass (electrode). The capacitance of the instrument varied in proportion to the overlapped area of the rod and brass sleeve.

The output of the transducer was obtained via a Southern Instruments M 700 L gauge oscillator and an MR 200 FF.M. gauge amplifier.

Data recording

The outputs from the load cell and displacement transducer were simultaneously recorded on a Tektronix 564 storage oscilloscope fitted with a Polaroid camera.

Results and discussion

Typical products of the extrusion forging of milling cutter blanks in high-speed steel are shown in figure 5.

Elementary energy balance

In an attempt to isolate the major factors influencing the extrusion of helical sections a simple energy balance was used where the work done/unit volume in extrusion was added to the torsional work done/unit volume expended producing an idealized helical product. The section under consideration (figure 2) was taken as approximately equivalent to a square, the full plastic torque for a square being known¹⁰.

TABLE 1

Outside diameter of cutter (mm)	Reduction in area	Interface diameter (mm)	Outside diameter (mm)	Diametral interference (mm)	Pressure capability (MPa)
25.4	1.65:1	59.5	165	0.30	1250
19	2.5:1	49.5	165	0.25	1250

TABLE 2

Material	Major alloying constituents (%)					Hardness V.P.N.
	C	W	Mo	Cr	V	
18:4:1 (T1)	0.7-0.8	17.5-18.5	0.7 max.	3.75-4.25	1.0-1.25	278
6:5:2 (M2)	0.82-0.87	6.0-6.75	4.75-5.50	3.75-4.25	1.75-2.05	263

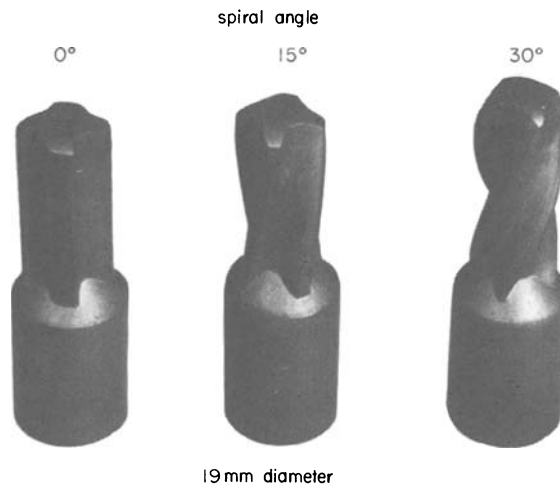
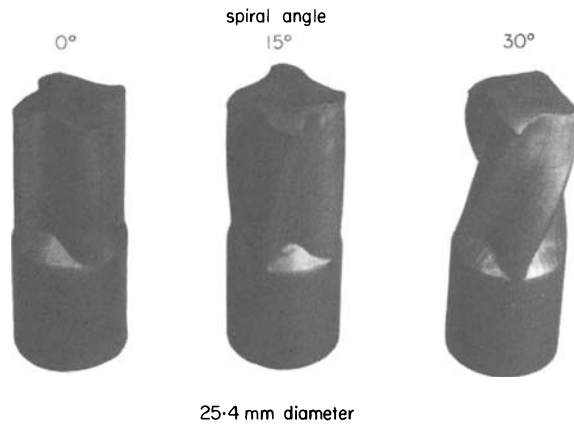


Figure 5. Extruded sections in high-speed steel.

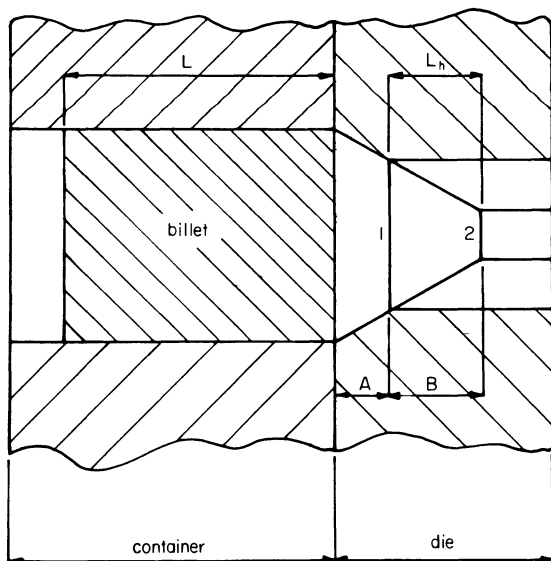


Figure 6. Interpenetration helical extrusion die.

Consider an interpenetration die of the type shown in figure 6, consisting of two elements:

(a) Pure reduction, terminating at 1 where the product is still circular.

(b) Reduction and twisting, from points 1 to 2 the product changes from a circular section to that of the desired twisted product.

Now extrusion pressure \equiv work done/unit volume and is conventionally expressed as¹¹

$$P = Y(A + B \ln R) \quad (1)$$

where $R = \frac{\pi r^2}{a^2}$

Plastic torque for circular section at 1 = $T_1 = \frac{\pi r_c^3 \cdot Y}{3}$ (2)

Plastic torque for square section at 2 = $T_2 = \frac{a^3 \cdot Y}{12}$ (3)

but $a = \sqrt{(2)}r_c$ (4)

Therefore $T_2 = 0.236 \cdot r_c^3 \cdot Y$ (5)

Arithmetic mean torque

$$T_m = \frac{T_1 + T_2}{2} = 0.643 \cdot r_c^3 \cdot Y \quad (6)$$

Arithmetic mean area

$$A_m = \frac{A_1 + A_2}{2} = \frac{\pi r_c^2 + a^2}{2} \quad (7)$$

Substituting from (4)

$$A_m = 2.571 \cdot r_c^3 \quad (8)$$

Therefore mean volume of die between 1 and 2 distance L_h apart is

$$V_m = 2.571 \cdot r_c^2 \cdot L_h \quad (9)$$

A helical section is defined by its helix angle and lead, therefore over the helical portion of the die from point 1 to point 2 the rotation of a section at 2 with respect to a section at 1 is

$$\theta = \frac{L_h \cdot \tan \beta}{r_c} \quad (10)$$

Therefore torsional work done/unit volume

$$P_T = \frac{T_m \cdot \theta}{V_m} \quad (11)$$

Substituting from (6), (8), (9) and (10) gives

$$P_T = 0.25 \cdot Y \cdot \tan \beta \quad (12)$$

In general $P_T = C \cdot Y \cdot \tan \beta$ (13)

Therefore summing the energy/unit volume requirements for extrusion and torsion gives the pressure for the extrusion of a helical section as

$$\frac{P}{Y} = A + B \cdot \ln R + C \cdot \tan \beta \quad (14)$$

EXTRUSION PRESSURES

The extrusion pressures required for the high and low velocity extrusion of 0° and 15° and 30° spiral fluted products at reductions of 1.65 : 1 and 2.6 : 1, and at temperatures between 800°C and 1100°C are plotted against reduction in figures 7(a) to (h).

Also included in figures 7(a) to (h) for comparison are the results for circular extrusions of high-speed steels over the same range of variables.

The extrusion pressures quoted throughout the following sections are the peak pressures, although these pressures are influenced by both container and die friction they are considered the most significant from a practical standpoint.

Results are only shown for 'high-speed steel' as the differences between the extrusion pressures for T1 and M2 were generally small, the results shown are the mean pressures for the two alloys. It is obvious from figures 7(a) to (h) that considerable scatter was encountered, however, this is unavoidable in hot-working processes.

For axisymmetric extrusion it has often been demonstrated that the relationship between the pressure required and reduction is of the semi-logarithmic form displayed by Equation (1).

The most recent and realistic theoretical treatment of extrusion is by Lambert and Kobayashi¹² who, using the upper bound approach including the effects of die friction, have estimated the coefficients in equation 1, for 90° included dies as

Die friction	A	B
0	0.45	0.90
Max.	0.60	1.38

In addition to die friction there are frictional losses at the billet/container interface. Allowing for container friction on the basis of a 'shear-factor'¹³ gives the actual extrusion pressure as

$$P_E = P + \frac{4mkL}{L} \quad (15)$$

The mean flow-stress-strain curves for the high-speed steels, obtained from compression tests¹⁴, are depicted in figure 8. Only one set of curves is shown as the strengths of the two high-speed steels considered (T1 and M2) are, within experimental error, identical. The strain rates for the extrusion operations were calculated using the root-mean-power method of Jonas and Chandra¹⁵.

It is obvious from figure 8 that the flow strength of high-speed steel continually decreases as the temperature is raised but, above approximately 800° ($0.55 T_m$), high velocities require higher pressures due to strain-rate effects.

To estimate container friction the pressure decrease during extrusion was noted and the shear factors (m) calculated from expression (15), the results are given in figure 9. As can be seen from figure 9, the performance of Copaslip as a lubricant in hot extrusion considerably deteriorates as the temperature exceeds 1000°C , both in high and low-speed working.

Combining the flow stress data in figure 8 with the frictional effects from figure 9 in Equation (15) gives the shaded zones in figures 7(a) to (h). In general it can be seen that the experimental extrusion pressures for the circular sections are within the predicted bands both at high and low speeds.

From the energy balance presented in the previous section the pressures required for the extrusion of the helical sections should conform to an expression of the type shown by Equation (14).

Comparing Equations (1) and (14) it is obvious that the additional pressures required for the spiral fluted sections should be independent of reduction and only influenced by the helix angle of the product. It has been demonstrated that in general the extrusion of non-circular sections require pressures little different from circular sections of the same reduction in area^{16,17}, implying that the coefficients in Table 2 apply also to non-circular sections.

Therefore from the above reasoning the pressure requirements for the fluted sections at the various spiral angles should be describable by lines drawn through the experimental results for the complex sections and parallel to the results for the circular sections, such lines are shown in figures 7(a) to (h).

The first point of note in figures 7(a) to (h) is that the results for the extrusion of the straight fluted shapes are lower than the equivalent pressures for the circular sections. This has been observed before¹⁶, but the discrepancies obtained here are larger than any reported figures.

It is evident from figures 7(a) to (h) that the majority of the data for the spiral fluted sections are consistent with the view that the results should be approximately parallel to the pressures required for the circular extrusion.

The other facet displayed by Equation (14) is that

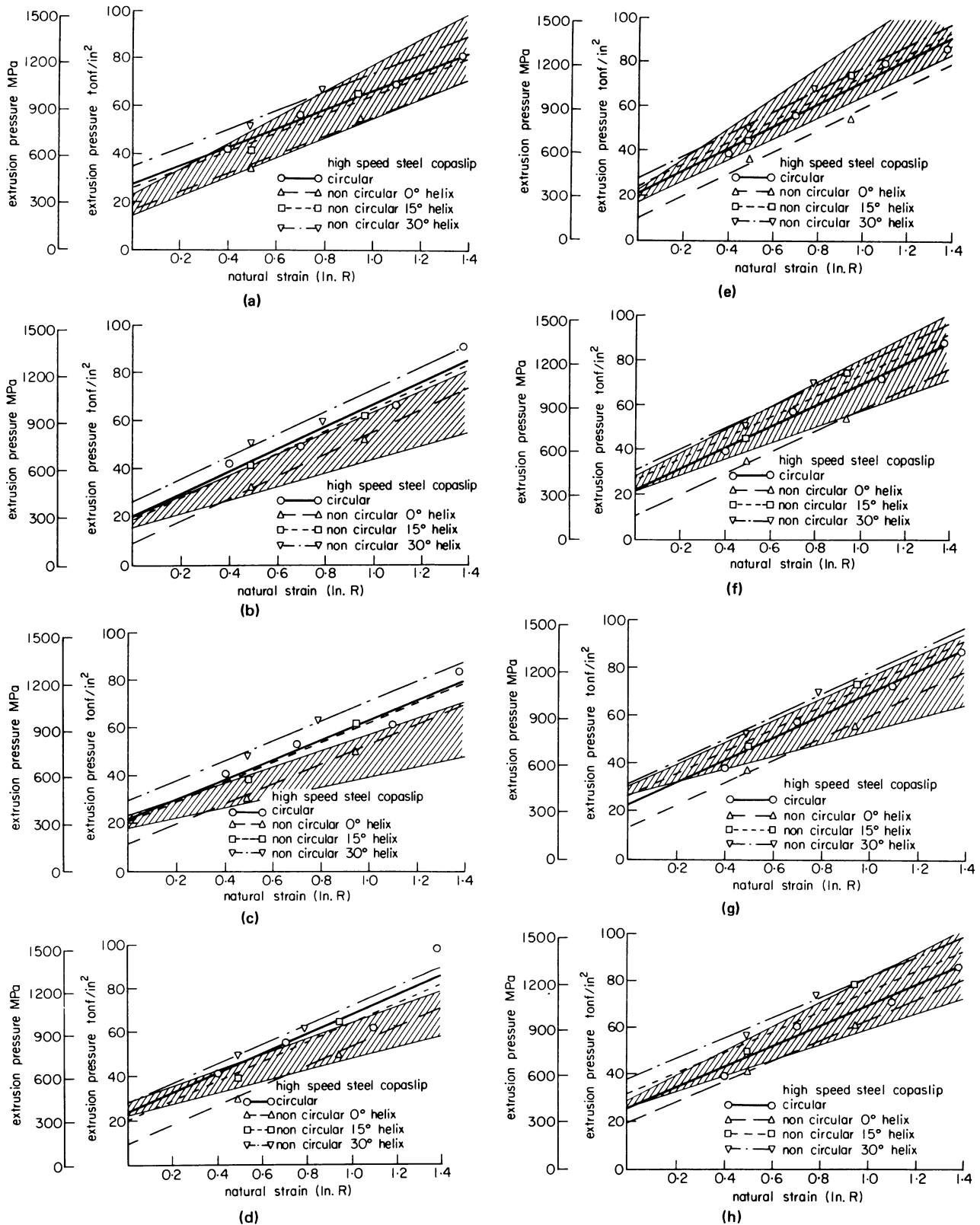


Figure 7. Extrusion pressures versus reduction for the extrusion of high-speed steels. (a) Low velocity 800°C. (b) Low velocity 900°C. (c) Low velocity 1000°C. (d) Low velocity 1100°C. (e) High velocity 800°C. (f) High velocity 900°C. (g) High velocity 1000°C. (h) High velocity 1100°C.

at a given reduction increasing the helix angle requires higher pressures, this can be seen to be true from the figures 7(a) to (h). However, for the particular tests conducted here, where helical products of 0°, 15° and 30° spiral angle were extruded, it would appear

from Equation (14) that there should be a slightly greater pressure difference moving from 15° to 30° products than from 0° to 15° spiral angles.

The coefficient in Equation (14) has been evaluated from figures 7(a) to (h), the values obtained are

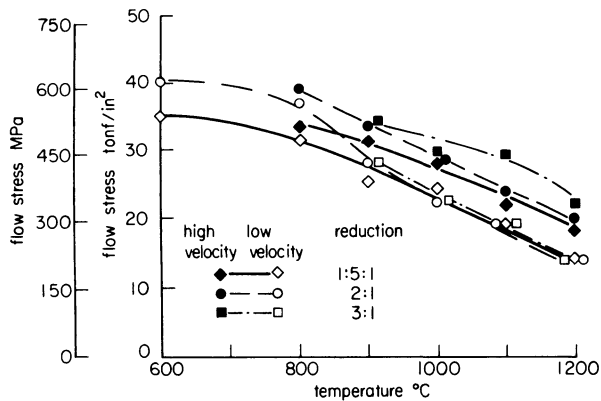


Figure 8. Stress-strain curves for high-speed steels at various temperatures and strain rates.

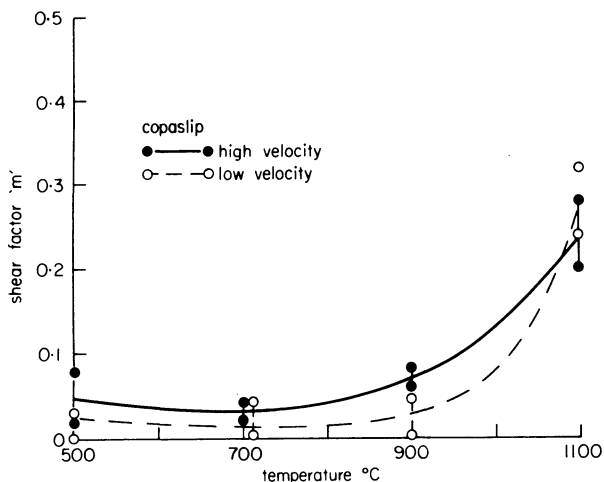


Figure 9. Shear factor (*m*) versus temperature for Copaslip lubricant when extruding high-speed steel.

shown in Table 3, from which it can be seen that 'C' is not a constant but varies with temperature, helix angle and speed.

No reasons can be advanced for the finding that higher coefficients are required for the 15° than for the 30° products or that the differences are more marked at high velocity. However, the variation of 'C' with temperature is thought to be largely due to die friction, omitted from the simple analysis, which would presumably vary with temperature in a manner similar to container friction as depicted in figure 9.

PRODUCT SOUNDNESS

At certain combinations of billet preheat temperature and extrusion speed the extrudates displayed extensive transverse cracking.

In figures 10(a) and (b) the cracking of 25.4-mm

diameter 30° helix components are illustrated, these results being representative of the remainder. Also included in figures 10(a) and (b) are the ductility data for 6 : 3 : 2 and 18 : 4 : 1, high-speed steels obtained by Berry¹⁸ from hot torsion tests. Although the ductility curve for the molybdenum steel is for a slightly different composition than that used here, it can be seen that the ductility exhibited by this alloy is roughly twice that of the tungsten material, this difference occurs throughout the two alloy series¹⁹.

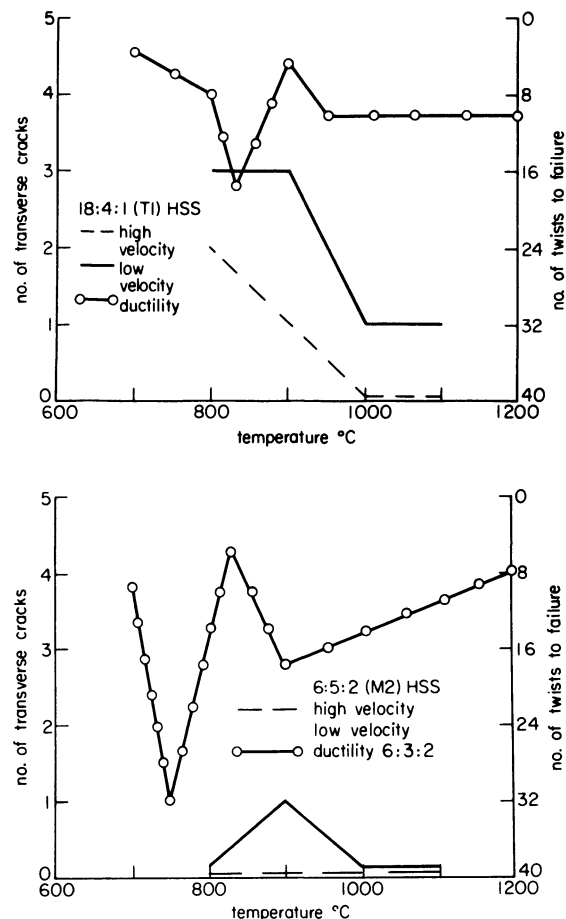


Figure 10. Incidence of cracked products of high-speed steel. (a) 18 : 4 : 1 (T1). (b) 6 : 5 : 2 (M2).

From figures 10(a) and (b) it is evident that the incidence of cracking is strongly related to the ductility of the material because the cracking is less for 6 : 5 : 2 than for 18 : 4 : 1, and the incidence of highest cracking occurs around 800°C when the material is very brittle. Furthermore the onset of cracking is dependent on the speed of working, for at high velocities the cracking is considerably reduced for 18 : 4 : 1, and completely eliminated for 6 : 5 : 2.

TABLE 3 Coefficient 'C'

Temperature (°C)	Spiral angle:	Low velocity		High velocity	
		15°	30°	15°	30°
800		0.98	0.92	1.45	0.88
900		1.22	1.08	1.74	1.08
1000		1.4	1.29	1.77	1.25
1100		2.06	1.73	1.73	1.20

The reasons for the reduction of cracking at high velocities are not clear but are in agreement with published results for hydrostatic extrusion. Oyane et al.²⁰ found that when extruding brittle materials by differential pressure extrusion the back pressure necessary to suppress cracking was reduced at high extrusion rates. These investigators attributed the crack suppression to reduced die friction at high speeds, unfortunately in the tests reported here no measures of die friction were possible.

MICROSTRUCTURES

Typical microstructures obtained from the materials at various temperatures are shown in figures 11(a), (b) and (c). No differences were observable between 18 : 4 : 1 and 6 : 5 : 2, or between shanks and flutes, nor were there any effects attributable to speed.

At 800°C (figure 11(a)) the annealed structure was obtained, consisting of carbide particles in a dark etching matrix of pearlite. At 900°C and 1100°C

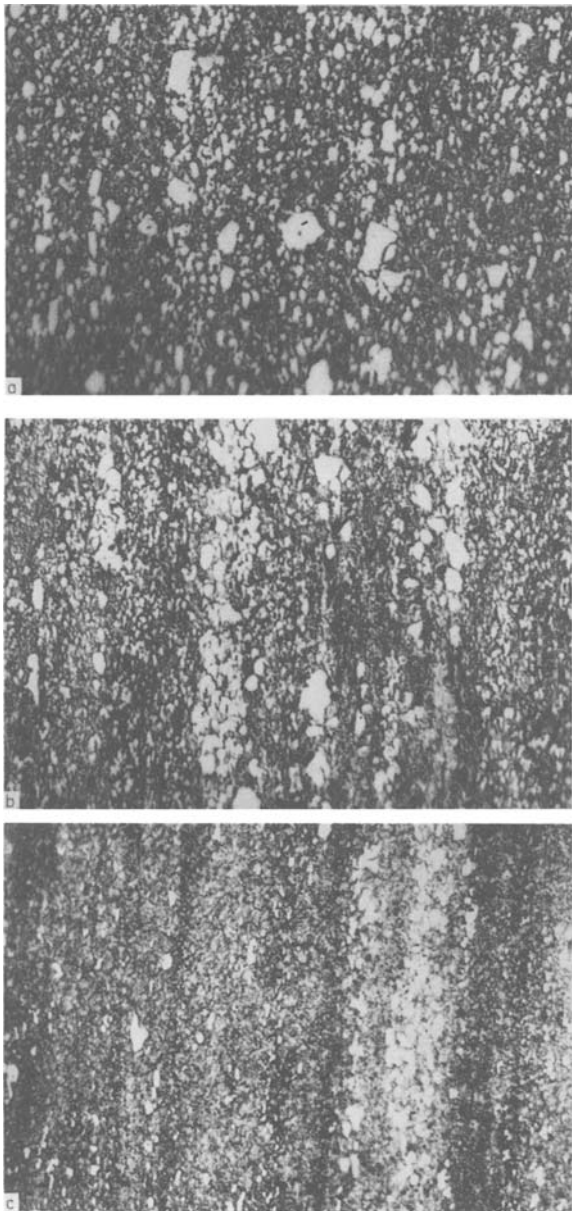


Figure 11. Typical microstructures of extruded high-speed steel blanks. (a) 800°C. (b) 900°C. (c) 1100°C.

(figures 11(b) and (c)) the pearlite was beginning to transform to austenite giving light areas against the dark residual pearlite.

PRODUCT HARDNESS

The variation of product hardness with temperature for the 25.4-mm diameter straight-fluted sections extruded at low velocity are shown in figure 12, these results are typical of all other test conditions.

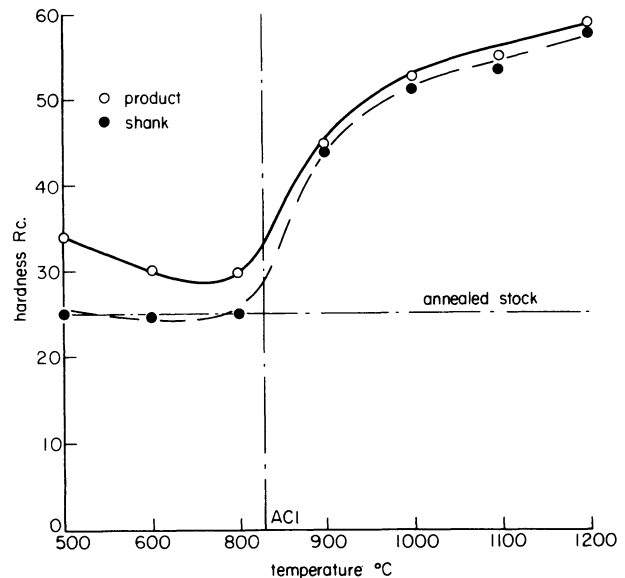


Figure 12. Hardness versus temperature for straight-fluted high-speed steel sections.

Up to the AC1 the hardness of the flutes is higher than that of the unextruded remainder due to a work hardening stress/strain curve (figure 8). Above the AC1 no real differences were observable between the shanks and flutes, the predominant influence being the billet preheat temperature.

It is apparent from figure 12 that if the AC1 temperature is exceeded during extrusion then the resultant blanks will require a full anneal unless subsequent machining is restricted to grinding.

PERFORMANCE OF TOOLING

Within the restricted nature of the trials it was obviously not possible to completely assess the production potential of the tooling, however, some comments are thought appropriate.

The 5 per cent chromium die materials used were subject to considerable erosion during the trials. In a production application of this process carbide dies, if practicable, would be required.

A simple pusher-type ejector was used to remove the spiral products from the dies. The ejection forces required immediately after forming were beyond the capabilities of the 50 kN installation used. The products had to be left to cool significantly in the dies before ejection was possible. Leaving the extrudates to cool in the dies probably considerably detracted from the performance of the tooling due to tempering back of the die materials, in a real situation quick ejection would obviously be necessary.

SUMMARY OF CONCLUSIONS

- (1) 18 : 4 : 1 (T1) and 6 : 5 : 2 (M2) high-speed steels can be extruded into spiral shapes, suitable for use as milling cutters, at temperatures between 800°C and 1100°C, at high and low velocities, with moderate extrusion pressures.
- (2) The pressures required for the extrusion of spiral sections were in overall agreement with an elementary energy balance proposed.
- (3) At a given temperature and reduction in area increasing the product spiral angle requires higher extrusion pressures.
- (4) At a given temperature and helix angle, the additional pressures required by the various spiral angles extrudates were roughly the same at the two reductions investigated.
- (5) Product cracking was encountered when extruding 18 : 4 : 1 (T1) and 6 : 5 : 2 (M2) at low velocities, and for 18 : 4 : 1 (T1) at high velocities, however, no cracking was obtained when extruding 6 : 5 : 2 at high velocities. The incidence of cracking was attributable to the ductility of the materials.
- (6) The metallurgical condition of the products was satisfactory, and no appreciable hardening was evident when extruding at temperatures below the AC1 (830°C).
- (7) From these investigations the optimum conditions for the extrusion of high-speed steel cutter blanks would appear to be the warm (800°C) extrusion of 6 : 5 : 2 (M2) high speed steel, at high velocities, to produce extrudates free from cracks and capable of further machining without annealing.

ACKNOWLEDGMENTS

The investigations reported are part of a large programme on warm forging supported by the SRC.

REFERENCES

1. Yu. A. Geller. Hot mechanical working of high-speed steels, *Instrumental'nye Stali Metallurgizdt*, (1955). p. 322. B.I.S.I. Translation no. 3165, 1963.
2. Now tool steel extrusions *Steel* (1956) **92**, 139, July 9.
3. P. J. Sukolski and G. Hoyle. The extrusion of high-speed steel sections, *J.I.S.I.*, (1959) **193**, November, p. 270.
4. H. Pugh. Hydrostatic extrusion, N.E.L. report no. 416.
5. High energy rate forming at General Dynamics Fort Worth, *Advanced High Energy Rate Forming A.S.T.M.E.* (1962) Book II, SP 62-15.
6. L. T. Chan, F. Bakhtar and S. A. Tobias. Design and development of Petro-Forge high-energy rate forming machines, *Proc. Inst. Mech. Engrs* (1965-66) **180**.
7. S. J. Becker and L. Mollick. The theory of the ideal design of a compound vessel, *Trans. A.S.M.E.* (1960).
8. T. A. Dean. A preliminary warm-forging essay, *Proc. 11th Int. M.T.D.R. Conf., Birmingham.* (1970).
9. R. K. Vinycombe and F. H. Martyr. Design of capacitance displacement transducers, *Instrument Practice* (1966) November.
10. W. Johnson and P. Mellor. *Plasticity For Mechanical Engineers*, Van Nostrand (1962).
11. W. Johnson and H. Kudo. *The Mechanics of Metal Extrusion* Manchester University Press (1962).
12. E. R. Lambert and S. Kobayashi. An approximate solution for the mechanics of axisymmetric extrusion, *Proc. 9th Int. M.T.D.R. Conf., Birmingham* (1968).
13. V. De Pierre. Experimental measurement of forces during extrusion and correlation with theory, *J. Lub. Tech.* (1970) July.
14. T. A. Dean and C. E. N. Sturgess. Stress-Strain characteristics of various steels over a wide range of strain rates and temperatures, *Proc. Inst. Mech. Engrs.* (1973) **187**, 523.
15. J. J. Jonas and T. Chandra. The extrusion force and the mean strain rate during the extrusion of strain-rate sensitive materials, *Metallurgical Transactions* (1970) August, **1**, 2079.
16. M. T. Watkins, J. McKenzie and E. Whitfield. A preliminary report on the effect of aperture shape in cold extrusion of steel, N.E.L. Report No. 383.
17. W. Johnson. Experiments in the cold extrusion of rods of non-circular section, *J. Mech. Phys. Solids*, (1958) **7**, 37.
18. J. T. Berry. *Recent Developments in the Processing of High-Speed Steels*, Climax Molybdenum Corporation, Inc. Greenwich, Connecticut.
19. G. A. Roberts, J. C. Hamaker and A. R. Johnson. Tool steels, *A.S.M.* (1962).
20. M. Oyane, A. Takeji, S. Kato, T. Yamada and N. Nishihara. Back pressure extrusion of brittle metals and metallic powder, *Bull. J.S.M.E.* (1969) **12**, no. 50, p. 377.

DIE LIFE IN HOT FORMING

CRACKING AND FRACTURE OF HOT-WORK DIE STEELS

by

A. THOMAS*

SUMMARY

High-strain low-endurance fatigue tests have been carried out on notched specimens of four hot-work die steels loaded in three-point bending at room temperature and 350°C at two different loading rates. The effect of several parameters on crack initiation, crack growth rate and fracture has been studied.

The number of loading cycles to cause crack initiation is not significantly influenced by material, position of the test piece in the die block, hardness or specimen orientation. Preheating accelerates crack initiation particularly at low loading rates.

At room temperature crack growth rates for all materials and test conditions were similar. Preheating increased growth rates at low loading rates but had little effect at the higher loading rate.

Material composition affected brittle fracture. Preheating reduced the occurrence of fracture but increased hardness and a higher loading rate increased susceptibility to fracture.

INTRODUCTION

A common mode of failure in forging dies is mechanical fatigue cracking which usually occurs at points of stress concentration such as internal fillet radii. During forging metal may be forced into cracks and trapped there by the elastic recovery of the die causing the forging to 'stick' in the die. This can reduce production rates to such an extent that the die must be removed for resinking. In addition fatigue cracks frequently lead to complete brittle fracture of dies.

Hammer dies are claimed to be more prone to cracking than press dies due to their impact loading and there is some evidence¹ that high energy rate forging (HERF) dies are particularly prone to early cracking.

When cracking is the principal mode of die failure, die lives tend to be low, usually of the order of a few thousand forgings rather than tens of thousands.

In the work described high-strain low-endurance fatigue tests were carried out on notched specimens cut from blocks of four hot-work die steels. The effect on crack initiation, crack growth rate and final

fracture of the following parameters was studied; loading rate, specimen orientation and position in the die block, preheating and hardness.

EXPERIMENTAL

Materials tested and preparation details

The chemical composition of the four materials investigated is shown in Table 1.

To produce the die blocks ingots were forged to 10 in square billet and sawn pieces 28 in long were then upset 50 per cent and worked to 14 in cubes. The ingot reductions used were 4 : 1 for No. 5 Electem and Jethete, 6·25 : 1 for Hydie and 9 : 1 for VWMC.

This method of forging is designed to reduce directional properties in the die blocks.

All material was received in the annealed condition and specimens were sawn from the blocks, rough machined and then heat treated as described in Table 2. Finally the test pieces were ground to size and notched to produce the dimensions shown in figure 1.

TABLE 1 Composition of hot-work die steels

	C	Si	Mn	S	P	Analysis					N
						Ni	Cr	Mo	W	V	
No. 5 Electem	0·54	0·23	0·73	0·014	0·010	1·47	0·75	0·28	—	—	—
Hydie	0·33	0·21	0·76	0·021	0·018	0·20	2·72	0·74	—	0·14	—
VWMC	0·36	1·00	0·39	0·012	0·014	—	5·07	1·44	1·47	0·52	—
Jethete	0·08	0·28	0·72	0·010	0·010	2·40	12·53	1·86	—	0·31	0·021

* Drop Forging Research Association, Sheffield

TABLE 2 Heat treatment details

Material	Heat treatment		
No. 5 Electem	850°C 2 h	Oil quenched	Tempered 2 h
Hydie	900°C 2 h	Oil quenched	Tempered 2 h
VWMC	1050°C 2 h	Air cooled	Tempered 2 h + 2 h
Jethete	1050°C 2 h	Air cooled	Tempered 2 h

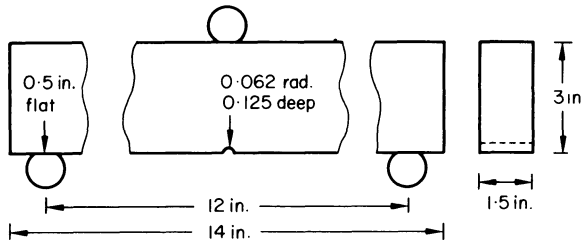


Figure 1. Fatigue test specimen dimensions.

Test procedure

Specimens were loaded in three-point bending as shown in figure 1 on a press and on a drop hammer. On the press the specimen deflection was controlled by adjusting the stroke of the press, whilst on the hammer stops were used to arrest the tup at the required deflection. During each test, displacement remained constant but the load reduced as crack growth proceeded. Loads were measured throughout each test. The loading rates (that is zero to peak load times) used were 0.045 s and 0.004 s for the press and hammer tests respectively. These are typical of loading rates found in hammer and high energy rate forging.

The surface crack length at each side of the test piece was measured with a travelling microscope and the average value was converted to a mean length across the full width of the specimen by means of a regression equation developed from measurements made on fractured specimens.

For tests where preheating was required specimens were initially heated for two hours at 350°C, transferred to the test rig and loaded for not more than 200 cycles. The specimens were then reheated for 15 min before further testing.

During loading a temperature drop of about 30°C occurred.

Identification of test conditions

The conditions under which specimens were tested are identified by the four-letter code shown in Table 3.

TABLE 3 Code for identifying test conditions

1st letter	L = longitudinal	T = transverse
2nd letter	I = inside	O = outside
3rd letter	S = slow (press) test	F = fast (hammer) test
4th letter	C = cold (room temperature) test	H = hot (350°C) test

TEST RESULTS AND DISCUSSION

When discussing the test results it is convenient to consider the crack initiation, crack growth and

fracture aspects of the fatigue process separately since the parameters investigated affected each in different ways.

Crack initiation

The crack initiation period is defined here as the number of loading cycles required to produce a crack 0.025 in long, which was the smallest crack which could be reliably detected. Some of the hardest materials tested at the highest load fractured before any cracking was measurable.

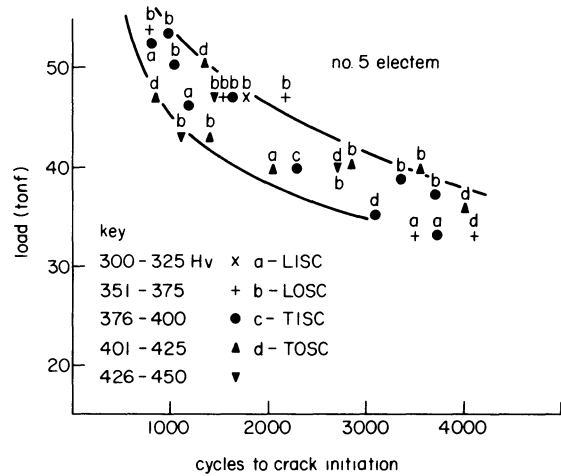


Figure 2. Initiation period as a function of load for No. 5 Electem press tests at room temperature.

Figure 2 shows initiation period as a function of initial applied load for No. 5 Electem specimens tested in the press at room temperature. It is clear that position in the die block, orientation and hardness have little effect on the initiation period. The same conclusion could be drawn from similar curves for other materials and test conditions.

To investigate the effect of material, loading rate and preheating, the test results for inside and outside and longitudinal and transverse specimens at all hardness levels were combined.

The relationship between the applied load L and the initiation period N_i was assumed to be of the form

$$N_i = aL^n$$

Least squares regression lines were calculated for plots of $\log N_i$ against $\log L$ and these were then used to calculate best values for N_i at three selected initial loads.

The correlation coefficient between $\log N_i$ and $\log L$ was always high, usually exceeding 0.9. The calculated values of initiation period are shown in Table 4.

The data in Table 4 were then used to assess the influence of die material, loading rate and preheating on crack initiation by applying t -tests to the difference of the mean values listed in Table 5.

Table 5 shows that crack initiation is not sensitive to material composition but that loading rate and preheating interact to influence crack formation. At room temperature loading rate does not affect the crack nucleation period but at 350°C a tenfold

TABLE 4 Calculated values of Ni

Material	No. 5 Electem				Hydie				VWMC				Jethete			
	LISC	LIFC	LISH	LIFH	LISC	LIFC	LISH	LIFH	LISC	LIFC	LISH	LIFH	LISC	LIFC	LISH	LIFH
Condition	LISC	LIFC	LISH	LIFH	LISC	LIFC	LISH	LIFH	LISC	LIFC	LISH	LIFH	LISC	LIFC	LISH	LIFH
	LOSC	LOFC	LOSH	LOFH	etc.
	TISC	TIFC	TISH	TIFH	etc.
	TOSC	TOFC	TOSH	TOFH	etc.
Load-tons																
50	1104	594	230	897	1191	2143	412	980	817	966	471	982	1285	1026	594	394
42	2065	1227	1227	1552	2070	3451	789	2275	1667	2455	929	1977	2223	1648	1194	1153
34	4365	2931	1734	2992	4018	6095	1718	6252	3908	7499	2099	4581	4285	2911	2761	4178

TABLE 5 Influence of various parameters on crack initiation

Parameter tested	Means compared	t-value	Degrees of freedom	Significance
Material	No. 5 Electem (1689) against grand mean (2160)	0.9184	58	> 0.3 < 0.4
Material	Hydie (2616) against grand mean (2160)	0.8207	58	> 0.4 < 0.5
Material	VWMC (2362) against grand mean (2160)	0.3606	58	> 0.7 < 0.8
Material	JETHETE (1971) against grand mean (2160)	0.3647	58	> 0.7 < 0.8
Preheating	'Cold' mean (2581) against 'Hot' mean (1738)	0.7967	46	> 0.05 < 0.1
Loading rate	'Slow' mean (1771) against 'Fast' mean (2548)	1.6486	46	> 0.1 < 0.2
Loading rate and preheat				
cold tests	'Slow' mean (2416) against 'Fast' mean (2745)	0.4548	22	> 0.6 < 0.7
hot tests	'Slow' mean (1126) against 'Fast' mean (2351)	2.1550	22	> 0.05 < 0.1

increase in the loading rate doubles the nucleation period. Other workers^{2,3} have noted that strain rate and load dwelling affect life in high-strain fatigue testing in a similar way.

The results presented suggest that the early crack formation in high energy rate forging dies compared with hammer and press dies is not due to the higher loading rate. Similarly the impact loading of hammer dies is not the cause of their early cracking compared with press dies. A more likely explanation of the observed behaviour lies in the excess energy transmitted to dies in different forging plant.

Since presses are stroke-restricted machines, the forging load builds up to the required value and overloading of the dies does not occur. In hammers and HERF machines, however, any excess energy in the tup after forging is completed must be absorbed by elastic deformation of the dies and machine, leading to increased loads on the die and consequently earlier cracking.

A further important practical observation is that the onset of cracking in forging dies cannot be delayed by material selection or preheating. The only effective method of retarding crack formation is by reducing the die load.

Crack growth rate

In low stress fatigue tests it has been well established that crack growth rate (dl/dN) can be related to the range of stress intensity parameter (ΔK) by an equation of the form

$$dl/dN = A \Delta K^n$$

Initially attempts were made to compare crack growth rates in the present tests on this basis. However for much of the data the net stress at the crack tip calculated from beam equations exceeded 0.8 times the material yield strength, thus invalidating

a linear elastic fracture mechanics approach. In addition it was found that in many tests the crack growth rate was influenced not only by the K value but also by the applied load, as shown in figure 3(a). This effect was most pronounced in brittle specimens. As toughness was increased by reducing hardness (figure 3(b)), preheating (figure 3(c)) or material change (figure 3(d)), the load effect was considerably diminished.

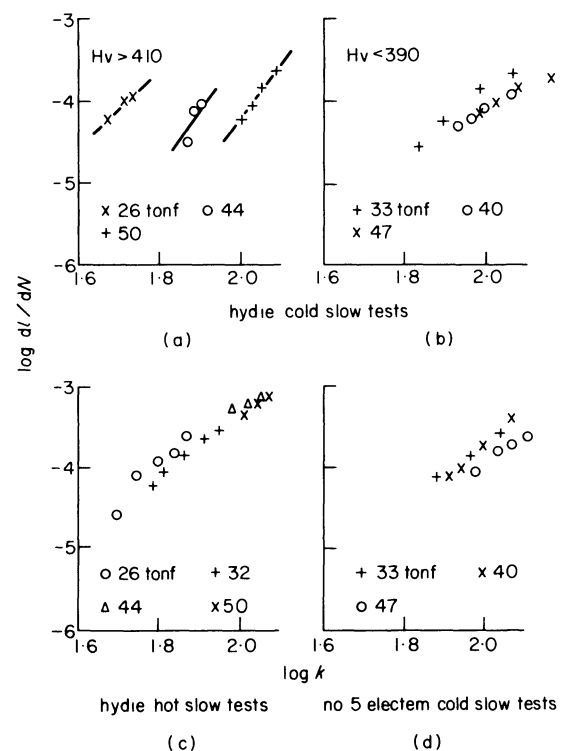


Figure 3. Effect of toughness on the influence of initial load on crack growth rate against stress intensity parameter relationship.

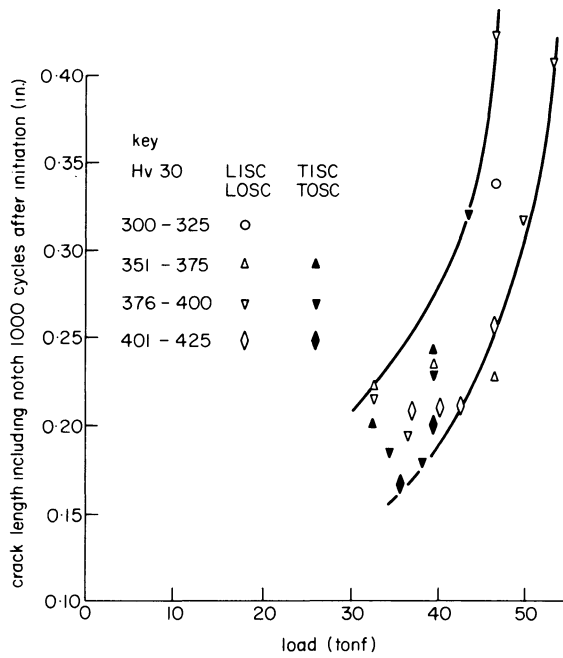


Figure 4. Crack growth rate as a function of initial load for No. 5 Electem press tests at room temperature.

Because of these difficulties the crack growth rate data were examined as follows. Curves were plotted to show the crack length one thousand cycles after initiation as a function of the initial applied load. Figure 4 shows such a plot for No. 5 Electem specimens tested at room temperature in the press.

Specimen position in the die block, orientation and hardness have little effect on the growth rate as shown by the common scatter band for all the test results. This was found to be the case for other materials and test conditions. Figure 5 shows data for other materials tested at room temperature in the press superimposed on the scatter band of figure 4. All the data lie close to that for No. 5 Electem,

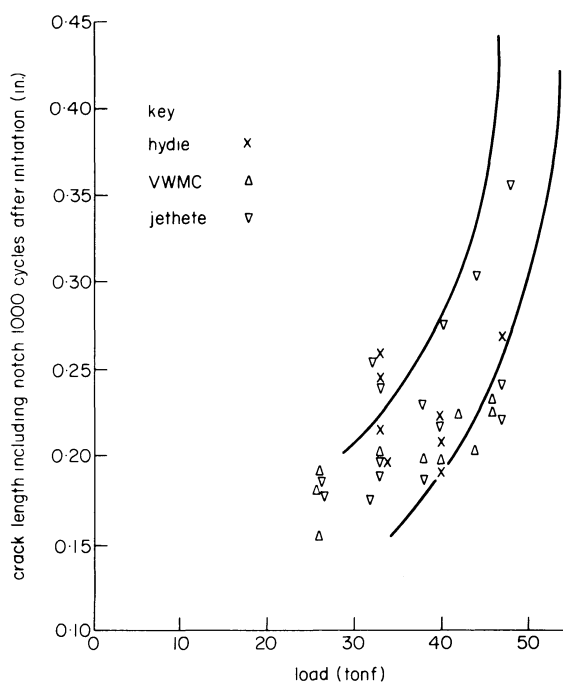


Figure 5. Effect of material on crack growth rate.

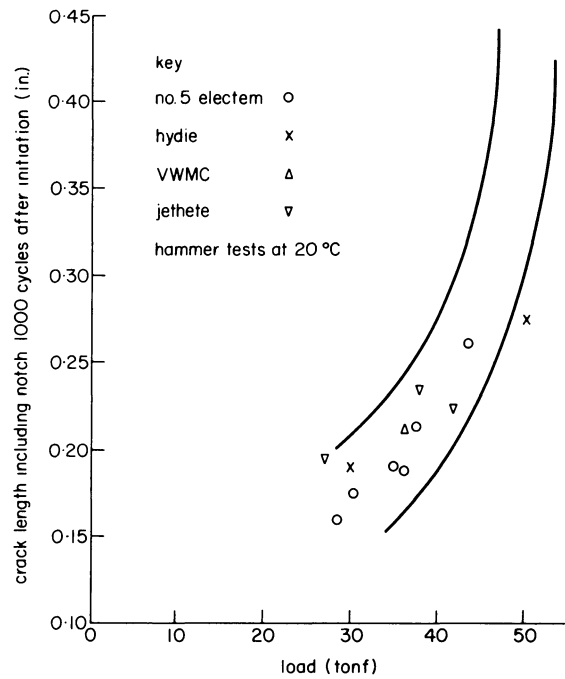


Figure 6. Effect of strain rate on crack growth rate at room temperature.

showing that material composition has little effect on crack growth rate.

In figure 6 data for room temperature tests at high strain rate are superimposed on the scatter band of figure 4 and show that at ambient temperature growth rate is not strain rate dependent.

The superimposed data in figure 7 are for preheated specimens at both high and low loading rates. At the low strain rate preheating leads to an increased crack growth rate, the magnitude of the effect increasing with increased initial load. At the higher loading rate, however, preheating does not lead to such a pronounced increase in crack growth rate except for No. 5 Electem specimens. The reason for

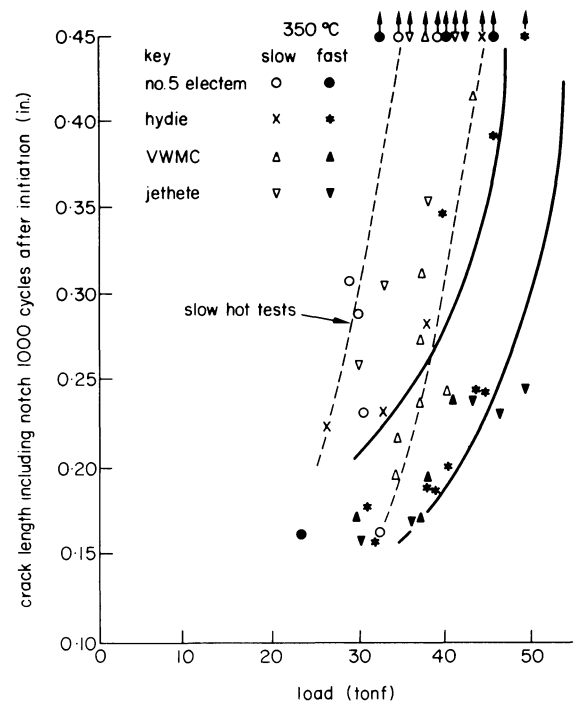


Figure 7. Effect of strain rate on crack growth rate at 350°C.

the anomalous behaviour of this material is not known.

In general the effects of various parameters on crack growth rate are the same as on crack nucleation. Thus again high strain rates are not responsible for rapid crack growth in hammer and HERF machine dies. Once more reduced loading appears to be the only effective method available to reduce crack growth.

Fracture

Not all of the specimens tested failed by fracture. In some tests, as very long cracks grew the load fell considerably and crack growth diminished. When crack lengths exceeded 0.5 in tests were discontinued.

The effects of the parameters studied on fracture may be assessed from figures 8(a)-8(d) and 9(a)-9(d). In these figures points have been plotted to show specimen hardness and initial load and the symbols used indicate whether fracture occurred or not. Also indicated on figure 8(a) are the hardness levels to which No. 5 Electem is usually heat-treated. Long experience has shown that for hammer dies the B level must not be exceeded except in very shallow dies or cracking is likely to occur.

It is interesting to note that the dividing line between fractured and unfractured specimens in slow tests (that is loading rates typical of hammer dies) is close to the empirically determined optimum level. This suggests that the test results for other materials may give a valid indication of safe hardness levels to avoid fracture.

At room temperature Hydrie has about the same upper hardness limit as No. 5 Electem but the

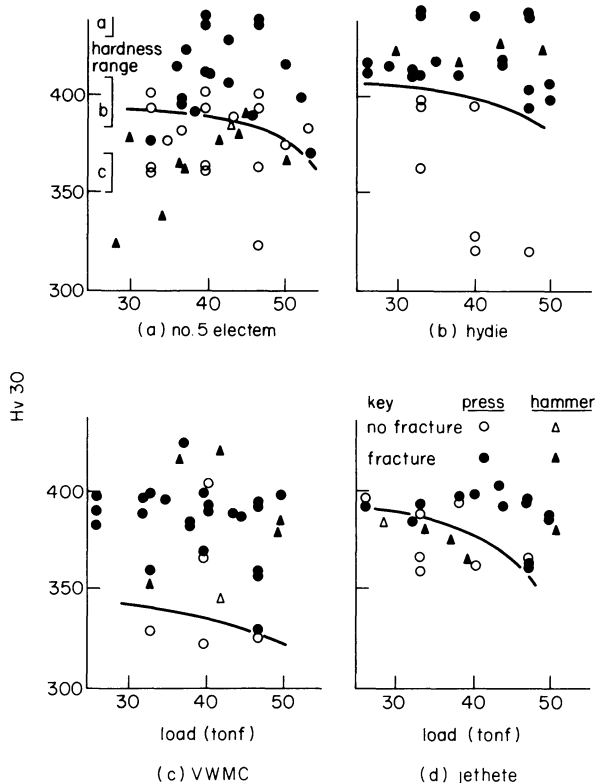


Figure 8. Effect of material, hardness and strain rate on fracture at room temperature.

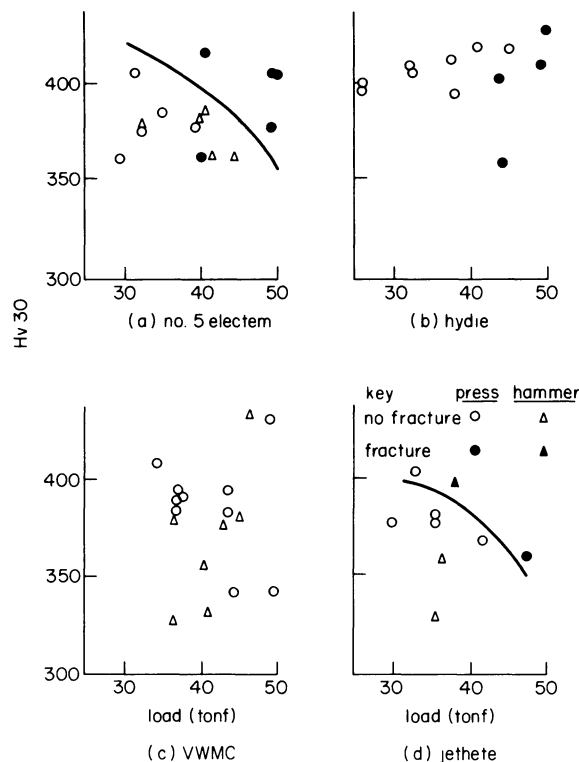


Figure 9. Effect of material, hardness and strain rate on fracture at 350°C.

hardness of Jethete and particularly VWMC must be reduced to avoid fracture.

Figure 8(a) shows that at ambient temperature an increased strain rate severely reduces fracture resistance of No. 5 Electem. It is not clear whether the other materials would behave in a similar manner since data on hammer tests at very low hardness levels are lacking.

The effect of preheating, as would be expected, is to increase the hardness levels at which fracture occurs and also to reduce or remove the detrimental effects of high loading rate. The improved fracture resistance of VWMC is very substantial (figure 9(c)) but for No. 5 Electem it is only marginal.

Bayliss⁴ has shown that the plane strain fracture toughness of No. 5 Electem increases from about 60 ksi/in² at room temperature to a value of 80 at 100°C. It remains steady at this value up to 300°C and then falls sharply again to its original value of 60. It seems from this evidence that the preheat temperature used for No. 5 Electem was too high to show significant differences in fracture resistance between hot and cold tests.

It may be inferred from the present results that the fracture toughness of VWMC is much higher at 350°C than at room temperature but unfortunately no data are available to support this hypothesis. A useful practical result of the tests is the fact that VWMC when preheated to 350°C has a fracture resistance as good as that of No. 5 Electem. This is important since it is known that VWMC is more wear resistant than No. 5 Electem but its use in hammers has been restricted due to its supposed brittleness. The present work indicates that with proper preheating this is not the case.

CONCLUSIONS

- (1) The formation and early growth of cracks in the four die steels investigated is insensitive to material or hardness within the range studied.

Preheating, especially at low loading rates, accelerates nucleation and growth rate.

- (2) Fracture resistance is dependent on material and strain rate at room temperature. Preheating increases fracture resistance but it appears that excessive preheating can reduce toughness back to room temperature levels.

Data on the variation of fracture toughness with temperature for hot-work die steels is urgently needed so that correct preheating temperatures can be selected.

- (3) Early cracking in hammer dies and HERF dies commonly attributed to impact loading is more likely to be due to overloading of dies caused by absorption of excess forging energy.

ACKNOWLEDGMENTS

Much of the work reported was financed by the Department of Trade and Industry under contract No. K78A/06/A72a. The author's thanks are due to Mr A. Cottam and Mr J. McKay who carried out the experimental work.

REFERENCES

1. A. Thomas and K. Denham. Die performance in high-energy rate forming. *Proc. 12th MTDR Conf.* Pergamon, 1971.
2. A. Coles *et al.* The high strain fatigue properties of low alloy creep-resisting steels. *Thermal and High Strain Fatigue*, Metals and Metallurgy Trust, 1967.
3. K. J. Miller. The effect of strain rate on low-endurance torsional fatigue in an alloy steel (En 25). *Ibid.*
4. R. N. Bayliss. Walter Somers Ltd. Private communication.

DIE LIFE ESTIMATION IN FORGING

by

S. STØREN*, J. EBBESEN†, J. SLUTÅS† and I. SÆTRET†

SUMMARY

Based on approximately calculated thermal and mechanical loads on forging dies, guidelines are given in this paper for die design and choice of die material in order to avoid critical phenomena, such as fatigue crack propagation, unstable crack growth or plastic deformation of the die surface. If the risk of the occurrence of these phenomena is sufficiently low, it is assumed that the die life is controlled by a stable abrasive wear process governed by a simple relation between contact time, local sliding velocity, local pressure and the hardness of the die material at working temperature. Limitations of the method and suggestions for further improvements are discussed.

NOTATION

a	depth of surface crack
b	width of flash
B	width of die
C_1, C_2, C_3, C_4	constants
d_i	internal diameter of spherical die
E	Young's modulus
f	tolerances of final product
G	weight of final product
h	die thickness
H_g, H_b	geometrical dimensions of die
H_v	Vickers hardness
H_B	Brinell hardness
k	coefficients of thermal conduction
K_I	stress intensity factor
K_{IC}	fracture toughness
K_f	yield stress of workpiece material
l	sliding length
N	number of cycles
o	circumference of flash
p	forging pressure
q	depth of penetration of thermal shock
r	radius of curvature
R	radius of workpiece
s	flash thickness
S	number of blows
t	time
T	temperature
V_{fl}	flash volume
w	ratio of weight of workpiece to weight of final product

α	coefficient of heat transfer
β	coefficient of thermal expansion
κ	thermal diffusivity
μ	coefficient of friction
ν	Poisson's ratio
ρ	density
$\sigma_{0,2}$	yield stress of die material
σ	stress, pressure

INTRODUCTION

At present, forging die costs can amount to between 10 and 20 per cent of the cost of the forged product¹. At the same time, it was found² that the life of dies used for the same product can vary five-fold, indicating an unsatisfactory control of the factors affecting die life. Although these facts resulted in some work on the life of forging dies, further work is still needed. Until now, most of the work was done on establishing the factors affecting die life and less on the influence of these factors on die life in order to be able to predict the life of forging dies. In this paper, the authors propose a procedure for determining the life of forging dies, based on a simple hypothesis in four parts.

PREVIOUS WORK

Forging handbooks^{3,4} list the following factors as influencing the life of dies used in closed-die forging: die material and hardness, flow-stress and temperature of workpiece, scale, workpiece design, tolerances

* Division of Metallurgy and Metals Working, The Norwegian Institute of Technology, University of Trondheim.

† The Foundation of Scientific and Industrial Research (SINTEF), Trondheim.

of the workpiece, and rapidity and intensity of the blow.

The same handbooks mention the following factors as the main causes for die failures: overloading, overheating, abrasive wear, heat checking and cold dies.

Tholander¹, in his detailed discussion of the different wear mechanisms, distinguishes between the following: plastic deformation of the die surface or the flash land mechanical wear, local melting due to explosive burning of the lubricants, formation of surface cracks due to thermal fatigue and formation of the surface cracks due to mechanical fatigue.

Kannapen⁵ gives a further extensive review of published work dealing with these wear mechanisms.

Thomas² and Aston and Muir⁶, contributed to some important research work carried out at the University of Aston in Birmingham and at the Drop Forging Research Association in Sheffield. Using statistical methods on observations of the wear process in industrial forging, they found many interesting relationships between important parameters by means of regression analysis. They clearly show the statistical nature of the problem. Following up this work, an extensive laboratory test programme, simulating the wear process was carried out for a variety of hammer die materials (see Aston⁷).

In Germany, efforts were made to calculate the thermal and mechanical loads on forging dies, as reported by Lange⁸, Bech⁹, Klafs¹⁰ and Pohl¹¹.

At the Batelle Columbus Laboratories in Ohio, USA, important research work was done dealing with load calculations and analysis of geometry by means of computerized procedures—see Altan¹², Akgerman and Altan¹³. In the design of forging dies, these authors point out the necessity of predicting: shape complexity and forging volume, number and geometry of preforms, design of the flash for preforming and finishing, and forging load, energy and centre of load.

Against this background, the authors felt that an effort was required to synthesize the results obtained at the different research centres. In their opinion, the final goal is a quantitative and practical procedure enabling a best choice of die material and die design in order to optimize the life of drop forging dies. The procedure presented here and used in a Norwegian forging handbook¹⁴, is not yet in its final form, but must be considered as a preliminary proposal, open for discussion and improvements.

HYPOTHESIS

The following hypothesis is put forward for determining the life of forging dies:

- (i) If temperature variations on the die surface cause thermal stresses of the same magnitude as the yield stress of the die material, cracking and accelerated wear will soon occur.
- (ii) If the combination of normal pressure and shear stresses (in terms of effective stress) on the die surface reaches the same magnitude as the yield stress of the surface material at the

appropriate temperature, local deformation and accelerated wear will occur.

- (iii) If the static and dynamic stresses in the die during forging cause stress intensity factors in microcracks on the die surface of a magnitude large enough to propagate these cracks, a casual overloading of the die can rapidly lead to a complete damage of the die (brittle fracture).
- (iv) If none of the foregoing forms of loading are critical, continuous abrasive wear will determine the die life. The main factors governing the wear velocity are the surface hardness of the die material, forging pressure and sliding length of forged material. Lubricant, microstructure and oxide scales are secondary factors, but should not be excluded.

The parameters governing these four forms of loading depend on the following process parameters: (a) workpiece—material; geometrical form, magnitude, weight and temperature when put into

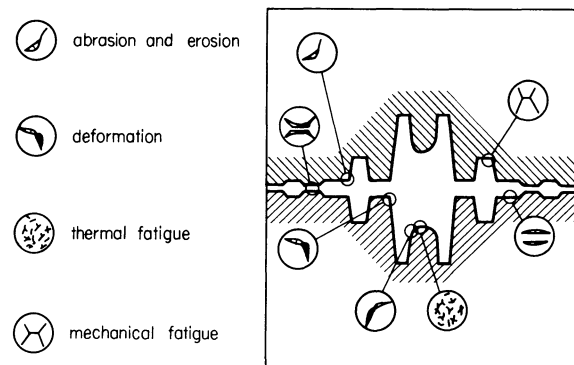


Figure 1.

the die, and geometrical form and weight of the product, tolerances and number. (b) Forging machine—number of blows, and impact velocity and contact time. (c) Forging die—design of the die and flash land; preheating of the die; die dimensions, and die material.

Figure 1 shows where the different wear mechanisms are likely to occur.

PRACTICAL PROCEDURE

The nature of the problem discussed here is of such a complexity that approximate methods must be used to obtain a result in a reasonable time. The authors' choice of approach is therefore one of several that are possible, laying stress on simplicity.

Temperature

Calculations of thermal loads may be divided into three parts: (a) estimation of the length of the transient period before quasi-stationary conditions occur; (b) calculation of heat generated within the workpiece during deformation and conducted to the die, and (c) calculation of temperatures due to the thermal shock occurring when the hot workpiece is put into the much colder die.

The expressions below are conclusions drawn from calculations on one-dimensional models with constant material properties. Despite essential restrictions, we believe that they have a value in enabling us to evaluate the effect of changing the different parameters.

Figure 2 shows a simple model of a die, the material having a thermal diffusivity κ , exposed to a hot workpiece with period t_0 . Assuming a constant initial temperature and no heat loss from the surface, the number of cycles N until quasi-stationary

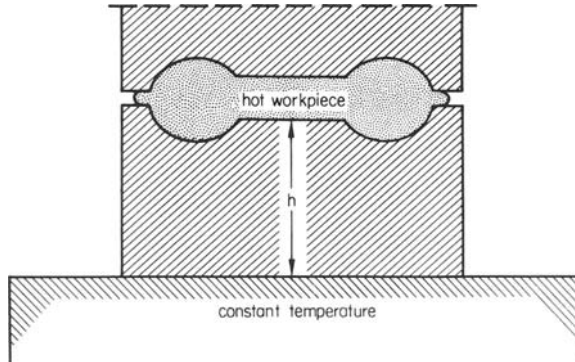


Figure 2.

conditions prevail in the centre of the die can be estimated from

$$N = h^2 / \kappa \times t_0 \quad (1)$$

This value is an absolute minimum. For real dies, N should be multiplied by a factor $C_1 > 1$, the magnitude depending on the ratio of the forging area to the total area of the die. The transient period is generally unwanted and it is therefore of interest to make N as small as possible. In addition to variation of the parameters in (1), N may be reduced considerably by suitable preheating of the die.

In both the transient and the quasi-stationary period, the temperature in the die may be divided into two parts, a basic temperature and a periodic temperature. The former increases monotonously in the transient period and is constant in the quasi-stationary. The periodic temperature varies within each load cycle and it is assumed that it can be superimposed on the basic temperature. It consists of the two temperatures mentioned in (b) and (c). The temperature caused by dissipation of heat in the workpiece during deformation is small in hot forging and is therefore neglected.

In the shock problem we are primarily interested in the maximum temperature, this being the most critical. If a workpiece with a temperature T_w is put into the die (shown in figure 2) with a surface temperature T_s , the maximum temperature in the centre of the die is given by¹⁵

$$T_{\max_1} \approx T_s + \frac{T_w - T_s}{1 + \epsilon} \quad (2)$$

where

$$\epsilon = \frac{\kappa_d}{\kappa_w} \times \left(\frac{\kappa_w}{\kappa_d} \right)^{1/2}$$

κ_d , κ_w , κ_d and κ_w are the coefficients of thermal conduction and the thermal diffusivity of the workpiece and the die, respectively. Equation (2) holds if the time of contact t_c between workpiece and the die exceeds the rise-time of T_{\max} .

$$t_c > \frac{36}{\kappa_d} \times \left(\frac{\kappa_d / \alpha}{1 + \epsilon} \right)^2 \quad (3)$$

α is the coefficient of heat transfer between the workpiece and the die. The rise-time of T_{\max} is defined as the time when T_{\max} has reached 90 per cent of its final value.

If Equation (3) is not satisfied, the maximum temperature is governed by α and may be estimated from

$$T_{\max_2} \approx T_s + \frac{T_w - T_s}{1 + \epsilon} \times \left[\frac{\alpha}{6\kappa_d} (1 + \epsilon) \times (\kappa_d \times t)^{1/2} \right]^{1/3} \quad (4)$$

where t is the time counted from the start of close contact.

The depth q to which the effect of the shock penetrates is estimated from¹⁶

$$q = C_2 \times (\kappa_d \times t)^{1/2} \quad (5)$$

where C_2 is a constant $\approx 2.0 - 3.0$.

Equation (5) is only valid for short times ($q \ll R$, where R is the radius of the workpiece) and best accuracy is achieved in the die centre. When the temperatures in the die are known, accelerated thermal fatigue is assumed to occur when

$$T_{\max} - T_s > 1.6 \times \frac{\sigma_{0.2} \times (1 - \nu)}{E \times \beta} \quad (6)$$

where $\sigma_{0.2}$ = yield stress of the die material $\sim 1/3 H_B$

ν = Poisson's ratio

E = Young's modulus

β = coefficient of thermal expansion.

All constants are evaluated at the appropriate temperature.

Plastic deformation

The most likely place for plastic deformation to occur is at the inlet to the flash land, see figure 1. A simple expression for the pressure σ_z on the die in this region is

$$\sigma_z = -K_f \times \left(1 + 2\mu \times \frac{b}{s} \right) \quad (7)$$

where K_f is the yield stress of the workpiece material, μ the frictional coefficient and b/s the ratio of width to thickness of the flash land.

To avoid plastic deformation or accelerated abrasive wear, the following criterion is used:

$$\sigma_z < 0.75 \times H_B \quad (8)$$

where H_B is the Brinell hardness of the die material at maximum temperature.

Unstable fracture

The surface of the die will always contain microcracks due to imperfect manufacturing or overloading. During forging, the forging pressures

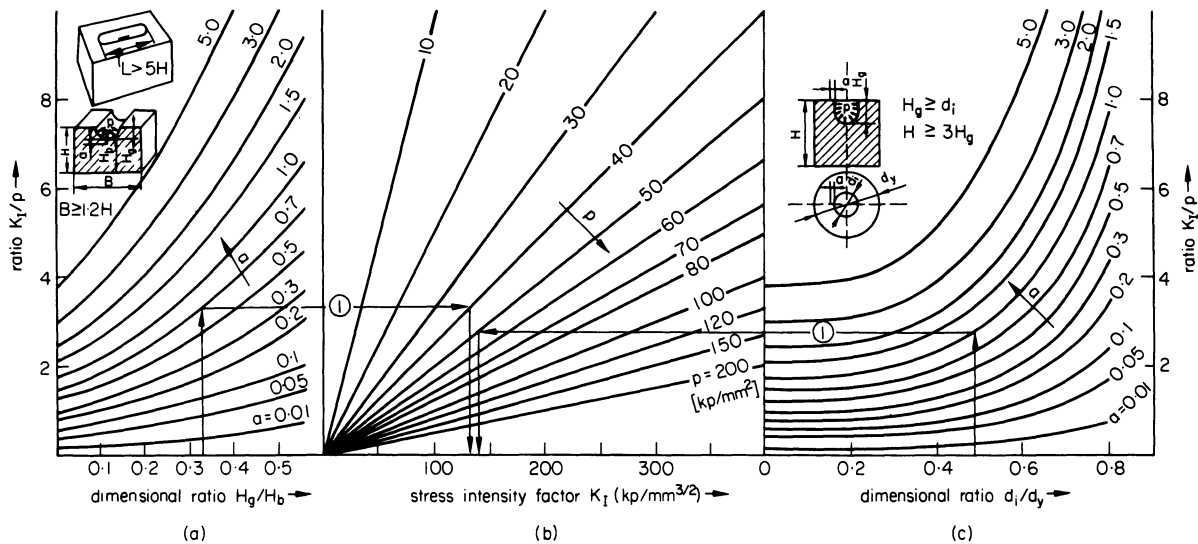


Figure 3.

cause a stress intensity factor K_I at the crack tip. If K_I exceeds the fracture toughness K_{IC} of the die material, unstable fracture will occur. If the surface of the die has the shape of a cylinder or a sphere, the stress intensity factor is approximated by

$$K_I = 1.7 \times \sigma \times (a)^{1/2} \quad (9)$$

where σ is the stress in the direction parallel to the surface and normal to the crack and a the depth of the crack. The calculation of σ is a problem of elasticity and is not explained here.

When the surface of the die is not a cylinder or a sphere, but close to it, corrected stress intensity factors are given by:

$$K_{ICyl} = K_I \times \left(\frac{H_g}{r+a} \right)^{1/2} \quad (10)$$

and

$$K_{Isp h} = K_I \times \left(\frac{d_i}{2(r+a)} \right)^{1/2} \quad (11)$$

for a cylinder and a sphere, respectively. H_g and d_i are shown in figure 3 and r is the radius of curvature of the surface where the crack lies.

For two simple cases, a cylinder and a sphere, with even internal pressure p , the stress intensity factor has been calculated for different values of p and a , and dimensions of the die. The results are shown in figure 3.

Through experiments (CT-standard test) the fracture toughness of the die material may be found as a function of the temperature. Likewise, the stable crack growth ($K_I < K_{IC}$) may be found as a function

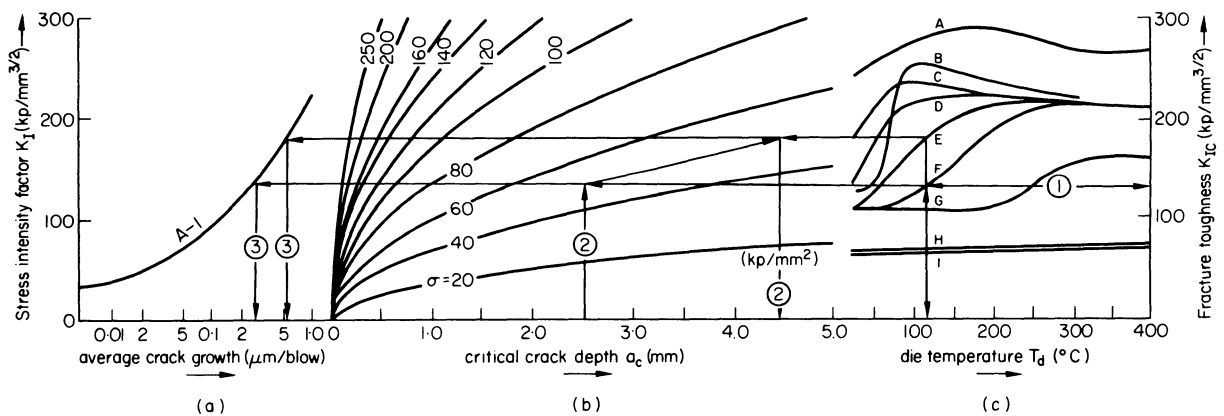


Figure 4.

TABLE 1 Analysis and heat treatment of materials A-I

Material	C	Mn	Si	Cr	Mo	V	W	Fe	Hardening	Annealing	Hardness HR _C
A	0.4			3.0	0.5			Rest	1050°C/air	2 h/540°C	46
B	0.33		1.0	5.0	1.1	0.5		"	"	2 h/590°C	50
C	0.4	0.4	1.0	5.3	1.4	1.0		"	"	1 h/635°C + 1 h/625°C	42
D	"	"	"	"	"	"		"	"	1 h/610°C + 2 h/600°C	48
E	"	"	"	"	"	"		"	"	2 h/800°C	50
F	"	"	"	"	"	"		"	"	1 h/570°C + 1 h/560°C	53.7
G	"	"	"	"	"	"		"	"	no annealing	55
H	0.87	0.3	0.3	4.2	5.0	1.9	6.4	"	1100°C/air	1 h/640°C + 1 h/620°C	50
I	"	"	"	"	"	"	"	"	1200°C/air	1 h/610°C + 1 h/600°C	60

of K_I . This has been done for some selected materials and the results are shown in figure 4. The analysis and heat treatment of the materials are shown in Table 1.

From figures 3 and 4, the required die temperature T_d may be found for a given material, or the optimum material may be found if the die temperature is given (example (1)).

Figure 4 may also be used for examining dies where unstable fracture has occurred. The depth of the initial crack and the critical crack depth a_c may be found from measurements on the fracture surface. The die temperature and the material being known, the fracture toughness and the stress σ in the die surface may be found (see figures 4c and 4b). The effect of choosing another material or another die temperature on the critical crack depth when the stress σ is kept constant may then be estimated (example (2)).

Figures 4a and 4b may be used for estimating the number of blows causing a given increase in crack depth. Example (3) in figure 4 shows the effect of an increased critical crack depth on the crack growth.

The maximum crack depth due to accelerated thermal fatigue may be estimated from

$$a_t = C \times (\kappa \times t_c)^{1/2} \times \left(\frac{T_{\max} - T_s}{300} - 1 \right) \quad (12)$$

where $C = 1700$ (for tool-steel). When a is known, the procedure is as already explained.

Abrasive wear

If none of the foregoing loading conditions has led to die failure, it is assumed that a stable abrasive wear mechanism will determine the life of the die. The number of blows a die can stand before the tolerances of the product are exceeded is estimated from¹⁷:

$$S = C_3 \times \frac{H_v \times f}{\sigma_z \times l} \quad (13)$$

C_3 is a constant determined by experiments in industrial use, H_v the Vickers hardness of the die material at the appropriate temperature, f the tolerance and l the sliding length.

By sliding length is meant the length of material that is pressed into the flash gap. It is calculated from:

$$l = C_4 \times \frac{V_{fl}}{o \times s} = \frac{(w - 1) \times G}{o \times s \times \rho} \quad (14)$$

C_4 is a constant, V_{fl} the volume of the flash, o the circumference of the forged part, s the thickness of the flash, ρ the density of the workpiece material, w the ratio of weight of workpiece and final product and G the weight of the final product. The values of s and w must be ascertained by reference to forging handbooks.

The width of the flash often varies along the circumference and some places are almost without grade. The value of l must then be corrected and instead of the circumference is inserted the length with flash formation.

DISCUSSION

As already stated in the introduction, this paper must be regarded as a proposal for the determination of the life of forging dies and for the selection of an optimum die material.

Fundamental material properties are used extensively. This implies that these properties are known for different materials at varying temperatures. However, such data are still difficult to find in the literature. For a few selected tool materials, thermal and fracture toughness data were determined in Trondheim. In Germany much work was done to establish the flow stress as a function of strain rate and temperature, but further research work has still to be carried out.

All expressions that were used above are approximate and must therefore be used critically. The temperatures have been calculated assuming one-dimensional conditions. For flat parts in the centre of the die such conditions prevail very closely and the calculated error is small. For webs and grooves, or in the region near the flash, the expressions are not correct and should be used with care. The same arguments apply to the calculation of plastic deformations.

The fracture toughness of the die material has been determined using the standard CT-test. This yields a value of K_{IC} that is valid as long as plain strain prevails in the test specimen as well as in the real crack.

Having established the hypothesis, the next step is to try it out under industrial forging conditions. This has been done to a limited extent in Norway with good results. To carry out a thorough test of the hypothesis, we believe that an investigation is needed of the kind reported by Aston and Muir⁶ using statistical methods. Such an investigation is however difficult to perform in Norway with her relatively small forging industry and should be carried out by one of the large industrialized countries. Instead of the variables used by Aston and Muir, we believe that the more fundamental variables presented in this paper should be used. These are temperature; thermal and mechanical properties of the workpiece and the die material; fracture toughness of the die material; initial crack depths on the die surface and complexity and weight of the forged product. Such investigations would require close control of variables such as temperature; weight of workpiece; die surface quality and fracture toughness of the die material. An investigation of the fracture surface should be carried out after each unstable fracture. For the control of temperatures close to the die surface, a special measuring system was developed in Trondheim and tried with very good results in the forging industry. The main characteristic of this system is that it enables a close determination of the depth of the thermocouple below the die surface. As a consequence, the temperature profile may be calculated fairly accurately. The effect of dynamic loads was not included in the expressions referred to in this paper. However, this effect should be included, both for drop hammers and the slower presses, in order to improve the method.

CONCLUSIONS

- (1) The life of forging dies depends on a set of fundamental parameters. An estimation of die life may be divided into four parts: Accelerated thermal fatigue; plastic deformation; brittle fracture, and abrasive wear.
- (2) A limited test of the hypothesis has given good results.
- (3) A thorough test of the hypothesis should be carried out using statistical methods.

ACKNOWLEDGMENTS

The present work was supported by the Royal Norwegian Council for Scientific and Industrial Research (NTNF) through three different projects: B-0645.2919, B-0645-2861 and B-0610.3186 and was carried out in close contact with the Norwegian forging industry.

REFERENCES

1. E. Tholander. Förslitning av smidessänken (Wear of forging dies). *MEKANRESULTAT*, Sveriges mekanförbund, Stockholm, Sweden, June 1967.
2. A. Thomas. Variability of life in drop forging dies. *Metal Forming*, 41-5, February 1971.
3. *Metals Handbook*. Vol. 5, Forging and Casting, 8th ed. *American Soc. for Metals*, 1970.
4. H. W. Haller. *Handbuch des Schmiedens* Carl Hanser Verlag, Munich, (1971).
5. A. Kannappan. Wear in forging dies. *Metal Forming* (1969 and 1970).
6. J. L. Aston and A. R. Muir. Factors affecting the life of drop forging dies. *J.I.S.I.* (Feb 1969) 167-76.
7. J. L. Aston et al. The wear testing of hot work die steels *Metallurgia and Metal Forming* (February 1972).
8. K. Lange, *Gesenkschmieden von Stahl*. Springer Verlag, 1958.
9. G. Bech. Wärmemässige Beanspruchung von Werkzeugstählen beim Warmstauchen und Schmieden in Gesenken. *Stahl u. Eisen* (1958) 78, 1556-63.
10. U. Klafs. Ein Beitrag zur Bestimmung der Temperaturverteilung in Werkzeug und Werkstück beim Warmumformen. *Dr.-Ing. Dissertation, Hannover* (1969).
11. W. Pohl. Ein Verfahren zur näherungsweise Berechnung der Wärmeentwicklung und der Temperaturverteilung beim Kaltstauchen von Metallen. *Berichte aus dem Institute für Umformtechnik, Universität Stuttgart (TH)*. (23) Verlag W. Girardet, Essen, 1972.
12. T. Altan and H. J. Henning. Closed die forging of round shapes, *Metallurgia and Metal Forming* (March 1972).
13. N. Akgerman and T. Altan. Modular analysis of geometry and stresses in closed-die forging: Application to a Structural Part. *J. Engrng Ind.* (1972) 1025-34.
14. *Senkesmiing av stål. Driftshåndbok*. (Forging handbook). (1973) NTNF project. B-0645.2919. NTH/SINTEF, Trondheim, Norway.
15. A. S. Carslaw and J. C. Jaeger. *Conduction of heat in solids*. Oxford University Press, (1959) p. 89.
16. Y.-C. Fung. *Foundations of Solid Mechanics*, pp. 408-409. Prentice-Hall, 1965.
17. S. Kalpakjian. *Mechanical Processing of Materials*, p 48. Van Nostrand, 1967.

WEAR IN DROP FORGING DIES

by

T. M. SILVA* and T. A. DEAN*

SUMMARY

The work presented in this paper describes a series of die-wear tests on cavity drop-forging dies, using an automatic forging unit. The damage occurring on both top and bottom dies is examined and compared.

INTRODUCTION

Recent years have seen extensive developments in component manufacturing processes. Casting methods have been improved so that products of good accuracy and moderate price can be made. The mechanical properties of powder metal parts have been improved and components can be used under conditions of high stress, both static and dynamic.

Progress in the techniques of steel-metal fabrication has resulted in significant reductions in costs of manufacturing by this method. In the field of metal forming highly accurate components of excellent surface finish can be made available by cold forging.

The industries producing by the above methods are all to a large extent competing for markets which were previously exclusively provided for by the drop forging industry. A major example is the automobile industry¹ to which about half the weight of drop forgings in the United Kingdom is destined. It has been apparent over the last two decades that the amount of forgings used by this industry has not increased in proportion to the output of vehicles, due to the adoption of components from other sources.

In the face of this fact the drop-forging industry is under pressure to produce either cheaper components or components of higher quality.

Dies can be seen to affect two facets of drop-forging production. Firstly the price, for about 10 per cent of the unit cost of a forged component has been attributed to the cost of the dies². Secondly the progressive degradation of drop forging tools, in particular the surface wear, reduces the quality of components by causing dimensional inaccuracies. These two features of drop-forging tools are not strictly separable as the deterioration and eventual demise has a direct bearing on production costs. The importance of tools in the drop forging process has resulted in much recent attention being paid to their

performance. The need to understand the mechanisms of die damage, before effective measures can be taken to prolong lives, has long been accepted. In order to examine the reaction of tool materials to what is virtually the simplest condition of forging, a number of researchers have carried out experimental investigations on the performance of flat dies in simple, hot, upsetting operations^{3,4}. Whilst some parallels can be drawn between the behaviour of flat dies and that of cavity dies, basic differences, such as the presence of radii and walls in the latter, cause various patterns of stock movement across the surfaces, which result in large variations of temperature within the tools and which also set up complex states of stress.

The examination of used industrial forging dies and correlation of results by multiple regression analyses has been a means⁵ of relating die damage to several process parameters such as draft angle, area of forging, and forging weight. The results have indicated that in an industrial environment, the main factor affecting die-lives is the weight of the forging, whilst the influence of draft angle and fillet radii is significant. However, little experimental evidence is available on the detailed mechanisms of the progress of the various forms of die damage. With this in mind a wide programme of investigations is being pursued, embracing all aspects of damage in drop-forging dies. This paper describes part of the results obtained so far, which focus attention on a comparison of wear occurring on the top and bottom dies of a tool set.

EXPERIMENTAL DETAILS

Forging equipment

The forging tests were carried out on a Mk IID Petro-Forge⁶, high-velocity forging machine. Nominally this has a maximum output of 10,000 ft-lbf at a platen closing speed of 35 ft/sec. Lower energy outputs are achieved with lower impact speeds.

* Department of Mechanical Engineering, University of Birmingham

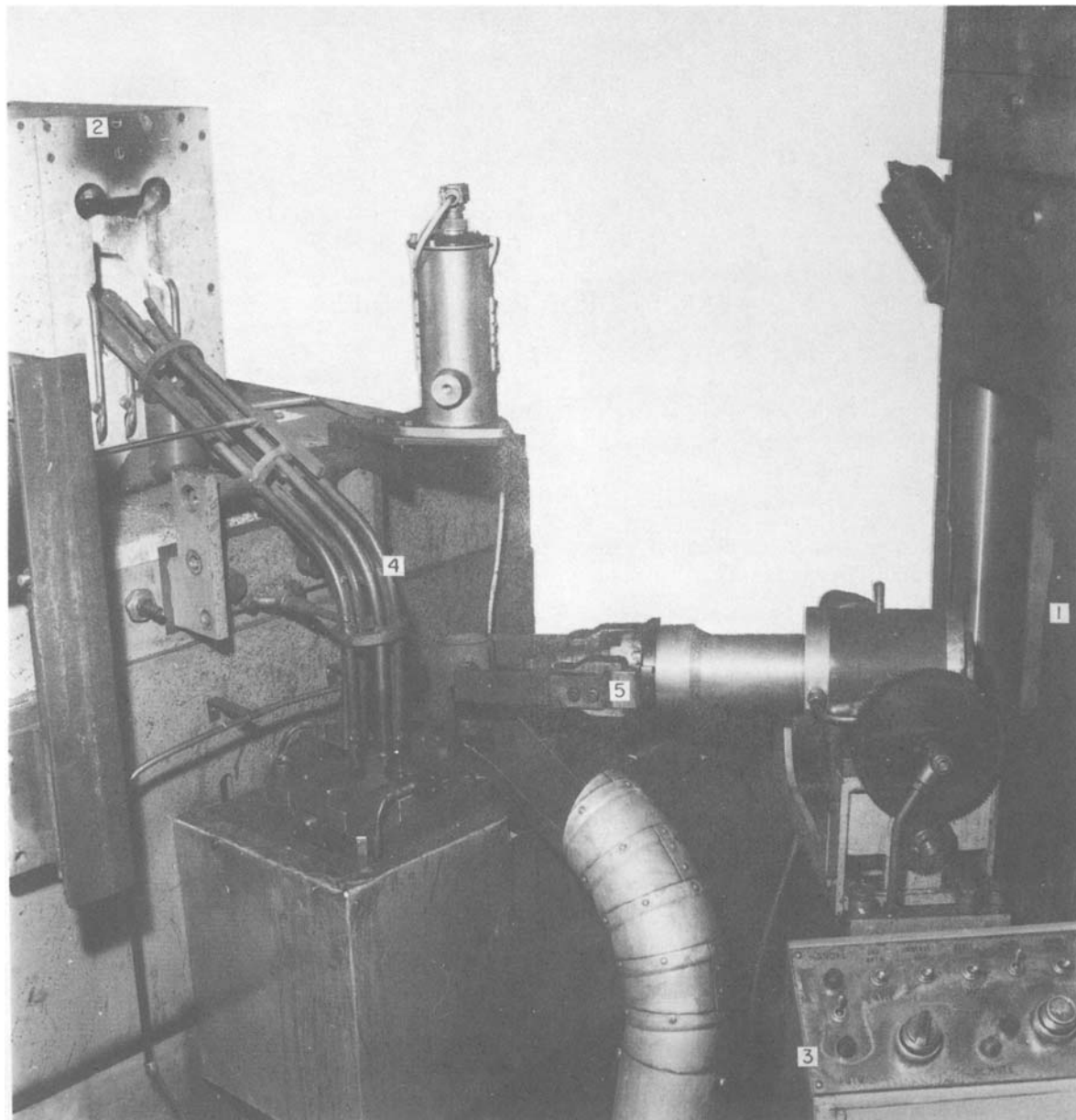


Figure 1. The mechanical handling unit.

The forging system, shown in figure 1, was fully automated. The complete unit consisted of the Mk IID Petro-Forge (1), the induction furnace (2), and the billet handling equipment (3). The hot billets were fed through the induction coil by vibration and slid down the chute (4). They were then picked up by the jaws (5), and were placed on the centre of the bottom die with axis vertical. The machine then delivered a blow, and after deformation the forged component was ejected clear of the die cavity by the hydraulic ejector. The component was then removed by the pick-up jaws.

The required preheat temperature dictated the rate of feed of billets through the induction coil and hence the cycle time. The cycle time fell within 8 ± 0.2 s more than 90 per cent of the time. The mean preheat temperature of the billets was 1040°C , with a variation of $\pm 20^{\circ}\text{C}$. The system had several safety devices incorporated in it as follows:

- (1) A pyrometer checked the temperature of billets on the bottom die and if it was outside

- the preset limits, the machine was not fired, although the rest of the cycle was completed.
- (2) No blow was delivered if the ejector was in the 'up' position.
- (3) No blow was delivered when either of the jaws was in the 'down' position on the dies.

Table 1 gives the times taken for various operations by the system. No die lubrication was used in these tests.

TABLE 1

Operation	Time (seconds)
Transfer from furnace to die	8
On the die prior to forging	2.5
On the die after forging	2.5
On the ejector alone before being picked up	2

Tooling

The die set used for the experiments is shown in figure 2(a). The cavities were machined in inserts

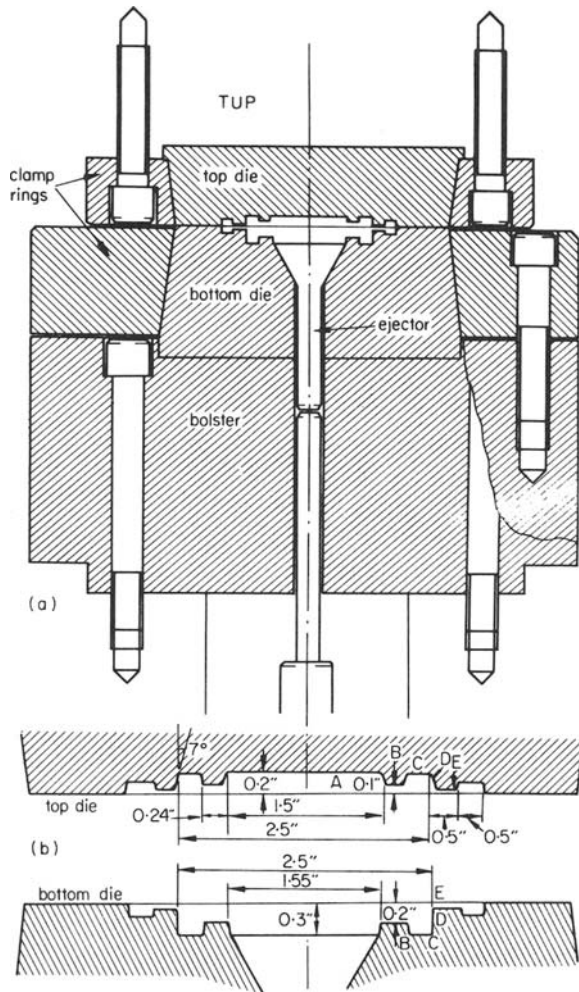


Figure 2. Tool assembly and die geometry.

which were held in bolsters bolted to both the machine bed and the tup. The bolsters were made from EN 24 steel heat-treated to 32–36 Rockwell C and each one comprised two components: a platen with a locating recess and a clamp ring which secured the inserts on their tapered sides. The inserts were made from a 5 per cent chromium hot-working die steel equivalent to AISI H13 supplied by Uddeholm. This is a popular material for many forging processes nowadays because of its good wear resistance and is favoured for many press forgings. The dies were turned to shape, hardened and double-tempered to 46–48 Rockwell C, and flash ground.

Details of cavity geometry are shown in figure 2(b). The shape was chosen after considerations of ease of manufacture which suggested a turned form and features which would provide information on all the aspects of damage. A similar shape of cavity was machined in both top and bottom die, differences between the two which are evident were provided to facilitate the mechanical handling of the process, that is to ensure that the forged shape was always on the lower die for the disposal jaws. These were a slightly deeper cavity in the lower die which was not provided with draft angles, in contrast to the top die which had 7° of draft.

The bottom die was provided with an ejector

which formed the flat central portion. This was seated on a conical face turned into the main body of the insert.

Stock material

The stock was 1½-in diameter EN8 bright bar sawn into billets 1⅓ in long. This material was chosen because of its wide use in the forging industry, and its good forgeability enabling the achievement of reductions of over 80 per cent. The dimensions selected for the billets produced forgings with an average flash thickness of 0.065 in. However, the flash did not extend to the full distance of the land and so did not enter the gutter.

Test procedure and wear measurement

The wear tests consisted essentially of using each die set for forging several thousand billets. These forgings were made up of forging 'runs', each run consisting of approximately 500 forgings. Tests on a particular die were discontinued before the end of the normal wear period if it cracked in a gross manner.

At the end of each forging run the dies were thoroughly cleansed using a wire brush to remove any adhered scale. The method of measurement of dimensions of the dies was by taking replicas of the cavities. Trial and error proved that the most satisfactory method was to cold-forging a bar of aluminium of 1½ in diameter and 3½ in long and to place it horizontally across the bottom die. Two blows on this at a reduced energy produced the best replica representing the parent dies very accurately. The walls of the flash gutter were used as a datum for measurements as no changes in dimensions occurred on these surfaces. The height measurements were taken by placing the replicas horizontally on two parallel bars, and then using a sensitive clock gauge, measuring up to 0.001 mm, to read the heights around ridges. The replicas were sectioned slightly off centre and were subsequently ground up to the centre line. They were then projected on to the screen of an optical contour projector with a magnification of 20 and the profiles were recorded.

RESULTS AND DISCUSSION

Changes in dimensions

Preliminary tests established a wear pattern for both top and bottom dies, which indicated that the ridge, B, and the wall, D, suffered most wear (see figure 2(b)). The region, A, of the top die showed signs of wear, the rate of wear being lower than that of ridge B. The ejectors assumed a concave shape after several thousand forgings. It was decided to monitor the dimensions of these regions, namely the height of ridge B and the diameter of wall D of both top and bottom dies, the height of region A of the top die, and the height at the centre of the ejector.

The height of ridge B was measured at its centre and the diameter of wall D was measured at the level of the base of the flash gutter (figure 2(b)). Figures 3(a) and (b) indicate a steady reduction of the height of ridge B of both top and bottom dies as the number of forgings increased, leading to a linear relationship. The height co-ordinate of ridge B measured around a

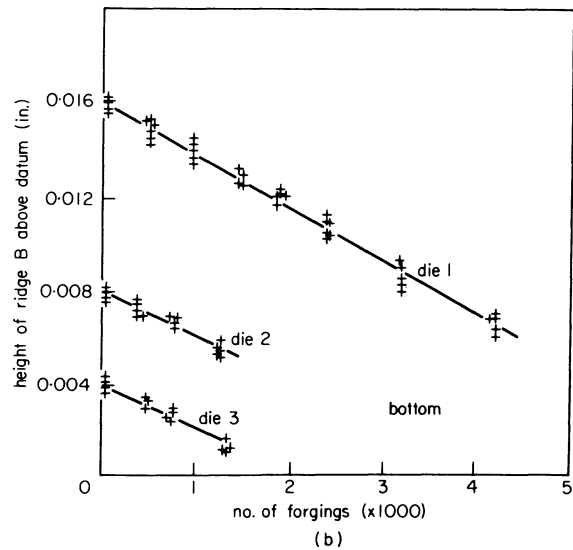
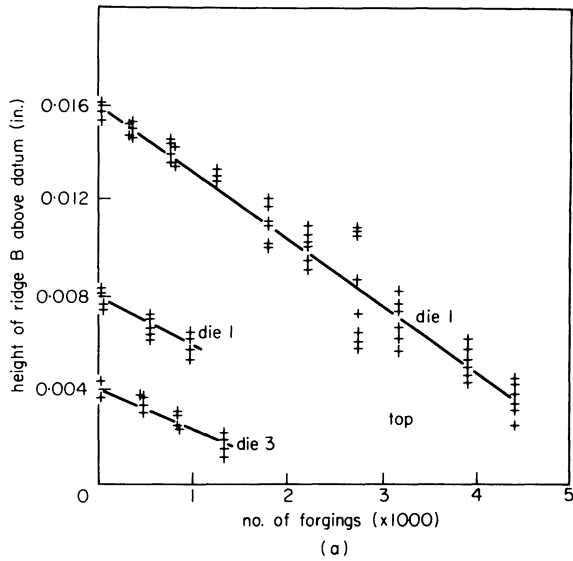


Figure 3. Graphs of height of ridge B versus number of forgings.

die for a particular number of forgings showed some degree of scatter. This can be attributed to two main causes. The first is the radial erosion marks on the ridge, B, giving rise to a rough surface texture. The second is the small amount of unevenness in wear that takes place around the dies, which is largely due to billets not being placed absolutely centrally. The best-fitting straight lines have been determined and the confidence limits have been found. The top die is found to wear at a slightly higher rate than the bottom die. The rates of wear are 0.0028 in (with 95 per cent confidence limits, ± 0.00046 in) per 1000 forgings for the top die and 0.0020 in (with 95 per cent confidence limits, ± 0.00067 in) per 1000 forgings for the bottom die.

Now the forging conditions experienced by the top and bottom dies are different, largely because of the differences in times of contact with the hot, stock material. The billet rests on the bottom die for about 2.5 s before and after forging, whilst the ridge, B, of the top die contacts the billet for about 2 ms every cycle. The consequences as regards the dies are twofold. Firstly the bulk of the bottom die is heated

more uniformly and to a higher temperature than the top die. Thermal gradients and the tendency for the propagation of fatigue cracks are reduced. Secondly the contact between the bottom die and the billet prior to deformation causes the lower surface of the billet to be chilled. Subsequent surface flow of the billet is restricted in the bottom die compared with that in the top.

The wear of die surfaces is a complex phenomenon but may largely be attributed to erosion due to the rubbing of the deforming stock surface, together with the loss of small particles broken away because of fatigue cracking. As explained above both the extent of interfacial sliding and the amplitude of the transient surface stresses are greater in the top die, which is shown to suffer a greater rate of wear.

Figures 4(a) and (b) indicate a steady increase in diameters at wall D for both top and bottom dies as forging progressed. In contrast with the height co-ordinate did not show much scatter, mainly due to the smoother eroded surface at D than at B. The rates of wear are 0.00175 (± 0.00006 in) for the top die and 0.0018 (± 0.00007 in) for the bottom

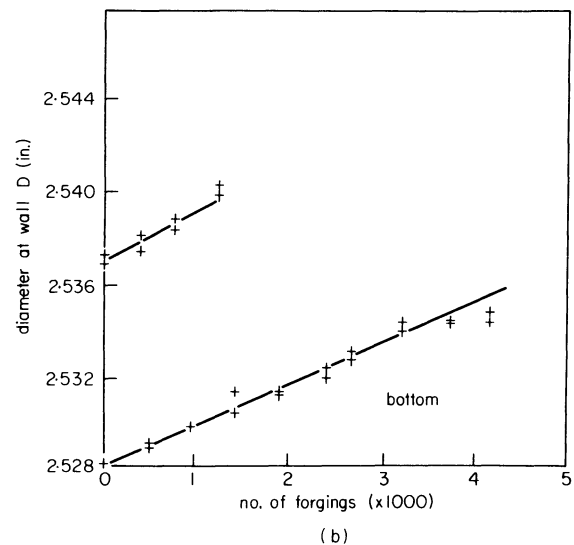
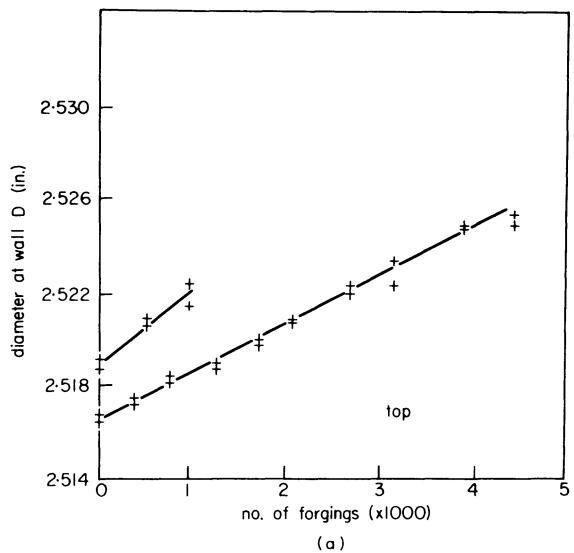


Figure 4. Graphs of diameter at wall D versus number of forgings.

die per 1000 forgings. The fact that the rate of wear in the top and the bottom dies was similar indicated that the conditions did not differ greatly. It is expected that little material flow would have occurred along the die wall in either top or bottom die. So in this respect the wear in the two dies will be similar and less than that on the ridge, B, which is subjected to extensive rubbing. However, similar temperature differences, as mentioned above, are to be expected for the respective dies in this region.

Far less erosion took place in the central region of the top die than on the ridge, B, and less also than on the wall, D. Surface profile measurements along a diameter produced traces showing a higher central region and two troughs and, at the edges, two higher

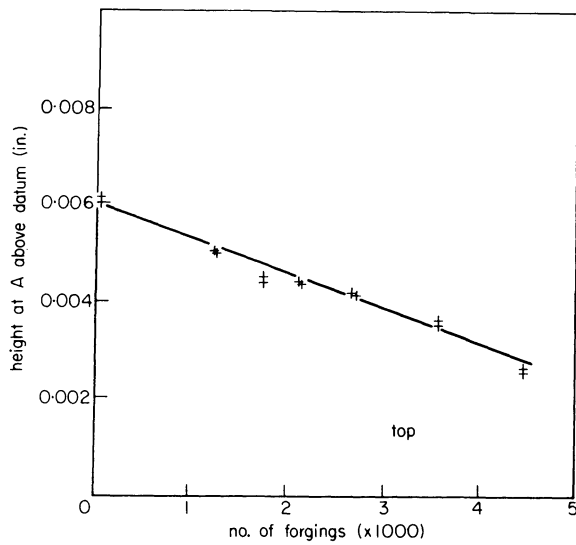


Figure 5. Graph of height at A versus number of forgings.

regions again. See figure 6(a) for the surface profile trace. The wear had obviously taken the form obtained in simple upsetting operations; a central region of comparatively little wear surrounded by an annular groove of maximum erosion. However, whereas in simple upsetting tests the wear groove is situated close to the periphery of the undeformed billet, in the cavity dies it was found approximately midway between the axis and the inner wall of ridge B. This fact indicates that billet surface movement at its edge was largely restrained by the wall, but that some sliding occurred nearer the axis. Figure 5 shows measurements taken of the greatest depth of wear in region A. This had a value of 0.0033 in after 4500 forgings.

The main cause of damage to the ejector was distortion due to its being forced further into the conical seat as forging progressed. The top face became concave, which prevented the accurate measurement of wear profiles. When ejectors had been in use for the manufacture of 3000 or more forgings it was noticed that the profile, characteristic of the central region, A, of the top die became superimposed on the concave distorted shape. The progressive change in the surface contour on a diameter of the ejector is shown in figure 6(b).

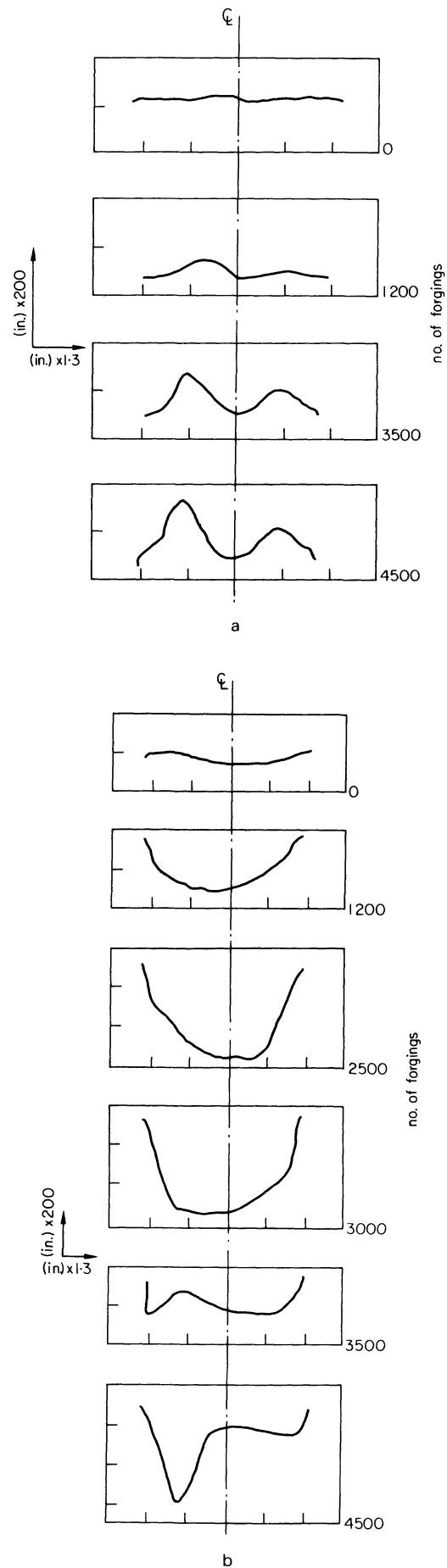


Figure 6. (a) Surface profile of the central flat of the top die. (b) Surface profile of the ejector.

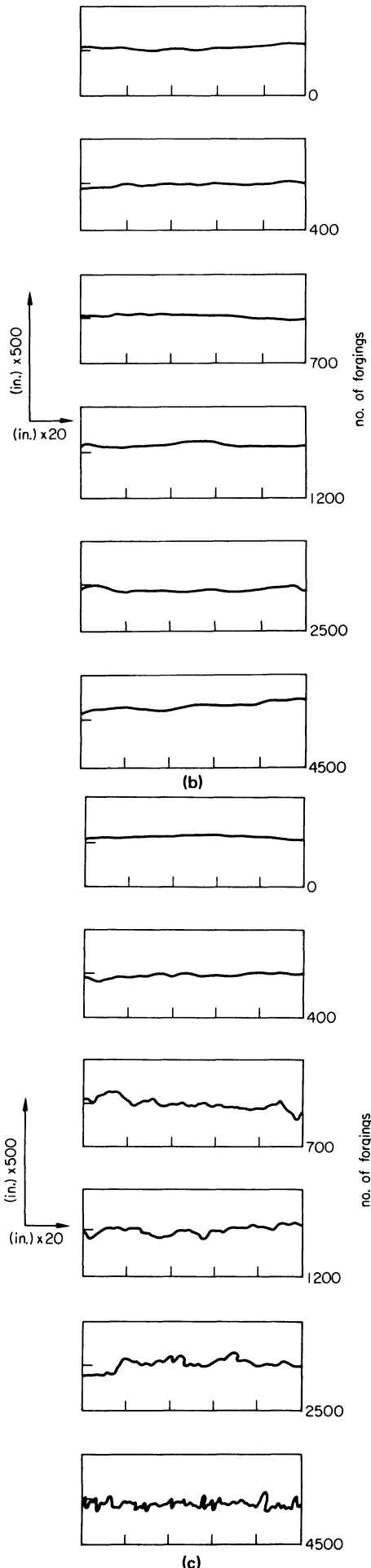
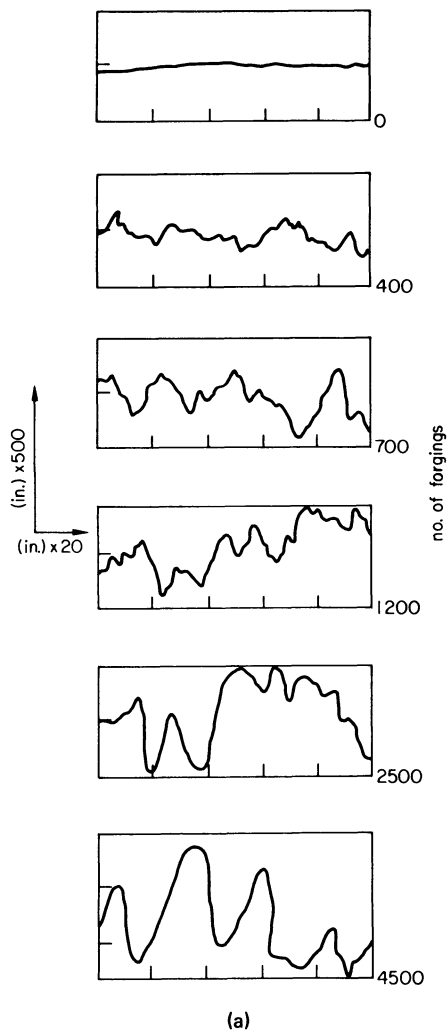
The recess, C, of both top and bottom dies showed very little sign of wear. In fact the original machining marks remained after several thousand forgings had been made.

The flash lands, E, suffered erosion marks similar to those of the top die, region A, but were much less severe. The deepest measured in this region was only 0.002 in, subsequent to the manufacture of 4500 forgings.

Wear tests on flat dies³ under similar process conditions and using the same materials have shown a maximum depth rate of wear of 0.00125 in per 1000 forgings, a figure of about half the value obtained on the ridge, B, in the cavity tests reported here. It is interesting to note that slightly greater wear on the bottom than on the top die was obtained in the flat tool set, whereas in the regions of greatest wear (ridges B) in the cavity dies the reverse has been found.

Surface texture measurements

Figure 7 shows the measurements taken of the surface texture of the dies at stages during a test run. As the results for top and bottom dies were similar, except for the central areas, only those for the top die are presented. In all cases surface-texture traverses were made in a circumferential direction, that is to say normal to the erosion marks where these were



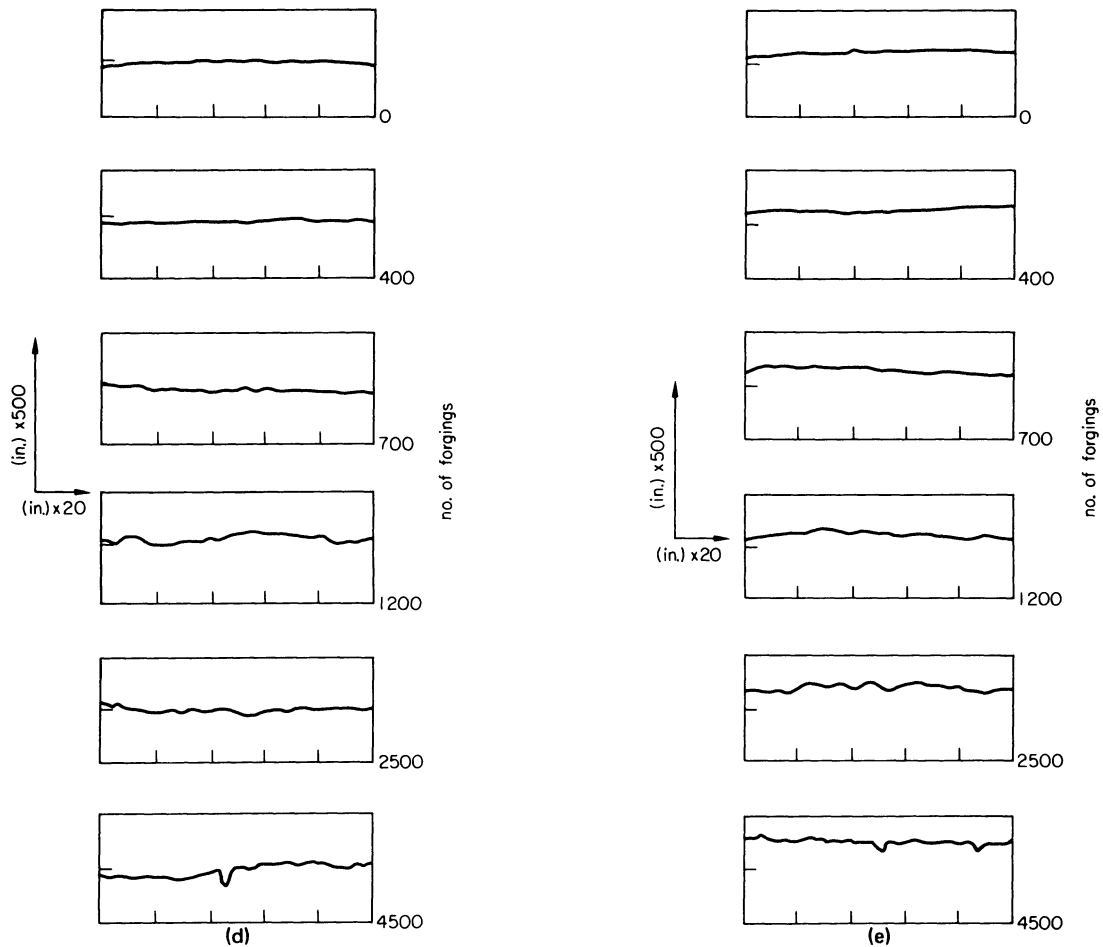


Figure 7. (a) Surface texture of the ridge, B. (b) Surface texture of the recess. (c) Surface texture of the flash land. (d) Surface texture of the centre of the top die. (e) Surface texture of the ejector.

present. See figure 8 for the erosion marks on various regions of the die. The largest measured changes in surface roughness are associated with the regions of

greatest wear, that is, the ridge, B. Similarly the least change in texture occurs in the recess, C, which suffered least wear. The surface roughness at the top of ridge B increased to a maximum from the start of a forging run until about 2000 forgings had been made and remained at that value for subsequent production. The roughness values at the flash lands reached a maximum at about the same time as for the ridge, but it was a lower figure. Changes in surface texture of the region A of the top die occurred in the area of maximum wear. The surface texture of the face of the ejector coarsened during forging, but this was due more to surface oxidation than abrasion as the blemishes showed no directionality.

CONCLUSIONS

The relative patterns of wear in the top and bottom dies were similar. The degree of wear occasioned in any one region of the cavity seems largely determined by the amount of interfacial flow experienced. If billets are not positioned centrally in the cavity more wear is experienced in the region where most flow occurs. Surface finish becomes degraded in the regions of most wear, the greater wear generates a rougher surface, although CLA readings reach a maximum for a particular region of the die. Similar amounts of wear occurred in top and bottom dies except in the regions of greatest rubbing and therefore greatest wear. In these places more erosion occurred on the top die than on the bottom.

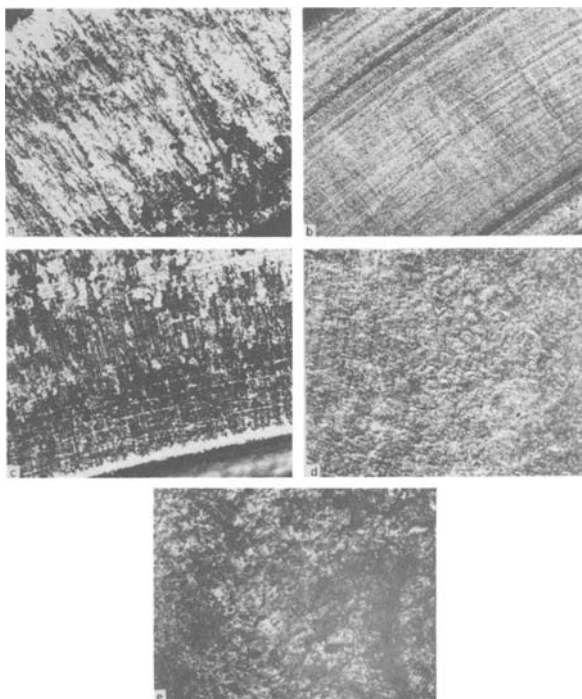


Figure 8. Appearance of various regions of the die after 4500 forgings. (a) Ridge B. (b) Recess C. (c) Flash land E. (d) Central region A. (e) Ejector.

ACKNOWLEDGMENT

This paper contains the results of part of a large research programme into aggregate die damage, sponsored by S.R.C.

REFERENCES

1. S. E. Rogers and A. Thomas. Drop forging—past achievements and future prospects, *Metallurgia and Metal Forming* (1973) **40**, 236.
2. F. Bakhtar. The economics of high-energy rate forming. *The Engineer* (1969) 7 March.
3. S. M. J. Ali. Die wear in high-speed hot forging. Ph.D. Thesis University of Birmingham.
4. A. Thomas. Investigation on die materials. Ph.D. Thesis Sheffield Polytechnic (1970).
5. J. L. Aston and A. R. Muir. Factors affecting the life of drop forging dies. *J.I.S.I.* (1969) **207**, 167.
6. L. T. Chan, F. Bakhtar and S. A. Tobias. Design and development of Petro-Forge high-energy rate forging machines. *Proc. Inst. Mech. Engrs* (1965–6) vol. 180, pt. 1, no. 29, p. 689.

THE EFFECT OF DIE TEMPERATURE ON METAL FLOW AND DIE WEAR DURING HIGH-SPEED HOT FORGING

by

B. W. ROOKS*

SUMMARY

This paper describes the results of tests carried out on a Mk IIA Petro-Forge high-speed hammer in order to determine the effects of the bulk die temperature on the die-wear rate during simple upsetting of hot steel (EN8) billets. Variations in the bulk die temperature are achieved by changes in cycle time and dwell time. Each test consists of upsetting 1000 billets to a reduction of 75 per cent, both with a colloidal graphite lubricant and dry, after which the dies are measured for wear with the aid of a Talysurf.

The frictional conditions at the die-billet interface are studied with the aid of the ring test, in which annular billets are upset between lubricated dies to various levels of deformation from 0 to 65 per cent. The amount of lubrication applied is the same as in the wear tests. Four die-temperature levels are used in order to simulate the interface temperature conditions appertaining in the wear tests.

The results of the wear test indicate that within the ranges tested cycle time has only a marginal influence on die temperature and wear, whereas dwell time has a significant effect on both temperature and wear. It is deduced from these results that the real independent parameter influencing wear is die temperature, dwell time and cycle time affecting wear only through changes in die temperature. The main conclusions are that under lubricated conditions wear is inversely proportional to bulk die temperature (and dwell time), whereas under dry conditions the reverse is true.

The evidence of the flow pattern of etched billet sections and of the appearance of the dies after testing indicate that an increase of die temperature causes a decrease in the amount of billet sliding over the die surface. This is further supported by the results of the ring test which show that as the die temperature is increased so is the interfacial coefficient of friction.

Using the values of the coefficient of friction obtained from the ring test, predictions are made of the interfacial metal flow during simple upsetting for various die temperatures. From these predictions theoretical wear profiles at each die temperature are constructed in arbitrary units. These profiles show a good qualitative agreement with experimentally obtained profiles. Finally, the wear volume (in arbitrary units) is obtained from the theoretical profiles at each die temperature. It is shown that the theoretical rate of change of wear volume with die temperature is very close to the experimentally derived ratio. Therefore it is concluded that the reduction in wear with increase of die temperature (and of dwell time) is mainly attributable to the change in the interfacial flow pattern.

INTRODUCTION

A number of phenomena contribute to the failure of dies in hot forging, notably mechanical fatigue, thermal fatigue, plastic deformation, and erosive wear. The topic of this paper is restricted to die damage caused by erosive wear, and in particular the influences of die temperature and metal flow. Erosive wear has been found to be affected by factors such as billet oxide scale¹, lubrication^{1,2}, and die-billet contact time³. From these studies has emerged the fact that a primary cause of wear is the sliding of the billet material and the embedded scale over the die

surface, although a thorough understanding of the mechanism on a microscopic scale is still sought.

The application of a lubricant reduces friction and thereby promotes interfacial metal flow, such that under certain conditions, e.g. during simple upsetting, lubrication leads to an increase in the rate of die wear.

Die-billet contact time or dwell time has been found to affect wear in conflicting ways. It can cause die softening due to the high temperatures generated and hence increase wear³. Also dwell time has been found to affect the metallurgical conditions of the die surface⁴. Under certain conditions phase trans-

* Department of Mechanical Engineering, University of Birmingham

formation to martensite producing hard layers in the die surface has occurred, which could be expected to reduce wear. However, this could be offset by thermal fatigue cracking which also is promoted by phase transformations. A further aspect of the dwell time is its influence on the ability of the lubricant to promote metal flow. As a consequence there is evidence of wear being reduced by increasing dwell time⁴.

The object of the work presented in this paper is to study the effects of cycle time and dwell on die temperature and metal flow and to relate these aspects to die wear. Metal flow and the frictional conditions at the die-billet interface are examined with the aid of the ring test⁵. From these tests the amount of interfacial metal sliding can be established during upsetting and the knowledge utilized to relate die wear to metal flow. The wear tests are carried out over a range of cycle times from 5.5 to 7.5 s, and dwell times from 5 to 1000 ms on a Mk IIA Petro-Forge machine specially programmed for the purpose. The test technique has been described previously²⁻⁴.

EXPERIMENTAL METHODS

Wear tests

The technique originally developed by the Drop Forging Research Association⁶, in which cylindrical billets are simply upset between flat dies, was used to evaluate the rate of die wear under various forging conditions. Steel billets (EN 8) were heated to 1100°C in an induction furnace and transferred to a Petro-Forge high-speed hammer by means of an automatic handling device⁷.

The billets, measuring 1 in in diameter by 1.125 in long were upset to a reduction of 75 per cent, which required a nominal energy input of 4500 ft-lbf, given by a platen impact speed of 32 ft/s. Each test consisted of upsetting 1000 billets, after which the dies were removed, degreased and descaled prior to being measured for the amount of wear produced. During the tests readings were taken of the bulk temperatures of top and bottom dies as sensed by thermocouples positioned $\frac{1}{4}$ in below the surface of each die. Wear on each die was detected with a Talysurf, the procedure having been described in a previous publication³.

Tests were carried out dry and lubricated by spraying with a colloidal graphite suspension in water. The main test parameters were those of cycle time and dwell time. The conditions covered were dwell times of 1000, 750, 250 and 5 ms, and cycle times of 7.5, 6.5 and 5.5 s. Two types of die material were used, H.50 and no. 5 die steels (see Table 1).

Ring tests

The annular billets used in the ring test were also of the EN 8 steel specification. Their dimensions followed the proportions used by previous investigators^{5,8}, and were

outer diameter = 2.0 in
inner diameter = 1.0 in
height = 0.667 in

Each test was carried out on the same Petro-Forge as was used for the wear tests, and the colloidal graphite lubricant was applied in exactly the same way and for the same duration as in the wear tests. Only lubricated tests were performed. The design of dies with thermocouples inserted $\frac{1}{4}$ in below the surface were similarly comparable to those used in the wear tests.

The test procedure consisted of heating the dies by means of gas burners to different levels of temperature, applying the lubricant and carrying out the forging operation. The actual temperature of the die was recorded just prior to the forging blow. The energy delivered at each blow was varied such that the amount of reduction fell in the range of 0 to 65 per cent. Measurements of the inner diameter and the height of each billet were taken after it had cooled to room temperature.

EXPERIMENTAL RESULTS

Wear tests

The wear profile appearing on all the dies followed the familiar 'annular' pattern common to flat dies, with wear of the top dies being greater than the bottom dies^{1,3}. Some of the surface profiles produced in the tests are shown in figure 1. From such profile traces the volume of material worn from the die during a test are computed.

The effects of cycle time on the wear volume and on the bulk die temperature under lubricated conditions are shown in figure 2(a) and (b), respectively. The die material for these tests was H.50, and the dwell time was 5 ms. The results show that the inner wear on the top die increases with cycle time whereas the reverse occurs for the bottom die. In figure 2(b) the temperature of the top die is seen to decrease with cycle time as expected, whereas the bottom die temperature actually increases with cycle time. This is due to the fact that an increase of cycle time causes delay in removal of the forging from the die and hence, an increase in the time that the forged part stays on the die.

The effects of dwell time on the wear volume under lubricated and unlubricated conditions are

TABLE 1 Die Material compositions

Material	Composition (%)						
	C	Si	Mn	Ni	Cr	Mo	Va
H.50	0.35	1.0	0.30	—	5.0	1.35	1.10
No. 5	0.60	0.30	0.60	1.5	0.6	0.25	—

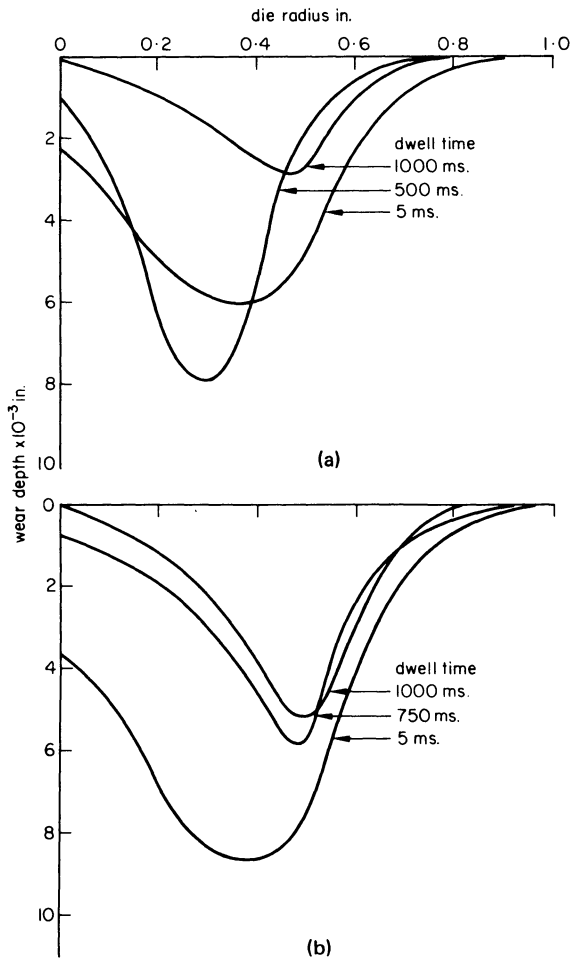


Figure 1. Wear profiles of top dies (lubricated dies, 6.5 s cycle time). (a) H.50 dies. (b) No. 5 dies.

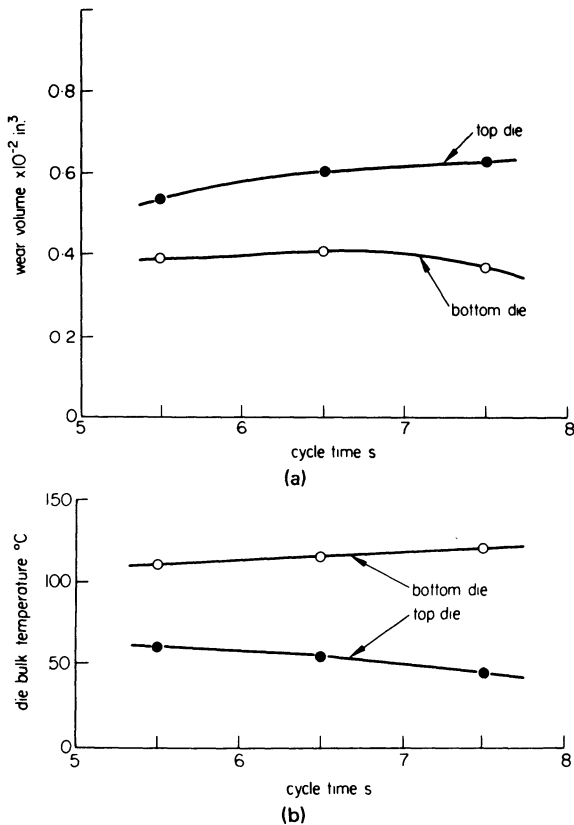


Figure 2. The effect of cycle time (lubricated dies, 6.5 ms dwell time). (a) Wear volume. (b) Die bulk temperature.

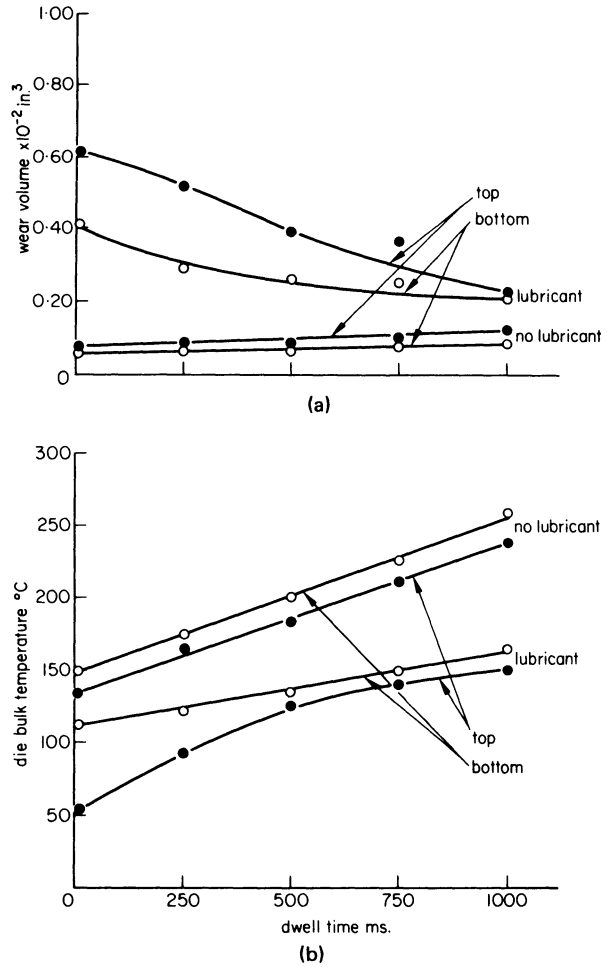


Figure 3. The effect of dwell time (6.5 cycle time, H.50 dies). (a) Wear volume. (b) Die bulk temperature.

shown in figure 3(a). These tests were carried out with H.50 dies and at a cycle time of 6.5 s. Dwell time is seen to affect wear to a more significant level than does cycle time. Under lubricated conditions an increase of dwell time causes a reduction in wear. For the top dies an increase of dwell time from 5 ms to 1000 ms causes a decrease in wear volume of 38.5 per cent, and for the bottom dies the corresponding figure is 49 per cent. Similar trends were found to occur at a cycle time of 7.5 s, and with no. 5 die steel, as shown in figure 4.

Under unlubricated conditions the reverse trend occurs, wear of the top dies increasing by almost 70 per cent over the range 5 ms to 1000 ms, although the level is much lower than in the lubricated case, e.g. for the top dies at a dwell time of 5 ms, the wear is less than 13 per cent of lubricated dies.

The appearance of top lubricated and unlubricated dies after testing are shown in figure 5, each at a cycle time of 6.5 s. It is noticeable that as the dwell time is increased the appearance of the lubricated dies changes and approaches that of the unlubricated dies. Without lubricant the dies' appearance did not change with dwell time.

Typical etched sections of deformed billets taken from each lubricated dwell-time test are shown in figure 6. A flow pattern changing with dwell time can be detected. The top surface at low dwell times indicates a high degree of slipping, that on the

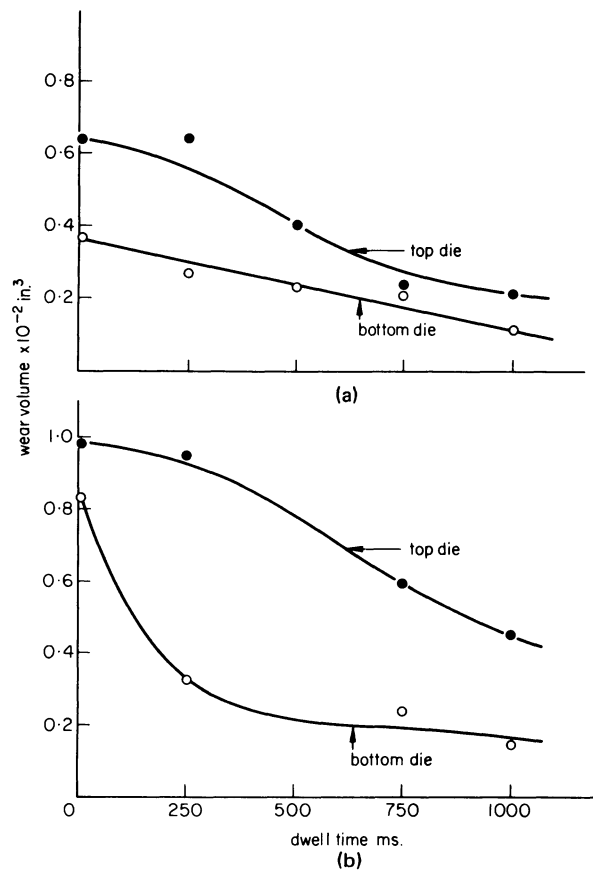


Figure 4. Further effects of dwell time (lubricated dies). (a) H.50 dies, 7.5 s cycle time. (b) No. 5 dies, 6.5 s cycle time.

bottom surface being less. However, as the dwell time is increased the degree of slipping on the top becomes less, such that at a dwell time of 1000 ms there is little evidence of slipping on either top or bottom faces, very similar to the pattern of the upset billets produced under unlubricated conditions.

The readings of temperature for top and bottom dies is plotted as a function of dwell time in figure 3(b) for lubricated and unlubricated conditions. The higher temperature of the bottom die relative to the top follows the trends of previous results obtained on Petro-Forge machines^{2,3}. The main feature of the results of figure 3(b), albeit an expected one, is that for all conditions the die temperatures increase with dwell time. It is noticeable that in comparison with figure 2(b) dwell time causes a more significant change in temperature than does cycle time.

If the wear and temperature results of figures 2(a), 2(b), 3(a) and 3(b) are replotted with wear as a function of temperature the effect is as shown in

figure 7. Generally this plot indicates that cycle time and dwell time effects are not distinguishable and that, with respect to wear, the independent variable is die temperature.

Ring test results

All the ring tests were carried out under lubricated conditions with parameters of percentage reduction and bulk die temperature. As it was not possible to control the die temperature to any great accuracy, the results were grouped into four ranges of temperature. It was found that within any one group an approximate empirical relationship could be applied to the results:

$$\log H = mR + b \quad (1)$$

where H = percentage reduction in the height of the ring,

R = percentage decrease in the internal diameter of the ring,

and m and b are constants dependent on the die temperature.

The values of m and b were regressed for each temperature group and are given in Table 2, together with the regression coefficient and the temperature standard deviation. These results are plotted in figure 8 with percentage change in internal diameter of the ring as a function of percentage reduction in height for the four temperature groups. Superimposed on this graph are theoretical constant-friction curves derived by the method of Hawkyard and Johnson⁹. With unlubricated dies, Jain and Bramley⁸ found that full sticking conditions generally applied, i.e. $\mu = 0.577$.

The experimental results for each temperature group indicate a change of friction with deformation. To some extent this change could be due to the fact that these results were not obtained under constant strain rate conditions, each point being obtained with a different blow energy and, therefore, a different impact velocity. However, Jain and Bramley also found changes in friction with deformation when the impact velocity was held constant.

The method of Hawkyard and Johnson⁹ enables an approximate estimate of the coefficient of friction to be made at any deformation point, by equating the slope of the experimental curve to that of a constant friction curve at the same deformation point. The estimates for the four temperature groups are shown also in figure 8.

The coefficient of friction is seen to increase with deformation following the trend found by Jain and Bramley. However, the significance of the present

TABLE 2 Regression analysis results of the ring test

Temperature (°C)		Regression coefficient r	Equation 1 constants	
mean	Standard deviation		m	b
58.8	4.3	0.89	0.06	-0.003
84.5	5.4	0.93	0.07	-0.72
112.3	5.7	0.94	0.07	-0.39
150.0	7.1	0.96	0.07	-0.05

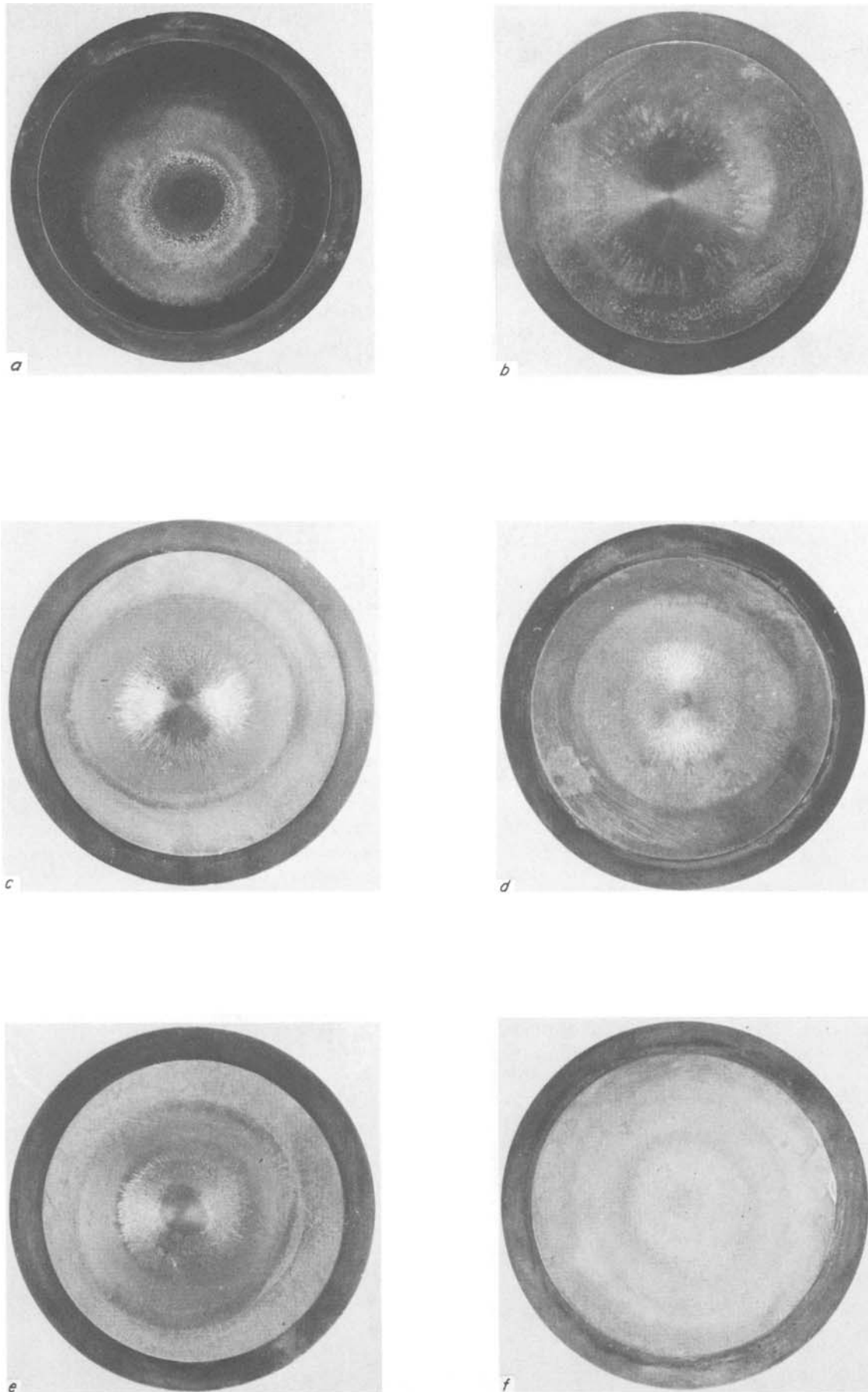


Figure 5. Appearance of top dies after wear tests at different dwell times (H.50 dies, 6.5 s cycle time). (a) 5 ms, unlubricated (b) 5 ms, lubricated (c) 250 ms, lubricated (d) 500 ms, lubricated (e) 750 ms, lubricated (f) 1000 ms, lubricated.

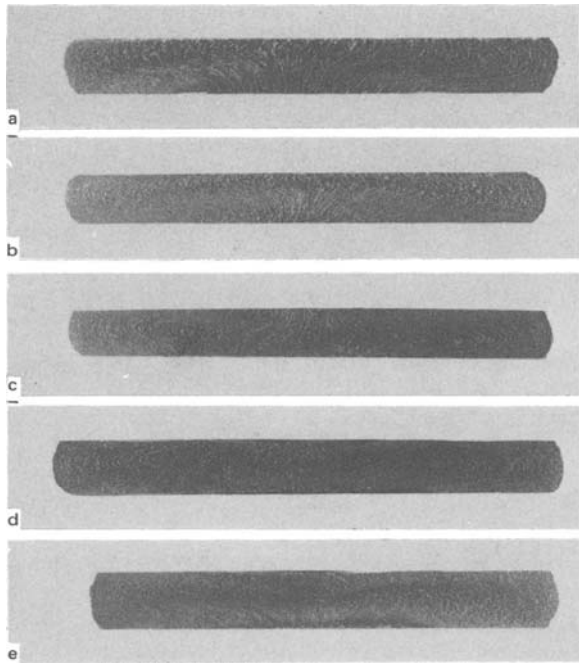


Figure 6. Etched sections of upset billets, at different dwell times (lubricated dies). (a) 5 ms. (b) 250 ms. (c) 500 ms. (d) 750 ms. (e) 1000 ms.

results is that the coefficient of friction increases with die temperature at each point of deformation. At the highest temperature (150°C) full sticking conditions are reached, i.e. $\mu = 0.577$, at a reduction of 45 per cent. Above this reduction the estimated coefficient exceeds the practical limit of 0.577 and this must be attributable to the inaccuracies of the assumed log-linear relationship at higher reductions, as given by Equation (1) and Table 2.

DISCUSSION OF RESULTS

The results of the tests indicate that wear is more dependent on the die temperature under lubricated conditions than under unlubricated conditions (see figure 7). In the latter case no change in the appearance of the dies or in the flow pattern of etched billets with variations in the die temperature was detected. Therefore it is concluded that the slight increases in wear with temperature with unlubricated dies is due to a reduction in wear resistance, probably caused by softening of the die surface. In the former lubricated case, however, the problem is more

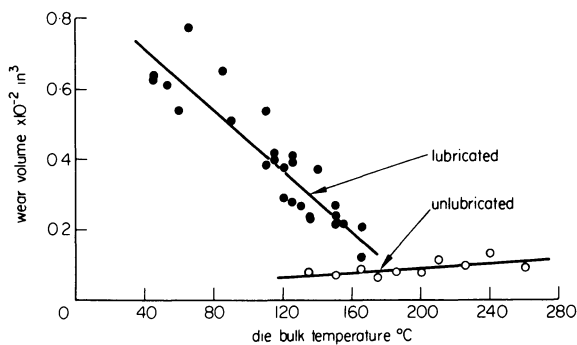


Figure 7. The effect of the die bulk temperature on wear volume (H.50 dies, 6.5 s cycle time).

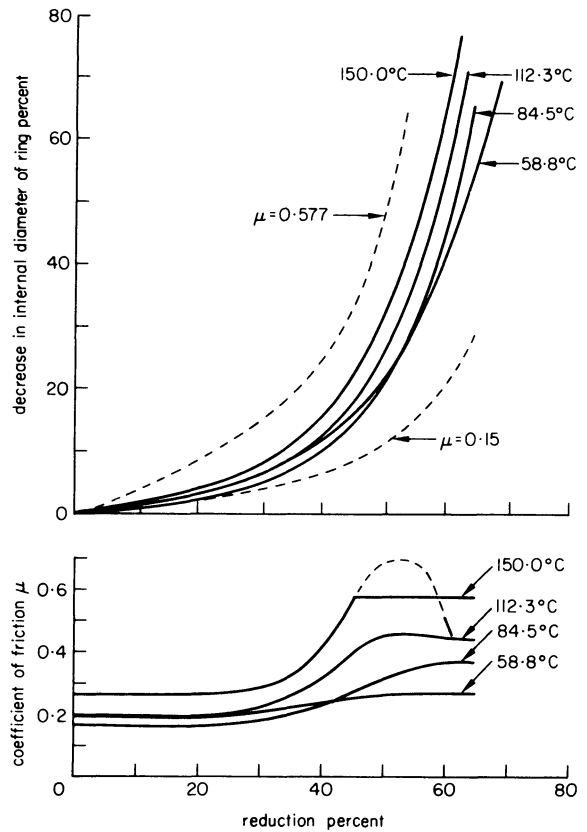


Figure 8. Ring test results at four die bulk temperatures (lubricated dies).

complex as there is evidence of a change in the interfacial frictional conditions as the temperature is varied. Thus to make a true assessment of the effect of die temperature on the wear of lubricated dies in the absence of changes in factors such as metal flow requires a closer examination of the relation between die temperature and metal flow at the die surface.

The results of the ring test (figure 8) show that as the die temperature is increased the coefficient increases, and at the highest temperature (150°C) and at reductions above 45 per cent, is equal to that under unlubricated conditions. It can be expected, therefore, that reduced interfacial metal sliding will occur as the die temperature rises, and this will result in less erosive wear. The extent by which wear is

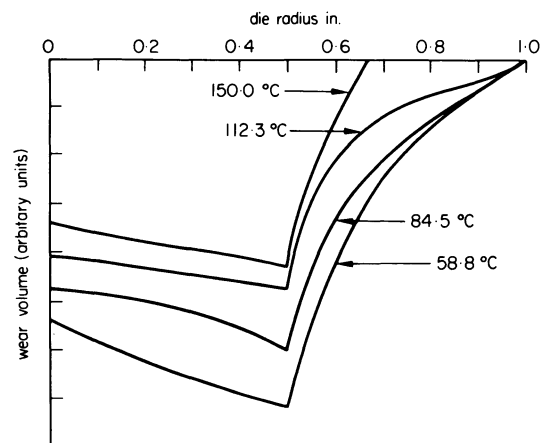


Figure 9. Estimated wear profiles at four die bulk temperatures.

reduced by the change in the flow pattern can be assessed with the aid of a method developed by Thomas¹, which he used to examine the true effect of a lubricant on the wear behaviour of dies during simple upsetting tests. With this method he was able to prove that the increase in wear caused by the application of a lubricant was due to an increase in the degree of interfacial metal sliding and that the lubricant was beneficial in providing protection to the die surface.

In this method it is assumed that the wear at any point on the die surface is proportional to the amount of material that slides over that point during the upsetting operation. Using this assumption curves of wear depth in arbitrary units, as a function of die radius can be constructed. For comparative purposes such a method is considered to be adequate and can be used to compare the wear behaviour of dies at different temperatures when interfacial metal sliding is not constant.

The pattern of metal flow in simple upsetting is such that there are two regions, an inner one where no relative movement between billet and die occurs (the sticking region), and an outer one where sliding occurs. The radius of this sticking zone is a function of the degree of deformation and of the coefficient of friction. It can be approximated to by the relation:

$$r_s = r_i - \frac{h_i}{2\mu} \ln \frac{1}{\mu\sqrt{3}} \quad (2)$$

where r_s = the radius of the sticking region,
 r_i = the instantaneous radius of the billet,
 h_i = the instantaneous height of the billet,
 and μ = the coefficient of friction at the interface.

Equation (2) was used to construct curves of the sticking-zone radius as a function of the instantaneous billet height for each die temperature: assuming that the coefficient of friction was the same as that obtained from the ring tests at each temperature and each reduction. Then, as described by Thomas, wear depth-die radius profiles were drawn, and these are shown in figure 9. The wear volume at each temperature, in arbitrary units, was computed from these curves and, together with the ratio of the wear volume to the wear volume at a die temperature of 58.8°C, are given in Table 3.

TABLE 3 Estimated wear volumes and wear volume ratios

Mean die temperature	Wear volume (arbitrary units)	Wear volume ratio
58.8	101.6	1.00
84.5	80.7	0.80
112.3	62.8	0.62
150.0	42.0	0.41

The predicted profiles in figure 9 can be compared with experimental profiles shown in figure 1, which are for the two die materials at different dwell times. It is seen that there is, at least, a qualitative agreement between the estimated and experimental profiles. To obtain a quantitative comparison the wear-volume ratios, given in Table 3, are plotted

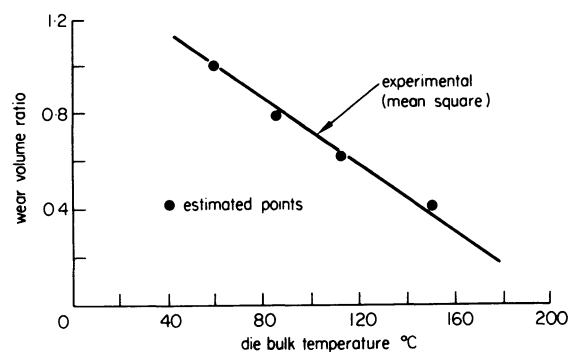


Figure 10. Experimental (lubricated H.50 dies) and estimated wear-volume ratios.

together with the experimental ratios, as a function of die temperature in figure 10. In view of the approximations and assumptions made the comparison is remarkably close. However, even allowing for errors in the approach the results do demonstrate quite clearly that the reduction in wear as a result of increasing the dwell time with lubricated dies is mainly due to the reduction in the interfacial metal sliding as the die temperature is increased. Why this latter phenomena should occur is not known at the present time.

It would seem from figure 10 that for lubricated dies, the die temperature and, therefore, the dwell time have no effect on the wear performance, which is a little surprising. It is certainly a situation that needs still further work. For example the pertinent parameter is the temperature of the die surface which was not measured in this investigation. Also the interfacial flow was not directly monitored and this must be known more accurately for a further understanding of die damage by erosive wear. In fact the main conclusion to be drawn from the presented work is that erosive wear is more sensitive to changes in the billet flow pattern than it is to other forging parameters such as cycle time and dwell time.

SUMMARY OF CONCLUSIONS

- (1) An increase of dwell time during simple upsetting of steel billets was found to significantly decrease the wear of dies lubricated with colloidal graphite, but to increase the wear of unlubricated dies. Under all conditions the bulk die temperature was increased by an increase of dwell time.
- (2) An increase of cycle time was found to marginally increase the wear of the top dies but to marginally decrease the wear of the bottom dies. The bulk temperature of the top dies decreased with cycle time but that of the bottom dies increased.
- (3) The effects of cycle time and dwell time on die wear were resolved primarily to be due to the effects of the bulk die temperature.
- (4) An increase of bulk die temperature lead to an increase in the interfacial coefficient of friction and hence to a decrease in the amount of billet movement over the die surface. At the highest die temperature (150°C) the friction approaches that of unlubricated dies.

- (5) The effect of dwell time reducing the wear of lubricated dies is attributed almost entirely to the decrease in the interfacial 'billet sliding caused by the increased die temperature.
- (6) Erosive wear was found to be more sensitive to changes in interfacial billet movement than to other forging parameters.

ACKNOWLEDGMENT

This work is financed by a grant from the Science Research Council awarded for investigations into die life in hot forging. The wear test results were obtained by Dr A. K. Singh, to whom the author is grateful.

REFERENCES

1. A. Thomas. The wear of drop forging dies. *Tribology in iron and steel works*, I.S.I. preprint 125, p. 55 (1969).
2. B. W. Rooks, S. A. Tobias and S. M. J. Ali. Some aspects of die wear in high-speed hot forging, *Advances in machine tool design and research*, Pergamon, p. 745 (1971).
3. S. M. J. Ali, B. W. Rooks and S. A. Tobias. The effect of dwell time on die wear in high speed hot forging, *Proc. Inst. Mech. Engrs.*, (1971) **185**, 1171.
4. A. K. Singh, B. W. Rooks and S. A. Tobias. Factors affecting die wear, *Wear*, (1973) **25**, no. 2.
5. A. T. Male and M. G. Cockcroft. A method for the determination of the coefficient of friction under conditions of bulk plastic deformation, *J. Inst. Metals* (1964) **93**, 38.
6. Anon. Final report on the influence of high rate forging on die life. Drop Forging Research Association, Sheffield (1969) (Ministry of Technology contract, ref. F/TP/01).
7. B. W. Rooks, K. O. Okpere and R. M. H. Cheng. Automatic handling in hot forging research. *Proc. 1st C.I.R.T.*, University of Nottingham, March 27-9 (1973).
8. S. C. Jain and A. N. Bramley. Speed and frictional effects in hot forging. *Proc. Inst. Mech. Engrs.*, (1968) **182**, pt. 1, p. 783.
9. J. B. Hawkyard and W. Johnson. An analysis of the changes in geometry of short hollow cylinder during axial compression. *Int. J. Mech. Sci.* (1967) **9**, 163.

ROLLING

BALL ROLLING: A LITERATURE SURVEY AND SOME EXPERIMENTAL RESULTS

by

N. R. CHITKARA*, W. JOHNSON* and J. R. S. UTTLEY*

SUMMARY

A preliminary study was made of the process of rolling balls from cylindrical bar stock using Plasticine and tellurium lead as model materials. This paper includes a comprehensive literature survey of previous work related to ball rolling and details of a short experimental investigation of the plastic deformation of Plasticine bars during rolling in a hand-driven mill. Roll separating force and torque were measured when bars of tellurium lead $\frac{1}{2}$ in. in diameter were rolled to balls of 9 mm nominal size on a power-driven helical ball rolling mill. The plastic deformation modes of the Plasticine bars as they progressed through the mill to their final ball shape are also discussed.

INTRODUCTION

Steel balls are employed in large quantities in ore-crushing, pulverizing and in ball mills throughout the world. They are also required in different sizes by the ball-bearing industry. To satisfy their increasing demand for balls in different stages of finish, the Russians appear to have been the first to develop the rolling of balls from bar stock in a specially designed mill and the first commercial installation of such a mill in the USSR was completed in 1950¹. Rolling balls (as opposed to forging and machining) results in a considerable saving of material and labour.

Although this process has been in commercial use in the USSR for the last 15–20 years, it is only recently that Kelco Automation Ltd.³, have started manufacturing and marketing such machines under licence. There is almost a complete absence of any good English language technical literature on the subject—only a few papers^{2–5} describe transverse and unconventional rolling. While the authors' main appreciation of this project came from reading an article by Tselikov⁶, containing an account of the development of helical rolling with its application to ball rolling, thread and form rolling, as carried out by the Union Research Institute of Metallurgical Machine Building (VNII METMAS) a systematic investigation aimed at understanding the basic mechanics of the ball rolling process and, in particular, the choice of the roll pass geometry is, nevertheless, lacking. Research or theoretical analyses of the complex geometry of this process are also difficult. The aim of this paper is therefore to survey the literature on ball rolling and on transverse or

helical rolling as far as it impinges on it. The authors include also details of some observations made on a model experimental mill, using both Plasticine and tellurium lead as model material, with special reference to the cavitation problem.

BALL ROLLING PROCESS

Ball rolling is basically a transverse rolling process. A diagrammatic representation of the ball rolling process is shown in figure 1(a). The essential features of the ball rolling mill are the two cylindrical rolls that have identical helical grooves of a special form cut on their periphery. These rolls rotate in the same sense with the same speed, with their axes set a pre-determined distance apart and inclined at a small angle to each other called the skew angle—normally less than 10° —which depends on the ball size required, is illustrated in figure 1(b). Contact with the rolls causes the bar to rotate in the opposite sense to that of the rolls while a ram pushes the bar forward axially until the helical thread on the rolls bites into the bar and produces an indentation. The bite causes intense local deformation the extent of which increases as the thread penetrates deeper with axial travel of the bar. The bar passes through the mill in a screwing motion and takes up the shape of the roll pass. In each revolution of the rolls a small mass of the bar is deformed into a ball shape; the bar is deformed to become a string of ball-like blocks joined by pips. The pip, however, becomes thinner at each pitch distance along the roll and is finally sheared through at the last sharp edged thread and the ball emerges from a portion of the bar.

* Division of Applied Mechanics, Department of Mechanical Engineering, University of Manchester Institute of Science and Technology

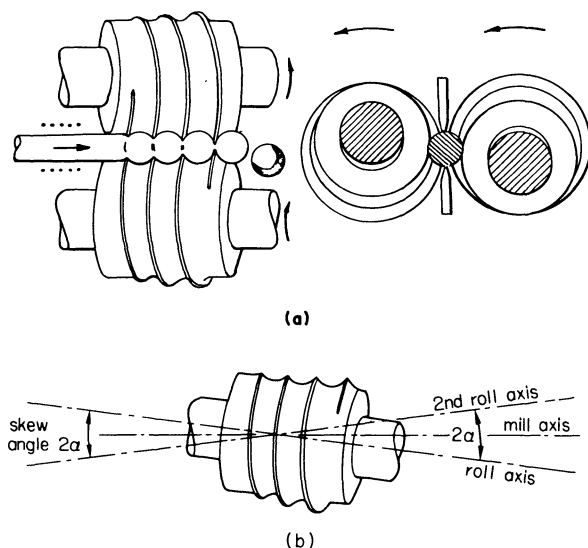


Figure 1. (a) Diagrammatic representation of the ball rolling process. (b) Relative alignment of roll axes.

LITERATURE SURVEY

Transverse and helical rolling: the cavitation phenomenon

Transverse or helical rolling is a term applied to rotary forming processes where the bar is drawn between rolls in a rotary motion transverse to the rotation of the mill rolls and along their barrels. The rolls may be plane cylindrical or double conical or may have special grooves, as in the case of ball rolling. The barrels are usually crossed or 'skewed' slightly and the mills may be two-roll or three-roll types, as used in tube manufacture. Products of great commercial importance, of axially symmetric shapes, such as fin tubes, shafts,* gears, rollers, balls, bicycle hubs, etc., are produced nowadays by one or the other of these rolling processes¹. The literature on the subject is extensive and only the more useful publications are reviewed here. References to most of these publications are listed in the bibliography.

Smirnov¹⁸ seems to be the first to have made a careful study, in 1947, of the deformation in cross-barrel rolling, primarily investigating the reason for the formation of a central, axial cavity down the centre of a cross-rolled bar. He recognized the existence of two schools of thought concerning the cause of cavitation, the first attributing it to shear stresses and the second holding the view that it was due to the normal pressure between the rolls that created tension. Smirnov and his colleagues supported the latter view-point basing their opinion on observations of ruptured tubes in which the diameter of the workpiece had increased after a cavity had formed in it. Ten years later, Mise et al.¹⁹ published the results of their study of the piercing of solid round billets (which were similar to those of Smirnov) and indicated that shear stresses, and not normal stresses, were the prime cause of cavitation; their investigation covered aspects of surface twists

and surface flaws. To study the plastic flow patterns, stainless steel pins were welded into the surface of a bar before rolling.

Tetrin's²⁰ useful contributions to the theory of cross and oblique rolling should be noted, especially his detailed kinematic studies of the relative slip in the tangential direction of rolled billets.

An attempt to calculate the forces set up during deformation and also the stress distribution in cross-rolling was made by Brovman²¹, who pointed out that the stressed condition of the billet was not axisymmetric, as assumed by Smirnov.

In 1960, Tetrin and Louzin¹² conducted tests on composite steel billets in order to investigate the true nature and cause of the axial rupture in cross-rolling. Their results showed a zone of intense plastic deformation at the centre of the billet that led to axial rupture in the composite discs, which was an indication of the development of intense transverse tensile stresses in the central zone. The authors also explained that the increase in critical reduction with a decrease in billet length in the mill resulted from the development of low transverse tensile stresses and verified the explanation by rolling solid billets under conditions similar to those of the composite ones. They observed that rupture occurred after the same amount of reduction as in the composite billets.

Blazynski and Cole²³ investigated redundant deformation in the rotary piercing process by conducting a series of experiments on wax models by rolling laminated wax billets on a two-rolled Mannesmann type piercer.

To prevent the cavitation phenomenon that was inherent in a two-roll piercer, the three-roll mill was developed. Earlier, Tselikov et al.²⁴ had analysed some of the problems associated with three-roll helical rolling mills. Blazynski^{25,26,27} has described experiments on a three-roll piercing and elongating mill using model wax and lead billets. Other work on tube manufacturing was carried out by Harvey and Palmer²⁸, Blazynski and Jubb²⁹, Zorowski and Holbrook³⁰ and, very recently, by Weinstein and Barash³¹.

To widen the field of applications of two-roll and three roll mills, the Russians developed the technique of profile rolling for products such as railway axles, which are axisymmetric in shape, but vary in profile. The three-roll process suffered from such defects as annular fissure, instead of the axial fissure observed in two-roll transverse rolling. An analysis of the rupture in the three-roll process has been attempted by Tselikov et al.³², and Johnson^{33a,b}, who applied the results of a slip line field solution originally proposed for forging in Nasmyth anvils in 1958; see also Tomlenov³⁴.

Kazanskaya et al.³⁵ studied the cause and development of annular cracks in the transverse helical rolling of round periodic sections and observed that the cavitation defects of two-roll transverse rolling could be removed by increasing the number of working rolls; reducing the length of workpiece, and/or by applying external axial tension. Nefdov et al.³⁶ attributed the surface defects in helical rolling to tangential deformation and endorsed Kazanskaya's observation that the application of an axial force to

* See W. Johnson and J. B. Hawkyard. Recent Developments in Metal Forming. *The Production Engineer*. (1971) 239-247. June.

the billet prevented the cavitation defect. Another suggestion was to replace conical contoured rolls by rolls of a parabolic shape.

The reasons for axial cavity formation or annular fissure development may to some extent be understood with the help of plane strain slip line field solutions for forging. Identifying rolls with forging tools, it is easily shown that, when the tool contact length to the bar diameter is small enough, tension is developed about the axis of a bar. The solution of the corresponding problem for a three-roll indentation or forging situation indicates a sound core and the possibility of annular fissure (see reference 33) and the picture given as an example* by Tomlenov³⁴. Annular fissure development is equally likely with three-rolls, as slip line analyses will show.

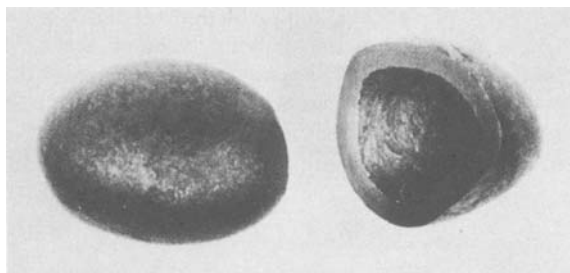


Figure 2. Ball sectioned showing surface of internal cavity.

A fascinating example of unintentional hollow ball manufacture is given in figure 2; this item was originally a ball-bearing failure from a high-speed turbine bearing, in which the ball had been helically rolled at high temperature⁷.

Ball rolling

Though the industrial process of ball rolling has been in practice for the last twenty years, the English language literature concerning it is mainly descriptive²⁻⁵; no detailed analytical work aimed at improving our understanding of the process involved appears to have been carried out. There are, however, a few papers on this subject, mostly by USSR and Japanese research workers and an account of these is included for completeness.

In 1956, Granovskii⁸ gave a resumé of the ball-rolling process and advanced the following reasons for the rolling of balls in helically flanged grooves. If the ball was forged or stamped in a non-continuous manner, then a considerable amount of scrap may be produced. Also, a stamped semi-product for a ball-bearing had a fin, or burr, which impeded subsequent processing. The production of large balls by these methods entailed further machining and grinding. Stamped balls for crusher mills did not possess a true geometric shape and thus led in turn to rapid wear which limited the output of the crusher mills. Granovskii's work does not offer any explanation for selecting the form of the thread for the rolls, but this author comments that a constant volume of material has to undergo deformation in each pass and that the grooves must

therefore have a variable pitch in order to accommodate the changes in profiles and dimensions consequent on the stretching of the compressed bar. He mentions however that, when the ball size rolled is doubled, the power required increases by three or four times and if the ball diameter is increased three times the power consumption increases twelve times.

In Japan, Awano and Ochiai^{9a}, carried out a series of investigations in 1958-59 to compare the properties of hot rolled and hot forged steel balls made of SUJ.2 commercial quality steel. Based on measurements of physical and mechanical properties, e.g. shape, surface roughness, wear, crushing strength etc., they presented results which were, in all cases, in favour of rolled balls. Using an approximate method based on plane strain metal flow conditions in the roll gap, roll force and torque in ball rolling were also estimated. Awano and Ochiai^{9b}, reviewing the salient points of their earlier work^{9a} presented some results of their investigation to study the plastic flow of Plasticine as a model material. Awano et al.^{9c} (1961) further compared the mechanical and other properties of rolled steel balls with those of die forged balls. Based again on measurements of static properties and on practical tests, they presented results in favour of rolled balls in all cases.

In their paper entitled, *The mastery of ball rolling machines*, Kurganov et al.¹⁰, in 1962, mention the possibility of doubling the rate of production of balls from round billets using double-threaded helical rolls. Balls with a diameter of 40 mm were made first and this was increased later to 50 and 60 mm.

Masharov and Nikolau¹¹ claim to have successfully rolled balls of different sizes in four-start threaded rolls. The minimum diameter of the ball produced was reduced from 127 to 62 mm. Apart from reviewing the ball rolling process, these authors describe the difficulties encountered during the evolution of a roll pass shape. One such pass shape showed the changes in the heights of the groove flanges and the changes in radii of the necks between the balls for a four-start thread form, when the ball diameter was 41.5 mm. However, they do not indicate clearly how the groove shapes were arrived at and to state merely that force and torque measurements were used to correct the design of the roll shape. Mention is made of the fact that the flange width was calculated on the basis of constancy of volume, but they say that they cannot predict correctly the elongation which takes place in the neck of the multi-threaded grooved rolls, which would help in selecting different pitch distances between the finishing and the forming zones of the pass. The process is also briefly mentioned by Holub³⁷. Tselikov and Javorankov¹² (1965), quote some figures for roll force and power for the hot rolling of balls of respectively 28, 55 and 80 mm diameter. Valueva et al.¹³ (1970), describe the rolling of 31.3 mm diameter grinding mill balls on a 620 skew-rolling mill at the Azovstal works. Rolling is said to have been done on a four-threaded pass in order to ensure the production of balls of a correct geometric shape and excellent surface quality. A special template was used to monitor the cutting of the pass. Tselikov et al.³⁸ (1970), discuss and compare the practical

* Reproduced in: W. Johnson and P. B. Mellor. *Engineering Plasticity* Van Nostrand-Reinhold p. 504 (1973).

characteristics of a special rolling mill for rolling periodic profiles. Their conclusions were that mills designed for special complex profiles were economical and saved about 500 000 tons of steel per annum in the USSR.

Two more papers of relevance to this subject are: *Thread and Form Rolling*³⁹ and *Analysis and Experiment in V-Groove Forming*⁴⁰.

In the USA, Greis⁴¹ (1970), using a computer, has extended the principle of thread rolling to develop a roll forming process, which is essentially an ultra-high speed rotary open die forging process that produces balls and roll pins at a rate of 600 to 4000 every minute.

Nussbaum⁴² (1971) describes the development of the Floturn process by the Lodge and Shipley Co. of Cincinnati, Ohio, which is, basically, a rotary metal forming process that produces parts, round in cross-section, but which may be straight-sided contoured cones or cylindrically shaped. The process consists of the cold plastic deformation of the material to the shape of a rotating arbor by means of power fed rolls.

To estimate the roll force P in ball rolling, the authors endeavoured to use

$$P = n \left(\frac{\pi}{4} d^2 \right) Y_m C$$

where n is the number of balls of diameter d in the mill at any one time, Y_m is the mean yield stress of the material at the appropriate strain rate and C is a constant, reflecting, among other things, the fraction of the total amount of possible contact (projected area) between the balls and the rolls. This suggestion is commented on again below. A similar approach was used by Kudo and Yokai¹⁴. Dobrucki¹⁵, however, suggests treating this case as a classical forging problem in plasticity, that is, of forcing two rigid and very large punches into a plastic cylinder, and suggests the following purely empirical function f for roll force,

$$P = \sum_{i=1}^{i=m} f(\kappa_f, d, D, \alpha) \quad (1)$$

where K_f is a material property, d the ball diameter, D the roll diameter and α , the semi-skew angle; $i = 1$ to $i = m$ refer to the number of disc slices into which the rolled bar within the profiled rolls is subdivided.

A very approximate estimate of roll load and torque in ball rolling

Consider a simple compression type analysis so that during rolling a peak force P occurs. Then

$$P = AY_m \quad (2)$$

where A is the average projected area of contact between one of the rolls and the bar and Y_m is the average yield stress of the material. If at any one time there are n number of balls in the pass, then the maximum projected area of the bar in contact with the rolls (assuming no increase in ball diameter), is

$$A_{\max} = n \frac{\pi}{4} d^2$$

Only a proportion c of this area, however, will contact the rolls at any one time. Thus, the projected area in contact may be written as

$$A = n \left(\frac{\pi}{4} d^2 \right) c. \quad (3)$$

Roll separating force is consequently

$$P = n \left(\frac{\pi}{4} d^2 \right) c Y_m \quad (4)$$

for a hot rolling condition. Similarly, as a first approximation, the roll-torque per roll relationship may be written as

$$T = ARk = AR(Y_m/2) \quad (5)$$

where R is the average radius of the roll and k is the yield shear stress of the material taken as $Y_m/2$. In Equation (5) it would be more appropriate to use the curved area of the ball in contact instead of the projected area. If A^* is the curved area of the ball in contact, the modified equation for roll-torque per roll is

$$T = A^*Rk = A^*R(Y_m/2). \quad (6)$$

The above approach, however, neglects redundant work, variation of roll radius, pressure distribution and change in pressure angle from section to section across the roll.

EXPERIMENTAL DETAILS

Equipment

Experimental work was carried out using Plasticine and tellurium lead as model materials. The preliminary set of experiments carried out with Plasticine bars had as its object the study of the plastic deformation of the material during the process with the aim of understanding the basis of roll shape design. The equipment was a hand-driven Mannesmann type, two-roll mandrel model mill, which was modified into a $\frac{1}{4}$ in diameter nominal size ball mill. This was done by replacing the wooden rollers in the model mill by 2.5 in nominal diameter mild steel rolls, having special grooves cut on their periphery.

Necessary mechanisms for the relative alignment of the rolls in the axial and angular directions were also incorporated in the apparatus and protractors were used to set the mill to any skew angle. The modified mill is illustrated in figure 3, showing a part-rolled Plasticine specimen undergoing deformation, but with the mill stopped.

A second group of investigations was carried out on $\frac{1}{2}$ -in diameter annealed tellurium lead bars in a power-driven 9-mm diameter nominal size ball mill, the rolling load and torque required being measured. An existing two-high conventional rolling mill, described elsewhere¹⁶, was modified to a two-roll transverse type ball mill. Thread grooves were cut into the rolls. Provision for feeding and guiding the stock—bottom and top guides—and measuring load and torque were also incorporated. Load cells and torque meters were calibrated and suitably connected

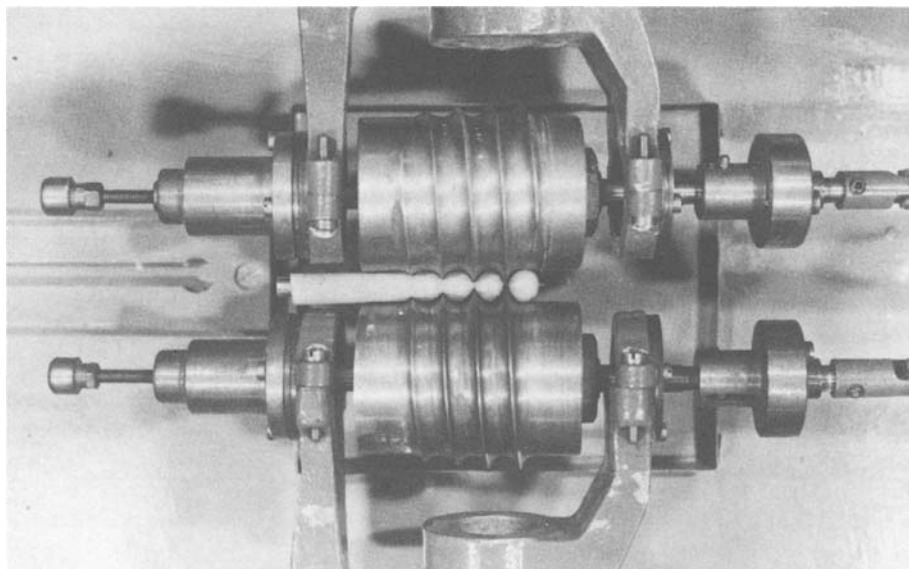


Figure 3. Model Plasticine ball rolling mill.

during the tests to a 6-channel ultra-violet galvo-recorder through switching and balancing units.

Details of the hand-driven ball mill for rolling Plasticine and of the power-driven mill for lead balls are given by Uttley¹⁷ who describes also the techniques used for the preparation of different types of Plasticine specimens and the special roll-turning procedures adopted for cutting threaded groove rolls.

Shape of the roll pass

A typical roll pass shape (shown in the article by Masharov and Nikolau¹¹, who reported on their investigations into ball rolling in a 41.5 mm diameter ball mill having multi-threaded rolls) was scaled down for investigations to the smaller ball size mentioned and a set of multi-threaded rolls was prepared and tried. It soon became clear that rolling with multi-thread rolls was difficult in the absence of more detailed data from the work of Masharov and Nikolau, such as the number of threads, the diameter of the rolls and the skew angle used. Hence, single thread rolls were prepared, the shape of the pass being shown in figure 4(a); this was done, firstly, by scaling down the Masharov and Nikolau shape to conform to the size of the ball and, secondly, by modifying the width and/or depth of the grooves so generated in order to give constant pitch and the requirement of constant volume of stock in each pass.

Experimental procedure

The procedure for rolling Plasticine balls was first to align the rolls and then to check thread by thread, using the mechanism provided for relative alignment in the axial and angular directions of either of the rolls and then by revolving slowly both rolls by hand. The rolls were set at a predetermined distance and adjusted to a required skew angle with the help of an angle protractor placed on each of the yokes. A well prepared bar-shaped specimen was then fed to the mill through the feeding tube and gently pushed while the rolls were revolved through the manually-operated driving mechanism. French chalk was used as lubricant and to prevent the Plasticine from sticking to the bottom guide. When the bite of the

rolls was adequate and the specimen started to revolve, the bar was allowed to go through the pass to its final shape. A trial and error procedure was adopted to select the right skew angle and to secure the correct ball shape. During the experimental work it was noticed that the bar had a tendency to lift up from the mill. Many guides were tried, but later it was observed that a small pressure applied from above was sufficient to hold down the specimen.

In the case of lead balls, since the stock size of the extruded bars was $\frac{1}{32}$ in. in diameter, the roll shape, as shown in figure 4(a), was first scaled-up and tried. Lubrication was provided by spraying PTFE on to the rolls. The experimental procedure for mill setting and aligning the rolls was similar to that used for the Plasticine balls.

When the lead bars were presented to the rolls the top guide was acting on the balls in such a manner that too much plastic work was done on the lead specimens early in the process and consequently the balls were sheared through earlier than anticipated. Modifications to the roll pass shape and a trial and error procedure were therefore adopted, namely, rolling one of the bars at a time, noting the effects on it and taking corrective measures—either alignment of the guides or modifications to the guides and/or rolls. The modified roll pass shape, shown in figure 4(b), was finally obtained when balls of reasonable shape and size were produced with a roll gap setting distance of 0.365/0.370 in and a skew angle of 7° . The subsequent experimental work was performed on this basis.

Rolling load and torque were measured in the tests, details of which appear in the next section, which contains also the results of the plastic deformation using Plasticine.

RESULTS

Some typical results of rolling Plasticine balls from bar stock are shown in figures 5, 6 and 7. In figure 5(a), a partly-rolled stone-coloured Plasticine bar is compared with an unrolled Plasticine bar. A fully-

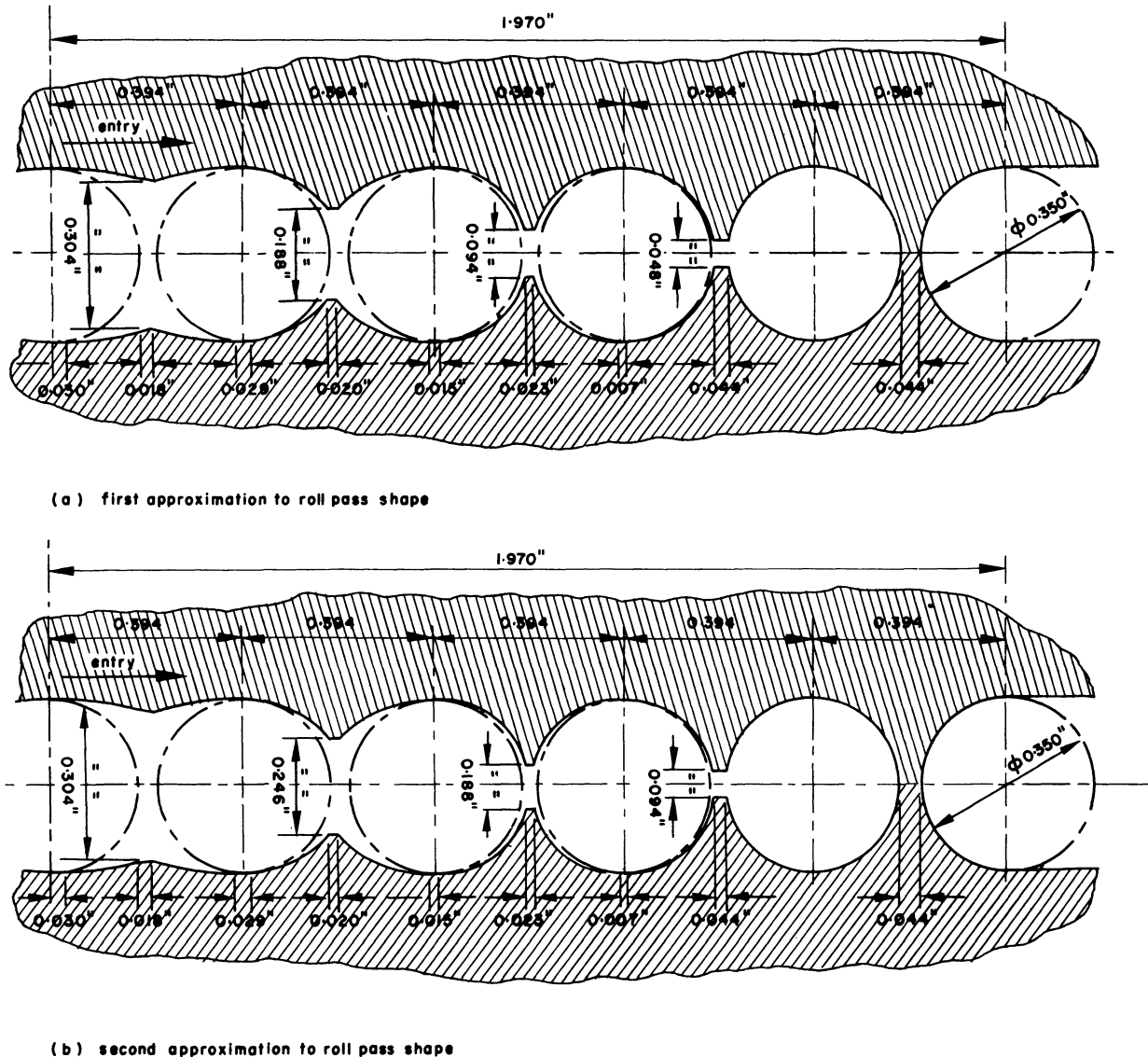


Figure 4. Design of roll pass shapes.

rolled Plasticine bar and a fully-made ball are also shown for comparison. Figures 5(b) and 5(c) show partly-rolled composite specimens made up of Plasticine of different colours when sliced into halves. Figure 6(a) shows a composite specimen made from a layer of a coloured Plasticine, sandwiched between segments of differently coloured Plasticine, part-rolled. An original specimen and a section of the stock is also shown for comparison. Transverse shear of the surface and surface twist are clearly visible. The twist modes, apparent in part-rolled Plasticine specimens made of stone-coloured Plasticine, with an ink line drawn on their surface before rolling, are also shown in the same figure. Figure 6(b) shows four of the Plasticine specimens when rolled at different mill settings and the progressive modes of deformation in the Plasticine bars during ball rolling (when the mill was stopped after gradually increasing rotational distance) are shown in figure 6(c). Figures 7(a) and 7(b) show the variation of mean average and the peak average values of rolling loads and torques, measured when specimens of tellurium lead were rolled at the roll gap settings and at various roll speeds with a roll skew angle set at 7° .

A selection of balls produced at different rolling speeds when the roll gap was 0.370 in were measured. The ball diameter varied from 0.370 to 0.386 in, with an average ball diameter of 0.379 in. The ball diameter was always found to be slightly larger than the initial roll gap. Details of the variation of ball size across diameters marked A and B (as shown in figure 12) are given in Table 1 in the Appendix. The increase in diameter of the ball compared to bar stock diameter was reported by Granovskii⁸ and Dobrucki¹⁵. To produce balls of good structure, Dobrucki recommends that the billet diameter should be about 97 per cent of the ball diameter. This compares well with our results.

Figure 8 shows the stress-strain behaviour of the tellurium lead material used in experimental investigation with different platen speeds. Figure 9 shows a random collection of lead balls produced (good and bad) and a one-half new pence coin and a finished 1.25 in diameter ball to facilitate size and accuracy comparisons. The balls shown with axial ruptures were produced when the roll gap setting was 0.355 in.

Figures 10(i) and (ii) show the imprints of the ink picked up by the previously rolled bar and by the

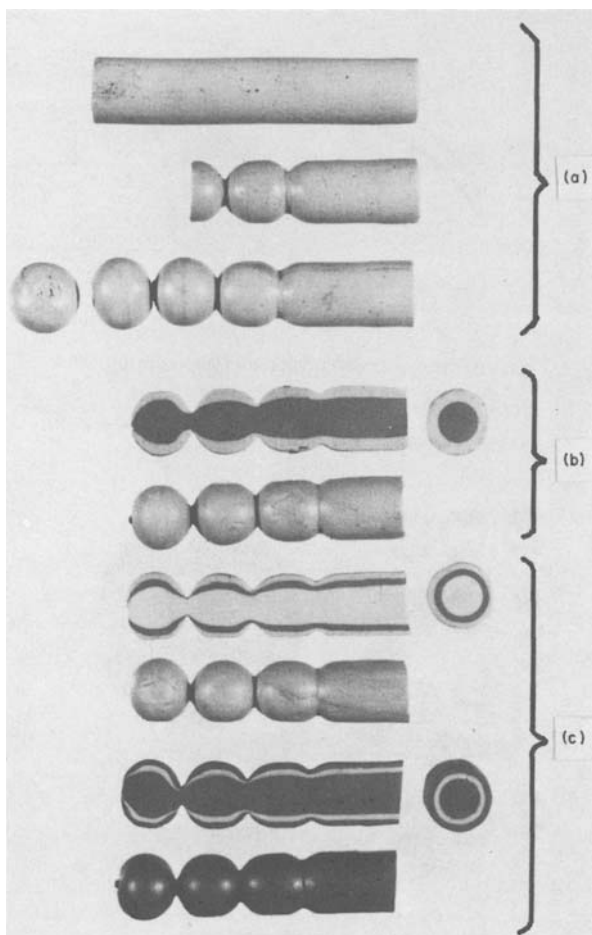


Figure 5. Rolled Plasticine bars.

unrolled Plasticine bar from the inked rolls when these bars were simply pressed between the rolls, the rolls being adjusted to the same original settings as used in rolling. An unrolled bar (iii) is also shown for comparison. In the figure, an engine-divided scale is also included to help to estimate correctly the projected area of contact between the bar and the rolls. A sample of calculations to estimate the roll load and torque for this case (on the assumption that lead is the material) is given in Appendix I. The procedure used to calculate the curved area in contact from the projected area is shown in figure 11.

DISCUSSION

Large amounts of local deformation occur in regions where a bar is indented by the roll threads, but the deformations do not penetrate right through the body, see figures 5(a)-5(c). The material, when indented, squeezes out on either side of the thread. This is also clear from figure 5(b), which shows that a small inner core remains almost undisturbed during its passage through the pass; when the ball diameters were compared with the original bar diameters, the balls were slightly larger.

Surface shear and twist are clearly illustrated in figure 6. Large surface shearing action takes place in the zone where a ball is finally separated from the bar stock; the ball spins and is wrung away from the body. Shearing action also increases as the skew angle of the roll increases, see figure 6(b). The load and

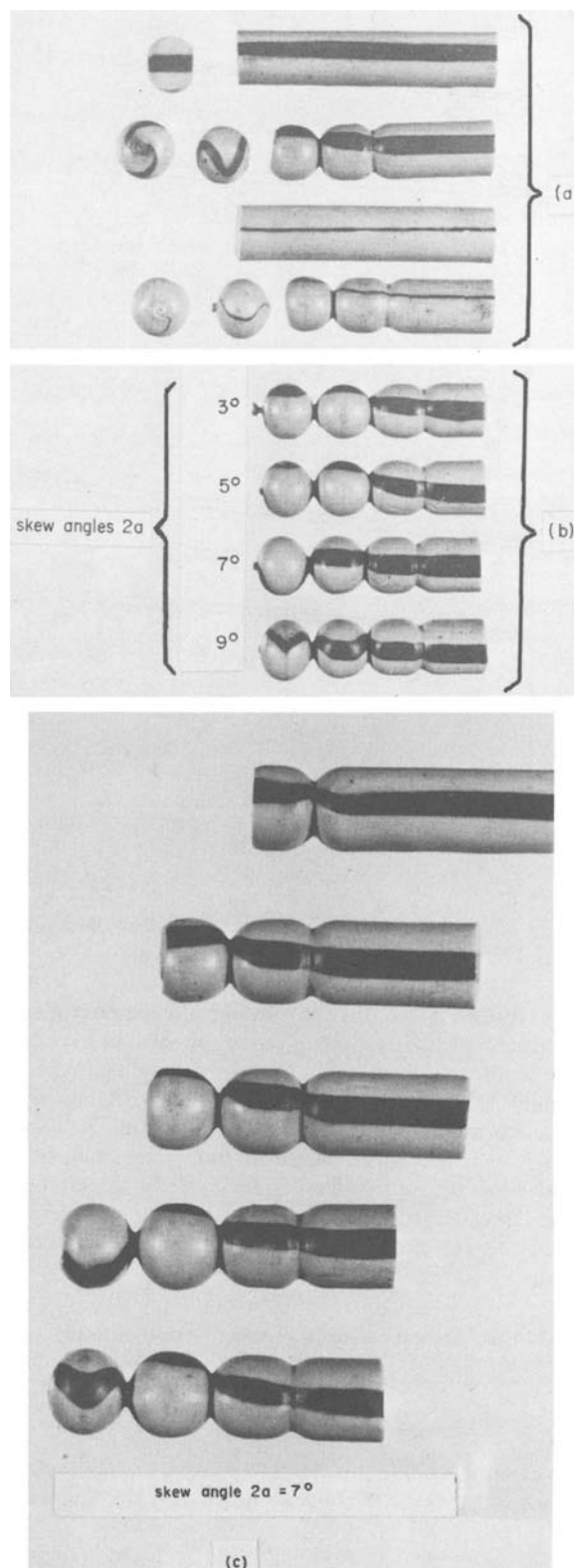
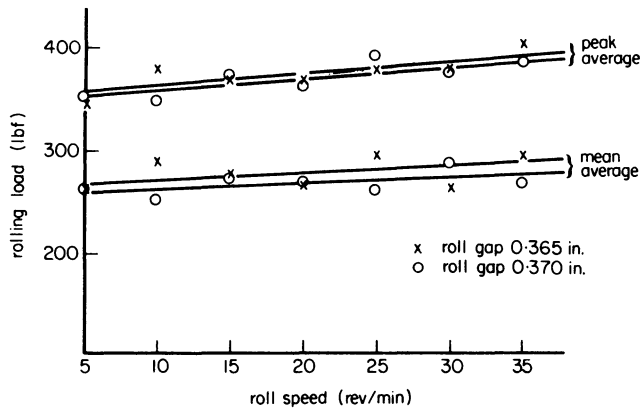


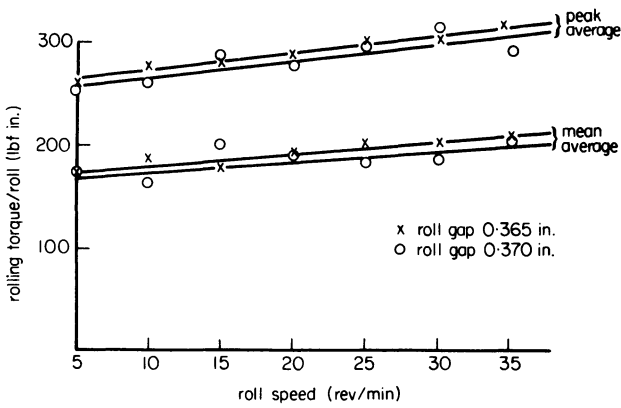
Figure 6. (a), (b) Shear in ball rolling. (c) Progressive deformation in ball rolling.

torque vary throughout each revolution of the rolls during the steady production of balls. The mean average and peak average values of the load and torque vary little with mill speed, as shown in figure 7. Strain rate and frictional characteristics would be expected to have small influences only on these two parameters.

The calculated value of the rolling load (280 lbf)



(a)



(b)

Figure 7. (a) Variation of average rolling load with roll speed. (b) Variation of average roll torque.

and torque (142 lbf. in), using the approximate projected area approach given in equations (4) and (5), compares well with the corresponding experimental values of the rolling load (270 lbf) and the mean torque (180 lbf. in) plotted in figures 7(a) and 7(b), for a roll speed of 20 rev/min. The calculated values of the torque, however, are slightly lower, but when the curved area of contact is used for calculation, the agreement between the two values is very close.

TABLE 1 Typical ball diameters for 7° skew angle and 0.370 in roll gap at different speeds (see figure 12).

Speed rev/min	* Dimensions in	Specimen Numbers			
		1	2	3	4
5	A	0.386	0.381	0.381	0.381
	B	0.385	0.380	0.379	0.380
10	A	0.382	0.382	0.380	0.374
	B	0.378	0.377	0.380	0.372
15	A	0.377	0.382	0.382	0.375
	B	0.377	0.384	0.380	0.374
20	A	0.370	0.386	0.378	0.377
	B	0.368	0.387	0.380	0.376
25	A	0.379	0.380	0.382	0.371
	B	0.381	0.378	0.379	0.371
30	A	0.383	0.372	0.377	0.383
	B	0.383	0.373	0.378	0.382
35	A	0.375	0.383	0.377	0.377
	B	0.373	0.383	0.377	0.379

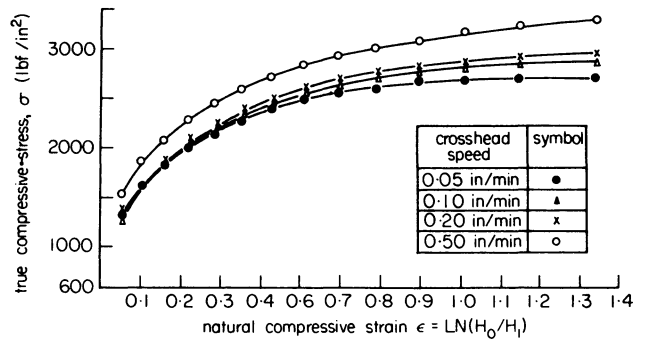


Figure 8. Stress-strain curve; material: tellurium lead.

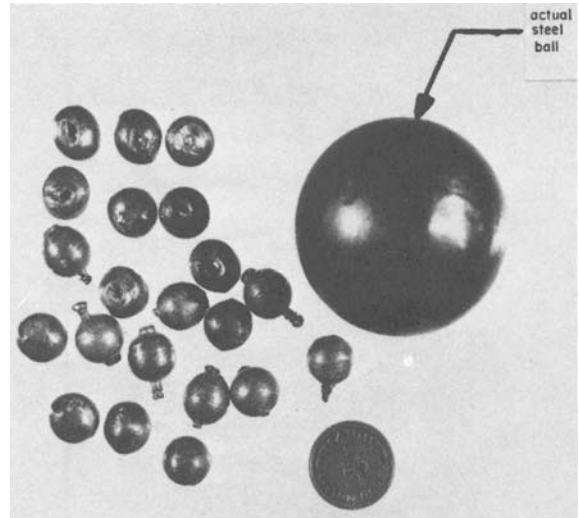


Figure 9. Ball mill products.

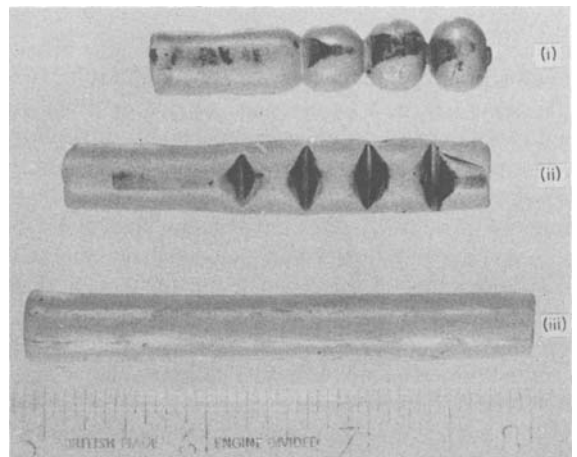


Figure 10. Imprint of roll on Plasticine bars in process and on an unrolled bar.

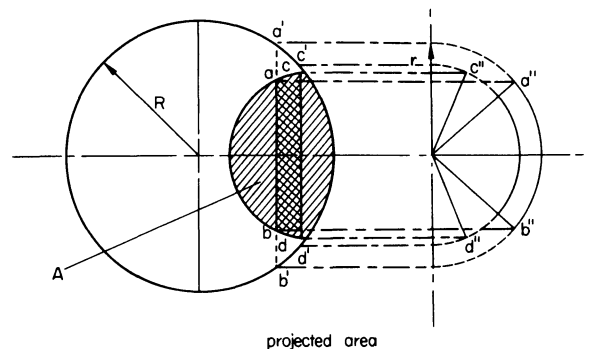


Figure 11. Graphical construction to evaluate curved area from projected area.

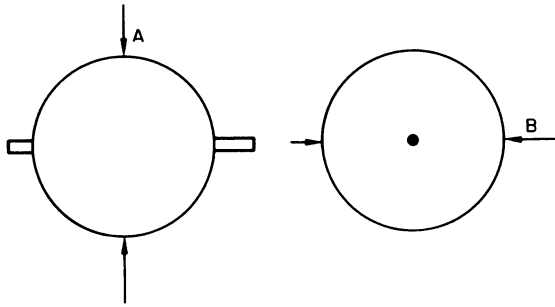


Figure 12. Dimensional measurement of balls.

Figure 9 shows that most balls have either pips or no pips, but some balls have axial ruptures. Similar observations on the formation of internal cracks are also reported by Dobrucki¹⁵, where the billets were of too large a diameter for normal roll settings. This clearly shows the possibility of certain conditions arising at certain stages that must be avoided if balls of the correct shape and integrity are to be produced.

CONCLUSIONS

From this limited preliminary report, the following remarks may be made in conclusion:

- (1) For the rolling of good balls, the shape of the pass and the skew angle of the roll setting have an important bearing. The billet diameter should be about 97-97.5 per cent of the ball diameter. The skew angle should be large enough to enable the axial component of roll velocity to equal the axial speed of the compressed material.
- (2) There is a critical roll setting position at which axial fracture is likely to occur in the balls.
- (3) Deformation in ball rolling is primarily an intense local indentation between roll threads, combined with heavy localized surface shear.
- (4) The breakage of the ball from the stock is a wringing action due to spin of the ball with the stock pressed between the threads.

There is evidently room for a more detailed experimental investigation of this process. It is doubtful if a theoretical study can do other than suggest rational approaches to an understanding of the process and the grouping of data.

APPENDIX I

Figure 11 shows the projected areas for a ball of a nominal diameter of 0.350 in contact as being $(0.051) + (0.038) + (0.051)$ in², i.e. a total area of 0.140 in². (A planimeter was used for this measurement.)

Assuming an average value for Y_m of 2000 lbf/in², then the roll force will be: $F = AY_m$ (modified Equation 2).

$$0.140 \times 2000 = 280 \text{ lb}$$

The torque on one roll will be:

$$T = ARY_m/2 \text{ (Equation 6)}$$

where A , A^* and R are as defined in the text. Then if $A = A^*$

$$T = 0.140 \times 1.013 \times \frac{2000}{2} = 142 \text{ lbf. in}$$

These values of load and torque may be compared with the observed values shown in figure 7. However, A^* is always larger than A and its value lies between $1.0 \leq A^*/A \leq 1.57$, depending upon the position of the area in contact. This ratio of A^*/A is on the high side if more of the projected area shows up near the circular boundary while the converse is the case if the area is more towards the centre. A more reasonable figure would be $A^*/A = 1.25$. Then the torque

$$T = 142 \times 1.25 = 178 \text{ lbf. in}$$

A graphical method of calculating curved area A^* from the projected area A is illustrated in figure 11 and is self-explanatory. In the figure, $a''b''$, $c''d''$ are the curved lengths corresponding to the projected lengths ab and cd .

ACKNOWLEDGMENTS

The authors wish to thank Mr G. Robinson, Mr J. Howe, Mr A. Bowers and the members of the workshop staff for their help.

REFERENCES

1. A. I. Tselikov and V. V. Smirnov. *Rolling Mills* (English translation of Russian work). Pergamon Press, 11, 1965.
2. C. Vaughan. Transverse rolling for the production of finished and semi-finished components. *I.S.I.* (1969).
3. F. Neuberger, L. Mockel and L. Rolz. Transverse Rolling. *Metal Forming* (1968), 35, 10.
4. P. F. Egan. Transverse rolling solves many problems. *Metal Forming* (1967), 34, 4.
5. Redman's Tools Ltd. Hot roll forming from bar stock. *Metal Forming* (1968), 35, 3.
6. A. I. Tselikov. Helical rolling of steel in the USSR. *Steel Times* (1965), November 26, 694-696.
7. Company Report. Hollow rolling of balls of a damaged rolling bearing. Tech. Rep. VIII, Brit. Eng. Boiler and Elect. Ins. Co. Ltd. (1968).
8. S. P. Granovskii. Russian method for rolling ball bearings and crusher balls. (In Russian). Translated from *Stal* (1956), 333-357.
9. T. Awano and M. Ochiai. (a) Technical report, Ngoya Kogyogijyutsu shikenjo Hokoku, (March, 1958) 7; (September, 1958), 7, 631; (October, 1958) 7, 709; (February, 1959) 8, 91; (March, 1959) 8, 151; (b) Ball Rolling. A review of work listed in 9(a), *J. of J.S.M.E.* (October, 1959) 62, 489, 1467-1473.
9. T. Awano, M. Ochiai, M. Ito and A. Kozak. (c) Metal flow in forging (XV) Properties of rolled steel balls. Rep. Gov. Ind. Res. Inst. (1961), 10, 757-774 (In Japanese.)
10. M. A. Kurganov, A. P. Nikolau and A. A. Masharov. The mastery of ball rolling machines. *Stal* in English, 277-279 (April, 1962).

11. A. A. Masharov and A. P. Nikolau. Rolling balls in multi-thread helical grooves. *Stal* (1962), 6, 537-538.
12. A. I. Tselikov and V. A. Javoronkov. New methods of rolling in production of machine parts. *Int. Conf. on Iron and Steel Making, Düsseldorf* (1965).
13. E. O. Valueva, A. A. Masharov and A. P. Nikolau. Rolling 31.3 mm diameter balls on a 620 skew rolling mill. *Stal* (in English), (December, 1970), 975-976.
14. H. Kudo and M. Yokai. Investigations into the helical rolling process. *Proc. of the 8th M.T.D.R. Conference* (1967).
15. W. Dobrucki. Lecture entitled: Investigation of the Ball Rolling Process, given at the University of Aston, Birmingham, 3 November, 1967.
16. W. Johnson and G. Needham. An experimental study of asymmetrical rolling. *App. Mech. Convention*, Cambridge, Inst. Mech. Engrs. April, 1966.
17. J. R. S. Uttley. M.Sc. Dissertation. Manchester University, October, 1971.
18. V. S. Smirnov. The deformation process in cross-rolling. (In Russian). Translated from *Stal* (1947), 7, (6).
19. S. Mise, T. Okamoto and A. Funachi. Metal flow in oblique rolling. (In Japanese), Translated from *Sumitoto Metals* (July, 1957), 9.
20. P. K. Tetrin. Contribution to problems concerning the kinematics of the cross and oblique rolling processes, (In Russian), *Stal* (October, 1958), 923-925. *I.S.I. Translation No. 1175*.
21. M. Y. A. Brovman. Deformation of the work-material in cross rolling, Translation No. 1525. Translated from *Stal*, Rolling and Tube Manufacture Supplement (1959), 22-23.
22. P. K. Tetrin and Yu. F. Louzin. On the mechanism of metal rupture during cross rolling. *Stal* (October, 1960), 10, 930-932.
23. T. Z. Blazynski and I. M. Cole. An analysis of redundant deformation in rotary piercing. *Proc. Inst. Mech. Eng.* (1963-64), 178, Pt. 1, (33).
24. A. I. Tselikov, V. M. Lugovski and E. M. Tretyakov. The theory of transverse cold rolling in three roll mills, *Russian Engineering Journal* (1961), (7).
25. T. Z. Blazynski. An assessment of a combined rotary piercing and elongating process for tube making. *J. Inst. of Metals* (1967), 95, 97-105.
26. T. Z. Blazynski and C. Jubb. (a) An assessment redundant shears in the combined rotary-piercing elongating process. *J. of Strain Analysis* (1968), 3, (4).
27. T. Z. Blazynski. Theoretical and experimental assessment of roll and plug forces and roll torques in the combined rotary piercing-elongating process. *Int. J. Mech. Sci.* (1968), 10, 903-925.
28. S. J. Harvey and W. B. Palmer. Analysis of tube relieving process. *Int. J. M.T.D.R.* (1967), 1059-1074.
29. T. Z. Blazynski and C. Jubb. (a) An assessment of the roll separating force and torque in the Assel tube elongating process. *Proc. of the 11th Int. M.T.D.R. Conference*, Manchester (1970); (b) Major defects in rotary tube making processes. *J. Inst. of Metals* (1969), 97, 363.
30. C. F. Zorowski and R. L. Holbrook. Influence of mill set-up on hollow geometry produced by rotary piercing. *Proc. of the 8th M.T.D.R. Conference*, Birmingham (1967).
31. J. S. Weinstein and Moshe M. Barash. A study of three-roll piercing. *Int. J. Mech. Sci.* (1971), 14.
32. A. I. Tselikov. Stress and strain in metal rolling. M.I.R. Publishers, Moscow (1967).
33. W. Johnson. (a) An approximate treatment of metal deformation in rolling, rolling contact and rotary forming, *Int. J. Prod. Res.* (1964) 3, (1); (b) Indentation and forging and the action of Nasmyth anvil. *The Engineer* (5328) (1958), 205.
34. A. Tomlenov. Mechanical Working of Metals. Moscow, 1963.
35. I. I. Kazanskaya, M. G. Panfilov and V. I. Ippolitov. Reason for the development of defects in transverse helical rolling of round periodic sections. *Stal* in English (September, 1962), 12, 709-711.
36. A. A. Nefdov, L. A. Brezhnev, A. P. Sichevoi, O. P. Bystrikov and E. A. Mil'man. A study of metal deformation under transverse helical rolling, *Stal* in English (May, 1964), 14, 371-374.
37. J. Holub. Transverse hot rolling. *Machinery* (January, 1963).
38. A. I. Tselikov, S. P. Milyutin, I. I. Kazanskaya and E. I. Levin. Technology and rolling mills for rolling round periodic profiles, *Metallurgy* (April, 1970), 40-43 (In Russian).
39. D. H. Seymour and F. M. Lomas. Thread and form rolling. *Conf. on Technology Eng. Mann.*, Paper 20. 1956, Inst. Mech. Engrs.
40. H. Kudo and K. Tamura. Analysis and experiment in V-groove forming, C.I.R.P., September, 1968.
41. H. A. Greis. Skewed axis roll forming. A new way to make small parts fast. *Machinery* (N.R.) (December, 1970), 76, (16) 42-45.
42. A. I. Nussbaum. Floturn shear forming process. *Sheet Met. Ind.* (June, 1971), 48, 471-475.
43. J. Jech. New Technology for manufacture of serious improved life of antifricition bearing balls. *Hutnik* (1971), 21, (2), 66-69 (in Czech).

BIBLIOGRAPHY

COLD ROLLING OF RING GEARS

by

J. B. HAWKYARD*, F. R. NAVARATNE* and W. JOHNSON*

SUMMARY

Investigations are described into the cold forming of spur gears using a ring rolling mill. The ring-shaped blank is placed in a ring die with the required tooth form on the inner periphery and is expanded into it by the rolling action of a central mandrel. Sliding between the workpiece and die is minimized so that die wear is kept low. The process has been applied to mild steel, to a Ni-Cr alloy steel (En 24) and to softer metals. The products are shown to be structurally sound and of good form relative to the die profile. A simple analysis is presented relating roll force to the degree of tooth formation.

INTRODUCTION

Existing methods for the cold rolling of gears generally involve a large number of rotations or reciprocations between the die and the workpiece and the resulting interface sliding can produce lubrication and die wear problems. A method which avoids any reciprocating or rolling action might offer advantages in this respect. In this paper, a ring gear rolling process is described in which the workpiece is impressed into a gear-shaped die without any reversals of motion.

The principle is illustrated in figures 1 and 2. The annular blank fits closely inside a ring die with an internal profile corresponding to the required tooth form. Material is displaced radially outward to fill the die impressions by the rolling action of an internal mandrel, over many revolutions. The annular die is supported in a substantial housing and the rolling is performed between the main roll and mandrel of a ring rolling mill.

The required roll force would be expected to increase with radial thickness of the blank and the method is obviously not suitable for the direct forming of solid gears. However, it might prove to be economically attractive for producing composite gears using a lower quality hub material.

It appears that this cold gear rolling method has not been attempted so far, although it is believed that a somewhat similar hot forming process has been used for preforming large gears. For an account of some of the existing gear rolling and forging processes see references 1-6.

EQUIPMENT

The tooling is designed to operate in the experimental ring rolling mill developed in the Mechanical Engin-

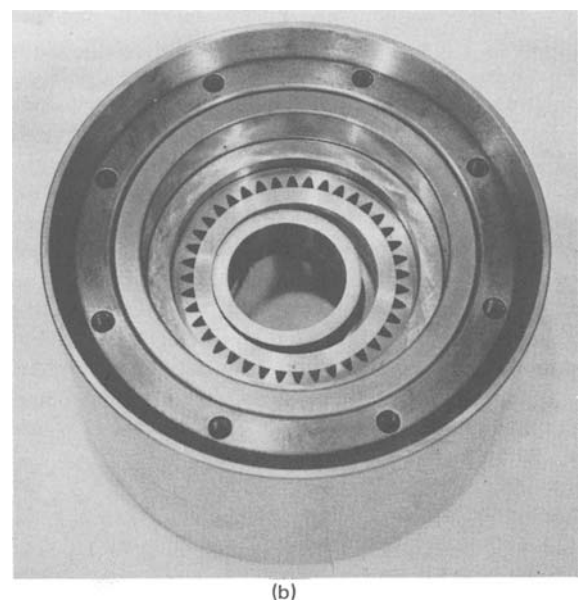
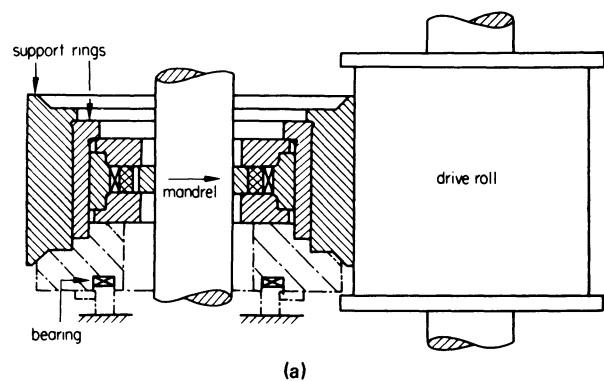


Figure 1. (a) Sectional view of tooling positioned between the rolls of a ring rolling mill. (b) Part-assembled tooling with workpiece and inner forming ring in position.

* University of Manchester Institute of Science and Technology (UMIST)

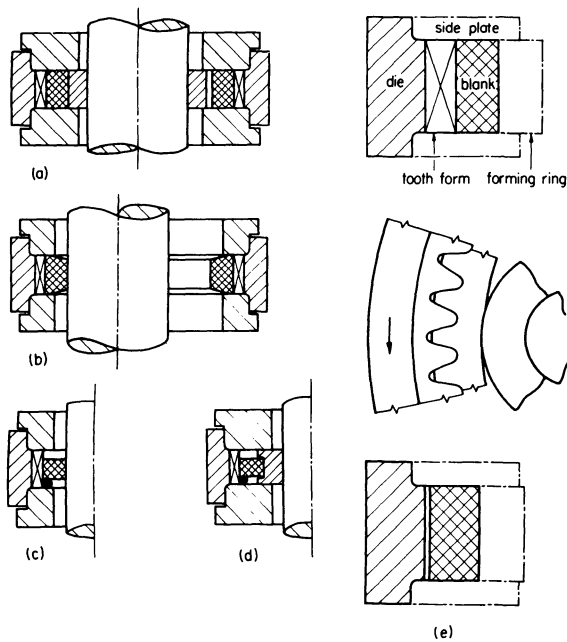


Figure 2. Workpiece shapes and tool arrangements. (a) Full-width workpiece and forming ring. (b) Tapered workpiece formed directly by mandrel. (c) Narrow workpiece supported centrally by a rubber O-ring. (d) Narrow workpiece rolled by a grooved forming ring.

engineering Department of UMIST and described in references 7 and 8. Sectional views and an illustration of the tooling appear in figures 1 and 2.

The ring die has the tooth form on the inner surface and is supported in relatively massive shrink rings of hardened Ni-Cr alloy steel (En 25). A 1.5 degree interface taper between the inner and outer support rings enables the inner ring and the die to be put into compression by axial clamping. The hoop compression in the die reduces tensile stresses in the die teeth during rolling and provides a fine adjustment to the die diameter, which might allow for die regrinding, for example. The amount of compression applied for the present work resulted in a 0.007 in reduction in the root circle diameter of the gear form.

The tapered interface between the rings is polished and well lubricated with molybdenum disulphide and, on release of the axial clamping force, the rings separate and relax, relieving the die compression and freeing the formed component so that little or no extraction force is required. This feature would be particularly useful for making helical spur gears.

The die assembly is rotated by the main roll of the ring mill and the roll pressure is applied between the main roll and the mandrel. A ball bearing thrust race in the lower clamping plate allows the assembly to rotate freely on the mill worktable.

Dies and workpieces

Results are presented for tests on two dies, of 5 and 8 d.p., 20° pressure angle and a pitch circle diameter of 5 in. The dies are in 1.5% C, 12% Cr, 0.8% Mo, 0.4% Va tool steel, shaped to involute tooth form in a Fellows' gear shaper before salt-bath hardening to Rc 53. The outer diameters and end faces of the dies were ground after hardening, but no further machining operation was applied to the teeth apart from the

application of corner radii to some of them, at positions corresponding to the roots of the formed gear teeth. The method of manufacture was relatively inexpensive and it was considered that, in the event of any slight distortion due to heat treatment, the accuracy of the process could be assessed to some useful degree by making direct comparison between formed teeth and the particular die cavities in which they were made.

Initial tests were carried out on lead and soft aluminium, using blanks of the full 1 in die width. Axial spread was prevented by side plates (see figure 2a), so that plane strain conditions were maintained, and a hardened steel forming ring fitted between the plates transferred the roll force to the workpiece. An alternative arrangement employed a tapered blank, shown in figure 2b, rolled directly by the mandrel, but the radial spread at the inner surface was generally too great and the material became trapped between the sideplates and the mandrel. The maximum roll force of the mill (18 tons) was insufficient for the rolling of 1 in wide steel gears and the blank widths used were 0.5 in for mild steel and 0.4 in for En 24. A rubber O-ring supported the blank in the centre of the die initially without influencing the spread during rolling, as illustrated in figure 2c.

Test conditions

The experiments were made with the speed of the main roll set at 30 rev/min, giving a die speed of 22 rev/min, assuming no slip between the roll and die assembly. The roll closure rate was controlled at 0.6 in/min and the relief valve on the roll force circuit was set to give an appropriate maximum load. The rolling cycle consisted of an initial period at a constant approach rate; progressively increasing force, and a final stage at a constant controlled force. The total rolling time was approximately one to two minutes. The die and ring blank were lubricated with molybdenum disulphide before each experiment so that, on unbolting the clamping flange and allowing the shrink rings to separate, the formed gear was easily removed.

RESULTS

Figure 3a illustrates typical products in mild steel and in En 24, with the end faces machined to the original blank width after rolling. The spread that occurs during the rolling of rings without end constraint is illustrated by the sectional profiles shown in figure 3b. Spread is negligible at the tips of the teeth and in fact there might be a slight contraction in axial length, probably associated with the development of the concave axial profile at the tooth tip. The spread increases inwards from the tips, being typically 40 to 50 per cent at the inner periphery when using a plain forming roll. The grooved forming ring, see figure 2d, limits the maximum spread to less than 40 per cent. Spread becomes greater with increase in radial blank thickness and the restraint on spreading increases with the depth of the groove in the roll. For maximum restraint the end flanges of the grooved roll would be just clear of the dieform at the final stage of rolling.

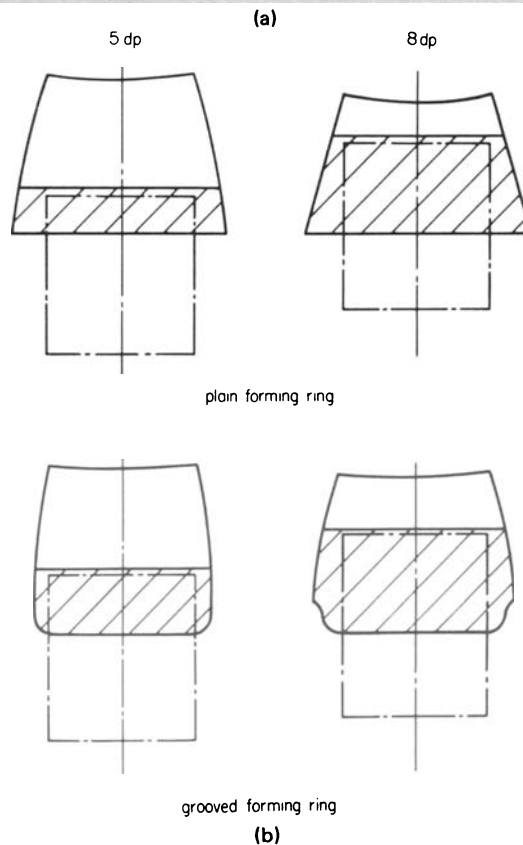
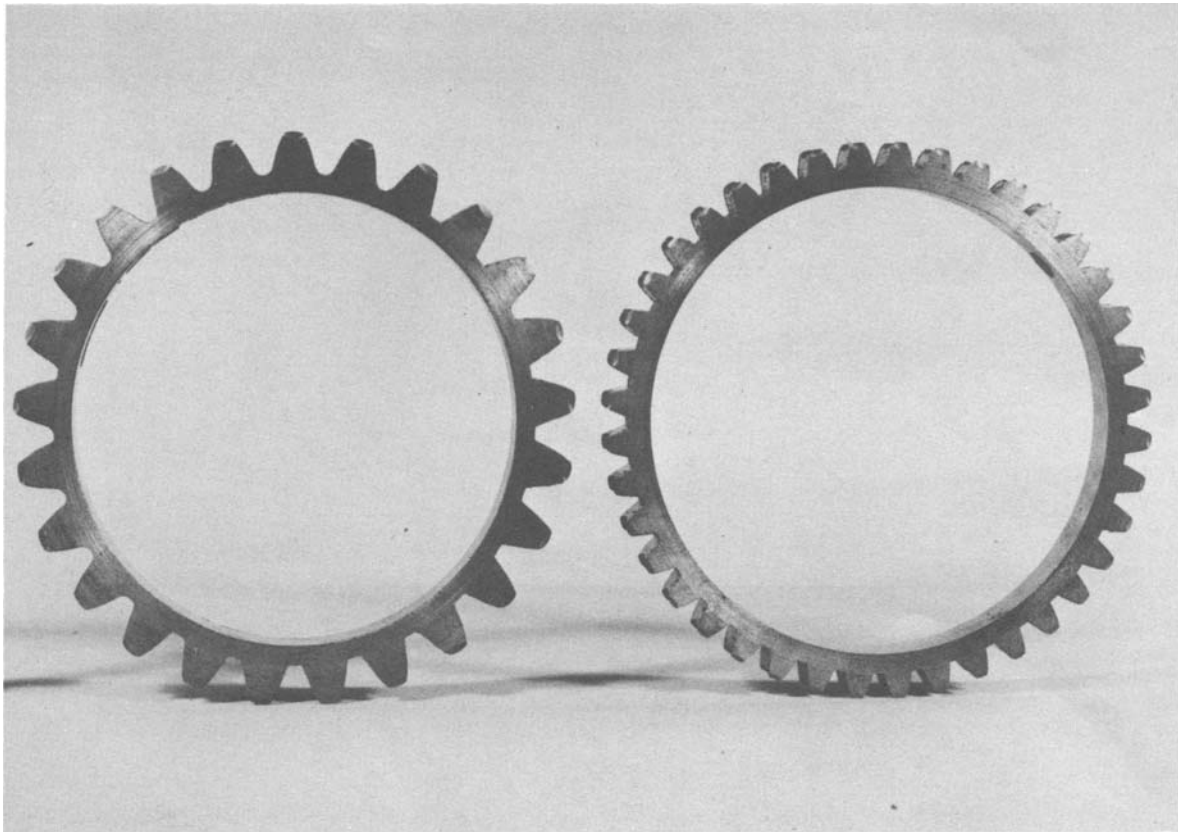


Figure 3. (a) Examples of cold-rolled steel gears (En 24 and mild steel). (b) Sectional profiles of steel ring gears compared with original blank shapes, showing axial spread.

Metal flow

Metal flow into the tooth cavities is shown by the macrographs in figure 4a and the pattern of surface markings (see figure 4b) is obtained by recording the distortion of radial lines scribed on the end face of a blank. The macrographs show the smooth flow of

metal into the tooth with the radiused root, avoiding the formation of surface cracks. Figure 4b indicates that surface extension along the tooth flanks is fairly uniform, supporting the observation that the lubricant is well distributed over the final toothform. The sharp-edged tooth root tends to produce a dead metal

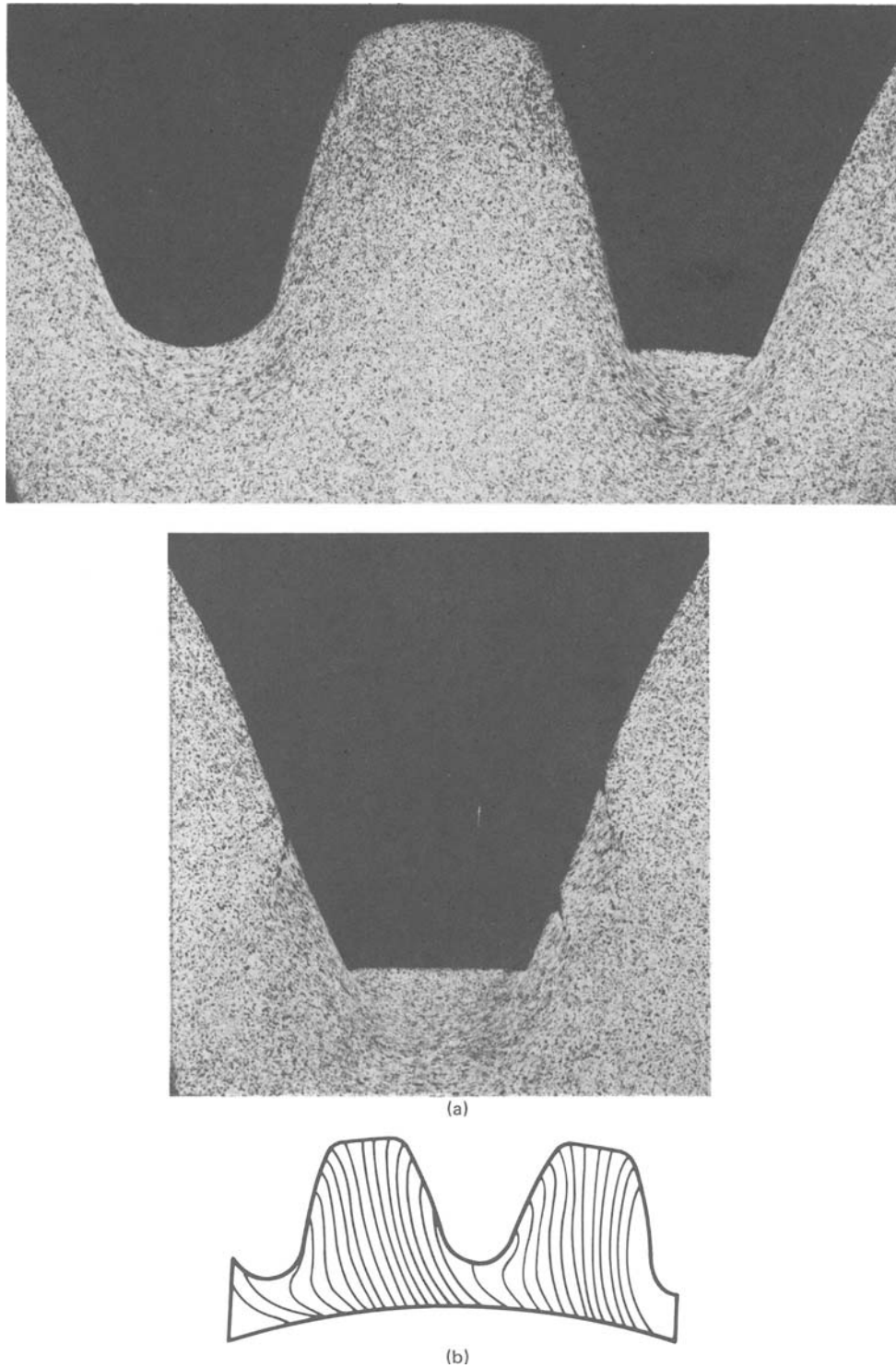


Figure 4. (a) Macrographs of gear teeth. Pair of adjacent 8 d.p. tooth spaces showing the smooth flow around the radiused root and the dead metal zone under the square root. The single 5 d.p. tooth space shows surface cracks which originate at sharp corners and slide along flanks. (b) Distortion of radial markings on the end face of a workpiece.

zone and cracking, which originates at a corner and moves along the flank as forming proceeds (see figure 4a). To reveal any concealed surface faults, gears were given an hydrochloric acid pickling treatment. The radiused teeth did not show any signs of surface defects.

Surface finish

The dies had not received any surface finishing treatment after hardening and it was considered that the most appropriate means of assessing surface finish

was to compare the formed surface with the corresponding portion of the die.

Surface finish measurements on a Taylor Hobson Talysurf gave readings on the flanks of formed teeth of 13 to 45 C.L.A. axially along the tooth and 28 to 42 C.L.A. radially, traversing towards the tip. For the corresponding surfaces of the die, the values were 36-50 and 30-37 C.L.A.

The surface finish of the formed product will be influenced by the type of lubricant and the relatively thick bonded coating of molybdenum disulphide used

for these tests might not give optimum results. It would be reasonable to conclude that the finish was determined mainly by the die and that an improvement in die finish would result in a smoother product, subject to the lubrication conditions being satisfactory.

Tooth form

The accuracy of the forming process was examined by comparing formed tooth profiles with the die cavities in which they were produced. Teeth were parted from the gears and trimmed in thickness so that, when fitted into the die space and illuminated from below under a toolmaker's microscope, any misfit would be revealed by light transmission through the gap. Thin shims of Melinex plastic film (down to 0.25×10^{-3} in thick) were interposed between the die and formed tooth to assist in assessing the widths of the lightgaps. It was estimated that the maximum gap width for steel gears of 5 d.p. and 8 d.p. was appreciably less than 10^{-3} in over the tooth flanks and probably in the region of 10^{-4} in. The thickness of the lubricant film appeared to be of this order and the misfit between die and workpiece could then be accounted for by lubricant thickness and elastic recovery strains. There was no evidence of any rocking of the formed tooth within the die cavity, for example.

Roundness

To investigate out-of-roundness and mean diameter variations in the formed gears, tooth-root diameter measurements were obtained and recorded against angular position, together with corresponding measurements from the dies. The final mean diameter of a gear was found to be within ± 0.0025 in of the die size and the out-of-roundness, i.e. the difference between minimum and maximum root diameter, varied from 0.0015 to 0.007 in for different gears. There was no obvious relationship between rolling conditions and out-of-roundness and mean size, although it appeared that a grooved forming ring resulted in an increased mean diameter, as compared with a plain roll.

It should be recalled that the die ring was compressed by about 0.007 in at the root diameter when assembled in the shrink rings, so that the final mean diameter of the product was greater than the die diameter when clamped by 0.007 in ± 0.0025 in.

Commercially, it would be necessary either to control the diameter more closely during forming or allow for correction of ovality and diameter by stretching over a central hub.

ROLL FORCE

Analysis for die cavity penetration at constant roll force

The final stage of forming occurs at constant roll force with a progressively decreasing penetration rate until flow into the die cavities finally ceases. There will then be only localized yielding under the roll.

A relationship can be derived giving the ultimate penetration under constant roll force by assuming simple plastic fields before and after termination of

flow. Equating the respective forces gives the transition condition. Plane strain deformation is assumed.

The flow model assumed during penetration consists of two plastic zones (see figure 5a), corresponding to the tooth cavities and the rim. The tooth form is approximated to a straight-sided convergent channel of included angle $2\alpha = 2\phi - \theta = 2\phi - 180/N$, the angle between tangents to the flanks at the pitch circle. ϕ is the pressure angle of the gear and N the number of teeth. Flow into this converging channel is considered as an extrusion process, with a pressure p_e at entry given by⁹

$$p_e = 2k \left(\frac{1 + \mu \cot \alpha}{\mu \cot \alpha} \right) \left[\left(\frac{y_0}{y} \right)^{\mu \cot \alpha} - 1 \right] \quad (1)$$

μ is the friction coefficient, α the channel semi-angle and y_0/y is the ratio of base width y_0 to current tip width y .

The annular rim of the gear is assumed to deform under the roll by simple radial shear. It is easily shown that an upper bound value for the force to cause radial shear without any change in thickness t is $2kwt$, where w is the axial length.

The total roll force F_1 required to cause shearing of the ring and extrusion into the tooth cavities over projected contact length l is then

$$F_1 = p_e lw + 2kwt$$

When penetration ceases, the roll causes local deformation only and it is assumed that the plastic field is then similar to that for indentation of a semi-infinite or laterally constrained block, with indentation pressure $p = 2k(1 + \pi/2)$ and force $F_2 = 2k(1 + \pi/2)wl$.

The transition point from flow to no-flow conditions gives

$$p_e lw + 2kwt = 2k(1 + \pi/2)wl = F \quad (2)$$

Combining (1) with (2) and rearranging gives

$$\frac{y_0}{y} = \left[1 + \frac{(1 + \pi/2)(1 - 2kwt/F)\mu \cot \alpha}{1 + \mu \cot \alpha} \right]^{1/\mu \cot \alpha}$$

This is represented in figure 5b for values of friction coefficient μ between 0 and 0.15 and $\alpha = 16.4^\circ$, corresponding to one of the experimental dies (20° pressure angle, 25 teeth). The degree of penetration into the tooth cavities $1 - y/y_0$ is recorded against $2kwt/F$. At $2kwt/F = 0$ the results relate to a workpiece with zero rim thickness, $t = 0$, so that the roll pressure is equal to the extrusion pressure p_e and at $2kwt/F = 1$ there is no penetration into the die cavities, the whole roll force being required to shear the rim. Hence $2kwt/F$ represents the proportion of roll force required to deform the rim and $(1 - 2kwt/F)$ is that available for the extrusion process. Also shown in figure 5b are values of $1 - y/y_0$ corresponding to a standard full depth tooth and a stub tooth.

This analysis is highly simplified in many respects other than those considered above. For instance, there is no consideration of the effects of rim and roll curvatures or of the roll contact length l in relation to the tooth cavity width y_0 . It appears that the present approach is not appropriate where the contact length

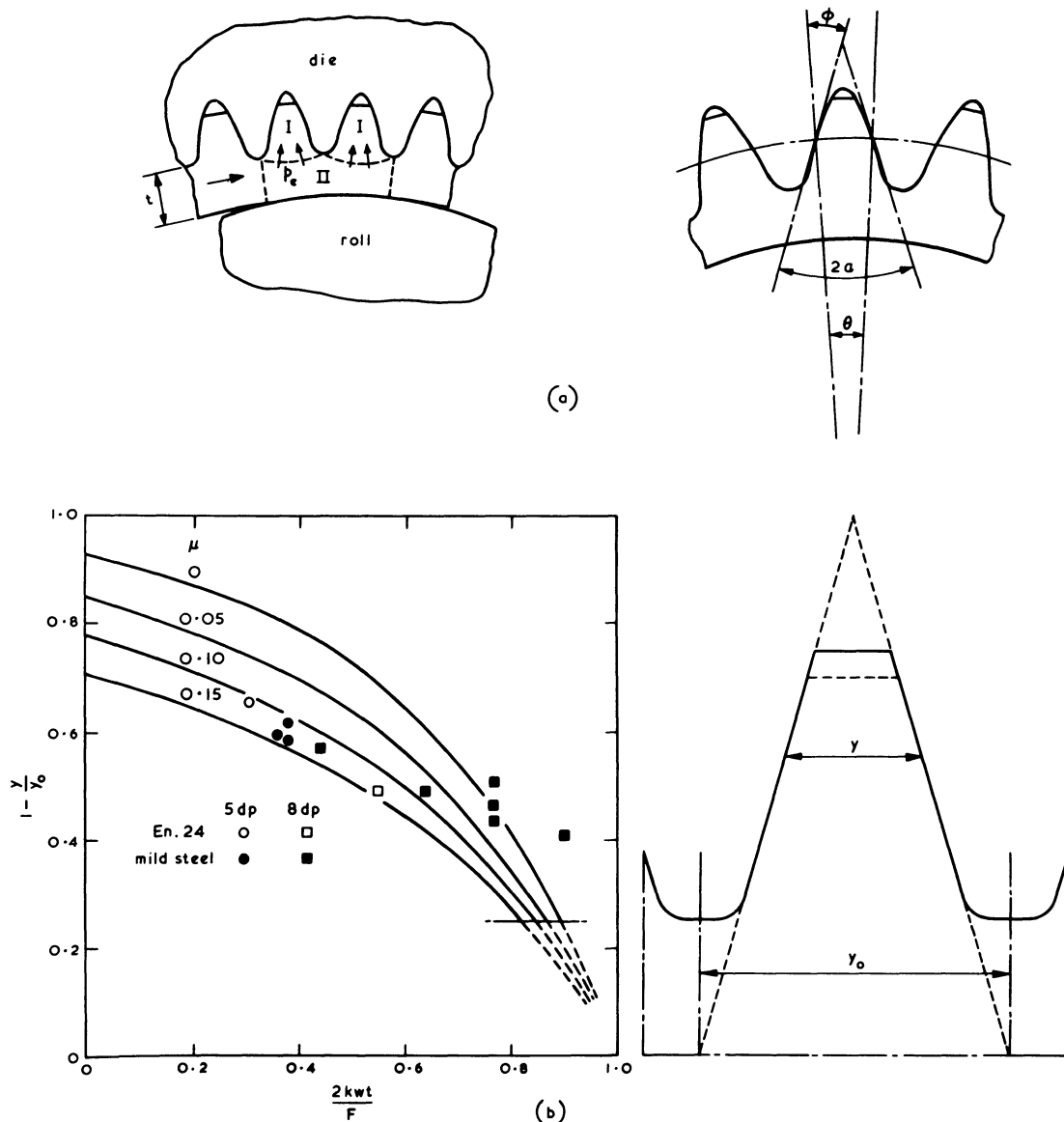


Figure 5. Relationship between roll force and tooth height. (a) Plastic fields assumed for analysis and simplified form of tooth cavity. (b) Theoretical curves and experimental results.

is appreciably less than the tooth cavity width. It is also evident that the solution is not applicable for $(1 - y/y_0) < 0.25$. However, the main region of interests is at low values of $2kwt/F$, where the model is likely to be more appropriate.

Comparison with experimental results

Results of force measurements taken during experiments using the 5 d.p. and 8 d.p. dies and rings in mild steel and En 24 are presented with the theoretical curves in figure 5b. Values of w and t used in the parameter $2kwt/F$ were taken from measurements on the finished gears, w being the mean width of the rim and t the thickness. The plane strain yield stress was determined from compression tests on small specimens machined from formed gear teeth, the stress-strain curve being extrapolated to zero strain to give the yield stress at the termination of the rolling process. The mild steel gave a uniaxial yield stress of 42 tonf/in² and the En 24 gave 50 tonf/in². The Mises yield criterion was used and no allowance was made for strain rate effects.

At values of $2kwt/F$ below 0.7, the experimental points lie in a band between the theoretical curves for $\mu = 0.1$ and 0.15. Divergence from the curves as $2kwt/F \rightarrow 1$ is due at least in part to the constant value assumed for the yield stress $2k$ for all stages of rolling. At small amounts of penetration lower values should be used which would displace the points to the left and give better correlation.

The experimental results indicate that the stub tooth is formed at $2kwt/F = 0.25$ or $F = 2k(4wt)$ and, by extrapolation, full depth is attained at about $F = 2k(10wt)$. Full depth teeth were made in lead and aluminium, but not in steel. To achieve full depth teeth in steel, with the limited roll force available (18 tonf), it would have been necessary to use deeper grooved rings than could be accommodated with the existing equipment, or otherwise reduce the face width of the dies and limit the spread with side plates. The theoretical curves suggest that there would be considerable benefit in reducing the friction coefficient by smoother die surfaces and an improved lubricating system.

CONCLUSIONS

The authors' experiments have confirmed in principle the feasibility of producing ring gears by cold rolling in an enclosing die. About 40 forming operations were made on steel gears and many more on softer metals without the occurrence of any faults or failures. Lubricant distribution appears to be favourable for giving low die wear and the prestressing of the die should reduce or eliminate tensile stresses between the die teeth. Bending stresses in the teeth might be significant, but it is likely that appreciable support against bending is provided by the workpiece.

The method of prestressing the die and releasing it to free the component after each operation proved to be effective for experimental purposes. A single application of lubricant between the polished surfaces of the shrink rings remained effective over most of the test programme on steel gears so that the rings were self-releasing on removal of the axial clamping force and no surface damage occurred at the interface.

For quantity production, it would be necessary to establish the reliability of this method over large numbers of assembly cycles and develop a more effective axial clamping system, possibly hydraulic. It is possible that release of the die is not necessary, at least in some cases, the gear being simply pressed out of the die after forming.

ACKNOWLEDGMENTS

The ring rolling research project described in this paper has been supported by S.R.C. grants which are gratefully acknowledged. Thanks are due to Dr S. J.

Hashmi, who performed some of the experimental work and to Mr D. A. Ryder and his colleagues for the metallurgical examinations on the formed gears.

REFERENCES

1. A. N. Norwicki. Comparison of gear production by rolling with traditional techniques. *Int. Mach. Tool Des. and Research Conf.*, Manchester, 1973.
2. A. R. O. Abdel-Rahman and T. A. Dean. Production considerations for the high speed forging of spur gear forms. *Int. Mach. Tool Des. and Research Conf.*, Manchester, 1973.
3. P. F. Egan, B. Jones, G. Connell. Production of gears by rolling. *Proc. Inst. Mech. Engrs.* (1970) Paper 20 184.
4. J. C. Crockett. Production of gears by cold rolling, *Machinery and Production Engineering*, (1971) 119 543.
5. E. P. De Garmo. *Materials and Processes in Manufacturing*, Collier-Macmillan Ltd., London, 1969.
6. Z. Marciniak. A rocking die technique for cold forming operations. *Machinery and Production Engineering*. (1970) 117, 792.
7. J. B. Hawkyard, E. Appleton and W. Johnson. An experimental wide ring rolling mill of novel design. *Int. Mach. Tool Des. and Research Conf.* Birmingham, 1972.
8. J. B. Hawkyard, W. Johnson, J. Kirkland and E. Appleton. Analyses for roll force and torque in ring rolling. *Int. J. Mech. Sci.* (1973) 15, 873.
9. G. W. Rowe. *An Introduction to the Principles of Metalworking*, E. Arnold, London, 1965.

AN EXPERIMENTAL INVESTIGATION INTO THE S-MILL ROLLING PROCESS

by

D. K. ROBERTSON* and D. H. SANSOME*

SUMMARY

The need to roll very thin, hard metal strip and sheet presented manufacturers with many difficulties over a period of years. The development of the Sendzimir and other small diameter multi-roll mills has solved many problems, but the capital costs involved are often high and some inherent process problems remain. Consequently, the S-mill was designed to overcome the difficulties associated with the high-speed production of thin, hard, wide strip of uniform gauge. This paper contains a description of the design and initial experimental investigation into this new rolling process and gives evidence of S-mill operation under optimum conditions.

INTRODUCTION

Demand for thinner, harder rolled metal strip and sheet increased during the last twenty years. This demand provided the stimulus for the development of more sophisticated rolling mills having work rolls of very small diameter in which each roll is supported by back-up rolls of larger diameter. Among mills developed in this category are the Sendzimir¹, Planetary², Pendulum³, C-B-S⁴ and S⁵-mills. The main advantages of using small diameter work rolls may be summarized as follows:

- (a) The elastic deformation of the roll surface forming the arc of contact, commonly described as roll flattening, is minimized resulting in a greater overall reduction. Ultimately, when deforming thin, hard metals, even conventional mills employing small diameter work rolls reach a point when no further reduction is possible^{6,7}.
- (b) For otherwise constant conditions, the smaller the roll diameter, the smaller the roll force. Thus, lighter and therefore less expensive mills can be specified.
- (c) Available evidence tends to support the claim that the smaller the diameter of the roll, the less the lateral and longitudinal gauge variation, provided that the small work roll is adequately supported.

The first patent, covering the S-mill, was filed by J. Franek and B. Grinstead of the Metal Box Company in 1965. Their mill, which followed from the development of the temper stretch leveller, was designed to reduce tinplate at high output compatible with subsequent processing while maintaining the

gauge both across the width and along the length. The S-mill is of novel construction as it does not depend on the conventional screw-down to achieve the required reduction and since it combines the stresses resulting from bending, tension and compression, it should be capable of rolling very thin strip.

An intensive programme of research into the mechanics of the process is at present under way in the Department of Production Engineering at the University of Aston and the first results obtained from that investigation are discussed in this paper.

THE PROCESS

S-mill rolling is a process which incorporates plastic bending under tension in conjunction with longitudinal tensions and roll pressure or nip to reduce the strip thickness. It relies upon the ratio between input and output strip speeds as a means for determining and controlling the reduction in place of the conventional screw-down at the roll gap. These features led to a mill concept that, apart from the C-B-S mill, represents a substantial departure from the methods employed at present in conventional rolling mills.

The work roll configuration of the S-mill is shown in figure 1. Work rolls W_1 and W_2 are free to float in the mill, but they are held in an equilibrium position by the three back-up rolls R_1 , R_2 and R_3 and the strip tensions. The only constraint on the work rolls is in the axial direction, where support arms prevent their movement out of the mill. Sufficient tensions are applied at entry and exit to prevent any slip between the driven rolls R_1 , R_2 and the strip and the reduction in thickness is thus achieved by the velocity

* Department of Production Engineering, University of Aston in Birmingham

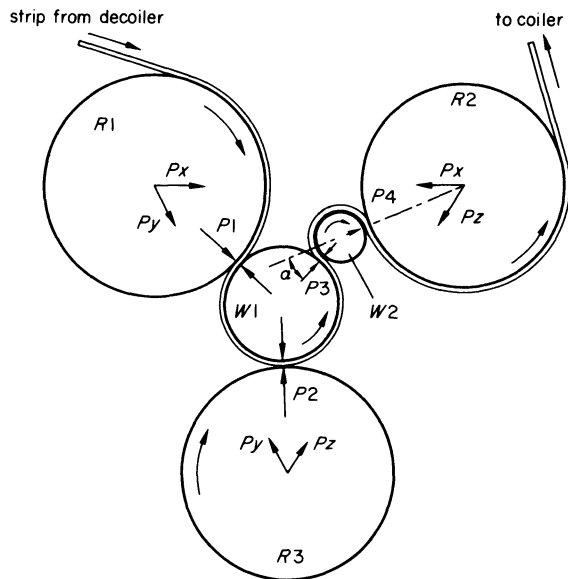


Figure 1. Diagrammatic representation of the S-mill work roll area. P_x , P_y , P_z —measured roll forces. P_1 , P_2 , P_3 , P_4 —actual roll separating forces.

ratio between the two driven rolls. For example, if the output driven roll R_2 rotates at twice the speed of the input driven roll R_1 , a 50 per cent reduction in strip thickness is achieved. This reduction is distributed between the four bite points, indicated in figure 1, and occurs as a consequence of the drawing, bending and stretching of the strip, forcing the work rolls into the gaps between the three back-up rolls.

The geometry of the mill can be altered by changing the distance between the three back-up rolls. This change is defined by the alpha angle (see figure 1). For large alpha angles, the gap between rolls R_2 and W_1 is small, while for small alpha angles, the

W_2 work roll is drawn right into this gap. Obviously, for stable operation of the mill, it is essential to have an alpha angle large enough to prevent the small work roll W_2 being pulled through by the high strip tensions.

EXPERIMENTAL EQUIPMENT

The mill on which the tests were conducted is illustrated in figure 2. It is a greatly modified version of the original mill on which the Metal Box Company carried out their early development work into S-mill rolling.

The distance between the three back-up rolls was made fully adjustable to take into account the variation in the alpha angle. This adjustment was achieved by means of variable length load cells, placed between the back-up rolls. The different diameters of the rolls in the work roll area are given in Table 1.

TABLE 1 Roll diameters in the work roll area

Roll	Diameter	
	(in)	(mm)
R_1	5.88	149.4
R_2	5.88	149.4
R_3	5.88	149.4
W_1	3.25	82.6
W_2	1.00	25.4

The tension necessary to prevent any slip between the driven rolls and the strip was supplied by wrapping the strip around the bridle rolls (see figure 2). The strip path through the mill is shown in figure 3. To help increase the tension, two pairs of pinch rolls were positioned in the bridle train, as shown. The drive to the input driven roll and the bridle train,

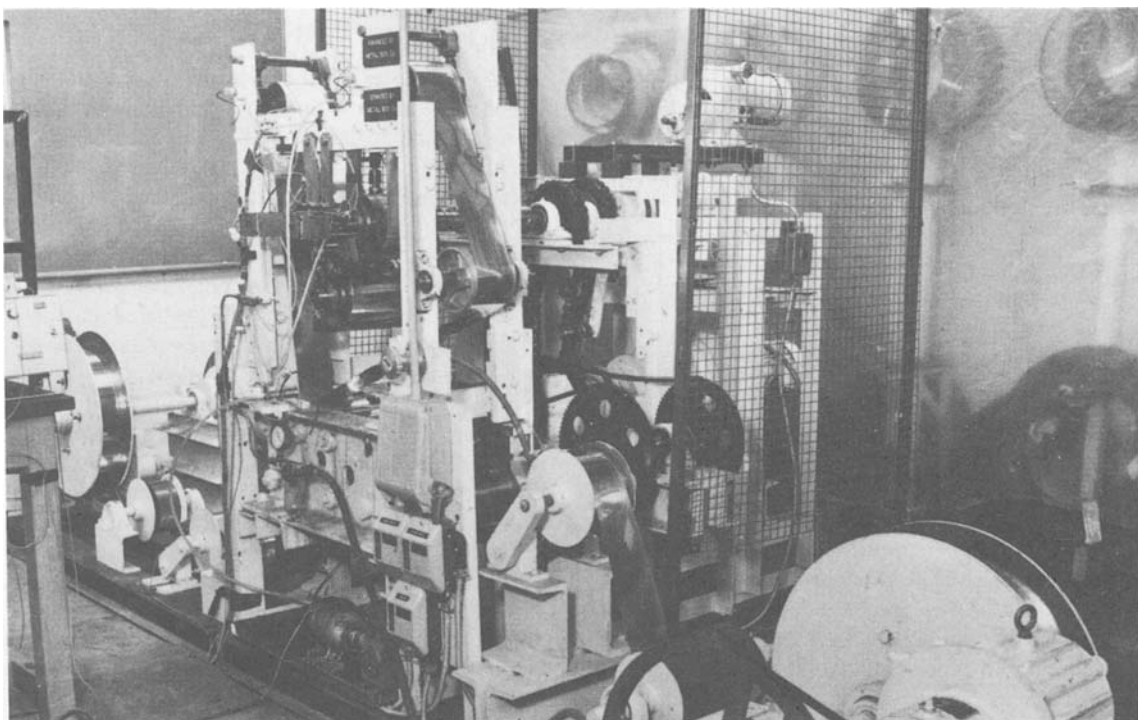


Figure 2. The experimental S-mill.

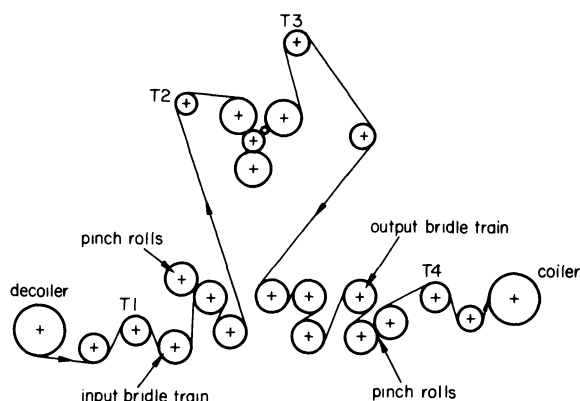


Figure 3. Strip path through the S-mill. T₁, T₂, T₃, T₄—tension measuring rolls.

was taken from a 2.24 kW (3 hp) motor through a 70:1 worm reduction gear box. The speed of the input roll and the bridle train were maintained constant at 10.8 rev/min. The drive to the output driven roll was from a 2.98 kW (4 hp) motor, through a Kopp variator and a 50:1 reduction gear box. Thus, the speed of the output train could be varied from 8 to 48 rev/min and the speed of the output bridle rolls automatically adjusted to suit any required reduction.

Early experimental work showed the need for sufficient back tension to preserve good tracking of the strip through the mill. This tension was built up by a disc brake on the decoiler. The amount of tension was adjusted by altering the pressure on the brake pad. Sufficient tension was supplied by the coiler simply to ensure the strip was closely wound on to it.

The composition of the lubricant used in all the rolling tests was: 2 per cent oleic acid, 2 per cent ammonia solution, 10 per cent white spirit and 86 per cent water. The oleic acid was incorporated because of the lubricating properties of long chain fatty acids and their ability to reduce surface tension and uniformly distribute the white spirit. The ammonia solution was used to counteract the corrosive properties of the acid. The white spirit acted as a lubricant and also as an emulsifying agent, while the water acted as a dilutant and coolant. The lubricant, supplied by a pump, was sprayed on to both sides of the strip as it entered the work roll area. To reduce friction and wear on the rolls and to improve the surface finish of the rolled material, all the rolls were specially chromium plated and polished to give a surface finish of the order of 0.0254 microns (1 micro inch).

The material rolled in all the experiments was 0.307 mm (0.0121 in) thick tinplate, 104.8 mm (4.125 in) wide. Details of the yield stress curve are as shown in Table 2.

TABLE 2 Details of the yield stress curve

% overall reduction	0	5	10	15	20	25	30	35
Yield stress tonf/in ²	21.0	23.5	25.6	27.2	28.5	29.9	31.3	32.6
Yield stress N/mm ²	324.6	363.2	395.7	420.4	440.5	462.1	483.8	503.9

INSTRUMENTATION

In each test, measurements were made of the roll separating forces, the roll torques and the strip tensions after leaving the decoiler and before and after leaving the work roll area. In order to preserve the floating work roll aspect of the S-mill design, it was necessary to measure the actual roll separating forces indirectly. This was accomplished by placing load cells between the three back-up rolls and from their signals the actual roll separating forces were calculated. These load cells had electrical resistance strain gauges bonded to their surfaces and connected into the normal Wheatstone bridge circuits. The torque meters were interposed between the rolls and their motors. They also had strain gauges bonded to their surfaces and connected into bridge circuits. Two types of tension meters were used. The first type was fully described by Afonja⁸. The second type used a load cell similar to those employed to measure the roll separating forces. The signals from all the strain gauge circuits were fed into two, twelve-channel ultraviolet recorders, whose paper speed was set at 254 mm/min (10 in/min). The load cells, torque meters and strip tension meters were all calibrated before and after the tests, but no detectable alteration in the calibration was observed.

TEST PROCEDURE

The instrumentation and recording equipment were allowed to warm up for at least one hour, and during this time a preparatory procedure was followed which included the lubrication of the mill bearings and the removal of the strip from around the tension meter rolls to permit any out of balance in the strain gauge circuits to be corrected. The distances between the three back-up rolls were accurately set and checked, using precision slip gauges. All rolls in the mill were thoroughly cleaned and degreased with Inhibisol and final circuit checks made. The strip was replaced around the tension meter rolls and the slack removed. These checks were repeated before each test.

The ingoing strip thickness was measured with a hand micrometer; the thickness, which was recorded, was determined by taking the average of ten tests, both from along the length and across the width of the strip. The final material thickness was also taken as the average of ten readings in a similar manner.

The pump was started and circulation of the lubricant was allowed to become established throughout the system. The mill was started and the speed of the output roll selected to give the required reduction. When the mill had reached steady-state conditions, the event markers on the ultraviolet

recorders were activated and the outgoing strip marked to indicate the roll speed. This procedure was repeated for increasing reductions up to 35 per cent. The tests were repeated for various positions of the back-up rolls, that is, for different alpha angles.

Initially, the reduction of the strip, at the four bite points, was determined by a hand micrometer. However, occasionally this led to large errors, due to the original strip thickness being only 0.307 mm (0.0121 in); hence the following method was devised to eliminate, or reduce, the error to a minimum. Marks were made on the ingoing strip, approximately 6.4 mm (0.25 in) apart, and the distances between them were measured accurately using a travelling micrometer microscope. The marked strip was rolled into the work roll area and the mill was stopped. The front and back tensions were removed as quickly as possible and the rolls taken out of the mill. The distance between the marks was measured using the travelling micrometer microscope and the elongations recorded.

DISCUSSION OF RESULTS

From a consideration of the forces on the small work roll, shown in figure 4, it can be seen that the

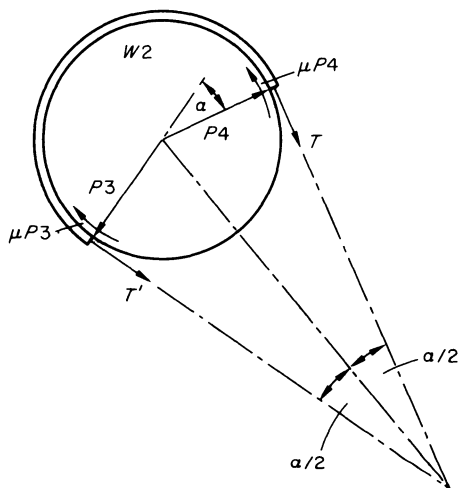


Figure 4. Diagrammatic representation of the forces on the small work roll, W2.

following relationship exists between the strip tension T , roll force P , alpha angle α and the coefficient of friction μ .

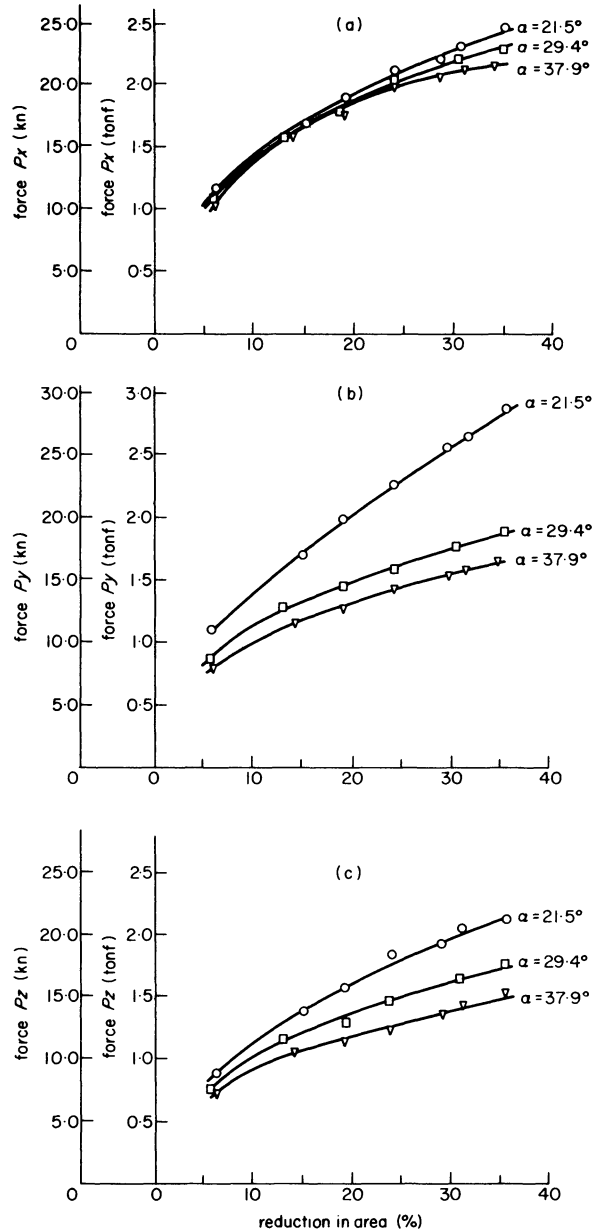
$$\dot{T} = P_4 \left(\tan \frac{\alpha}{2} + \mu \right) \quad (1)$$

and

$$P_3 = P_4 \quad (2)$$

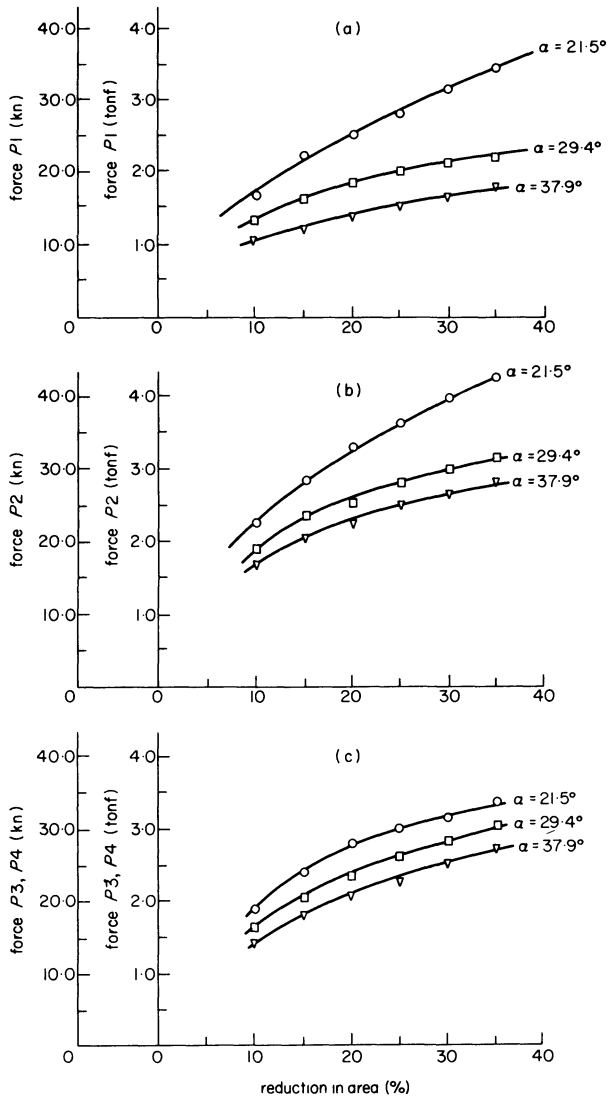
All the results obtained are shown in graphical form in figures 5 to 12. They give the variation of the measured forces, actual roll separating forces, roll torques and strip tensions against reduction and alpha angle.

From an analysis of the graphs, the following conclusions are drawn.

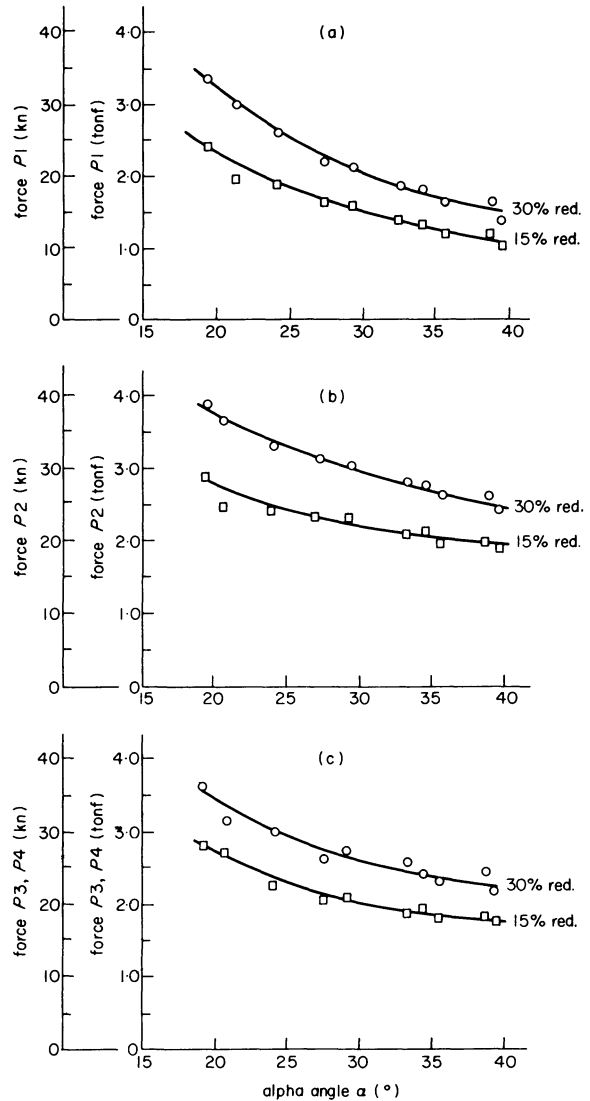


Figures 5(a), (b), (c). Variation of measured forces P_x , P_y , P_z with percentage reduction for set alpha angles.

Figures 5(a), (b) and (c) show how the three measured forces P_x , P_y and P_z increase with reduction and that, at a given reduction of area the forces are least for the greatest alpha angle. In general, P_x is the greatest and P_z the least force. These experimental measured forces can be converted to give the actual roll separating forces at bites 1, 2, 3 and 4 and, from a consideration of the forces on the small work roll, it can be shown that P_3 must at all times be equal to P_4 . Figures 6(a), (b) and (c) show the variation of actual roll separating forces against reduction of area for various alpha angles. Again, it will be seen that, for a given reduction, the roll force is greater for small alpha angles. As confirmation of these conclusions, figures 7(a), (b) and (c) show, for 15 and 30 per cent overall reductions, how the roll forces vary with alpha angle. These results also show that the roll force at bite 2 is always the greatest, while the force at bite 1 is always the least, for all reductions and alpha angles. This is not



Figures 6(a), (b), (c). Variation of actual roll separating forces P_1 , P_2 , P_3 , P_4 with percentage reduction for set alpha angles.



Figures 7(a), (b), (c). Variation of actual roll separating forces P_1 , P_2 , P_3 , P_4 with alpha angle for set reductions.

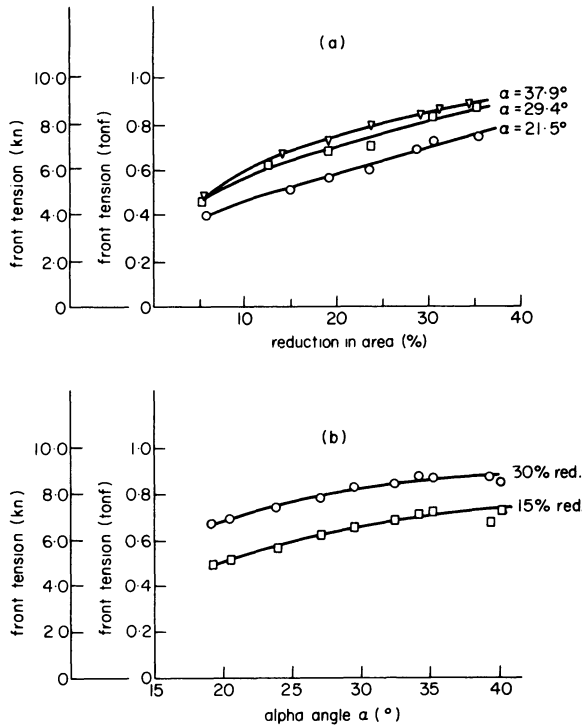
surprising when the equilibria of forces on the mill are considered. However, when the reduction of the strip at the different points is considered (see Table 3) it can be seen that the proportion of the total reduction at bite 2 is, on average, only approximately 6 per cent. The reasons for the greatest force producing the least reduction are probably twofold. Firstly, the strip tension is likely to be at its lowest at this bite; hence, from a consideration of the yield criterion, the roll force is greatest. Secondly, from observations of bite 2, the

length of the arc of contact would appear greater than expected for the reduction achieved.

While the percentage reduction at each bite varies slightly with overall reduction, it will be realized that for bites 1 and 2 the appearance of variation is probably due to small experimental errors. However, this variation is quite marked for bites 3 and 4. It will be seen that as the overall reduction increased the proportion of the reduction at bite 3 decreases and correspondingly increases at bite 4. The reason for this difference, it will be realized, is due to the strip

TABLE 3 Strip reduction at different bite points

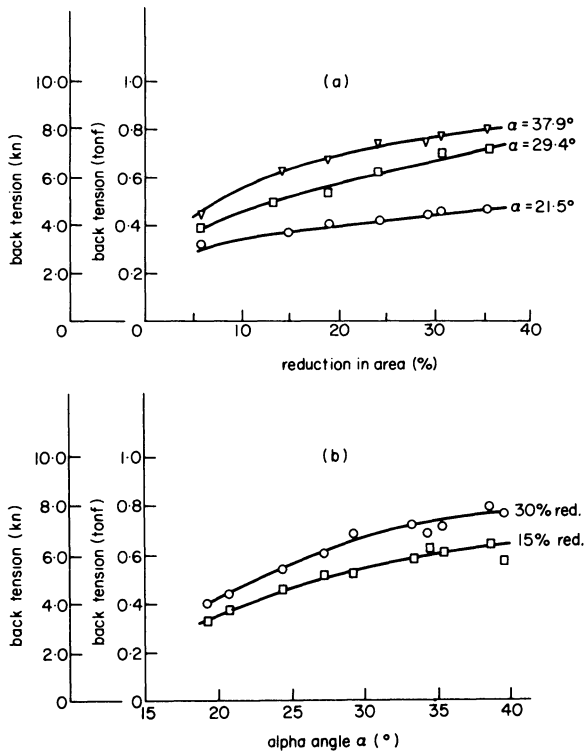
Bite No.	Total reduction of area					
	12.5%	17.3%	22.25%	26.9%	31.8%	34.75%
	Proportion of redn: at bite %	Proportion of redn: at bite %	Proportion of redn: at bite %	Proportion of redn: at bite %	Proportion of redn: at bite %	Proportion of redn: at bite %
1	13.6	11.6	12.1	13.0	11.8	10.7
2	6.4	7.5	6.9	5.6	6.3	5.7
3	46.0	44.4	47.2	44.6	36.9	38.6
4	34.0	36.5	33.8	36.8	45.0	45.0



Figures 8(a), (b). Variation of front tension with percentage reduction and alpha angle.

tensions increasing with reduction, while the roll forces P_3 and P_4 remain equal. Consequently, the proportion of the reduction at bites 3 and 4 must alter progressively to keep the small work roll in equilibrium.

Figures 8(a) and 8(b) show the variation of front tension with reduction and alpha angle. The front



Figures 9(a), (b). Variation of back tension with percentage reduction and alpha angle.

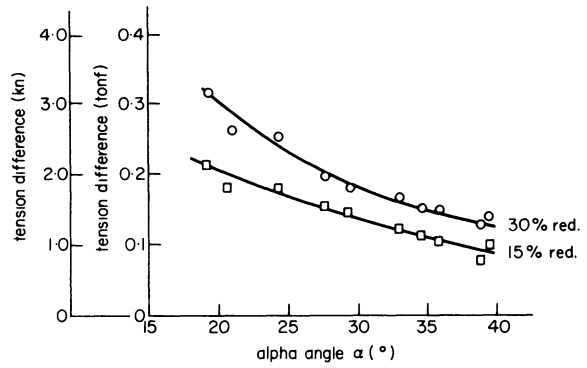


Figure 10. Variation of tension difference with alpha angle for set reductions.

tension rises rapidly for reductions up to approximately 15 per cent, after which the tension tends to level off, with only a small increase occurring as the reduction increases. The variation in front tension with the alpha angle is quite small. Therefore, from Equation (1), since the front tension varies only slightly with the alpha angle, the roll force P_4 must vary with the alpha angle to preserve the equilibrium.

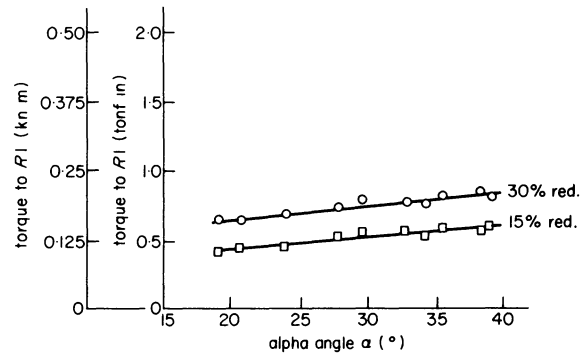


Figure 11. Variation of input torque with alpha angle for set reductions.

Figures 9(a) and 9(b) show the variation of back tension with reduction of area and alpha angle. The back tension is seen to increase more with the alpha angle than the front tension. For an explanation of this observation, consider figure 10, in which is plotted the variation of tension difference ($T_{entry} - T_{exit}$) with alpha angle. The smaller the alpha angle, the greater the tension difference, that is, the greater the work input. It can be seen from figures 11 and 12

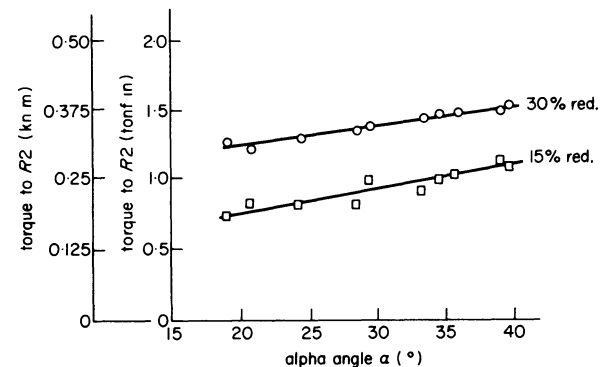


Figure 12. Variation of output torque with alpha angle for set reductions.

that both the input and the output torques to rolls R_1 and R_2 , respectively, increase with increasing alpha angle. Experimental evidence shows that the input roll acts as a brake and therefore removes work from the mill, that is, it is driven by the strip tension. From this it can be deduced that the total power input necessary to achieve a given reduction is relatively constant for all alpha angles, that is, the strip tension difference decreases with increasing alpha angle while the torque difference increases.

CONCLUSIONS

Since the S-mill process, essentially, is one in which strip is drawn through rotatable dies, it may be appreciated that the limiting condition is the front tension. Therefore, it is very important to select an arrangement of the variables and mill geometry which produce the minimum tensions to do the work and yet are adequate to prevent any slip between the rolls and the strip. The experimental evidence shows that to keep the tensions as low as possible and hence maximize the reduction achievable, the alpha angle must be as small as possible.

Further work on the mill is at present taking place at the University of Aston in order to obtain a full understanding of the mechanics of this process. A comprehensive theory and the experimental results will be published later⁹.

ACKNOWLEDGMENTS

The authors wish to thank most sincerely Mr M. Richard of the Metal Box Company Limited for his

helpful comments and Mr G. P. Clay and The Metal Box Company Limited for their financial and practical help with the project. Also, their thanks are due to Mr J. Hiron, Mr H. Pratt, Mr M. Cox and Mr J. Smith for their assistance with the manufacture and instrumentation of the experimental mill.

REFERENCES

1. M. G. Sendzimir and L. Zdanowicz. A mill for cold rolling metals to close tolerances. *Iron Steel Eng.* (November, 1956) 65.
2. T. Sendzimir. The Planetary mill and its uses. *Iron Steel Eng.* (January, 1958) 95.
3. K. Saxl. The Pendulum mill—a new method of rolling metals. *Proc. Inst. Mech. Eng.* (1964–65) 179, 453.
4. L. F. Coffin. Status of contact-bend-stretch rolling. *J. of Metals* (August, 1967) 14.
5. J. T. Franek and B. Grinstead. *British Patent Specification* 1,087,097 (1965).
6. H. Ford and J. Alexander. Rolling hard materials to thin gauges. *J. Inst. of Metals.* (1959–60) 88, 193.
7. H. Kuhn and A. Weinstein. The minimum gauge problem in thin strip rolling. *Trans. A.S.M.E. J. of Lub. Tech.* (July, 1971) 331.
8. A. A. Afonja. The sandwich rolling of thin, hard strip. *Ph.D. thesis, University of Aston in Birmingham*, 1969.
9. D. K. Robertson. The mechanics of drawing tinplate through rotatable dies. *Ph.D. thesis, University of Aston in Birmingham*. To be published.

DRAWING, TUBE AND SHEET FORMING

A PRELIMINARY INVESTIGATION OF THE DIE-LESS DRAWING OF TITANIUM AND SOME STEELS

by

J. M. ALEXANDER* and T. W. TURNER*

SUMMARY

Results are described of an experimental investigation of the die-less drawing of titanium alloy, mild, chrome and stainless steel bars of round and rectangular cross section. The main factors required to achieve constant cross section of drawn material were investigated.

The finite element method and other techniques of theoretical prediction of the process were used and are compared with the experimental results.

NOTATION

A	Cross-sectional area
A_I	Cross-sectional area of undrawn bar
A_F	Cross-sectional area of drawn bar
A_i	Cross-sectional area of i^{th} section
B	Constant
$[B]$	Strain displacement matrix
$[D]$	Stress strain matrix
E	Young's modulus
$\{F\}$	Force matrix
K, K'', K_2	Constant
L	Load
Q	Activation energy
Q'	Quantity of heat
R	Gas constant
$R' = V_D/V_C$	Ratio of draw and coil velocities
T	Temperature
T^*, T_i^*	Constant
V	Velocity
V_C	Velocity of the induction coil
V_D	Velocity of the draw head
$V_C + V_D$	Velocity of the draw
e	Engineering strain
e	Superscript to denote single element
f, g, h	Function
$[k]$	Element stiffness matrix
k'	Conductivity
m	Strain rate sensitivity index
n	Constant
r	Radius
u', v'	Components of displacement
z	Direction of specimen axis
α, β, γ	Constant
δ	Displacement
ϵ	Natural strain

$\dot{\epsilon}$	Strain rate
ν	Poisson's ratio
σ, τ	Stress
$\bar{\sigma}$	Equivalent stress
Δ	Area of element

INTRODUCTION

The phenomenon of superplasticity, comprehensively reviewed by Johnson¹, is exhibited by many metals and alloys, both ferrous and non-ferrous, and is either a fine-grain-size or a transformation effect. In recent years a great amount of interest has developed in the phenomenon because of the possibility of commercial application. It is from this background that the idea of die-less drawing emerged with experimentation by Johnson² and Weiss and Kot³.

Die-less drawing is a process in which a longitudinal tensile force is applied to a bar, a small length of which is heated. After a neck has formed the heat source is moved along the bar, the neck cooled and a reduced cross-section drawn from the bar. Only that part of the bar which is being deformed is heated. A schematic diagram of the process is shown in figure 1.

EXPERIMENTAL APPROACH

The foundations of any metal working process are laid by the experimental evaluation of the variables which control the material behaviour. A four-dimensional relationship between flow stress, strain, strain rate, and temperature is the requirement for investigating die-less drawing. Unfortunately the speeds used so far in the process give strain rates in the region 10^{-1} to 10^{-3} s^{-1} . These are very slow

* Department of Mechanical Engineering, Imperial College of Science and Technology

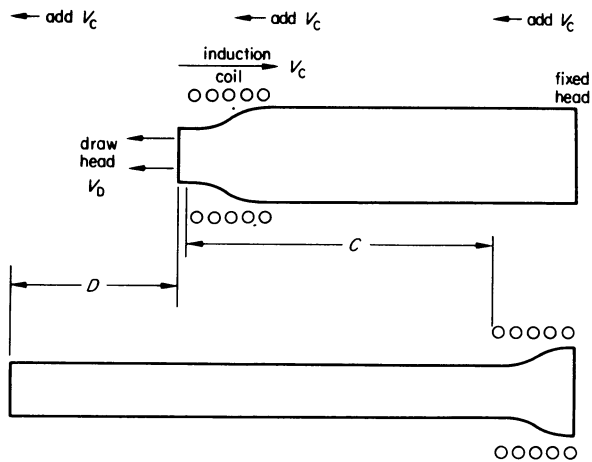


Figure 1. Schematic diagram of die-less drawing.

compared with conventional forming methods such as rolling and extrusion but very fast when compared with creep. There is, therefore, a distinct lack of data, although the position is changing as many workers investigate superplastic phenomena.

Weiss and Kot³, wishing to take advantage of the considerable increase in ductility during phase transformation, set up an experiment for die-less drawing. A lathe was modified to give a draw head driven by a spindle speed using a threaded rod. Steel specimens were heated using an induction coil, which gave temperatures of 705°C and 815°C. The plain carbon steel, AISI 1018, was reduced by up to 50% in area at draw head speeds in the range 1.2 cm/min (0.5 in/min) to 10.9 cm/min (4.3 in/min). Specimen diameters were 0.635 cm (0.250 in) diameter. Commercially pure titanium of the same diameter was also drawn at 1.2 cm/min by up to 50 per cent reduction in area at temperatures of the order of 900°C, and Ti-6Al-4V, 0.95 cm (0.375 in) diameter, drawn at 2.5 cm/min (1.0 in/min) by up to 54.5 per cent reduction in area at a temperature of 930°C. They concluded that die-less drawing is feasible for steel, titanium, and Ti-6Al-4V and, noting that their calculations of strain rate for AISI 1018 are two or three orders of magnitude higher than the strain rates necessary to cause transformation plasticity, they consider the die-less process might be used on materials which do not undergo a transformation.

Johnson² carried out similar experiments on a range of materials using a modified lathe and concludes that the process has commercial possibilities.

The present experimental programme was started with the intention of demonstrating the viability of the process for bars of the order of 5.0 cm (2 in) diameter and investigating the following parameters:

Stress	Strain rate sensitivity index
Strain	Material
Strain rate	Induction coil shape
Temperature	Temperature profile
	Specimen cross-section

During the testing it was found that other parameters played an important part in stable drawing: they are:

Grain size	Position of zone in coil
------------	--------------------------

Rate of heat extraction	Mass of moving head
Induction coil length	Heating technique

PROCEDURE

Testing each specimen was carried out as follows:

- (1) The bar was set up in the rig with the heating coil positioned a distance of 2.5 cm (1 in) from the start end. Reference to figure 2 shows the position of each of the components.
- (2) The setting of the gear box was adjusted to give the appropriate speed ratio, R' .
- (3) Having balanced the Wheatstone Bridge circuit of the load cell, the cell was calibrated using the steelyard of the Buckton testing machine set at a load close to the expected drawing load.
- (4) The power to the induction coil was switched on and the operating current adjusted to give the chosen maximum temperature for the draw.
- (5) The chart recorder was switched on and the draw head set in motion by opening the hydraulic drive valve.
- (6) When the neck represented about 30 per cent reduction in area, the coil position was adjusted to place the developing deformation zone correctly. This adjustment is unnecessary for superplastic materials like titanium.
- (7) Draw speed was set at the desired value by adjusting the needle valve in the by-pass line of the hydraulic drive to the drawing head.
- (8) The air pressure supplying the cooling coil was adjusted. NB. Several starting procedures were tried such that adjustments (6), (7) and (8) had to be made several times.
- (9) The successful draws were stopped and the load relieved immediately.
- (10) After cooling, the diameters of the deformation zone were measured every 0.127 cm (0.05 in) or 0.254 cm (0.10 in) along the axis. A travelling microscope was used.

In the tests the undeformed bar is quite stationary and the heating coil moves along it in the opposite direction to the drawn material. It is convenient to view this by fixing the frame of reference so that it moves with the coil. To do this it is necessary to add a velocity of $-V_C$ to all parts of the system. This effectively fixes the coil in space so that material is apparently fed into the coil at a velocity V_C and drawn from the coil in the same direction at a velocity $V_C + V_D$ (see figure 1). A simple relation exists between the velocity ratio, R' , and the percentage reduction in area.

$$\text{Volume of bar moving into coil} = A_I V_C,$$

$$\text{Volume of bar drawn from coil} = A_F (V_C + V_D)$$

$$\text{For no volume change } \frac{A_F}{A_I} = \frac{V_C}{V_C + V_D} = \frac{1}{1 + R'}$$

$$\begin{aligned} \text{Percentage reduction in area} &= 100 \left(1 - \frac{1}{1 + R'} \right) \\ &= \frac{100 R'}{1 + R'} \end{aligned}$$

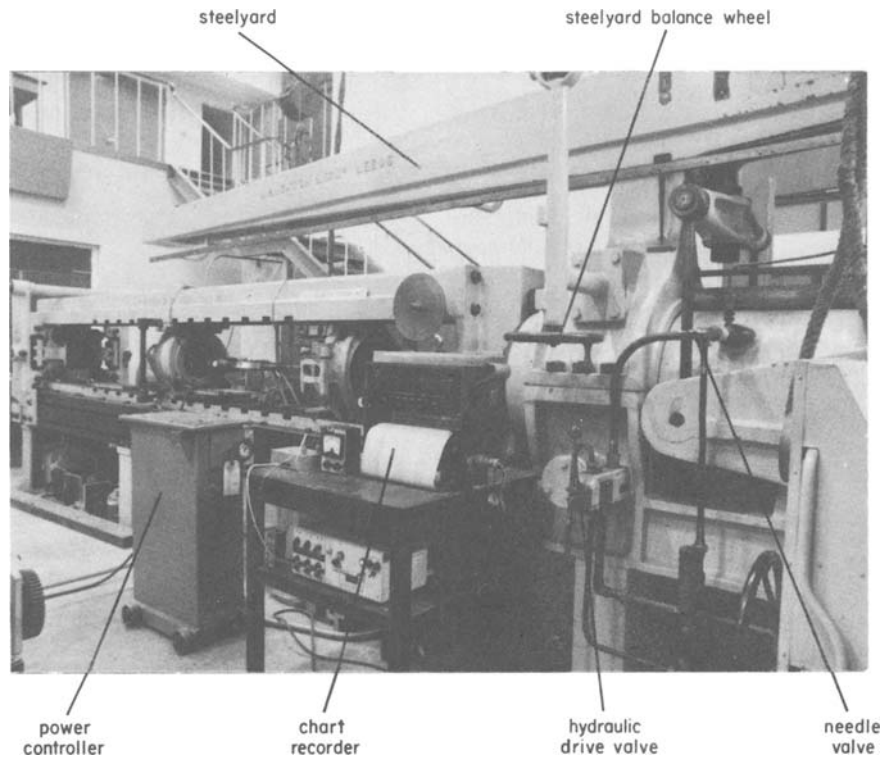
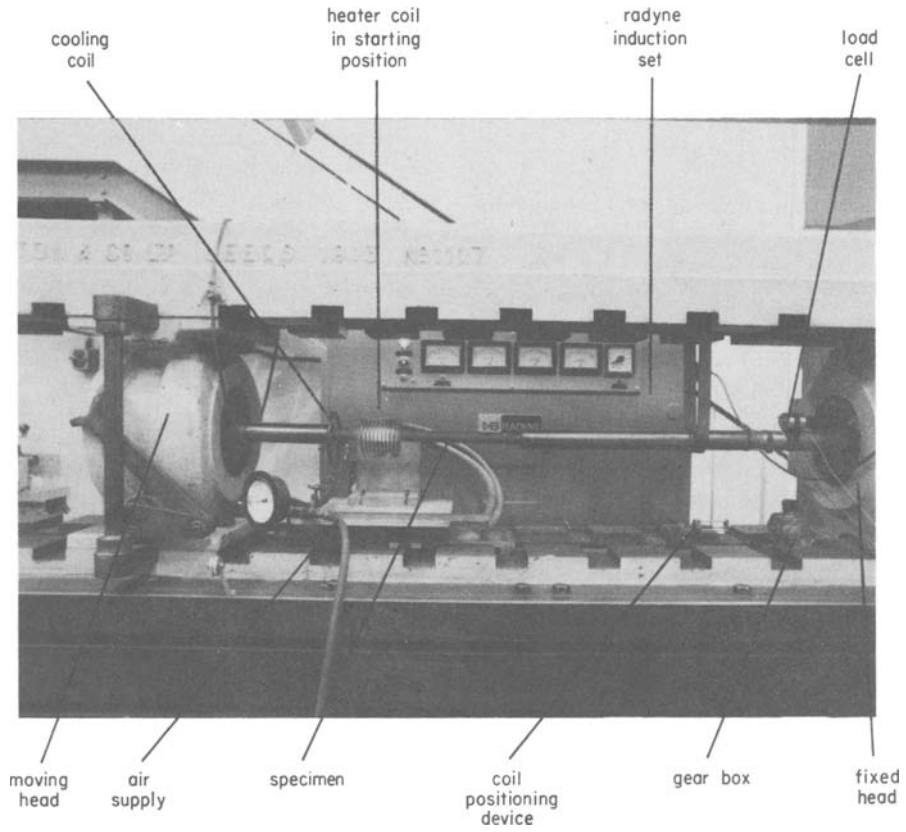


Figure 2. General arrangement.

Many of the experiments were carried out with $V_C = V_D$, i.e. $R' = 1$, giving a reduction in area of 50 per cent.

MATERIALS

The materials tested in the experimental programme were as follows:

(1) Mild steel supplied by BISRA

7.6 cm (3 in) diameter
5.0 cm (2 in) diameter
7.6 cm x 3.8 cm (3 in x 1.5 in)
3.8 cm x 3.8 cm (1.5 in x 1.5 in)

(2) Niobium mild steel supplied by BISRA

0.07 per cent C, 1.4 per cent Mn, 0.03 per cent Nb, Remainder Fe.

2.5 cm (1 in) diameter

Casts 2807 and 2808 were soaked at 1250°C for 1 hour then rolled under controlled conditions to a finishing temperature of 825°C and then air cooled.

Cast 2809 was finish rolled at 950°C and quenched directly into water.

(3) Stainless steel supplied by BISRA

EN 58 J: 0.12 per cent C, 0.20 per cent Si, 2.00 per cent Mn, 8.0/12.0 per cent Ni, 17.0/20.0 per cent Cr, 2.50/3.50 per cent Mo, 0.04 per cent S, 0.045 per cent P, Remainder Fe.

5.0 cm (2 in) diameter
3.8 cm (1.5 in) diameter
3.2 cm x 3.2 cm (1.25 x 1.25 in)

(4) Chrome steel

0.19 per cent C, 0.25 per cent Si, 0.51 per cent Mn, 3.20 per cent Cr, 0.0055 per cent S, 0.012 per cent P, Remainder Fe.

5.7 cm (2.25 in) diameter

(5) Chrome steel supplied by Kayser Ellison

KEA 162: 1.0 per cent C, 0.65 per cent Mn, 5.25 per cent Cr, 1.0 per cent Mo, 0.2 per cent V, Remainder Fe.

5.0 cm (2 in) diameter
3.8 cm (1.5 in) diameter
5.0 cm x 3.8 cm (2 in x 1.5 in)

(6) Titanium alloy supplied by Rolls Royce (1971) Ltd.

6 per cent Al, 4 per cent V, Remainder Ti.
2.5 cm (1 in) diameter

EXPERIMENTAL APPARATUS

Buckton 100 ton testing machine

The machine is massively built on a horizontal bed 9 metres (28 feet) long, having a capacity for loads up to 1000 kN (100 tons), and a stroke of over 60 cm (2 feet). The general arrangement of the testing machine is shown in the photographs, figure 2.

Radyne induction set

The set, Model RD 150, obtained from Radyne Limited, is a medium frequency valve generator operating between 2 and 5 kHz with a maximum continuously rated output of 15 kW.

Six working coils were used with eight to twelve turns ranging from 5 cm (2 in) to 10 cm (4 in) internal diameter. The first coil to be used was supplied by Radyne Limited and consists of twelve turns of square section copper tubing. The other coils were made from 0.95 cm ($\frac{3}{8}$ in) or 0.80 cm ($\frac{5}{16}$ in) diameter copper tubing.

Coil carriage

The coil carriage is an arrangement to enable the heating coil and air cooling coil to move in the opposite direction to the draw head.

The sliding member of the Buckton is arranged to pull a wire wound round the pulley of a variable speed gear box. A wire wound round the output pulley in the opposite direction to the winding on the input is attached to the coil carriage. Changing the input and output pulley diameter ratio extends the range of the velocity ratio R' .

The coil is mounted on the carriage using insulating board. The board also carries a cooling coil positioned about 7.5 cm (3 in) from the trailing end of the induction coil. Twelve holes 0.10 cm (0.040 in) diameter drilled on the inside of a single turn of copper tubing 10 cm (4 in) diameter allow air to blow on to the neck of the specimen.

Setting up the draw

Figure 3 shows how the specimen is supported in the rig. Most of the specimens drawn were 46 cm (18 in) long, but different lengths can be accommodated by keying the moving head at different positions on the Buckton sliding member.

The starting position of the coil is indicated in figure 3, and after the necking phase, the coil position can be adjusted by a screw thread wire shortening device in the wire connecting the draw head to the gear box.

INSTRUMENTATION

Load measurement

The load cell used consists of a standard Wheatstone Bridge arrangement. The load cell was situated at the remote end of the extension bar, i.e. 70 cm (27 in) from the heated work piece, so that it always operated at room temperature.

Displacement measurement

Two displacement transducers were incorporated in the rig. One is a simple visual device which is essential for observation at slow speeds and a second is designed to give a displacement trace on the chart recorder, thus enabling velocities to be read.

Temperature measurement

An optical pyrometer was used to measure the temperature for all of the tests. Maximum temperatures are easily taken in this way. Observation of the

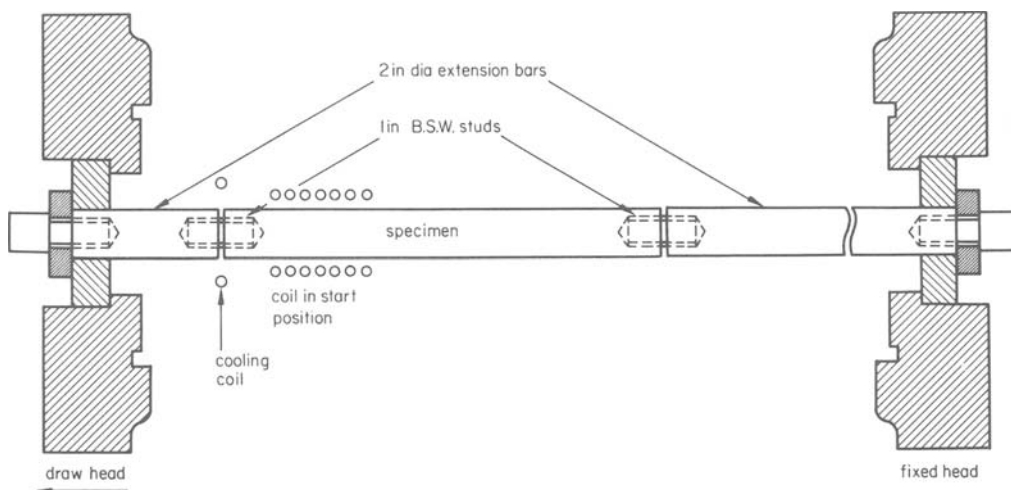


Figure 3. Specimen support in the rig.

steady state over several minutes, with several pyrometer readings of the maximum temperature recorded, indicate an accuracy of better than $\pm 5^{\circ}\text{C}$.

EXPERIMENTS

The first pull was a hot tensile test on mild steel of rectangular cross-section. Sections are shown in the photograph, figure 4. Although this was not a draw, it has been included because it shows an interesting variation in the cross-section and also suggests that substantial reductions might be achieved

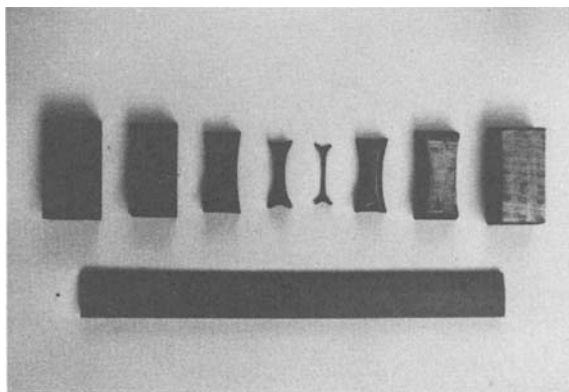


Figure 4. Hot tensile test cross-sections. Mild steel.

The first set of drawn bars were all of mild steel with the coil speed equal to the drawn head speed to give a 50 per cent reduction.

All the specimens were painted with an Acheson Colloids Dag material, Dag 1946, to stop scaling. Scale forms rapidly over 1000°C and does not allow heat to escape from the lower end of the deformation zone, by creating a kind of hot box (figure 5).

A selection of the specimens is photographed in figures 6 to 11. Figure 6(a) shows four 7.6 cm (3 in) diameter mild steel bars which were unsuccessfully drawn at the beginning of the programme. They are representative of a number of bars which failed to draw a constant cross-section, in the temperature range 1000°C to 1200°C , mainly the result of bad scaling.

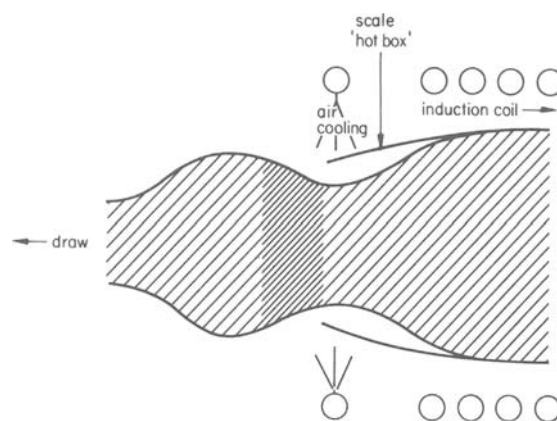


Figure 5. 'Hot box' created by mild steel draw.

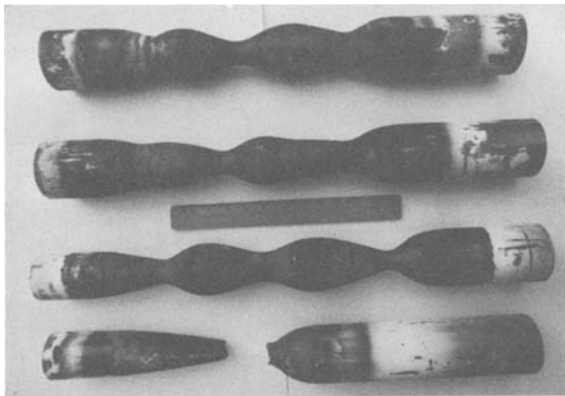
Figure 6(b) shows three 7.6 cm x 3.8 cm (3 in x 1.5 in) mild steel specimens, two of which were chromized.

Corner cracking does not occur in the three 3.8 cm x 3.8 cm (1.5 in x 1.5 in) mild steel specimens shown in figure 7(a). The bottom bar is chromized and apparently starts well, but a stable draw could not be established at this stage of the experimental work.

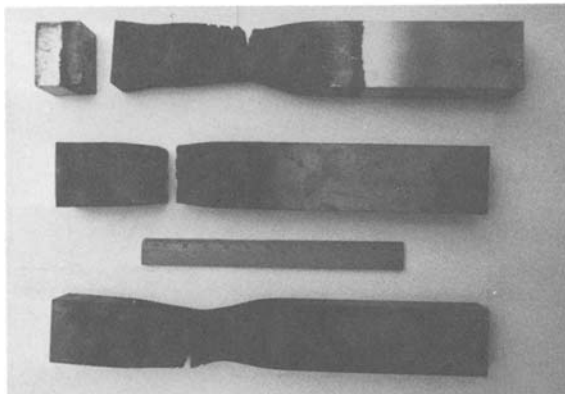
A selection of 2.5 cm (1 in) diameter niobium mild steel specimens appear in figure 7(b). Despite repeated efforts, this fine grained material, which could be drawn down to a very large reduction in area (on the right of the bottom bar) always failed to give a stable draw.

The three 5 cm (2 in) diameter and one 3.8 cm (1.5 in) diameter EN 58 J stainless steel specimens, figure 8(a), show the difficulty of establishing a stable draw, the bad surface cracks and the good surface obtainable in the 3.8 cm diameter bar.

The top 5 cm (2 in) diameter chrome steel specimens shown in figure 8(b) are good examples of stable draws. The surface is good and the damped oscillation is clearly seen. Both were reduced by 50 per cent in area. The third bar on the bottom was one of the first attempts to draw a taper. Oscillations of the diameter are evident, but the draw is apparently stable when the reduction in area is 66 per cent on the right of the bar.

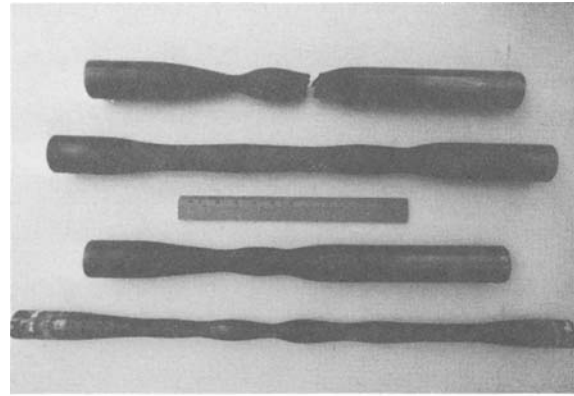


(a)

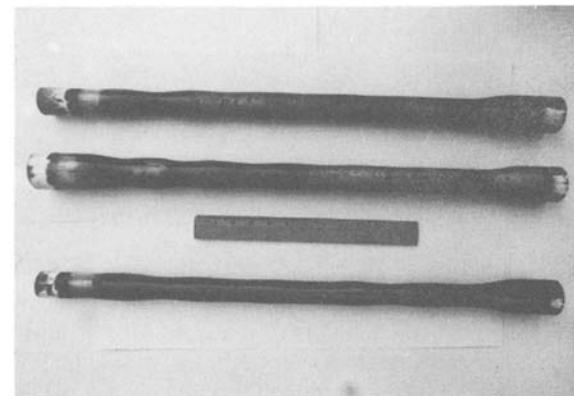


(b)

Figure 6. (a) 7.6 cm (3 in) diameter mild steel specimens. (b) 7.6 cm x 3.8 cm (3 x 1.5 in) mild steel specimens.

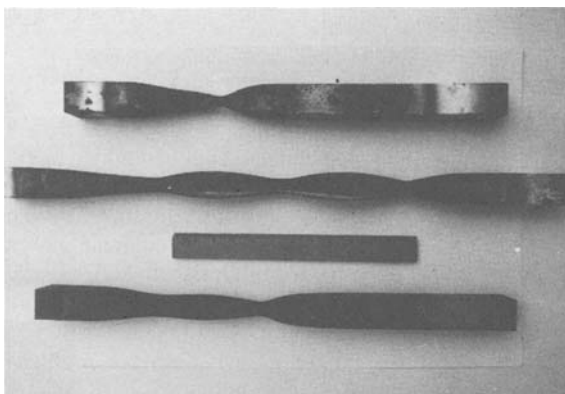


(a)

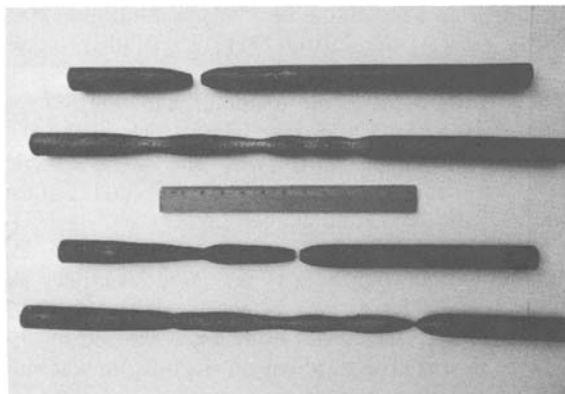


(b)

Figure 8. (a) 5 cm (2 in) and 3.8 cm (1.5 in) diameter EN 58 J stainless steel specimens. (b) 5 cm (2 in) diameter chrome steel specimens.



(a)



(b)

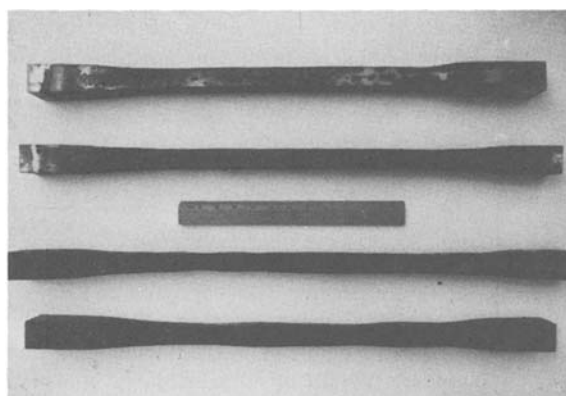
Figure 7. (a) 3.8 cm x 3.8 cm (1.5 x 1.5 in) mild steel specimens. (b) 2.5 cm (1 in) diameter niobium mild steel specimens.

Figure 9 shows some good results with chromized mild steel specimens. All the surfaces are good, but oscillations in the 7.6 cm (3 in) diameter bars are not fully controlled.

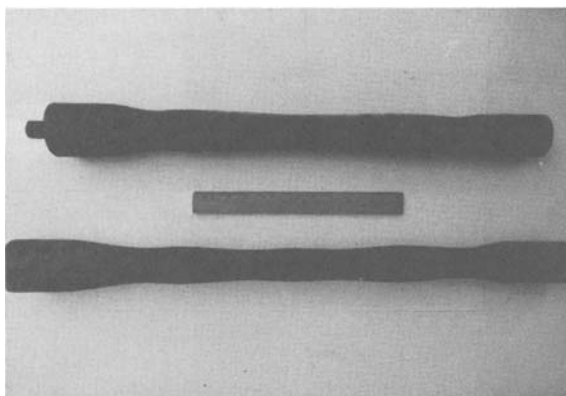
Some more good results obtained with 3.8 cm (1.5 in) diameter EN 58 J stainless steel specimens appear in figure 10. The top specimen in figure 10(a) has a good finish, the middle bar of figure 10(a) has a good surface finish and is reduced by more than 75 per cent in area. The third bar in figure 10(a) has been included because the neck on the right of the bar represents a percentage reduction in area of 88 per cent. However, no stable draw was ever achieved at this percentage reduction in area. Other examples are shown in figure 10(b).

Some intentionally tapered specimens are shown in Figure 11(a). The chrome steel specimen starts at 33 per cent and tapers to 66 per cent reduction in area. The tear on the right of the bar is the result of a step increase in the diameter which allowed a crack to propagate from the effective 'notch' produced half way down the deformation zone. The middle and lower titanium bars are tapered from 33 per cent to 66 per cent and 50 per cent to 75 per cent reduction in area respectively.

Some examples of the 2.5 cm (1 in) diameter titanium specimens appear in the photograph of figure 11(b).

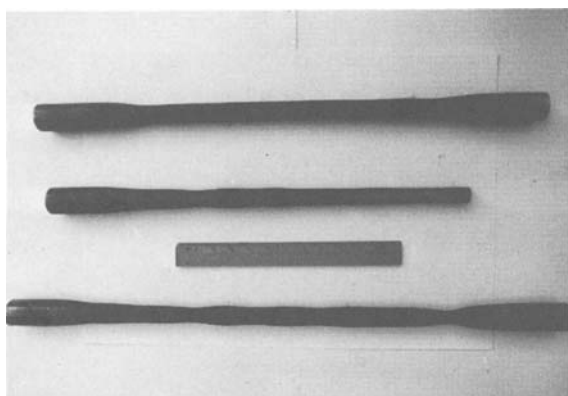


(a)

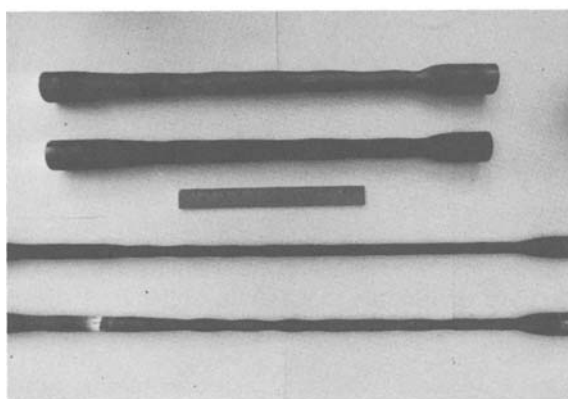


(b)

Figure 9. (a) 3.8 x 3.8 cm (1.5 x 1.5 in) chromized mild steel specimens. (b) 7.6 cm (3 in) diameter chromized mild steel specimens.

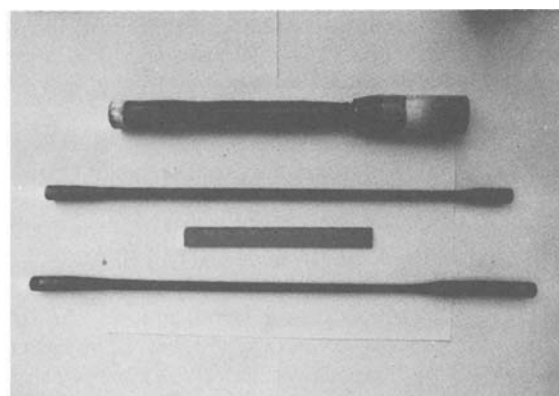


(a)

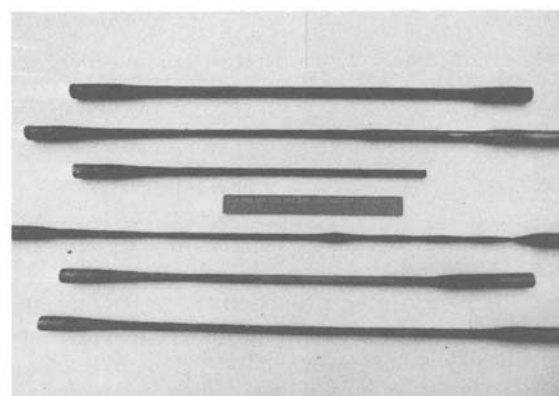


(b)

Figure 10. (a) 3.8 cm (1.5 in) diameter EN 58 J stainless steel specimens. (b) 3.8 cm (1.5 in) diameter EN 58 J stainless steel specimens.



(a)



(b)

Figure 11. (a) (i) 5.7 cm (2.25 in) chrome steel tapered specimen. (ii) 2.5 cm (1 in) diameter titanium tapered specimen. (iii) 2.5 cm (1 in) diameter titanium tapered specimen. (b) 2.5 cm (1 in) diameter titanium specimens.

Discussion of the experimental results

A summary of the conclusions which can be drawn is as follows:

- (1) Die-less drawing is possible with mild steel, chrome steels, EN 58 J stainless steel and titanium.
- (2) Percentage reductions in area up to 75 per cent for steel and 84 per cent for titanium are possible.
- (3) Round and square cross-sections may be drawn.
- (4) Speeds, up to the highest velocities possible for the maximum rate at which heat can be applied, of the order of 7.6 cm/min (3 in/min) give satisfactory stable draws.
- (5) Good surface finishes can be obtained on the drawn bars.
- (6) Although the specimens photographed in figure 6(a) represent unsatisfactory draws, we can nevertheless draw the conclusion that non-constant cross-section drawing under controlled conditions is possible.

THE THEORETICAL APPROACH

The ideal theoretical model is one which enables the temperature distribution to be analysed in terms of the heat input, the conduction, convection and radiation, and the convection resulting from the material flow. Having obtained the distribution of

temperature, the model should enable the profile of the bar and deformation zone, the stress distribution and the flow lines to be evaluated given the material properties.

The effect of strain hardening is reduced at elevated temperatures and the experiments of Cook⁴, giving true stress-strain curves for steel in compression, show an appreciable reduction of the effect as the temperature rises from 900°C to 1100°C.

Strain rate sensitivity is not only a much bigger effect but is also playing a significant role at the narrow end of the zone where there is no strain hardening. It is the narrow part of the zone which is critical when considering stability, and the neglect of strain hardening is quite justified here.

Stress-strain-strain rate-temperature relationships

The secondary creep equation

$$\dot{\epsilon} = \frac{2}{3} K'' \sigma^n \exp \left[\frac{-Q}{RT} \right] \quad (1)$$

has been used extensively to account for a temperature dependence⁵⁻⁹, but the activation energy is generally found to be a function of temperature, and the stress function, replaced by a hyperbolic sine function, sometimes gives a better representation.

The closest approach to an empirical expression which fits the available data for stainless steel is:

$$\sigma = K_2 \left\{ \frac{T^* - T}{\alpha + 2/\epsilon} \right\}^{n\epsilon^m} \quad (2)$$

where K_2 , α , T^* , n and the strain rate sensitivity index, m , are all constant.

A value of 0.2 was taken for the strain rate sensitivity index, m , as an upper bound. Cook's⁴ data at strain rates from 1.5 to 100 s⁻¹ gives a range of m values from 0.07 to 0.11 between 900°C and 1100°C rising to 0.146 at 1200°C.

Near parabolic curves are obtained when stress is plotted against the log of strain rate for titanium. These curves plotted for a number of temperatures led to the expression:

$$\sigma = \frac{T_1^* - \gamma \log \dot{\epsilon} - T^2}{\beta(-\log \dot{\epsilon})} + 8\dot{\epsilon}^{0.301} \quad (3)$$

where T_1^* , β and γ are constants.

Both Equations (2) and (3) proved to be clumsy when used and it was decided that the secondary creep Equation (1) should be used to develop both the simple tension and finite element approach.

Temperature distribution

The steady state heat conduction equation:

$$\frac{\partial}{\partial r} \left(k'r \frac{\partial T}{\partial r} \right) + \frac{\partial}{\partial z} \left(k'r \frac{\partial T}{\partial z} \right) + rQ' = 0 \quad (4)$$

together with the boundary conditions defines the problem, and a finite element analysis is possible. Zienkiewicz¹⁰ describes how the analysis is developed, where, with the aid of the calculus of variation,

a volume integral is minimized and a 'stiffness' matrix set up. Experimental surface temperatures give the boundary conditions and Q' is assessed at 25 per cent of the energy input per unit volume. Temperature contours may be obtained in this way. The above analysis does not take into account the small convection term resulting from the material flow.

This temperature distribution, calculated from the experimental surface temperatures and the finite element technique, forms the basis for the simple tension and finite element solution of the drawing problem.

The boundary conditions specified for this temperature distribution and the remaining analytical work were set up in the normal way for the assumed meshes.

Approximate solution assuming plane sections remain plane

The simplifying assumptions that plane sections remain plane and that the temperature of a thin section is constant, enable some progress to be made in understanding the process.

Accurate measurement of the deformation zone gives the cross-sectional areas and because the material is considered to be incompressible,

$$\frac{A_1}{A_i} = 1 + \frac{\delta z}{z}$$

At each section we can therefore obtain the engineering strain $e_z = \delta z/z$, from which we calculate the natural strain:

$$\epsilon_z = \ln(1 + e_z)$$

With material fed into the deformation zone at an axial velocity, V_C , the axial velocity of each section is determined from

$$V_i = \frac{A_1}{A_i} \cdot V_C$$

The slope of the velocity curve along the z axis gives the axial strain rate:

$$\dot{\epsilon}_z = \frac{V_i - V_{i-1}}{\delta z}$$

Since the load, L , is constant the stress at each section is

$$\sigma_i = \frac{L}{A_i}$$

A simple computer program has been developed to obtain tables of the parameters for each specimen from the measured diameters of the deformation zone.

Gunasekera¹¹ has taken the program a step further, for, since σ_i and ϵ_i are evaluated, a point can be obtained on the stress-strain-temperature curves which gives a temperature by straightforward linear interpolation.

The dynamic yield stress can be expressed as the product of a strain hardening function and a strain rate sensitivity function:

$$\sigma = \sigma_0 g(\epsilon) f(\dot{\epsilon})$$

If we now take the temperature into account by making σ_0 a function of temperature only we obtain

$$\sigma = h(T)g(\epsilon)f(\dot{\epsilon})$$

then

$$L = A\sigma = Ah(T)g(\epsilon)f(\dot{\epsilon}) \tag{5}$$

This may be written in differential form

$$\frac{dL}{L} = \frac{dA}{A} + \frac{dh}{h} + \frac{dg}{g} + \frac{df}{f} \tag{6}$$

If we write finite differences in place of the differentials and consider that each function varies with respect to one variable only, we may write:

$$\frac{\delta L}{L} = \frac{\delta A}{A} + \frac{(\partial h/\partial T)\delta T}{h} + \frac{(\partial g/\partial \epsilon)\delta \epsilon}{g} + \frac{(\partial f/\partial \dot{\epsilon})\delta \dot{\epsilon}}{f} \tag{7}$$

A necessary but not sufficient condition for stability is given by $\delta L/L = 0$.

All the finite differences are known from the experimental data, i.e. δA , δT , $\delta \epsilon$, and $\delta \dot{\epsilon}$. The strain hardening function can be taken as:

$$g = (1 + B\epsilon)^n$$

or taken from the experimental data of Cook⁴.

The rate of change of stress with temperature is constant for stainless steel, the only material for which we have useful data. Thus three of the four terms can be evaluated and the strain rate sensitivity term set at the values which make the summation of the four terms zero, i.e. $\delta L/L = 0$. When each term is plotted against z , an indication of the relative effects of reduction, strain hardening, temperature, and strain rate sensitivity is obtained.

Figure 12 shows these relative effects where the 'strain rate sensitivity' effect is assumed to be the dotted curve. The dotted curve in fact represents values such that $\delta L/L = 0$ is satisfied at every point along the zone.

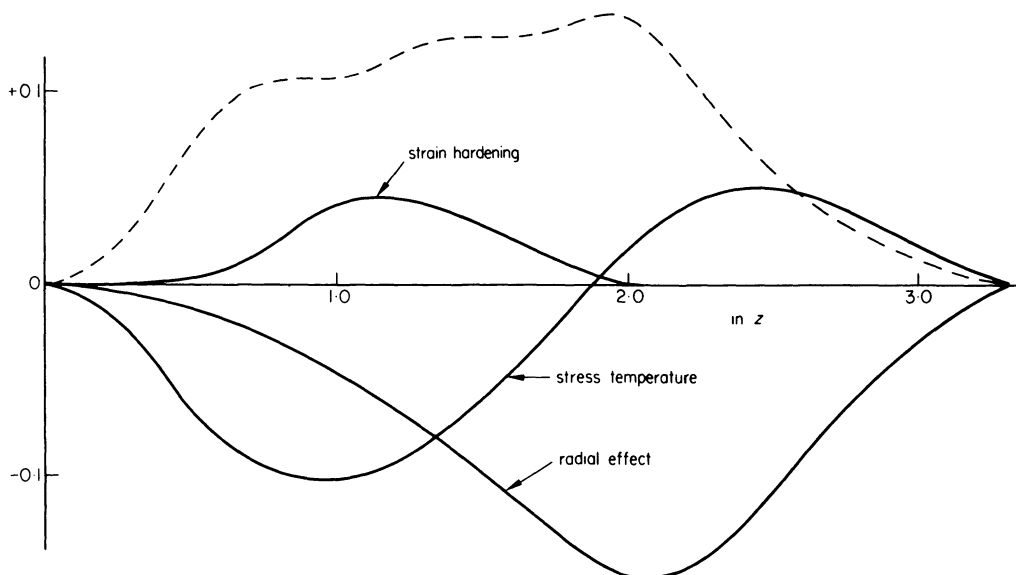


Figure 12. Relative effects of reducing area, stress-temperature dependence, and strain hardening.

Solution in simple tension

At the stage when the finite element method appeared to be intractable, it was decided to develop a solution where the rod is considered to be an assemblage of thin tubes.

The mesh is generated in terms of the number of nodal z and nodal r points for an axisymmetrical constant cross-section bar. The equivalent stiffness matrix is set up and the temperature distribution found using the conduction equation:

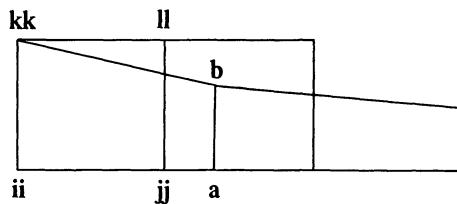
$$\frac{\partial}{\partial r} \left(k'r \frac{\partial T}{\partial r} \right) + \frac{\partial}{\partial z} \left(k'r \frac{\partial T}{\partial z} \right) + rQ' = 0.$$

The same mesh is used for 'plastically deforming' the bar, but instead of using triangular ring elements rectangular ring elements are employed. Initialization is carried out setting a constant stress (equivalent to a constant load), the velocity V_C at every nodal point, and the average absolute temperature is calculated for each rectangular element. Strain and strain rates are effectively set at zero.

The elemental tubes are now considered to be subjected to a pure tension and plastically deformed according to the equation of state

$$\dot{\epsilon} = K\sigma^n \exp \left[\frac{-Q}{RT} \right]$$

(compare Equation (1)), in the following way.



The element stress is taken at point **ii** and the strain rate calculated using the average absolute temperature and the known constants.

From the known strain rate, referenced to point *ii* and assumed constant throughout the element, the velocity of the point *jj* and *ll* is found relative to point *ii* and *kk* respectively. Hence we may calculate new *z* coordinates for the points *jj* and *ll*. The condition of constant volume enables the new *r* coordinate of points *jj* and *ll* to be found, and the engineering strain can be calculated. The new coordinates of *jj* and *ll* are shown above in the points *a* and *b* respectively, and these become the *ii* and *kk* for the next element.

The stress for the next element is calculated from:

$$\sigma_{jj} = \frac{r_{kk}^2 - r_{ii}^2}{r_{ll}^2 - r_{jj}^2} \cdot \sigma_{ii}$$

and the computer program continues for each element in the tube and then proceeds with the remaining elements in precisely the same way.

The mesh is now advanced a step into the 'heated zone'. Several elements of 'undeformed bar' are added to simulate the motion into the coil and an equal number of elements removed from the 'drawn bar'. The new nodal points are initialized and a new temperature distribution found. The temperature distribution on the second iteration is appreciably modified because, instead of 'seeing' a constant cross-section as in the first iteration, the coordinates of the mesh are appreciably changed. The program now recalculates a new set of coordinates, stresses, strains, and strain rates.

It was found that after only five iterations a 'stable draw' is established and very little computer time is needed to obtain the profile of the bar.

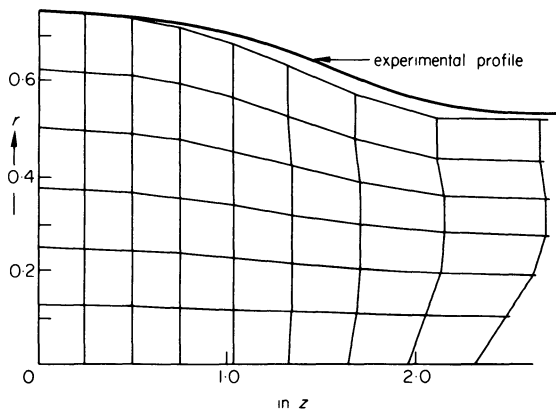


Figure 13. Mesh deformation obtained from simple tension analysis.

A plot of the stabilized mesh coordinates is shown in figure 13, using a typical temperature profile obtained when drawing EN 58 J stainless steel. The initial stress value, coil velocity and limiting or final velocity were taken from drawing experiments where the percentage reduction in area was 50 per cent and the constants from Sellars and Tegart⁵. An experimental profile is also shown and makes a fair comparison.

Finite element method

The main effort into a finite element analysis of the process developed from the techniques discussed in the Appendix using the displacement method of the

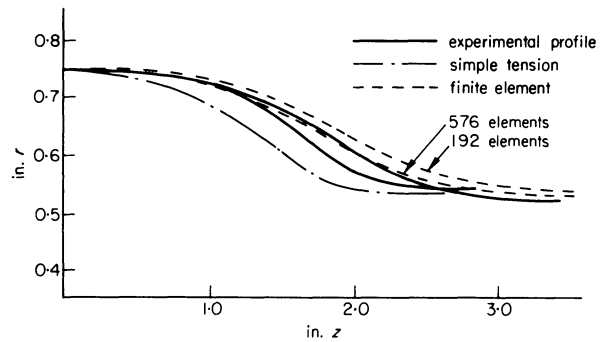


Figure 14. Deformation zone profiles for EN 58 J stainless steel 50 per cent reduction in area.

elastic stress-strain program. Comparisons between the predictions of this finite element approach and two of the experimental profiles obtained are shown in figure 14, together with the prediction from the 'simple tension' approach.

DISCUSSION

Experimental method

The photographs, figures 6 to 11, give a representation of the 160 drawn specimens which were tested during the experimental programme. Approximately one third of the total number of bars tested were stable draws and sufficient data has been extracted to make subsequent drawing experiments with titanium, EN 58 J stainless steel, and mild steel very nearly 100 per cent successful.

The titanium alloy Ti-6Al-4V presents no difficulties and it would appear that many superplastic alloys behave in a similar way. Johnson¹² gives a list of many metals and alloys which exhibit the phenomenon, having strain rate sensitivity indices greater than 0.5.

Success with the titanium alloy is all the more remarkable because loads less than 10 kN (1 ton) were obtained from a 1000 kN (100 ton) testing machine; loads at which the machine is least likely to give a steady pull.

There is little doubt that the successful draws with stainless steel and chromized mild steel have nothing to do with superplastic phenomena. Calculations from the available data⁴ give *m* values less than 0.15, and another mechanism to explain the equilibrium of the process must be found. Hardening due to the temperature gradient offers a satisfactory explanation, since there is no evidence of strain hardening. In point of fact Cook's data⁴ show strain softening at the strains encountered at the thinner section of the deformation zone. Stainless steel shows sensitivity to a slight increase or decrease in the rate of heat extraction, and figure 12 shows the stress temperature effect to be most important at the drawn end of the zone. Here the experimentally determined axial temperature gradient is of the order of 20°C/cm (50°C/in) which may be compared with 12°C/cm (30°C/in) computed by the plane sections approach. The unstable draws give computed axial temperature gradients of 120°C/cm (300°C/in) which could not be established in the experiments and which account for the instability.

A similar argument is not applicable to the titanium alloy. Firstly there was no example of instability and secondly the very rapid increase in the m value as the strain rate decreases causes the necessary hardening of the thinnest sections of the zone.

Chromizing was a satisfactory method of stopping scale formation, thus enabling a steady rate of cooling to establish the stable drawing of mild steel. Unfortunately it is an expensive process and allows grain growth to occur with deleterious effects on the material drawing properties. Berkatekt manufactured by Berk Limited, and a new lubricant¹³ is reported to be effective when used on mild steel at temperatures in excess of 1000°C, but the simplest method would be to find some method of extracting the heat at a constant rate. This might be a simple scraping device which removes the scale at the cooling sections thus allowing a constant rate of heat loss.

Although the experimental work on temperature distribution was sufficient to achieve the success reported, there is a good deal that could be done to develop the experimental method. For example, a secondary heat source warming the skin of the material, reflectors including aluminizing the inside of the heating coil, or lagging the bar with glass fibre tape are possible methods of minimizing the radial temperature gradients, which can produce surface cracks.

THEORETICAL METHOD

Plane sections

There is experimental and some theoretical work which shows that plane sections remain plane over the first half of the deformation zone at least. Figure 13 gives the resultant mesh deformation from simple tension analysis. This shows that plane sections remain substantially plane for much of the deformation zone and this was confirmed by experiments using visio-plasticity techniques. It is reasonable therefore, to use the method of plane sections to give an analysis of the varying parameters along the z axis of the heated zone. Favourable comparison between the computed profiles and experimental temperature profiles is further evidence confirming plane sections.

Simple tension

The advantage of using this simple approach lies in the speed of computation. Approximately 10 seconds of computer time is required to obtain a solution compared with over 1000 seconds for the finite element program (refined mesh of 576 elements) and about 400 seconds for the program using 192 elements.

The computed diameters are within 10 per cent of those obtained experimentally, and the strain rate variation compares favourably with the variations computed using the plane sections approach together with the experimental profiles.

Finite elements

Fair agreement was obtained between the deformed shape predicted by the simple finite element

approach outlined and the experimentally measured profiles. It seems from this that stress and strain distribution obtained from the finite element solution will give a reasonable prediction of the actual situation.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the Science Research Council and are grateful to the Metalworking Division of the British Steel Corporation (BISRA Laboratories, Sheffield) and to Rolls Royce (1971) Ltd (Bristol Engines Division) for the supply of specimens.

They also wish to acknowledge helpful discussions with Dr R. H. Johnson of the Electricity Council Research Centre and Dr A. H. Meleka of Rolls Royce (1971) Ltd.

APPENDIX

The strain-displacement relationship is derived from a linear representation of the r and z displacements in terms of the coordinates.

$$\{\epsilon\} = [B] \{\delta\}^e$$

where

$$\{\epsilon\} = \begin{Bmatrix} \epsilon_z \\ \epsilon_r \\ \epsilon_\theta \\ \epsilon_{rz} \end{Bmatrix} = \begin{Bmatrix} \frac{\partial v'}{\partial z} \\ \frac{\partial u'}{\partial r} \\ \frac{u'}{r} \\ \frac{\partial u'}{\partial z} + \frac{\partial v'}{\partial r} \end{Bmatrix} \quad (8)$$

$$[B] = [B_i, B_j, B_p]$$

where

$$[B_i] = \frac{1}{2\Delta} \begin{bmatrix} 0 & c_i \\ b_i & 0 \\ a_i/r + b_i + c_i z/r & 0 \\ c_i & b_i \end{bmatrix}$$

$$a_i = r_j z_p - r_p z_j$$

$$b_i = z_j - z_p$$

$$c_i = r_p - r_j$$

Equation (8) may be differentiated with respect to time to give

$$\{\dot{\epsilon}\} = [B] \{\dot{\delta}\}^e \quad (9)$$

where the matrix $[B]$, composed of elements related only to the r and z coordinates of the mesh, remains unchanged.

The elasticity matrix $[D]$ which relates the stresses and strains is derived from the elastic equations for an isotropic material:—

$$\begin{aligned}\epsilon_z &= \frac{1}{E} \{\sigma_z - \nu(\sigma_r + \sigma_\theta)\} \\ \epsilon_r &= \frac{1}{E} \{\sigma_r - \nu(\sigma_\theta + \sigma_z)\} \\ \epsilon_\theta &= \frac{1}{E} \{\sigma_\theta - \nu(\sigma_z + \sigma_r)\} \\ \epsilon_{rz} &= \frac{2(1+\nu)}{E} \cdot \tau_{rz}\end{aligned}\quad (10)$$

We have, on solving for the stresses

$$\{\sigma\} = \begin{Bmatrix} \sigma_z \\ \sigma_r \\ \sigma_\theta \\ \sigma_{rz} \end{Bmatrix} = [D] \{\epsilon\} \quad (11)$$

where $D = \frac{E(1-\nu)}{(1+\nu)(1-2\nu)}$

$$\begin{bmatrix} 1 & \frac{\nu}{1-\nu} & \frac{\nu}{1-\nu} & 0 \\ \frac{\nu}{1-\nu} & 1 & \frac{\nu}{1-\nu} & 0 \\ \frac{\nu}{1-\nu} & \frac{\nu}{1-\nu} & 1 & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2(1-\nu)} \end{bmatrix} \quad (12)$$

Using a typical secondary creep equation we can write the axial component of the strain rate

$$\begin{aligned}\dot{\epsilon}_z &= K'' \bar{\sigma}^{n-1} \exp\left[\frac{-Q}{RT}\right] \sigma'_z \\ &= \frac{2}{3} K'' \bar{\sigma}^{n-1} \exp\left[\frac{-Q}{RT}\right] \left[\sigma_z - \frac{1}{2}(\sigma_r + \sigma_\theta)\right]\end{aligned}\quad (13)$$

and once again we can see the similarity between Equations (10) and (13). We must replace ν by 0.5 and E by $\{(2/3)K''\bar{\sigma}^{n-1} \exp[-Q/RT]\}^{-1}$ to make the Equations (10) and (13) identical if the strains are replaced by strain rates and Equation (11) above becomes

$$\{\sigma\} = [D] \{\dot{\epsilon}\} \quad (14)$$

where the stress-strain matrix, $[D]$, becomes a stress-strain rate matrix the components of which are identical, i.e. Equation (12) defining $[D]$ above is unchanged.

The element stiffness is now found, using an approximation to simplify the computation of the volume integral which must be taken over the complete ring of material in the axisymmetrical analysis.

$$[k] = 2\pi [B]^T [D] [B] \bar{r} \Delta \quad (15)$$

where

$$\bar{r} = \frac{(r_i + r_j + r_p)}{3}$$

and the $[B]$ matrix is evaluated using the coordinates of the centroid.

The mathematical model which reduces the elements to a set of stiffness, solves the elastic problem by setting the forces equal to the stiffnesses multiplied by the displacements,

$$\{F\}^e = \Delta \cdot 2\pi \bar{r} [B]^T [D] [B] \{\delta\}^e \quad (16)$$

When applied to elastic problems only very small displacements of the mesh are encountered, but noting that the strain-displacement relationship can be differentiated to give a strain rate-velocity relation, Equation (9), it would appear possible to rewrite equation (16) above relating forces to velocities:

$$\{F\}^e = \Delta \cdot 2\pi \bar{r} [B]^T [D] [B] \{\delta\}^e \quad (17)$$

The only difference is that strains are replaced by strain rates and displacements are replaced by velocities.

Application of plasticity problems with large displacements appears straightforward but Poisson's ratio, ν , for an incompressible material is 0.5 and the $[D]$ matrix becomes singular because of the $(1-2\nu)$ term in the denominator of Equation (12). Hermann¹⁴ resolves the problem using a variational theorem which is valid for both compressible and incompressible materials. He formulates elastic field equations by utilizing displacements and a function of the mean pressure, but the computation time is appreciably increased by the necessity of solving for the mean pressure function in each element. By putting a value close but not equal to 0.5, the components of the $[D]$ matrix can be evaluated and some progress can nevertheless be made.

The mesh which is generated represents a constant cross-section specimen and the program is run to obtain a set of axial and radial velocities at each nodal point. Since the direction of material flow is defined by these velocities, the data may be used to modify the coordinates of the mesh in the following way. In figure 15 the arrowed lines represent the directions of the material flow which we express as a slope, S .

$$S = \frac{V_r}{V_z}$$

The intermediate nodal point, I, has a slope, S_I , represented by the dotted line IA. IB is the slope S_J , of the adjacent nodal point, and the new radial coordinate of nodal point, J, is modified to a position midway between points A and B.

$$r_J = r_I + \frac{(S_J + S_I) \cdot Z_n}{2} \quad (18)$$

A single modification of the mesh causes appreciable 'plastic deformation' and if the mesh is refined (figure 15 represents a very coarse mesh) the profile of the bar and flow lines in the material are found in very

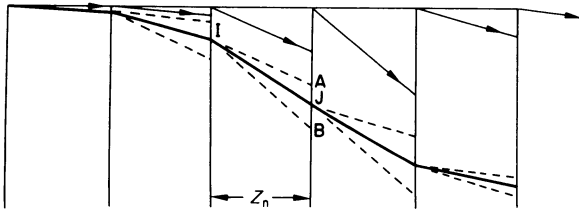


Figure 15. Modification of coordinates using velocity slopes.

few iterations. With the fine mesh used in the present work it was necessary to use an under-relaxation factor of 10 for several iterations to avoid crossed flow lines. The factor was removed after ten iterations.

Having modified the coordinates of the mesh, the program proceeds to iterate on the stresses using the equation of state.

The absolute temperatures at each nodal point were calculated using the finite element technique separately and read in as data. New velocities were calculated again and the program continued to give sufficient iterations for close tolerances.

Summarizing the program we have:

- (1) Initialize and generate the constant cross-section mesh. Read in the temperatures.
- (2) Evaluate the strain rate-displacement matrix, $[B]$.
- (3) Calculate the areas of the elements.
- (4) Evaluate the stress matrix, $[D]$, and $[B]^T [D] [B]$.
- (5) Initialize the loads.
- (6) Invert the stiffness matrix.
- (7) Set the boundary conditions, initial and final velocities and zero radial velocities at entry and along the centre line.
- (8) Iterate to give axial and radial velocities at each nodal point. Between 200 and 1500 iterations are required.
- (9) Modify the coordinates to fit the slopes determined from the r and z velocity ratios.
- (10) Compute the stresses and strain rates.
- (11) Iterate equivalent stresses. (approximately 50 iterations)
- (12) Return to 3.

- (13) Calculate the forces.
- (14) Print the solution.

REFERENCES

1. R. H. Johnson. Superplasticity. *Met. Rev.* (1970) 146.
2. R. H. Johnson. Private communication.
3. V. Weiss and R. A. Kot. Die-less wire drawing with transformation plasticity. *Wire Journal* (Sept 1969).
4. P. M. Cook. True stress-strain curves for steel in compression at high temperatures and strain rates, for application to the calculation of load and torque in hot rolling. *Proc. Conf. Properties of Materials at High Rates of Strain, Inst. Mech. Eng.* (1957) 86.
5. C. M. Sellars and W. J. McG. Tegart. La relation entre la resistance et la structure dans la deformation a chaud. *Memoires Scientifiques Rev. Met.* (1966) 63 No. 9, 731.
6. G. C. Cornfield and R. H. Johnson. Theoretical predictions of plastic flow in hot rolling including the effect of various temperature distributions. *Electricity Council Research Centre Report* (Oct 1971) M427.
7. F. W. Sharman. Analysis of Metal Deformation-Extrusion. *Electricity Council Research Centre Report* (Sept 1972) M537.
8. G. C. Cornfield. An investigation of the growth of instabilities in the die-less drawing of rods and tubes. *Electricity Council Research Centre Report* 1971.
9. J. A. Bailey, S. L. Haas and M. K. Shah. Effect of strain rate and temperature on the resistance to torsional deformation of several aluminium alloys. *Int. J. Mech. Sci.* (1972) 14, 735.
10. O. C. Zienkiewicz. *The finite element method in Engineering Science*, Ch. 15, McGraw-Hill, 1971.
11. J. S. Gunasekera. Private communication.
12. R. H. Johnson. Superplasticity in metals and alloys. *Sheffield University Metallurgical Society Journal* (1970) 9.
13. A. R. Globus. A new high temperature lubricant for hot rolling. *Iron and Steel Engineer* (Aug. 1970).
14. L. R. Herrmann. Elasticity equations for incompressible and nearly incompressible materials by a variational theorem. *AIAA Journal* (1965) 3, 1896.

A FUNDAMENTAL STUDY ON DIELESS DRAWING

by

HIDEO SEKIGUCHI*, KOJI KOBATAKE* and KOZO OSAKADA†

SUMMARY

A new metal forming process termed *dieless drawing* which uses no die or container has been developed. The metal bar is fixed at one end, and locally heated to a high temperature by an induction coil. It is then pulled, at the other end, with a constant velocity V_1 , while the induction coil is moved in the opposite, or same, direction at a constant velocity V_2 .

Larger reductions of area can be obtained even if the material has low workability, since the process is frictionless and is a type of hot working. The reduction of area is fixed by the ratio of the velocities, V_1/V_2 .

Some applications of the process are demonstrated experimentally; for example the production of smooth uniform bars (solid and hollow), tapered bars, and bars of various cross-section.

INTRODUCTION

Dieless drawing is a new technique of metal forming which does not use dies. The drawing bar (solid or hollow) is fixed at one end, and is locally heated to a high temperature by an induction heating coil. It is then pulled at the other end with a constant axial velocity V_1 , while the induction coil is moved in the opposite or same direction at a constant velocity V_2 . Since the heated zone of the bar has a low flow stress, necking occurs only in this region. Necking, however, is diffused out by the continuous motion of the induction coil, yielding a straight uniform bar. Dieless drawing has many advantages:

- (1) the absence of a die makes the process less expensive as compared to conventional processes
- (2) the process is suitable for a material which has a high strength and/or a high frictional resistance, since it is a frictionless process
- (3) a large reduction of area can be obtained in a single pass
- (4) the reduction of area is determined solely by the ratio of the velocities V_1 and V_2
- (5) if the velocity, V_1 or V_2 , is continuously or abruptly changed during the process, various cross-sections can be produced.

There are many factors that determine the success of this process: the temperature, the cooling rate, the reduction of area, the drawing velocity and the material properties. It has been found experimentally, however, that it is very easy to produce smooth

uniform bars with the large reductions of area characteristic of this process. Carbon steel and stainless steel in particular behave well. Some applications of dieless drawing, to produce various cross-sections, are demonstrated experimentally.

EXPERIMENTAL PROCEDURE

The materials used in these experiments on dieless drawing are 0.10 per cent and 0.45 per cent carbon steel, 18 per cent Cr-8 per cent Ni stainless steel, and pure titanium. Each drawing specimen is 250 mm long and 8 mm in diameter. No preliminary heat treatment was done. An Instron testing machine with a capacity of 5000 kg was used as the pulling equipment.

Induction heating is superior to any other heat source in view of the requirement for local heating, the ability to produce high heating rates, and the controlling of temperature. A high frequency induction heating device, used in this experiment, has a maximum power of 3 kW and a frequency of 2 MHz. The temperature of the material depends on factors such as the power of the induction heating source, the heating time, the shape of the induction coil, the distance between the coil and the material, and the material's physical constants (specific heat, density, thermal conductivity, permeability). The shape of the induction coil in this experiment is shown in figure 1. The coil can be moved at speeds of 0 ~ 300 mm/min by a servomotor transferring system.

* Department of Mechanical Engineering, Nara Technical College

† Department of Mechanical Engineering, Kobe University

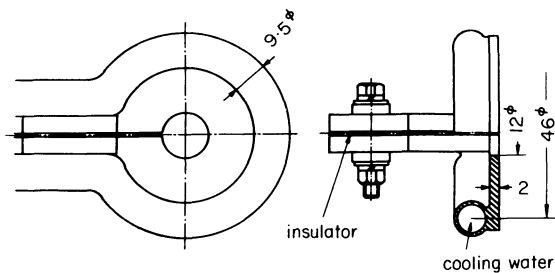


Figure 1. The shape of the induction coil.

It is necessary to cool the deformed region rapidly, and so an air-cooler, shown in figure 2, is attached under the induction coil. The distance between the coil and the nozzle of the air-cooler is kept at 11 mm, and the nozzle is located at a point 2 mm from the surface of the specimen. Air pressures from 0.25 to 2 kg/cm² are used.

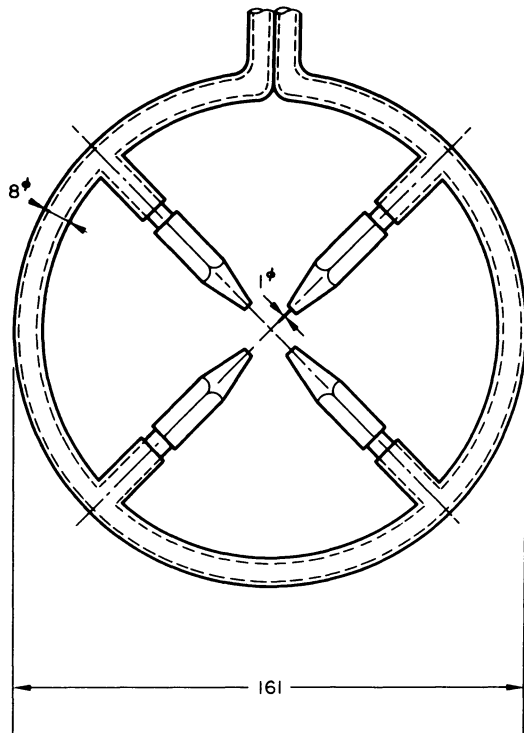


Figure 2. The shape of the air-cooler.

The variation of temperature along the length of the bar, and the radial direction of the bar, are both obtained during drawing, using chromel-alumel thermocouples of 0.3 mm diameter. These thermocouples are fixed to the surface or interior of the bar by percussion welding, and the measured temperature is recorded on an oscillograph through a low pass filter, to eliminate inductive noise.

RESULTS AND DISCUSSION

Smooth uniform bars

Figure 3 represents two typical arrangements for dieless drawing. A_1 and A_2 are, respectively, the area before and after deformation. While the bar is pulled with a velocity V_1 , the induction coil is transferred in the opposite, or same, direction at a constant velocity

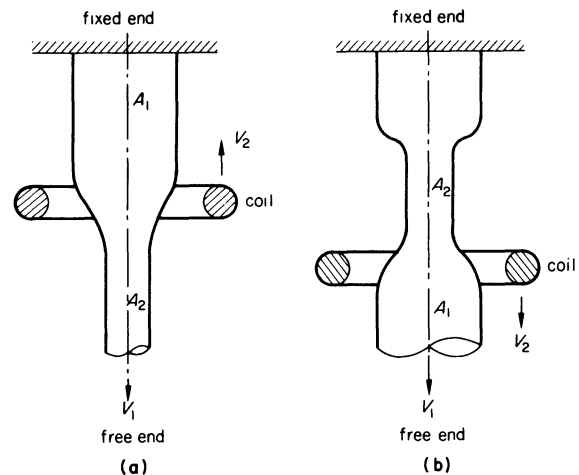


Figure 3. Two typical arrangements for dieless drawing.

V_2 . From the law of the constant volume, the relation

$$A_1 V_2 = A_2 (V_1 + V_2) \quad (1)$$

holds, when steady state is reached in the arrangement in figure 3(a). The reduction of area, r , is given as

$$r = 1 - A_2/A_1 \quad (2)$$

$$= V_1/(V_1 + V_2) \quad (3)$$

In the arrangement in figure 3(b), the velocities are related to the areas as

$$A_1 (V_2 - V_1) = A_2 V_2 \quad (4)$$

and the reduction of area is given by

$$r = V_1/V_2 \quad (5)$$

The arrangement in figure 3(a) is more effective in obtaining a large reduction of area, but the latter seems to be a unique process. Though these two processes have been accomplished successfully, the process involved in figure 3(a) is mainly described in this paper.

Other factors, such as the temperature distribution, affect the success of the process: an insufficient cooling rate produces a broken bar; and an insufficient temperature causes cracks to form on the surface of the bar. That is to say, the transition from an unstable state to a steady state is affected by many factors. If the selection of the conditions is not suitable, necking will not diffuse uniformly, and a broken or non-uniform bar will be obtained.

The suitable conditions have been found experimentally for carbon steels, stainless steel and titanium. For stainless steel and titanium, the maximum power of the induction heating device is supplied, and a power of 1.2 kW (40 per cent of its capacity) is used for carbon steels. The experiments are conducted under the tensile velocities from 10 mm/min to 50 mm/min. The pre-heat period before the deformation is 30 sec, and, in the case of stainless steel, temperature increases to about 900°C in this time period.

Figure 4 shows the relationship between the reduction of area, r , and the velocity ratio, V_1/V_2 .

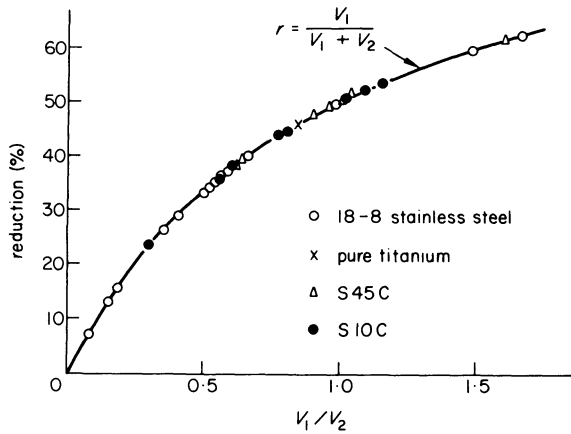


Figure 4. The relationship between the reduction of area, r , and the velocity ratio, V_1/V_2 .

All the experimental results are in good agreement with the values calculated from Equation (3). The maximum values of the reduction in a single pass are 61 per cent and 60.2 per cent, respectively, for stainless steel and 0.45 per cent carbon steel.

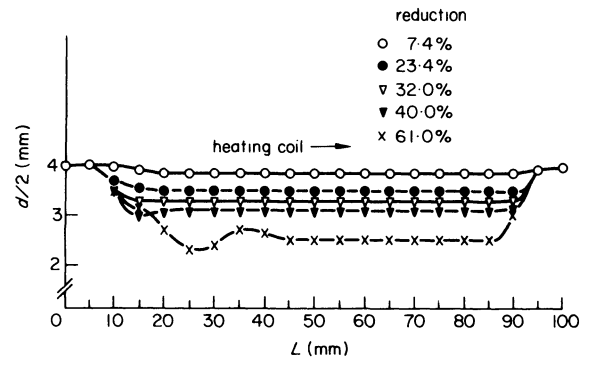


Figure 5. The profiles of the deformed specimens in stainless steel.

The surfaces of the drawn steel bars are very clean with little oxide films, but titanium has a rough surface. Figure 5 represents profiles of some specimens after deformation. In the range of the steady state, there is no variation in diameter. The actual profiles of the smooth uniform bars can be seen in figure 6.

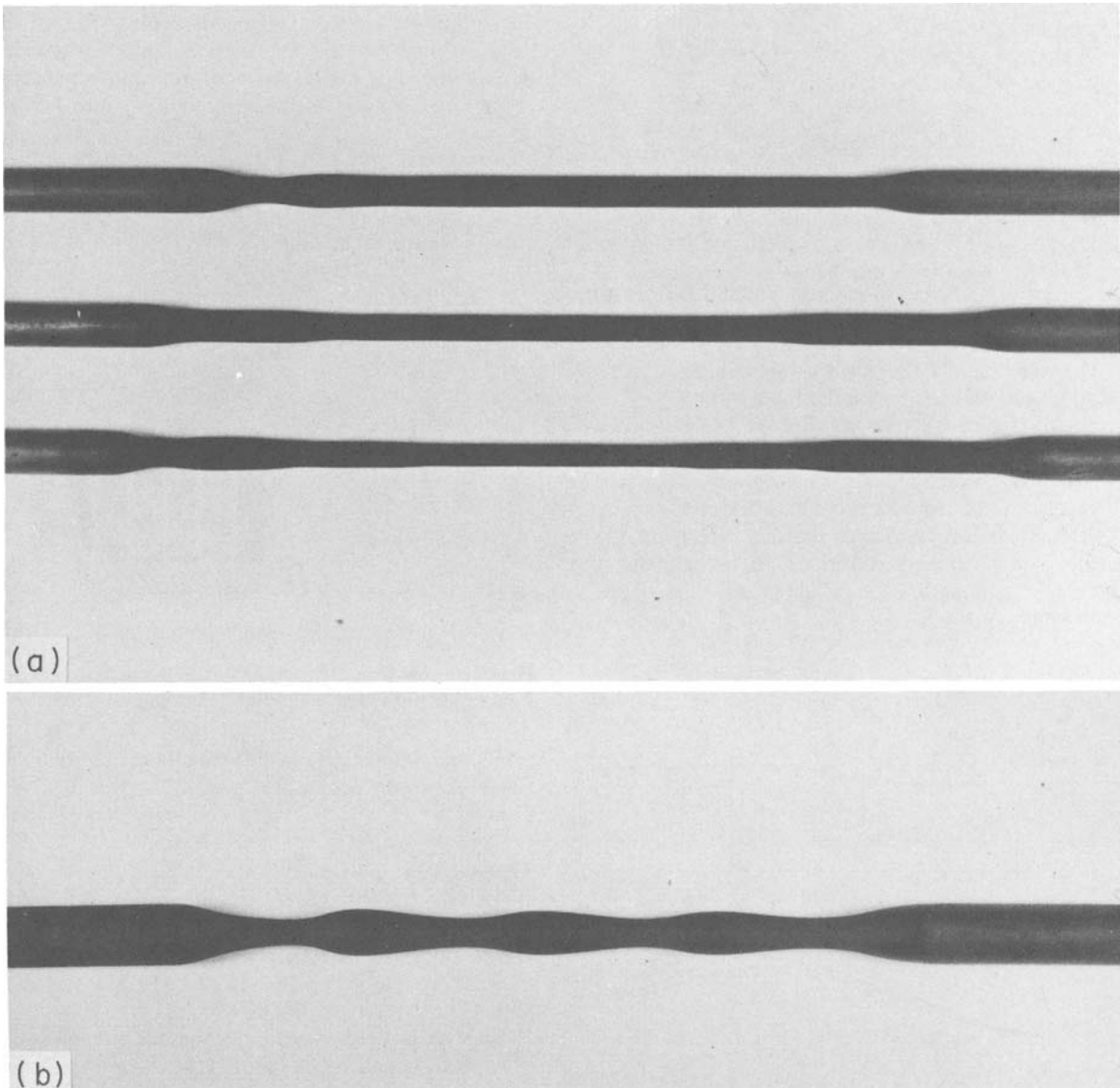


Figure 6. The drawing bars in 0.45 per cent carbon steel. (a) From upper, 1st pass: $r = 51.0$ per cent, 2nd pass: $r = 66.2$ per cent, and 3rd pass: $r = 75.0$ per cent, (b) a non-uniform bar ($V_1 = 15$ mm/min, $V_2 = 8.3$ mm/min).

The more suitable the operational parameters, the larger is the reduction of area obtained in a single pass. The specimen of carbon steel shown in Figure 6(b), for example, was expected to produce a reduction of 64 per cent, but instead the diameter varies as a sinusoidal function. If a larger value of the velocity V_2 is set at the beginning, and then the velocity is continuously changed to a smaller value, a smooth uniform bar is obtained.

Figures 7(a) and (b) show the temperature distribution in stainless steel specimens at time intervals in the pre-heating period. The pressure of the cooling air is 1 kg/cm^2 and the coil stays at the point $z = 0$. The maximum temperatures increase with the pre-heating time, and the surface temperatures are lower than the inner temperatures because of radiant heat loss.

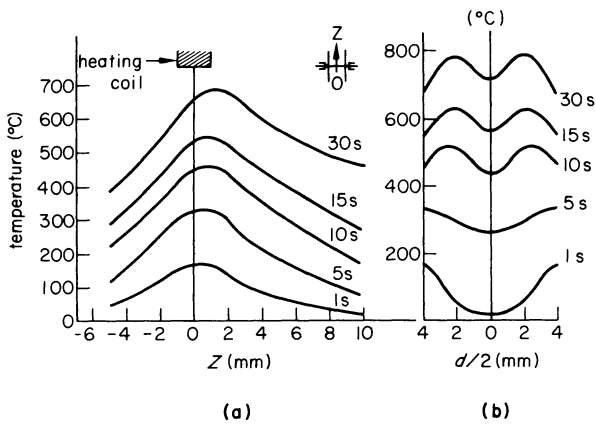


Figure 7. The temperature distribution in stainless steel. (a) The surface temperatures in the longitudinal direction, (b) the inner temperatures in the radial direction.

The variation of the surface temperature, during deformation, was also measured for a reduction of 36 per cent (figure 8). The profiles of the deformed specimen and the temperature distribution in the longitudinal direction are presented. A slight difference in the temperature distribution, caused by a different tensile velocity, has no effect on the profile in the necked region of the specimen. A suitable temperature configuration, however, determined by the shape of the coil and the cooling

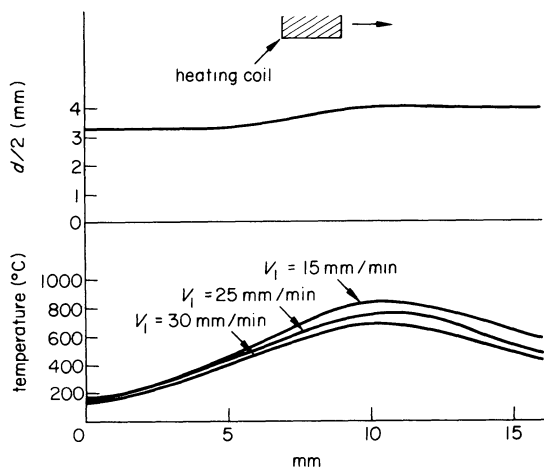


Figure 8. The variation of the surface temperature during deformation in stainless steel ($r = 36$ per cent).

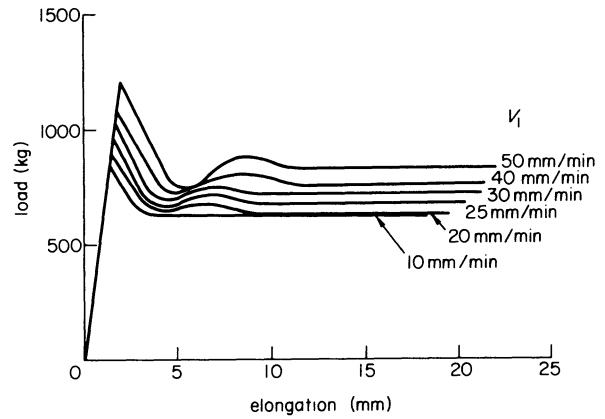


Figure 9. The load-elongation diagrams in 0.45 per cent carbon steel ($r = 40$ per cent, air pressure 1 kg/cm^2).

process, is essential for establishing a stable deformation.

As mentioned above, the same value of V_1/V_2 can be obtained with different velocities of V_1 and V_2 . A higher tensile speed results in an increase in the working force during deformation (figure 9). It is important to supply a sufficient cooling rate in the case of high speed deformation, since, when the tensile speed is higher than 50 mm/min , under the same conditions as in figure 9, the specimen breaks, as a result of too short a cooling period after deformation.

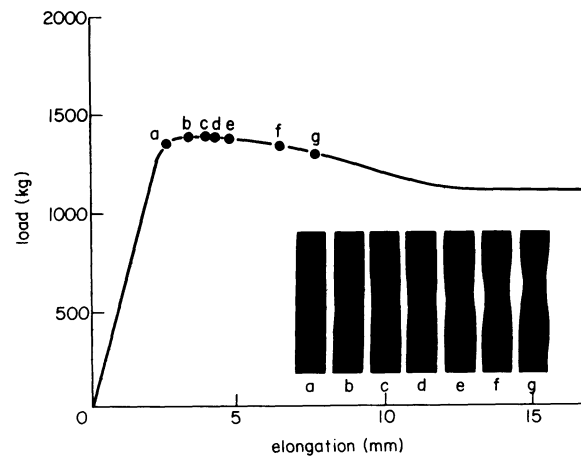


Figure 10. The load-elongation diagram in stainless steel ($V_1 = 10 \text{ mm/min}$, $V_2 = 17.8 \text{ mm/min}$, $r = 36$ per cent).

Figure 10 is a typical load-elongation diagram which shows the progress of necking, and its diffusing process.

Tapered bars

If the velocity V_1 and V_2 is continuously changed at a constant rate during the deformation, a tapered bar will be obtained. A controller is employed to change the coil velocity V_2 to a desired rate, using a photo-transistor and a DC motor.

Smooth tapered bars are easily produced and some examples are shown in figure 11(a). The conditions producing the desired degree of taper can be theoretically established, in terms of a function of the increment of the velocity V_2 .

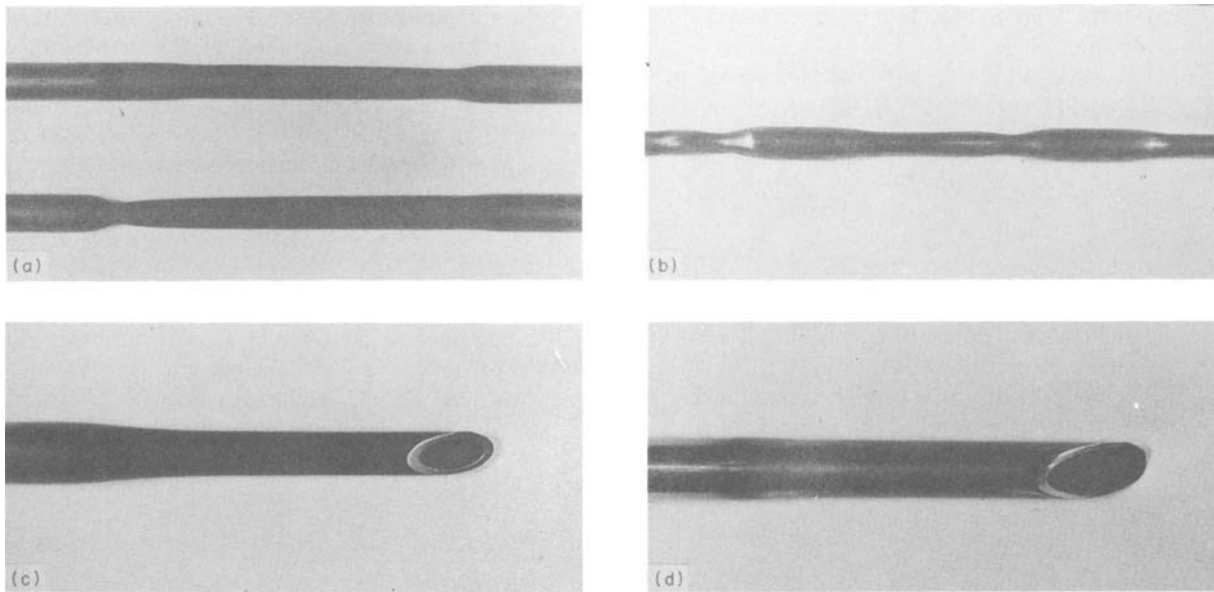


Figure 11. Some applications of dieless drawing in carbon steel. (a) Tapered bars (upper: $V_1 = 15$ mm/min, $V_2 = 60 \rightarrow 6.7$ mm/min, lower: $V_1 = 15$ mm/min, $V_2 = 15 \rightarrow 68.2$ mm/min), (b) a varied cross-section bar, (c) a smooth hollow bar ($V_1 = 15$ mm/min, $V_2 = 30$ mm/min), (d) a clad wire with stainless steel.

Varied cross-section bars

Bars of varied cross-section can be produced in dieless drawing, when

- (a) a bar is drawn partially or in several passes
- (b) the velocity, V_1 or V_2 , is abruptly or intermittently changed
- (c) an unstable deformation occurs regularly

The first and second methods are useful for making a varied cross-section bar such as a spindle, and it is easy to control the diameter with a prescribed programme. Sometimes an unsuitable condition for making a uniform bar produces an interesting shape, as in figure 6(b). An insufficient cooling rate cannot restrict the deformation in the necked region, and necking does not diffuse uniformly. Wavy bars, for example, are useful for wire used in producing pre-stress concrete, and for some decorations.

The working conditions required to produce varied cross-section bars can be found theoretically and experimentally, and their mass production is possible.

Uniform hollow bars

Dieless drawing can also be applied to producing a hollow bar (figure 11(c)). Steel pipes, which have an outer diameter of 8 mm and an inner diameter of 6 mm, are used for the specimens. There is a straight line relationship between the thickness and the outer diameter after deformation, and the surface is extremely smooth.

Drawn bars of composite materials

A stainless steel hollow bar and a carbon steel solid bar were drawn simultaneously by this process. The outer diameter of the pipe is 8.0 mm and the inner diameter is 6.0 mm. The solid bar has a diameter of 5.8 mm, which has been drawn from the original diameter of 8.0 mm in this process.

The composite materials are deformed to the outer diameter of 7.4 mm at a tensile velocity of 15

mm/min, and the good result is seen in figure 11(d). In order to examine the strength of the composite of these two materials, a shearing test is conducted. The composite bar has little resistance to separation by shear, but this method is useful in obtaining a clad wire which has resistance to corrosion.

Other applications

The mechanical properties of the dieless drawing bars have not been extensively investigated. Hardness tests are carried out. Drawn steels have almost the same hardness values as those of the as-received materials, and no variation in hardness is seen in the radial direction. This means they have been annealed after deformation.

It is possible to give a heat treatment to the material during or after drawing, because both the heating rate and the cooling rate are controllable; direct quenching, annealing, ausforming and so on can be applied effectively in this process.

Since a small hole on a bar is elongated longitudinally, the process is useful for making a long, narrow hole on the drawn bar. It is also confirmed that a steel plate can be shaped by this process.

CONCLUSIONS

Dieless drawing has been developed as a new metal forming process, and it has the following advantages

- (1) dieless drawing is suitable for a material which has a high strength and/or a high frictional resistance, since it is a type of hot working and a frictionless process
- (2) the working force in the process is very small
- (3) a large reduction of area can be obtained in a single pass
- (4) the reduction of area is determined solely by the ratio of the axial velocity and the coil velocity

- (5) there is little oxide film on the surface of the drawing bar
- (6) it is possible to give some kinds of heat treatments to the materials during and after drawing
- (7) the process can be applied to produce a tapered bar, a varied cross-section bar and others

However, it has also the following disadvantages

- (1) the suitable temperature distribution must be found experimentally
- (2) in the present position, high speed drawing is not realized because of the time required to heat and cool the material.

More applications of dieless drawing are possible and the mechanical properties of the drawing bars should be extensively investigated.

From a manufacturing point of view, the process seems to be available, and a practical dieless drawing machine will be made in the near future.

ACKNOWLEDGMENTS

The authors would like to thank Professor M. Oyane of Kyoto University for his advice, and Professor J. W. Spretnak of The Ohio State University for help in the preparation of the manuscript. They are also indebted to Mr K. Kubo of Nagoya Industrial Research Institute for his contributions by way of discussion.

THEORETICAL STUDY OF THE DRAWING OF BI-METAL ROD AND TUBE

by

H. T. CHIA* and D. H. SANSOME*

SUMMARY

A theory for axisymmetric drawing of tube and rod is presented and compared with experimental results. The equations derived enable drawing force, die pressure and interlayer shear stress between the different material combinations to be determined. It was found that the magnitude and direction of the shear stress at the bi-metal interface and the bi-metal thickness ratio had a marked effect on draw load.

NOTATION

B	$\mu \cot \theta$
H_c	thickness of tube before drawing
H_m	radius of rod before drawing
Q	$= h_c / (h_m + h_c)$
Y	mean yield stress
c	clad
e	exit plane
h_c	thickness of an element in the tube
h_m	radius of an element in the rod
m	matrix
p	die pressure
p_m	pressure between the tube and rod
r	overall reduction of area
θ	die semi-angle
θ_m	angle between the tube and rod interface and the centre line of rod
ϕ	redundant work factor
σ	stress
τ_i	shear stress at interface between tube and rod
τ_s	shear stress between die and tube

INTRODUCTION

In recent years there has been a marked increase in the use of bi-metal rod and tube in electrical, thermal, nuclear and aerospace engineering. With an appropriate selection of materials and rod diameters and tube thicknesses it is possible to achieve the required properties of strength combined with good thermal or electrical conductivity or good corrosion resistance. In the cable manufacturing industry, for example, high tensile wire is clad with high conductivity copper or aluminium. To produce bi-metal rod and tube combinations by drawing it is essential to investigate how the parameters such as material combination, rod diameter and tube thickness affect the mechanics

of the draw process. With the equations developed by the authors in this paper it is possible to determine the draw force, die pressure and interlayer shear stress of bi-metal rod for any dimension and material combination provided the tube wall is thin compared with the rod diameter.

Some workers have published studies on the mechanics of rolling and drawing of sandwich strip^{1,2}, but there remains a lack of understanding of the mechanics of bi-metal rod and tube-drawing. Consequently, this work was undertaken to enable the manufacturers to establish rational draw schedules founded on proved theoretical and experimental studies. The work reported here is part of an extensive theoretical and experimental programme involving studies of the drawing of bi-metal rod and tube and bi-metal tube on tube combinations. It is hoped to present further experimental and theoretical work on bi-metal rod and tube-drawing at the 16th MTDR Conference.

EQUILIBRIUM ANALYSIS FOR THE DRAWING OF BI-METAL TUBE ON ROD

Figure 1(a) shows the state of stress when the rod and tube are being drawn. The element shown in figure 1(a) is enlarged in figure 1(b). In this analysis it is assumed that the tube has a thin wall, that is $Q < 15$ per cent. The materials used are assumed to be rigid plastic so that there is no work-hardening effect. The axial and radial stresses are assumed to be uniformly distributed and principal stresses. In this analysis the die semi-angle is small and $\mu < 0.1$. Since drawing occurs through a conical die, the clad matrix interface is assumed to converge toward the virtual apex.

* Department of Production Engineering, The University of Aston in Birmingham

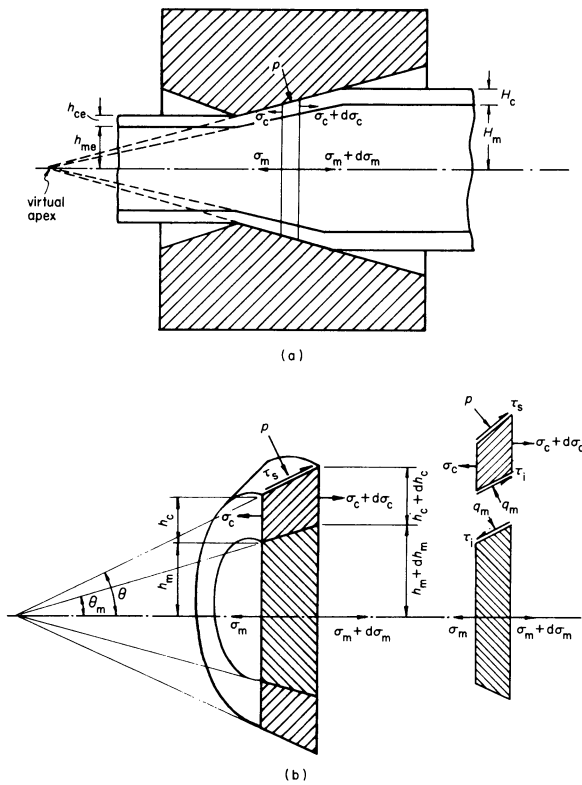


Figure 1. (a) State of stress in the bi-metal rod and tube. (b) State of stress for an element in the bi-metal rod and tube.

Considering first the equilibrium of horizontal forces in the matrix only,

$$\frac{d(\sigma_m h_m^2)}{dx} + 2h_m \tan \theta q_m - 2h_m \tau_i = 0 \quad (1)$$

Similarly, considering the equilibrium of horizontal forces in the clad,

$$\frac{d(\sigma_c h_c h_m)}{dx} + \tau_s (h_m + h_c) + h_m \tau_i + p(h_m + h_c) \tan \theta - h_m q_m \tan \theta_m = 0 \quad (2)$$

As the forces in the direction perpendicular to the die surface are in equilibrium,

$$p = q_m + \tau_s \tan \theta + \tau_i \tan \theta_m \quad (3)$$

and since the frictional contribution to the die pressure and the die semi-angle are assumed to be small,

$$p = q_m$$

Using the Tresca Yield criterion,

$$\sigma_1 - \sigma_3 = Y$$

For matrix,

$$\sigma_m + q_m = Y_m$$

For clad,

$$\sigma_c + q_m = Y_c$$

$$\therefore \sigma_m - \sigma_c = Y_m - Y_c \quad (4)$$

Since the materials for tube and rod are non-work-hardening, Y_m and Y_c are constant. Differentiating (4) gives

$$\frac{d\sigma_m}{dx} = \frac{d\sigma_c}{dx}$$

From the geometry of figure 1(b),

$$\tan \theta_m = \frac{dh_m}{dx} \quad (5)$$

$$\tan \theta = \frac{dh_c}{dx} + \tan \theta_m \quad (6)$$

$$\frac{dh_c}{dx} = \tan \theta - \tan \theta_m \quad (7)$$

Putting (6) into (1) gives

$$\frac{d\sigma_m}{dx} = \frac{2}{h_m} \tau_i - \frac{2}{h_m} \tan \theta_m (q_m + \sigma_m) \quad (8)$$

Substituting (6) and (7) into (2) and letting

$$Q = \frac{h_c}{h_c + h_m}$$

gives

$$\frac{d\sigma_c}{dx} = \frac{\tan \theta_m}{h_m} \left\{ -\frac{\tau_s}{Q} \left[\frac{1}{\tan \theta_m} + \frac{\tan \theta}{1-Q} \right] - 2\sigma_c - \frac{\tau_i}{Q} \left[\frac{1-Q}{\tan \theta_m} + \frac{\tan \theta_m}{1-Q} \right] - q_m \left[\frac{2-Q}{1-Q} \right] \right\} \quad (9)$$

When Q is small,

$$\frac{2-Q}{1-Q} \approx 2$$

that is

$$Q = \frac{h_c}{h_m + h_c} < 15 \text{ per cent}$$

and equating (9) and (8) yields

$$\tau_i = \frac{2Q(1-Q) \tan \theta (Y_m - Y_c) - \tau_s \sec^2 \theta}{2Q + (1-Q) \sec^2 \theta} \quad (10)$$

In figure 2, Q is plotted against h_c , for $Q = 0.15$, h_c/h_m is about 16 per cent. Hence the maximum tube thickness varies from about 0.020 in for 0.125 in diameter rod to about 0.16 in for 1 in diameter rod.

Equation (10) does not hold as $Q \rightarrow 1$ since it was assumed in establishing the equilibrium equations that Q is small.

The variation of τ_i with small values of Q appears in figure 3.

From equation (1) it is possible to obtain

$$\sigma_m = \int \frac{2}{h_m} (\tau_i - Y_m \tan \theta_m) dx$$

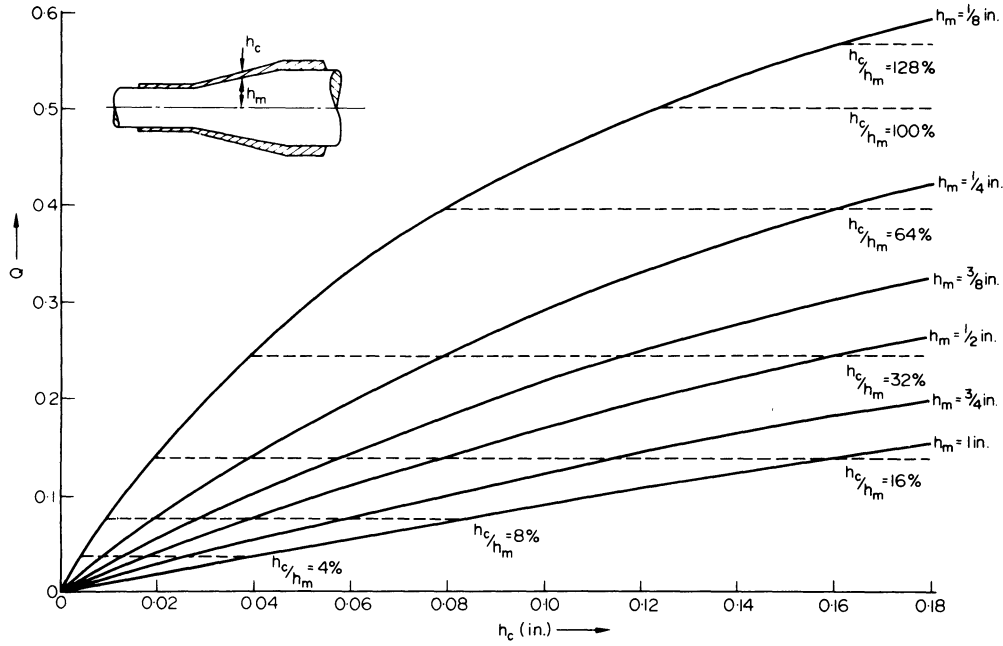


Figure 2. Variation of Q against tube thickness.

Substituting τ_i from (10) and integrating,

$$\sigma_m = \frac{2}{(1-Q)\tan\theta} \times \left\{ \left[\frac{2Q(1-Q)\tan\theta(Y_m - Y_c) - \tau_s \sec^2\theta}{2Q + (1-Q)\sec^2\theta} \right] - Y_m(1-Q)\tan\theta \right\} \ln h_m + C_m = A_1 \ln h_m + c_m$$

where

$$\frac{2}{(1-Q)\tan\theta} \left\{ \left[\frac{2Q(1-Q)\tan\theta(Y_m - Y_c) - \tau_s \sec^2\theta}{2Q + (1-Q)\sec^2\theta} \right] - Y_m(1-Q)\tan\theta \right\} = A_1$$

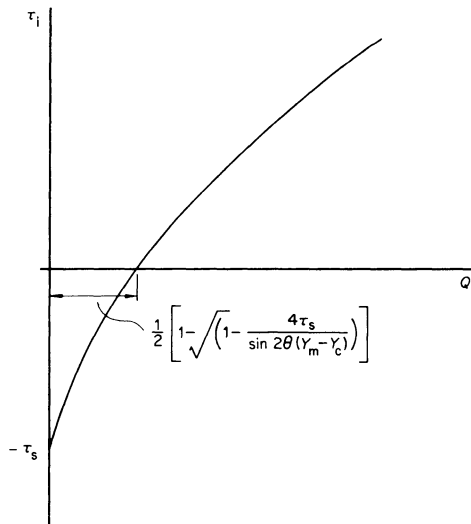


Figure 3. Variation of τ_i against small value of Q .

Similarly,

$$\sigma_c = \left[-\frac{\tau_s \sec^2\theta}{Q(1-Q)\tan\theta} - 2Y_c - \tau_i \frac{\sec^2\theta}{Q\tan\theta} \right] \times \ln h_c + C_c = A_2 \ln h_c + C_c$$

where

$$A_2 = \left[-\frac{\tau_s \sec^2\theta}{Q(1-Q)\tan\theta} - 2Y_c - \tau_i \frac{\sec^2\theta}{Q\tan\theta} \right]$$

Summing the forces at the entry plane in the absence of back tension,

$$\sigma_m = A_1 \ln \frac{h_m}{H_m} - (Y_c - Y_m) \frac{2H_c}{H_m + 2H_c}$$

$$\sigma_c = A_2 \ln \frac{h_c}{H_c} + (Y_c - Y_m) \frac{H_m}{H_m + 2H_c}$$

From these two expressions, σ_{total} can be determined.

$$\sigma_{total} = \frac{\ln\sqrt{(1-r)}}{1+Q} \left[-\frac{4\tau_s}{(1-Q)\sin 2\theta} - 4Y_cQ - \frac{4\tau_i}{\sin 2\theta} + \frac{4Q(1-Q)(Y_m - Y_c)\sin 2\theta - 4\tau_s}{2Q\sin 2\theta + 2(1-Q)\tan\theta} - 2Y_m(1-Q) \right] \quad (11)$$

Equation (11) clearly indicates that for a certain die angle and lubricant used, the magnitude and direction of the shear stress at the bi-metal interface

and the bi-metal thickness ratio have a marked effect on draw load.

When $\cos^2 \theta \approx 1$, Equation (11) can be further simplified.

$$\frac{\sigma_{total}}{Y_m} = \frac{1}{1+Q} \left[1 - Q + 2Q \frac{Y_c}{Y_m} + \frac{2\tau_s}{(1-Q) \sin 2\theta Y_m} \right] \times \ln \left(\frac{1}{1-r} \right) \quad (12)$$

Considering equilibrium of forces at the die face,

$$\frac{p_{mean}}{Y_m} = \frac{\frac{(1-r)}{r} \left\{ (1-Q) + 2Q \frac{Y_c}{Y_m} \right\} \ln \left(\frac{1}{1-r} \right)}{1+Q - \frac{\mu}{\sin \theta} \left[\frac{1-r}{r(1-Q) \cos \theta} \ln \left(\frac{1}{1-r} \right) - (1+Q) \cos \theta \right]}$$

Taking the work-hardening effect into consideration, when $Q = 0$,

$$p_{mean} = \frac{\phi Y_m \left(\frac{1-r}{r} \right) \ln \left(\frac{1}{1-r} \right)}{1 - \frac{\mu}{\sin \theta} \left[\frac{1-r}{r \cos \theta} \ln \left(\frac{1}{1-r} \right) - \cos \theta \right]} \quad (13)$$

Putting $\tau_s = \mu p_{mean}$, Equation (12) can be expressed as

$$\frac{\sigma_{total}}{\phi Y_m} = \frac{1}{1+Q} \left\{ (1-Q) + 2Q \frac{Y_c}{Y_m} + \frac{\mu(1-r) \ln \left(\frac{1}{1-r} \right)}{(1-Q) \left[r \sin \theta \cos \theta - \mu \left(\frac{1-r}{1-r} \right) - r \cos^2 \theta \right]} \right\} \times \ln \left(\frac{1}{1-r} \right) \quad (14)$$

OBSERVATIONS ON THE ANALYSIS

From Equations (12) and (14) it is clear that the drawing stress can be determined when either the coefficient of friction μ or the shear stress τ_s between the die and clad interface are known. Equation (12) requires a knowledge of μ and p_{mean} and so it is not as useful as Equation (14). Figure 4 shows the graph of $\sigma_{total}/\phi Y_m$ against reduction of area using Equations (12) and (14) for $Q = 0$, that is when drawing rod alone. The values of μ and p_{mean} were obtained from Weistreich's experimental results³ and substituted into Equation (12).

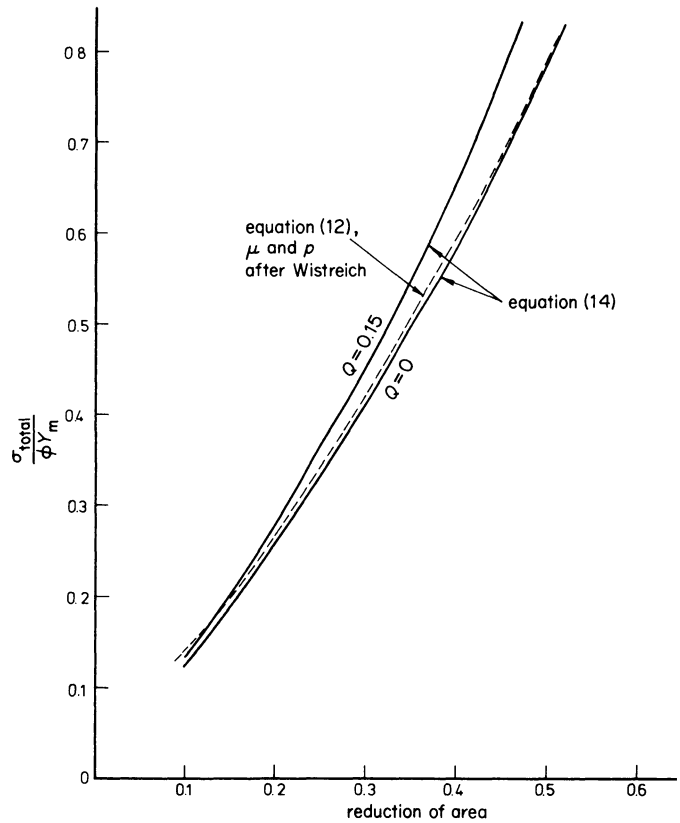


Figure 4. Change of $\sigma_{total}/\phi Y_m$ against reduction of area for different values of Q .

Figures 5 and 6 show the graphs of draw force and pressure plotted against reduction of area with Weistreich's experimental results and Equations (13) and (14). The value of redundant work factor ϕ for $Q = 0$ was obtained from the nomogram by Rowe⁴. The

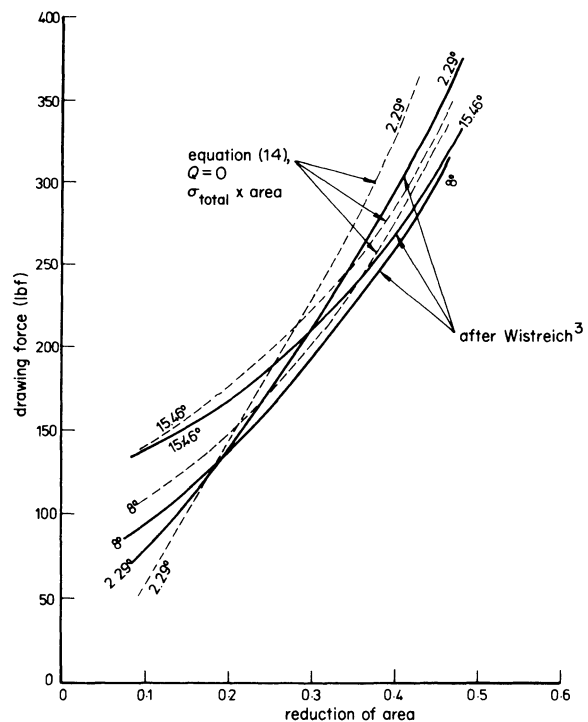


Figure 5. Drawing force against reduction of area for conventional rod drawing.

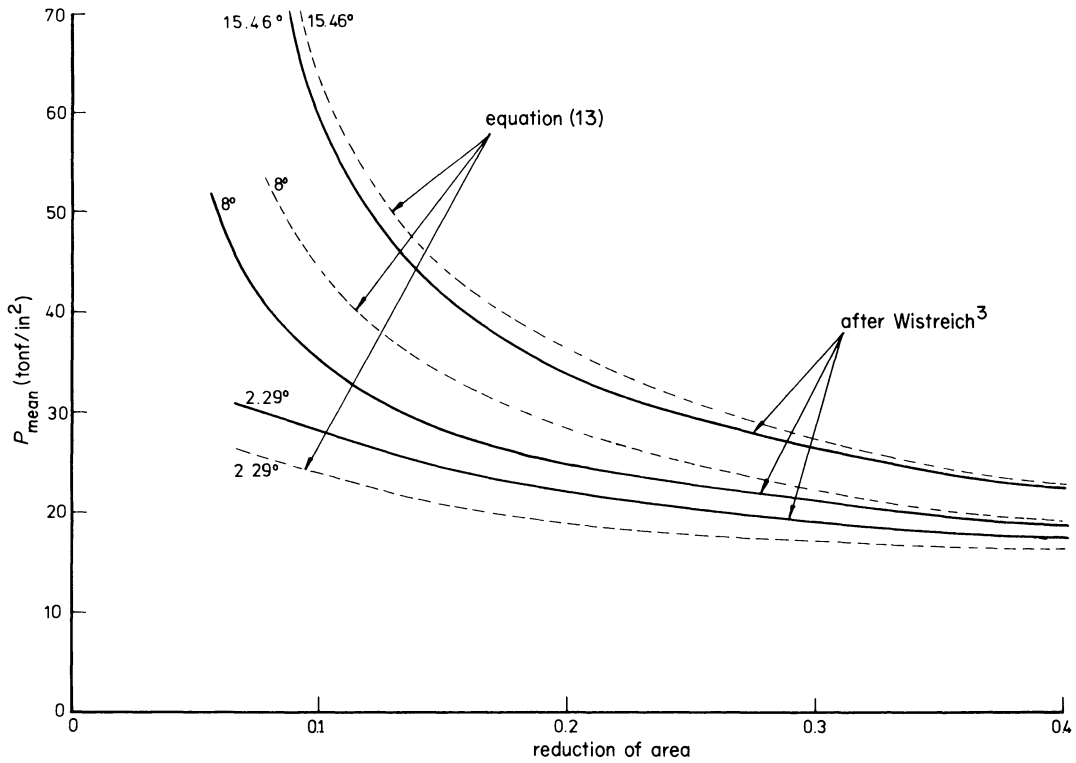


Figure 6. Die pressure against reduction of area for conventional rod drawing.

value of ϕ for Q other than zero can be obtained experimentally by plotting

$$\frac{\sigma_{total}}{Y_m \left(1 - Q + 2Q \frac{Y_c}{Y_m} + \frac{2\tau_s}{(1 - Q) \sin 2\theta Y_m} \right)}$$

against $\ln(1/1 - r)$.

Figures 7 and 8 are graphs of $\sigma_{total}/\phi Y_m$ and p_{mean} versus $\theta(1 - r)/r$.

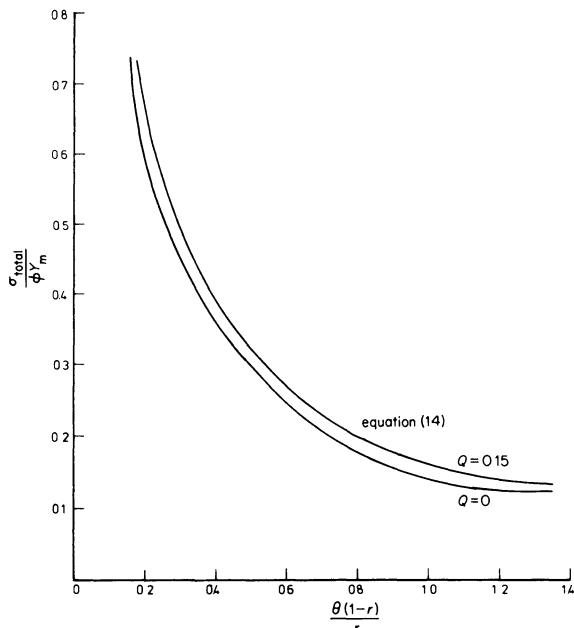


Figure 7. $\sigma_{total}/\phi Y_m$ against $\theta(1 - r)/r$ for bi-metal drawing.

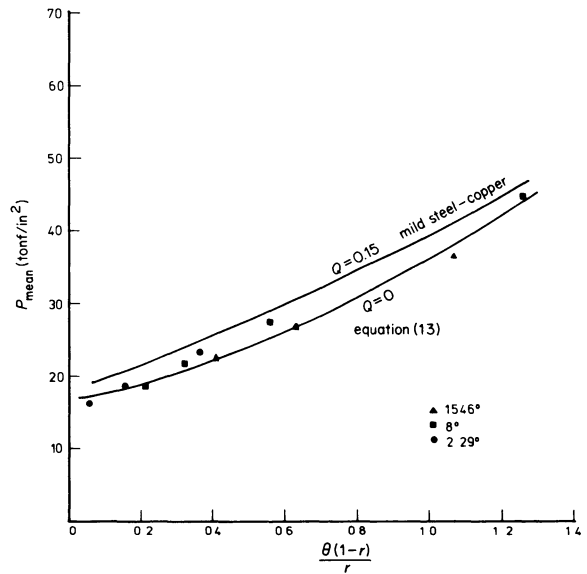


Figure 8. Mean die pressure against $\theta(1 - r)/r$ for bi-metal drawing.

Figure 9 compares the σ_{total} obtained by experiments and Equation (14) for stainless steel tube on $\frac{1}{2}$ in aluminium rod.

EXPERIMENTAL EQUIPMENT AND PROCEDURE

The tests were performed on a draw-bench made in the University of Aston in Birmingham⁵. The draw load and die load were measured by (1) a load meter at the gripping jaw and (2) a load meter between the die and the frame of the draw-bench. These load meters employed Wheatstone bridge circuits whose output was measured by a Southern Instrument ultra-

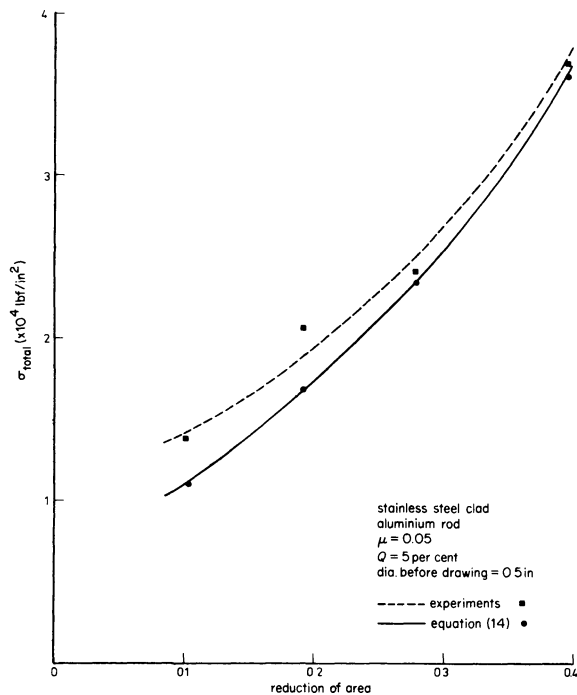


Figure 9. Comparison of theoretical and experimental drawing stress against reduction of area for $Q = 5$ per cent.

violet recorder. The dimensions of the bi-metal rod and tube were very closely controlled. The tubes and rods were cold drawn to size, pickled, neutralized, washed thoroughly and dried in air. The tubes and rods were fitted together. The dies used have a semi-angle of 7° . The lubricant used was Droyt drawing oil 7117.

DISCUSSION OF RESULTS

Figure 9 shows that the experimental and theoretical results are in good agreement. The slight under-estimation of the draw-load results from the use of industrial dies which incorporated both a radius profile and a parallel land. The theoretical curve in figure 9 from Equation (14) was based on the assumption that the die profile was conical and that no land was included in the design. When estimating the draw force using radiused dies the equivalent semi die-angle should be used. To obtain an accurate prediction of the draw load it is essential to have an accurate measurement of the bi-metal thickness ratio and reliable stress-strain curves for the materials before drawing. Consequently, having reliable data on these parameters, Equations (13) and (14) give a reasonably good estimation of draw load.

CONCLUSIONS

The derivation of equations for bi-metal tube and rod drawing is based on an assumption that the tube wall is thin by comparison with the rod diameter. This is indeed the case in the bi-metal rod drawing results reported in this paper. The expressions are applicable to the drawing of thin tube on rod and also to the drawing of conventional rod. Although drawing loads, and to some extent die pressures, are of practical interest to the industry in producing drawing schedules, the prediction of interlayer shear stress with different material and thickness ratio combinations are also possible from Equation (10), and this information is of value to establish optimum draw conditions. The effect of interlayer shear stress on drawing load can be easily seen from Equation (11). With the good agreement between the theoretical and practical results, the drawing stress and die pressure can be confidently predicted for a variety of bi-metal tube and rod combinations.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to the Commonwealth Scholarship Commission for their support of Mr Chia and to the Departments of Mechanical and Production Engineering for the use of their facilities. Also their thanks are due to Mr S. Twamley and Mr G. M. Jones for their assistance in the provision of experimental equipment and instrumentation. The authors are also indebted to Mr R. Beasley of Accles and Pollock Limited for his generosity in supplying tube, and to Mr T. Milner of the Delta Metal Company Limited who loaned some dies.

REFERENCES

1. R. R. Arnold and P. W. Whitton. Stress and deformation studies for sandwich rolling hard metals. *Proc. Inst. Mech. Engrs.* (1959), **173**, no. 8, p. 241.
2. A. A. Afonja and D. H. Sansome. A theoretical analysis of the sandwich rolling process. *International Journal of Mechanical Sciences* (1973), **15**, no. 1, pp. 1-14.
3. J. G. Wistreich. Investigation of the mechanics of wire drawing. *Proc. Instn. Mech. Engrs.* (1955), **169**, 654-5.
4. G. W. Rowe. *An Introduction to the Principles of Metalworking*. Edward Arnold Ltd. 1971, p. 293.
5. C. E. Winsper and D. H. Sansome. Application of ultrasonic vibration to the plug drawing of tubes. *Metal Forming* (March 1971), p. 71.

AN OSCILLATORY DEEP-DRAWING ANALOGUE

by

M. J. R. YOUNG* and D. H. SANSOME†

SUMMARY

Consideration is given to the application of ultrasonic vibration to the deep-drawing process. It is argued that from an understanding of the mechanics of cup-drawing, excitation of the die in a radial mode is likely to yield greatest benefit.

The desirability of quantifying both the surface and volume effects of ultrasonic excitation prompted the design of a simulative oscillatory test based on the Sach's wedge-drawing test. Careful design of the tool configuration, using a wedge test-piece having bi-axial symmetry, permitted measurement of the frictional forces associated with both the blank-holder plate and die profile radius.

Experimental results obtained with commercially pure aluminium indicated that 13 kHz, die vibration at an amplitude of 5×10^{-5} in reduced the effective blank-holder friction force to zero. Reductions in punch load of up to 40 per cent were obtained with increased die amplitude and the limiting draw ratio, under these particular test conditions, was increased from 1.90 to 2.00. It was inferred from the results of oscillatory wedge tests that significant benefits could be obtained by the correct application of vibrations to the axisymmetric process of cup-drawing.

INTRODUCTION

Since the early 1960s considerable scientific effort has been devoted to studying the application of both high and low-frequency vibrations to a wide range of metal forming processes. Consequently it may seem surprising that until recently the commercially important deep-drawing process has received scant attention. This omission is not due to any lack of appreciation of its importance on the part of the research worker; it is more a reflection of the difficulty of vibrating the relatively large drawing dies commonly used in deep-drawing processes. Renewed activity in the deep-drawing field has, to some extent, resulted from the commercial success achieved in ultrasonic tube-drawing. An appraisal of the technical problems of applying high-frequency vibrations to deep-drawing dies has led to a better understanding of the oscillatory principles involved and indicated the need to develop efficient vibrators based on a radial mode operation.

It was clear also that this technology could develop only in the light of a thorough understanding of the mechanics of the oscillatory deep-drawing process. To this end a programme of research was defined which was planned to establish a mechanical analogue capable of providing reliable simulative test data whilst permitting a complete theoretical analysis of oscillatory deep-drawing.

This paper outlines the basic philosophy behind the design of the mechanical analogue and describes the experimental procedure and test results obtained with commercially pure aluminium drawn at an excitation frequency of 13 kHz.

THE APPLICATION OF ULTRASONICS TO DEEP DRAWING

General principles of oscillatory metal forming

The objectives of applying vibrations to a metal-working operation are threefold: to reduce the process forces, to influence the nature and effect of interfacial friction, and to improve the surface finish of the drawn product. The effectiveness of the process depends upon correct design of the oscillatory system. It is important to select the correct vibrational mode and to transmit adequate acoustic power to the forming tools.

Forming forces are reduced when an oscillatory stress is induced in the deformation zone. This arises from a superposition mechanism, among other effects, in which the mean deformation load is reduced by an amount equal to the applied oscillatory force amplitude. This effect has been clearly illustrated by Winsper and Sansome¹, and later by Dawson², in wire-drawing with longitudinal vibrations at both low and high frequencies.

* Metal Forming Division, Ultrasonics Ltd., Shipley, Yorks.

† Department of Production Engineering, University of Aston in Birmingham

Generally the principal surface effect which influences the frictional conditions at the tool-workpiece interface, and which is of considerable importance in reducing forming forces, occurs as a result of the reversal of the friction vector. This is achieved when the oscillatory axis is parallel to the line of action of the friction force and when the particle velocity amplitude exceeds the linear velocity of the deforming metal. This mechanism had not been adequately demonstrated under conditions of plastic deformation until the research programme described in this paper was undertaken.

Most of the metal-forming processes examined in researches have utilized longitudinal oscillatory systems. Although the state of knowledge concerning the design of such systems is well advanced, evidence is emerging which indicates that in appropriate circumstances greater benefits can be achieved from a radial mode vibration. Thus the reduction in draw force relies on the fact that it is possible to vibrate a drawing die so that it exerts a radial swaging action on the deforming workpiece, producing a stress variation which lowers the peak drawing load. Also, the radial oscillatory configuration permits good acoustic coupling of the vibratory system to the work tool (in most cases a drawing die) and yet, due to the vibrational mode induced in the deformation zone, little energy is lost by transmission along the drawn rod or tube compared with the equivalent axial system. Oscillatory power is therefore absorbed in the deformation zone where it is required. These factors, together with others, were borne in mind when considering the application of an ultrasonic vibration to the deep-drawing process. The arguments influencing the choice of vibrational mode are discussed in greater depth in the following section.

Vibration mode

The choice of the vibrational mode in the present case was dictated by the following considerations:

- (1) Research into ultrasonic tube-drawing, with an axially oscillated plug, consistently indicated significant force reductions. These were claimed to result in an increase in the maximum area reduction possible, for a given material. If the die or blank-holder plate used in cup-drawing is vibrated in a radial mode, the relative motion between the tools and workpiece becomes analogous to that achieved in tube-drawing with an axially excited plug.
- (2) It is believed that reduction of friction can be achieved by establishing a condition in which the friction vector acting at the interface between tool and workpiece is cyclically reversed. This occurs when relative cyclic motion is set up at the friction interface. A radial oscillatory mode would permit such a mechanism to operate at the test piece-die and test piece-blank-holder interfaces.
- (3) The use of a longitudinal mode applied to either the punch or die was felt likely to result in a superposition effect similar to that observed in wire-drawing at both low and high frequencies. This has been found to produce limited benefits due to the transmission of

stress waves into the drawn product, resulting in a peak stress which is equal to (and sometimes exceeds) the non-oscillatory drawing stress. Under these conditions no increase in area reduction would be anticipated.

- (4) It was considered likely that a longitudinal vibration of the punch would not significantly affect the friction conditions in the cupping die or at the blank-holder-workpiece interface. Friction reduction was likely, however, between the punch and inner cup wall. This would have a detrimental effect on the drawing process and would be expected to lower the limiting draw ratio.
- (5) Previous analysis of the cup-drawing process indicated that the radial drawing-in stress constitutes the major proportion of the cup wall stress. It was therefore considered desirable to adopt a vibrational mode which offered the possibility of influencing either the stress system or material properties during this stage of drawing.

It was clear from the above arguments that a radial mode excitation of the deep-drawing die or blank-holder plate was likely to yield greater benefits than an axial mode of oscillation of the punch or of the die. A simulative test was therefore required which could readily incorporate the appropriate oscillatory system whilst retaining the essential characteristics of the deep-drawing process.

THE 'WEDGE TEST' AS AN OSCILLATORY DEEP DRAWING ANALOGUE

Characteristics of the Sach's 'wedge test'

In the wedge-drawing test a tapered sample of strip is drawn through a die of similar profile, producing an essentially parallel-drawn strip. The initial shape of the test-piece is a close approximation to a sector from a plane circular disc, and the deformation effected during drawing is therefore closely representative of that undergone by the blank during the 'cupping' test. The maximum length or width of wedge drawable under given conditions is taken as the criterion of drawability, and this may be expressed as the ratio of the initial to final width, or as the ratio of the radii defining the circular extremities of the wedge. This is illustrated in figure 1.

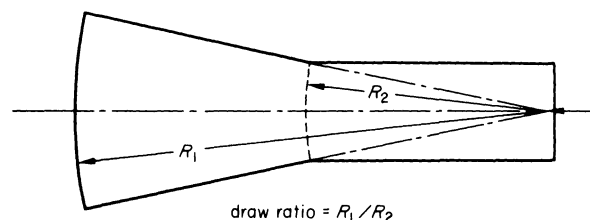


Figure 1. Wedge test-piece.

The wedge test has received classical examination by a number of researchers in the past³ including Loxley and Swift⁴ who, it may be recalled, undertook an exhaustive programme of tests in order to provide a reliable comparison with cupping test data.

Their work revealed a significant discrepancy between the values of the length to diameter ratio determined for aluminium and steel by the respective tests. They attributed this to the presence of edge friction in the wedge test which resulted, in the case of mild steel, in an increased drawing stress and a consequential reduction in the limiting draw ratio from 2.18 to 1.81.

It was argued that the increase in wedge drawing load due to edge friction was to some extent offset by the absence of load increments arising from the motion of the corresponding cupping blank over the die radius. Frictionless wedge-drawing would, therefore, be expected to result in a higher length to diameter ratio relative to the cupping test. However, even the best solid-phase lubricants were unlikely to eliminate edge friction and it was concluded that the wedge test results would always result in a lower value of the limiting draw ratio than the equivalent cupping test.

Oscillatory 'wedge-drawing'

One might be tempted to conclude from the previous consideration that the fundamental problem of edge friction disqualifies the wedge test as a deep-drawing analogue. However, it is apparent on examination that this test should prove a most convenient means of indicating, at least qualitatively, how the process forces are influenced by the application of vibrations to the tools.

That objective may be achieved by adapting the wedge-drawing test so that it includes each stage of drawing present in the axisymmetric process. This is done by designing the test-piece so that it represents two diametrically opposed sectors from a plane circular blank. The bi-axial symmetry so produced then permits simultaneous drawing of the two sectors through individual wedge dies by applying a force in a direction normal to the radial axis of the test-piece. This arrangement—shown schematically in figure 2—may readily be employed to examine the effects of

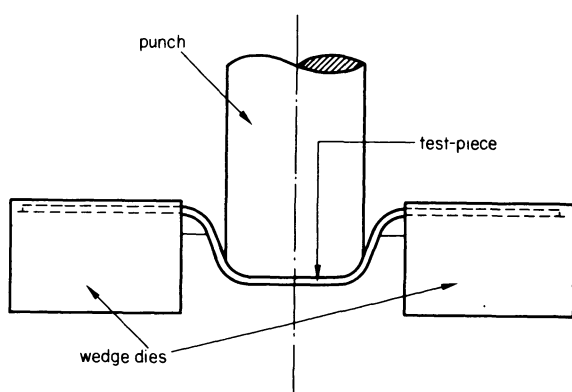


Figure 2. Oscillatory wedge test.

different vibrational modes on the drawing load. In particular, radial mode vibrations of an annular cup-drawing die may be simulated by vibrating both wedge dies longitudinally. Also, a resonant system based on axial symmetry offers considerable simplification of both experimental and analytical techniques compared with the corresponding radial system. In addition to simplifying the problem of acoustic

design, the relatively simple bi-axial test readily permits the direct measurement of the blank-holder friction force.

In the oscillatory wedge-drawing test two basic tenets may be postulated:

- (1) Axial oscillation of the dies, or blank-holder plate, or both, might result in the partial or complete removal of edge friction and in addition have some influence upon the remaining punch load components common to both cup and wedge-drawing tests. If this were the case the ratio of length to diameter would be expected to be equal to or greater than the corresponding value for cup-drawing. Consequently, the oscillatory form of the wedge test might therefore be a very close approximation to the cupping test.
- (2) Alternatively, a reduction in the forming load may be produced without the friction conditions at the edges of the test-piece being affected. Under these circumstances, although the results could not be compared directly with the corresponding cup-drawing data, any reduction in punch load would be associated with that stage of drawing which has its counterpart in the cupping test. If this were the case the ratio of length to diameter obtained in the oscillatory wedge-drawing test would be greater than that in the classical non-oscillatory wedge-drawing test but less than the corresponding value for cup-drawing. The results would, therefore, be expected to indicate qualitatively the effects produced in cup-drawing with radially resonant tooling.

From experience gained in oscillatory wire-drawing and tube-sinking it became apparent that the first postulate was untenable but that the second, if supported by experimental data, would confirm the fundamental principles of oscillatory metal deformation. Consequently it was decided to adopt the wedge-drawing analogue as a most critical method of testing principles, oscillatory technique and analytical procedure.

EXPERIMENTAL EQUIPMENT

General design

The wedge drawing subpress and instrumentation are shown schematically in figure 3. The subpress consisted essentially of a horizontally opposed pair of wedge dies, (1) driven at 13 kHz by two conical horns (2). The ends of the bi-symmetric test blank were held in their respective wedge dies by blank holder plates (3) having a similar but reverse profile machined on the lower surface. The hold-down force was provided by two hydraulic cylinders (4) pressurized from a common source. The complete assembly (shown in figure 4) was contained in a subpress which was attached to the ram of a hydraulic testing machine. A punch (5) of rectangular section was mounted on the underside of the fixed crosshead, and drawing was effected by raising the main ram with the subpress attached.

The friction force acting at the blank holder

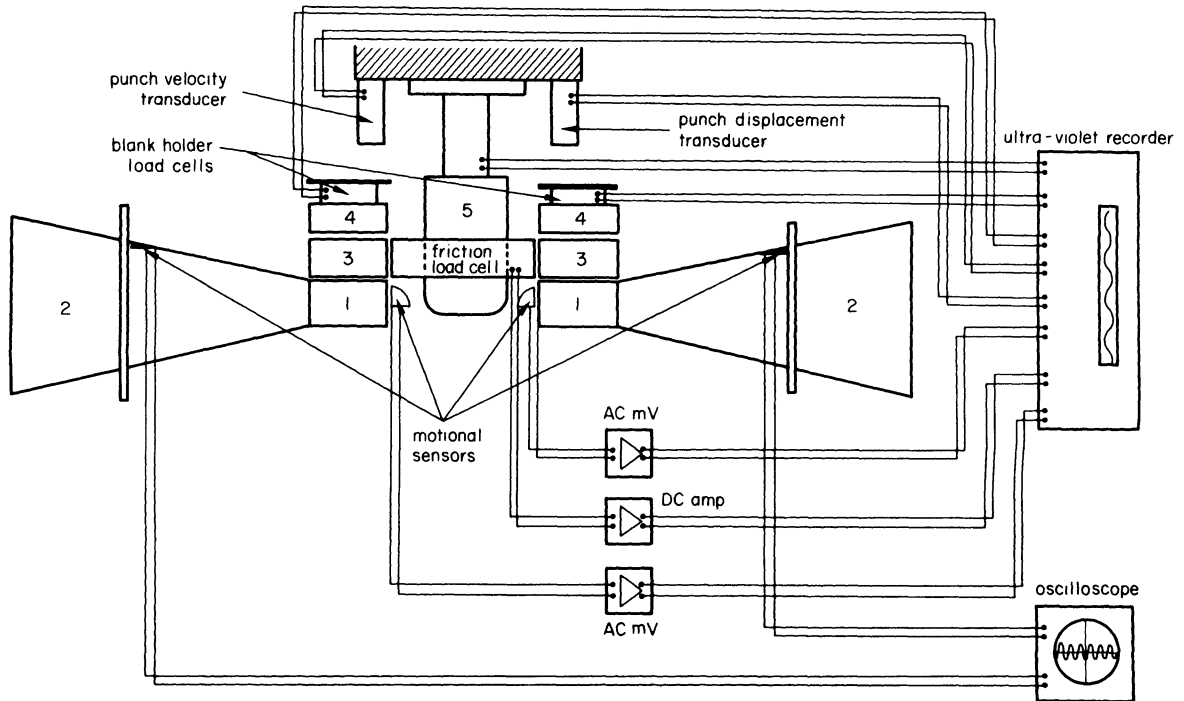


Figure 3. Schematic illustrations of the wedge-drawing subpress and instrumentation.

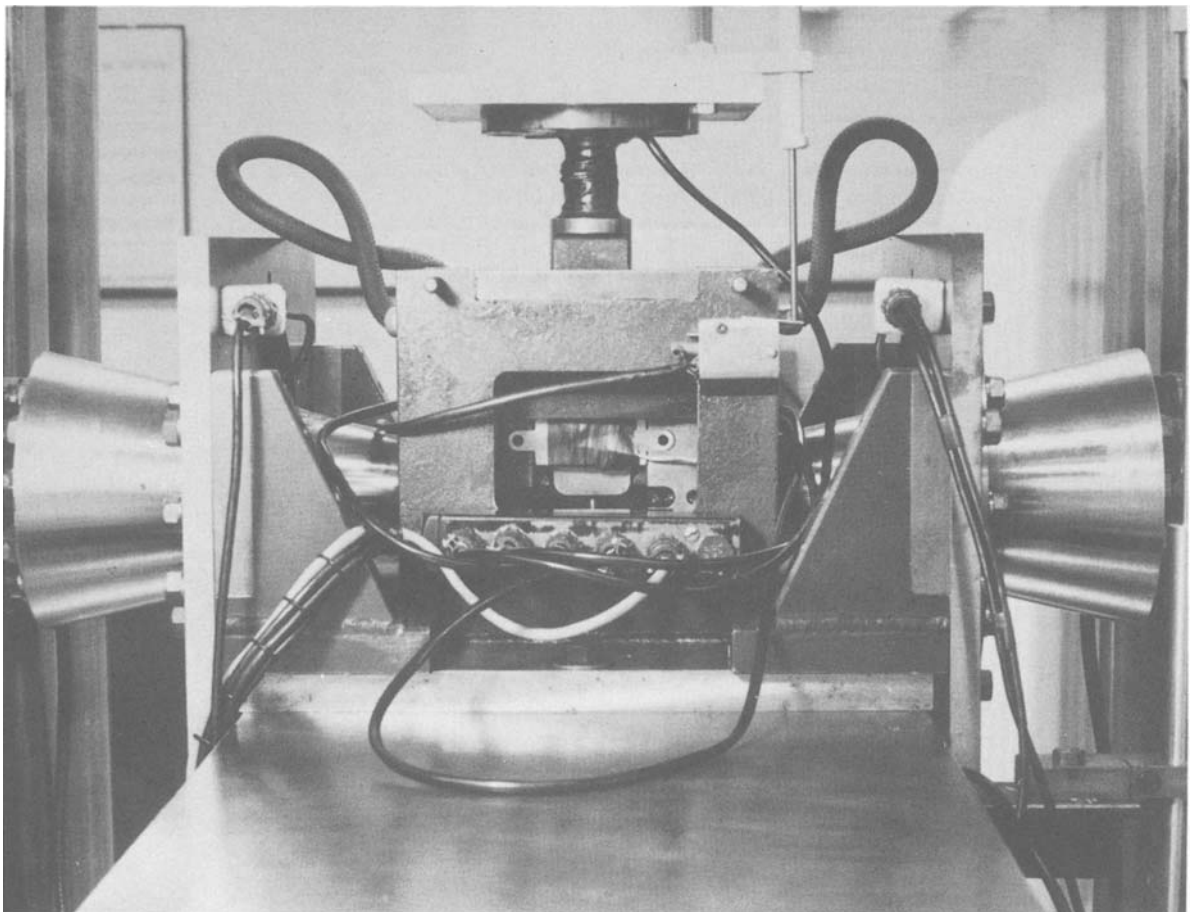


Figure 4. Oscillatory wedge-drawing test rig.

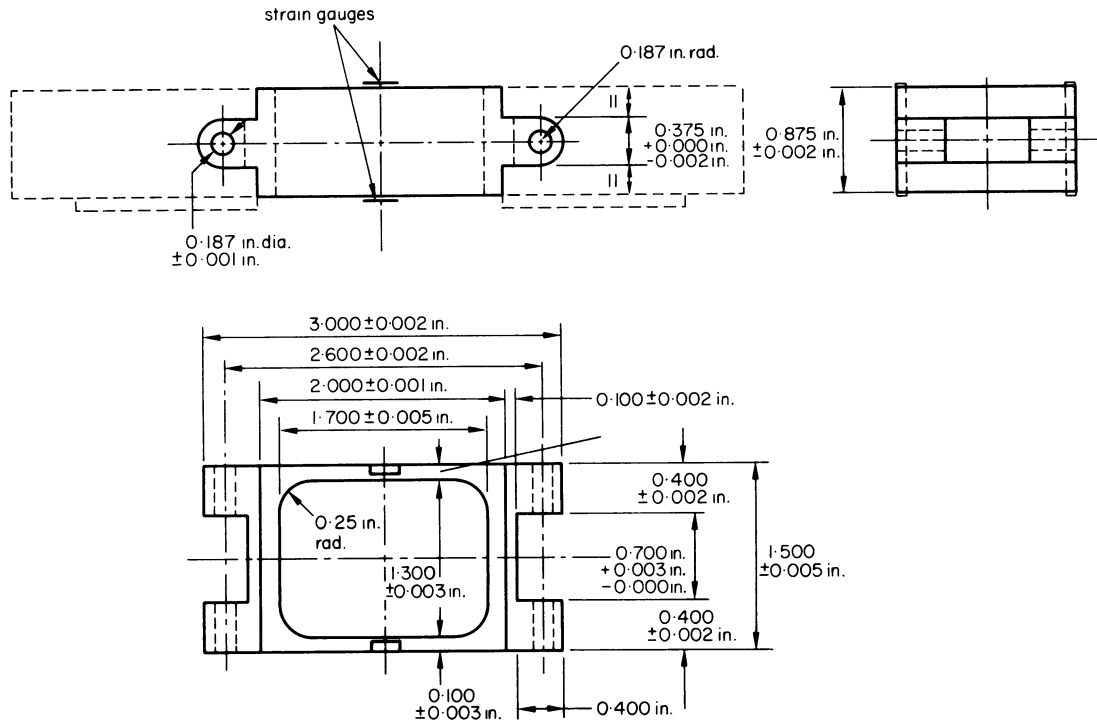


Figure 5. Blank-holder friction load cell.

test-piece interface was measured by placing a load cell between the two hold-down plates (figure 5). The whole assembly was designed to float on the upper surface of the test-piece with the load cell subjected to a compressive force due to the inward movement of the hold-down plates. Great care was taken in the design of the load cell to eliminate the effects of spurious bending and torsional distortion. This was achieved by accurate positioning of the strain gauges comprising the full bridge circuit and by providing a loose clearance fit in the locating socket between the hold-down plates and load cell.

Ultrasonic energy was provided by a 3 kW generator driving thirty-two nickel alloy magneto-strictive transducers which were attached to the conical horns by thread connections. Power input to the horns was controlled by a Variac transformer. The output frequency of the generator was variable over the range 10–14 kHz.

Instrumentation

In addition to measuring the blank-holder friction force, load cells were incorporated to measure the punch load and the hold-down force in each cylinder. Each load cell incorporated a strain gauge bridge; the output was displayed on an ultra-violet recorder. The oscillatory amplitude for each die was monitored by piezo-electric discs bonded to the free end of each drawing die. The output from these was rectified, giving a root mean square voltage which was fed into the ultra-violet recorder. Punch displacement and velocity were monitored using potentiometric and LDV transducers respectively, and the output displayed by the recorder.

Test-piece—design and manufacture

The dimensions of the wedge-drawing test-piece are shown in figure 6. With the inner radius made equal

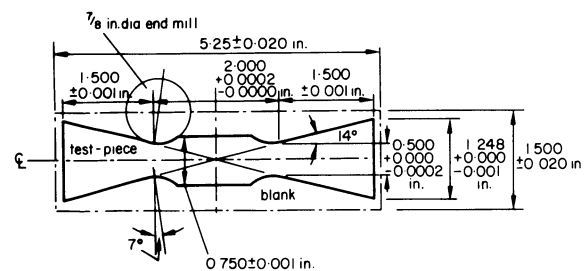


Figure 6. Dimensions of wedge-drawing test-piece.

to 1.0 in, the draw ratio R_1/R_2 was numerically equal to the outer radius R_1 .

In order to prevent uneven deformation of the wedges resulting in lateral motion of the test-piece relative to the bottom face of the punch, a $\frac{3}{8}$ in pin was attached to the base of the punch. This was located in a corresponding hole in the centre of the test-piece. Premature failure at the site of the locating hole was prevented by increasing the test piece width in this region.

The test pieces were manufactured in batches of ten on an NC milling machine. A range of draw ratios was provided by milling the outer radii to the appropriate dimension.

EXPERIMENTAL WORK

Test procedure

Each test-piece was lubricated with Macoma 275 mineral oil by dipping both ends in the lubricant and allowing the excess to drain away. The test piece was positioned so that both ends located appropriately in the drawing dies and a pre-set blank-holder force was applied through the hydraulic hold down mechanism.

The line pressure was adjusted for each draw ratio so that the initial pressure applied to the test-piece was constant. The value of pressure was determined

in preliminary tests as the minimum required to prevent wrinkling.

The majority of tests were carried out at a constant punch speed of 0.32 in/s using intermittently applied vibrations. Due to the relatively low punch speed it was possible at the onset of drawing to make minor tuning adjustments to ensure that resonance was established in both oscillatory systems. The resonant frequency of the horns was found to remain virtually constant under load and therefore further tuning adjustments were not required during the intermittent manual interruption of the ultrasonic power supply to the transducers. This was verified by repeating the tests under similar conditions but with a constant oscillatory power input: the displacement amplitude traces obtained in each case were identical within experimental variation.

Test results

Examples of typical ultra-violet recorder traces representing non-oscillatory and interrupted oscillatory tests are shown in figures 7 and 8 respectively. During interrupted tests the punch load and the friction force are seen to increase abruptly when the ultrasound is switched off and to fall immediately when the sound is switched on.

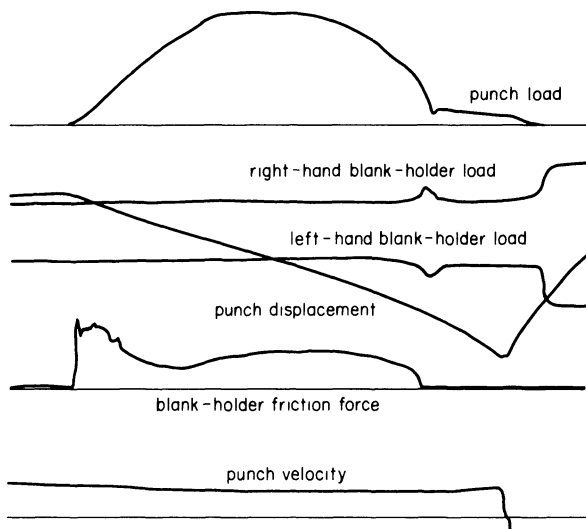


Figure 7. Ultra-violet recorder trace for non-oscillatory test.

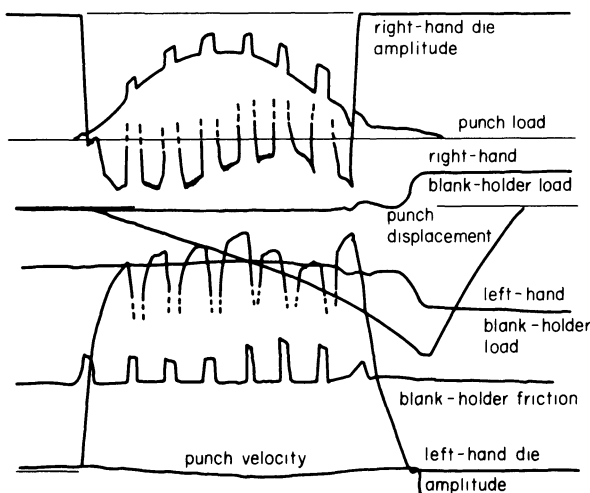


Figure 8. Ultra-violet recorder trace for interrupted oscillatory test.

oscillatory power is re-applied. Discontinuities in the die amplitude traces define the duration of the non-oscillatory intervals. By constructing an envelope of the peak values of punch load and friction force, complete traces representing both oscillatory and non-oscillatory conditions are obtained during a single test, thus permitting a direct evaluation of the effect of vibration. A total of some 800 tests were performed in which the effects of die amplitude on punch load and blank-holder friction force were examined over a range of draw ratios extending to the limit, under both oscillatory and non-oscillatory conditions.

The effect of die amplitude on blank-holder friction force is illustrated in figure 9, in which the

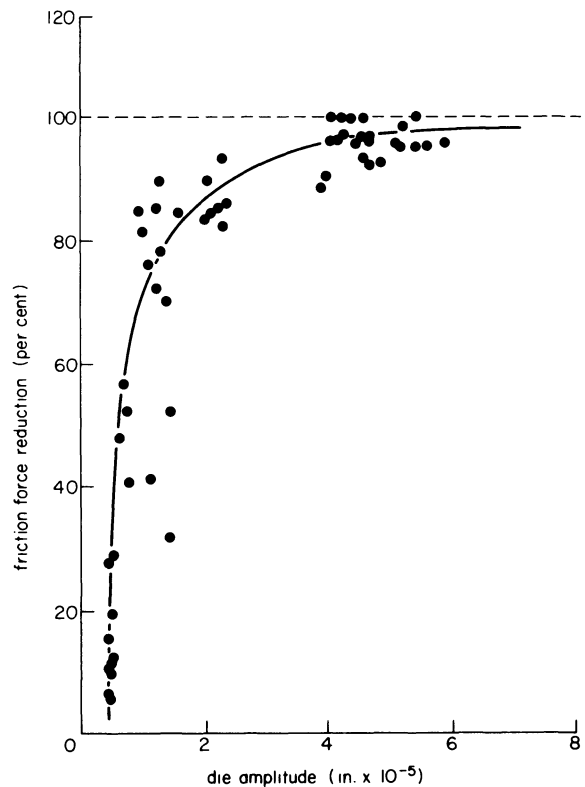


Figure 9. The effect of die amplitude on the reduction in blank holder friction force for a draw ratio of 1.70.

percentage reduction in friction force is plotted against die amplitude for a draw ratio of 1.70:1. This pattern was repeated over the range of ratios examined. The reduction in maximum punch load under the same conditions is shown in figure 10,

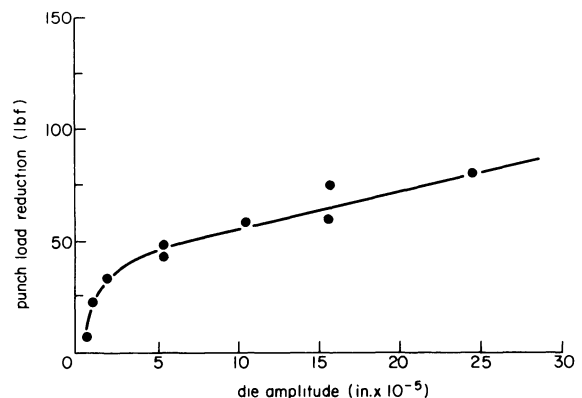


Figure 10. The effect of die amplitude on the reduction in maximum punch load for a draw ratio of 1.70.

where punch load reduction is plotted as a function of the die amplitude. Figure 11 shows the variation of maximum punch load with draw ratio under both oscillatory and non-oscillatory conditions.

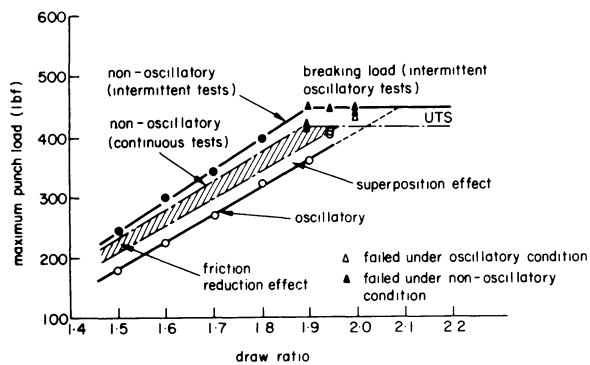


Figure 11. Variation of maximum punch load with increasing draw ratio.

GENERAL DISCUSSION

Several interesting features emerge from the graphs of friction force and punch load reduction. The reduction in blank-holder friction force with increasing die amplitude—(figure 9)—shows a marked effect at lower amplitudes. This reduces asymptotically as the amplitude tends to infinity and the reduction in friction force approaches 100 per cent. Extrapolation of the graph to zero friction reduction indicates a threshold amplitude below which no effect is produced. This was found to coincide with a velocity ratio equal

to unity, indicating that the oscillatory velocity amplitude must be greater than the punch velocity or, more accurately, the mean drawing velocity in order to affect the friction conditions.

Examination of the corresponding reduction in maximum punch load (figure 10) showed two distinct effects: a non-linear reduction in punch load at amplitudes up to 5×10^{-5} in and a linear effect at higher amplitudes up to 25×10^{-5} in, which was the limit examined in these tests.

The amplitude range of the non-linear effect was seen to coincide with the threshold and saturation values of amplitude characterizing the friction reduction effect illustrated in figure 9. The magnitude of the friction force reduction was however less than the measured value of punch load reduction which suggested that an additional effect was present at low amplitudes. It was considered likely that friction between the test blank and die profile radius would be affected under oscillatory conditions. This possibility was examined by comparing punch load-displacement curves obtained under oscillatory and non-oscillatory conditions with:

- (1) the original wedge dies, and
- (2) an assembly in which the fixed profile radius was effectively replaced by the 'frictionless' roller shown in figure 12.

This permitted an indirect measurement of the friction force associated with the motion of the test-piece around the die profile radius. Typical curves illustrating the punch load-displacement character-

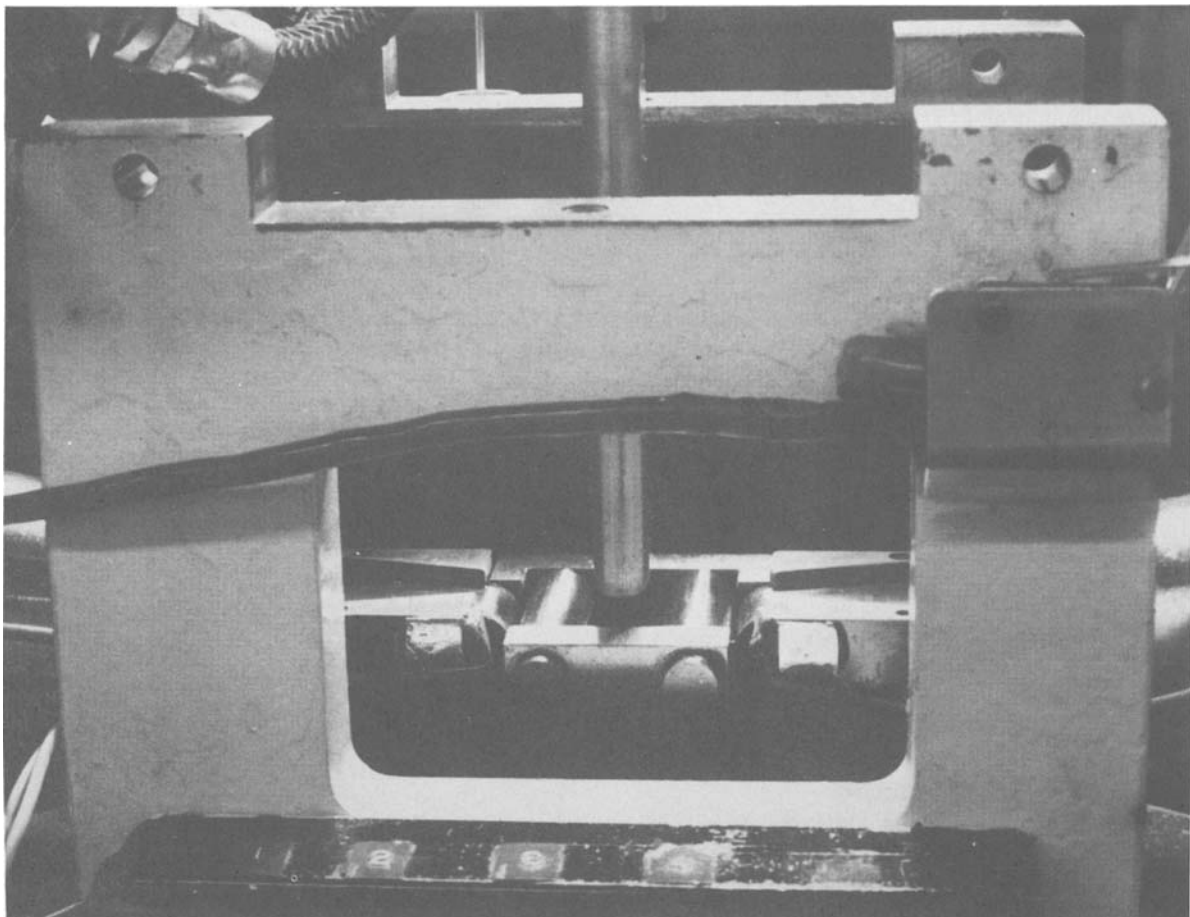


Figure 12. 'Frictionless' roller assembly.

istics under these conditions are shown in figure 13. It will be noted that the general shape of curve 2, relating to tests with the roller assembly inserted, approximates closely to that obtained with the standard dies under oscillating conditions, curve 3. The magnitude of the profile radius friction is indicated by the shaded area between curve 1 for non-oscillatory drawing through normal wedge dies, and curve 2, the corresponding relationships for 'frictionless' radius tests.

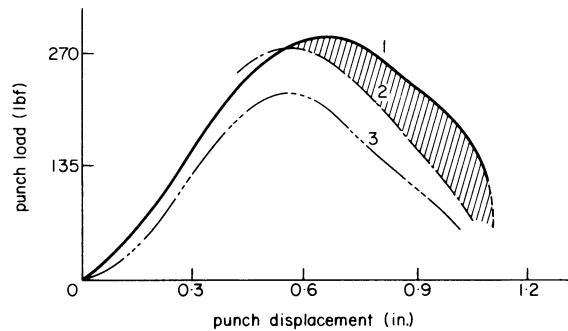


Figure 13. Punch load-displacement curves under oscillatory and non-oscillatory conditions.

The apparent increase in punch load during the initial stages of tests using the 'frictionless' radius was considered to be due to minor changes in the punch design which were necessary to accommodate the roller assembly. The die radius friction force was, therefore, determined from graphs of strip tension against punch displacement. This allowed for the effect of punch radius on the orientation of the test-piece to the punch axis and consequently upon the punch load.

Values of die radius and blank-holder friction force are summarized in Table 1 where the sum of the measured friction effects is compared with the magnitude of the non-linear punch load reduction at the saturation amplitude 5×10^{-5} in. Good agreement was found for draw ratios between 1.50 and 1.80,

TABLE 1 Comparison of the non-linear punch load reduction with measured values of blank-holder and die profile radius friction force, at the position of maximum punch load

Draw ratio	Blank-holder friction force (lbf)	Die radius friction force (lbf)	Magnitude of non-linear punch load reduction (lbf)	Sum of blank holder and die radius friction forces (lbf)
1.50	11.0	15	25	26.0
1.60	12.5	14	29	26.5
1.70	15.5	20	35	35.5
1.80	15.0	20	33	35.0

which suggested that both friction force components of the punch load were effectively eliminated at relatively low vibration amplitudes.

It is now widely accepted that oscillatory force super-position during axisymmetric drawing reduces the draw force by an amount which is directly proportional to the displacement amplitude induced in the tooling⁵. It was therefore considered probable

that the linear relationship between punch load and die amplitude observed during oscillatory wedge-drawing was attributable to this effect. However, due to the restricted access it was not possible to measure the dynamic stress in the deformation zone and this could not be confirmed experimentally.

The practical significance of the recorded punch load reduction was illustrated by determining the ratio of length to diameter under oscillatory and non-oscillatory conditions. The graph showing the relationship between maximum punch load and draw ratio (figure 11) indicated that under interrupted oscillatory tests failure occurred consistently at a draw ratio of 1.90 and above during the non-oscillatory interval.

Under continuous oscillatory conditions at a die amplitude of 20×10^{-5} in the maximum punch-load curve was displaced to the right and its slope was reduced, indicating that failure would be expected at a higher draw ratio of 2.10. However, this was not achieved in the present test series since the maximum die amplitude that was possible at the higher draw ratios was approximately 15×10^{-5} in. It was noted that failure occurred under oscillatory conditions at a higher load than the corresponding value recorded under continuous non-oscillatory tests. The latter was found to coincide closely with the UTS of the test material which suggested that the mechanical properties of the test-piece were affected by ultrasound, producing an effective increase in strength and delaying the onset of tensile instability in the drawn portion of the test piece.

It was apparent also that the maximum punch load recorded during intermittent oscillatory tests was higher than the corresponding non-oscillatory value. This was accounted for by a measured increase in the blank-holder friction force under oscillatory conditions.

A probable explanation of this phenomenon is provided by considering the effect of vibrations on the true contact area at the test-piece-blank-holder interface. If it is assumed that the relative cyclic motion at the interface leads to local heating of asperities with a consequent decrease in yield strength, then during the oscillatory interval the true contact area will increase in order to support the constant blank-holder force. When the vibrations are switched off the additional heat input ceases and the contacting regions are rapidly chilled. This raises the shear strength of the asperities and, since the increased contact area is maintained, causes an increase in the force required to shear the junctions.

Four effects are therefore illustrated by the graphs of figure 11, namely:

- (1) a reduction in punch load related to the effective removal of friction between the test-piece and blank-holder and between the test-piece and die profile radius;
- (2) a further reduction in punch load due to oscillatory force superposition;
- (3) an increase in the maximum non-oscillatory punch load which was characteristic of the intermittent oscillatory test;
- (4) an increase in the breaking load under oscillatory conditions.

It is postulated that similar effects would occur in axisymmetric cup-drawing with radial excitation of the drawing die. Although the relative magnitudes of the friction and superposition effects would be expected to differ from those observed in the wedge test, it is anticipated in cup-drawing that there would be an increase in the ratio of length to diameter. The greatest advantage would be expected when drawing thin-walled cups, since in this case a high blank-holder load is required to prevent wrinkling and the friction force would represent a higher proportion of the drawing load.

CONCLUSIONS

Wedge testing with axially activated drawing dies effectively eliminated the friction force associated with the blank-holder and die profile radius. A corresponding reduction in punch load was observed at low die amplitudes which was replaced by an additional linear effect as the die amplitude was increased. This was attributed to force superposition.

The cumulative effect of these mechanisms resulted in an increased limiting draw ratio relative to the non-oscillatory values. Furthermore, it was observed that failure occurred under oscillatory conditions at a higher load than in equivalent non-oscillatory tests, which was thought to be due to an oscillatory hardening mechanism.

It was concluded that similar effects would result from radial vibration of a cup-drawing die leading to a significant increase in the length to diameter ratio.

ACKNOWLEDGMENTS

The authors wish to thank The Science Research Council for their financial support, and The Department of Mechanical Engineering at the University of Aston where the research was undertaken.

Thanks are due also to Ultrasonics Ltd for their assistance in providing electrical generators and to Messrs H. Pratt and G. M. Jones for their help in constructing the test rig.

REFERENCES

1. C. E. Winsper and D. H. Sansome. The influence of oscillatory energy on the stresses during plastic deformation. *Journal Institute Metals* (1968) **96**, 353.
2. G. R. Dawson. *Ph.D. Thesis, University of Aston in Birmingham*, 1972.
3. G. Sachs. Ein Neues Prüfgerät für Tiefziehblech. *Metallwirtschaft* (1930), **9**, 213.
4. E. M. Loxley and H. W. Swift. The wedge drawing test. *Engineering*, (Jan. 1945), 38.
5. G. R. Dawson, C. E. Winsper and D. H. Sansome. Application of high and low frequency oscillations to the Plastic Deformation of Metals. *Metal Forming* (1970), 234.

RIGID-PLASTIC ANALYSIS OF BORE EXPANDING AND FLANGE DRAWING WITH ANISOTROPIC SHEET METALS BY THE MATRIX METHOD

by

SANG H. LEE* AND SHIRO KOBAYASHI*

SUMMARY

The matrix method of analysis for rigid-plastic materials was applied to the problems of plane-stress bore expanding and flange drawing. The analysis was performed for two materials: one in which the anisotropy exists only in the direction of sheet thickness (*normal anisotropy*); and the other possessing the two mutually perpendicular axes of anisotropy in the plane of the sheet and the other axis in the direction of thickness (*planar anisotropy*). The results revealed the effect of the state of anisotropy on the deformation characteristics in these processes, assuming that the coefficient of anisotropy remains constant and the axes of anisotropy do not rotate during the deformation.

INTRODUCTION

Many materials, employed in engineering applications, possess mechanical properties which are direction-dependent. This property, termed anisotropy, stems from a metallurgical structure of the material that depends on alloying elements and the conditions of mechanical and thermal treatments. Since the metal sheets are usually cold-rolled, it is in the sheet-metal forming process, among various plastic deformation processes, that the effect of anisotropy on the deformation characteristics is most appreciable and important.

In the past the calculation of the detailed mechanics of large plastic deformation of metal sheets has been achieved with some success by numerical methods¹⁻¹². However, without exception, these works have dealt with deformations which possess a high degree of symmetry, and, with regard to mechanical properties of sheet materials, were concerned with the anisotropy existing only in the direction of sheet thickness (*normal anisotropy*). Therefore, methods which are capable of solving nonaxisymmetric problems in sheet-metal forming have long been sought.

The finite-element method is one of the numerical methods, and it was applied to the elastic-plastic analysis of nonaxisymmetric configuration of sheet stretching with normal anisotropy by Mehta and Kobayashi¹³. Yamada¹⁴ presented a stress-strain matrix for a material which was elastically isotropic, and which obeys Hill's anisotropic yield criterion in the plastic range. He treated the incipient deformation of a circular blank of anisotropic material in the

flange drawing process. While the finite-element analysis found applications to many problems where plastic deformation is still small and of the order comparable to elastic deformation, a need for other methods, suitable for problems of large strain, became apparent because of the ineffectiveness of computation. Thus, Wang¹⁵ proposed a method of solution for problems of sheet deformation in which two spatial independent variables are required to define the geometry. The method is based on a variational procedure, and assumes that the material of the sheet is rigid-plastic and satisfies the Levy-Mises stress-strain equations. In his concluding remarks, Wang states that the success of this method relies heavily on optimization techniques for nonlinear problems containing a large number of variables. Based on the same variational principle for rigid-plastic materials, but adopting a different approach to the solution, Lee and Kobayashi¹⁶ developed a matrix method in which nonlinear relationships are reduced to linear perturbation equations. The method was applied to plane-stress bore expanding and flange drawing with a metal sheet having three mutually perpendicular axes of anisotropy^{16,17}. In these studies, emphasis was placed on demonstrating the problem-solving capability of the method and the efficiency of the computation, and on examining the accuracies of the computed solutions. Therefore, the analysis of the problems was carried out only for the range of relatively small plastic deformation.

In this paper, the mechanics of bore expanding and flange drawing are examined over the large deformation range, using the matrix method. The

* Department of Mechanical Engineering, University of California

analysis was performed for two materials: one in which the anisotropy exists only in the direction of sheet thickness (*normal anisotropy*) and the other possessing the two mutually perpendicular axes of anisotropy in the plane of the sheet with the other axis in the direction of the thickness (*planar anisotropy*, or *orthotropy*).

METHOD OF ANALYSIS

The basic concepts, of the matrix method for rigid-plastic deformation problems, were the use of the Lagrange multiplier in a variational formulation and linearization of nonlinear stiffness equations.

According to the variational principle^{1,18} for a body, V , composed of a rigid-plastic material, under boundary conditions such that the entire body is deforming plastically, the functional (Equation (1)) becomes stationary for the actual solution. This is with respect to the multiplier λ and the velocity fields which satisfy the velocity boundary conditions, but not necessarily the incompressibility condition

$$\Phi = \int_V \bar{\sigma} \dot{\bar{\epsilon}} dV + \int_V \lambda C^T \dot{\bar{\epsilon}} dV - \int_{S_T} T^T U dS \quad (1)$$

where $\bar{\sigma}$ and $\dot{\bar{\epsilon}}$ are the effective stress and effective strain rate, respectively; C is the proper matrix notation of the Kronecker delta such that $C^T \dot{\bar{\epsilon}} = 0$ implies the incompressibility condition; $\dot{\bar{\epsilon}}$, the strain-rate vector; T , the traction vector specified on the boundary S_T ; U , the velocity vector.

The formulation of the discrete variational problem follows the same procedure used in the elastic-plastic, finite-element method^{19,20}. The region, V , of the rigid-plastic material is divided into M elements interconnected at N nodal points. If $\Phi^{(m)}$ is the value of the functional Φ evaluated over the m -th element, then

$$\Phi = \sum_m^M \Phi^{(m)} \quad (2)$$

The discretization of the variational problem is performed on the elemental level by approximating the functional $\Phi^{(m)}$ by a function $\phi^{(m)}$ of the m -th element. This approximation is accomplished by replacing U with a class of kinematically complete velocity distributions which are expressed in terms of element nodal point values. With this distribution, the functional Φ is approximated by ϕ which is a function of the nodal point values

$$\begin{aligned} \Phi &\simeq \phi(u_1, u_2, \dots, u_k, \dots, u_N, \lambda^{(1)}, \dots, \lambda^{(m)}, \dots, \lambda^{(M)}) \\ &= \sum_m^M \phi^{(m)}(u^{(m)}, \lambda^{(m)}) \end{aligned} \quad (3)$$

where $u^{(m)}$ is the column matrix whose components are the velocities at nodal points associated with the m -th element; u_k is the velocity vector at the k -th nodal point and $\lambda^{(m)}$ is the Lagrange multiplier of the m -th element. Therefore, the initial variational problem is reduced to determining the set of u_k and $\lambda^{(m)}$ that gives a stationary value of ϕ . The nodal values and the elemental distributions then provide

the best approximate solution (within the class of distribution considered) to the actual distribution U , and the values of λ , together with the Levy-Mises equations, result in the corresponding stress distribution σ . Formulations of the stiffness relations, perturbation relations, and the element matrices are given in the appendix of the accompanying paper²¹.

PROBLEM STATEMENTS AND COMPUTATIONAL CONDITIONS

The problems considered are bore expanding and flange drawing under plane stress, and are shown schematically in figure 1.

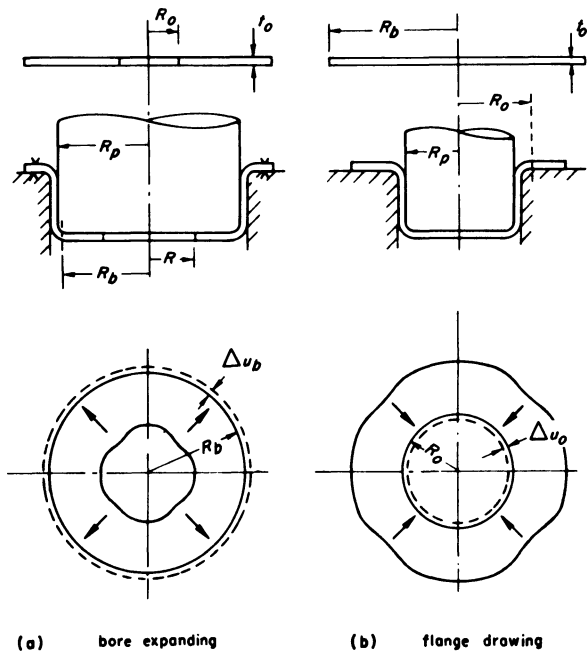


Figure 1. Processes of bore expanding and flange drawing.

Bore expanding

In the bore-expanding process a circular blank having a concentric hole is stretched radially. The stretching is accomplished by drawing the blank through a die opening, using a flat-bottomed punch (figure 1(a)). The outer periphery of the blank is fixed by a holder.

To analyze the expansion of the hole of radius R_0 in the blank of initial thickness t_0 , a gradually increasing radial displacement u_b is applied uniformly over the circular boundary of the constant radius R_b (\approx punch radius R_p). This boundary condition does not represent the actual situation exactly, but is assumed to approximate the bore-expanding process shown in figure 1(a). In the numerical procedure, the uniform incremental radial displacement Δu_b was applied along the boundary of radius R_b at each incremental deformation step. The specimen dimension selected for the analysis is $R_b/R_0 = 4.4$.

The exact solution is known for this problem if the sheet is assumed to be rigid-plastic and isotropic, and obeys the von Mises criterion. This solution was used as an initial guess of the velocity $u_{(0)}$. The computation was carried out up to a total radial

displacement of $u_b/R_0 = 0.5$ with an increment of $\Delta u_b/R_0 = 0.05$.

Flange drawing

The waviness on the rim of cylindrical cups, deep drawn from circular metal blanks cut from anisotropic sheets, is known as *earing*. Earing begins while the blank is being drawn towards the shoulder of the die (figure 1(b)). This earing process is analyzed by considering a circular sheet, of initial radius R_b , which is subjected to uniform radial displacement u_0 , towards its centre, along the circular boundary of constant radius R_0 (\approx punch radius R_p).

Similar to bore expanding, this boundary condition is imposed by applying an incremental displacement Δu_0 along the boundary of radius R_0 at each deformation step. The drawing ratio selected for the analysis was $R_b/R_0 = 2.5$. If the material is rigid-plastic and isotropic, the exact solution at the initial yielding is also known, and used as the initial guess for the first incremental solution. Once the solution for the first increment is obtained, the solution at a step is utilized as the initial guess for the next increment. The analysis was performed up to a total displacement of $u_0/R_0 = 1.33$ with an increment of $\Delta u_0/R_0 = 0.04$.

Material properties

The stress and strain property of the material used for the computation is expressed by

$$\bar{\sigma} = Y_0 \left(1 + \frac{\bar{\epsilon}}{\bar{\epsilon}_0} \right)^n \quad (4)$$

with $Y_0 = 23\,600$ psi, $\bar{\epsilon}_0 = 0.008$, $n = 0.24$, where $\bar{\sigma}$ and $\bar{\epsilon}$ are the effective stress and the effective strain, respectively. When the principal axes of anisotropy are the axes of a reference coordinate system, the effective stress, according to Hill¹, is given by

$$\bar{\sigma} \left(\frac{3}{2} \frac{(G+H)\sigma_x^2 - 2H\sigma_x\sigma_y + (H+F)\sigma_y^2 + 2N\tau_{xy}^2}{F+G+H} \right)^{1/2} \quad (5)$$

under the condition of plane stress. The parameters F , G , H , and N in Equation (5) characterize the current state of anisotropy. Frequently, the anisotropy of material is represented by the strain ratios known as *r-values*. These parameters have the following relationships:

$$\frac{G}{H} = \frac{1}{r_x}, \frac{F}{H} = \frac{1}{r_y} \quad \text{and} \quad \frac{N}{G} = \left(r_{45} + \frac{1}{2} \right) \left(1 + \frac{r_x}{r_y} \right) \quad (6)$$

where r_x , r_{45} , and r_y are the ratios of the width strain and thickness strain in simple tension of the specimen cut in the x direction, the 45° direction to the x axis, and in the y direction, respectively. The x reference axis is chosen to be parallel to the direction

of rolling of the sheet. The corresponding effective strain rate is given by

$$\dot{\bar{\epsilon}} = \left(\frac{2}{3} (F+G+H) \times \left(\frac{(H+G)\dot{\epsilon}_y^2 + (F+H)\dot{\epsilon}_x^2 + 2H\dot{\epsilon}_x\dot{\epsilon}_y}{(FG+GH+HF)} \right) + \frac{\dot{\gamma}_{xy}^2}{2N} \right)^{1/2} \quad (7)$$

and the effective strain in Equation (4) is defined by

$$\bar{\epsilon} = \int \dot{\bar{\epsilon}} dt \quad (8)$$

Two materials were considered for the calculation. One is the material of planar anisotropy whose r -values are given by $r_x = 1.69$, $r_{45} = 1.43$, and $r_y = 2.24$, which correspond to those of A1-killed steel¹⁷, and the other is a material of normal anisotropy with $r_{ave} = \frac{1}{4} (r_x + 2r_{45} + r_y) = 1.70$ and having the identical effective stress-strain curve given by Equation (4).

It must be noted that in simple tension under the given anisotropy the yield stress has its maximum in the direction α_y to the rolling direction such that

$$\tan^2 \alpha_y = \frac{r_{45}(r_x + r_y) - 2r_x r_y + \frac{1}{2}(r_x - r_y)}{r_{45}(r_x + r_y) - 2r_x r_y - \frac{1}{2}(r_x - r_y)}$$

and the r -value attains its minimum in the direction α_r given by

$$\sin^2 \alpha_r = \frac{1}{a} \{ b - (b^2 - ac)^{1/2} \}$$

where

$$a = \frac{1}{r_x} - \frac{1}{r_y}, \quad b = \frac{1}{r_x},$$

$$c = \frac{1}{r_x} - \frac{1/r_y - 1/r_x}{2r_{45}(1/r_x + 1/r_y) - 4}$$

With the r -values in the present calculation, $\alpha_y = 49^\circ$ and $\alpha_r = 39^\circ$.

RESULTS AND DISCUSSION

Bore expanding

Load-displacement relations show that for both materials the load increases first and then decreases, indicating the maximum at $R/R_0 = 1.3$. The load is the resultant radial force acting along the boundary of radius R_b , and may be considered approximately as the punch load. The results also show that the simplification of *planar* anisotropy by *normal* anisotropy underestimates the punch load by approximately 5 per cent.

The total radial displacements are shown in figure 2(a). The trends of the distributions are almost the same for both states of anisotropy. The total radial displacement along the bore periphery indicates that the noncircular expansion of the bore due to the planar anisotropy is ineffective and the bore remains closely round after expansion. However, the complexity of deformation due to planar anisotropy

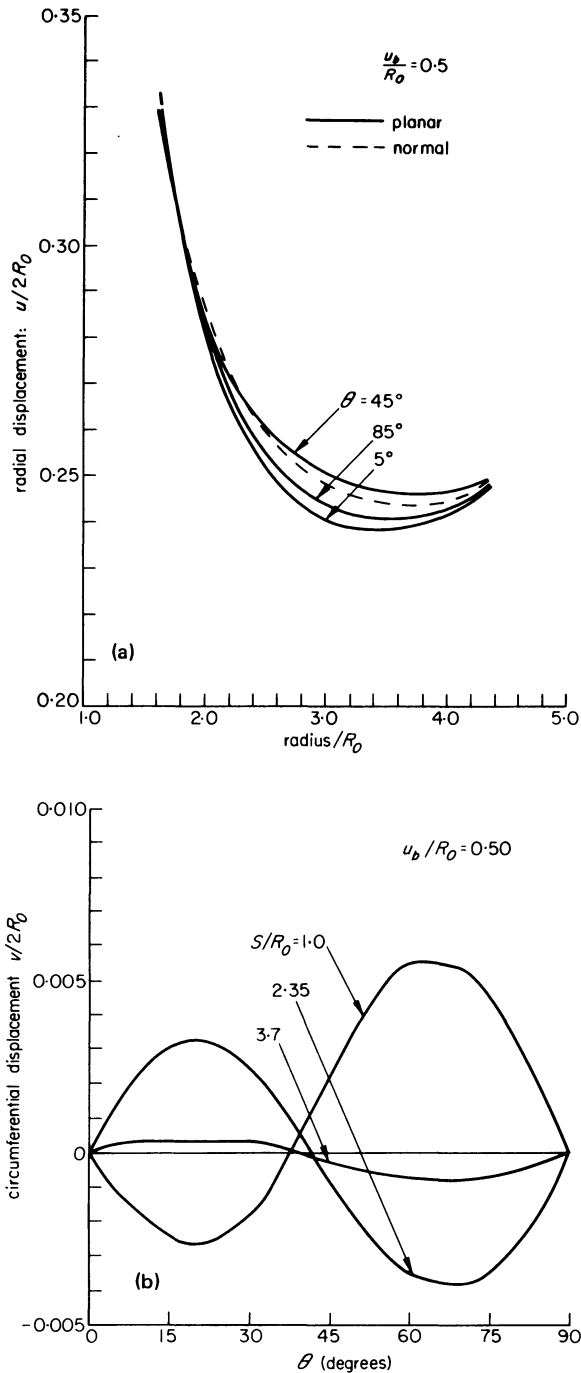


Figure 2. (a) Radial displacement distributions as functions of the radius (b) The circumferential displacement distributions.

becomes evident when the circumferential displacements in figure 2(b) are examined.

For normal anisotropy, the circumferential displacement is zero throughout the sheet. For planar anisotropy, however, non-zero circumferential displacements are observed, although the magnitude of the displacements is much smaller than that of the radial displacement. The angle θ in figure 2(b) is measured counterclockwise from the direction of the x axis (rolling direction). Along the bore periphery $S/R_0 = 1.0$ (S is the radial position of an element in the initial blank), the circumferential displacement is zero at $\theta = 37^\circ$ and is clockwise for smaller angles and counterclockwise for larger angles. However, for the

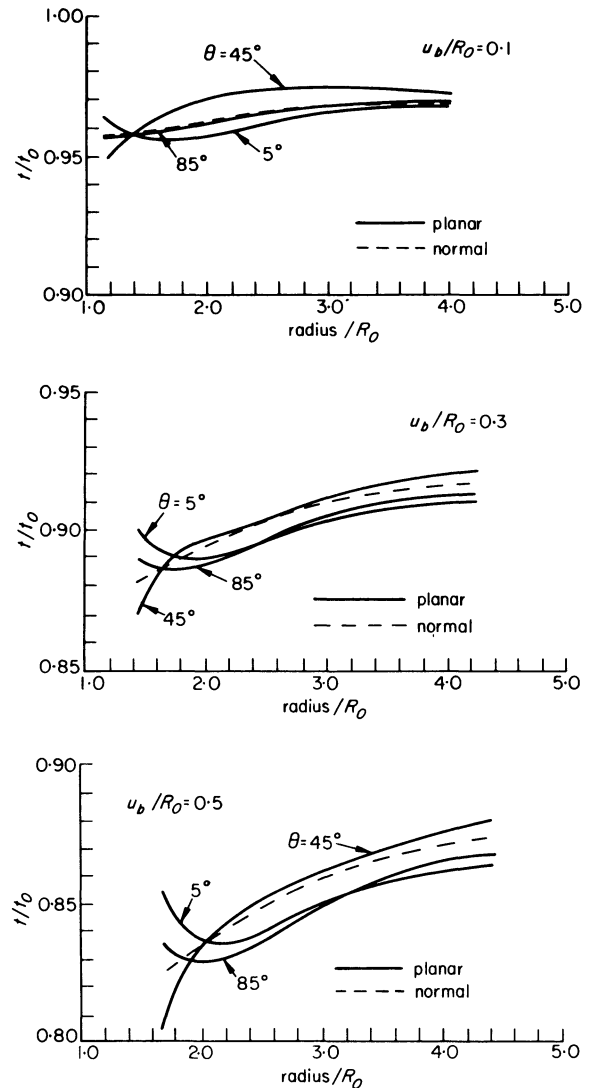


Figure 3. Distributions of sheet thickness in bore expanding.

element inside the blank corresponding to $S/R_0 = 2.35$, the direction of the displacement is reversed, with zero displacement at $\theta = 41^\circ$. Along the bore periphery the stress state is uniaxial tension. Therefore, the direction for maximum yield stress ($\alpha \equiv 49^\circ$) is $\theta = 90^\circ - \alpha_y = 41^\circ$, and the direction for r_{\min} ($\alpha_r = 39^\circ$) is $\theta = 90^\circ - \alpha_r = 51^\circ$. It may be of interest to compare these values with $\theta = 52^\circ$ for minimum radial displacement, and $\theta = 37^\circ$ for zero circumferential displacement.

Figure 3 shows the thickness distributions. The *normal* anisotropic solution has the thickness increasing smoothly towards the outer periphery, being thinnest at the bore. For *planar* anisotropy the same trend of the distribution is found along $\theta = 45^\circ$, but the gradient is much steeper near the bore. The magnitude is almost the same except at the bore, where *planar* anisotropic solution indicates more thinning than for *normal* anisotropy. Along the elements in the directions of anisotropic axes ($\theta = 5^\circ$, $\theta = 85^\circ$), however, the maximum thinning occurs some distance away from the bore. Since the stress state is uniaxial tension along the bore periphery, the maximum thinning should occur in the direction of minimum r -value ($\theta = 51^\circ$). Toward the outer peri-

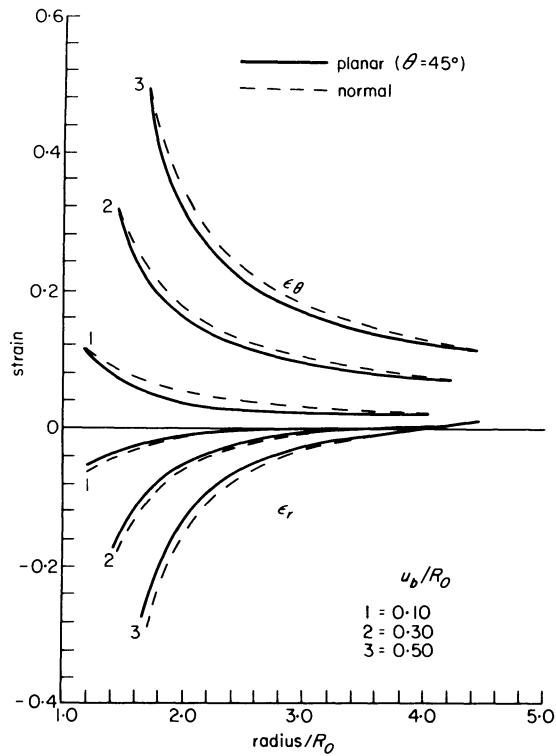


Figure 4. Strain distributions at various deformation stages of bore expanding.

phery in that direction, however, the thickness becomes larger than those in other directions. This nonuniform distribution of thickness due to *planar* anisotropy near the bore should have significant implications on the determination of forming limits, since the localized neck and eventual fracture occurs in this region in bore expanding.

Although figure 3 reveals considerable differences in the distributions of thickness, the difference in the distribution of strain components does not appear to be as drastic. In figure 4 the strain distributions for *normal* anisotropy, and those in the direction $\theta = 45^\circ$ for *planar* anisotropy are compared. The distributions

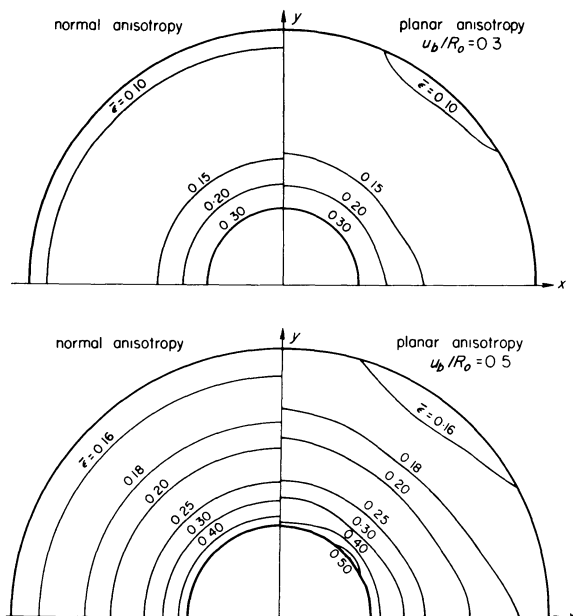


Figure 5. The effective strain contours in bore expanding.

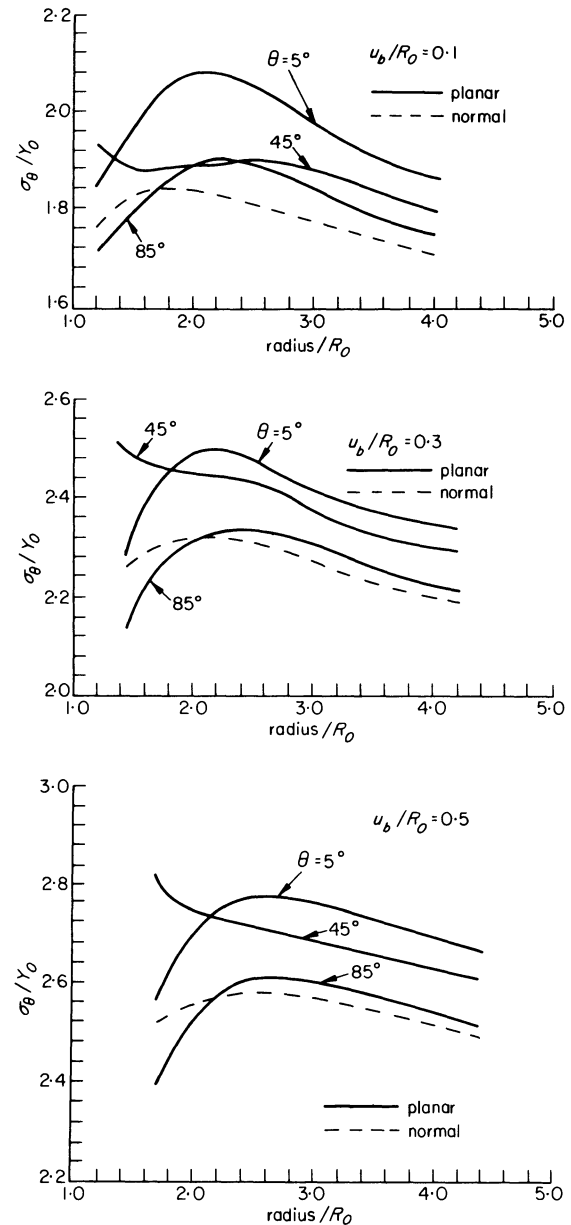


Figure 6. Distributions of the circumferential stress component at various deformation stages of bore expanding.

in the directions of $\theta = 5^\circ$ and $\theta = 85^\circ$ are almost identical to those for *normal* anisotropy. Non-uniformity of the strain distributions in the circumferential direction, due to *planar* anisotropy, is clearly seen in figure 5 where the effective strain contours are compared for the two cases at various deformation states. It is observed that the effect of *planar* anisotropy, on the effective strain distribution, is minor in bore expanding. Figure 6 shows the circumferential stress (σ_θ) distributions. The distributions for *normal* anisotropy, and those at $\theta = 5^\circ$ and 85° for *planar* anisotropy, are similar. The stress increases first and then decreases from the bore periphery towards the outer boundary, indicating the maximum is some distance way from the bore.

The stress distribution at $\theta = 45^\circ$ is interesting. The greatest stress occurs at the bore, and the stress distribution shows the minimum and the maximum along the radial direction, at early stages of the process. As the deformation continues, the maximum

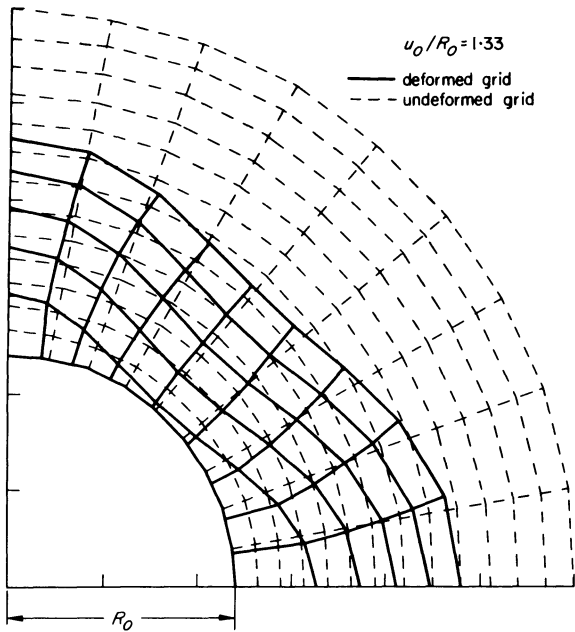


Figure 7. Deformed grid pattern in flange drawing.

and minimum disappear, and the stress decreases monotonously toward the outer boundary of the blank. Again, along the bore periphery the maximum stress occurs at the direction of maximum yield stress $\theta = 41^\circ$.

Flange drawing

As for bore expanding, the load is higher for *planar* anisotropy. The load was taken as the resultant radial force acting along the boundary of radius R_0 . A general picture of the complex deformation involved in the formation of ears and hollows can be seen in the grid-distortion in flange drawing, shown in figure 7. It is evident that in the region near the inner boundary the circumferential displacement is negative (clockwise) near the x axis and becomes positive near the y axis. It is also clear that the direction of the circumferential displacement is reversed near the outer boundary. This variation of the circumferential displacements is qualitatively the same as that observed in bore expanding (figure 2(b)). Furthermore, the formation of ears and hollows is clearly visible in figure 7. The ears appear along the axes of anisotropy and a hollow is formed approximately at $\theta = 47^\circ$ at an early stage, and changes its location slightly to $\theta = 49^\circ$ near the end of the process. According to Hill, the ears are formed at $\theta = 0^\circ$ and 90° , and a hollow at $(90^\circ - \bar{\alpha}) = 41^\circ$. The location of hollow, shown in figure 7, is not in agreement with Hill's prediction, but is closer to that for minimum r -value ($\theta = 51^\circ$).

The thickness distributions are given in figure 8. For normal anisotropy, the thickness increases toward the rim of the flange, and the flange becomes thicker than the initial blank as the drawing process proceeds. The effect of *planar* anisotropy on the

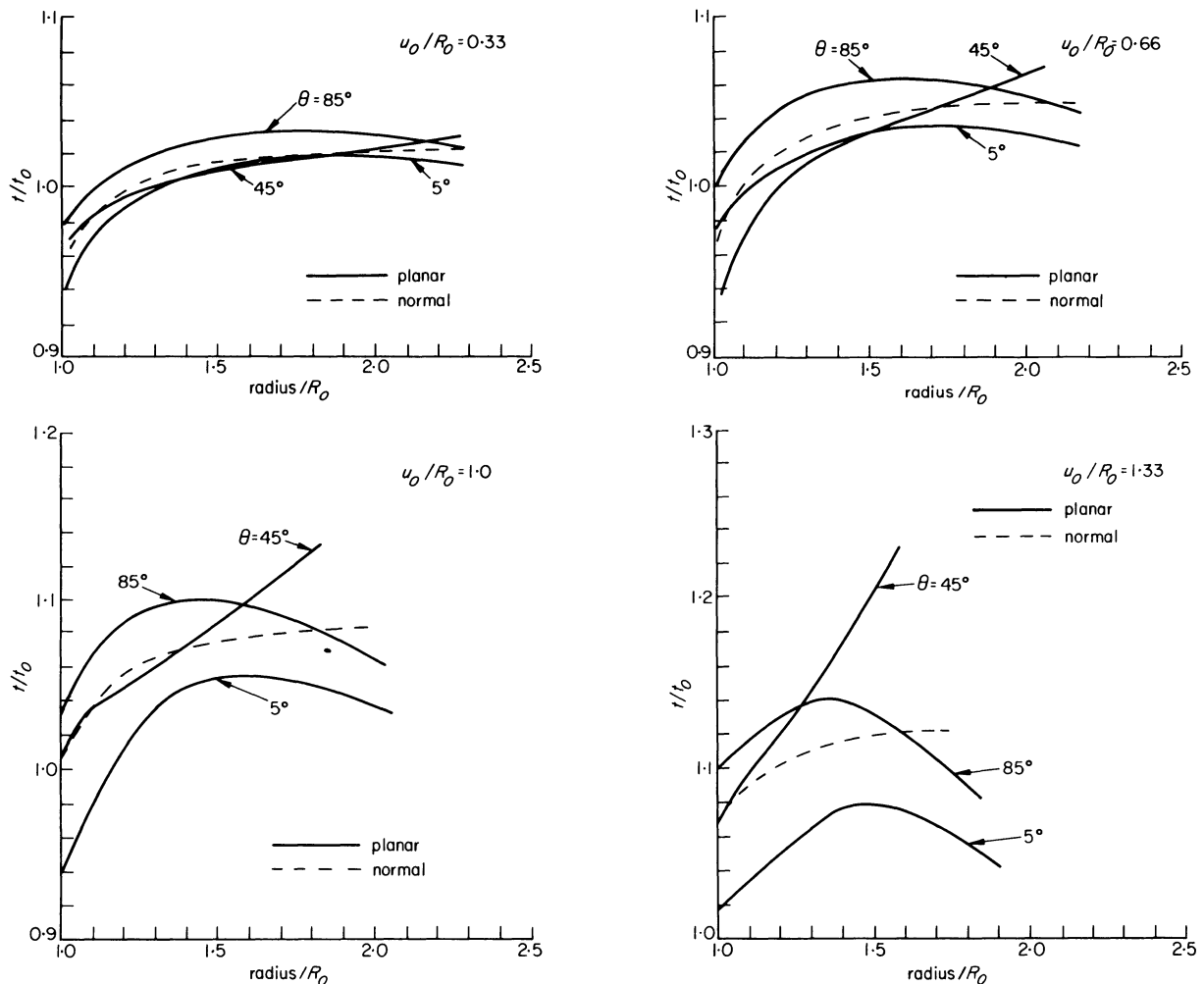


Figure 8. Thickness distributions in flange drawing.

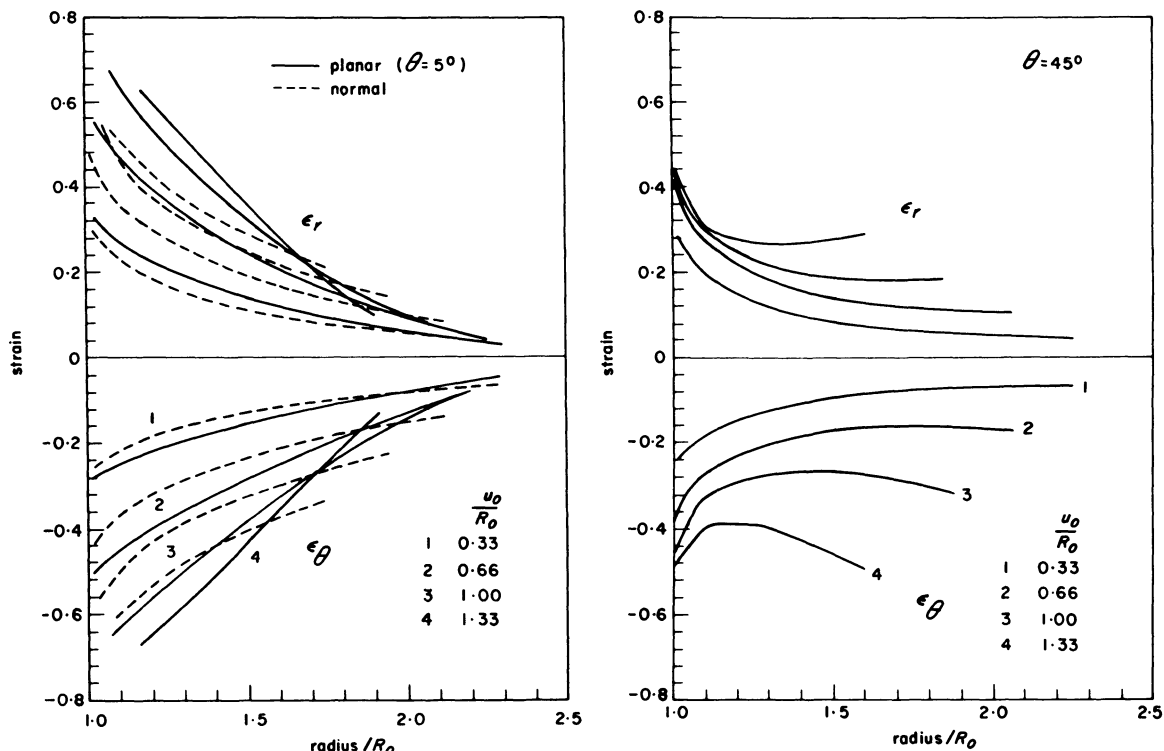


Figure 9. Strain distributions in flange drawing.

thickness distributions is greater in flange drawing than in bore expanding. Along the axes of anisotropy, as the deformation increases, the thickness distributions begin to show the maximum; first increasing from the inner boundary and then decreasing toward the rim of the flange. Along $\theta = 45^\circ$, the distribution is completely different from those at $\theta = 5^\circ$ and 85° . The thickness increases toward the rim as it does for the *normal* anisotropy, but the gradient is much larger. Along the rim of the flange, the maximum thickening occurs in the direction of minimum r -value, $\theta = 51^\circ$. In flange drawing, when the blank holder is used, the contact between the flange and the holder begins along the direction of minimum r -value rather than along the rim of the flange. Since the friction due to the contact between the blank and the holder affects the subsequent deformation, it is likely that the appearance of the ears and hollows would be less pronounced when the blank holder is used.

The strain distributions (ϵ_r and ϵ_θ) in the directions $\theta = 5^\circ$ and $\theta = 85^\circ$ are almost identical, and the comparison with the *normal* anisotropic solution is shown in figure 9. Except for minor differences in magnitude, the distributions are very similar for both cases of anisotropy. However, the strain distributions in the $\theta = 45^\circ$ direction are quite different from others.

Following a large change in strain near the inner boundary, the strains remain more or less uniform over the rest of the flange. The effective strain distributions, shown in figure 10, appear to reflect the distributions of ϵ_r in figure 9. Along the axes of anisotropy the effective strain decreases toward the rim of the flange, and the distributions are approximately the same as for *normal* anisotropy. Along $\theta = 45^\circ$, the strain decreases near the inner boundary in a

similar manner, but then the variation of the strains remains small toward the rim.

The distributions of the circumferential stress for both cases of anisotropy are given in figure 11. Again, the distributions along the axes of anisotropy, in the *planar* anisotropic case, are the same in magnitude, as well as in trend, as those for the *normal* anisotropic case. Difference between the two cases appears in the stress distribution along $\theta = 45^\circ$, where the stresses are much larger compressive stresses, and the gradients of the distribution are steeper.

CONCLUSIONS

The application of the matrix method to the plane stress problem, taking into account the planar anisotropy, revealed some interesting results.

In bore expanding, the circular hole changes its shape as it expands due to *planar* anisotropy. The noncircular expansion, however, is ineffective and the bore remains closely round, even after considerable expansion. Comparing the solutions for *planar* anisotropy and for *normal* anisotropy, the effect of the

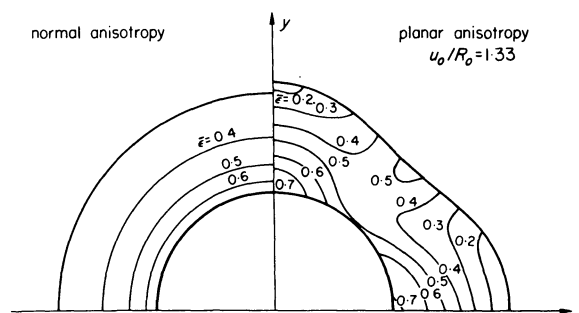


Figure 10. The effective strain contours in flange drawing.

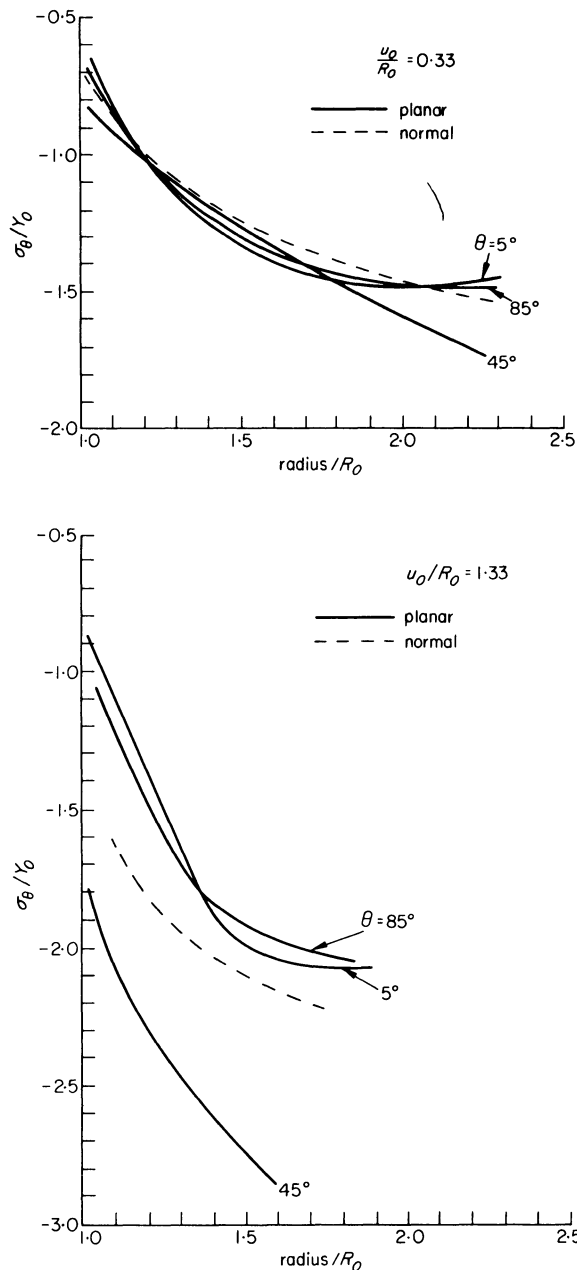


Figure 11. Distributions of the circumferential stress component.

former, on the stress and strain distributions, was markedly exhibited in the region near the bore.

In flange drawing, the solutions for *planar* anisotropy are similar to those for *normal* anisotropy along the anisotropic axes. The considerable effect of *planar* anisotropy was indicated in the solutions along the 45° direction to the axes of anisotropy. Also, the method enabled us to analyze the formation of ears and hollows along the rim of the sheet blank. It was revealed that the radial direction of a hollow did not coincide with that of the principal strain-rate axis.

In general, *planar* anisotropy has a more significant effect on the deformation in flange drawing than in bore expanding. A striking result in both problems was the effect of *planar* anisotropy on the sheet thickness distributions. The thickness varies considerably not only in the radial direction but also in the circumferential direction. These nonuniform

thickness distributions, due to *planar* anisotropy, appear to have important implications on the occurrence of instability, such as necking in bore expanding and wrinkling in flange drawing.

ACKNOWLEDGMENTS

The authors wish to thank the National Science Foundation for its grant GH-37120, and the General Motors corporation for its grant, under which the present work was possible. They also wish to thank Dr. C. H. Lee for his continued assistance in carrying out the work.

REFERENCES

1. R. Hill. *Mathematical Theory of Plasticity*. Oxford University Press, 1950.
2. S. Y. Chung and H. W. Swift. Cup-drawing from a flat blank: part I, experimental investigation; part II, analytical investigation. *Proceedings of Inst. of Mech. Engrs.* (1951) 165, 199.
3. Y. Yamada. Studies on formability of sheet metals. *Rep. Inst. Ind. Sci., Tokyo University* (1961) 11, 240.
4. D. M. Woo. Analysis of cup-drawing process. *J. Mech. Eng. Science* (1964) 6, 116.
5. D. M. Woo. The analysis of axisymmetric forming of sheet metal and the hydrostatic bulging process. *Int. J. Mech. Sci.* (1964) 6, 303.
6. D. M. Woo. The stretch-forming test. *The Engineer* (1965) 220, 876.
7. D. C. Chiang and S. Kobayashi. The effect of anisotropy and work-hardening characteristics on the stress and strain distribution in deep drawing. *J. Engrg. Ind., Trans. ASME* (1966) 88, 101.
8. B. Budiansky and N. M. Wang. On the swift cup test. *J. of Mechs. Phys. Solids* (1966) 14, 357.
9. B. Storakers. Finite plastic deformation of a circular membrane under hydrostatic pressure. *Int. J. Mech. Sci.* (1966) 8, 619.
10. Y. Yamada and M. Koide. Analysis of the bore-expanding test by the incremental theory of plasticity. *Int. J. Mech. Sci.* (1968) 10, 1.
11. N. M. Wang and M. R. Shammany. On the plastic bulging of a circular diaphragm by hydrostatic pressure. *J. of Mechs. and Phys. of Solids* (1969) 17, 43.
12. N. M. Wang. Large plastic deformation of a circular sheet caused by punch stretching. *General Motors Research Publication GMR-862*, 1969.
13. H. S. Mehta and S. Kobayashi. Finite-element analysis and experimental investigation of sheet-metal stretching. *Rep. No. MD 71-2, University of California, Berkeley*, 1971.
14. Y. Yamada. Recent Japanese developments in matrix displacement method for elasto-plastic problems. *Paper presented at Japan-U.S. Seminar on Matrix Methods of Structural Analysis and Design, Tokyo, Japan*, 1969.
15. N. M. Wang. A variational method for problems of large plastic deformation of metal sheets. *General Motors Research Publication GMR-1038*, 1970.
16. C. H. Lee and S. Kobayashi. New solutions to

- rigid-plastic deformation problems using a matrix method. *J. Engrg. Ind., Trans. ASME* (1973) **95**, 865.
17. S. H. Lee, C. H. Lee, and Shiro Kobayashi. Analysis of Plane-stress bore expanding with anisotropic sheet metals. *Report to the Mathematics Department, Research Laboratories, General Motors Technical Center, Warren, Michigan*, 1972.
 18. K. Washizu. *Variational Methods in Elasticity and Plasticity*. Pergamon Press, London, 1950.
 19. C. H. Lee and Shiro Kobayashi. Elastoplastic analysis of plane-strain and axisymmetric flat punch indentation by the finite-element method. *Int. J. Mech. Sci.* (1970) **12**, 349.
 20. C. H. Lee and Shiro Kobayashi. Analysis of axisymmetric upsetting and plane-strain sidepressing of solid cylinders by the finite-element method. *J. Engrg. Ind., Trans. ASME, Series B* (1971) **93**, 445.
 21. S. N. Shah and Shiro Kobayashi. Rigid-plastic analysis of cold heading by the matrix method. To be presented at the 15th International Conference of MTDR, Birmingham, Sept. 1974.

TUBE INVERSION: A PROCESS FOR OBTAINING GROSS EXPANSIONS IN THE DIAMETER OF THIN-WALL TUBES

by

S. T. S. AL-HASSANI*

SUMMARY

Circumferential strains much in excess of the instability strains in internally pressurized tubes may be obtained by axially compressing the tube over a radiused die. As the tube walls slide over the curved parts of the die, simultaneous bending, stretching and unbending takes place. The resulting tube is of larger diameter, same wall thickness, and shorter in length.

The phenomena associated with the inversion process of strain-hardening tubes are described and load-displacement curves are presented.

Experimental results to establish the optimum radial expansion are given and conditions for buckling and tearing are identified.

A simple theoretical analysis of the process which allows for strain-hardening is given.

NOTATION

A, B and n	Strain-hardening constants.
D	Diameter of tube.
E_t	Tangent modulus.
P	Inverting load.
W	Plastic work done.
Y	Yield stress.
r	die radius.
t	Wall thickness.
v	Velocity.
$\bar{\sigma}$	Representative stress.
$\bar{\epsilon}$	Representative strain.
$\dot{\epsilon}$	Strain rate.

Suffixes

B	Bending.
E	Extensional.
m	Mean.

INTRODUCTION

In simple internal pressurization of a thin-wall tube, instability is reached at a hoop strain given by¹ $\epsilon_\theta = \frac{1}{2}[n - \sqrt{(3)B}]$, where n and B are the strain-hardening constants of a material whose representative stress-strain curve is given by $\bar{\sigma} = A(B + \bar{\epsilon})^n$.

Usually, in most metals other than super-plastic, the instability strain is quite small and it becomes rather difficult to expand the diameter of a tube without gross thinning and subsequent fracturing of the wall.

In this paper a process which obtains large diametral expansions much in excess of the instability strain is described. Basically this is a process which allows the tube to be turned inside-out by pushing it over a die which contains a fillet radius as shown in figure 1. With a die of an appropriate fillet radius, the

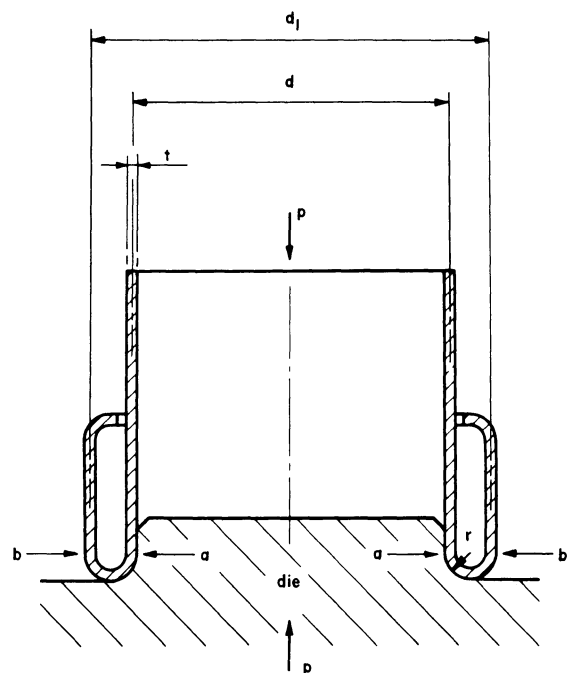


Figure 1. Line diagram showing the configuration of an inverted tube.

* Department of Mechanical Engineering, University of Manchester Institute of Science and Technology.

tube wall slides over the curved part of the die and, of its own accord, inverts back on itself resulting in a tube with a large diameter. The original inside wall now is on the outside.

Of course inversion is not the only mode of deformation possible. When the fillet radius is smaller, uniform compression develops which may well cause a brittle fracture or buckling of the tube. Three possible buckling modes, depending upon the tube's dimensions, are possible. Diamond, concertina (axisymmetric) and column (Euler) buckling modes are shown in figure 2. For a larger die radius, the tube's tip tears before it can roll over the die as shown in figure 2.

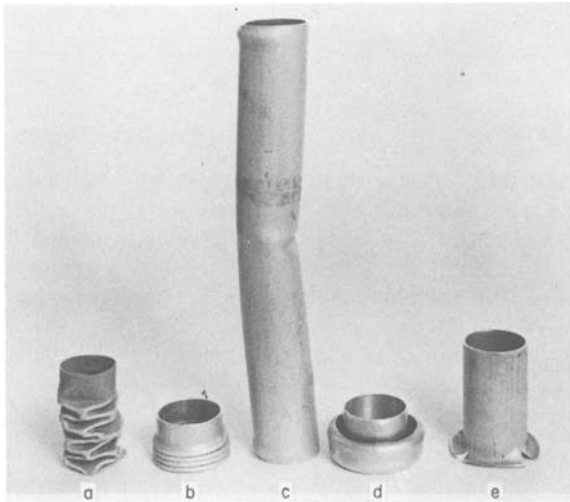


Figure 2. Photograph of possible modes of deformation: (a) Diamond buckling, (b) Concertina or axisymmetric buckling, (c) Column or Euler buckling, (d) Inversion, (e) Tearing.

The process of turning thin-wall tubes inside-out or outside-in is popular in the field of safety of vehicles and other fields where steady dissipation of kinetic energy during impact is required. Recently Mellor² employed an inside-out inversion tube to act as a load limiter in compression tests. The main feature of the limiter is that it deforms only when the compressive force pushing the specimen exceeds the inverting load.

So far this process has been considered in relation to the design of collapsible steering columns, for

facilitating lunar and planetary air drop cargo, design of buffer elements for nuclear-reactor control rods, and for aircraft seat anchors. The particular feature which made this process so popular in the field of safety is the rectangular force-displacement characteristic which ensures a uniform retardation of the colliding object. An investigation into the important parameters of inversion tubes and a detailed analysis is given by Al-Hassani, Johnson and Lowe³ and related and subsequent work is summarized by Johnson⁴.

It is intended, in this paper, to focus attention on the metal-forming aspect of the tube inversion process and to reproduce the analysis in a manner useful to the metal-forming engineer. The analysis is, then, used to identify the conditions for successful inversion.

THE INVERSION PROCESS

A ciné film is taken of the tube as it slides and turns over when axially compressed by an Instron cross-head moving with a constant speed of 2 in/min. A typical load-shortening curve for a half-hard aluminium tube, 2.0-in o.d., 0.064-in wall thickness, 3.5-in long, inverted over a die with a $\frac{5}{32}$ -in radius is given in figure 3. The shape of the tube at each stage, from A to G in figure 3, is given by the photographs shown in figure 4. In this figure, each tube represents

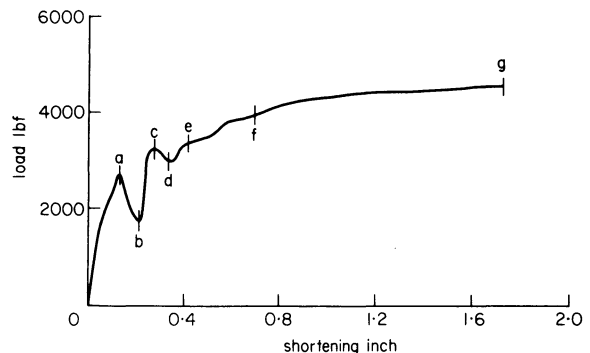


Figure 3. Load-shortening curve for a half-hard aluminium tube 2.0-in o.d., 0.064-in wall thickness and 3.5-in long, inverting over a die with $\frac{5}{32}$ -in fillet radius; letters refer to figure 4.

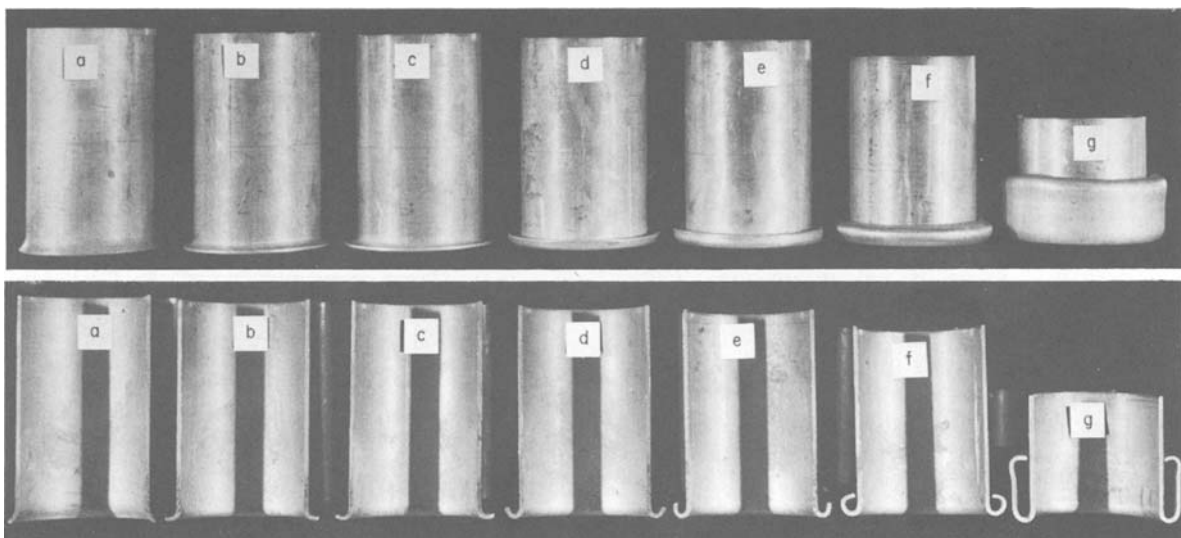


Figure 4. Stages of inversion; letters refer to figure 3.

a separate test with a shortening up to the stage indicated by the corresponding letter in figure 3. These photographs are identical to those identified in the ciné film. Also shown in figure 4 are sectioned tubes at each stage of inversion.

It is seen that the load initially rises steeply up to a maximum at A, when the tube begins to flare into a funnel shape. After the peak, there is a rapid fall in load accompanied by full conformity to the die. When the radius is fully formed on the tube, the load again rises as the tube is pushed around the die. A second peak, C, is then reached, and from this stage the flared end begins to roll up. As the tube completes a movement of 180° around the die radius, a further fall in the load is noted at D. Subsequently the load rises to a maximum steady state value when the tube tip has traversed a total angle of 270° as shown by section F in figure 3. The curling of the tube is terminated just before contact is made with the undeformed surface of the straight tube. During the final equilibrium part of the process, the tube continues to flow around the radius, but straightens to form another parallel-sided tube outside the original tube. The fully inverted tube is indicated by stage G.

RESULTS OF TESTS ON THE MAJOR FEATURES OF THE PROCESS

Results of inversion tests on half-hard aluminium tubes of several sizes showed that, provided an appropriate die fillet radius is used, inversion is possible. Figure 5 shows 2.0-in, 2.5-in and 7.0-in o.d.

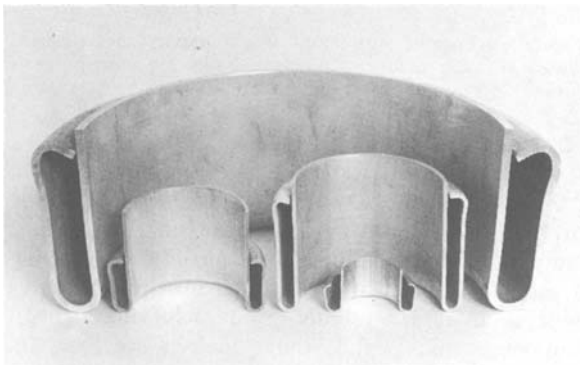


Figure 5. Sectioned Aluminium inverted tubes of 1.0, 2.0, 2.5 and 7.0-in outside diameter.

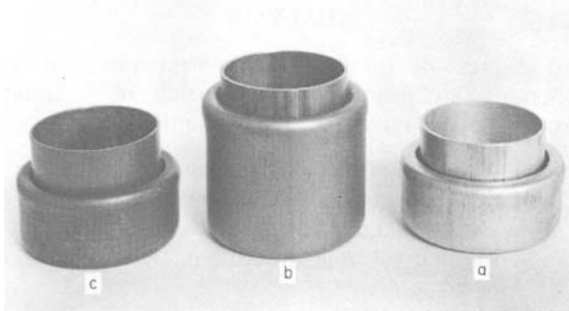


Figure 6. Inverted 2.0-in and 0.064-in wall thickness aluminium (A), brass (B) and copper (C) tubes.

half-hard aluminium tubes sectioned after inversion. A common feature noticed in all these tubes is the unchanging wall thickness during inversion. A close examination of the tubes revealed that the resulting tubes are shorter in length with no change in thickness. It is this feature which makes the process rather useful to the metal-forming industry. Metals other than aluminium also invert in the same manner, provided they are ductile enough and an appropriate die radius is used. Figure 6 shows typical inverted 2.0-in o.d., 0.064-in wall thickness half-hard aluminium, brass and copper tubes.

The most important parameter in this process is the fillet radius. Tests on aluminium tubes inverted on dies of different radius showed a range of values to this radius below and above which buckling or inversion result in respectively. Figure 7 shows a plot

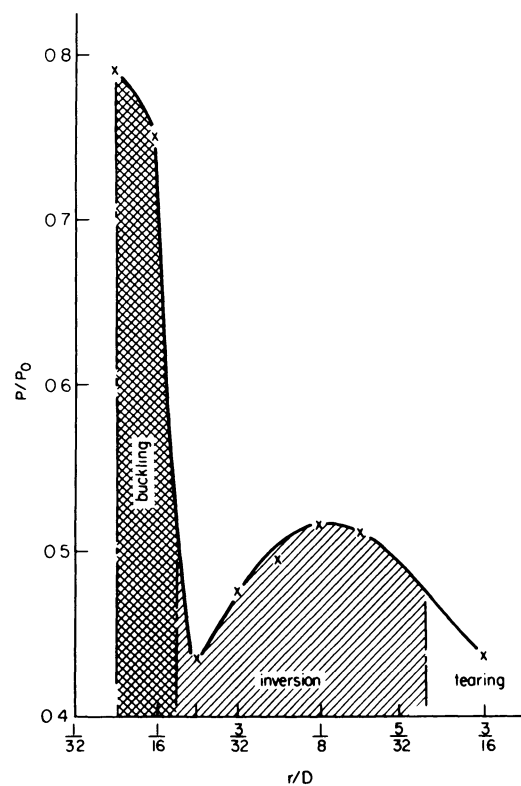


Figure 7. Variation of inverting or maximum load P/P_0 with die fillet radius r/D . P_0 is the yield compressive load.

of inverting or maximum load P divided by the compressive yield load P_0 versus the die radius divided by the tube diameter. P_0 is taken to be $\pi D t Y$ where Y is the yield stress at about 20 per cent strain. It is found that there exists a particular die radius which produces a minimum inverting load. In this case, the tube wall will conform completely to the curved part of the die.

The extent of the region of successful inversion varies from metal to metal, but the critical radius is found to be independent of the material. At large values of r/D , the inverted tube wall begins to thin. It is possible, therefore, to use a large fillet radius if it is acceptable to have a larger tube diameter with a relatively thinner wall. In inverting with a critical die radius, 25 per cent diametral expansion is achieved. However, inverting with the largest possible die

radius, it is possible to achieve 56 per cent diametral expansion before tearing.

An alternative method of obtaining larger diameters is to use a die of a critical radius but with an angle larger than 90° . This die enables the tube to slide over the slanted edge of the die whilst inverting. Figure 8 shows the influence on the maximum

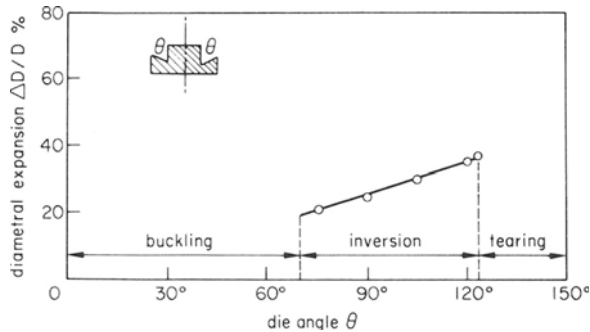


Figure 8. Variation of percentage diametral expansion $\Delta D/D$ with the die angle. The die radius kept at $\frac{3}{8}$ -in.

diametral expansion of the angle θ in a die with a critical radius. It is found that below 70° die angle, the tube cannot flow around the die and buckling develops. Beyond 120° , however, the tube reaches instability, where neither inversion nor tearing is produced and at large values of θ tearing dominates.

It is interesting to note that by cutting the curly collar from the inverted tube and then repeating the inversion a very steady inversion is produced. Figure 9 shows a 2.0-in o.d., 0.064-in wall thickness,

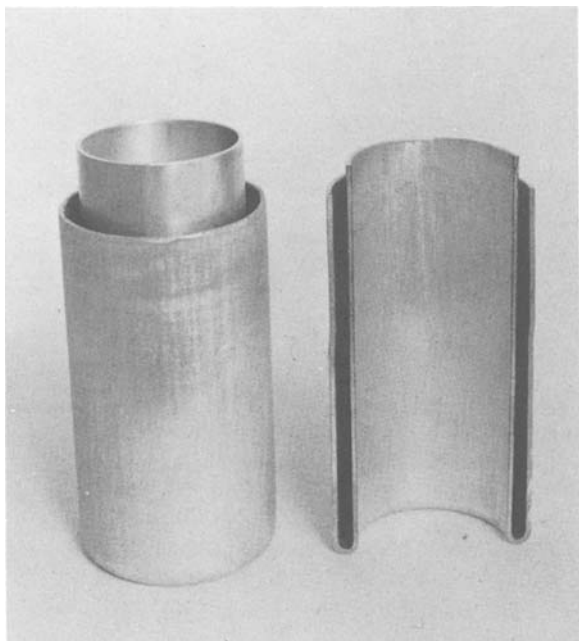


Figure 9. Half-hard aluminium tube 2.0-in o.d., 0.064-in wall thickness steadily inverted after removal of the collar.

aluminium tube with its collar sawn off. The resulting tube is reasonably uniform both in diameter and wall thickness. The load-shortening curve for the collarless tube indicated a constant load throughout. When the outer tube was cut and again reinverted, the load-shortening curve was similar to that of the first

inversion. However, the steady inversion load was in this case slightly higher. It should be noted that for successful reinversion the once inverted tube must have a uniform diameter throughout its length.

Chemical etching and microscopic examination of the surface of the original, once-inverted and twice-inverted tubes revealed that grains are appreciably reduced in each inversion.

It is worth mentioning here that successful inversion of two coaxial tubes, originally made to slide fit one inside the other, was achieved by Mellor². Figure 10, shows a combination of 2.0-in o.d.,

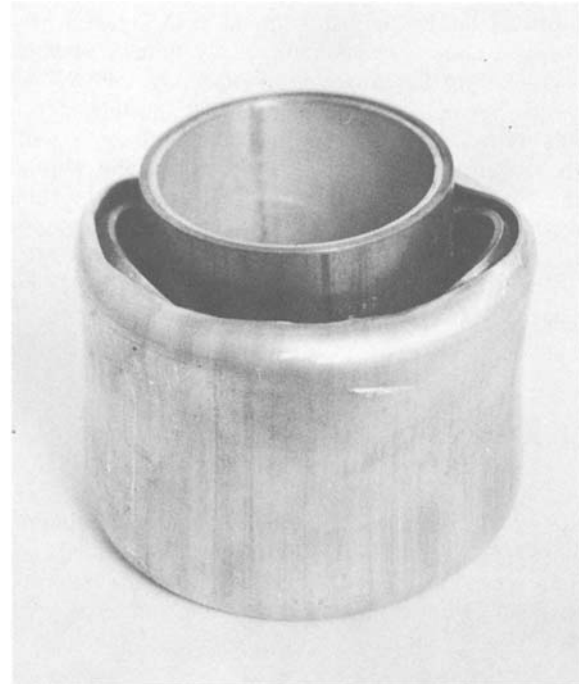


Figure 10. Coaxial aluminium inside brass, tubes inverted with a 90° die of fillet radius = 0.25 in. Mean diameter of tubes = 2.0 in and both tubes have wall thickness = 0.064 in.

0.064-in wall thickness half-hard aluminium inside a 2.0-in o.d., 0.064-in wall thickness 70/30 brass tube inverted with a 90° die with a $\frac{1}{4}$ -in fillet radius. Notwithstanding the difficulties involved in selecting a suitable die radius and material for the tubes, this process is rather promising for achieving tight-fit laminated tubes and thus deserves further investigation.

ANALYSIS

An expression for the steady compressive load necessary to maintain the inversion of a strain hardening tube is arrived at³ by assuming that,

- (1) the tube wall thickness remains constant,
- (2) the deformation process involves bending as well as enlarging the tube diameter, and
- (3) Bauehinger effects, buckling, and tearing are not present.

If the radius of inversion is r and the tube-wall thickness is t , see figure 1, the average strain in

bending a straight element of the tube is $\simeq \frac{1}{2} \ln[1 + (t/2r)]$. The total strain in both bending and unbending is, therefore

$$\epsilon_B = \ln\left(1 + \frac{t}{2r}\right) \quad (1)$$

For no thickness change, the representative strain $\bar{\epsilon}$, from volume constancy is

$$\bar{\epsilon} = \frac{2}{\sqrt{3}} \epsilon_B = \frac{2}{\sqrt{3}} \ln\left(1 + \frac{t}{2r}\right) \quad (2)$$

The work done per unit time in bending and unbending the tube wall is

$$W_B = \pi D t v \int \bar{\sigma} d\bar{\epsilon}$$

which, for $\bar{\sigma} = A(B + \bar{\epsilon})^n$ becomes

$$W_B = \frac{\pi D t v A (B + \bar{\epsilon})^{n+1}}{n+1} \quad (3)$$

where v is the crosshead speed and D the original diameter of the tube. From Equations (2) and (3) we have,

$$W_B \simeq \frac{\pi D t v A}{n+1} \left[B + \frac{2}{\sqrt{3}} \ln\left(1 + \frac{t}{2r}\right) \right]^{n+1} \quad (4)$$

which for small $t/2r$ becomes

$$W_B \simeq \frac{\pi D t v A}{n+1} \left[B + \frac{t}{\sqrt{(3)r}} \right]^{n+1} \quad (5)$$

The hoop strain in expanding the tube diameter D to $D + 4r$ is $\ln[1 + (4r/D)]$ and for no change in thickness, volume constancy gives, in this case

$$\bar{\epsilon} = \frac{2}{\sqrt{3}} \ln\left(1 + \frac{4r}{D}\right) \quad (6)$$

The amount of work done per unit time, W_E , in expansion is obtained from (3) and (6). Thus

$$W_E = \frac{\pi D t v A}{n+1} \left[B + \frac{2}{\sqrt{3}} \ln\left(1 + \frac{4r}{D}\right) \right]^{n+1} \quad (7)$$

which for small $4r/D$ reduces to

$$W_E = \frac{\pi D t v A}{n+1} \left[B + \frac{8r}{\sqrt{(3)D}} \right]^{n+1} \quad (8)$$

The rate at which work is done by the compressive force, neglecting friction, is from Equations (5) and (8)

$$\begin{aligned} P v &= W_B + W_E \\ &= \frac{\pi D t v A}{n+1} \left\{ \left[B + \frac{t}{\sqrt{(3)r}} \right]^{n+1} + \left[B + \frac{8r}{\sqrt{(3)D}} \right]^{n+1} \right\} \end{aligned} \quad (9)$$

The critical value of r is found by supposing it to

acquire a value which makes P a minimum. Hence, for $dP/dr = 0$, we have

$$\frac{8r^2}{Dt} = \left\{ \left[B + \frac{t}{\sqrt{(3)r}} \right] \left[B + \frac{8r}{\sqrt{(3)D}} \right] \right\}^n \quad (10)$$

which is satisfied by

$$r = \sqrt{\left(\frac{Dt}{8}\right)} \quad \text{or} \quad \frac{r}{D} = \sqrt{\left(\frac{t}{8D}\right)} \quad (11)$$

Putting this value of r into Equation (9), we obtain

$$P = \frac{2\pi D t A}{n+1} \left[B + \sqrt{\left(\frac{8t}{3D}\right)} \right]^{n+1} \quad (12)$$

For a perfectly rigid plastic tube, $n = B = 0$ and in plane strain $A = (\sqrt{3}/2)Y$. Putting these in Equation (12) we get

$$P = \pi D t Y \sqrt{\left(\frac{8t}{D}\right)} \quad (13)$$

which is identical to the expression derived by Johnson⁴ from a rigid plastic analysis.

In order to check for tearing we assume that the tip of the tube will not tear provided the hoop strain (as it reaches the straight edge of the die), $\ln[1 + (2r/D)]$, is less than the instability strain in axial tension. It may be easily shown¹ that in simple tension instability is reached when $\epsilon = n - B$. Thus, for no tearing we must have

$$n - B > \ln\left(1 + \frac{2r}{D}\right) \quad (14)$$

or, for small r/D

$$\frac{r}{D} < \frac{1}{2}(n - B) \quad (15)$$

To avoid buckling, the radius of the die must be such that the inverting load remains below the axial buckling load for the tube P_B . It may be shown⁵ that for a tube whose tangent modulus is E_t .

$$P_B = \frac{4}{3} \pi t^2 E_t \quad (16)$$

For $P < P_B$, we must have, using Equation (13)

$$\frac{r}{D} > \frac{3Y}{4E_t} \quad (17)$$

where Y and E_t may be taken at, say, 0.1 strain in a compression test.

Another limitation to the inverting radius may be obtained from the geometrical condition, that the critical radius of the die must be larger than the tube wall thickness. Otherwise, even if the tube tip turns over, it will touch the original straight tube and may enhance buckling. Thus only tubes which have $\sqrt{(Dt/8)} > t$ may fully invert without buckling. This requires

$$\frac{D}{t} > 8 \quad (18)$$

Combining conditions (15) and (17), we have, for tubes satisfying conditions (18), successful inversion may be achieved provided the die radius falls in the range

$$\frac{3Y}{4E_t} < \frac{r}{D} < \frac{1}{2}(n - B) \quad (19)$$

In high-speed inversion, the rate of strain may be large. An estimate of the strain rate may be obtained from dividing the mean strain by the time an element takes to bend and unbend, i.e. to traverse a distance πr . The mean strain in each element is the work done on the tube per unit time, divided by the volume deformed times Y . Hence, from equation (13)

$$\epsilon_m = \frac{Pv}{\pi DtYv} = \sqrt{\left(\frac{8t}{D}\right)}$$

The mean strain rate, therefore, is approximately

$$\dot{\epsilon}_m \approx \sqrt{\left(\frac{8t}{D}\right)} \cdot \frac{v}{\pi r} \approx \frac{v}{\pi D} \quad (20)$$

It is not surprising, therefore, that the same tubes could not be inverted by explosive impact. At speeds up to 200 ft/sec, a 2.0-in diameter tube inverts with a strain rate ≈ 400 per second and at this value the mean yield stress for aluminium is likely to be between 20 and 60 per cent greater than the quasi-static yield stress. It is possible for the inverting load to increase to a large enough value to produce buckling.

Under the drop-hammer and in the collision of vehicles, $v \approx 30$ mile/hr (44 ft/sec), so that the strain rate is approximately 90 per second and the yield stress does not increase sufficiently to influence the inversion load.

DISCUSSION AND CONCLUSION

Thin-walled tubes of different materials and different sizes have been successfully inverted inside-out using dies of different fillet radius. It is found that the theoretically predicted inversion radius agrees very well with the experimental value. Inversion is precariously balanced between Euler column, axisymmetric and diamond buckling, and tearing modes. A range of values of the fillet radius are predicted for successful inversion. These are based on the condition that the die radius must be less than that which causes tearing of the tip of the tube, as it slides over the curved part of the die, and larger than that which causes the inverting load to exceed the axial

buckling load of the tube. A cursory examination of this range indicated reasonable agreement for some metals and did not agree for others. The discrepancy is expected, however, to be due to the error involved in the theoretical expressions used for both the buckling load and the tensile instability strain. Better agreement is reached when experimental values for these were used in Equation (19). In this case, Equation (19) should be rewritten in the form,

$$P_B/8P_0 < \frac{r}{D} < \frac{1}{2} \epsilon_t$$

where P_B is the experimentally measured axial buckling load, P_0 is the yield load and ϵ_t is the experimentally determined tensile fracture strain.

It has been further deduced that only tubes with a diameter larger than eight times the wall thickness could be considered for inversion. In tubes with larger wall thickness the tip touches the original straight wall of the tube, crimps on it and, consequently, causes the tube to buckle.

Due to the neglect of friction and to the other simplifying assumptions made in the analysis, the theoretical inversion load given by Equations (12) or (13) is expected to underestimate the actual value. Experiments showed that for aluminium, copper and brass tubes, Equation (12) predicted values of between 80 and 90 per cent of the experimentally obtained steady inversion loads.

The uniformity in diameter and the constancy in wall thickness of the inverted tubes suggests this process as an alternative metal forming technique to internal pressurization of tubes.

ACKNOWLEDGEMENTS

The author would like to thank Professor W. Johnson for his useful comments. He would also like to thank Mr W. T. Lowe and Mr R. J. Mellor whose post-graduate work provided the background to this paper.

REFERENCES

1. W. Johnson and P. B. Mellor. *Engineering Plasticity*, Van Nostrand, 1973.
2. R. J. Mellor. The axial impact of tubes, MSc. Dissertation, University of Manchester, 1971.
3. S. T. S. Al-Hassani, W. Johnson and W. T. Lowe. Characteristics of inversion tubes under axial loading, *J.M.E.S.* (1972) 14, no. 6.
4. W. Johnson. *Impact Strength of Materials*, Arnolds, 1972.
5. S. Timoshenko. *Theory of Elastic Stability*, McGraw-Hill, New York, p. 442, 1939.

COLD FORMING, BLANKING, PIERCING, CROPPING

THE STRENGTH OF THE MATERIAL IN COLD FORMED PARTS

by

JOSEPH DATSKO*

SUMMARY

Cold formed parts are extensively used in the manufacture of automobiles, aeroplanes, appliances, and all other mechanical devices. In order to reliably design such parts it is necessary for the engineer to be able to calculate the strength of the material in the fabricated part at those locations where the stresses due to the applied loads are high.

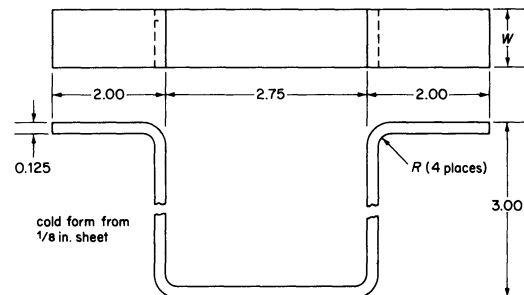
The mechanical design engineer of today has been well educated in the techniques of stress analysis. In a typical stress analysis problem the mechanical properties of the material are considered as handbook information. It is assumed that the design engineer gets the properties from the materials engineer who in turn has access to the appropriate tabular data. Unfortunately, the data listed in the handbooks are for the material in its original condition as a bar or plate, whereas the properties that are needed in order to make the stress analysis valid are the properties of the part after it has been fabricated from the bar or plate.

This paper presents a new method of designating the strength of the material in a formed part that includes both the directionality and the sense of the particular strength being analysed. The techniques for calculating the numerical values of the strength on the basis of the material's original properties as well as the subsequent mechanical processing are discussed. The calculations of the strength of the cold formed material are made on the basis of the 'Apparent Rules of Strengthening' that are presented in this paper.

INTRODUCTION

Formed parts are extensively used in the manufacture of automobiles, aeroplanes, home appliances, and all other mechanical devices. In order to design such parts so that they have a mechanical reliability along with a low weight and cost it is necessary for the engineer to be able to determine the strength of the material in the fabricated part. This is particularly true at those locations where the stresses due to the applied loads are high. The traditional experimental method of determining the strength of the part by testing a prototype is much too expensive to be widely used and it appears very 'unengineering-wise' in the light of today's sophisticated space-age design achievements.

In addition, to insure the reliability of a mechanical component the design engineer must specify, and have control over, the important details of the actual forming processes that are used to fabricate the part. For example, the engineer designing a hat section as shown in figure 1 can no longer simply state on the blueprint: 'cold form from x in. thick sheets'. If he designs the part on the basis of the material having a certain strength in the legs of the hat section, then he must also indicate whether the section should be formed on a press brake, a roll



REQD	W (in.)	R (in.)	UNIVERSITY of MICHIGAN		NAME -	HAT SECTION
50	1.00	0.156	TOLERANCES	DRAWN	R. N.	FOR - RESEARCH
50	4.00	0.156	DEC. ± 0.020	APPD.	J. D.	REQD - SCALE - FULL
100	36.00	0.218	FRACT. ± 0.030	DATE	1-22-71	MATERIAL PART NO.
			ANG. ± 30'			1100 ALUMINIUM 2-100

Figure 1. Blueprint of a hat section.

forming machine, or in a draw die on a press. If it is formed in a draw die, then he must also specify the die radius inasmuch as the die radius has a significant effect on the strength of the formed part. This is illustrated in the example at the end of the paper.

The mechanical design engineer of today has been well educated in the techniques of stress analysis. In a typical stress analysis problem the mechanical pro-

* Mechanical Engineering Department, University of Michigan.

properties of the material are considered as handbook information. It is assumed that the design engineer gets the properties from the material manufacturer's handbook, or else he obtains the properties from a materials engineer. Unfortunately, the data listed in the handbooks are for the material in its original condition as a bar or plate, whereas the properties that are needed in order to make the stress analysis valid are the properties of the part after it has been fabricated from the bar or plate.

This paper presents the concepts and theories of strength analysis, which is an emerging engineering discipline that enables the designers to calculate the strength of the fabricated parts. The calculations are made on the basis of an accurate knowledge of a material's original properties as well as the effect that subsequent mechanical processing has on those properties. Presented first is a new method of designating strength that includes both directionality and sense. Then the techniques for calculating the numerical values of the strength in all directions in parts made by cold-working processes are discussed.

STRENGTH DESIGNATION

A material that has been cold worked has mechanical properties that are different in both the sense of the strength (tension or compression) and in the direction within the part (longitudinal transverse). Therefore it is necessary, when referring to the strength of a cold formed part, to specify both the sense and the direction of the strength that is being studied. In addition, when making a strength analysis as discussed in the following section, it is necessary to include the prior strain history at the location where the strength is being studied. The traditional manner of discussing the yield strength of a cold worked material without being this specific is completely inappropriate in today's sophisticated design age where product reliability and liability are so important. Nearly all of the data that one finds in the textbooks and handbooks on materials which show the variation in yield strength or tensile strength as a function of cold work do not indicate either the sense of the direction of the strength. Some of this data is presented in the form of graphs or curves with the percentage of cold work as the abscissa and the strength as the ordinate. More often the data is presented in tabular form as illustrated in Table 1 for 304 stainless steel. Nothing is stated in either presentation concerning the sense (tensile, compressive

TABLE 1 Effect of cold work on mechanical properties* (304 stainless steel)

Per cent cold reduction	Yield strength (psi 0.2% offset)	Tensile strength (psi)	Elongation (per cent in 2 in.)
0	32 660	86 300	54.5
10	70 000	97 900	36.5
20	96 200	113 200	24.0
30	118 500	131 100	15.5
40	135 000	145 700	12.0
50	145 000	158 300	6.0

* Allegheny Ludlum Steel Company

or shear) of the strength or its orientation in the part. How does an engineer use this data if he is designing a shaft of 304 stainless steel that will have one of its ends upset by 30 per cent cold work to form a gear blank? Since the load on a gear tooth is applied in the transverse or circumferential direction, the strength in the transverse direction is the appropriate one to use. In this particular case, is either the tensile yield strength or the compressive yield strength equal to 118 500 psi? It will be shown in the following section of this paper that the compressive yield strength in this situation is only 75 per cent of the tensile yield strength. Since there is such a large difference between these two values it is necessary for the designer to include this type of an analysis in his calculations for the fabricated mechanism to be reliable.

Before describing this method of designating strength, it is beneficial to first distinguish between stress and strength since these two terms are sometimes confused with each other and the same symbols are frequently used for both. This confusion probably arises because the tensile test not only determines the strength of the material, but stresses are also induced in the tensile specimen during loading. In the context of strength analysis, strength is a *mechanical property* of a material whereas stress is a *condition* a material is in due to a load. A part has its strength prior to the application of a load but it has a stress only when the load is applied. Strength is a particular load (such as yield, ultimate or fracture) divided by the original area of the specimen, and is designated by the capital letter S . Stress is any load divided by the instantaneous area of the part and is designated by the Greek letter σ . Thus σ_u is the stress on a tensile specimen when the ultimate (maximum) load was applied, and S_u is the ultimate strength of the material.

After several years of teaching experience with this subject the author has developed a very simple and useful notation¹ for designating the strength of a cold formed part. In this notation strength is represented by the capital letter S followed by four or more letters in the form of subscripts, with parentheses placed around the S and first letter. A typical example of this notation is $(S_u)_{tLC}$, which reads 'the ultimate tensile strength in the longitudinal direction after prior axial compression'.

To explain this designation, it is convenient to replace the letter subscripts with sequential numbers. In this form the standard designation is $(S_1)_{234}$. The first subscript indicates the type of strength and may be any of the following lower case letters: y , for yield; u , for ultimate; f , for fracture; e , for endurance limit.

The second subscript refers to the sense of the strength and it may be any of the following three lower case letters: t , for tensile; c , for compression; s , for shear.

The third subscript indicates the direction or orientation of the test specimen within the formed part. It may be L , for a specimen whose axis is parallel to the longitudinal direction of the part. For symmetrical cross sections such as rounds or squares it may be T , for transverse; R , for radial; or C , for

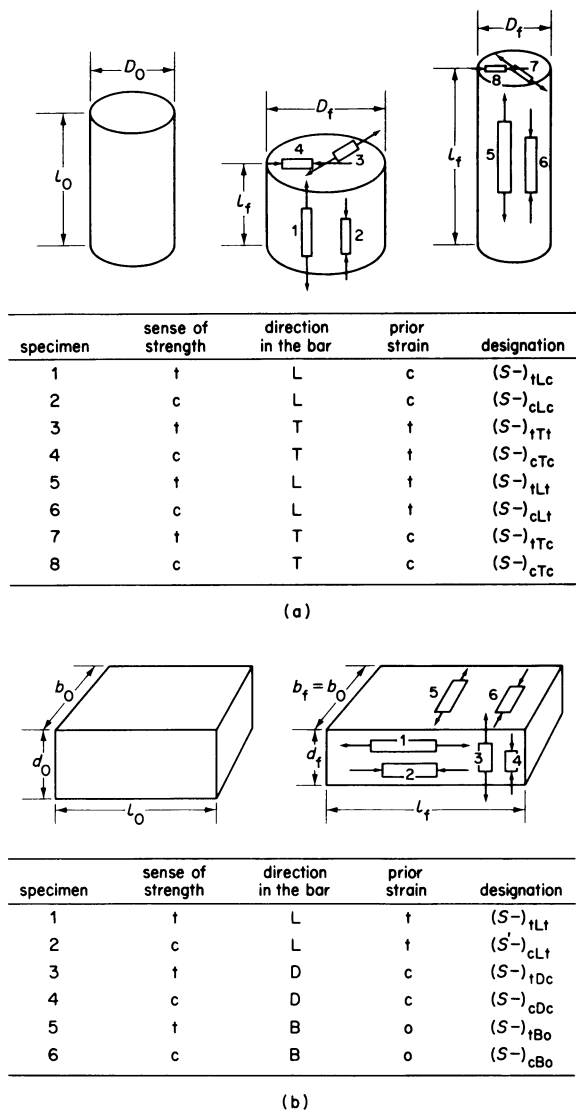


Figure 2. Strength designation after (a) uniaxial deformation; (b) rolling a plate.

circumferential. These latter three are all equivalent. In the case of plates or rectangular bars two transverse directions must be specified: the long transverse can be indicated by B (breadth); the short transverse can be designated D (depth).

The fourth subscript refers to the sense of the prior strain *in the axial direction of the test specimen*, and not in the direction of the deforming force. The notation here is the same as for the second subscript. That is: t, for tension; c, for compression; or s, for shear.

Also, a fifth (and sixth) subscript could be added if a part were deformed more than once. For example, a bar could be first stretched and then compressed. This would be true in the case of cold heading a portion of a previously cold drawn bar.

A second example is given to illustrate this designation. Consider the notation $(S_y)_{cTc}$ which means 'the compressive yield strength in the transverse direction after prior compression in the transverse direction' (equivalent to an extension in the longitudinal or axial direction). Figure 2 illustrates this method of designating the strength in round bars and plates. It should be noted in figure 2(b) for the

rolled plate that since the final breadth is equal to the initial breadth, the strain in this direction is 0.

Consider the previous problem where a head is cold upset on the end of a previously cold drawn bar. The notation for the transverse tensile yield strength in this case is $(S_y)_{tTct}$. In similar fashion the strength at any location of any cold formed part can be specified.

The following section gives the concepts and techniques for calculating the numerical values of the strengths of cold formed parts on the basis of the material's original plastic stress-strain properties and the amount of the plastic deformation.

THE APPARENT RULES OF STRAIN STRENGTHENING

The uniaxial stress-strain characteristics of a material are determined by means of either a tensile or compressive test on a sample of the material. For most non-cold worked materials the results of the two tests are identical. In order to simplify the use of plastic stress-strain data it is beneficial to reduce it to a mathematical expression that is reasonably accurate. The two most commonly used expressions² are:

$$(1) \sigma = \sigma_0 \epsilon^m;$$

$$(2) \sigma = A + B\epsilon.$$

The author found the first expression to be more reliable over a wide range of materials and it will be used exclusively in the following concepts and rules. However, these concepts and rules are general in that any valid mathematical expression of the material's stress-strain characteristics can be employed. It should be pointed out that the first expression is frequently written with other symbols, such as:

$$\sigma = K\epsilon^m \quad \sigma = K\epsilon^n \quad \sigma = A\delta^m$$

and so on.

The terms strain hardening and strain strengthening are frequently considered to be synonymous since, in a general way, strength and hardness are related to each other. However, in order to apply the rules of strain strengthening to solve problems in the field of strength analysis it is necessary to distinguish between the strain hardening equation (or exponent) and the strain strengthening equation (or exponent). The original strain hardening exponent, frequently referred to as Meyer's strain hardening exponent, is designated by the letter *n* in the equation $L = Ad^n$ where *L* is the load on a ball of given diameter and *d* is the diameter of the resulting indentation³. The numerical values of *n* vary between 2.0 and 2.7 over the entire range of metals.

Since the uniaxial stress-strain curve of a material is also the locus of yield strengths of the material after prior tensile deformation¹, its mathematical expression, $\sigma = \sigma_0 \epsilon^m$, is really a strain strengthening equation. Likewise it is appropriate to call *m* the strain strengthening exponent. The numerical values of *m* vary between 0 and 0.7 over the entire range of metals. The relationship between the strain hardening exponent and the strain strengthening exponent³, namely $m = n - 2$, is very useful in the field of strength

analysis. Thus if the strain hardening exponent of a material is known, its strain strengthening exponent can be predicted. This is particularly useful since it is much easier to experimentally determine the value of n .

The rules of strain strengthening can now be presented on the basis of the above notation of strength and the stress-strain characteristics of materials. These rules have been developed by the author during the past fifteen years while attempting to present to engineering students in a systematic manner the results of a large number of individual student research projects in the common field of plastic deformation and mechanical properties. The author refers to these rules as the 'apparent' rules because they portray what the mechanical properties of a cold worked material appear to be on the basis of broad based experimental studies.

Rule 1. Strain strengthening is a 'bulk' mechanism. Even a deformation load that is applied in only one direction causes deformation and strengthening in all directions. For example, in the cold rolling of a plate as illustrated in figure 2, even though there is no increase in the breadth of the plate both the tensile and compressive yield strengths are significantly increased in this direction. In fact, with as little as 10 per cent cold work, both the tensile and compressive yield strengths in the breadth direction are increased by approximately 50-100 per cent even though $\epsilon_b = 0$.

Rule 2. The maximum deformation that can be given a material during the forming of a part is that deformation which induces a tensile strain that is numerically equal to the fracture strain of that material when tested in the same direction and identical environment. In other words, fracture will occur when the induced $\epsilon_t = \epsilon_f$. This has previously been referred to by the author as *the failure theory for plastic working*¹. However, from a practical approach, a few forming operations are terminated when the induced tensile strain is equal to the ultimate load strain, which is also numerically equal to the strain strengthening exponent. This is true for those forming operations in which the condition known as necking occurs.

Thus, in compressing a cylinder or in upsetting a head on the end of a bar, the maximum longitudinal strain that can be given to that part is equal to two times the transverse fracture strain of the original bar. The factor two is necessary because the transverse strains are numerically equal to one-half the longitudinal strains during plastic deformation where the volume of the material remains constant.

Rule 3. The numerical value of σ_w in the strain strengthening equation $\sigma_w = \sigma_0 \epsilon_w^m$ is always determined by using for ϵ_w the largest value of strain, which is usually the longitudinal or axial strain. The subscript w denotes cold work. Thus in determining the transverse strengths of a part deformed by a uniaxial load, as shown in figure 2(a), the longitudinal strain ϵ_L is used and not the transverse strain ϵ_T which is equal to one half ϵ_T .

Rule 4. After upsetting, stretching or rolling round the square bars, the yield strengths that have

the second and fourth subscripts of the same sense are equal to σ_w . Thus

$$(\mathcal{S}_y)_{tLt} = (\mathcal{S}_y)_{cLc} = (\mathcal{S}_y)_{tTt} = (\mathcal{S}_y)_{cTc} = \sigma_w = \sigma_0 \epsilon_w^m$$

Rule 5. After upsetting, stretching or rolling round and square bars, the yield strengths that have the second and fourth subscripts of the opposite sense are equal to $(\frac{1}{2})^m$ times σ_w . Thus

$$(\mathcal{S}_y)_{tLc} = (\mathcal{S}_y)_{cLt} = (\mathcal{S}_y)_{tTc} = (\mathcal{S}_y)_{cTt} = 0.5^m \sigma_w$$

Rule 6. After rolling or bending flat bars and plates the yield strengths that have the second and fourth subscripts of the same sense or else have the fourth subscript equal to zero are equal to σ_w . Thus

$$(\mathcal{S}_y)_{tLt} = (\mathcal{S}_y)_{cDc} = (\mathcal{S}_y)_{tB0} = (\mathcal{S}_y)_{cB0} = \sigma_w$$

Rule 7. After rolling or bending flat bars and plates the yield strengths that have the second and fourth subscripts of the opposite sense are equal to $(\frac{1}{2})^m$ times σ_w . Thus

$$(\mathcal{S}_y)_{cLt} = (\mathcal{S}_y)_{tDc} = 0.5^m \sigma_w$$

Rule 8. One reversal of equal but opposite strain of magnitude equal to ϵ_w makes all $(\mathcal{S}_y)_w$'s equal to σ_w obtained by using ϵ_w in the strain strengthening equation. Thus for a tensile strain of +0.1 followed by a compressive strain of -0.1, the value of σ_w is determined by using $\epsilon_w = 0.1$, and not 0 or 0.2.

Rule 9. The tensile ultimate strength of a material after prior tensile deformation in the direction of the specimen is equal to the original non-cold worked tensile ultimate strength divided by one minus the fraction of cold work. The fraction of cold work is the percentage of cold work divided by 100. In terms of the above described strength notation,

$$(\mathcal{S}_u)_{t-t} = \frac{(\mathcal{S}_u)_0}{1 - W_f}$$

This relationship is valid, as derived in reference 1, for deformation strains equal to or less than the ultimate load strain. Thus the tensile ultimate strength in the transverse direction of a cylinder that has been cold upset by 20 per cent cold work is 1/0.8 or 1.25 times its original tensile strength.

Rule 10. The tensile ultimate strength of a material after prior compressive deformation in the direction of the specimen is equal to the original non-cold worked tensile ultimate strength divided by one minus half the fraction of cold work. In terms of our strength notation,

$$(\mathcal{S}_u)_{t-c} = \frac{(\mathcal{S}_u)_0}{1 - 0.5W_f}$$

This relationship is also valid only for cold work deformation strains not exceeding the ultimate load strain of the original material. Thus the tensile ultimate strength in the axial direction of a cylinder previously compressed 20 per cent is 1/0.9 or 11 per cent higher than its original strength.

Rule 11. The tensile ultimate strength of a material after cold deformation strains greater than the ultimate load strain of the original material is equal to the tensile yield strength of the deformed material.

Rule 12. The fatigue strength of a material is increased by cold working in proportion to the increase in the tensile strength.

EXAMPLE OF STRENGTH ANALYSIS

The application of the above concepts and relationships can best be illustrated by considering a typical strength analysis problem. Suppose a hat section as shown in figure 1 is made by deep drawing $\frac{1}{8}$ in thick x 10 in wide x 20 in long blanks over a $\frac{1}{8}$ in die radius and that the material is 1100-0 aluminium strip having the following isotropic mechanical properties: yield strength, 4.8 ksi; tensile strength, 12.6 ksi; strength coefficient, 23 ksi; strain strengthening exponent, 0.25; reduction of area, 85 per cent.

The calculated values for the tensile yield strength, the tensile ultimate strength, and the reversed load endurance limit in the axial direction of the hat section are shown in figure 3. The calculations are made on the basis of the 'Apparent Rules of Strain Strengthening' presented in the preceding section. It is also assumed that the neutral axis moves inward from the mid-thickness during bending to a location that results in the outer fibre and inner fibre strains being equal. The strains are calculated by means of the expression

$$\epsilon = \ln \frac{R + Y}{R + N}$$

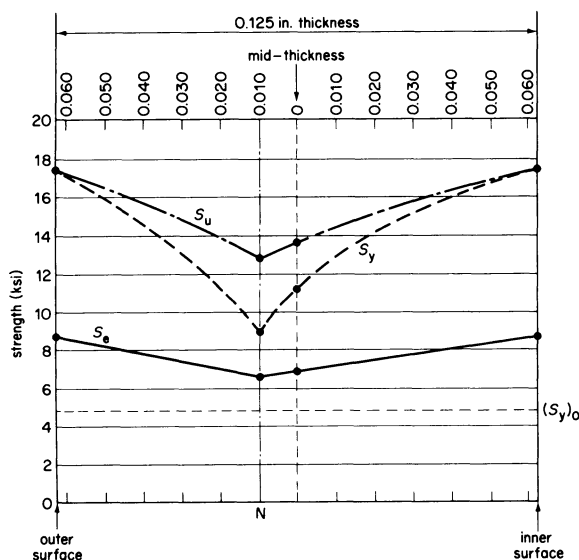


Figure 3. Strength distribution through the thickness of a leg of a deep-drawn hat section.

where R is the inside (or die) radius, Y is the distance above the inside radius, and N is the distance to the neutral axis above R .

Thus for this problem the final neutral axis is found to be 0.052 in above the inside radius and the inner and outer fibre strains are equal to 0.348 which is equal to 29 per cent cold work. Since the strain due to cold work is greater than the strain strengthening exponent, the tensile yield strength will be only slightly less than the tensile ultimate strength which are calculated from the strain strengthening equation as follows. For the outside surface of the channel legs

$$(S_y)_{tLct} = \sigma_w = \sigma_0(\epsilon_w)^m = 23(0.348)^{0.25} = 17.6 \text{ ksi}$$

At the inside surface of the channel legs the notation of the strength is $(S_y)_{tLtc}$ and it is also equal to 17.6 ksi.

At the mid-thickness when the neutral axis is at 0.052 in above R , the strain is 0.0594 or 6 per cent cold work. By substituting this value into the strain strengthening equation, the tensile yield strength is determined to be 11.3 ksi.

One further interesting point is the strength at the location where the final neutral axis lies after the bending operation. Since the neutral axis originally is at the mid-thickness, the 0.052 in location receives compressive strains during the early part of the bending operation. But since the final strain is zero it is obvious that there was a maximum compressive strain when the bend radius was somewhere between infinity and the final value of $\frac{1}{8}$ in. This radius can be calculated and the resulting maximum strain at N is found to be 0.0237 or slightly more than 2 per cent cold work. This gives a resulting tensile yield strength of 9 ksi which is nearly twice the original value.

The endurance limit shown in figure 3 was calculated by assuming that it is one-half the value of the tensile strength, as is customarily done for aluminium.

ACKNOWLEDGMENTS

The forming projects in which these concepts and relationships were derived were financially supported by the National Science Foundation, the Mechanical Engineering Department of the University of Michigan, and the O. W. Boston Grant In Aid.

REFERENCES

1. J. Datsko. *Material Properties and Manufacturing Process*, John Wiley and Sons, Inc., New York, 1966, p. 342.
2. Oscar Hoffman and George Sachs. *Theory of Plasticity*, McGraw-Hill Company, New York, 1953.
3. Hugh O'Neill. *The Hardness of Metals and its Measurement*, Chapman Hall, London, 1934.

THE DEVELOPMENT OF A COLD FORMING PROCESS ON AN AUTOMATIC HIGH-SPEED FORMING MACHINE

by

T. SCAIFE*, B. PARSONS* and B. N. COLE*

SUMMARY

The design and initial development of an automatic high-speed forming machine was presented in an earlier paper¹. In the present paper further development of this machine, including the tooling and instrumentation for use in the cold can-extrusion of steel, is described.

The problems encountered in performing the process of high-speed impact can-extrusion are described and defined in terms of their dependence on the machine and/or process. The methods which have been introduced to overcome these problems are discussed.

Assessments of machine performance and experimental measurements of tooling loads are given and compared with measurements for slow-speed operation. Estimates of the part played by tool friction and strain rate sensitivity of the workpiece are also given.

Successful production-type runs, in which the high-speed machine has operated on automatic cycling for a few hours, are described. Reference is made to the heat generation in tooling and products, and the dimensional accuracy of the products obtained from the continuous run.

INTRODUCTION

The production of components by cold-forming processes is of considerable interest to industry since such processes often produce products with sound mechanical properties and of such surface finish and dimensional accuracy that subsequent machining is greatly reduced. Materials which are particularly amenable to such processes are low carbon steels. Economic justification of the high capital cost of the presses and tooling can only be obtained by using large batch production of components.

A three-stage cold-forming method is generally adopted by industry to produce tubular steel components of relatively short lengths from solid bar stock. Firstly a lubricated billet is backward-extruded into a can. The base wall of the can is subsequently pierced and the resulting tube is annealed and relubricated. The third stage is the forward ironing of the short tube into the tubular product, which can be either fully ironed or backwardly ejected, and needs only a small amount of machining.

Of the claims which have been made regarding the benefits to be achieved by using high extrusion speeds, the more relevant would seem to be the expectancy of achieving an improved quality of shear in a blanking operation², and the possibility of better material flow leading to enhanced product properties

including surface finish³. The main disadvantages of high-speed working are reckoned to be the increased tool loading and the risk of damage to the product as a consequence of inertial forces, though the latter can be avoided by careful design⁴.

Numerous papers³⁻⁶ have been presented describing cold-forming extrusion processes with high-speed machines. These have been concerned mainly with single-shot type experiments under laboratory conditions. In particular Davies and Singh⁵ reported on the high-speed cold extrusion of steel (En25) and Aluminium (NE4) and concluded that backward extrusion, forward extrusion and ironing were feasible and that the high-speed extrusion products were comparable with those produced using more conventional speeds. In most cases the tool loading during the high-speed processes was found to be marginally lower than that during the corresponding slow-speed process. This latter observation was explained to some extent by producing evidence that suggested a reduction in frictional forces between the deforming material and dies as the extruding speed increased. Also Davies and Dhawan² conducted a comparison of cold blanking at high and low speeds on mild steel, high carbon steel, brass, copper and aluminium. A marked improvement in the quality of shear was observed for the ferrous materials as the speed of the blanking operation was

* Department of Mechanical Engineering, University of Leeds

increased from quasi-static to 30-40 ft/s. No improvement in blanking characteristics was observed above this speed range. The main disadvantage of the high-speed blanking operation appeared to be the increased peak loads, which in some cases were twice that required for blanking at more conventional speeds.

The present paper contains the results of process development work performed on an automatic high-speed machine which was designed specifically for mass production of tubular steel products using cold-forming processes. The process reported is the backward extrusion of a solid EN32B steel billet into a can having a base wall/side wall thickness ratio less than 1.0.

EXPERIMENTAL MACHINE TOOLING AND MATERIALS

The high-speed forming machine used during this investigation is shown complete with instrumentation, in figure 1.

The machine is horizontally mounted and is activated by high-pressure air from a storage vessel. An automatic billet feed system comprising mechanical slides and low-pressure pneumatic cylinders is incorporated. Ejection of the finished product is performed either by a hydraulic piston or a mechanical spring-loaded stripper plate. The machine is recoiled by a pair of hydraulic cylinders, which recompress the air almost to the original firing pressure; thus wastage of compressed air per cycle is kept to a minimum. Manual or automatic operation of the machine is possible. A system of electrical interlocks prevents the machine from misfiring during

an operation, and manual override at any stage in the cycle has been incorporated. Cycle times of the order of 10 seconds have been reported during preliminary development of the machine⁷.

Although the design energy rating of the machine is 25 000 ft lbf per blow, during the present investigation it has operated only in the energy range 726 to 6540 ft lbf per blow, that is an impact velocity range of 20 to 60 ft/s.

Tooling

The tooling arrangement used is shown in figure 2. The duplex container A, which is mounted in the head of the machine and rests on the back-up piece B, is prevented from moving by the clamp ring C. The anvil D, which can be a standard, performing or combined extrusion type, is a sliding fit in the container bore and rests on the back-up piece B. The anvil was designed so that if required it could be used as a load cell. The punch E, which rests directly on the punch load cell F, is held in position by the retaining nut G, which screws into the moving ram head. Ejection of the product is via the ejection rod I which passes through the machine head, or by the spring-loaded stripper plate H, which is attached to the ram head.

A shock-resisting steel (Pax 2), a high chromium steel (Vital) and a high-speed steel (M 2) at various hardnesses, have been used for the anvil and punches. Satisfactory performance has been obtained using M 2 at 62-64R°C for the punches, and at 58-60R°C for the anvils. The container insert was made from Vital tool steel at 60-62R°C with an outer shrink ring of EN30B at 47-49R°C.

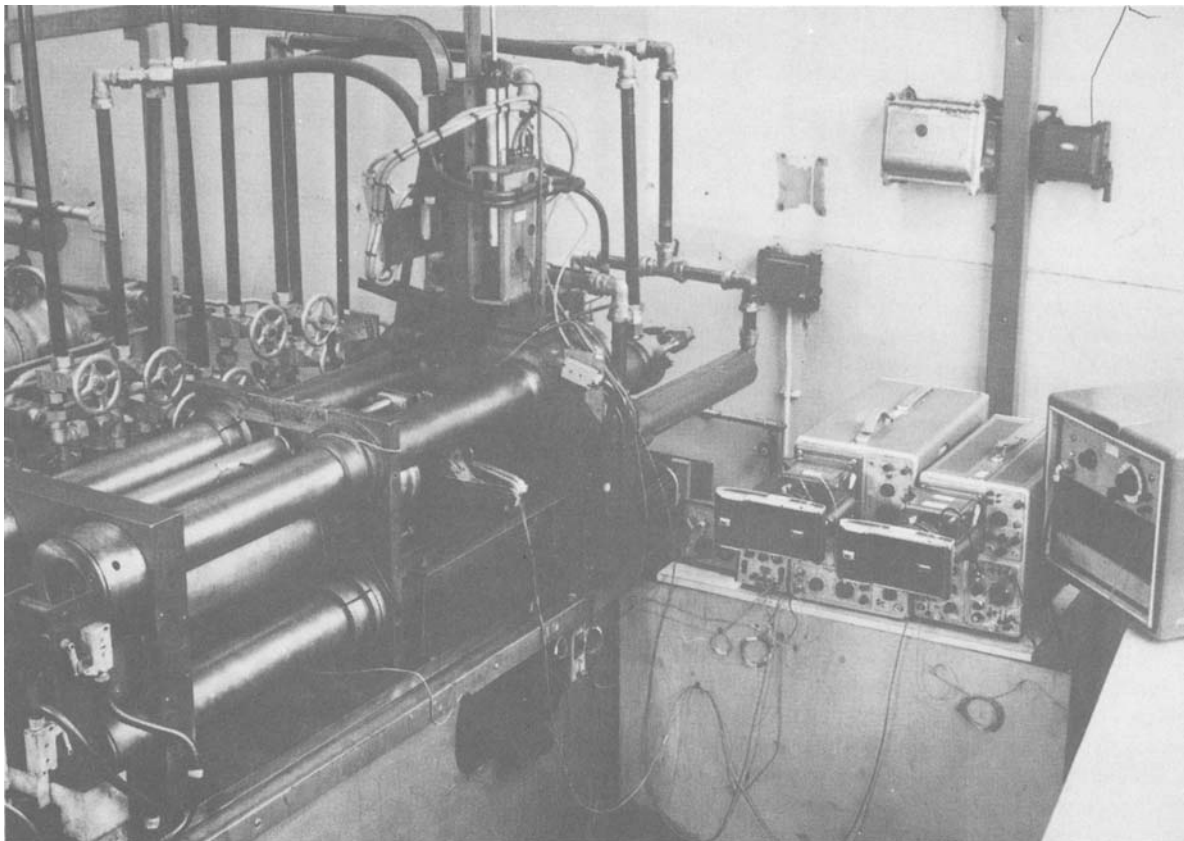


Figure 1. The high-speed machine complete with instrumentation.

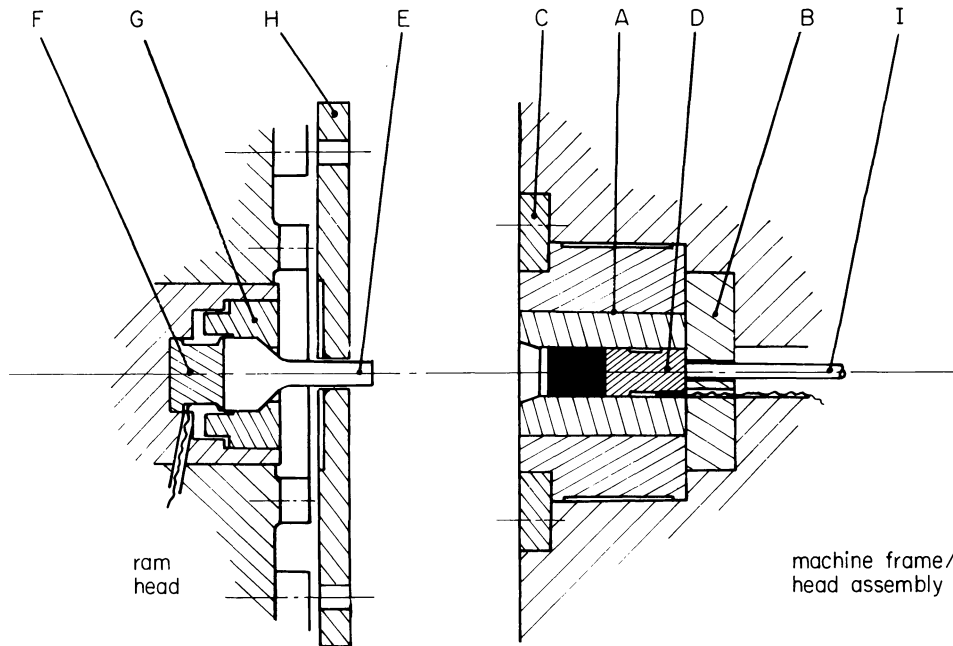


Figure 2. Tooling.

In order to assess the products produced at high speeds, and also to provide a datum for the comparison of tool loading, the tests have been repeated at slow speed. A 300-ton hydraulic press with an incorporated sub-press, was used at a forming speed of 0.0041 ft/s.

The billet material extruded was EN32B steel supplied annealed, phosphated and lubricated with Bonderlube 235 by an industrial firm. The billets had a nominal diameter and length of 1 in and 1.125 in respectively.

Instrumentation

The instrumentation used to record the high-speed data is shown graphically in figure 3.

Load measurement was obtained using two load cells, one mounted directly behind the punch and the other behind the billet. Four 120 Ω strain gauges were mounted on each load cell in a full bridge arrangement, and supplied with a constant d.c. voltage. The load cells were statically calibrated, by transferring the tooling to a 300 ton hydraulic press, at several periods during the investigation.

The ram head displacement, relative to the machine frame/head assembly, was indicated using a capacitance transducer, a resistance potentiometer and/or a high-speed camera.

The capacitance transducer, a coaxial type, was designed according to the principles advocated by Vinycomb and Martyr⁸. It was horizontally connected to the machine immediately above the expansion chamber in which the rear end of the piston cylinder moves. The outer casing was connected to the machine frame/head assembly and the inner rod was connected to the rear of the ram head. A change in capacitance was produced when the inner rod moved relative to the dielectric tubing, which was housed in the outer casing. The change in capacitance was converted to a voltage charge of 1.825 V/in of relative movement, by an oscillator and frequency modulated amplifier. The maximum

relative displacement which the device could measure was 8 in.

The resistance potentiometer was similar in basic design to that used by Kawada et al⁹. An electrical contact attached to the ram head moved along a high-resistance wire which was stretched between the machine frame.

The wire, supplied with a small constant d.c. voltage, formed one arm of a Wheatstone bridge and

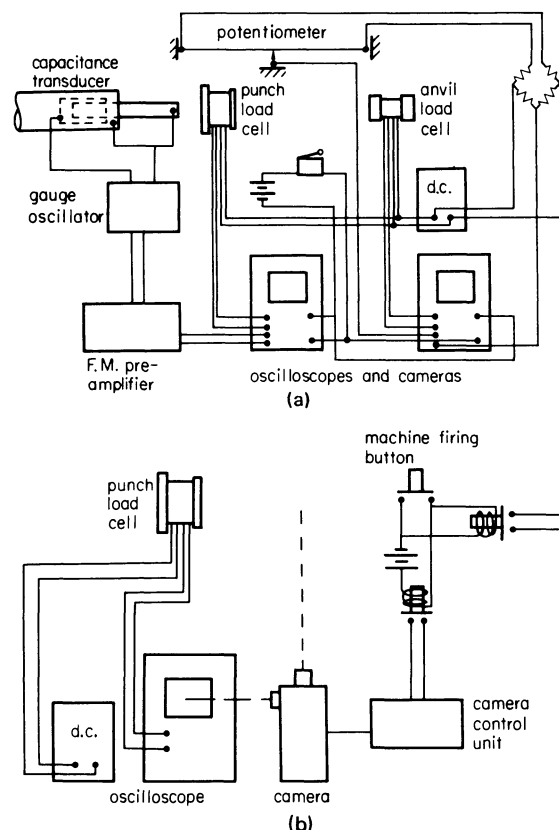


Figure 3. Instrumentation. (a) Instrumentation used to obtain load against displacement traces; (b) instrumentation used during high-speed filming of the machine.

hence a change in the ram position relative to the frame could be converted into a voltage output. This device was used in the initial stages of the research and subsequently for cross reference only.

Both the displacement devices were calibrated in situ, using slip gauges, before each test. Variation in the calibration curve for the capacitance transducer was found to be negligibly small.

The signals from the load cells and the displacement devices were fed to the vertical displays of two oscilloscopes. Hence each oscilloscope gave a trace of load against time and one of displacement against time which were recorded using polaroid film. The system was triggered by the ram head activating a limit switch, and thereby closing a circuit, just prior to impact. Typical punch load and ram displacement traces on a time base are shown in figure 4.

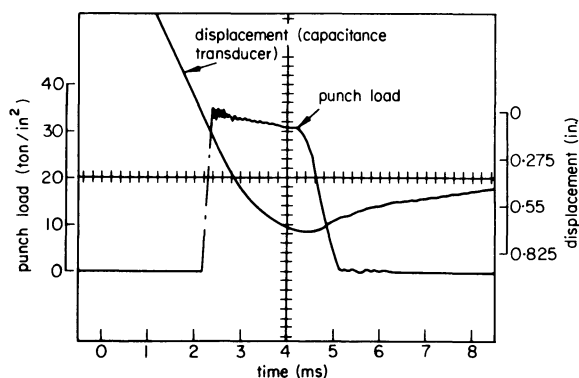


Figure 4. Typical punch-load and ram displacement traces (backward extrusion at a reduction ratio of 1.37).

During part of the investigation, high-speed filming techniques were used to obtain a complete visual record of the ram displacement just prior to, during, and after the extrusion process. The system, which was triggered by the firing button on the machine instrumentation panel, is shown graphically in figure 3(b). The relevant punch-load cell signal was connected to the vertical trace of an oscilloscope with the time base removed, and hence superimposed on the high-speed films.

EXPERIMENTAL RESULTS

The initial research was concerned with obtaining the basic characteristics of the high-speed machine, that is the curve of impact velocity against firing pressure which is shown in figure 5. The velocity was obtained by graphical differentiation of the displacement curve. Also shown plotted in figure 5 is the curve obtained by Jobbins *et al*¹ using inductance probes, and that predicted theoretically for unrestricted adiabatic expansion of the compressed air. It has been suggested that the difference between the experimental and theoretical lines was due to the choking of air through the firing valve and the friction on the ram guide pads and piston seal¹. The slight difference between the two curves derived from experimental data is almost certainly due to the friction on the ram guide pads and piston seal, since these had been renewed before the present authors obtained their experimental curve.

Early process development work was concerned with producing a can from solid 'as received' EN32B billet, by cold backward extrusion at a reduction ratio of 1.37.

A complete visual record of the piston and ram head arm displacement relative to the machine frame/head assembly whilst the machine performed the above operation was obtained by using a high-speed camera. Photographic targets and a scale were attached to relevant points on the machine as shown in figure 6(a). The machine was fired at various firing pressures ranging from 150 psi to 350 psi, giving an impact velocity range of 40 to 54 ft/s. The subsequent films were analysed and, after correction for parallax, the relative displacements were obtained. The impact point was obtained using a punch-load trace superimposed on the film. The piston displacement relative to the machine head

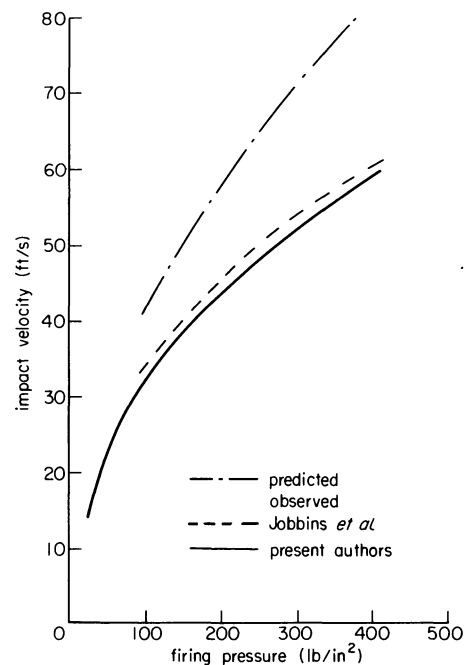
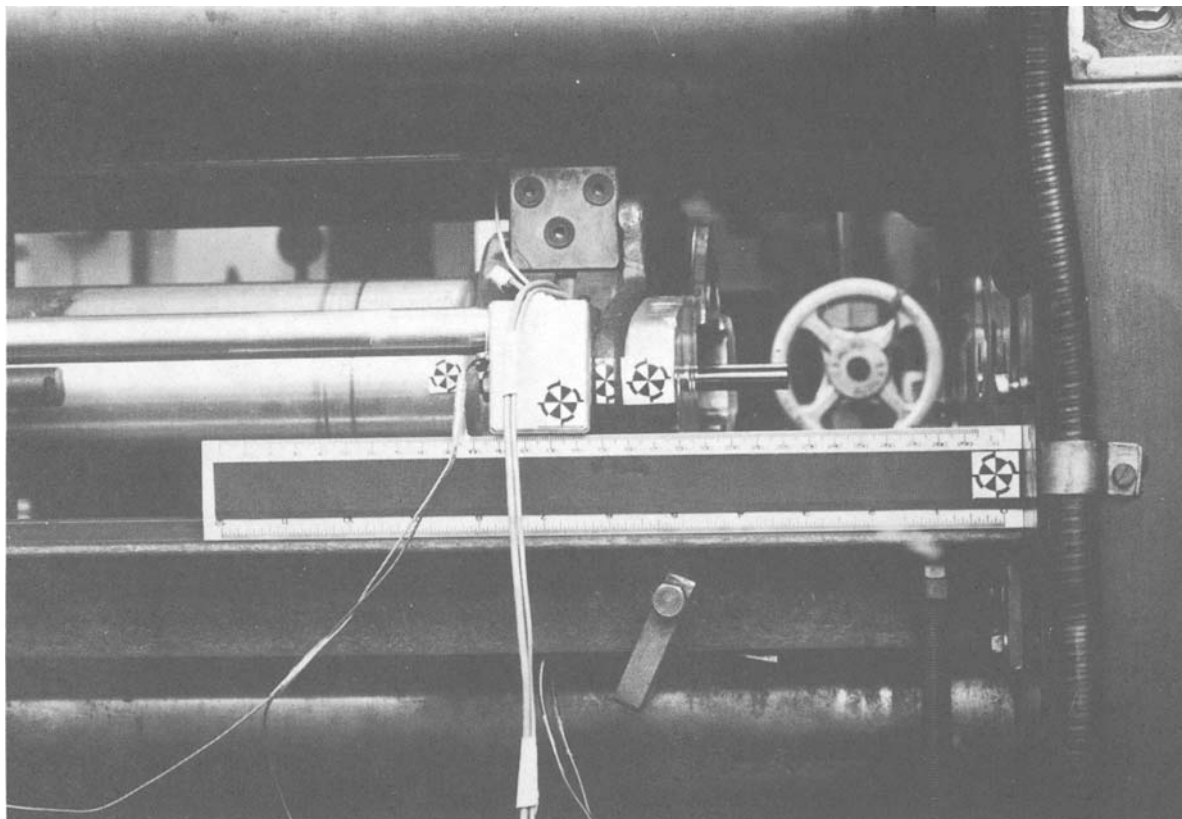


Figure 5.

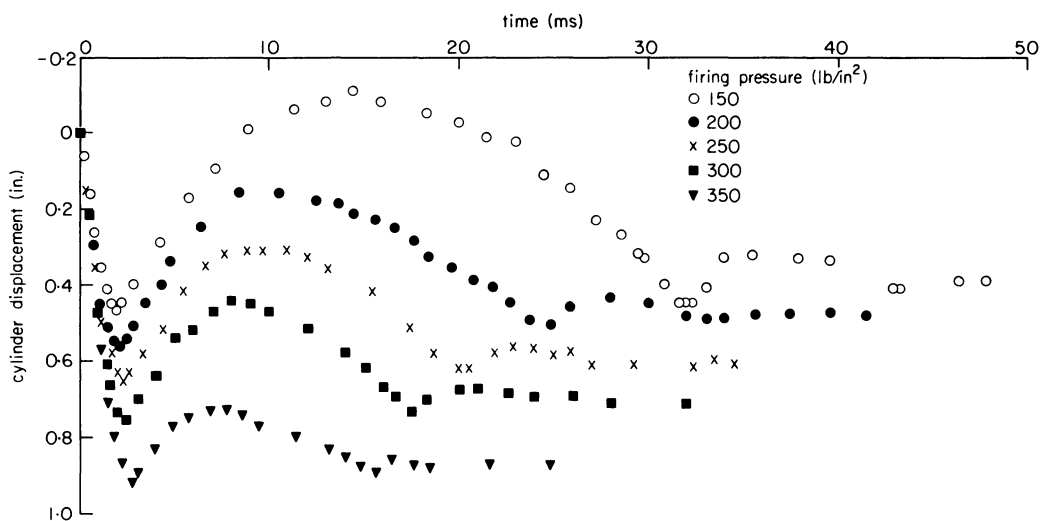
after impact, plotted against time after impact for increasing firing pressures is shown in figure 6(b).

The rebound of the piston is clearly shown, especially at the lowest firing pressure where the magnitude of the initial rebound is greater than the piston displacement during the extrusion process. The effect of the firing pressure on the initial and subsequent rebounds is shown. Increasing the firing pressure causes a decrease in the magnitude and number of rebounds. It is suggested that this is due to an increase in the residual pressure behind the ram in the expansion chamber, as reported by Jobbins *et al*.¹ rather than an increase of the frictional force on the punch nose.

The early process development was plagued by tooling failures, a typical example of which is shown in figure 7(a). As can be seen, a transverse fracture of the punch occurred at a distance approximately $\frac{1}{8}$ in from the tip. On the original tooling this coincided with the edge of the punch land.



(a)



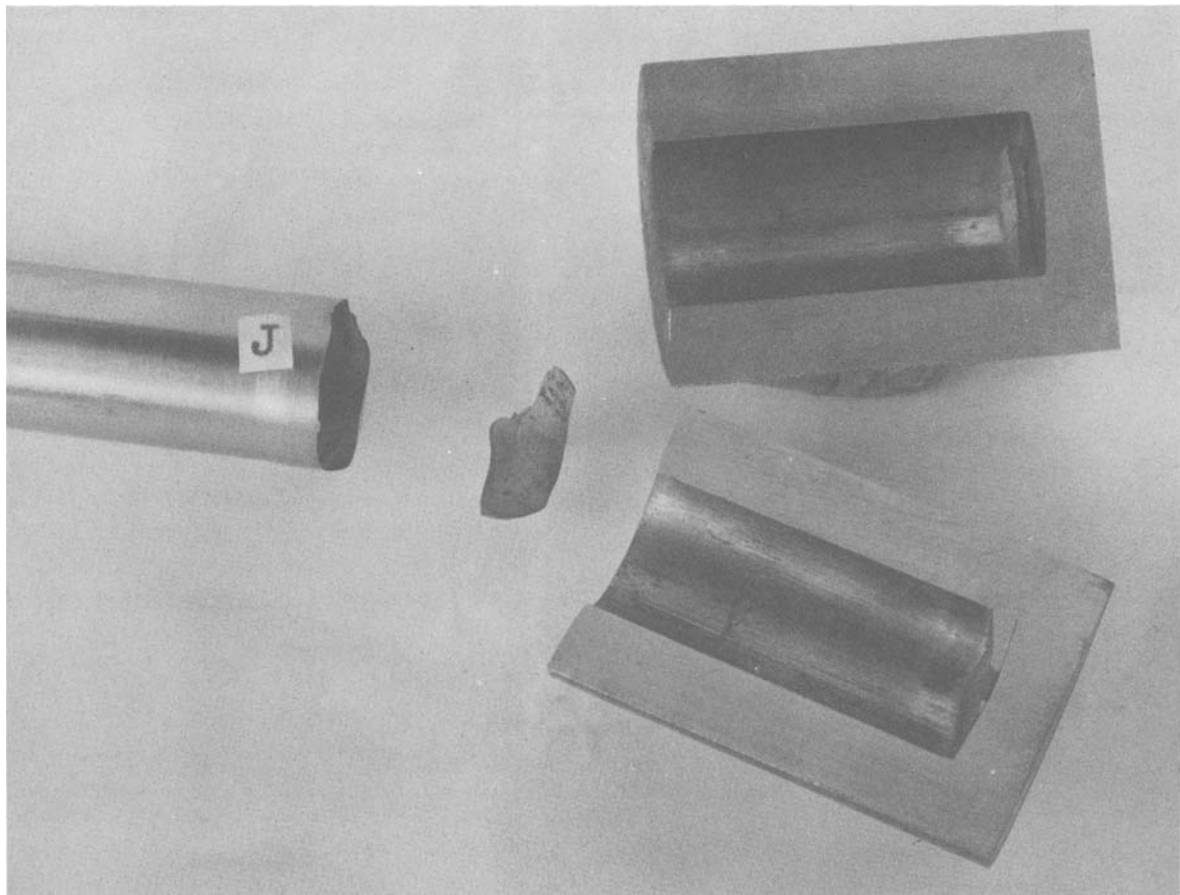
(b)

Figure 6. High-speed filming. (a) A view of the ram head and machine frame head assembly showing the position of the photographic targets and scale; (b) cylinder displacement against time after impact (backward extrusion at a reduction ratio of 1.37).

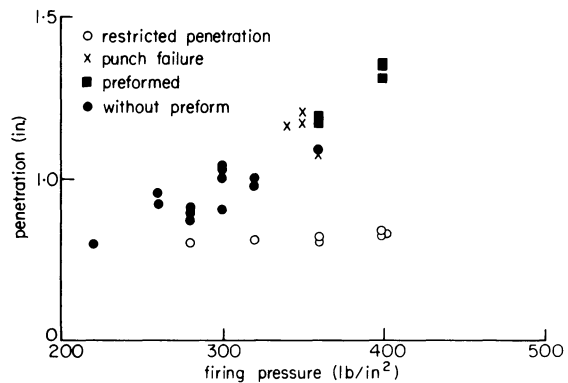
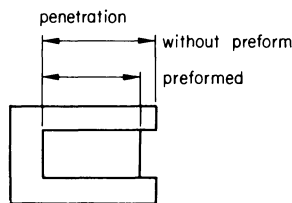
Subsequent tests, using punches with larger lands and varying hardnesses, concluded that neither parameter had any discernible effect on the fracture. The use of hemispherical and conical-nosed punches, although resulting in lower compressive tool loading without tool fracture, were abandoned due to local pick-up on the punch nose and poor concentricity of the resulting cans. Numerous failures of the type mentioned above were experienced, and an indication of their incidence in terms of the process conditions is given in figure 7(b). The plot shows a region of penetration/firing pressure where failure would occur. Tests restricting the punch penetration by using the

clash rings confirmed that a penetration limit rather than an impact velocity limit was being encountered.

Billets whose initial length prevented the critical penetration from being reached were backward-extruded into cans with very thin base walls, using high speeds. The load traces showed no rise above their respective initial peaks, even though cans with base-wall thicknesses less than 0.05 in were produced. Typical products, initial billets and punch-load traces are shown in figure 8. This figure also shows a product produced by combined backward and forward extrusion at high speed, without tool failure or excessive tool loading.



(a)



(b)

Figure 7. Punch failure. (a) Failed punch and relevant-sectioned can. Part of the punch tip can be seen at the base of the product bore. (b) Incidence of punch failure.

It was concluded therefore that the depth of penetration and not the base wall/side wall thickness ratio was the critical tool fracture parameter. This was to some extent confirmed when visual inspection of the punch load/billet interface showed a break-down in lubrication. Not surprisingly there was an indication of lubrication starvation as the depth of penetration increased above a limiting factor. In order to produce the required length of can, improved lubrication was obviously necessary.

Experiments showed that increasing the lubricant thickness on the billet face did not improve the quality of the lubrication since the majority of the lubricant was squeezed out at impact. However, preformed indented billets packed with lubricant gave

improved results, allowing cans of the required length with thin end walls to be produced. To achieve this the 'as received' billets were dumped onto an anvil held in the container so as to create an indent in the billet face, which was then packed with undiluted Molydag 15. Tests to determine a satisfactory indent profile concluded that an indent identical in shape to the punch nose with a depth of 0.1 in was desirable. This allowed a hardened slug of lubricant to cover the punch nose completely, thus giving maximum protection. Interrupted tests indicated that the lubricant trapped below the punch nose was gradually squeezed out as the penetration increased. A typical preformed billet and product with an end wall thickness of 0.1 in. are shown sectioned in figure 8(c).

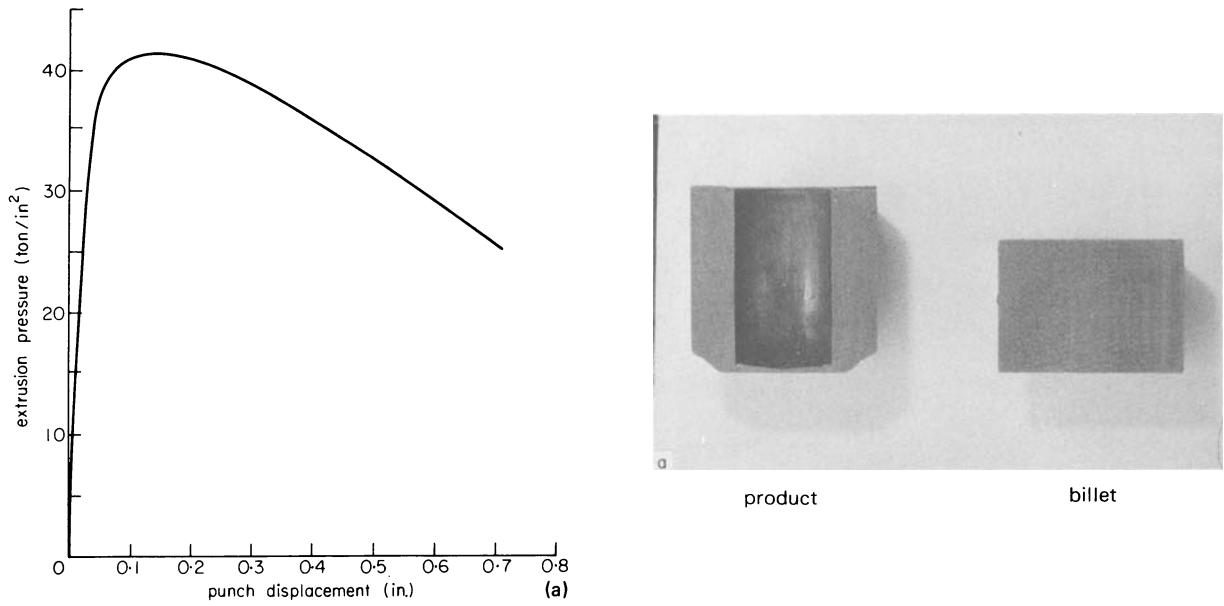
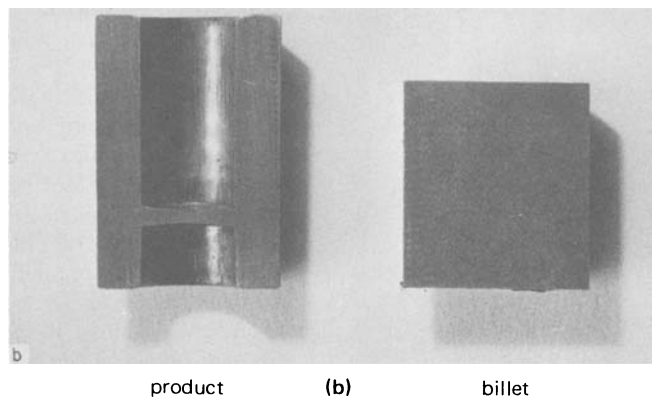
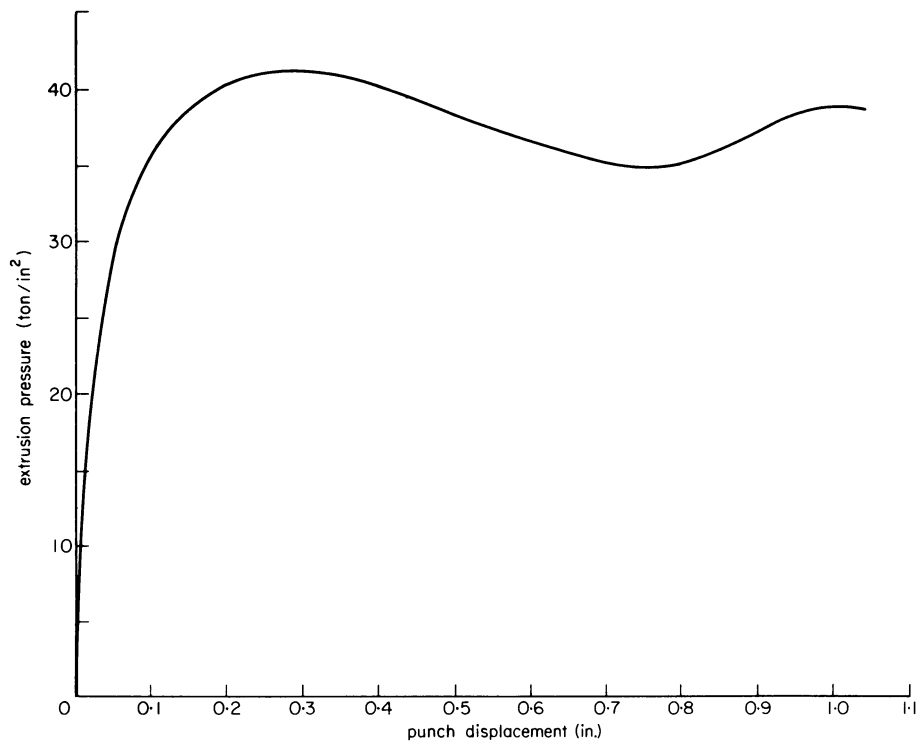
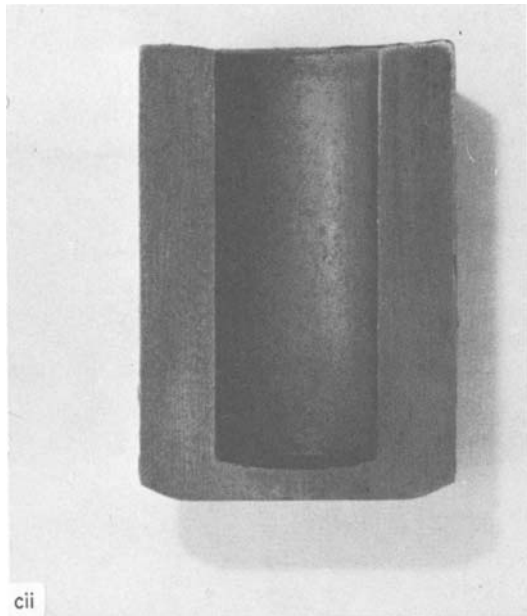
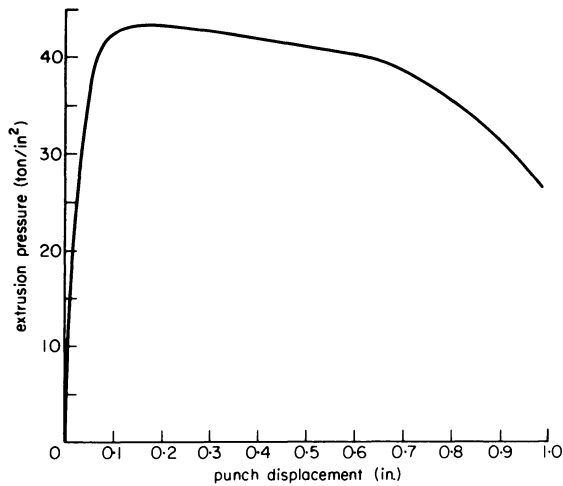


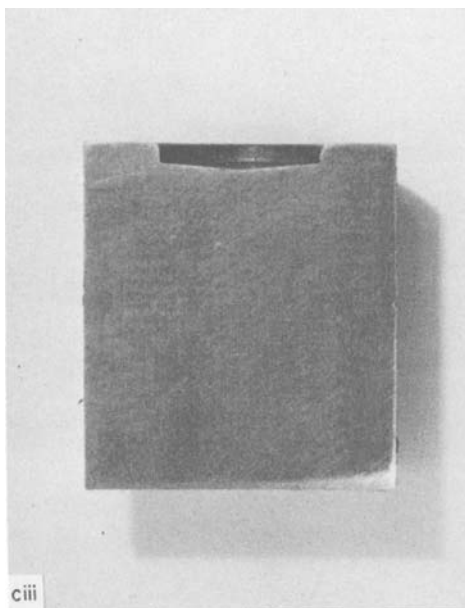
Figure 8. Typical components produced on the high-speed machine. Material: EN32B; Reduction ratio = 1.37. (a) backward extrusion of 'as received' billet. Lubricant: Bonderlube 235. Billet height = 0.725 in.



(b) Combined extrusion of an 'as-received' billet. Lubricant: Bonderlube 235. Billet height = 1.125 in.



product



billet
(c)

(c) Backward extrusion of a preformed billet (with added lubricant). Lubricant: Bonderlube 235 + Molydag 15. Billet height = 1.125 in (before preforming operation).

Comparison of compressive punch loads

Using 'as received' and preformed billets, the indent of the latter being packed with undiluted Molydag 15, cans were backward-extruded at various reduction ratios up to 2.05 using both the high and slow-speed machines.

A typical set of load against displacement traces, obtained when backward-extruding 'as received' billets at a reduction ratio of 1.37, is shown in figure 9. As can be seen there is an obvious difference in shape since the peak on the trace obtained for high-speed extrusion occurs much earlier than the corresponding slow-speed one. Similar differences were noticed when comparing data obtained during the backward extrusion of *preformed* billets, indents of which were packed with lubricant.

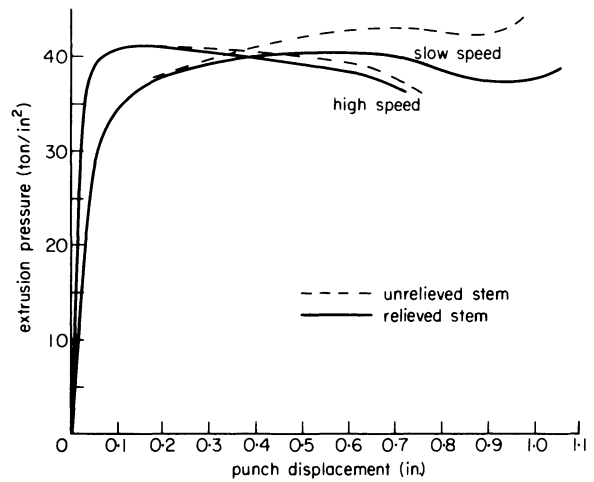


Figure 9. Typical traces of extrusion pressure against displacement traces at high and slow speeds with backward extrusion at a reduction ratio of 1.37. Lubricant: Bonderlube 235. Billet height = 1.125 in.

The initial peak extrusion pressures required when backward extruding 'as received' and preformed and lubricated billets at high and slow speeds are plotted in figure 10. The general trend is for the peak extrusion pressures needed at high speeds to be similar to, or slightly less than, those required at more conventional speeds. It is suggested that the slight variation at a given reduction ratio is due to the small variations in the billet hardness. The initial peak extrusion pressures required for the high-speed backward extrusion of preformed billets with added lubricant were greater than those required to extrude 'as received' billets at a corresponding speed. It is possible that the slight work-hardening of the billet during the preforming operation, as indicated by its increase in hardness, was responsible. At more conventional speeds the initial peak pressures required for the backward extrusion of preformed billets with added lubricant are similar to those required for 'as received' billets at low reduction ratios, and tend to be less at high reduction ratios. The work-hardening of the billet during preforming and the decrease in frictional forces during the backward extrusion process caused by increased lubrication, tend to oppose each other.

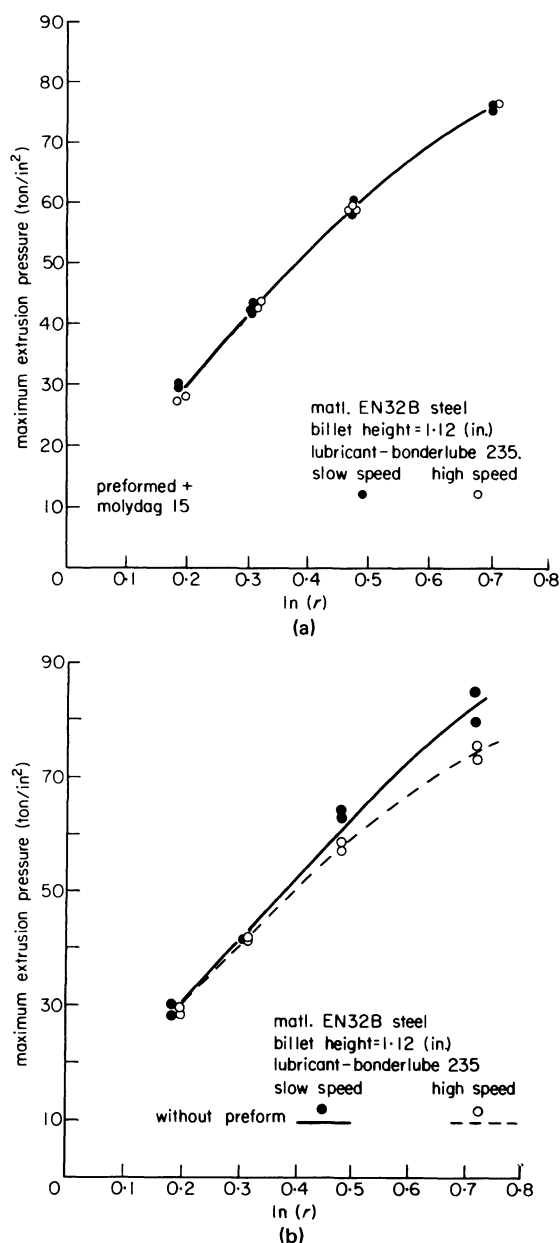


Figure 10. Initial peak extrusion pressures for backward extrusion at high and slow speeds. (a) Preformed indented billets with added lubricant; (b) 'as received' billets.

The effect of the original billet length on the peak extrusion load was negligibly small when extruding at high speeds, but a slight increase with increasing initial billet length was noticed whilst extruding at slow speeds.

A typical trace of extrusion pressure against displacement obtained for the high-speed combined forward and backward extrusion process mentioned earlier is shown in figure 8(b). This can be compared with figure 9 which shows a similar trace for the high-speed backward extrusion of an 'as received' billet at the same reduction ratio. As may be expected the initial rate of increase in the extrusion pressure is lower for the combined process, but the initial peaks are similar.

Velocity effects on friction

In order to assess the effect of velocity on the frictional forces acting at the billet/tooling interface,

an unrelieved punch was used. At a given penetration the difference in punch force required when using an unrelieved punch with that required when using a relieved one was taken as an indication of the friction present.

'As received' billets phosphated and lubricated with Bonderlube 235 were backward-extruded at a reduction ratio 1.37, at both high and low speeds. The data obtained are plotted in figure 9. After the initial rise is completed, the difference between the pair of high-speed traces is seen to be minimal compared with that between the slow-speed pair. This fact was evident throughout the velocity range investigated, between 45 and 53 ft/s.

The effect of velocity on the flow pattern of the material below the punch nose is shown in figure 11. Careful examination shows that the high-speed extruded specimen has a more definite flow pattern and a more intense hemispherical shear zone than that extruded at slow speed.

Strain rate sensitivity of the material

In the case of small extrusion ratios and impact velocities the value of the inertia force is negligibly small^{4,9}. By neglecting the elasticity of the tooling and friction, the observed punch pressure can be considered to be an estimate of the extrusion resistance⁹ and hence to indicate the strain rate sensitivity of the extruding material.

'As received' billets were backward-extruded at various reduction ratios and impact velocities. The peak extrusion pressures (P) and instantaneous velocities (V) obtained are shown in figure 12. The data can be approximated by the equation $P = P_1 V^n$ where P_1 is the extrusion pressure at a velocity of 1 ft/s.

Automatic cycling

Having solved the initial tooling problems attempts were made to use the high-speed machine in its automatic mode.

'As received' En32B billets, phosphated and lubricated with Bonderlube 235, were preformed on the high-speed machine and subsequently lubricated with undiluted Molydag 15. The preformed billets were then backward-extruded into cans using the automatic cycle of the machine. An average cycle time of 15 seconds was obtained on the machine for a few hours. Cans having an end wall thickness of less than 0.1 in and a reduction ratio of 1.37 were produced using a nominal energy of 6540 ft lbf per blow. Within the range allowed by the variation in billet properties and dimensions, cans of consistent dimensional accuracy was produced. A product similar to those produced by automatic mass production is shown in figure 8(c). The exact depth of penetration of cans was controlled by the clash rings on the ram head and machine frame which helped dissipate excess kinetic energy. In order to minimize the excess energy it was essential that the variation in billet properties and dimensions in any one batch were as small as possible. In general the cans remained in the container when the punch was withdrawn, and hence were ejected by the hydraulic ram and not by the stripper plate.

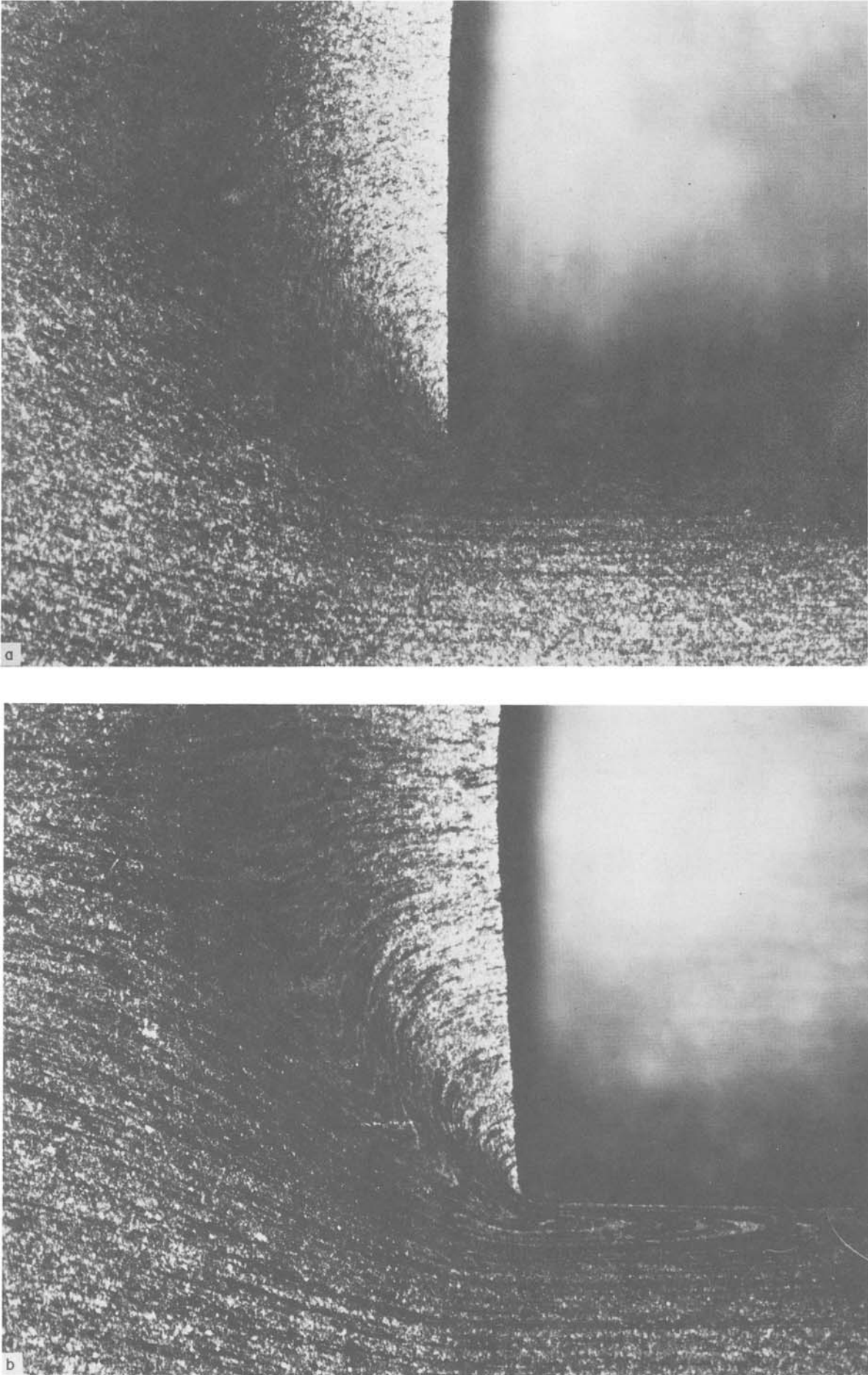


Figure 11. Flow patterns at the intersection of the can base and side wall during backward extrusion. (a) Slow speed, (b) high speed.

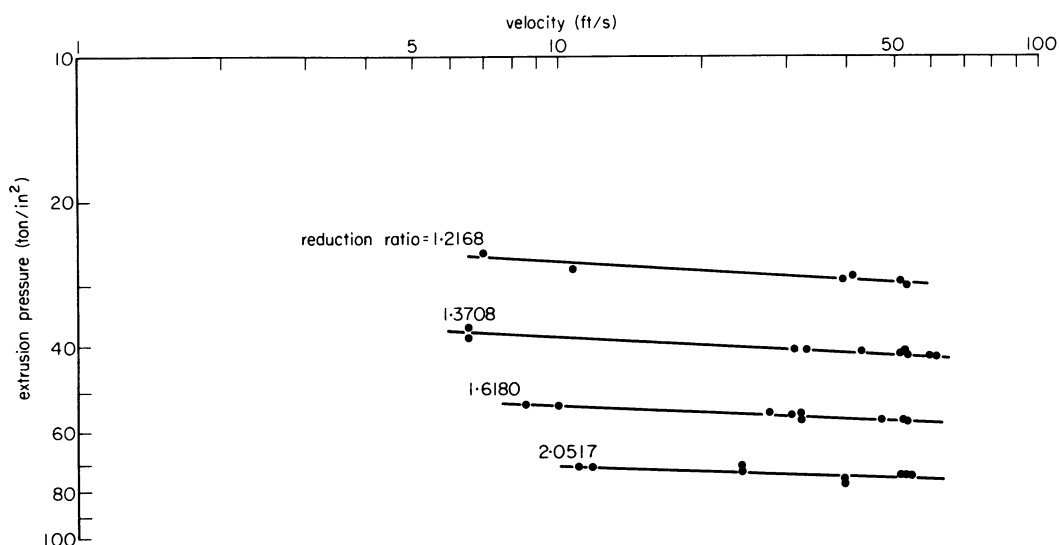


Figure 12. The effect of ram velocity (V) on the extrusion pressure (P) for backward extrusion. Lubricant: Bonderlube 235. Billet height = 1.125 in.

The temperature of the finished product ejected from the container was checked periodically using heat-sensitive crayons. An attempt at using heat-sensitive paints on the tooling was abandoned due to the formation of a covering layer of lubricant. The temperature of the product was found to increase with increasing punch penetration and, up to a limit, with the number of products produced. The temperature of the twentieth product was between 75 and 100°C, and remained constant from then on.

Provided that accurately preformed billets were used, the concentricity of the finished product was comparable with those at slow speed under test conditions. No apparent deterioration of the tooling was noticed during the automatic run and in particular the land remained free of scoring.

DISCUSSION OF RESULTS

The feasibility of using high-speed machines to produce cans with base walls less than 0.05 in thick without undue punch pressure has been adequately demonstrated. The use of preformed billets to improve the efficiency of the lubricant and hence allow longer cans to be produced has also been adequately demonstrated. Combined forward and backward extrusion, to produce a tube with a thin centre web, has been shown to be feasible at high speeds.

In the backward extrusion of steel the initial peak tooling pressure is important since in general it indicates the limit of the process. Therefore it is of practical significance that the initial peak tooling pressures required for high-speed backward extrusion are similar to or slightly lower than those required at conventional speeds. Similar comparisons have been made previously and similar conclusions drawn, except in the case of backward extrusion of EN2E⁵ where a reverse trend was evident.

Under certain conditions of extruding velocity and reduction ratio, inertia forces required to accelerate the extruding material can be significant⁴. Within the range of impact velocities and reduction ratios used

during the reported investigation, the inertia forces were negligibly small. For this case the basic differences in loads required for high and low-speed backward extrusion are concerned with the extruding material and the area of contact between the material and tooling where relative motion occurs.

For a given geometric configuration, increasing the velocity of deformation produces increased strain rates, and in general, especially for low carbon steels, this causes an increase in the resistance to deformation of the extruding material¹⁰. Tests on 0.03 per cent carbon steel within the strain rate range 2–200/s concluded that the ratio of the dynamic flow stress to the static flow stress at 2 per cent strain was 1.2 to 1.6, and the ratio of the dynamic yield stress to the static yield stress was 1.6 to 2.5¹⁰. The information presented in the present paper, figure 12, as an assessment of the strain rate sensitivity of the billet material suggests that within the strain rate range used the increase in deformation resistance is of the order mentioned above. The data can be approximated by the equation $P = P_1 V^n$ where P is the extrusion pressure at velocity V and P_1 is the extrusion pressure at 1 ft/s. A similar conclusion was obtained by Kawanda et al.⁹ for 0.25, 0.5 and 1.1 per cent carbon steels in the temperature range 900–1150°C.

Reports on high-speed forging processes^{3,5,6,9,11,12} have concluded that frictional forces are reduced as the relative speed of sliding between the tooling and deforming material is increased. Osakada et al.¹¹ using simple compression, reported a 50 per cent decrease in the coefficient of friction for stearic soap on a phosphate base when the speed of deformation was increased from 0.0033 ft/s to 33 ft/s. This information tends to support our findings, indicating a reduction of friction with increasing speed of deformation. It is suggested that the very small differences in the flow patterns were a result of the reduced friction and more localized heat generation which are apparent in high-speed forging⁶.

Therefore it is evident that, under the experimental conditions reported in this paper, the

increase in punch loading that would occur due to the increase in flow stress of the deforming material is counteracted by the decrease in frictional forces in the area of relative motion as the extrusion velocity increases. Whether the tool loads encountered during the high-speed extrusion processes are greater than or less than those required at more conventional speeds depends on the relative magnitude of the change in the above parameters as the extrusion velocity is increased.

The feasibility of using the high-speed machine to mass-produce large quantities of steel cans with very thin base walls from preformed and lubricated billets has been clearly demonstrated. The use of accurately preformed billets ensured that cans of comparable concentricity to the products produced at more conventional speeds were obtained. The added lubricant (undiluted Molydag 15) prevented any appreciable deterioration of the tooling. Heat generation in the product and tooling was slightly lower than that previously noted on a conventional machine¹³ and therefore did not pose any problems.

SUMMARY OF CONCLUSIONS

- (1) The use of preformed billets to improve the lubrication and concentricity of the product, and hence to allow cans with thin base walls to be automatically mass-produced using a high-speed machine, has been demonstrated.
- (2) The use of combined forward and backward extrusion to produce a steel tube with a thin centre web has been shown to be feasible at high speeds.
- (3) The surface finish and concentricity of the products from the high-speed processes were comparable with those produced at conventional speeds.
- (4) In the backward extrusion of steel the initial peak extrusion loads required at high speeds were similar to, or slightly lower than those required at conventional speeds.
- (5) The initial peak extrusion loads required for high-speed backward extrusion of preformed

billets with added lubricants were slightly higher than those required when using 'as-received' billets.

REFERENCES

1. B. Jobbins, B. Parsons and B. N. Cole. An automatic high-speed forming machine for cold-forming studies. *Proc. 11th Int. MTDR Conf.*, Pergamon, 1970.
2. R. Davies and S. M. Dhawan. Comparison of some strain rate effects in slow and high speed blanking, *Proc. 9th Int. MTDR Conf.*, Pergamon, 1969.
3. E. R. Austin, R. Davies and F. Bakhtar. High speed impact forming of metals. *Proc. Inst. Mech. Engrs.* (1967-8), no. 1, 182, 188.
4. B. N. Cole and F. Bakhtar. Dynamic effects in very high speed impact extrusion. *Int. J. Mach. Tool Des. Res.* (1963) 3, 77.
5. R. Davies and A. Singh. The production of components by cold extrusion using high energy rate techniques. *Advances in Mach. Tool Des. Res.*, Pergamon, 1971.
6. M. G. Jones, R. Davies and A. Singh. Some high speed cold forging operations, *Proc. 9th Int. MTDR Conf.*, Pergamon, 1969.
7. B. Parsons, B. Jobbins and B. N. Cole. *Prototype automatic H.E.R.F. machine for cold forming operations*, Internal Dept. Research Report, ref. no. 1, 1970.
8. R. K. Vinycomb and F. E. Martyr. Design of capacitance displacement transducers. *Instrument Practice* (1956) 10, 985.
9. T. Kawada, M. Suzuki, S. Ikoda and T. Dendo. Fundamental experiments and an application to the impact extrusion of steel. *Proc. 7th Int. MTDR Conf.*, Pergamon, 1966.
10. K. Tamara, T. Matsuo, M. Kinoshita and T. Maode. Strength of mild steel at high strain rates. *JSME* (1966) 9, no. 33, 21.
11. K. Osakada and M. Oyane. The effect of deformation speed on friction and lubrication in cold forging, *JSME* (1970) 13, no. 66, 1504.
12. A. T. Male. Variations in friction coefficients of metals during compressive deformation. *J. Inst. Metals* (1966) 94, 121.
13. A. K. Crudon. *Temperature measurements during cold extrusion of steel*. NEL. Report, no. 52, 1962.

FORWARD COLD EXTRUSION OF ALUMINIUM—ESTIMATION OF THE MAXIMUM EXTRUSION FORCE

by

L. E. FARMER* and B. Z. PIROG†

SUMMARY

The design of an experiment is discussed for assessing the effect of reduction ratio, die angle and billet length to diameter ratio on the maximum extrusion pressure for forward cold extruding aluminium alloy 6063 in its zero temper condition. It is shown that the main effects, and the first order interactions of each of these factors, have a significant effect on the maximum extrusion pressure. An equation is derived using orthogonal polynomials which include terms for all the statistically significant effects for predicting this pressure. A design chart is also constructed which consists of correction factor curves for each of the above three factors. It is shown that both methods predict the maximum extrusion pressure with a similar accuracy, even though the latter model is considerably simpler than the first.

INTRODUCTION

The usual tooling arrangement, employed to forward cold extrude metals, is shown in figure 1, together with a diagram indicating the expected variations in punch load as the billet is extruded. The maximum reached by this force in the early stages of the extrusion is of considerable practical significance as it indicates the minimum size of press required for the operation as well as the stress conditions that are most likely to cause the tooling to fail. Methods of predicting this force have been the subject of numerous investigations. The results of this work are usually presented in the form of design charts or equations. Examples of the equations are given in the paper by Boes and Pouw¹ while two of the better known design charts are by McKenzie² and Feldmann³.

A large proportion of this work has been related to the cold extrusion of steel and the relatively pure grades of aluminium. While the expected behaviour of the more highly alloyed types of aluminium during hot extrusion is well documented, published information relating to their cold extrusion performance is very limited.

This paper describes an investigation into the forward cold extrusion of one of these alloys. An experiment is planned for determining the effect of die angle, θ , extrusion ratio, r , and billet length to diameter ratio, L_0/D_0 , as defined in figure 1, on the maximum extrusion pressure, P . A statistical analysis is made of the data to assess whether the above

mentioned factors have a significant effect on the maximum extrusion pressure. An equation, based on the significant factors, is then derived for predicting the maximum extrusion pressure.

Finally a design chart is constructed which consists of a series of correction factor curves for each significant effect and the accuracy of the pressures calculated by both methods are compared.

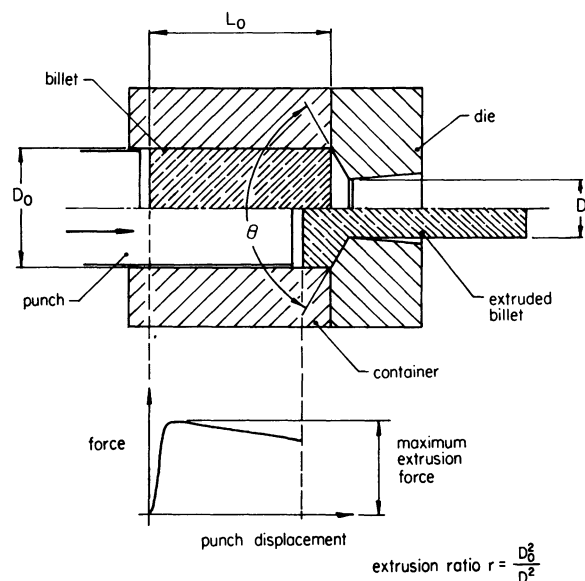


Figure 1. Tooling layout and force diagram for forward cold extrusion.

* School of Mechanical and Industrial Engineering, The University of New South Wales, Australia

† Industrial Engineering Department, A.C.I., Sydney, Australia

PLAN OF EXPERIMENT

If all other factors remain constant then it is assumed that the maximum extrusion pressure, P , is a function of three variables; extrusion ratio, r , included die angle, θ , and the billet length to diameter ratio L_0/D_0 .

$$P = f(r, \theta, L_0/D_0) \quad (1)$$

The form of Equation (1) is unknown, however in an investigation at NEL into the cold extrusion of steel it was shown by McKenzie² that P could be estimated from an equation of the form

$$P = q_\theta q_{L_0/D_0} q_{LUB} P_{basic} \quad (2)$$

where q_θ and q_{L_0/D_0} were obtained from correction factor curves for given values of θ and L_0/D_0 , and q_{LUB} was a correction factor for the type of lubricant. The basic extrusion pressure, P_{basic} , was also derived from a curve as a function of the extrusion ratio. Die angle, L_0/D_0 ratio and lubrication conditions were defined for this curve, and these levels had a value of 1.0 on their respective correction factor scales. As it is planned to use only one lubricant in these experiments, the correction factor q_{LUB} in Equation (2) was assigned a value of unity in this paper.

An important feature of the above approach is that only the main effect of each factor is considered when calculating the maximum extrusion pressure. This assumes that other effects, such as interactions between the factors, are included by the way in which the main effects are combined in Equation (2). Therefore, prior to constructing such charts, an analysis was made to determine the significance of the main effects being plotted and whether interactions exist between the factors.

A factorial design of experiment is well suited to the above requirements because not only can the significance of the main effects of the factors and their interactions be easily assessed, but the data is presented in a convenient form for deriving correction factor curves. To illustrate this latter point, consider the factorial experiment where the maximum extrusion pressure P is recorded for each combination of the factors r , θ and L_0/D_0 which have l , m and n levels respectively. The correction factor, $q_{\theta i}$ for the i th level of θ is

$$q_{\theta i} = \frac{\bar{P}_{\theta i}}{P_{mean}} \quad (3)$$

where $\bar{P}_{\theta i}$ is the mean of the ln values of P in the table that were recorded with θ at the level i , and P_{mean} is the mean of all the tabulated values. Plotting the m values of q_θ against their respective levels of θ results in the die angle correction factor curve. The same procedure is followed for the L_0/D_0 ratio factor. Finally the P_{basic} curve is a plot of the l values of \bar{P}_{r_i} against r_i .

An obvious modification to this model is to replace the basic extrusion pressure by a correction factor q_r and multiply the three factors by the mean pressure P_{mean}

$$P = q_r q_\theta q_{L_0/D_0} P_{mean} \quad (4)$$

The model for predicting a value of P from the factorial experiment described above is

$$P_{ijk} = P_{mean} + Mr_i + M\theta_j + M \frac{L_0}{D_0} + Fr_i\theta_j + Fr_i \frac{L_0}{D_0} + F\theta_j \frac{L_0}{D_0} + Sr_i\theta_j \frac{L_0}{D_0} \quad (5)$$

where terms beginning with M are the main effect contributions for factors r , θ and L_0/D_0 at levels i , j and k , and the remaining F and S terms are the contributions from the first and second order interactions.

The expected accuracy of Equation (4) would be inferior to that of Equation (5) where there are significant interactions between factors. Assuming the interactions are not significant, then the main effects are independent, and Equation (5) reduces to

$$P = P_{mean} (1 + (q_r - 1) + (q_\theta - 1) + (q_{L_0/D_0} - 1)) \quad (6)$$

Equation (4) and (6) give approximately the same results for correction factors close to 1.0 but the expected accuracy of Equation (4) would be inferior to (6) where there are no significant interactions. Equation (5) can be represented by

$$P = P_{mean} + \sum_{ri=1}^l \beta_{ri} \phi_{ri}(Xr) + \sum_{\theta i=1}^m \beta_{\theta i} \phi_{\theta i}(X_\theta) + \sum_{(L_0/D_0)i=1}^n \beta_{(L_0/D_0)i} \phi_{(L_0/D_0)i}(X_{L_0/D_0}) + \sum_{ri=1}^l \sum_{\theta i=1}^m \beta_{ri, \theta i} \phi_{ri}(Xr) \phi_{\theta i}(X_\theta) + \sum_{ri=1}^l \sum_{(L_0/D_0)i=1}^n \beta_{ri, (L_0/D_0)i} \phi_{ri}(Xr) \phi_{(L_0/D_0)i}(X_{L_0/D_0}) + \sum_{\theta i=1}^m \sum_{(L_0/D_0)i=1}^n \beta_{\theta i, (L_0/D_0)i} \phi_{\theta i}(X_\theta) \phi_{(L_0/D_0)i}(X_{L_0/D_0}) + \sum_{ri=1}^l \sum_{\theta i=1}^m \sum_{(L_0/D_0)i=1}^n \beta_{ri, \theta i, (L_0/D_0)i} \phi_{ri}(Xr) \phi_{\theta i}(X_\theta) \phi_{(L_0/D_0)i}(X_{L_0/D_0}) \quad (7)$$

where P is the expected response, and P_{mean} the mean of the tabulated values. The single, double and triple summations represent the main effects and the first and second order interactions respectively. The first summation term consists of the coefficient β_{ri} and the i th order orthogonal polynomial $\phi_{ri}(Xr)$ for the factor r . If the experiment is designed so that there are equal intervals between the levels of the factors then tabulated⁴ values of these polynomials are available to simplify calculations.

To determine the significant orders of these polynomials, an analysis of variance is first made of the data to indicate the effects in Equation (5) that make a significant contribution to the maximum extrusion pressure. The sums of squares of the significant effects are partitioned⁵ into linear, quadratic, cubic, etc., components which correspond to the first,

second and third order orthogonal polynomials. The significant orthogonal polynomial components and their coefficients are evaluated as described by Davies⁵ and substituted into Equation (7) to give a relationship between P and the original variables r , θ and L_0/D_0 . It is possible to construct correction factor curves for the main effects and interactions from this equation.

The factorial design that was chosen consisted of 4 levels of extrusion ratio and 3 levels of die angle and L_0/D_0 ratio. Each experiment was replicated twice to give a separate estimate of the experimental error. The levels selected for each factor are listed below and they cover the range normally encountered in practical extrusion operations.

Extrusion ratio $r = 1.5625, 2.0408, 2.7778, 4.0000$

Die angle $\theta = 180^\circ, 135^\circ, 90^\circ$

Billet length to diameter ratio $L_0/D_0 = 3, 2, 1$

The uneven spacing between the extrusion ratio levels resulted from dies that had equal increments between their diameters. To retain the generality of the term, extrusion ratio, the square root of the reciprocal of r was used in the orthogonal polynomial calculations. That is r becomes $(1/r)^{1/2}$ in Equation (7), and has levels of 0.8, 0.7, 0.6, and 0.5 in the experiment.

EXPERIMENTAL METHOD

The work material, supplied in 6 m lengths of 25.4 mm diameter extruded bar, was aluminium alloy 6063 that has a nominal composition of

Si = 0.3–0.6 per cent, Fe = 0.35 per cent, Cu = 0.10 per cent, Mn = 0.10 per cent, Mg = 0.45–0.9 per cent, Cr = 0.10 per cent, Zn = 0.10 per cent, Ti = 0.10 per cent, other elements = 0.15 per cent, Al = remainder.

These bars were cut into billets of the required length and numbered consecutively. This allowed the billets to be randomly allocated to extrusion ratio and die angle combinations so as to reduce any possible effect due to variations in material properties in the bars. The billets were annealed to a hardness of 21.5 HV₅ by soaking for 3 hours at 410°C, then cooling at a maximum rate of 20°C per hour to 290°C. Below this temperature the cooling rate was uncontrolled.

After annealing, the billets were lubricated with the Bonderite 170X/Bonderlube 234 process. The porous crystalline phosphate coating, deposited by this process, tightly adheres to the surface of the billet and provides not only a good base for carrying the stearate lubricant, but also a barrier that prevents the aluminium coming into contact with the die and hence reduces the occurrence of 'Pick Up'. This process is similar to the phosphate/stearate lubricants that are extensively used for cold extruding steel.

The extrusion die inserts were made from a high carbon-high chrome steel (2 per cent C, 12 per cent Cr, 0.3 per cent Si, 2 per cent Mn) that was hardened and tempered to 60 Rc. Details of the die design, which is similar to that of NEL⁶ are given in figure 2. Support rings were not fitted to the inserts because the stresses developed in the dies were relatively low.

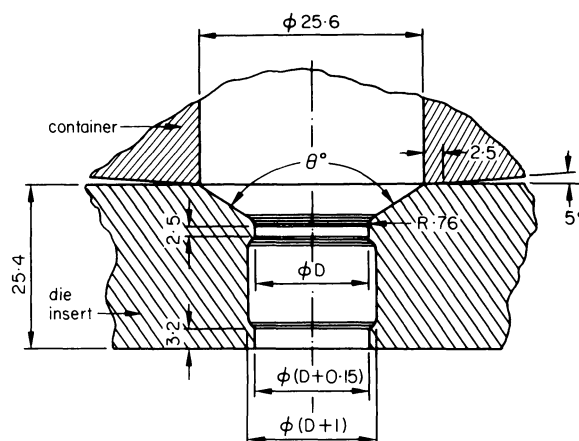


Figure 2. Details of the die profile.

To enable easy cleaning, inspection and changing of the die inserts and container, a sub assembly, consisting of these two components locked in a holder, was designed⁷ so that it could be easily attached and removed from its mounting on a load cell in a die set. The die set was located in a 500 kN capacity hydraulic press which had a ram speed set at 15.24×10^{-3} m/s under no load conditions. This speed decreased with load to approximately 10.15×10^{-3} m/s under maximum load. It was shown by Pugh et al⁸ that variations in speed over this range should have little effect on the maximum extrusion pressure.

The punch force, velocity and displacement were recorded on the X , X_1 , and Y axes of a 2 pen EAI graph plotter. The output from semiconductor strain gauges on the load cell, and an LVDT displacement transducer mounted between the dieset platens, were passed through an SEL 4000 carrier system to the X and Y axes of the plotter while the Hewlett Packard linear velocity transducer attached to the press was connected directly to the X_1 axis of the plotter.

EXPERIMENTAL RESULTS AND ANALYSIS

The maximum extrusion forces for the 72 tests were divided by the cross section area of the billet and recorded in Table 1. Defects such as *chevrons* and *pickup* on the dies were not encountered during the tests. Dead metal zones occurred only on billets extruded through dies with included angles of 180° .

The results of an analysis of variance of the data is given in Table 2 where the numbers in brackets are the total sums of squares and mean squares for each main effect and interaction. All three main effects and their first order interactions were significant at the 1 per cent level and the second order interaction sum of squares was pooled with the residual sum of squares for the remaining calculations.

The sums of squares of the significant effects were partitioned into their linear and quadratic etc. components. Further variance ratio tests between the mean squares of these components and the pooled residual mean square were made, and the significant components are indicated by an asterisk in Table 2. The linear, quadratic and cubic components of the main effect of $(1/r)^{1/2}$ were significant. Generally

TABLE 1 Maximum extrusion pressure P (MPa)

$(1/r)^{1/2}$	Repl.	L_0/D_0 Ratio								
		3 Die angle θ			2 Die angle θ			1 Die angle θ		
		180°	135°	90°	180°	135°	90°	180°	135°	90°
0.8	1	242	249	193	224	236	186	208	211	165
	2	249	230	193	224	218	180	205	199	165
0.7	1	329	317	261	305	301	242	280	273	224
	2	361	308	267	311	295	242	280	261	224
0.6	1	429	398	329	385	373	305	348	323	292
	2	429	404	342	398	373	305	357	323	280
0.5	1	541	497	423	500	491	398	451	410	373
	2	541	528	435	497	475	407	451	419	370

when $\log r$ is plotted, this relationship is sensibly linear. Had the spacing between the levels of r allowed equal increments of $\log r$ to be used instead of $(1/r)^{1/2}$ then this effect might have been considerably simplified.

Each term in Equation (7) was evaluated by applying the tabulated⁴ orthogonal polynomials to the significant components of the effects in Table 2. The significant terms of Equation (7) are given in

Table 3 while graphs of the three main effects are plotted in figures 3, 4 and 5. The three surfaces for the interactions are not shown because they are of little practical significance in this case.

Correction factor scales have been added to figures 3, 4 and 5 to enable them to be used directly with the model given by Equation (4). Correction factor values were also calculated from the means of the levels of each factor and plotted with those given by the

TABLE 2 Analysis of variance of maximum extrusion pressure data

Source of variance	Degrees of Freedom	Sum of squares	Mean square	Variance ratio
<i>Main effects:</i>				
$(1/r)^{1/2} = R$				
Linear R_1	1	592 273.3	592 273.3	9647.9*
Quadratic R_2	1	3612.5	3612.5	58.8*
Cubic R_3	1	677.9	677.9	11.0*
Sum	3	(596 563.7)	(198 854.6)	3239.3
θ				
Linear θ_1	1	63 365.3	63 365.3	1032.2*
Quadratic θ_2	1	5353.4	5353.4	87.2*
Sum	2	(68 718.7)	(34 359.3)	559.7
$L_0/D_0 = L$				
Linear L_1	1	41 008.5	41 008.5	668.0*
Quadratic L_2	1	166.8	166.8	2.7
Sum	2	(41 175.4)	(20 587.7)	335.4
<i>First order interactions:</i>				
$R \times \theta$				
$R_1 \times \theta_1$	1	4183.4	4183.4	68.1*
$R_1 \times \theta_2$	1	0.1	0.1	0.001
$R_2 \times \theta_1$	1	60.8	60.8	1.0
$R_2 \times \theta_2$	1	121.0	121.0	2.0
$R_3 \times \theta_1$	1	8.1	8.1	0.1
$R_3 \times \theta_2$	1	0.3	0.3	0.004
Sum	6	(4373.5)	(728.9)	11.9
$R \times L$				
$R_1 \times L_1$	1	3928.5	3928.5	64.0*
$R_1 \times L_2$	1	71.6	71.6	1.2
$R_2 \times L_1$	1	4.7	4.7	0.1
$R_2 \times L_2$	1	60.1	60.1	1.0
$R_3 \times L_1$	1	4.5	4.5	0.1
$R_3 \times L_2$	1	1.0	1.0	0.02
Sum	6	(4069.9)	(678.3)	11.0
$\theta \times L$				
$\theta_1 \times L_1$	1	1140.0	1140.0	18.6*
$\theta_2 \times L_2$	1	0.5	0.5	0.01
$\theta_1 \times L_1$	1	184.3	184.3	3.0
$\theta_2 \times L_2$	1	467.7	467.7	7.6
Sum	4	(1792.5)	(448.1)	7.3
<i>Second order interaction:</i>				
$R \times \theta \times L$				
Sum	12	(657.6)	(54.8)	0.9
<i>Residual (expt. error):</i>				
Sum	36	(2210.0)	(61.4)	

TABLE 3 Significant terms in equation (7)

Term	Evaluated form
P_{mean}	325.8056
$\sum_{r_i=1}^l \beta_{ri} \phi_{ri}(Xr)$	$2012.708 - 7432.67 \left(\frac{1}{r}\right)^{1/2} + 9626.2999 \frac{1}{r} - 4573.333 \left(\frac{1}{r}\right)^{3/2}$
$\sum_{\theta_i=1}^m \dots$	$-261.424 + 3.2462\theta - 0.00903\theta^2$
$\sum_{(L_0/D_0)_i=1}^n \dots$	$-58.458 + 29.229 L_0/D_0$
$\sum_{r_i=1}^l \sum_{\theta_i=1}^m \dots$	$-162.825 + 1.206\theta + 250.5 \left(\frac{1}{r}\right)^{1/2} - 1.8555\theta \left(\frac{1}{r}\right)^{1/2}$
$\sum_{r_i=1}^l \sum_{(L_0/D_0)_i=1}^n \dots$	$-105.196 + 161.84 \left(\frac{1}{r}\right)^{1/2} + 52.598 L_0/D_0 - 80.92 \left(\frac{1}{r}\right)^{1/2} L_0/D_0$
$\sum_{\theta_i=1}^m \sum_{(L_0/D_0)_i=1}^n \dots$	$354.314 - 5.3613\theta - 400.107 L_0/D_0 + 6.2478\theta L_0/D_0 + 0.189\theta^2 + 95.55 \left(\frac{L_0}{D_0}\right)^2 - 0.0227 \frac{L_0}{D_0} \theta^2 - 1.5288\theta \left(\frac{L_0}{D_0}\right)^2 + 0.00566\theta^2 \left(\frac{L_0}{D_0}\right)^2$

orthogonal polynomial analysis in figures 3, 4 and 5. As expected there is very little difference between the two sets of results.

All three models, given by Equations (4), (6) and (7), can be evaluated from the results of the above analysis, and it now remains to assess the expected accuracy of the pressures predicted by them. Considering Equation (7) where an estimate of its error

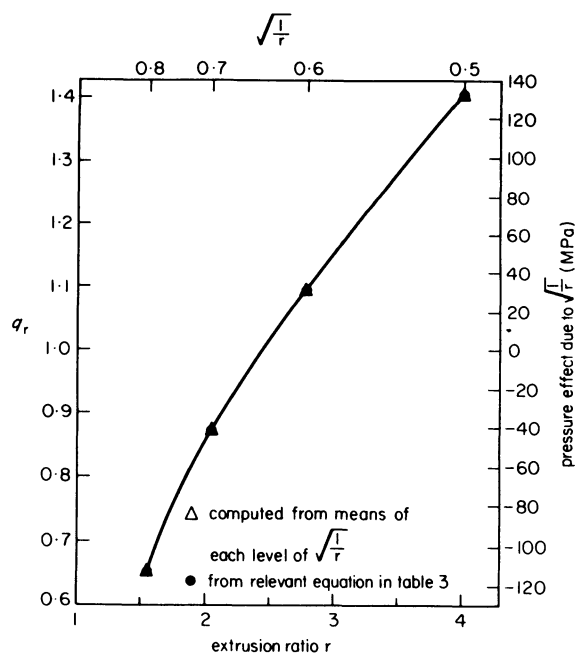


Figure 3. Pressure effect due to $(1/r)^{1/2}$ and extrusion ratio correction factor q_r .

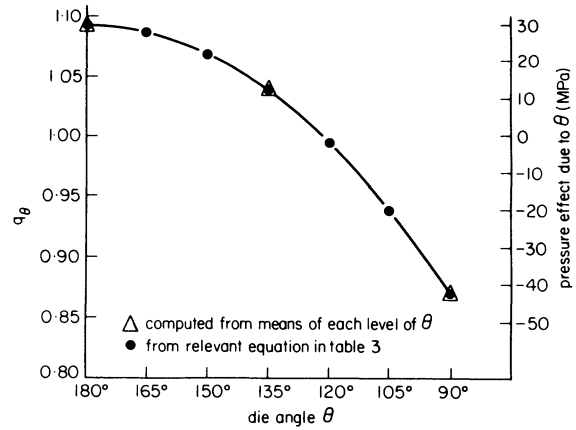


Figure 4. Pressure effect due to θ and die angle correction factor q_θ .

variance is given by the pooled second order interaction and residual mean squares in Table 2. It can be shown that the 95% confidence limits to be attached to an observation in Table 1 are ± 15.5 . Hence it would be expected with a probability of 0.95 that a maximum pressure observed during an experiment would lie within ± 15.5 MPa of the value predicted by Equation (7). Assuming that the first order interactions in Table 2 are not significant, as is the case for the model given by Equation (6), and pooled with the above variance the 95 per cent confidence limits become ± 28.6 MPa. Therefore a considerable reduction in accuracy can be expected when the factors are assumed to be independent and pressures are predicted with Equation (6).

To estimate the expected accuracy of the remaining model, pressures were calculated from the correction factor curves in figures 3, 4 and 5, and the differences between these values and the corresponding experimental results in Table 1 were tabulated for each of the two replications. The variances for the two sets of differences were calculated and pooled to give a variance of 67.3 with 70 degrees of freedom. An F ratio test was made on the above variance and the pooled error variance of the Equation (7) model. The results of this test indicated that the difference between these two variances was not statistically significant. To complete the analysis a t test on the mean of the differences calculated above showed that there was no significant bias in the

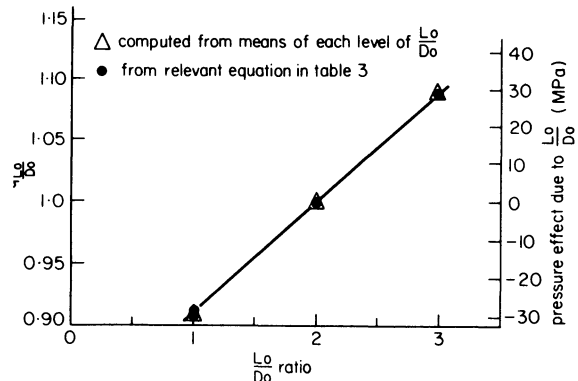


Figure 5. Pressure effect due to L_0/D_0 and the L_0/D_0 correction factor q_{L_0/D_0} .

results obtained from Equation (4). Hence this relatively simple model can be expected to predict maximum extrusion pressures to a similar degree of accuracy as the more complex model described by Equation (7). The 95 per cent confidence limits for Equation (4) was calculated to be ± 16.4 MPa. Naturally this accuracy is dependent on the care taken in selecting values from the correction factor curves.

CONCLUDING REMARKS

Annealed aluminium alloy 6063 was successfully forward cold extruded without the occurrence of defects such as *pick up* and *chevrons*. An analysis of the data from the experiments showed that the main effects and first order interactions of the reduction ratio, die angle and billet length to diameter ratio had significant effects on the maximum extrusion pressure.

The relatively simple model applied by NEL to predicting the maximum extrusion pressures for cold extruding steel was found to have a similar accuracy as a model that was derived using orthogonal polynomials. The equation for this latter model included terms for all the statistically significant effects whereas the model of NEL used the product of correction factors for the three main effects and did not include terms for significant interactions. The accuracy obtained from such a simple model for this particular set of experiments is very encouraging for developing a simple accurate method of predicting

maximum extrusion pressures. However further experiments need to be conducted over larger ranges of the variables and different materials so that a better estimate can be obtained of its performance.

REFERENCES

1. P. J. M. Boes and H. P. Pouw. A practical calculation method for extrusion pressures. *Sheet metal Industries* (May, 1966) P377-390.
2. J. McKenzie. Design charts to determine press loads for the cold extrusion of steel. *NEL Report No. 118* (1963) East Kilbride, Glasgow: National Engineering Laboratory.
3. H. D. Feldmann. *Cold Forging of Steel*. Hutchinson & Co, London 1961.
4. E. S. Pearson and H. O. Hartley. *Biometrika Tables for Statisticians* (1966) 1, Cambridge University Press.
5. O. L. Davies. *Design and Analysis of Industrial Experiments*. Hafner Publishing Company, New York, 1971.
6. J. McKenzie. Factors in the design of cold extrusion tooling. *NEL Report No. 117* (1963) East Kilbride, Glasgow: National Engineering Laboratory.
7. P. J. White. *Preliminary Investigation into the Cold Extrusion of Aluminium Rod*. B.E. Thesis, School of Mechanical and Industrial Engineering, The University of New South Wales 1971.
8. H. Ll. D. Pugh, M. T. Watkins and G. S. Lawson. Cold extrusion of aluminium and aluminium alloys. *Metal Treatment and Drop Forging* (1964) 31(3), 82-7, 31(4), 156-64.

RIGID-PLASTIC ANALYSIS OF COLD HEADING BY THE MATRIX METHOD

by

S. N. SHAH* and SHIRO KOBAYASHI*

SUMMARY

An effective treatment of rigid zones was introduced into the computational procedure of the matrix method. This improvement made it possible to solve, by the matrix method, problems where rigid zones are involved and their boundaries are not known. With this improved version of the method, the cold-heading process was analysed in detail. The computed results show excellent agreement with those by experiment in terms of load-displacement relationship, bulge contours, flow lines, and rigid-zone boundaries.

INTRODUCTION

Plastic deformation of metal wires or rods, without the use of heat, has found its greatest field of application in the form of heads, as in the manufacture of screws, bolts, and rivets. Although cold heading is generally identified with the production of standard fasteners, it is also adaptable to the manufacture of an endless variety of special fasteners and small machine elements like gears, cams, locking arms for automatic transmissions, etc., especially when used in conjunction with second-operation machines. Cold heading can also occur at the centre or at several points of the workpiece rather than at its ends. The heading process is described in length in the literature¹⁻⁵, including its advantages and disadvantages.

Although cold heading is a very common metal-working operation, only a few theoretical studies have been undertaken to reveal the details of the process⁶⁻⁸. The rules used in the industry are rules of thumb based on experience. However, these rules do not provide proper guidelines when rapidly work-hardening metals or alloys are worked on, or when high-temperature bolts are manufactured. Also, a wide range of failures has been reported in the literature⁹⁻¹¹. The high quality requirement for the feed wire, and the limits experienced for the dimensions to avoid failure, are the major problems facing the manufacturers. To provide guidelines for solutions to some of these problems an attempt is made to analyse the deformation characteristics of the heading process.

The matrix method, introduced some time ago by Lee and Kobayashi¹² for a rigid-plastic material, has been successful in studies of metal-working problems where the entire body is deforming plastically¹³⁻¹⁵. However, in many metal-working processes, the entire body is not deforming plastically, and the rigid-plastic boundary is unknown. To solve these problems a technique, where one would determine this boundary and then apply the method after eliminating rigid zones, was introduced recently¹⁶. This technique, along with a revised linearization process, was incorporated in the matrix method, and the improved version was applied in the analysis of compression of tall cylinders. In this problem the rigid zone appears at an early stage and vanishes after a small percentage reduction in height. To expand the applicability of the method, it is necessary to test the method for the problems where the rigid zone exists throughout the process. The analysis of heading was undertaken for this purpose, also. Hence, dynamic considerations are not included in the present analysis. The calculations were carried out in terms of the velocity distribution, grid distortion, stresses and strains at intervals of 1 per cent up to 60 per cent reduction in height. The results were compared with experimental observations.

MATRIX METHOD

Consider a body, V , whose surface, S , consists of S_T and S_U , under boundary conditions such that the entire body is deforming plastically. The body is

* Department of Mechanical Engineering, University of California

composed of a rigid-plastic material which obeys the von Mises yield criterion and its associated flow rule. Body forces are assumed to be absent in the region, V . Then, according to a variational principle^{17,18} for the actual solution, the functional (Equation (1)) becomes stationary with respect to the Lagrange multiplier λ and the velocity fields which satisfy the velocity boundary conditions on S_U , but not necessarily the incompressibility condition:

$$\Phi = \int_V \bar{\sigma} \dot{\epsilon} dV + \int_V \lambda C^T \dot{\epsilon} dV - \int_{S_T} T^T U dS \quad (1)$$

where $\bar{\sigma}$ is the effective stress; $\dot{\epsilon}$, the effective strain-rate; C is the proper matrix notation of the Kronecker delta, such that $C^T \dot{\epsilon}$ gives the first invariant of the strain-rate tensor; T is the traction vector specified on the boundary S_T ; and U is the velocity vector.

The effective strain-rate and the effective stress are defined respectively as

$$\dot{\epsilon} = \left(\frac{2}{3} \dot{\epsilon}^T D \dot{\epsilon} \right)^{1/2} \quad (2)$$

and

$$\bar{\sigma} = \left(\frac{3}{2} \sigma'^T D^{-1} \sigma' \right)^{1/2} \quad (3)$$

and the deviatorial stress vector, σ' , is related to the strain-rate vector, $\dot{\epsilon}$ according to the Levy-Mises equations, by

$$\sigma' = \frac{2}{3} \frac{\bar{\sigma}}{\dot{\epsilon}} D \dot{\epsilon} \quad (4)$$

The superscripts T and -1 , in these equations, denote the transpose and the inverse, respectively. It has been shown that λ is equal to the hydrostatic component of the stress vector, σ , for the actual solution¹⁹.

For discretization of the problem, the body, V , is divided into M elements, interconnected at N nodal points. The functional, Φ , is approximated by a function, ϕ , on the elemental level by replacing U with a kinematically complete distribution, given by

$$U = Gu \quad (5)$$

where G is the interpolation function and u is the vector whose components are the velocities at nodal points associated with the element. The strain-rate vector is derivable in the form

$$\dot{\epsilon} = Bu \quad (6)$$

Substitution of these values into the functional (Equation (1)) approximates Φ by ϕ . Applying the stationary-value requirement on ϕ , the equations for the solution of u , which are essentially nonlinear in nature, are obtained. For solution by the iterative method, these equations can be linearized by considering a small perturbation, Δu , in the velocity vector, u , such that

$$u_n = u_{n-1} + \Delta u_n \quad (7)$$

The relationship thus obtained on the elemental level can be assembled to form the global perturbation equation for the n -th iteration.

$$S_{n-1} \left\{ \frac{\Delta u}{\lambda} \right\}_n = R_{n-1} \quad (8)$$

This system of linear equations is solved for Δu and λ , and the iteration process may be terminated when the desired accuracy is obtained. From the values of u and λ , the strain-rate distribution and the stress distribution can be computed. Derivations of the perturbation equation and the formulation of matrices are given in the appendix.

RIGID ZONE ELIMINATION TECHNIQUE

The matrix method described in the foregoing section applies only if the entire body is deforming plastically, and no rigid zone or unloading exists in the deformation process. In practical problems, however, situations do arise where the rigid zone, as well as rigid unloading, are involved. If these regions of no deformation are contained within the control volume, V , the extremum principles do not apply to the problem of obtaining internal distributions. This difficulty can be handled by a technique introduced previously¹⁶.

The technique involves the determination of the boundary of a nearly rigid zone, and the elimination of this rigid zone from the computation. A nearly rigid zone can be characterized by its very low value of effective strain-rate in comparison with the deforming body. During the iteration process for an incremental solution over the entire body, the effective strain-rate in the possible rigid region approaches zero as the solution converges. As shown in the appendix, the effective strain-rate appears in the denominator of the stiffness matrix, S . Hence, the diagonal component of the normalized stiffness matrix will tend to become infinity if the nodal point associated with this component is contained in the rigid zone. Thus, when the solution converges sufficiently close to the actual solution (say, $\|\Delta u\|/\|u\| \leq 0.01$; see the next section for the measure of solution convergence), the nodal points, for which the diagonal component of the normalized stiffness matrix reaches a value greater than a critical value (say, 100), are considered to be in the rigid zone. At this stage of convergence, these nodal point values lying inside the rigid zone are eliminated from the perturbation relationship. Assuming that the rigid boundary thus obtained is correct, the iteration is continued for the solution in the plastically deforming region until a desired convergence is achieved. In the next step of deformation, the nodal point values, eliminated previously, are brought back into the computation, and the same iteration procedure is repeated. It should be noted, of course, that the converged solution gives the stress distribution only in the plastically deforming region.

COMPUTATIONAL CONDITIONS AND PROCEDURE

The analysis of cold heading of circular, solid cylinders was carried out under static conditions. This implies that neither the inertia effects on the forces, nor the strain-rate effects on the material properties were considered in the computation.

The calculations were performed for cylinders having an initial head height to diameter ratio of 1.5 with the static stress-strain relationship expressed by

$$\frac{\bar{\sigma}}{Y_0} = \left(1 + \frac{\bar{\epsilon}}{0.012}\right)^{0.1085}$$

with initial yield stress $Y_0 = 85\,000$ psi. This expression fits the stress-strain curve for SAE1040 steel in as-received condition at room temperature. The friction condition, at the die-workpiece interface, was specified in terms of a friction factor, f , defined by

$$\frac{1}{A} \int \tau \, dA = fY_0 \quad 0 \leq f \leq 1/\sqrt{3}$$

where τ is the frictional stress and A is the current contact area. It was assumed that f remains constant during the heading process and that the distribution of τ is linear in the radial direction (zero at the centre of the contact surface) as long as τ is less than the shear strength $Y_0/\sqrt{3}$, and remains constant after it reaches the shear strength. In the present calculations, $f = 0.25$ was taken at both upper and lower dies. This value corresponds approximately to the dry friction condition.

The problem was analysed in a step-by-step manner by treating it quasi-linearly during each incremental deformation. The cylinder was divided into 184 quadrilateral ring elements, and a 1 per cent reduction in height of the head was selected as the size of an increment. The initial guess for the velocity distribution for the first increment was taken from the solution for the upsetting of cylinders under similar conditions. The step-by-step procedure was continued using the velocity distribution corresponding to the previous step as an initial guess for the subsequent solution. The convergence was measured by the quantity, $\|\Delta\mathbf{u}\|/\|\mathbf{u}\|$, where the Euclidean vector norm is defined by

$$\|\mathbf{u}\| = \left(\sum_{i=1}^N u_i^2\right)^{1/2}$$

The solution with $\|\Delta\mathbf{u}\|/\|\mathbf{u}\| < 0.00001$ was considered acceptable at every step. The calculations were continued up to 60 per cent reduction in height.

This problem was solved using the computer facilities at Lawrence Berkeley Laboratories, Berkeley, California. The storage capacity required for this problem was about 100 000 octal words, without any use of tapes or other external peripheral devices. The number of iterations required for desired accuracy was about 30 for the first three steps, and about 6 or

7 afterwards. Total time required for the calculations was about 11 min or about 1 sec per iteration on the CDC7600 computer†.

EXPERIMENTAL PROCEDURE

To examine the correlation between the computed results and the experimental observations, a static heading operation of circular, solid cylinders was carried out. The lower die was provided with a hole and a plug of variable length, to fit the hole, in order to be able to vary the shank length. The specimens were prepared from SAE1040 steel in as-received condition, and the original geometry of the specimen was selected as 0.5 in diameter and 1.25 in length of which 0.5 in was the shank length. Experiments were carried out under the dry friction condition which was ensured by cleaning the specimen and the die surfaces with acetone. The Tinius Olsen universal testing machine with a capacity of 160 000 lb was used at a constant speed of 0.05 in/min. The load-displacement curves were recorded and different specimens were cold headed to approximately 15 per cent, 30 per cent, 45 per cent, and 60 per cent reduction in height of the head. The bulge profiles and the radii of the contact surfaces were measured at these stages. Then the specimens were sectioned in half along a diametric plane. The sectioned surfaces were ground and etched with a hot solution of 50 per cent hydrochloric acid in water for about 30 min. These surfaces were then washed with water to reveal the flow lines.

RESULTS AND DISCUSSION

The results obtained from the theoretical analysis are presented in terms of load-displacement relationship, strain distributions, stress distributions, and grid pattern distortions. These results were compared with the experiments and the agreement between the two was found to be excellent.

The load-displacement curve is shown in figure 1. The figure also shows the variation of average heading stress as a function of the reduction in height. The theoretical analysis reveals that the average pressure increases sharply during the early phases of deformation, but tends to level off as the reduction increases. The load, on the other hand, increases steadily until the beginning of folding, after which it starts increasing rather sharply due to the increase in contact area. These results are almost identical to those obtained for compression of the same material under the same friction conditions¹⁶. The experimental values are also shown in the figure and the agreement with theory is excellent.

The matrix method makes it possible to obtain not only the average forming pressure, but also the detailed pressure distribution at the die-workpiece interface. This distribution at the top die is shown in figure 2 at various stages of deformation. The significant influence of the height to diameter ratio, on the distribution of the pressure, is revealed by the

† The CDC7600 is about 5 times faster than the CDC6600, and about 20 times faster than the CDC6400.

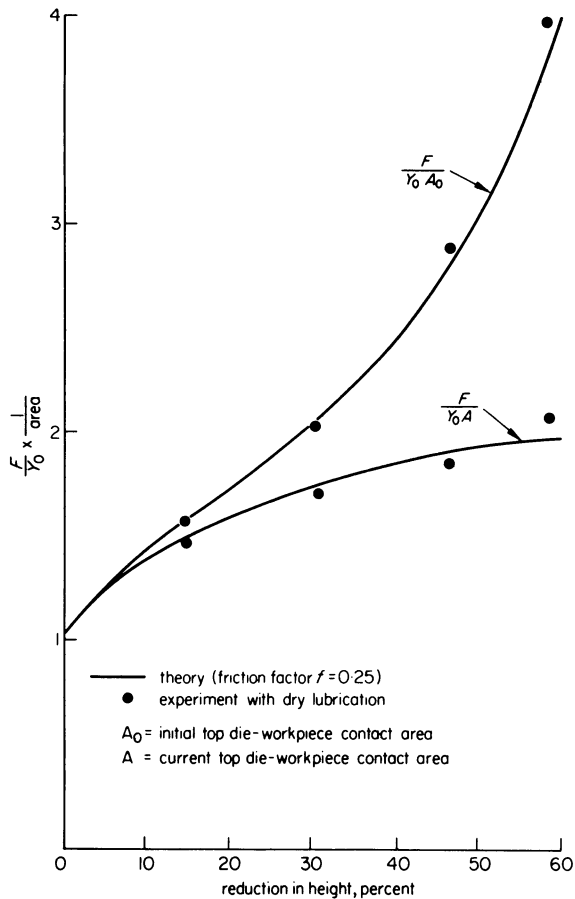


Figure 1. Load-displacement relationship.

analysis. The pressure is minimum at the centre and increases radially outwards, when the height to average diameter ratio is greater than 0.6, but this trend is reversed when this ratio is less than 0.6. These results are qualitatively similar to the experimental results for upsetting by Takahashi²⁰ and by Backofen and Van Rooyen²¹. It is also worth noting that the stress at the edge of the contact surface is always high.

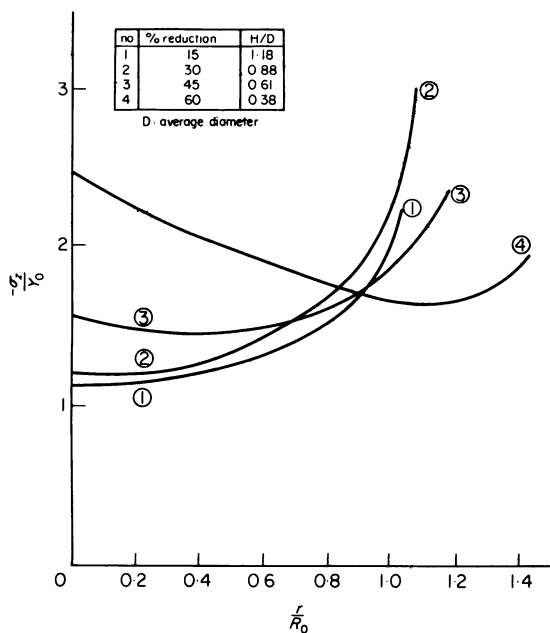


Figure 2. Heading pressure distributions along the top die during the process.

Folding is a common phenomenon in upsetting, and is characterized by a portion of the original free surface coming into contact with the die during deformation. The amount of folding is examined by the difference between the radius of the current contact surface and the current radius of the initial contact surface. In figure 3 these details are plotted for top and bottom dies. The differences between curves A₁ and A₂ and B₁ and B₂ indicate the amount of folding at the top die and the bottom die, respectively.

It is of interest to note that, after folding starts, a point on the initial contact surface on the top die still has outward velocity, and this radius continues to increase with increasing reduction. On the contrary, on the bottom die, a point on the initial contact surface almost stops moving once folding has

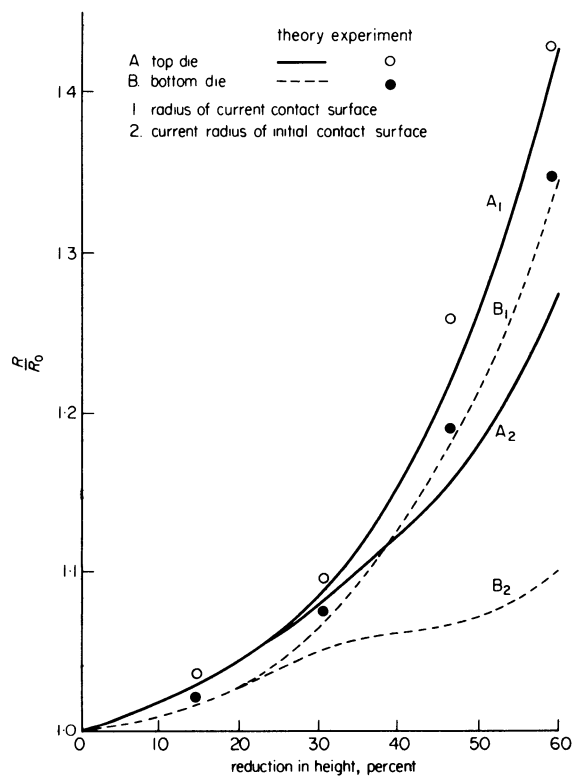


Figure 3. Folding curves.

occurred. Hence, although the curves for the radius of the current contact surface are parallel for the top and bottom dies, there is a greater amount of folding on the bottom die. Experimental results showing the radius of the current contact surface at the top and bottom dies, are also given in the figure and the agreement is once again very good.

The deformation pattern during the heading operation can be observed in detail by examining the effective strain distributions which are plotted on the right sides in figure 4 at various stages of deformation. An important feature of these distributions is that the pattern does not seem to change much as the reduction in the height increases. This is in contrast to the results of compression where the deformation pattern changes considerably¹⁶. As seen in the figure, there is a strain concentration at the edge of the contact surface on the top die, and on

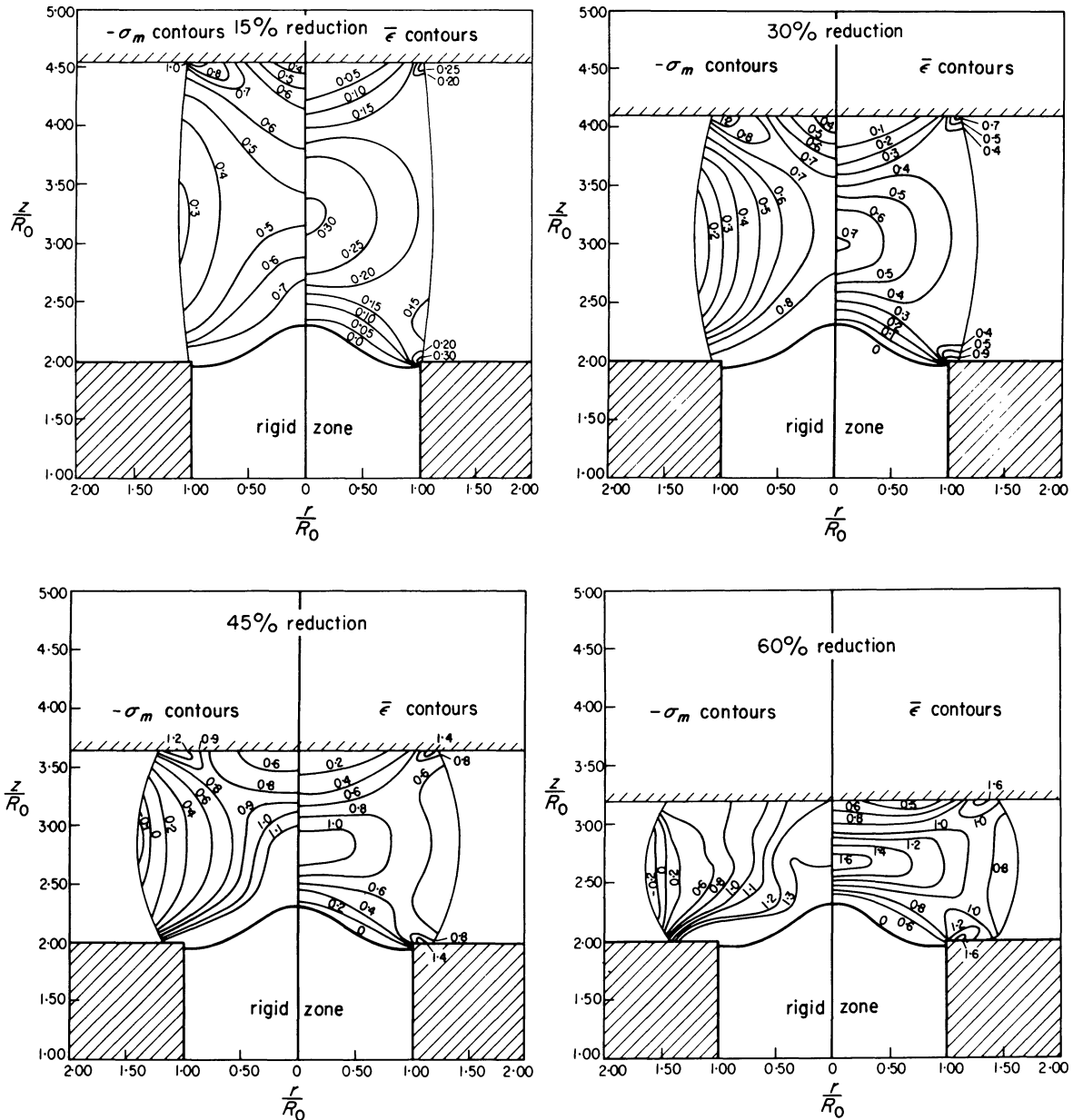


Figure 4. Effective strain ($\bar{\epsilon}$) contours and hydrostatic pressure ($-\sigma_m/Y_0$) contours at various reductions in height.

the contact surface at the bottom die. Otherwise, the maximum strain is near the axis, and almost at mid-length of the head. There is great strain variation along the axis, whereas along the free surface the strain value is almost the same. At the die surface the strain is always small compared to that in the rest of the deforming body. There is a large gradient of strain at the lower part, near the rigid zone, which is indicative of an intense shear between the rigid and the deforming zones. It is noteworthy that a rigid zone exists in the shank, and in the bottom of the head, throughout the deformation process, as indicated by the zero strain contour. The size of this rigid zone remains the same throughout the deformation process.

The introduction of the Lagrange multiplier to the matrix method permits the evaluation of the mean or hydrostatic stress component, σ_m . Since the hydrostatic pressure seems to play an important role in ductile fracture²²⁻²⁴, it is instructive to examine

the distribution of σ_m . These distributions are plotted on the left sides of figure 4 at different reductions in height. Although the trend of the effective strain distribution remains almost the same, the hydrostatic stress distribution changes a great deal with reduction in height. At the die surface the pressure is minimum in the centre and increases radially outwards for up to 45 per cent reduction in height, but this pattern reverses at higher reductions. This trend is similar to that observed for σ_z distribution in figure 2. The general pattern of σ_m distribution also changes considerably with the reduction in height, and, at high reductions, constant hydrostatic stress contours become almost vertical for most of the deforming region. A very important observation, from the fracture point of view, is that with the increase in reduction in height, the hydrostatic pressure near the axis of the specimen increases, whereas it decreases near the free surface, and indeed becomes tensile at higher reductions. This

suggests a reason for dominant surface cracking rather than internal cracks in heading operations.

A relevant comparison between the theory and experiments can be made in terms of the grid distortion pattern. This comparison is significant since it relates to the details of the internal deformation. As described earlier, the flow patterns were obtained experimentally by the etching technique and are shown in figure 5 at approximately 31 per cent, 47 per cent, and 63 per cent reductions in height. The figure also shows the grid patterns obtained theoretically at 31 per cent, 47 per cent,

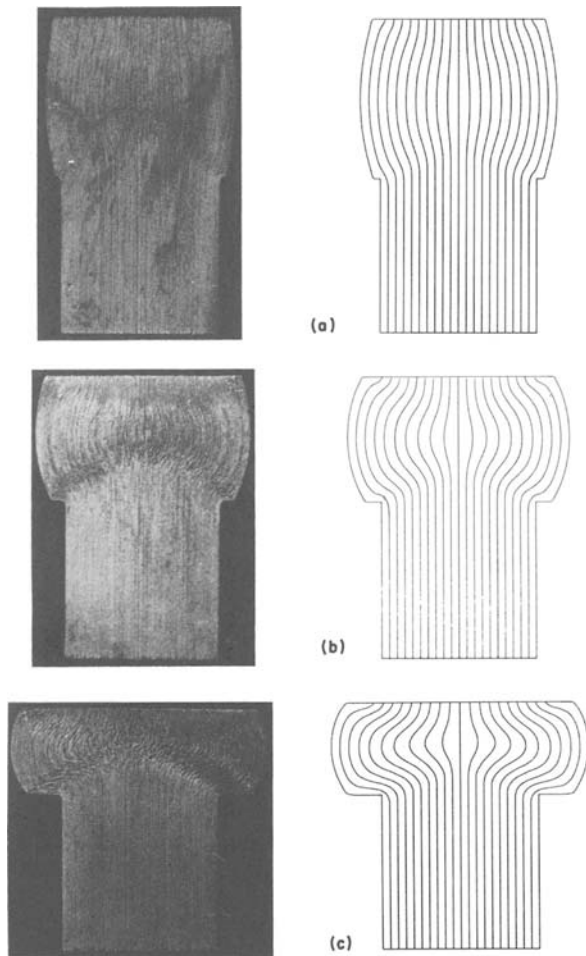


Figure 5. Experimental and theoretical flow lines: (a) 31 per cent reduction; (b) 47 per cent reduction; (c) 63 per cent reduction, experiment, and 60 per cent, theory.

and 60 per cent reductions. It can be seen that, even at 60 per cent reduction, the vertical line distortion predicted by the theory is almost identical to the vertical flow lines observed experimentally, except for a small region adjacent to the dies. Thus, not only the bulge profile contours of the free surface, but also the deformation of the internal region is predicted conclusively. The experiments also reveal the rigid-plastic boundary in the lower part of the head and an intense shear along this boundary. This rigid-plastic boundary matches very well with the computed zero strain contours in figure 4 and the intense shear is also seen in the theoretical results.

CONCLUSION

It was demonstrated clearly that the applicability of the method has been expanded to the problems where exact rigid zone boundaries are not known *a priori*. Agreements between the computed results and the experimental observations are excellent, not only over-all, but also in detail. In particular, an excellent agreement of flow patterns is impressive. When the results of the heading process are compared with those of a corresponding simple upsetting, it can be seen that the load-displacement relations are identical and the folding curves at the top die are almost the same. It appears that the deformation near the top die is independent of the condition at the bottom die. Also, it is of interest to note that the extent of the rigid zone in the shank remains the same throughout the process. The excellent agreement of flow lines suggests that the calculation of stress and strains is accurate. It is then possible to predict the limits on the dimensions to avoid the failure in the heading process if the fracture criterion for the material is known.

Although the dynamic considerations were not given in the present investigation, it is not difficult to include the dynamic effects in the formulations of the matrix method. The work along this line is now under way.

ACKNOWLEDGMENTS

The authors wish to thank the National Science Foundation for its grant GH-37120 under which the present investigation was possible. They also wish to thank Mr. Shokrai for his assistance in performing the experiments.

APPENDIX

A. Derivation of perturbation relationship

The approximation of the functional (equation (1)) in the text is given by

$$\phi^m = \int_{V^m} \bar{\sigma} \left(\frac{2}{3} u^T K u \right)^{1/2} dV + \lambda u^T Q - u^T F \quad (A1)$$

where

$$K = B^T D B; \quad Q = \int_{V^m} B^T C dV$$

and

$$F = \int_{S_T^m} G^T T dS;$$

and λ is assumed to be uniform over each element.

Applying the stationary condition to the function ϕ

$$\sum_m^M \frac{\partial \phi^m}{\partial u_k} = 0, \quad k = 1, \dots, N \quad (A2)$$

and

$$\sum_m^M \frac{\partial \phi^m}{\partial \lambda^i} = 0, \quad i = 1, \dots, M$$

where u_k is the velocity vector at the k -th nodal point and λ^i is the Lagrange multiplier of the i -th element.

Using the linearization technique described in the text, and neglecting second and higher order terms of Δu_n , we obtain

$$\frac{\partial \phi^m}{\partial u^m} = \bar{\sigma} P_{n-1} \Delta u_n + \lambda_n Q - F + \bar{\sigma} H_{n-1} \quad (A3)$$

$$\frac{\partial \phi^m}{\partial \lambda^m} = Q^T \Delta u_n + Q^T u_{n-1} \quad (A4)$$

In equations (A3) and (A4)

$$P_{n-1} = \frac{2}{3} \int_{\sqrt{m}} \frac{1}{\dot{\epsilon}_{n-1}} \left[K - \frac{2/3}{\dot{\epsilon}_{n-1}^2} b_{n-1} b_{n-1}^T \right] dV$$

$$H_{n-1} = \frac{2}{3} \int_{\sqrt{m}} \frac{1}{\dot{\epsilon}_{n-1}} b_{n-1} dV$$

$$b_{n-1} = K u_{n-1} \quad (A5)$$

and

$$\dot{\epsilon}_{n-1} = \left[\frac{2}{3} u_{n-1}^T K u_{n-1} \right]^{1/2}$$

Equations (A3) and (A4) can be symbolically represented as

$$\left\{ \frac{\partial \phi / \partial u}{\partial \phi / \partial \lambda} \right\}^m = S_{n-1}^m \left\{ \frac{\Delta u}{\lambda} \right\}_n - R_{n-1}^m \quad (A6)$$

for the m -th element. Using Equation (A2) and carrying out the assembly operation

$$S_{n-1} \left\{ \frac{\Delta u}{\lambda} \right\}_n = R_{n-1} \quad (A7)$$

where

$$S = \sum_m S^m; \quad R = \sum_m R^m; \quad \Delta u = \sum_m \Delta u^m; \quad \lambda = \sum_m \lambda^m$$

with Σ being the assembly operator.

B. Matrices for axisymmetric deformation

The stress and strain-rate vector components in the cylindrical co-ordinate system are listed as

$$\sigma = \begin{Bmatrix} \sigma'_r \\ \sigma'_z \\ \sigma'_\theta \\ \tau_{rz} \end{Bmatrix} \quad \text{and} \quad \dot{\epsilon} = \begin{Bmatrix} \dot{\epsilon}_r \\ \dot{\epsilon}_z \\ \dot{\epsilon}_\theta \\ \dot{\gamma}_{rz} \end{Bmatrix} \quad (B1)$$

Then, the flow matrix D in Equation (2) and the vector C of the Kronecker delta in equation (1) become

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \quad \text{and} \quad C = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \quad (B2)$$

In a quadrilateral element with four nodal points, the velocity components u and v , in the r and z direction, respectively, are assumed to be

$$u(s, t) = \sum_{i=1}^4 q_i u_i \quad \text{and} \quad v(s, t) = \sum_{i=1}^4 q_i v_i \quad (B3)$$

where (s, t) represents a natural coordinate system and the transformation from the (s, t) coordinates into the (r, z) co-ordinates are given by

$$r(s, t) = \sum_{i=1}^4 q_i r_i \quad \text{and} \quad z(s, t) = \sum_{i=1}^4 q_i z_i \quad (B4)$$

In Equations (B3) and (B4), u_i and v_i are the velocity components at the nodal points of the element whose (r, z) coordinates are r_i and z_i , and

$$q_i = \frac{(1 + s_i s)(1 + t_i t)}{4} \quad (B5)$$

where s_i and t_i are the (s, t) coordinates of node i given by $(-1, -1)$, $(1, -1)$, $(1, 1)$, and $(-1, 1)$ for nodes 1, 2, 3, and 4, respectively.

When the listings of the vectors in Equation (5) are given by

$$\begin{Bmatrix} u \\ v \end{Bmatrix} = U \simeq G \begin{Bmatrix} u_1 \\ v_1 \\ \vdots \\ u_4 \\ v_4 \end{Bmatrix}$$

the interpolation function G becomes

$$G = \begin{bmatrix} q_1 & 0 & q_2 & 0 & q_3 & 0 & q_4 & 0 \\ 0 & q_1 & 0 & q_2 & 0 & q_3 & 0 & q_4 \end{bmatrix} \quad (B6)$$

Since the strain-rate vector is defined by

$$\begin{Bmatrix} \dot{\epsilon}_r \\ \dot{\epsilon}_z \\ \dot{\epsilon}_\theta \\ \dot{\gamma}_{rz} \end{Bmatrix} = \begin{bmatrix} \frac{\partial}{\partial r} & 0 \\ 0 & \frac{\partial}{\partial z} \\ \frac{1}{r} & 0 \\ \frac{\partial}{\partial z} & \frac{\partial}{\partial r} \end{bmatrix} \begin{Bmatrix} u \\ v \end{Bmatrix}$$

The matrix B in equation (6) is obtained as

$$B = \begin{bmatrix} Y_1 & 0 & Y_2 & 0 & Y_3 & 0 & Y_4 & 0 \\ 0 & X_1 & 0 & X_2 & 0 & X_3 & 0 & X_4 \\ G_1 & 0 & G_2 & 0 & G_3 & 0 & G_4 & 0 \\ X_1 & Y_1 & X_2 & Y_2 & X_3 & Y_3 & X_4 & Y_4 \end{bmatrix} \quad (B7)$$

where

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} = \frac{1}{8|J|} \begin{pmatrix} -r_{24} + r_{34}s + r_{23}t \\ r_{13} - r_{34}s - r_{14}t \\ r_{24} - r_{12}s + r_{14}t \\ -r_{13} + r_{12}s - r_{23}t \end{pmatrix}$$

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \frac{1}{8|J|} \begin{pmatrix} z_{24} - z_{34}s - z_{23}t \\ -z_{13} + z_{34}s + z_{14}t \\ -z_{24} + z_{12}s - z_{14}t \\ z_{13} - z_{12}s + z_{23}t \end{pmatrix}$$

and $G_i = \frac{q_i}{r}$, $i = 1, \dots, 4$.

In Equation (B7)

$$r_{ij} = r_i - r_j \quad z_{ij} = z_i - z_j$$

and $|J|$ is the Jacobian of the transformation defined by

$$|J| = \frac{\partial r}{\partial s} \frac{\partial z}{\partial t} - \frac{\partial r}{\partial t} \frac{\partial z}{\partial s}$$

REFERENCES

1. T. B. Smith. Design for cold heading. *Wire and Wire Products* (1960) **35**, p. 1113.
2. Precision bolts and studs. *Automobile Engineer* (1945) **35**, p. 313.
3. H. E. Linsley. Fundamentals of cold heading. *American Machinist* (1948) **92**, p. 99.
4. C. S. Ricker. Cold heading cuts manufacturing costs. *American Machinist* (1948) **92**, p. 89.
5. R. W. Bolz. Production processes; their influence on design. *Machine Design* (1946) **18**, p. 119.
6. A. G. Lecocq. Stresses in the shank of a bolt during cold heading. *Wire (English version of Draht) Coburg, Germany* (1971) **115**, p. 197.
7. P. F. Thomason. The effect of heat treatment on the ductility in a cold heading process. *Proc. Inst. of Mech. Eng.* (1969-70) **184**, p. 875.
8. F. L. Gill and W. M. Baldwin. Proper wire drawing improves cold heading. *Metal Progress* (1964) **85**, p. 83.
9. L. R. Franks and J. C. Harrigan. Maintenance and improvement of quality cold heading wire. *Wire and Wire Products* (1951) **26**, p. 888.
10. W. E. Hill. Rod and wire quality vs. cold heading quality. *Wire and Wire Products* (1953) **28**, p. 1291.
11. F. W. Nobbs. Some unusual conditions encountered in cold heading products. *Wire and Wire Products* (1946) **21** p. 794.
12. C. H. Lee and S. Kobayashi. New solutions to rigid-plastic deformation problems using a matrix method. *Trans. ASME, J. Engrg. Ind.* (1973) **95**, p. 865.
13. C. H. Lee and S. Kobayashi. Deformation mechanics and workability in upsetting solid circular cylinders. *Proc. North Amer. Metalworking Res. Conf., Hamilton, Canada* (1973) **1**, p. 185.
14. S. H. Lee and S. Kobayashi. Rigid-plastic analysis of bore expanding and flange drawing with anisotropic sheet metals by the matrix method. To be presented at the 15th Intl. MTDR Conference, September 1974.
15. S. Kobayashi, C. H. Lee, and S. N. Shah. Analysis of rigid-plastic deformation problems by the matrix method. *J. Japan Soc. Tech. Plasticity* (1973) **14**, p. 770.
16. S. N. Shah, C. H. Lee, and S. Kobayashi. Compression of tall, circular, solid cylinders between parallel flat dies. To be presented at the Intl. Conf. on Prod. Engrg., Tokyo, August 1974.
17. R. Hill. *The Mathematical Theory of Plasticity*. Oxford University Press (1950).
18. K. Washizu. *Variational Methods in Elasticity and Plasticity*. Pergamon Press (1968).
19. S. Kobayashi, C. H. Lee, and S. I. Oh. Workability theory of materials in deformation processes. *USAF Technical Report AFML-TR-73-192*, 1973.
20. S. Takahashi. The distribution of contact pressures in compressing cylindrical specimens. *J. Japan Soc. Tech. Plasticity* (1965) **6**, p. 271.
21. G. T. Van Rooyen and W. A. Backofen. A study of interface friction in plastic compression. *Int. J. Mech. Sci.* (1960) **1**, p. 1.
22. F. A. McClintock. A criterion for ductile fracture by the growth of holes. *J. of Appl. Mech., Trans. ASME* (1968) **35**, p. 363.
23. J. R. Rice and D. M. Tracey. On the ductile enlargement of voids in triaxial stress fields. *J. of Mech. and Phys. of Solids* (1969) **17**, p. 201.
24. P. W. Bridgman. *Studies in Large Plastic Flow and Fracture*. McGraw-Hill Book Company, (1952).

PRODUCTION OF CRANKSHAFTS BY CLAMP SHEARING

by

TAKEO NAKAGAWA* and KOICHI SHIOMORI†

SUMMARY

A crankshaft can be produced from a bar by stopping the clamp-shearing operation in the middle of the press stroke. This paper deals with the technical development of the production process. In the preliminary shearing test of 15 mm bars of copper and mild steel, the optimum shear-forming conditions and the strength of the connected shear-formed zone were experimentally investigated. This was successfully applied to the actual production of a crankshaft of Al-63S in an electrical business machine. The newly developed clamp-shearing tool for making crankshafts has been working for the last three years and its construction is discussed in detail in this paper.

INTRODUCTION

It was pointed out in a previous paper¹ that crankshafts can be produced easily by modifying the process of clamp-shearing a bar. A considerable number of small crankshafts are used in accounting, sewing and electrical machines; recently we have succeeded in changing the production process of crankshafts for business machines from machining to the simpler and faster clamp-shearing method. This paper deals with the production of crankshafts by the clamp-shearing method and some experimental results obtained during the developmental stage.

CLAMP-SHEARING METHOD

In the conventional shearing-without-clamping device, a sheared bar moves freely in the direction of its axis away from the cutting edges of the shearing tool. If the bar is clamped firmly as illustrated in figure 1(a), so that it cannot move in the direction of its axis, an axial compressive force is induced which is large enough to shear the bar without fracturing the sheared surface. This process has been used for the last five years for the precision shearing of non-ferrous soft metal bar.

In the simultaneous shearing of multi-planes, which is as a simpler application of the clamp-shearing method, figure 1(b), fracture-free sheared surfaces can be obtained without using a special clamping device because the axial movement of each sheared bar is restrained by the others.

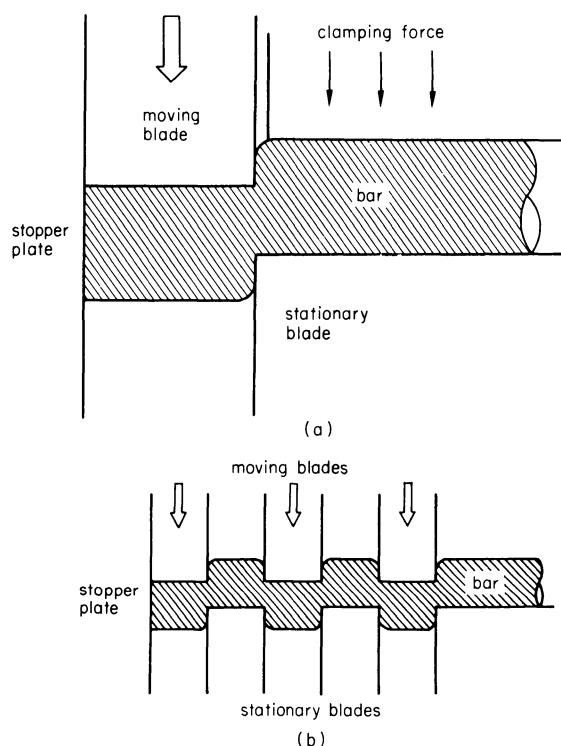


Figure 1. Clamp shearing method. (a) Clamp shearing of bar, (b) simultaneous shearing of multi-planes.

PRELIMINARY TEST BY EXPERIMENTAL TOOL

Design of experimental tool

The optimum tool construction, the effect of forming and the strength of the crankshaft were investigated

* Institute of Industrial Science, University of Tokyo, Roppongi, Minato-ku, Tokyo 106, Japan

† Tokyo Shibaura Electric Co. Ltd., Yanagi-cho, Kawasaki-shi 210, Japan

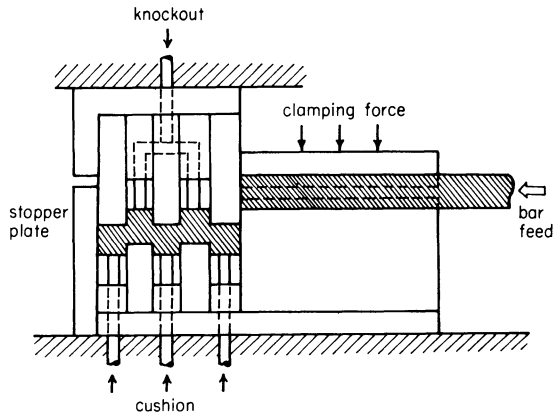


Figure 2. Experimental tool for shear-forming a crankshaft.

using the experimental clamp-shearing tool in figure 2. The design specifications of this tool are as follows:

- (1) The tool can be attached to a conventional mechanical press.
- (2) By a single press stroke the bar is clamp-sheared, the sheared slug is formed to the shape of the crankshaft and the shear-formed crankshaft is ejected.
- (3) The maximum clamping force is 20 tonnes supplied by a hydraulic source. The clamping pressure plate has a semi-circular groove in it.
- (4) The zero clearance needed between the two cutting edges is achieved by using spacers of the same thickness as the cutting blades.
- (5) The bar is circular in cross section with a diameter of 15 mm. Both the stationary and movable blades have semi-circular cutting edges with a diameter 0.1 mm larger than that of the bar.

The experimental shearing and shear-forming procedure is as follows. The bar is fed from the right-hand side as shown in figure 2 and clamped by a hydraulic force after the left-hand end of the bar

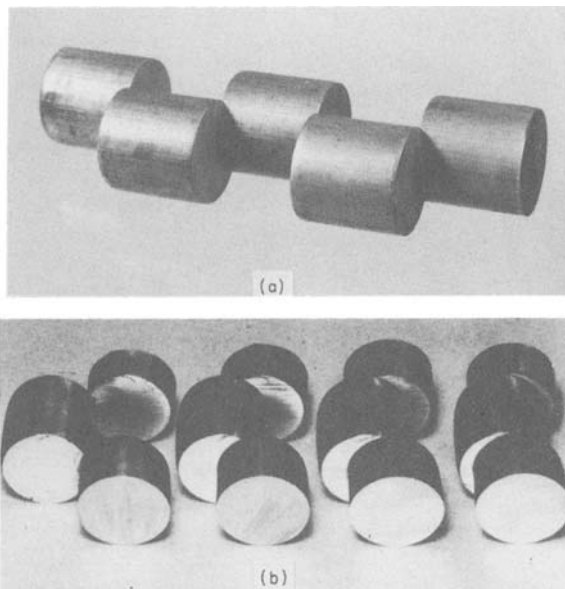
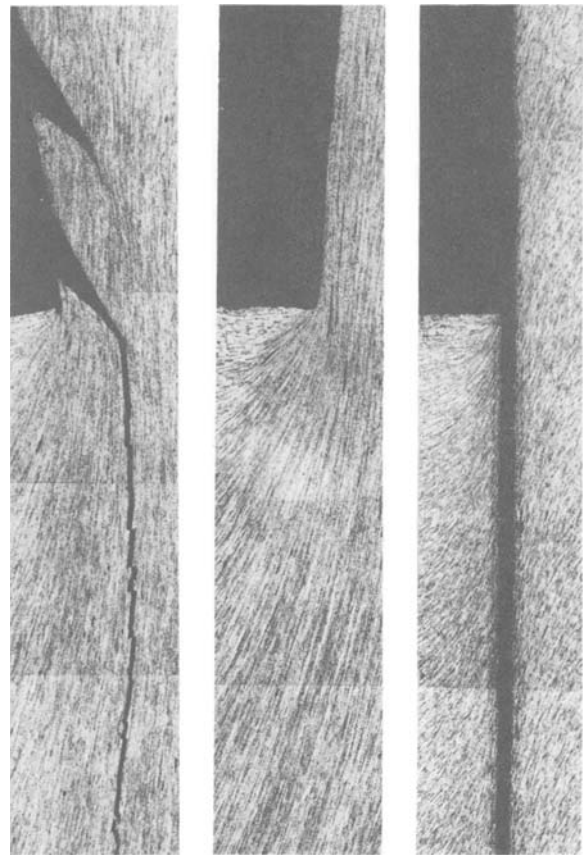


Figure 3. Crankshaft produced by experimental tool. (a) Copper, (b) 0.15 per cent carbon steel.

reaches the stopper plate. As the upper tool moves downwards the bar is sheared to the required length by a pair of shearing blades located on the right-hand side. During this operation, the sheared bar under the forming blades is pressed upwards by cushion pins in order to prevent cantilever bending of the sheared bar. After separation of the bar the shear-forming of the crankshaft begins, and at the lower dead point of the press stroke the shear-forming is completed. The shear-formed crankshaft is ejected from the upper tool by a knockout pin near the upper dead point. The clamping force is released just after the completion of shearing to prevent the sheared surfaces being damaged by rubbing against each other in the return stroke.

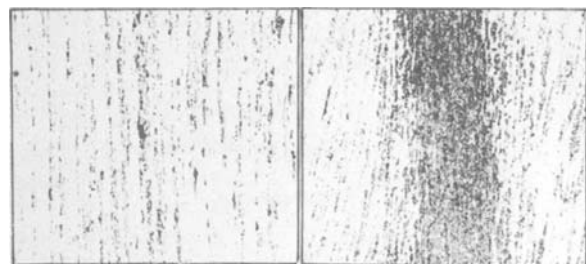


shearing speed
 $V \approx 0.001$ m/s
material testing
machine

$V \approx 0.1$ m/s
crank press

$V \approx 6.5$ m/s
Dynapak

(a)



$V \approx 0.1$ m/s
crank press

0.04 mm

$V \approx 6.5$ m/s
dynapak

(b)

Figure 4. Microstructure of shear-formed zone in 0.15 per cent carbon steel. (a) Flow and fracture patterns in three different speeds, (b) Microstructure in the flow line.

radius, it is possible to achieve 56 per cent diametral expansion before tearing.

An alternative method of obtaining larger diameters is to use a die of a critical radius but with an angle larger than 90° . This die enables the tube to slide over the slanted edge of the die whilst inverting. Figure 8 shows the influence on the maximum

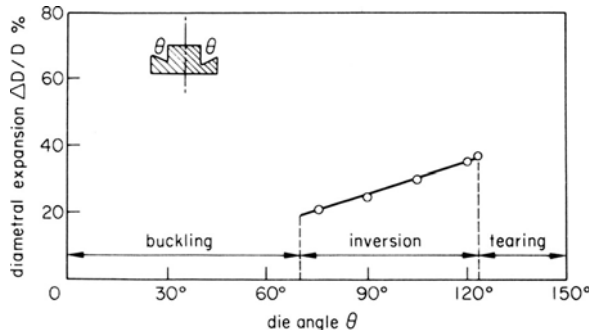


Figure 8. Variation of percentage diametral expansion $\Delta D/D$ with the die angle. The die radius kept at $\frac{3}{8}$ -in.

diametral expansion of the angle θ in a die with a critical radius. It is found that below 70° die angle, the tube cannot flow around the die and buckling develops. Beyond 120° , however, the tube reaches instability, where neither inversion nor tearing is produced and at large values of θ tearing dominates.

It is interesting to note that by cutting the curly collar from the inverted tube and then repeating the inversion a very steady inversion is produced. Figure 9 shows a 2.0-in o.d., 0.064-in wall thickness,

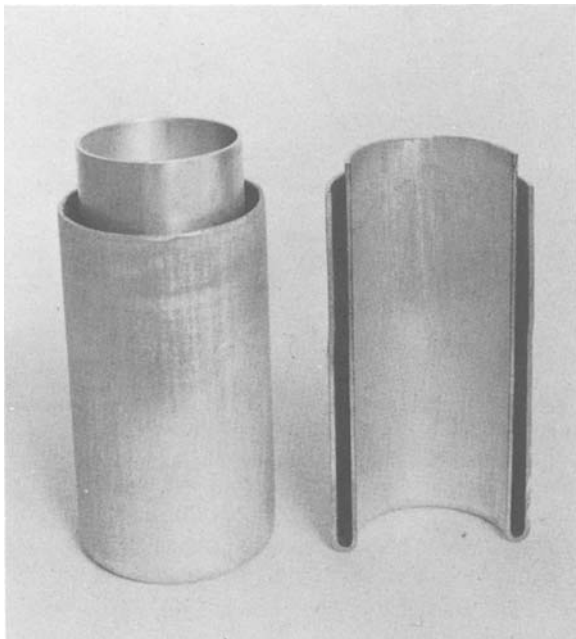


Figure 9. Half-hard aluminium tube 2.0-in o.d., 0.064-in wall thickness steadily inverted after removal of the collar.

aluminium tube with its collar sawn off. The resulting tube is reasonably uniform both in diameter and wall thickness. The load-shortening curve for the collarless tube indicated a constant load throughout. When the outer tube was cut and again reinverted, the load-shortening curve was similar to that of the first

inversion. However, the steady inversion load was in this case slightly higher. It should be noted that for successful reinversion the once inverted tube must have a uniform diameter throughout its length.

Chemical etching and microscopic examination of the surface of the original, once-inverted and twice-inverted tubes revealed that grains are appreciably reduced in each inversion.

It is worth mentioning here that successful inversion of two coaxial tubes, originally made to slide fit one inside the other, was achieved by Mellor². Figure 10, shows a combination of 2.0-in o.d.,

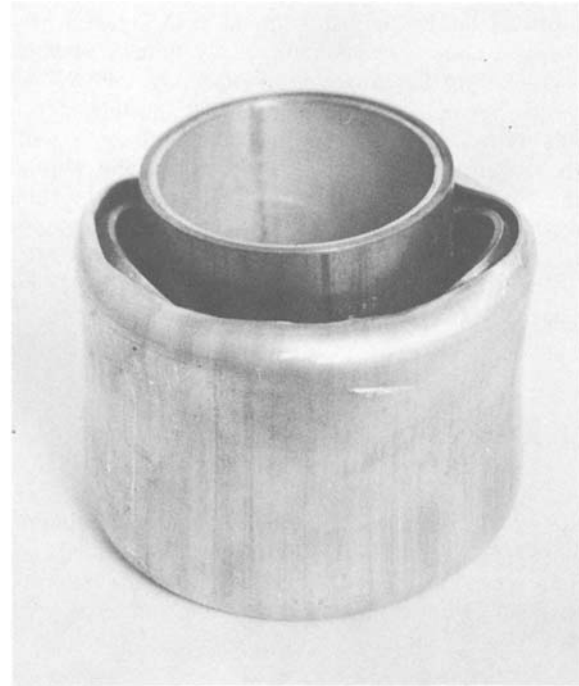


Figure 10. Coaxial aluminium inside brass, tubes inverted with a 90° die of fillet radius = 0.25 in. Mean diameter of tubes = 2.0 in and both tubes have wall thickness = 0.064 in.

0.064-in wall thickness half-hard aluminium inside a 2.0-in o.d., 0.064-in wall thickness 70/30 brass tube inverted with a 90° die with a $\frac{1}{4}$ -in fillet radius. Notwithstanding the difficulties involved in selecting a suitable die radius and material for the tubes, this process is rather promising for achieving tight-fit laminated tubes and thus deserves further investigation.

ANALYSIS

An expression for the steady compressive load necessary to maintain the inversion of a strain hardening tube is arrived at³ by assuming that,

- (1) the tube wall thickness remains constant,
- (2) the deformation process involves bending as well as enlarging the tube diameter, and
- (3) Bauehinger effects, buckling, and tearing are not present.

If the radius of inversion is r and the tube-wall thickness is t , see figure 1, the average strain in

TABLE 1 Strength of shear-formed zone

Material	Without notch (original material) (kg/mm ²)	With notch (kg/mm ²)
Cu-O	23.2	20.0
Cu-H	35.8	32.1
0.15 per cent carbon steel	42.5	33.2

of the test specimen with notch is 10-20 per cent less than that of the original material due to the notch effect but this decrease of strength does not seem to affect it in use.

Dimensional accuracy of the shear-formed crankshaft

The fracture-free sheared surface is smooth and the roughness height is less than 10 μm . The roundness of the shear-formed section is better than that in conventional clamp-shearing due to mutual restraint by the multi-planes. Figure 8(a) shows the distortion

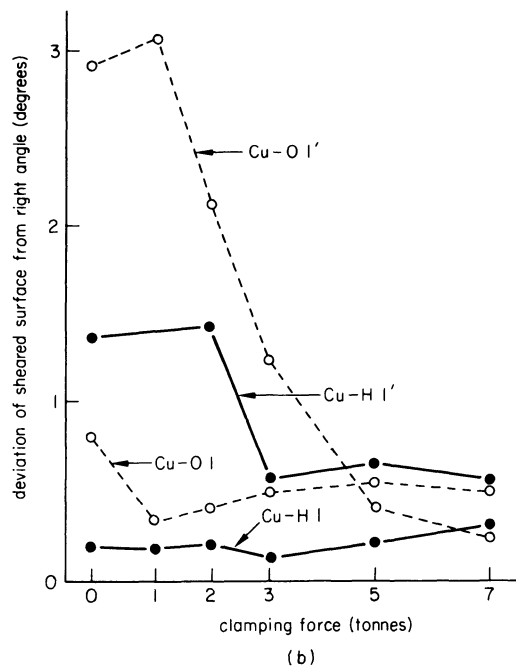
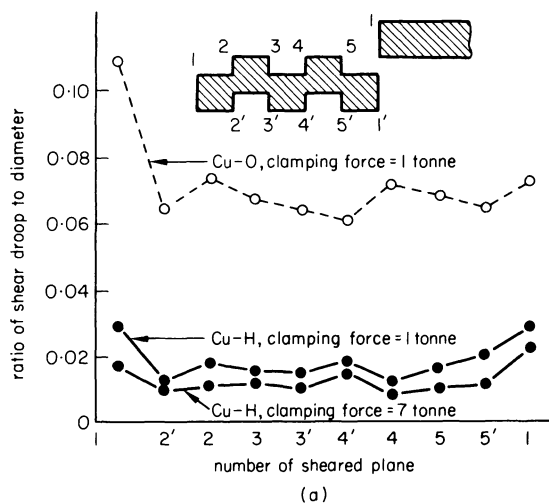


Figure 8. Dimensional accuracy of the shear-formed copper crankshaft. (a) Distortion of the shear-formed section by the shear droop. (b) Deviation of the sheared surface from the perpendicular.

from the real circular shape. The distortions of the shear-formed sections 2-5 are less than those of the sheared-off surfaces 1 which correspond to conventional clamp-shearing. This distortion is notable with soft or annealed material. The perpendicularity to the bar axis is perfectly good for the shear-formed section, but the sheared-off sections deviate from the perpendicular. This deviation can be decreased by increasing the clamping force as shown in figure 8(b).

PRODUCTION OF CRANKSHAFT BY CLAMP SHEARING

Shape of crankshaft

The above results show that a crankshaft can be produced without any difficulty by the clamp-shearing method, and the development of a production tool will now be discussed. Figure 9

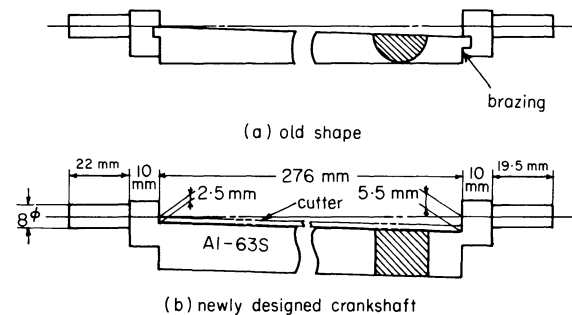


Figure 9. Old and newly designed crankshafts.

shows the old type of crankshaft made by brazing two parts together, and also a newly designed crankshaft produced by clamp-shearing. This crankshaft is used for the rotating shearing paper-cutter in an electric copying machine. The eccentricity of the crankshaft in this case is different on each side. The new crankshaft consists of a quenched die steel cutting blade and Al-63S die holder, and the cutting blade is fastened to the crankshaft holder. The holder is shear-formed from aluminium bar of square section by the clamp-shearing method and both ends of the bearing are machined after forming. The material Al-63S of the holder was selected because it is ductile enough to be clamp-sheared without fracture.

Tool construction and design for production

The newly designed tool shown in figure 10 is not essentially different from the previous experimental tool except that the clamping forces are supplied by urethane and pneumatic cushions, whereas a hydraulic source and a pneumatic cushion were used in the experimental tool. The tool in figure 10 can also be attached to the conventional mechanical press with a pneumatic cushion. The lower tool has a square groove which fits the bar material, and as a result the lateral expansion of the bar is restricted at the time of shearing.

The shearing and forming operation is as follows. The upper tool comes down with the ram of the press and the bar is half sheared in the clamped condition by four urethane cushions. After the middle part of

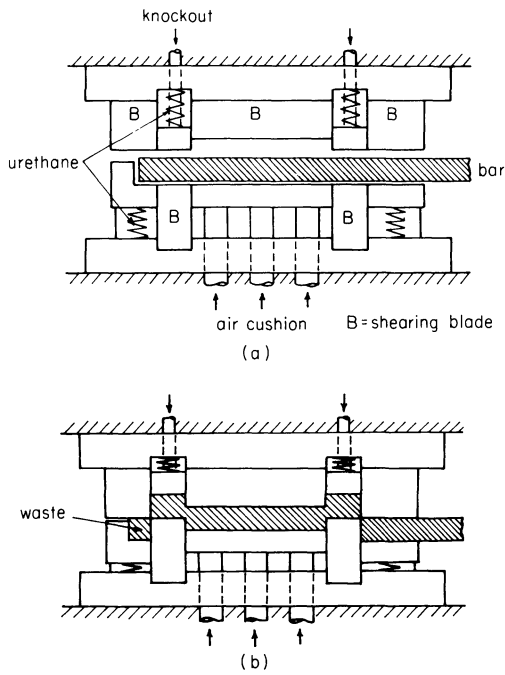


Figure 10. Production tool for mechanical press. (a) Before shearing operation, (b) end of shear-forming operation.

the bar touches the pneumatic cushion plate in the lower die, the clamping force of the pneumatic cushion starts to work. The press ram goes down further and both ends of the bar are separated off by shearing. Thus the sheared bar is shear-formed to the crankshaft.

The part sheared off the left-hand end of the bar is the waste material; this is sheared to avoid fracture by the shearing of simultaneous multi-planes in the case of insufficient clamping force.

Results and some modifications of the tool

The most important disadvantage of the tool in figure 10 when used with a mechanical press is an insufficient clamping force. In clamp-shearing a desirable clamping force of one to four times the shearing force is desirable. In this tool, however, the clamping force of the pneumatic cushion is at most 5 tonnes and of each urethane cushion, only 100 kg.

Due to the insufficient clamping force of the urethane cushions, both sheared ends of the crankshaft were not fracture-free, but this is not a production problem because these ends can be machined off afterwards. In the sheared surface of the shear-formed portion of crankshaft, small scale-like fractures were observed owing to the insufficient clamping force. These fractures were not harmful so far as the strength of the crankshaft was concerned unless and until they entered the connecting portion.

These scale-like fractures were eliminated by modifying the tool so that it could be used with a double-action hydraulic press. The modified tool is attached to the press just upside down, as shown in figure 11. The clamping force of the inner ram was 20 tonnes, enough to eliminate the fracture in the connecting portion. However several problems arose with the use of the hydraulic press, such as low production speed, difficulty of adjustment of the

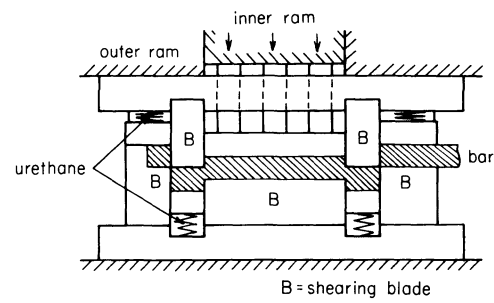


Figure 11. Production tool for double-action hydraulic press.

starting time of the inner ram and an insufficient ejecting force for the lower die.

A satisfactory fracture-free crankshaft was finally obtained by adopting stepped negative clearance of the shear-forming cutting edges of the tool for the mechanical press as illustrated in figure 12. The

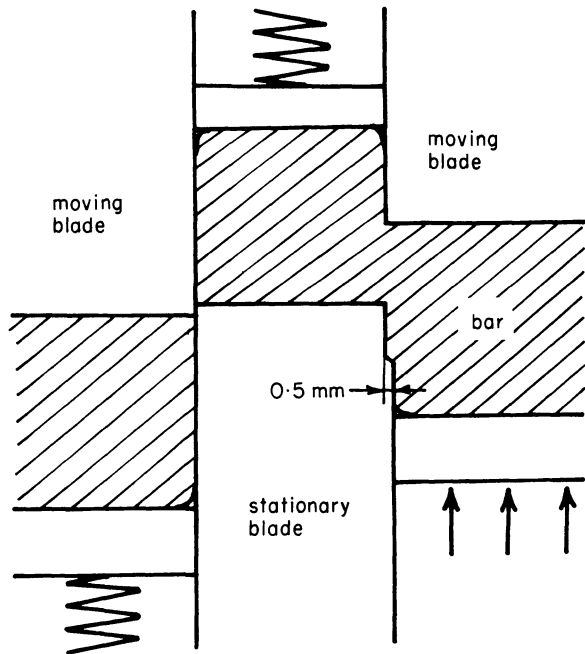


Figure 12. Stepped negative clearance for prevention of crack.

surplus material in the negative clearance zone protrudes as a burr on the lateral side of the bar, but this was only a problem of appearance. The next problem was that of ejecting the shear-formed crankshaft from the lower tool. Even with the pneumatic cushion the ejection was not stable because of friction between the crankshaft and the side walls of the groove. The section of the bar was changed to a slightly tapered section.

Figure 13 shows the appearances of the shear-formed products and the final product with the cutter blade. The shear-formed ferrous crankshaft produced by the same tool is also shown in figure 13.

The advantages of this process are that the production cost is decreased by 30 per cent as shown in Table 2, the unstableness of brazing and the inaccuracy resulting from brazing are eliminated, and the weight of the crankshaft is decreased.

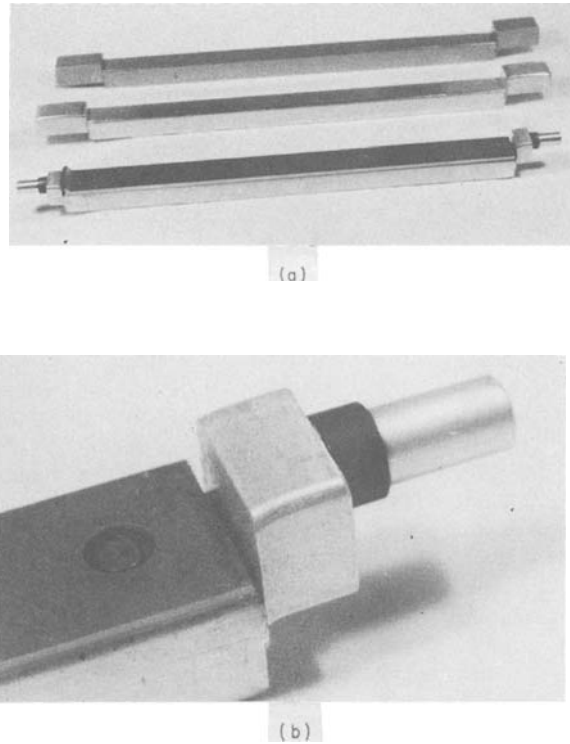


Figure 13. Appearances of shear-formed products (a) Upper: pure iron; middle: Al-63S; lower: final product of Al-63S. (b) Shear-formed part of final product.

TABLE 2 Cost comparison of two types of crankshaft shown in figure 9

	(a) Machining and brazing	(b) Shear-forming and machining
Material cost	7	5
Manufacturing cost	93	67
Total cost	100	72

CONCLUSIONS

Clamp-shearing of a bar was successfully applied to the production of a crankshaft. This method can also be used for soft steel such as an annealed low-carbon steel. The life of the tool² in the clamp-shearing of steel is not fatal here because of a shorter shearing length, and tool wear does not directly affect the burr. Modifications for suppressing fracture of the sheared surface such as negative clearance and roundness of the cutting edges, can easily be adopted.

The fatigue limit of the connecting portion decrease may be because of the notch effect and fibre flow of the shear-formed portion perpendicular to the bar axis. Generally this is not a problem, but if the fatigue consideration is more important, then the cutting edges of the blades can be rounded or chamfered to remove the notch effect and heat treatment can be given to the shear-formed crankshaft to eliminate fibre flow and work-hardening.

ACKNOWLEDGMENT

The authors would like to thank Mr K. Koizumi for his kind assistance in the experiment with the preliminary tool, Mr T. Yagi, Tool Design Section of Toshiba Electric Co., for his assistance in tool design, and Mr K. Fushimi, Managing Director of Toshiba Electric Co., for his encouragement of this development.

A part of this research was sponsored by the Japanese Ministry of Education from the research fund of 1971.

REFERENCES

1. T. Nakagawa and T. Maeda. Billet production by clamp shearing method. *Preprint 4th Int. Cold Forging Conf., Düsseldorf, 1970*, p. 351.
2. T. Nakagawa and K. Miyamoto. Tool-life test in clamp shearing of steel bar and wire. *Proc. 12th Int. MTDR Conf., 1971*, p. 567.

ANALYSIS OF STRESSES IN BAR CROPPING

by

V. G. WONG* and M. K. DAS†

SUMMARY

Expansion of the use of metal forming operations has resulted in an increasing demand for high quality billets. Cropping, because of the lower material wastage, compared with traditional machining operations, can result in much reduced costs for billet production. An understanding of the mechanics of cropping is therefore of primary interest. Analysis of the process is complicated by its unstable nature, with boundary conditions and internal stresses progressively changing until fracture occurs. However, a number of attempts have been made to analyse the stresses which warrant consideration.

The present paper examines critically some of these analytical attempts with the aim of assessing their practical usefulness and contribution to the overall understanding of the cropping process.

NOTATION

C, C_x, C_y	clearances ($x =$ axial, $y =$ radial)
$C_1 \dots C_6$	coefficients
E	modulus of elasticity
F	total friction load
G	modulus of rigidity
P	total cropping load
R	force at ends of bar
S_0, S_1, S_2	function
$2a$	length of the beam
$2b$	depth of the beam
$2c$	breadth of the beam
d	length of contact tool-bar
$2d_0$	axial clearance
$2l$	distance between the line of action of the loads
n	counter
u, v	displacements in x and y directions
Δ, ∂	operator
λ	Lame's constant
μ	coefficient of friction
π	constant
$\bar{\sigma}$	equivalent stress
σ_x, σ_y	normal component of stress parallel to x and y direction
τ_{xy}	shear-stress components
ϕ	stress function

INTRODUCTION

Many slow and uneconomic chip removal processes are gradually being replaced by more productive and economic operations. It is recognized that in this

respect cropping, because of reduced material wastage and higher production rates, can result in substantial economy for billet production^{1,2,3}. However, the resulting economic advantages are often offset by the requirement for secondary sizing and stress relieving operations to improve the quality of as-cropped billets. For this reason, there is a sustained effort towards producing billets of acceptable quality in as-cropped condition without these secondary operations through such methods as improved tool design, superposition of hydrostatic pressure ('cold flow' shearing) or the introduction of high cropping speeds^{4,5,6}.

Improved tool design for cropping has evolved through the process of trial and error and practical experiences over a long period of time^{7,8}. A well-designed tool usually has provisions for off-cut support to reduce bending, good alignment and rigidity, and facility for setting tool parameters like blade face clearance and so on. Such adjustability is of primary interest for conventional slow-speed cropping as the billet quality depends on a sensitive balance between tool parameters and properties of the bar material^{9,10}. With higher cropping speeds, however, the need for such sensitive adjustment is much reduced and a vastly superior billet is obtained in most engineering materials⁶. The exceptions are very ductile materials which do not readily produce well-defined cracks and within this group the technique of 'cold flow' shearing is effective in producing a mirror like surface⁵. The adoption of the latter, after initial enthusiasm, has been slow due to the complicated toolings and poor die life.

Even a cursory look at the metal forming literature

* University of Zulia, Maracaibo, Venezuela. Formerly of Department of Mechanical Engineering, Birmingham University

† Department of Mechanical Engineering, Birmingham University

suggests that very few serious studies have been undertaken on the cropping process in spite of its economic implications. In order to understand why billets of so vastly differing qualities are found in practice, it is necessary to examine the internal stresses developed in the bar due to the prevailing boundary conditions and its mode of failure. Some such studies in this direction are discussed in the present paper.

THE CROPPING PROCESS

In its simplest form, the cropping process consists of the bulk separation of the bar due to a relative movement of a pair of hardened blades in the transverse direction as shown in figure 1. Such a

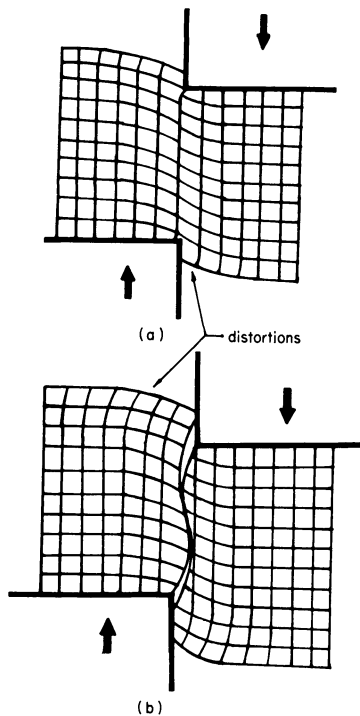


Figure 1. Types of cropped surfaces. (a) Ductile material. (b) Crack-sensitive material.

movement, due to resistance offered by the bar material, generates load at the boundaries. The problem is of a dynamic nature as the conditions of load and displacement change with the progressive movement of the blades. However, pressure very rapidly builds up at the contact surfaces, creating an internal stress field of considerable magnitude in the bar. Under these stresses the material suffers some initial deformation giving the appearance of barreling (figure 1). It is observed that these particular distortions are a function of material properties (decreasing with increasing tensile strength¹¹) and are always present unless very high axial pressure is superimposed making it impossible for any large distortion to occur^{5,12}. With a further rise of pressure the material beneath the cutting edges gives way to penetration through localized yielding. After a further penetration of the blades one of the following two modes of separation takes place. In the case of ductile materials the section fails though metal flow in the plane of tool motion, without pronounced

surface irregularities (figure 1(a)) provided that the tool clearance is optimum. Even in such a case, an instability ensues near the end of the process and a small portion of the section severed may actually be torn. However, with less ductile materials two crack fronts propagate from the cutting edges and the billet is severed when these meet. The fractured surface has a characteristic S-shape as indicated in figure 1(b), and severance takes place long before the cutting edges meet. In fact, the stroke necessary to initiate cracks is only a fraction of the lateral dimension of the bar.

It is clear from this that the mechanism of the cropping process is relatively simple for highly ductile materials, where the material experiences a full plastic stress and for separation to occur the blades must penetrate nearly the complete depth. Such a process (which includes hot cropping and shearing of very thin sheets) can be analysed by applying slip-line field theory with the simplifying assumption of rigid-plastic material^{13,14,15}. These solutions are essentially qualitative in nature, and suffer from an oversimplification.

The other group of materials for which cropping is over within a short stroke of tool movement, showing a characteristic S-shaped profile, is of greater importance as it includes most of the engineering materials. The internal stresses in such materials pass from a complete elastic state at the early stage to a mixed elastic-plastic state as the blade movement progresses.

After a certain amount of blade penetration the cracks are initiated at the cutting edges and the path taken by these crack fronts are not dependent on chance but are actually guided by the prevailing stress field. It will therefore be useful to start the discussion on the cropping process with the work on elastic stresses. Incidentally, it may be noted that in the case of brittle materials the mode of failure occurs within the elastic ranges of stresses by earlier initiation of cracks.

ELASTIC SOLUTIONS

Analysis of cropping and similar shearing processes has been attempted^{16,17} in the conventional manner by using the three considerations:

- (1) forces must be in equilibrium,
- (2) strains are related to the stresses through the generalized Hooke's law, and
- (3) compatibility conditions must be satisfied to ensure the existence of single valued continuous functions for displacements.

Once the equations have been formulated and reduced to a form where solution is possible, boundary conditions are introduced to obtain the particular solution. Like many other processes cropping has been analysed by using a two-dimensional model.

1 Solution with concentrated loads

Filon¹⁶ considered a rectangular beam subjected to two concentrated loads of equal magnitude as shown

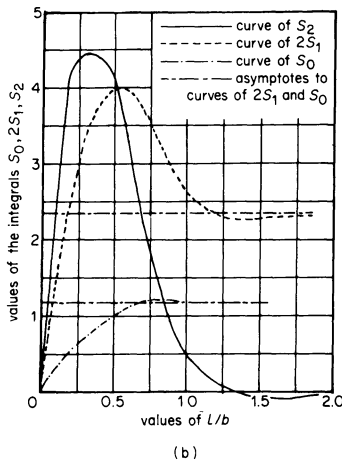
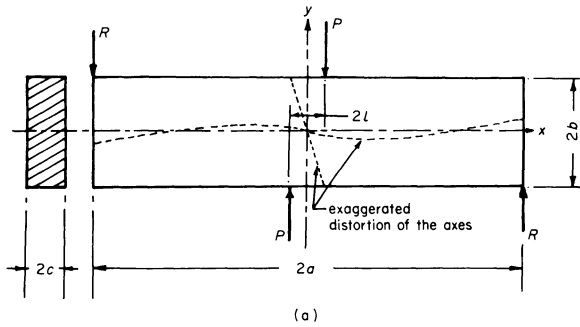


Figure 2. Shearing with concentrated loads. (a) The model and rotation of axes. (b) Variation of coefficients.

in figure 2(a). The dimensions of the beam are $2a, 2b$ and $2c$ and the concentrated loads P are separated by a distance of $2l$, with balancing forces located at the ends of the beam.

For plane strain conditions the equilibrium equations are

$$(\lambda + G) \frac{\delta e}{\delta x} + G \nabla^2 u = 0$$

$$(\lambda + G) \frac{\delta e}{\delta y} + G \nabla^2 v = 0$$

where

$$e = \frac{\delta u}{\delta x} + \frac{\delta v}{\delta y}, \quad \nabla^2 = \frac{\delta^2}{\delta x^2} + \frac{\delta^2}{\delta y^2} \quad (1)$$

and u, v are the displacements in x and y directions, λ is Lamé's constant, G is the modulus of rigidity.

Filon¹⁶ solved Equation (1) by using complex variables to obtain stresses σ_x, σ_y and τ_{xy} as some arbitrary functions. By assuming these arbitrary functions to be hyperbolic and circular functions, and prescribing the boundary conditions in the form of a Fourier series, he was able to obtain a complete solution in terms of the physical properties of the material and the coefficients of the Fourier series representing the boundary conditions. Although a solution was obtained it was not explicit in terms of the length parameter a . He therefore solved the case

of a beam of *infinite* length, which yielded τ_{xy} along the y -axis as

$$\tau_{xy} = -\frac{2P}{\pi b} S_0 - \frac{2P}{\pi b} \cdot \frac{y^2}{b^2} \cdot \left(\frac{S_2 - 2S_1}{2} \right) \quad (2)$$

where S_0, S_1 and S_2 are functions of the ratio l/b , as shown in figure 2(b). It is interesting to note that S_0, S_1 and S_2 vary greatly for $l/b < 1$, but for $l/b \geq 1$ they reach nearly constant values.

Equation (2) is plotted in figure 3 in non-dimensional form. The ordinate represents the location in the depth of the beam, whilst the abscissa

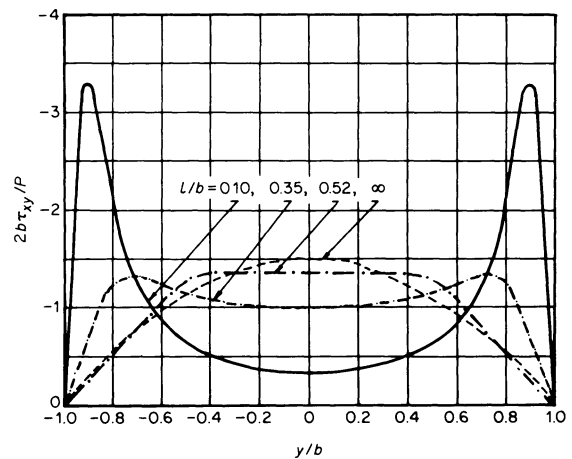


Figure 3. Variation of shear stress along the depth of the beam¹⁶.

represents the ratio of the actual shearing stress to the specified shearing force. As may be seen from figure 3, Filon was able to demonstrate that the popular conception of shearing stress being maximum at the centre of the beam is *only valid* for high ratios of l/b ($l/b \geq 1$). The actual location and value of the maximum shear stress in the beam is dependent on the ratio l/b . If l/b is very small, as in the case of cropping, the maximum shear stresses tend to occur near the top and the bottom. This is an important observation and is widely supported by the experimental observation that failure first occurs near the cutting edges in a cropping type of situation. It may be pointed out that Filon was concerned with this type of problem with a view to a better guide for the calculation of load carrying capacity of rivets¹⁶.

The work of Filon¹⁶ also indicated an interesting phenomenon concerning the *rotation* suffered by the axes of the beam when subjected to an opposing pair of loads P . The equation for these displacements for an infinitely long beam are as follows:

$$v = -\frac{Px l}{Eb^2} \left[2.80 - 4.96 \cdot \frac{(l^2 + x^2)}{b^2} \right]$$

$$u = -\frac{Pyl}{\pi Eb^2} \left[-5.292 + 0.492 \frac{y^2}{b^2} \right] \quad (3)$$

where E is the modulus of elasticity.

Equation (3) is shown graphically in figure 2(a), and it can be seen that the axes of the beam suffer a

rotation in the *anti-clockwise* direction when in fact the loading couple is *clockwise*.

2 Solution with distributed loads

Hojo¹⁷ considered the elastic solution for the cropping problem using distributed cropping loads of the type shown in figure 4(a). Such considerations represent a more realistic boundary condition than do concentrated loads. Furthermore, he introduced a distributed frictional load on the tool contact surface. He attempted to solve the problem by following the conventional procedure for the analysis of stress in the two-dimensional case, that is by considering the equation of the Airy's stress function:

$$\frac{\partial^4 \phi}{\partial x^4} + 2 \frac{\partial^4 \phi}{\partial x^2 \partial y^2} + \frac{\partial^4 \phi}{\partial y^4} = 0 \quad (4)$$

The integration of equation (4) yields the stresses as:

$$\sigma_x = \frac{\partial^2 \phi}{\partial y^2}, \quad \sigma_y = \frac{\partial^2 \phi}{\partial x^2}, \quad \tau_{xy} = -\frac{\partial^2 \phi}{\partial x \partial y} \quad (5)$$

The chosen stress function is a polynomial of x and y as follows.

$$\begin{aligned} \phi(x, y) = & [C_1 \cos \alpha x + C_2 \sin \alpha x] \\ & \cdot [C_3 \cosh \alpha y + C_4 \sinh \alpha y \\ & + C_5 y \cosh \alpha y + C_6 y \sinh \alpha y] \end{aligned} \quad (6)$$

where $\alpha = \pi\eta/a$.

The coefficients C_1 to C_6 are found from the boundary conditions. The problem is further simplified by reducing the length of the bar to points of minimum normal stress as shown in figure 4(b). The detailed manner of derivation of stresses from the boundary conditions (represented by a Fourier series) can be seen in reference 17. As the expressions for the stresses are too long, only the graphical representation of the same is shown in figure 5, where once again a non-dimensional form has been used. This diagram is strikingly like that of figure 2 from Filon's results, except that the additional stresses have also been incorporated. The shear stress distribution across the depth of the beam (along the y -axis) has the same pattern, that is the maximum values occur near the cutting edges with the minimum value occurring near the centre of the bar.

ELASTIC-PLASTIC SOLUTION

It was argued by Wong¹⁸ that theoretical studies, which are based on either elastic or plastic solutions alone, cannot give a true picture, as a typical bar material not only has important elastic-plastic properties, but also exhibits work-hardening characteristics. To obtain a realistic picture, he continued, an analysis of the cropping process should consist of the following:

- (1) A realistic model for the representation of the boundary conditions.
- (2) A solution for the complete stress-field under such boundary conditions by considering the actual behaviour of the work material.
- (3) Assumption of a logical hypothesis to predict the cropped profile.

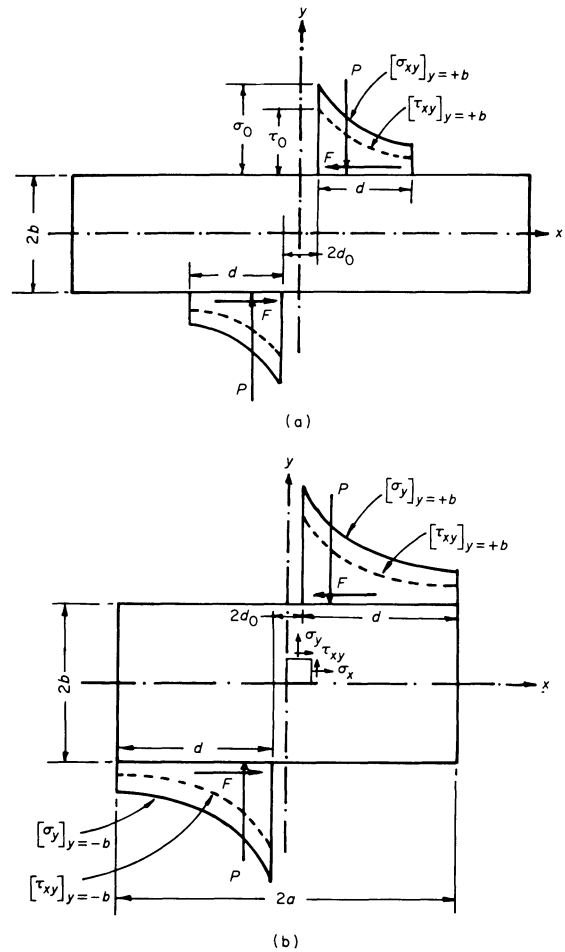


Figure 4. Cropping with distributed loads. (a) Bar length greater than the contact length. (b) Bar length equal to the contact length.

The details of an analysis based on the above guide lines is being published elsewhere¹⁹. Basically, the cropping load distribution is found from the properties of tool-bar interface and the solution for the stresses is found by using the finite element method. Once the stress field is established the expected fracture path is predicted on the basis of strain energy in the cropping zone.

1 The model

The case of a rectangular bar under closed type tooling is considered where conditions at the static and moving blade are taken to be similar (figure 6). The material is taken to be elastic work-hardening as is the case with common billet materials. Figure 6(a) shows the initial stage of the process. At the beginning of the process the whole surface of the bar makes contact with the tool surface, but when slight pressure is applied the bar suffers a rotation which causes the contact length to change, and the loads become concentrated near the cutting edges (A and B). The ends of the bar, however, now make contact with the flat surfaces at (1) and (2) respectively, as shown in figure 6(b). The magnitude of rotation is dependent on the clearance C_y and the length of the bar. With a further downward motion of the moving blade the concentrated loads at A and B get distributed over a finite length d . This line of contact is

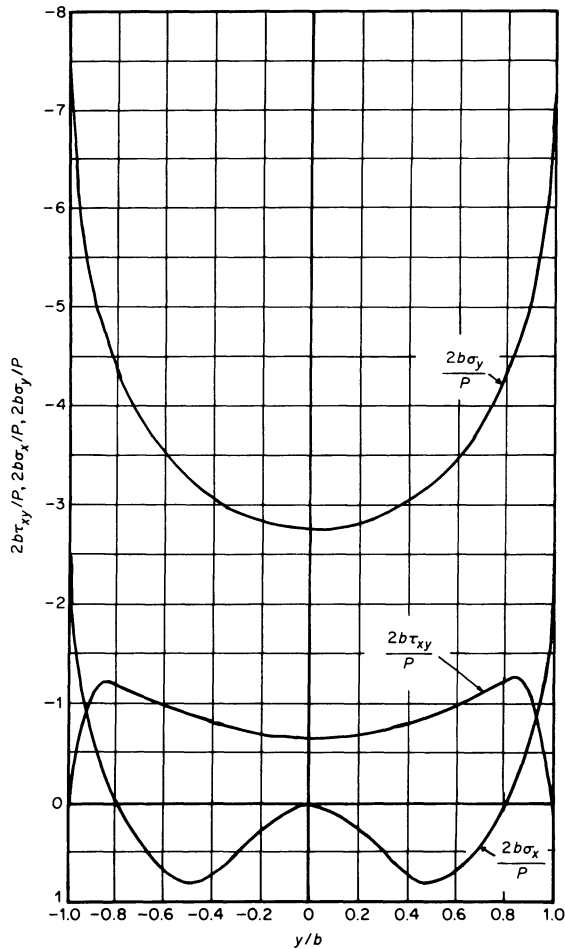


Figure 5. Variation of stresses along the depth¹⁷.

horizontal only if the tool suffers no rotation*. From the consideration of this rigid-elastic contact^{20,21}, the pressure function is obtained as

$$[\sigma_y]_{y=\pm b} = \frac{P}{2\pi\sqrt{[d^2 - (x - d - C_x/2)^2]}} \quad (7)$$

where σ_y is the pressure at the boundary (width in z-direction being unity),
 d is the length of contact,
 C_x is the clearance,
 P is total pressure.

This function is indeterminate at tip of the tool ($x = C_x/2$ in Equation (6)). However, Love²² established that this value is of considerable magnitude but finite. For all practical purposes, the value at the tip is found by numerical integration of the function bypassing the indeterminate point.

As a result of the vertical loads, the bar experiences a tendency for outward flow in the axial direction, causing frictional forces on the contact surfaces. This frictional pressure will have a similar function to that of vertical ones, except of reduced magnitude, that is

$$[\tau_{xy}]_{y=\pm b} = \mu[\sigma_y] \quad (8)$$

where μ is the coefficient of friction.

* This is effectively achieved when there is provision for good guidance. In the case of high speed cropping, this is further helped by the momentum of the moving blade.

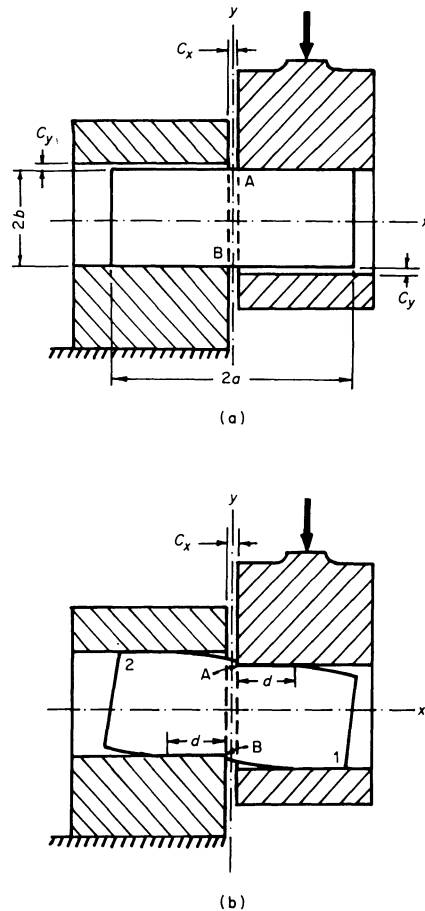


Figure 6. Single cropping arrangement. (a) Initial stage. (b) Advanced stage.

Integration of this distributed pressure gives the total frictional resistance F , and the force R at the ends of the bar is easily found from the equilibrium conditions as

$$R = \frac{P}{a}(l - \mu b) \quad (9)$$

where $2l$ is the distance between the line of action of the loads,
 $2a, 2b$ are the length and depth of the bar respectively.

Thus the complete boundary conditions are arrived at in a logical way and the final model is shown figure 7.

2 The elastic-plastic solution

It is extremely difficult analytically to solve the governing equations for the elastic region, the plastic region and the elastic-plastic interface. This is partly because the compatibility and the stress-strain relations are difficult to handle. Additionally, the location of the elastic-plastic interface is one of the unknowns and is usually of such an awkward shape that even the distribution in the elastic region can only be obtained by laborious numerical methods, as pointed out by Hill²³. However, the necessary stress field has been obtained by the use of the finite element method (FEM)¹⁹ by extending the conventional elastic programmes through the use of the Mises yield criterion and its associated flow rule (Prandtl-Reuss equation)²⁴.

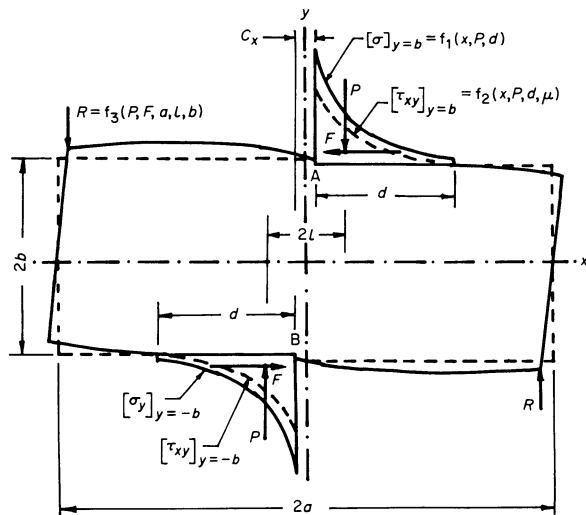


Figure 7. A model for the arrangement shown in figure 6.

The general procedure in the FEM using the displacement approach is as follows:

- (1) *The idealization of the structure.* The continuum is divided into a finite number of elements interconnected at a finite number of nodal points. For bar cropping the mesh is shown in figure 8.
- (2) *The formulation of elemental stiffness matrix.* This is done by considering the geometrical characteristics of the element and the physical properties of the material. Depending on whether or not an element has yielded, an elastic or a plastic stiffness matrix is used¹⁹.
- (3) *The assemblage process.* The continuum is unified by the assemblage of the individual elements. This yields which is known as the master stiffness matrix, the size of which depends on the number of nodal points and the allowed degrees of freedom for the elements.
- (4) *The solution.* The application of the loads and other boundary conditions at the nodal points permit the use of master stiffness matrix to form a very large number of simultaneous equations where the only unknowns are the nodal displacements. An accurate method of

solution (for example the Gauss-Siedel method) is adopted for solving the simultaneous equations leading to the displacement of each nodal point.

Once the displacements are known, various stresses are easily computed for each element. For the cropping model shown in figures 7 and 8, the stresses are calculated with progressively increasing loads¹⁹, that is the stresses are updated by adding the stress increments due to the added incremental load to stresses obtained from the previous loading. The equivalent stress for each element is also calculated for checking whether any element has yielded or in obtaining the stress field needed for the calculation of the fracture path.

Some of the stresses, when the whole field is totally elastic, are plotted in figure 9 for the mid-plane $x = 0$. The results are given in non-dimensional form by dividing the respective stress by the yield stress (Y). Some of these curves can be compared with figures 3 and 5 and the similarity can be noted for small values of l/b . In fact, from the results of the FEM analysis all the stresses can be plotted for any given plane, that is for $x = 0$ to $x = a$, at suitable intervals.

The sequence of the elements suffering plastic failure is shown in figure 10, where the darkened triangular elements are those where the equivalent stress ($\bar{\sigma}$) is greater than (or equal to) the initial yield stress (Y). It is interesting to note that the spread of the plastic zone starts from the first element in front of the cutting edge. The next element to fail is the triangle immediately below the first triangle, and so on. Obviously, a better picture of the spread of the plastic zone would emerge if a still finer mesh around the cutting edges was employed.

3 The cropped profile

In general it is difficult to evaluate the tool penetration after yielding has occurred as it is a complex function of the tool edge condition, angle of penetration due to bar rotation, influence of the progressive depth of penetration and so on. However Wong and Das¹⁹ assumed that crack initiation occurs at the cutting edges at some stage of tool penetration

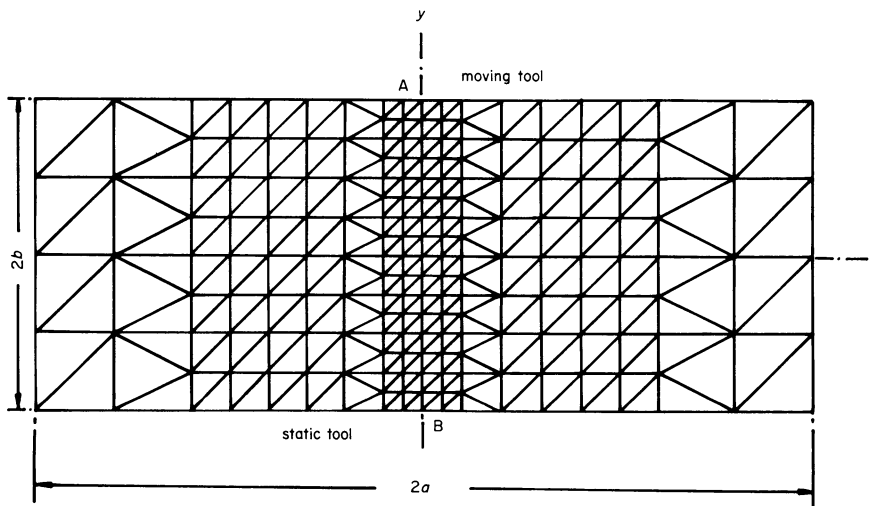


Figure 8. A mesh for single cropping.

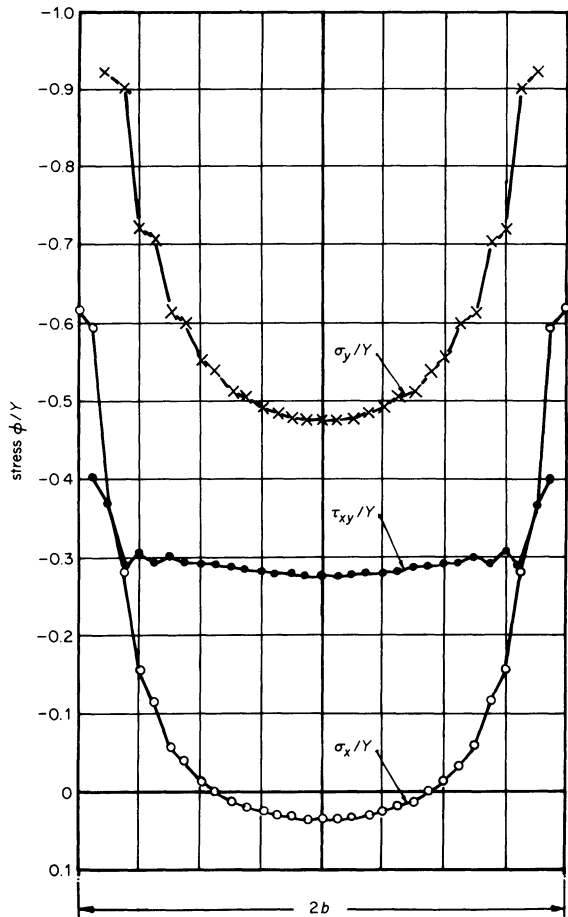


Figure 9. Variation of stresses by the finite element method¹⁹.

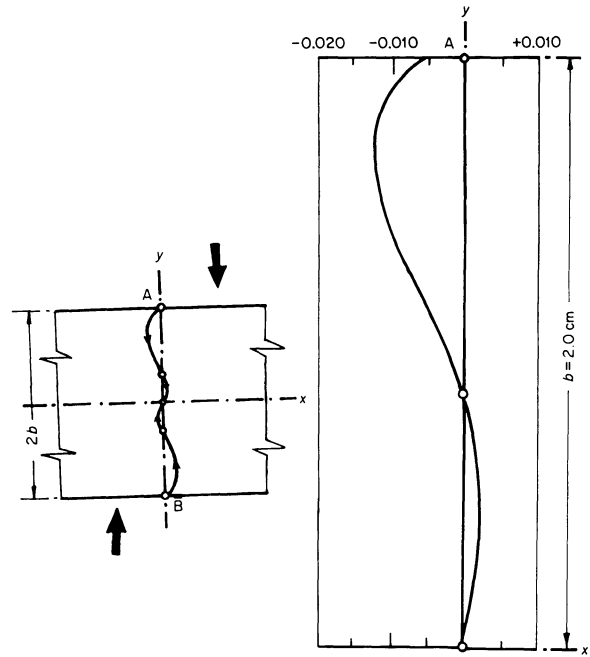


Figure 11. A typical predicted cropped profile.

(see figure 1(b)). In practice, however, the condition on either side of the centre of the tool may not be similar and the plain-strain condition may have been violated, thus resulting in a distorted double S profile. Figure 12 shows a recorded profile at the centre of a round bar cropped at high speed, which indeed shows the presence of a double S-shape.

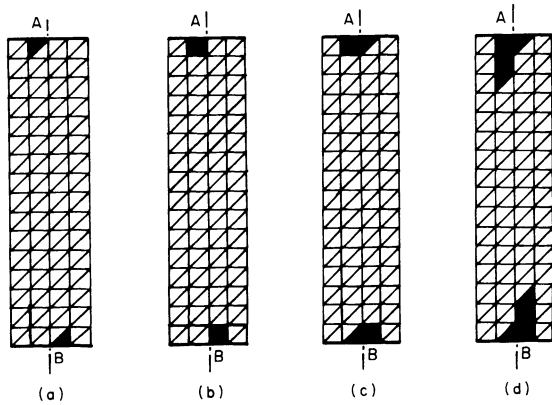


Figure 10. Progressive spread of the plastic region.

and proceeded to predict the path such cracks would follow, as guided by the stress field. In fact the strain energy stored at any part of the body is represented by the equivalent stress prevailing at that part of the body²³. Thus the fracture path to be followed by the cracks initiated at the cutting edges is given by the curve joining the locations of maximum equivalent stresses, as the *maximum energy releases* can occur only along this trajectory. A typical cropped profile predicted from such considerations for their model of single cropping is shown in figure 11. It is interesting to see that the predicted profile is *double S-shaped* with smaller width at the central part, rather than a single S-shape as is commonly thought to be the case

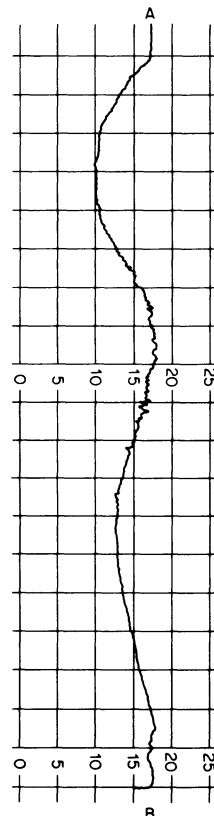


Figure 12. An observed cropped profile showing double S characteristics (Al Alloy, high-speed cropping).

The other important observations from this study include:

- (1) the quality of cropped profile deteriorates rapidly with worn-out tools, that is delayed penetration, and
- (2) the cropped profile is marginally affected by axial blade clearance as long as the tool guidance is good enough to ensure that the tool contact is horizontal, giving an asymptotic load distribution.

This latter finding is relevant as in high-speed cropping, where the load distribution may have been maintained by the momentum of the tool, the billet quality is relatively unaffected by blade clearance.

CONCLUSIONS

Some studies in connection with the analysis of stresses in cropping or similar shearing processes have been examined in order to obtain a better insight on the process. In particular the relative magnitudes of the internal stresses in the bar facilitate explanation of commonly observed phenomena such as initial failure at the cutting edges, when the centre of the bar may still be elastic. The elastic-plastic analysis using the finite element method is of great significance as this permits, for the first time, an accurate prediction of the cropping profile from the realistic presentation of the prevailing boundary conditions.

ACKNOWLEDGMENTS

The authors thankfully acknowledge the support and advice given by Professor S. A. Tobias, under whose general guidance the cropping work in the Department is being carried out.

REFERENCES

1. J. C. Widmont. The cheapest cold forming slug. *Metal-Working Production* (6 December 1961), p. 78.
2. F. Bakhtar and E. R. Austin. Relative economics of conventional and high strain-rate forming. *Int. J. Mach. Tool. Des. and Res.*, (1965), 5, 139.
3. A. J. Organ. Economic aspects of metal billet production. *Proc. 9th MTDR Conference* (1968), Birmingham.
4. Anon 'DIN 8588' Begriffeder der Fertigungstechnik Zerteilen, Entwurf (September 1963).
5. A. Szamosyolgyr. Wasteless production of intermediates for cold flow process. *Hungarian Heavy Industries*, First Quarter (1965), no. 48.
6. M. K. Das and S. A. Tobias. The petro-forge HERF cropping system. *Society of Manufacturing Engineers* (1972), MF 72-143.
7. E. G. Izod. Behaviour of materials of construction under pure shear, *Proc. Instn. Mech. Engrs.* (1906) 5.
8. Anon, Bar cropping billet defects. *British Cold Forging Group*, Data Sheet No. 72004 (1972).
9. O. Kienze and H. Zabel. Zerteilen metallischer Stangen durch Abscheren Opalden: *Westdeutscher Verlag* (1965).
10. T. M. Chang and H. W. Swift. Shearing of metal bars *J. Inst. Metals* (1950), 78, 119.
11. G. V. Marston. Shearing of bar and billet forging stock: hydraulic counter-pressure shear, *Metal Treatment and Drop Forging* (December 1963) p. 497.
12. S. Fukui, T. Maeda and T. Nakagawa. Cold extrusion of fine sheared slugs. *Scientific paper of the Institute of Physical and Chemical Research* (1970), 64, no. 1.
13. T. Jima. The theoretical research on the blanking of a sheet material. *Bulletin of JSME* (1963), 6, 568.
14. L. J. Kronjo. Some extrusion and indentation problem in plane plastic strain, *Ph.D. Thesis*, University of Birmingham, 1967.
15. W. Dos Santos and A. J. Organ. Deformation in the ductile fracture processes examined by the viscoplasticity method *Int. J. Mach. Tool Des. Res.* (1973), 13, 217.
16. L. N. G. Filon. On an approximate solution for the bending of a beam of rectangular cross-section under any system of load, with special reference to points of concentrated or discontinuous loading, *Trans. Roy. Soc. (London)*, ser. A (1903), 201, 67.
17. H. Hojo. Shearing mechanism of phenolic paper base laminated sheets, *Annals of the CIRP* (1967), XIV, 409.
18. V. G. Wong. Some aspects of bar cropping. *Ph.D. Thesis*, University of Birmingham, 1973.
19. V. G. Wong and M. K. Das. An Analysis of the Bar Cropping Process by Finite Element Method. To be published.
20. M. Sadowsky. Zqueidimensionale Probleme der Elastizitatstheorie, *Z. Angew Math. Mech.*, (1928), 8, p. 107.
21. S. P. Timoshenko and J. N. Goodier. *Theory of Elasticity*, McGraw Hill, 1970.
22. A. E. H. Love. The stress produced in a semi-infinite solid by pressure on part of the boundary, *Trans. Roy. Soc. Series A* (1929), 228, 337.
23. R. Hill. *The Mathematical Theory of Plasticity*. Clarendon Press, Oxford, 1971.
24. Y. Yamada and N. Yoshimura. Plastic stress-strain matrix and its application for the solution of elastic plastic problems by the finite element method. *Int. J. Mech. Sci.* (1968), 10, 343.

COMPACTION AND SINTER FORGING

THE DYNAMIC COMPACTION OF ALUMINIUM AND IRON POWDER

by

D. RAYBOULD*

SUMMARY

The paper presents data on the compaction of aluminium and iron powders at impact velocities of up to 500 m/sec. A brief description of the compaction machine and the experimental techniques employed is given.

A theoretical analysis of the pressure build up during dynamic compaction is presented and its predictions are compared with the observed data. The resultant tensile strength and the density of the compacts are reported and compared with the properties of the compacts produced quasistatically. It is observed that at high densities dynamic compaction increases the strength and the density of the specimens.

INTRODUCTION

The high velocity compaction of powders has now been under investigation for approximately 15 years. Although the process is not yet fully understood it is now feasible to take a more informed look at its possible advantages.

An important advantage of dynamic compaction is the ability to produce very high density homogeneous parts using smaller and less rigid machines than those required by quasistatic techniques. According to Davies and Austin¹ and Vityaz and Roman² the resultant compacts are between 10 and 30 per cent stronger than compacts of the same density produced quasistatically. This increase in strength is caused both by the larger areas of intermetallic contact between the grains² and the increased dislocation density. The larger areas of intermetallic contact together with the greater number of structural defects, is also responsible for the accelerated diffusion that occurs during sintering². Other advantages are: - improved dimensional accuracy³, reduced friction between the compact and the die wall¹ and a lower ejection pressure than after quasistatic compaction³, however, the latter advantage may not prevail under production conditions.

There are several disadvantages which must also be considered. For instance, it is difficult to obtain accurate dimensions in the direction of punching. Also, and perhaps of more importance, are the uncertainties surrounding the die life. This is a problem common to all High Speed Forming (HSF) operations, although in the case of powder compaction, the shock absorbing capacity of the powder ensures that the pressure is built-up much more

gradually than when the punch impacts a solid. In addition, it is possible that the introduction of high speed visco-elastic techniques⁴ could in many cases reduce tooling costs below those for quasistatic techniques.

In the last 15 years, a second group of workers has taken a keen interest in the compaction process, primarily because powders and porous solids have been found to be effective shock attenuators⁵. While the impact velocities studied by these workers are far above those employed by such machines as the Petro-Forge, they do coincide with the velocities employed during explosive compaction.

The aim of this work is to investigate quantitatively the region between normal High Energy Rate Forming (HERF) compaction velocities and the much higher velocities characteristic of pressure attenuation studies, with the objective of optimizing some aspects of dynamic compaction. Elwakil and Davies³ for instance, concluded that advantages could be obtained by using velocities higher than those available from the Petro-Forge and this appears to be supported by the information presented by Vityaz & Roman².

EXPERIMENTAL TECHNIQUE

Compaction machines

To conduct the high speed compactions a simple compressed air 'cannon' with a maximum piston velocity of 600 m/sec was constructed. It consisted of a 300 bar 4 litre reservoir, a specially designed fast action valve and a tube to guide the piston (punch) into the compaction chamber, as shown in figure 1. The compressed air for the 'cannon' was obtained

* Institut CERAC S.A., 1024 Ecublens, Switzerland

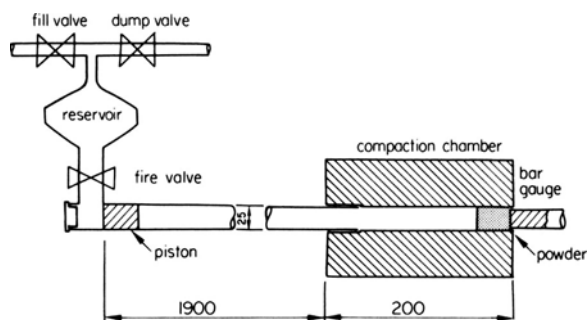


Figure 1. Schematic side view of the dynamic compaction machine.

from an Atlas Copco BP3 compressor. The possibility existed of drawing a vacuum of 5×10^{-5} bars in the guide tube and compaction chamber and hence in the powder.

The static experiments were carried out in a 2 MN hydraulic press, which was instrumented to record pressure and punch displacement.

Tooling

The same 25 mm internal diameter compaction chamber was used for both the dynamic and quasi-static compactions. The chamber was made from EN24 steel tempered to give a Hardness Vickers of 400, and, for simplicity, a mono-block construction was employed. The chamber was designed to take a maximum quasistatic stress of 100 N/mm^2 , but this was exceeded dynamically many times without any ensuing damage.

The dies, punches and projectiles were also made from EN24 steel, though some of the latter were made from aluminium and PVC.

Materials

Höganäs sponge iron powder MH 100.28 and Dural atomized aluminium powder AS/71 were the powders used. The properties of the aluminium powder are: Tap density 45 per cent of the theoretical density, free fall density 44 per cent of the theoretical density, average grain size around 100 microns. The properties of the iron powder were as specified by Höganäs.

The die walls were lubricated with PTFE spray, no experiments were conducted with admixed lubricants.

Instrumentation

The instrumentation was designed to measure both the velocity of the piston (projectile) just prior to impact and the pressure transmitted through the powder to the support wall at the opposite end to that of impact. Complex experimental techniques, as employed by Linde & Schmidt⁵ and their co-workers⁶, were not employed mainly because it was desired to keep the experimental techniques relatively simple so that a large range of variables could be investigated quickly.

Velocity measurements

The velocity of the projectile (punch) was obtained by measuring the length of time that the projectile interrupted a light beam from a quartz-halogen lamp

directed onto a photo-multiplier. This provided not only a reliable velocity measurement, but also a reliable external trigger for the pressure recorders.

Pressure measurements

A Hopkinson type bar gauge was used to measure the pressure. The one metre long bar was made of maraging steel (DIN X2NiCoMo 18,9,5). One end of the bar was positioned just inside the compaction chamber and the other end was supported by a heavy flange.

Strain gauges forming two legs of a bridge circuit were mounted on the bar a few centimetres from the powder-bar interface and the output from the circuit was fed to a high quality amplifier. Aluminium and steel projectiles were impacted onto the bar so that it could be calibrated dynamically. The velocity of these projectiles was measured and from this the pressure created at impact was calculated. Cross reference with the pressure predicted by the gauge factor indicated that the latter underestimated the dynamic pressure; the difference apparently depended on the type of glue used, but was constant for a particular glue.

Manganin foil piezoresistive stress gauges were also used to measure the pressure in a few of the experiments. The manganin gauge was cemented to a 50 mm thick steel disc which was bolted to the back of the compaction chamber. A protective layer of 0.03 mm thick mylar was cemented over the gauge's active area which measured 0.53 by 0.63 mm. The gauge had a nominal pressure calibration of 0.027 kbar^{-1} and the bridge circuit containing the gauge was powered by 50 volt pulses with a duration of 3 msec.

Temperature measurements

The mean temperature rise caused by compaction was measured by using mineral insulated Chromel-Alumel thermocouples which were clad in stainless steel and had bounded hot junctions. The necessity of a very fast response time together with the small increase in temperature, indicates the desirability of a small diameter thermocouple. However, the high pressure results in the powder extruding out through the thermocouple access hole and, in the process, breaking the thermocouple. To avoid this, thermocouples 1 mm in diameter were employed.

Recording instruments

For most of the aluminium experiments standard high frequency oscilloscopes were used. The velocity measurements were recorded on a storage oscilloscope. The pressure and temperature were recorded by camera on polaroid film. For all the iron experiments, and some of the aluminium, a Biomation, model 8100, two channel digital recorder was used. This allowed the simultaneous recording of velocity and pressure.

Analysis of the compacts

The properties of the green compacts were investigated by the use of micro-hardness tests, density measurements and optical microscopy with the aim of showing the variation of properties within the compacts, as well as with the compaction conditions.

A 300 grm weight was used for the micro-hardness measurements which were carried out on polished surfaces. The density measurements were carried out by Archimedes' method. Diametral compression tests were also conducted using an Instron testing machine.

THEORETICAL ANALYSIS

The data obtained from the compacts differs on several points from previous high speed compaction data¹. The density and hardness of the compacts were found to vary with length in a manner similar to that reported by Vityaz and Roman². They explained this by the generation of a reflected shock wave from the rigid base of the container. In addition, the compaction pressure increases in a series of steps, as shown in figure 2, rather than by a smooth curve or

quasistatic pressure density relationship. This assumption is justified by the reported strain rate insensitivity of the pressure density relationship of porous solids⁷. However, as that work was conducted at very high velocities and pressures, the properties of the solid, rather than those of the powder, predominated. Many equations have been put forward as representing the static pressure density relationship. It was found that up to the maximum pressures attainable with the static press, Equation (1) represents the experimental relationship found for aluminium and iron powder reasonably well. An additional advantage of the equation is its relative simplicity.

$$\rho = \rho_0 \left(a + \frac{1-a}{1+bP} \right)^{-1} \tag{1}$$

where *a* is a constant representing the initial porosity of the powder and *b* is a constant representing the stiffness of the powder.

It was also assumed that the specific internal energy of the porous material is the same as that of the compacted solid under identical conditions of pressure and temperature. This simplification was also employed by Herrman⁸. In effect, this states that the surface area of an individual powder particle is the same in the loose and compacted state. Obviously, during compaction, a small decrease in surface area may occur and the energy per unit area may decrease, but these are probably fairly well compensated for by the stored energy of deformation. This assumption, together with the fact that only relatively low pressures were of interest, permitted the changes in internal energy across the shock to be ignored. However, those changes in the internal energy that are included in the quasistatic pressure-density relationship were taken into account. Actually, as the changes in the particle's surface area are included in the quasistatic curve, the assumption that the internal energy of the loose powder and compacted solid are equal may be unnecessary.

Finally, it was assumed that for the pressures of interest in this investigation, the solid is incompressible, apart from the volume changes included in the quasistatic pressure-density relationship.

Considering a shock wave moving through the powder, as shown in figure 2. Across the shock mass and momentum must be conserved^{6,9} so

$$\delta(m) = \delta(\rho u) = 0 \tag{2}$$

and

$$\delta(P + mu) = 0 \tag{3}$$

where *m* is the mass per unit volume, ρ is the density and *u* is the velocity of the material relative to the shock.

Substituting the constitutive equation into these two conservation equations allows the velocity of the shock wave and the velocity of the material behind the shock wave to be calculated.

From Equations (2) and (3) it follows that

$$u_1^2 = \frac{P_2 - P_1}{\rho_1 \left(1 - \frac{\rho_1}{\rho_2} \right)} \tag{4}$$

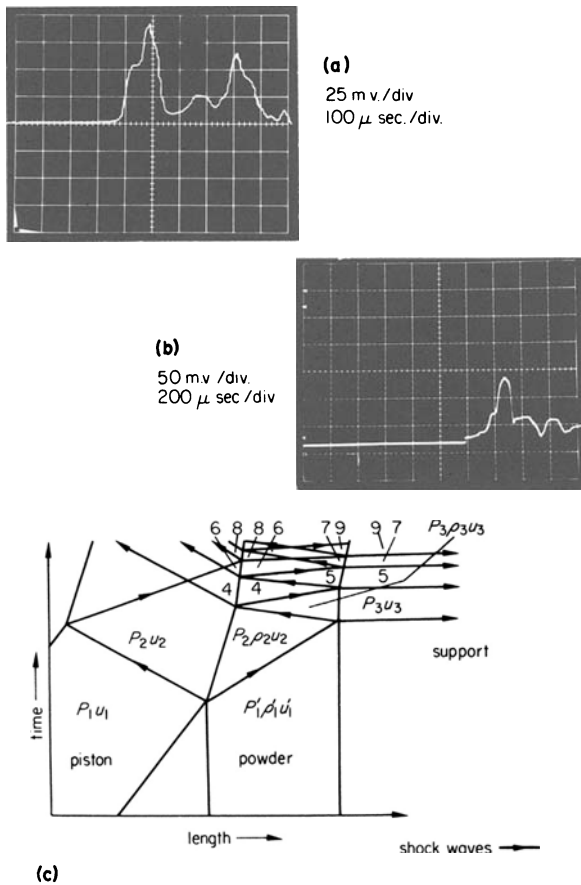


Figure 2. Pressure variations within the powder. (a) Pressure time trace recorded by the bar gauge, with a vacuum in the guide tube. (b) Pressure time trace recorded by a manganin gauge, with no vacuum. (c) Schematic representation of the shock waves motion within the powder.

by a single step. To explain these phenomena, a detailed study of the shock wave interactions within the powder, the projectile and the die supporting the powder is necessary. This requires knowledge of the dynamic pressure density relationship for the powder, the velocity of the shock wave in the powder and velocity of the powder itself.

Shock wave conditions

The constitutive equation for the powder was assumed to be an equation of the same type as the

where subscript (1) indicates the conditions in front of the shock and subscript (2) the conditions behind the shock.

If the shock is moving into stationary material then

$$u_1 = U$$

where U is the shock velocity. Substituting this into Equation (4) and rewriting in terms of pressure by using Equation (1) gives

$$U^2 = (1 + aP)^2 \frac{1 + P_2}{1 + P_1} \tag{5}$$

where

$$U = \frac{U}{c_0}$$

$$P = bP$$

and

$$c_0^2 = \frac{1}{(1 - a)b\rho_0}$$

For aluminium $c_0 = 193$ m/sec; while for iron $c_0 = 170$ m/sec. The particle velocity u is given by the change in velocity across the shock. The passage of the shock gives the material a velocity u , from which it follows directly that

$$u_2 = U - u$$

Substituting this in Equation (2) yields

$$\rho_1 U_1 = \rho_2 (U_1 - u)$$

So

$$u = U \left(1 - \frac{\rho_1}{\rho_2} \right)$$

Using Equations (1) and (3) this result can be rewritten in terms of the pressure to give

$$u = \frac{(1 - a)(P_2 - P_1)}{[(1 + P_1)(1 + P_2)]^{1/2}} \tag{6}$$

where

$$u = \frac{u}{c_0}$$

Rewriting Equation (6) and letting $P_1 = 0$ gives

$$\frac{u}{(1 - a)} = \frac{P_2}{(1 + P_2)^{1/2}} \tag{7}$$

A plot of Equation (7) is shown in figure 3, this can be considered as a dimensionless Hugoniot which is probably applicable to metal powders, which obey Equation (1).

It is now possible to calculate the pressure created upon impact, the pressure created after reflection at an interface, and the pressure resulting from shock wave interactions within the powder. Details are presented in Appendix 1.

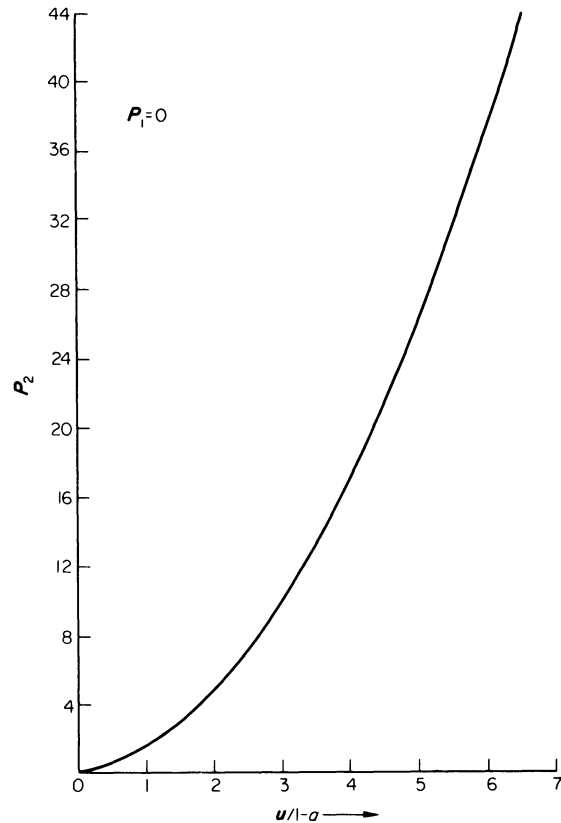


Figure 3. Proposed Dimensionless Hugoniot for metal powders.

The work done and the average temperature rise

The work required to compact the powder can be obtained from the equation

$$\text{The Work Done} = \int_{v_0}^v P dv \tag{8}$$

where v_0 is the powder's initial volume and v is its final volume at pressure P . The projectile of course must also provide the elastic and kinetic energy stored in itself, the powder and the die. If Equation (2) is substituted into Equation (8) then

$$WD = \int_{\rho_0}^{\rho} \frac{\rho_0 v_0}{\rho^2} P d\rho \tag{9}$$

Using Equation (1) to rewrite Equation (9) in terms of the pressure and integrating gives

$$WD = v_0 \frac{(1 - a)(1 + b)}{b} \left[\ln(1 + P) - \frac{P}{1 + P} \right] \tag{10}$$

This work must equal the heat necessary to produce the average temperature rise within the powder $\Delta T^\circ\text{C}$. Therefore, as the process can be assumed to be adiabatic

$$WD = av_0\rho c_p \Delta T \tag{11}$$

where a is the constant in the pressure density relationship and c_p is the specific heat of the solid. Combining Equations (10) and (11) gives

$$\Delta T = \frac{(1-a)}{\rho c_p b} \left[\ln(1+P) - \frac{P}{1+P} \right] \quad (12)$$

However, this equation underestimates the temperature rise, partly because the frictional heating that is not lost to the die walls has not been accounted for. In addition, the pressure density relationship is not followed exactly as was assumed, but is followed by a series of jumps from one equilibrium condition to another. The jumps would be expected to follow the Rayleigh line—which is the straight line connecting the two equilibrium points, but strain rate effects could result in the jump following a path above the Rayleigh line. However, for the present analysis, this error in Equations (10) and (12) may be ignored.

DISCUSSION OF THE RESULTS

Constitutive equation

In order to test the validity of using Equation (1), to describe the behaviour of the powder during dynamic compaction, the pressure observed for region 3 and when possible region 5 (see figure 2), was compared with that predicted for the compaction conditions using the constitutive equation as described in Appendix 1. For aluminium, the relief waves from the back of the projectile have little effect upon P_3 and can be ignored. However, for iron this is not the case and, unless the ratio of the powder to projectile length avoids the arrival of the relief waves before P_3 is recorded, they have to be accounted for in the calculation.

In figures 4 and 5 plots are shown for aluminium and iron powder impacted by steel and PVC projectiles. It can be seen that the agreement is reasonably good, despite the maximum discrepancy between the predicted and observed pressures being as high as 60 per cent.

Changing the value of b in Equation (1) from that determined from the quasistatic data did not improve the agreement.

Property variations of the compact

Vityaz and Roman² reported that a 'non-rigid', as opposed to a 'rigid', base resulted in appreciable variations in the properties of the compacts. In the present work the Hopkinson bar was used as the support, which was therefore reasonably rigid and long enough to avoid waves reflected from the back end of the bar returning before the compaction was complete. However, the length of the projectile and of the powder was varied. As figure 6 shows, the resultant variation in the hardness is very similar to that reported by Vityaz and Roman. The construction of wave diagrams shows that the variation in properties is caused by relief waves from the free end of the projectile, where the pressure, of course, must always be zero. Basically, three different cases have been identified. The first and also the simplest occurs when the pressure has already attained a value sufficient to cause complete compaction before the

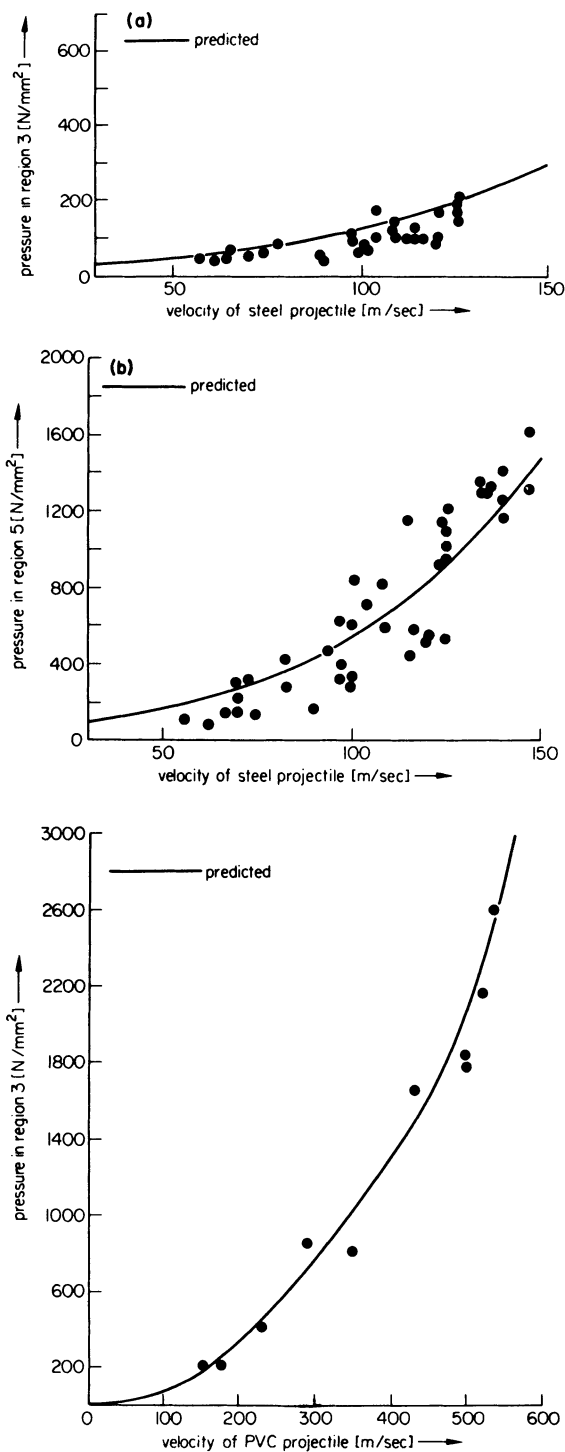


Figure 4. The variation of the pressure created in aluminium powder with the impact velocity. (a) The first reflected pressure (P_3) versus the velocity of a steel projectile. (b) The second reflected pressure (P_5) versus the velocity of a steel projectile. (c) The first reflected pressure (P_3) versus the velocity of a PVC projectile.

arrival of the relief wave. The compact then shows uniform properties. In the second case, the shock wave in the powder does not reach the rigid base before it is overtaken by the relief wave and its magnitude is reduced. This results in a reduction of the compact's properties from the point of impact. The third case occurs when the shock wave in the compact is reflected at the rigid base, but before it can return to the point of impact it is met by the

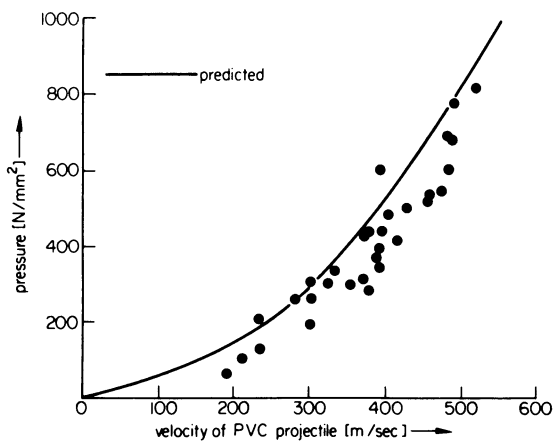
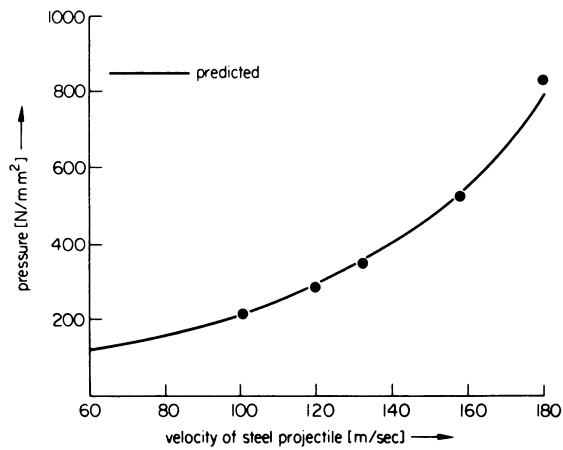


Figure 5. The variation of the first reflected pressure (P_3) created in iron powder with the impact velocity. (a) The reflected pressure (P_3) versus the velocity of a steel projectile. (b) The reflected pressure (P_3) versus the velocity of a PVC projectile.

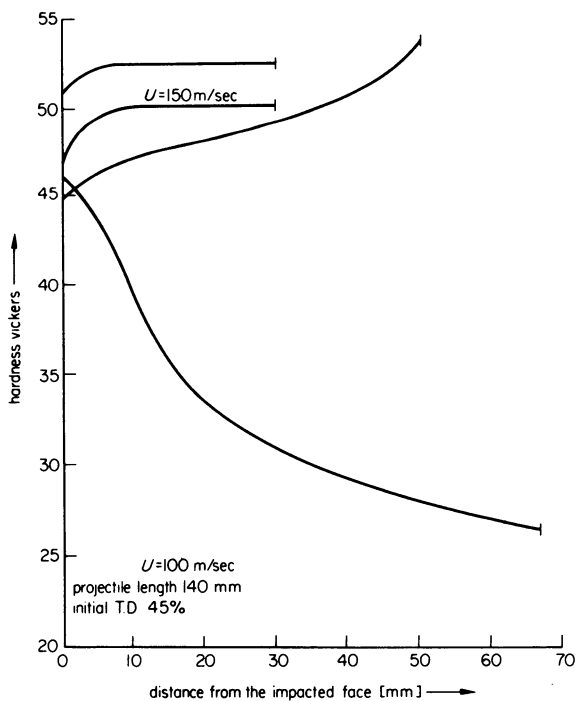


Figure 6. The variation of hardness with the powder's length for varying initial lengths of aluminium powder.

relief wave and its magnitude is diminished. Therefore, the properties of the compact increase with the distance from the point of contact. This is, of course, a simplification of the actual process as the construction of wave diagrams shows.

The constitutive equation did not always predict the correct shock wave velocity. Discrepancies when they did occur were usually around 15 per cent, but in one or two cases, for the aluminium, they were much more appreciable. Despite this, the wave diagrams predicted the variations in pressure and density within the compacts to a reasonable degree of accuracy.

Pressure attenuation

In order to demonstrate the effectiveness of powders as shock absorbers, experiments were conducted in which the impact velocity and projectile length were held constant, while the volume of powder was varied. As well as being measured, the variation of the reflected pressure (P_3) with powder volume was calculated by taking into account the relief waves

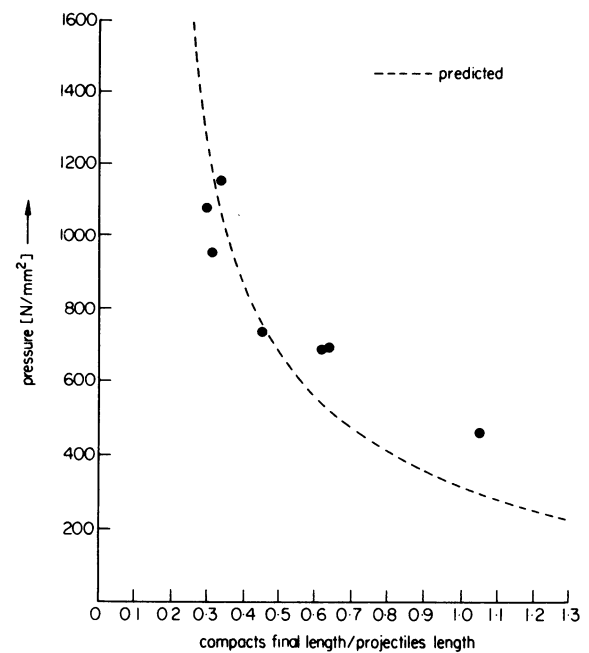
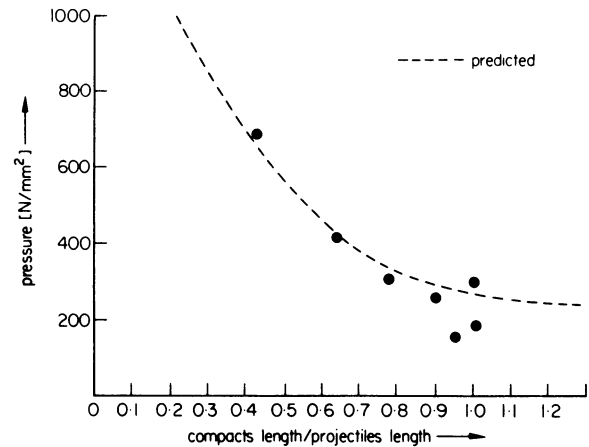


Figure 7. The variation of the maximum reflected pressure with the ratio of the powder to projectile length. (a) Iron powder impacted by PVC projectiles at 300 m/sec. (b) Iron powder impacted by aluminium projectiles at 480 m/sec.

originating from the back, free surface, of the projectile. The resultant wave diagrams show the important factor to be the ratio of the powder to the projectile length. This is plotted in figure 7 against the reflected pressure P_3 . The effectiveness of the powder as a shock absorber can be readily seen—as can the similarity of the calculated and observed pressures.

In order to check the uniqueness of the ratio of the powder to projectile length, experiments were conducted in which this factor was held constant while the powder length was increased. It was found that increasing the length of the powder from 20 to 50 mm had little, if any, effect upon the pressure. This indicates the prime importance of the ratio of the lengths. However, the experimental errors, resulting principally from variations in the projectile velocity, did not permit the conclusion that no dispersion of the shock occurred.

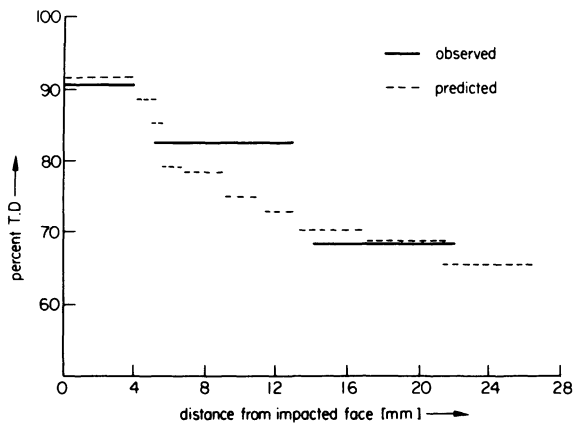


Figure 8. The variation in density along the length of a compact of iron powder impacted by a 25 mm long aluminium projectile at 480 m/sec, the powder volume ensured that no reflected waves were created.

Finally, an experiment was conducted in which the ratio of powder to the projectile length was such that the reflected pressure registered by the bar gauge was negligible. The density variation within this specimen was experimentally determined and compared with that predicted. As figure 8 shows, the difference between the two is small. This supported the assumption that frictional restraints on the powder or the bar gauge from the die walls have little effect on the measured values of P_3 .

Average temperature rise

The thermocouples used were clad in thin stainless steel sheets of thickness x and therefore recorded a temperature $\Delta T_{(xt)}$ at a time t while the true surface temperature of the thermocouple was $\Delta T_{(s)}$. In order to obtain this temperature from that recorded, use was made of Hill's integral profile solution¹⁰

$$\Delta T_{(xt)} = T_{(s)} \left(1 - \frac{x}{S_p} \right)^2$$

where

$$S_p = 1.01t^{1/2}$$

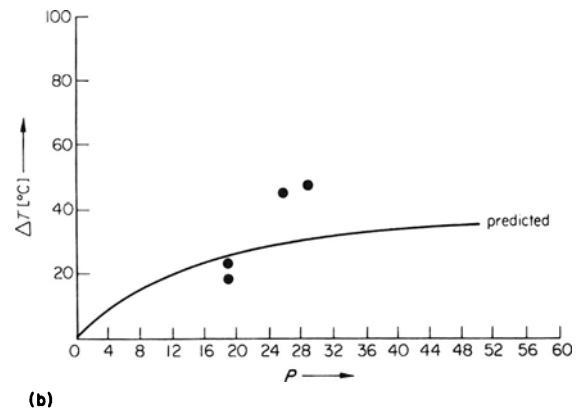
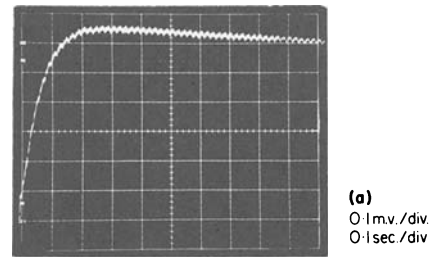


Figure 9. The variation of the temperature rise. (a) Observed variation of the thermocouples temperature with time, for aluminium powder. (b) The variation of the powder's average temperature rise with the maximum dimensionless pressure, for aluminium powder.

A temperature time trace is shown in figure 9. The average temperature rise $\Delta T_{(s)}$ was determined by taking from the traces the time t and $\Delta T_{(xt)}$ at two different points. This gave two values of $\Delta T_{(s)}$, which ideally should be equal. The average of the two temperatures is plotted in figure 9. It can be seen to be of the same order of magnitude and to show approximately the same variation with pressure as the theoretical curve. Several experimental errors could explain the discrepancies; for instance—incorrect positioning of the thermocouple, poor thermal contact between the thermocouple and the powder or an error in the value of x .

Effect of vacuum

Experiments were conducted with and without a vacuum of $\sim 5 \times 10^{-5}$ bars in the guide tube and powder. It was found that the pressure and hence the density were not affected by the presence of the vacuum, but the rise time of the pressure waves was reduced. Of greater importance was the appreciable retardation of the light-weight projectile by the air in the guide tube which had to be compressed before the powder was impacted. It can, however, be concluded that air within the powder can escape during dynamic compaction.

It is believed that the density increase caused by the first shock wave creates a displacement of air which moves towards the low pressure region in front of the shock wave. This air finally arrives at the support die where it may remain as a thin layer between the powder and the die, or, if the sealing permits—as it usually does—the air may escape into the atmosphere. The same sequence of events probably occurs with the reflected wave thus resulting, perhaps, in a thin layer of air between the

projectile and the powder, but it is more probable that this air escapes past the projectile.

PROPERTIES OF THE COMPACTS

Pressing characteristics

The pressure density curves of quasistatic and dynamically compacted aluminium and iron powder are compared in figure 10. The dynamic data was obtained using steel and PVC projectiles of various lengths and velocities impacting on varying powder volumes. Compacts with properties which varied along their length were excluded from the figure. However, there can be little doubt that some inhomogeneous compacts were inadvertently included and are responsible for the spread in the dynamic data.

For aluminium, the dynamic data can be seen to be consistently below the quasistatic curve. The dynamic data for iron is above the quasistatic curve at low densities, but at high densities passes below it. Comparison with the dynamic data obtained for iron on the Petro-Forge by Elwakil and Davies³ shows that, despite the different conditions, the two sets of dynamic data are remarkably similar.

A possible explanation of the relative strain rate insensitivity of the pressure-density relationship is that even though the impact velocities range from 50 to 500 m/sec or, as in the case of Elwakil and Davies, are around 10 m/sec, the density still increases by a series of jumps.

Deformation, therefore, consists of a series of jumps across which the strain rate is very high and these jumps are separated by relatively long periods of time. This is in direct contrast with the more conventional ideal of deformation occurring in one continuous process where the strain rate is proportionally lower.

The divergence shown in figure 10 between the quasistatic and dynamic compacts might be thought to invalidate the use of the quasistatic curve as the constitutive equation. However, for aluminium in particular, the dynamic curve does not give agreement with the measured pressures. The differences shown in figure 10 occur therefore, after the shock wave has passed.

The green strength of the compacts

The green tensile strength of the quasistatic and dynamically produced compacts was determined by diametral compression tests. For the aluminium the dynamic compacts produced using PVC projectiles showed an increase in strength with compaction velocity up to 400 m/sec. They were then 75 per cent stronger than the quasistatic compacts, but subsequently showed a decrease in strength with increasing compaction velocity. With steel projectiles the strength of the compacts increased with the impact velocity, to give a 50 per cent increase at the maximum velocity used.

For iron, the situation was complicated because it was necessary to ensure that compacts of similar density were compared. However, it appears that the increase in strength is less than with aluminium, as the increase was only 40 per cent.

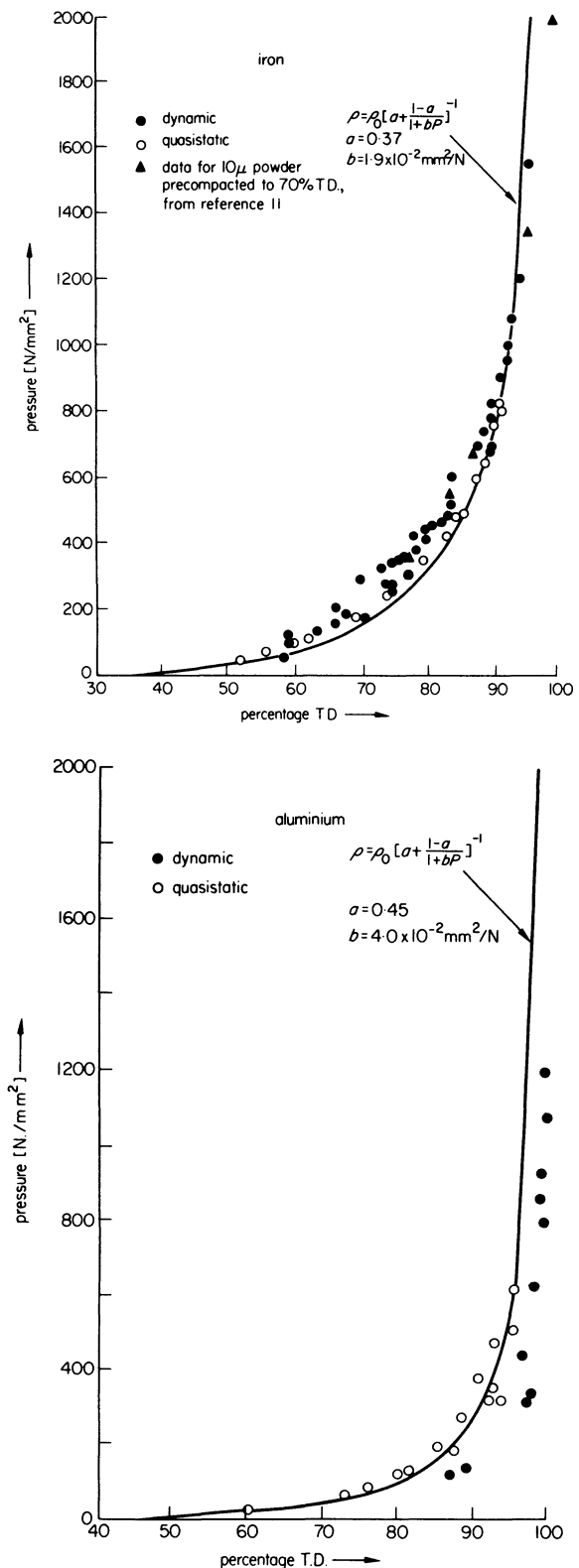


Figure 10. Comparison of the quasistatic and the dynamic pressure-final density relationships. (a) For iron. (b) For aluminium.

The complexity of the dependencies is further illustrated by the fact that in addition to varying with the material, the green strength of the compacts was also found to vary with the volume of powder and length of projectile.

The maximum increase found for the iron was 2, while for the aluminium it was over a factor of 3.

CONCLUSIONS

Dynamic compaction is an extremely complex process in which the problems of quasistatic compaction are increased, not only by variations in strain rate and temperature, but, more importantly, by the presence of shock waves, whose behaviour is of a more complex nature than that of shock waves in solids. Nevertheless, certain conclusions regarding the process can be drawn

- (1) Simple shock wave considerations can be used with reasonable accuracy to predict the dynamic compaction history and the final density of aluminium and iron powders for impact speeds up to 500 m/sec.
- (2) For aluminium and iron powder, the static pressure density relationship can be used as a good approximation of the constitutive relation up to impact velocities of at least 500 m/sec.
- (3) As previously reported, dynamic compaction does increase the tensile strength of the green compacts. The actual increase apparently depends upon the material compacted, the impact velocity of the projectile and the ratio of the projectile length to the length of the initial compact.
- (4) The density of iron after dynamic compaction shows the same relationship with the quasistatic data as the dynamic densities found by Elwakil and Davies using the Petro-Forge. That is for densities higher than approximately 90 per cent lower pressures are required for dynamic than for quasistatic compaction. For aluminium the density after dynamic compaction is always higher than the density after quasistatic compaction at the same pressure.
- (5) Compacts produced by dynamic compaction have homogeneous properties even when their length is relatively long in comparison with their diameter. Inhomogeneity occurs when the piston length is too short with respect to the powder length.

ACKNOWLEDGMENTS

In the course of this work valuable assistance was given by many of the members of Institut CERAC to whom gratitude is expressed. In particular, thanks are due to Dr B. Lemcke and Dr C. Young for their invaluable assistance and encouragement.

APPENDIX 1

Impact on the powder

Consider a projectile travelling at a velocity u_1 impacting upon powder so that the powder has a velocity u_2 and a pressure P_2 . The projectile must also have a velocity u_2 and P_2 , otherwise equilibrium would not occur across the projectile powder interface.

If P_2 is not above the elastic limit for the projectile, then

$$P_2 = \rho_p c_p (u_1 - u_2) \quad (A1)$$

For the powder when $P_1 = 0$ Equation (6) can be used to give

$$\frac{u_2}{c_0} = \frac{(1-a)P_2}{(1+P_2)^{1/2}} \quad (A2)$$

Using Equation (A1) to rewrite Equation (A2) in terms of pressure results in an equation which can be solved to give P_2 .

Determination of the pressure after the reflection of a shock

Consider the case of a shock moving into region one from region two. Across this shock the constitutive Equations (2) and (3) must be obeyed. After being reflected at an interface, the shock moves back into region 2 leaving behind it a new region 3. As before, regions two and three must obey Equations (2) and (3).

so

$$u_2 = \frac{(1-a)(P_2 - P_1)}{\{(1+P_1)(1+P_2)\}^{1/2}} \quad (A3)$$

and

$$u_2 = \frac{(1-a)(P_3 - P_2)}{\{(1+P_2)(1+P_3)\}^{1/2}} \quad (A4)$$

equating these, expanding and assuming that P_1 very much less than 1 or P_2 gives the interesting expression

$$P_3 = P_2(2 + P_2) \quad (A5)$$

However, this expression is only valid at low pressures for the support material has been assumed to remain stationary. For the impedance of the support material to be taken into account it simply requires Equations (A1) and (A2) be rewritten in full for the support

$$P_3 - P_1 = \rho_s c_s (u_3 - u_1) \quad (A6)$$

for the powder

$$u_3 - u_2 = \frac{c_0(1-a)(P_3 - P_2)}{\{(1+P_2)(1+P_3)\}^{1/2}} \quad (A7)$$

As before, Equation (A6) is substituted into Equation (A7) in order to remove one of the unknowns, u_3 and the equation is then solved for P_3 .

REFERENCES

1. R. Davies and E. R. Austin. *Developments in high speed metal forming*. The Machinery Publishing Co Ltd. Brighton, 1970.
2. P. A. Vityaz and O. V. Roman. Impulse compacting of powder materials. *The 13th Int. MTDR Conf. Birmingham*. Macmillan Press, London, 1972.
3. S. Elwakil and R. Davies. (1972) *ibid*.
4. H. A. Al-Qureshi. High speed punching of thick plates using visco-elastic dies. *The 4th Int. Conference of the Center of High Energy Forming*, University of Denver, Colorado, 1973.
5. R. K. Linde and D. N. Schmidt. Shock propagation in non-reactive porous solids. *JAP* (1966) 37, no. 8, 3259.

6. G. R. Fowles. Experimental techniques and instrumentation. Chapter 8 of *Dynamic Response of Materials to Intense Impulsive Loading*. Editors C. Chou and A. Hopkins. Air Force Materials Laboratory, Wright Patterson A.F.B., Ohio, 1973.
7. L. Davison. Shock wave structure in porous solids. *JAP* (1971) **42**, no. 13, 5503.
8. W. Herrman. Constitutive equation for the dynamic compaction of ductile porous material. *JAP* (1969) **40**, no. 6, 2490.
9. W. Johnson. *Impact Strength of Materials*. Edward Arnold Ltd. London, 1972.
10. T. Sheppard and D. Raybould. On load and temperature rise during axisymmetric extrusion *JIM* (1973) **101**, 33.
11. R. K. Linde, L. Seaman and D. N. Schmidt. Shock response of porous copper, iron, tungsten & polyurethane. *JAP* (1972) **43**, no. 8, 3367.

THE HOT EXTRUSION OF METAL POWDER PREFORMS

by

M. NEGM* and R. DAVIES*

SUMMARY

This paper contains a report of an investigation into the possibilities and limitations of hot extruding preforms which have been pressed, sintered and extruded under various conditions from iron powder. The variables involved in this work were initial density, sintering and extruding temperature, sintering time and area reduction ratio. Forward and backward extrusions were performed at slow and high speed. The effect of speed on the extrudability of the powder preforms has been studied. Density and hardness distributions of the hot extruded products were also investigated in detail.

INTRODUCTION

The forging of powder preforms has developed considerably during the past few years. As a result of the development of powder forging techniques, it is possible to achieve near full densities in hot forged parts. Marx and Davies¹ carried out some work on the hot forging properties of powder preforms under conditions of simple upsetting and Guest, Negm and Davies² studied the hot deformation and densification of powder preforms under conditions of recompaction and closed die upsetting. Some papers^{4,5} were also published on the possibility of cold extruding preforms from sintered iron powder, but little information is available on the hot extrusion of powder preforms. The aim of the present work is to extend the range of powder forging techniques to the forming of the preforms by hot extrusion. In this paper, the authors are mainly concerned with the limitations of the process and with the properties of hot extruded iron powder preforms, pressed, sintered and extruded under various conditions.

EQUIPMENT

Compaction and extrusion machines

The powder preforms were pressed on a 600 kN hydraulic press. High speed forward and backward extrusions were carried out on a Mk IIA Petro-Forge machine³, having a velocity range, at impact, of 4-5 m/sec. Slow speed forward and backward extrusions were produced on a 2000 KN eccentric press. The velocity at the beginning of extrusion was 0.2-0.25 m/sec.

Tooling

The compaction die for producing cylindrical powder preforms is shown in figure 1(a). The nominal diameter of container bore and punch was 25.4 mm. Both die and punch were made from Carrs 69 S steel, hardened to 58-60 Rc.

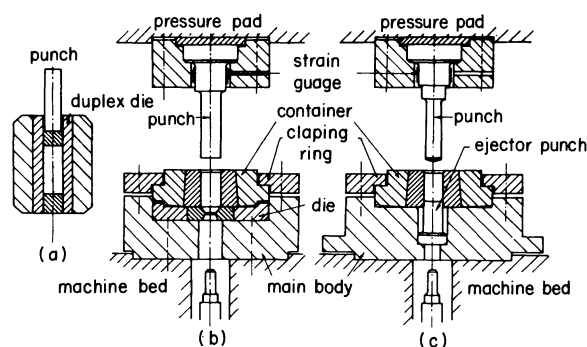


Figure 1. General layout of tooling: (a) compaction die; (b) forward extrusion die and (c) backward extrusion die.

The general tooling layout for forward extrusion is shown in figure 1(b). The container and die were located in the main body and clamped by a ring. An ejector was operated by a hydraulic cylinder mounted under the machine bed. A duplex type container with a nominal internal diameter of 25.75 mm was used. The extrusion punch, backed by a pressure pad, was clamped by the holder fixed to the moving platen. The dies were provided with a 90° die entry angle with reduction ratios, R , of 2, 3, 4 and 7. Dies and punch were made from Jessop-Saville Ultimar 100 steel, hardened to 46-48 Rc.

* Department of Mechanical Engineering, University of Birmingham

For backward extrusion, the container of nominal diameter 25.75 mm, was located in the main body and clamped by a ring—see figure 1(c). The ejector punch, guided by the main body, was operated in the same manner as described above. The backward extrusion punch was clamped to the moving platen in an identical manner to the forward extrusion punch. The backward extrusion punches gave reduction ratios, R , of 1.72, 1.96, 2.3, and 3 and 5.

Sintering furnace

All sintering operations were carried out in a 100 mm diameter tube furnace having a 230 mm long uniform heating zone. The atmosphere was cracked ammonia.

Materials used

Hoganas iron powder NC100 (see Table 1) was used throughout the tests. This was blended with 1 per cent Acrawax as a lubricant.

TABLE 1 Details of Hoganas iron powder NC100

Chemical analysis	% (approx)
iron	98 +
carbon	0.1
oxygen (loss of weight in hydrogen)	0.6
silica	0.3
sulphur	0.015
phosphorous	0.015
Physical properties	%
Sieve analysis (BS sieves)	
+ 100	1% max.
– 100 + 150	15 25%
– 150 + 200	20 30%
– 200 + 240	5 15%
– 240 + 350	15 30%
– 350	15 30%
apparent density 2.4 g/cm^3	
theoretical density 7.84 g/cm^3	

Lubricants

Copaslip, a mixture of copper and lead particles in a base of bentone grease was used for the extrusion tests.

Instrumentation

The extrusion load was measured by mounting strain gauges directly to the extrusion punches, and connecting to an a.c. bridge system. The output from the strain gauge bridge was fed to the vertical display of an oscilloscope. A capacitance displacement transducer was used to follow the movement of the upper punch. The output from the displacement transducer was converted to a voltage by a 2 mc/s oscillator and F.M. preamplifier. This was fed to the horizontal display of the oscilloscope and direct load-displacement characteristics were therefore obtained.

The impact speed of the upper platen was determined using a magnetic induction type velocity transducer. The output from the velocity transducer was fed to the vertical display of the oscilloscope.

EXPERIMENTAL RESULTS AND DISCUSSION

The tests described were performed to determine the characteristics and properties of hot extruded powder preforms. A considerable number of variables is involved in such an analysis: the type of powder, initial preform density, sintering temperature, sintering time, sintering atmosphere, extrusion temperature, lubricant (compaction and extrusion), type of extrusion (forward or backward), reduction ratios R , and impact velocity, all play their part. Some of these variables were readily eliminated by selecting type NC100 iron powder, a sintering gas of cracked ammonia, die wall lubricant of Acrawax for compaction and Copaslip for hot extrusion. For the sake of simplicity, it was decided to sinter and extrude at the same temperature.

Effect of sintering and extruding temperature on extrusion pressure

All other variables were held constant. A sintering time of 25 min was chosen since it was considered that this gave sufficient time for the billet to react to the furnace temperature. The preforms had an initial density of 6.85 g/cm^3 (87% of theoretical). All the preforms were extruded at both high and slow speeds to a reduction ratios R of 7 and 3 for forward and backward extrusion, respectively. Figure 2(a) shows the variation of maximum extrusion pressure for forward extrusion over the temperature range and indicates local minima of 880 and 920°C, and a local maxima of 970 and 1030° for high speed and slow speed, respectively. The decrease in extrusion

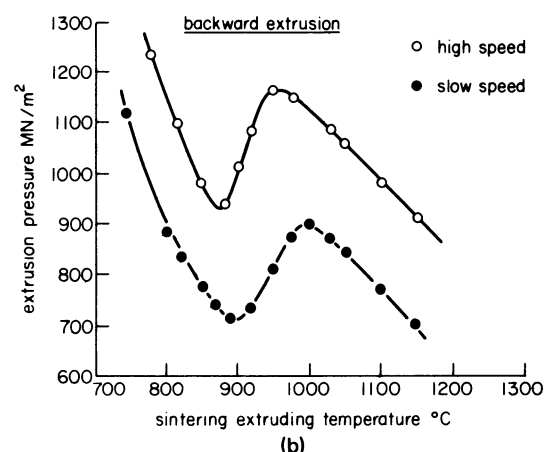
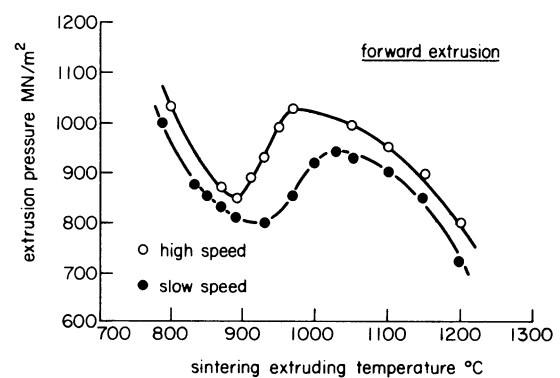


Figure 2. Variation of extrusion pressure with sintering and extruding temperature for (a) forward extrusion and (b) backward extrusion.

pressure from 750 to 880°C is readily explained by the fact that the warmer material is easier to extrude than the colder one. The rise in extrusion pressure from 880 to 970°C is due to the formation of a two-phase structure of $\alpha + \gamma$. At 970°C, an entirely austenitic structure is obtained which requires progressively less extrusion pressure as the temperature is increased to 1200°C. All the preforms extruded below the local minimum temperatures were cracked, the severity of cracking decreasing from 750°C upwards. The cracks disappeared near the local maxima. These results give good agreement with previous work by Marx and Davies¹ using MH100 iron powder. The difference between local minima and maxima for high and slow speeds, also shown in figure 2(a), are due to heat generation during the extrusion process. The values of extrusion pressure for high speed were higher than those for slow speed, probably because pure irons are highly strain-rate sensitive. The difference in forward extrusion—see figure 2(a)—was about 7 per cent, but in backward extrusion—see figure 2(b)—it was about 25 per cent.

Effect of sintering time on extrusion pressure

The sintering time was varied from 6 to 35 min for the same extrusion temperature of 1100°C. All other variables were the same as in the previous series of tests. Figures 3(a) and 3(b) show the maximum extrusion pressure for forward and backward extrusion plotted against sintering time in the furnace, for both high and slow speeds. The extrusion pressure

reduced with sintering time up to 15 min, after which time very little change in extrusion pressure took place. For backward extrusion, see figure 3(b), the reductions of extrusion pressure up to 15 min were small when compared with forward extrusion—this is possibly due to shattering of the backward extruded parts into many pieces at low sintering times. The difference of extrusion pressure between high and slow speeds was 7 and 25 per cent for forward and backward extrusion, respectively. This difference is due to the high strain rate sensitivity of pure irons and to the friction effect between extruded parts and tools. All the extruded preforms were cracked up to sintering time of 15 min, and cracking increased in severity as the sintering time was decreased, the grain size of the extruded parts increased with sintering time up to 15 min after which time very little change of structure was evident. Thus, the sintering time of 15 min was considered to be sufficient at a temperature of 1100°C.

Effect of initial density and reduction ratio 'R' on extrusion pressure

Preforms of equal mass having initial densities between 5.4 g/cm³ (68 per cent) and 7.2 g/cm³ (91 per cent) were sintered at 1100°C for 15 min and then extruded through various reduction ratios. Forward and backward extrusion was carried out at both high and slow speeds. The results of this series of tests appear in figure 4(a) for forward extrusion, where the maximum extrusion pressure was plotted against the initial density for various reduction ratios at high and slow speeds. Similarly, the results relating to backward extrusion are shown in figure 4(b).

In both cases of forward and backward extrusion, the results show that the extrusion pressure

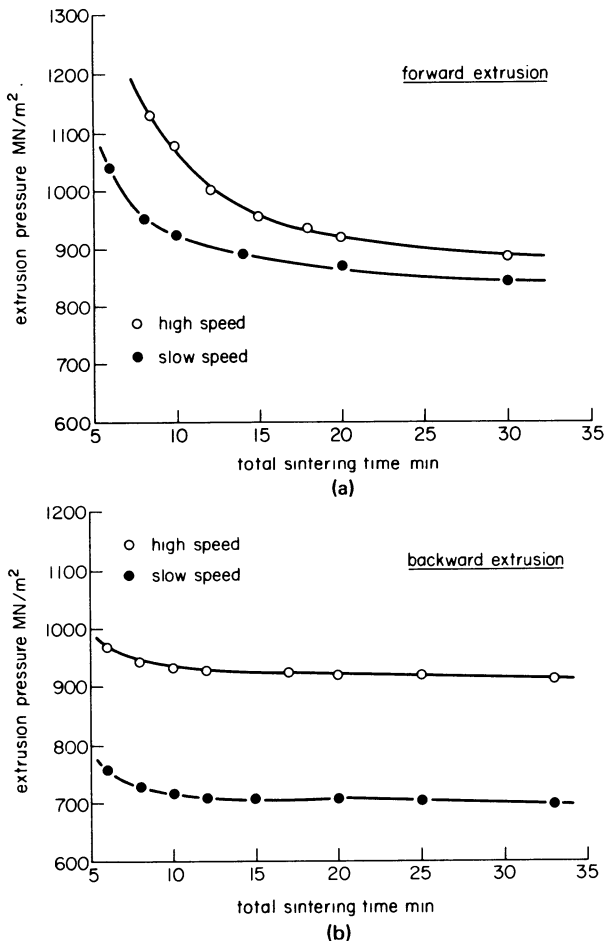


Figure 3. Variation of extrusion pressure with sintering time for (a) forward extrusion and (b) backward extrusion.

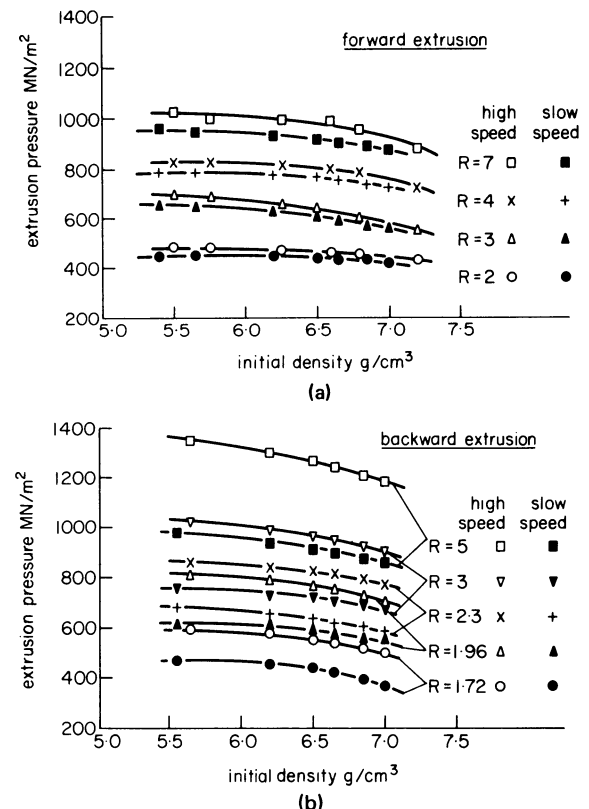


Figure 4. Variation of extrusion pressure with initial density for (a) forward extrusion and (b) backward extrusion.

decreases as the initial density increases. To explain this phenomenon, it must be borne in mind that higher initial density means shorter initial length of preform for the same mass. Hence, the frictional force between the preform and container for higher initial densities was less than that for lower initial densities. Moreover, the extrusion pressure at high speed was higher compared with that at slow speed—this can be attributed to the strain rate effect. However, the speed effect was more significant in backward extrusion than in forward extrusion. For example, in backward extrusion, the extrusion pressure at high speed was 25 per cent in excess of that at slow speed, but in forward extrusion, the difference was only 7 per cent. This phenomenon has also been observed by Singh and Davies⁵. This may suggest excessive friction occurring at high speed in backward extrusion. The excessive friction can be noticed from the load-displacement diagrams shown in figures 6(a) and 6(b). In the case of forward extrusion shown in figure 6(a), the extrusion pressure falls during the extrusion process for both high and slow speeds. However, in the case of backward extrusion, shown in figure 6(b), the extrusion pressure falls for slow speed and remains almost constant in the case of high speed. Furthermore, the occurrence of this excessive friction was also confirmed by the sticking of the extruded part of the punch, resulting in necking of the specimen, as shown in figure 7(a). However, figure 7(b) shows a specimen extruded at slow speed without the occurrence of necking. The excessive friction in backward extrusion at high speed can possibly be caused by the residual stresses tending to contract the extruded part on to the punch. Figures 5(a) and 5(b) show the extrusion pressure plotted against the reduction ratios for forward and backward extrusion at slow and high speeds, respectively. The extrusion pressure increased with the reduction ratio.

Effect of initial density on average final density and limits of cracking

Figures 8(a) and 8(b) show the average final density plotted against the initial density for different reduction ratios for both forward and backward extrusion. The average final density increased with the initial density and the value of average density increased with increasing extrusion ratio, R , for the same initial density. The effect of impact speed was to cause an increase in the average final density under the same conditions, but it was slightly greater in backward extrusion than in forward extrusion. The increase for an initial density of 5.75 g/cm³ was about 33 per cent, and about 9 per cent for an initial density of 7 g/cm³. Figures 9(a) and 9(b) show the average final density for various reduction ratios, for an initial density ranging from 5.75 to 7.2 g/cm³. The cracking limits for these conditions are also shown. All the preforms having an average final density above the density indicated by the 'cracked line' were undamaged. Below that line, the extrusions were cracked and the severity of cracking increased at lower average densities. The 'cracked line' for forward extrusion, shown in figure 9(a), occurred at an average final density of 7.70 g/cm³ for slow and high speed, respectively. For backward extrusion,

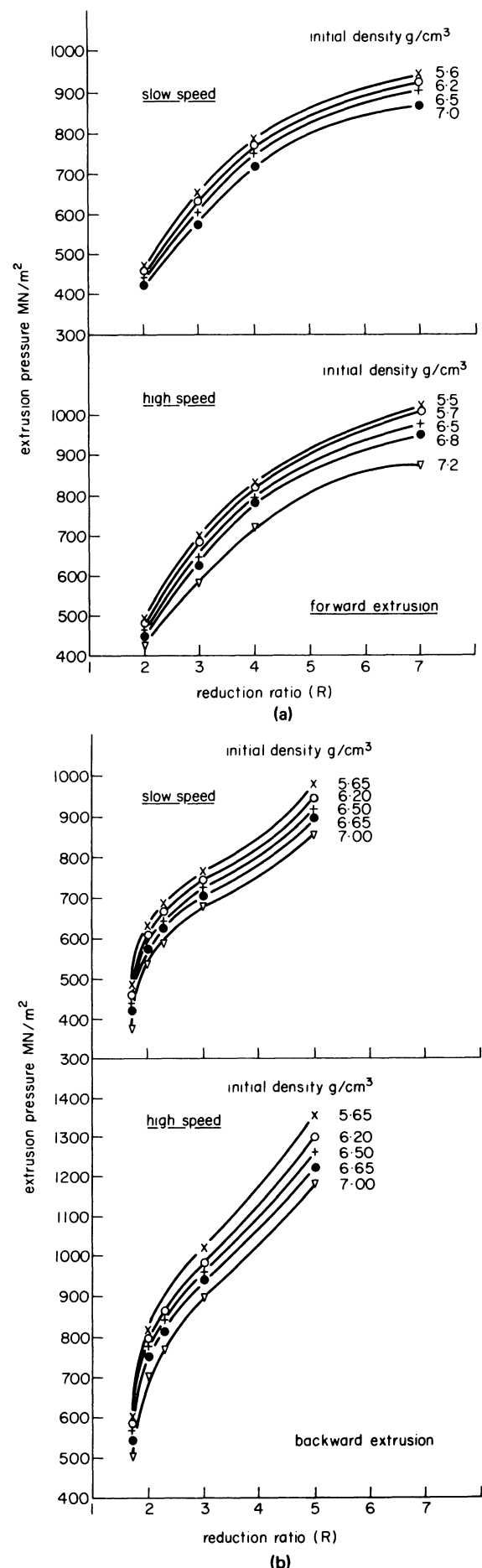


Figure 5. Variation of extrusion pressure with extrusion ratio for (a) forward extrusion and (b) backward extrusion.

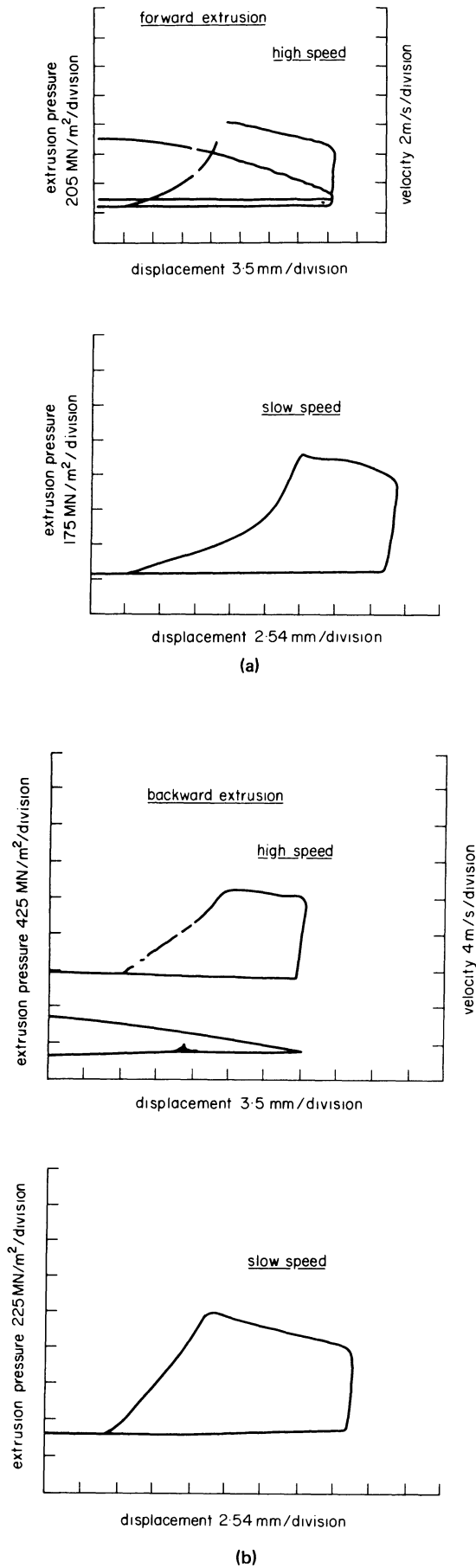


Figure 6. Typical oscilloscope record showing the variation of extrusion pressure and velocity with deformation at high and slow speeds for (a) forward extrusion and (b) backward extrusion.

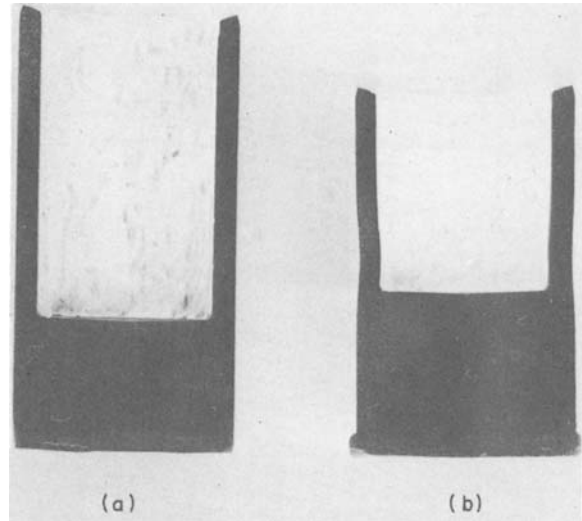


Figure 7. Backward extrusion of reduction ratio 3 for (a) high speed and (b) slow speed.

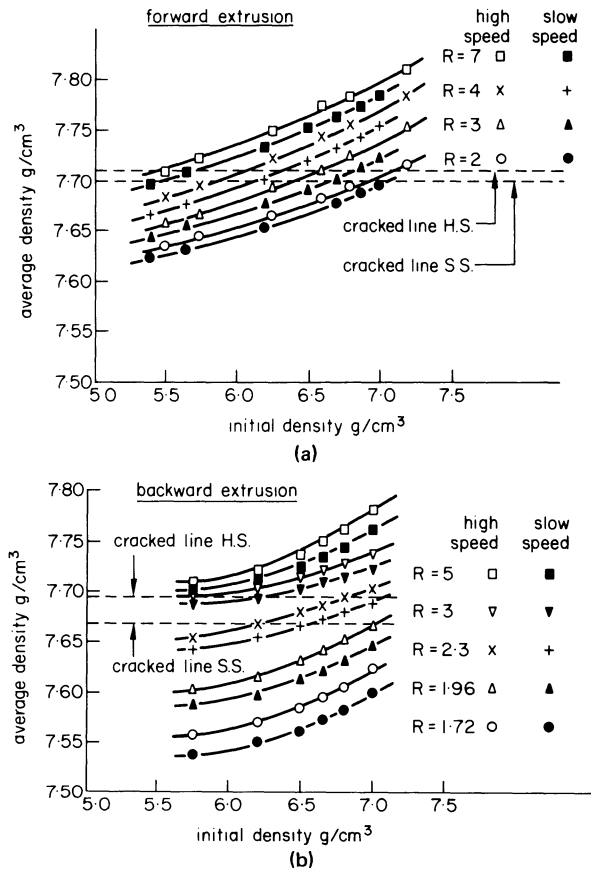


Figure 8. Variation of average final density with initial density, and limits of cracking for (a) forward extrusion and (b) backward extrusion.

illustrated in figure 9(b), the value of average final density at the "cracked line" was 7.67 g/cm^3 for slow speed.

In order to avoid this type of cracking of the extrusions, the preforms should be sintered at 1100°C for at least 15 min, and the combination of initial density and reduction ratio should be above the limits given by the cracked line, as shown in figures 8 and 9.

The density distribution of extruded products

The density distribution shown in figures 10(a) and 10(b) was measured by weighing the various parts of

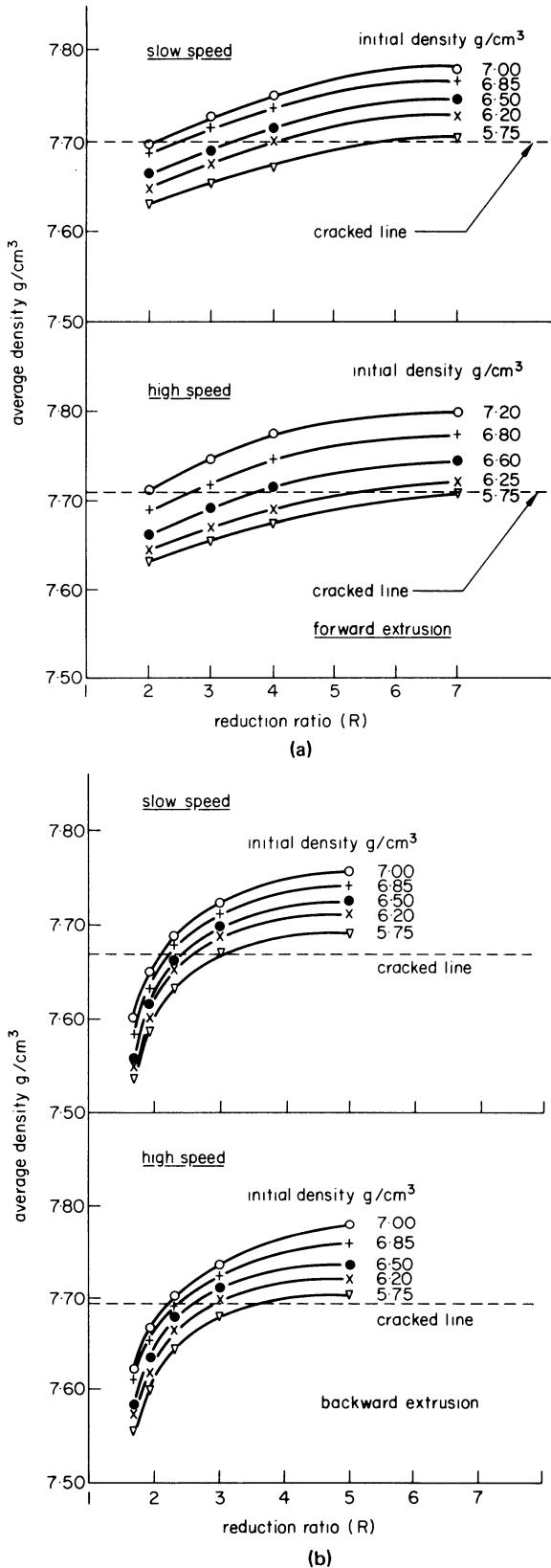


Figure 9. Variation of average final density with extrusion ratio and limits of cracking for (a) forward extrusion and (b) backward extrusion.

the extrusions in air and water. The first portion to be extruded has a comparatively low density, but other extruded parts have a constant high density. The undeformed part has a slightly lower density than the extruded parts. The density values for high speed extrusion were slightly greater than at slow

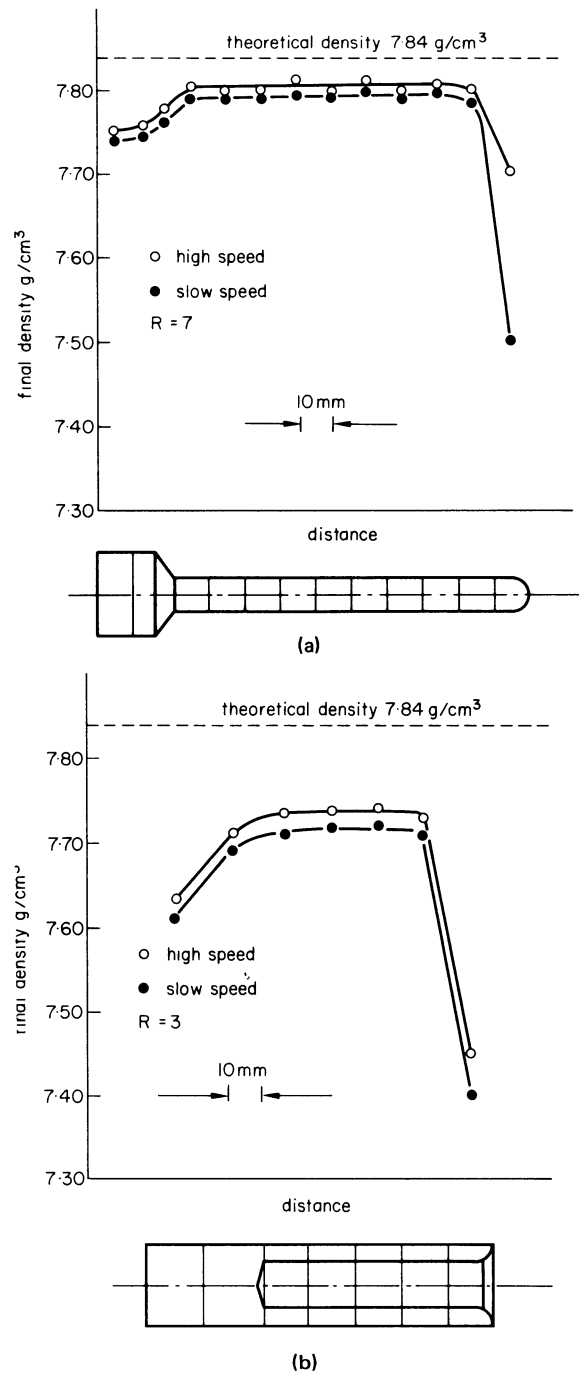


Figure 10. Density distribution for (a) forward extrusion and (b) backward extrusion.

speed for both forward and backward extrusion. This is due to the strain rate effect and higher extrusion pressures required. Similar results were obtained by Nakagawa et al.⁴ for the cold extrusion of sintered iron powder preforms.

Material flow and Vickers' hardness distribution

Four preforms of 4605 alloy steel of initial density 6.85 g/cm³ were sintered at 1150°C for 20 min and extruded. Two preforms were forward extruded through a reduction ratio of 7 at high and slow speeds and two preforms backward extruded through a reduction ratio of 3 at high and slow speeds. The results of hardness traverses across sectioned extrusions are shown in contour form in figures 11(a), (b),

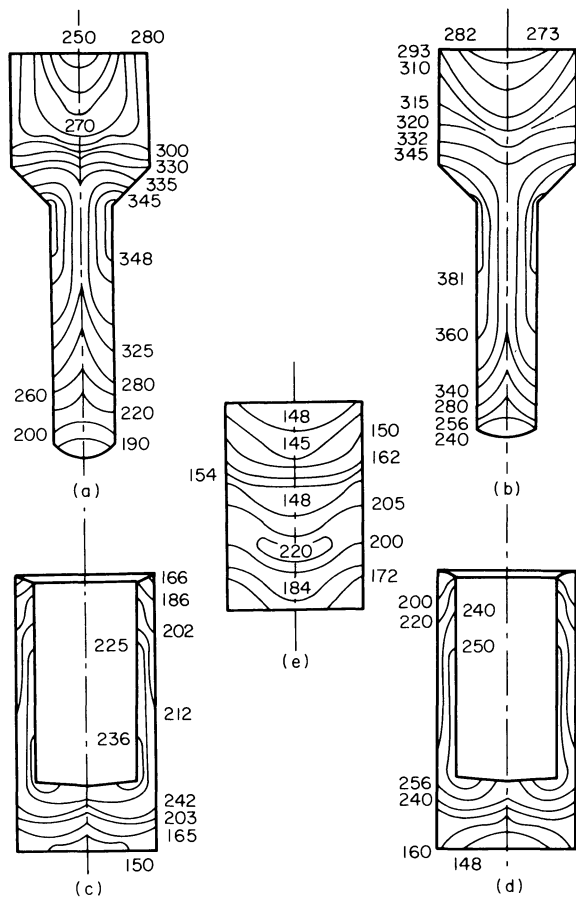


Figure 11. Hardness distributions of 460 S alloy steel for (a) forward extrusion at slow speed; (b) forward extrusion at high speed; (c) backward extrusion at slow speed; (d) backward extrusion at high speed and (e) unextruded sintered preform for comparison.

(c) and (d). The hardness contours for an unextruded preform are also shown for comparison in figure 11(e). The general tendency, particularly for forward extrusion, is a more uniform distribution of hardness, and a somewhat greater general level of hardness in the extrusions produced at high speed. In forward extrusion the more uniform hardness distribution was probably due to reduced frictional resistance at the workpiece tool interfaces under high speed conditions. The hardness distribution shows the same general trend as the density distribution.

CONCLUSIONS

- (1) The extrusion pressures were found to be higher at high speed than at slow speed. The

difference in forward extrusion was approximately 7 per cent and in backward extrusion approximately 25 per cent.

- (2) The average final density in high speed was more than that of slow speed for both forward and backward extrusion.
- (3) As far as cracking is concerned: (a) the optimum sintering and extruding temperature was selected at 1100°C and the optimum sintering time at 15 min; (b) to avoid cracking in both forward and backward extrusion, it was observed that there was a lower limit of the average final density of about 7.7 g/cm^3 ; below that limit, it would be difficult to extrude successfully. This final density can be achieved by a suitable combination of initial density and reduction ratio.
- (4) The portion extruded at the earlier stage has a comparatively low density, but other portions have a constant higher density.
- (5) In high-speed extrusion, the hardness distribution was more uniform and a somewhat greater general level of hardness was obtained when compared with slow-speed extrusion.

REFERENCES

1. J. B. Marx, R. Davis and T. L. Guest. Some considerations of the hot forging of powder preforms. *Proc. of the 11th Internat. M.T.D.R. Conference*, Birmingham, Vol. B, 729, (1970).
2. T. L. Guest, M. Negm and R. Davies. Metal flow and densification in the die forging of porous preforms. *Powder Met.* (1973), 314.
3. L. T. Chang and S. A. Tobias. Performance characteristics of Petro-Forge Mk I and Mk II machines. *Proc. of the 9th Internat. M.T.D.R. Conference*, Birmingham, 1968.
4. T. Nakagawa *et al.* Cold forging of ferrous M/P billet, *27th Symp. Japan Soc. Powder and Powder Metallurgy*, Tokyo, 1971.
5. A. Singh and R. Davies. Preliminary investigations of the cold extrusion of powder preforms. *Proc. of the 13th Internat. M.T.D.R. Conference*, Birmingham, 1972.
6. A. D. Sheikh, T. A. Dean, M. K. Das and S. A. Tobias. The effect of impact speed and lubricant in hot forging. *Proc. of the 13th Internat. M.T.D.R. Conference*, Birmingham, 1972.

PLASTICITY OF SINTERED IRON-POWDER COMPACTS

by

SHERIF D. EI WAKIL*

SUMMARY

The plastic behaviour of sintered NC 100-24 sponge iron powder compacts has been investigated. Cylindrical sintered billets were subjected to uniaxial homogeneous compression tests at room temperature. From the results obtained, the effects of the initial as-sintered density on the cold formability and the work-hardening characteristics of the compacts are given.

INTRODUCTION

The powder metallurgy process has recently had widespread industrial application. This is mainly due to the fact that it eliminates machining operations and offers more efficient utilization of materials and that components can be produced at high rates. However, this process has some disadvantages and limitations, such as, the high capital cost of pressing machines and tools. A further disadvantage is the relatively poor mechanical properties of sintered compacts due to porosity. For the latter reason, recent developments in powder metallurgy are aimed at increasing both the density achieved and the strength of the sintered components. The final aim is to produce large parts able to withstand high stress which can now only be produced by conventional metal forming processes.

The hot forging and the cold forming of powder preforms are the main processes being developed for producing parts with high density and greater strength. It has been reported¹ that approximately the full density was achieved in hot forged parts, with properties comparable to those obtained in parts made from solid material. However, there are problems connected with this process such as the excessive die wear, the difficulty of obtaining good dimensional accuracy and the oxidation of the preforms, as they are exposed to air in their hot state. Many of these problems are eliminated by using the cold forming process, and the preheating of sintered billets is unnecessary. A further advantage of the cold forming process is the improved strength of the parts due to work-hardening.

For successful utilization of the process of cold forming of powder preforms, it is necessary to gain knowledge of the densification and deformation

characteristics of the preform material at room temperature. These characteristics are usually determined from homogeneous compression tests. The results obtained can then be used as the basis of a plasticity theory for sintered porous materials. Another approach which was followed by some research workers²⁻⁴, involves investigating the behaviour of sintered compacts in actual cold forming processes. In this case, empirical equations and experimental graphs are obtained and can be used for practical purposes.

Some interesting work has been carried out by Antes⁵ who studied the deformation and densification characteristics of iron and steel powder preforms under three types of loading, namely uniaxial compression, plane strain and repressing, at room temperature. It was found that the rate of densification increased as the degree of restraint increased from uniaxial compression (no lateral constraint), to plane strain compression (restraint in one of the lateral directions) and finally to repressing (total constraint). However, the differences in densification rates between the uniaxial compression and plane strain compression decreased with increasing preform density; at a preform density of 7.2 g/cm³ the rates of densification were approximately the same. In addition, a linear relationship existed between preform density and final density for a given amount of straining, for each of the above-mentioned types of loading. It was also found that increasing the carbon content in the range of 0.06 to 0.41 per cent had virtually no effect on the rate of densification but increased the load required to achieve a given density.

Another investigation has been carried out by Kuhn and Downey⁶ who studied the basic densification and deformation behaviour of sintered sponge

* Department of Mechanical Engineering, Al-Azhar University, Cairo, Egypt.

iron powder compacts under simple uniaxial homogeneous compression tests at room temperature. The resulting experimental data relating yield stress and Poisson's ratio to preform density formed the basis of a plasticity theory for sintered powder compacts. In addition, a yield criterion for sintered powder compacts was proposed which accurately predicted the forming stresses in repressing and in plane strain compression.

Three separate, but interrelated studies have been undertaken by Kuhn et al⁷. They included a study of the deformation of sintered sponge iron powder preforms and its relation to porosity and contaminant level, a model study of the impurity-base material system aimed at determining the mechanism of densification and a transmission electron microscopy study as a further aid to the study of densification and fracture. The basic deformation and densification behaviour of the sintered compacts, particularly their work-hardening and lateral spread characteristics, was determined. In the second study, preliminary tests showed that a model of a void space surrounded by a matrix material accurately simulated the macroscopic deformation behaviour of a powder compact. The electron micrographs, taken before and after compressing the compacts, indicated that the density of dislocation increased after compression.

Nakagawa et al² extruded reduced iron powder compacts at room temperature. The cold forgeability of the sintered compacts and the effects of the reduction ratio were discussed. The density of the forward and backward extrusions with reduction ratios of 9 and 2.75 respectively was found to vary from 96 to 99 per cent of theoretical density when the initial density of the preform was in the range of 83 to 86 per cent.

Another investigation has been carried out by Nakagawa et al³ who studied the cold forgeability of sintered iron powder preforms. In order to investigate the restriction of the shape in the cold forming process, seven kinds of standard forging were studied, namely upsetting, upsetting with forward extrusion, forward extrusion, backward extrusion and two kinds of complex extrusion and tube extrusion. The effects of the reduction ratio were investigated in the case of forward and backward extrusion. It was found that low density preforms exhibited a tendency to crack. This was attributed to the existence of pores and weak grain boundaries in those preforms. Cracking took place also in almost all preforms extruded with low reduction ratios. Another important point in Nakagawa's work was to do with the forging load. It was found that the forging load of sintered iron powder billets was always lower than that of conventional wrought iron billets. The density distribution of the extruded preforms was also investigated. The results indicated that the portion extruded at the earlier stage had a comparatively low density, whereas other portions had a very high density, which was in some cases very close to the theoretical density.

Preliminary experiments were carried out by Singh and Davies⁴. Their work was mainly concerned with the feasibility and limitations of cold extrusion of preforms pressed from iron powders (electrolytic iron

powder. Höganäs sponge iron powder MH 100-24, Höganäs atomized iron powder AHC 100-29 and MP 32 reduced and atomized iron powder). Both forward and backward extrusion were performed at low and high speeds and the effect of speed on the extrudability of these powder compacts was studied. In the forward extrusion at slow and high speeds it was observed that there was a lower limiting reduction ratio below which it would be difficult to extrude successfully. It was also found that the density of the underformed part was always slightly higher than that of the extruded part. Highest densities (about 99 per cent of the theoretical density) were obtained with electrolytic iron powder, whereas the lowest densities (95 per cent of theoretical density) were observed when sponge iron powder was used. In addition, the effect of impurities in the latter powder was to make the flow less uniform.

The present work aims at extending the range of results obtained by previous workers and obtaining more detailed information regarding the cold formability and work-hardening characteristics of sintered compacts. For simplicity, the deformation and densification characteristics of the sintered billets were determined from uniaxial homogeneous compression tests. The results obtained can be used in developing plasticity equations for sintered porous materials. These equations can in turn be applied to the various cold forming processes.

EXPERIMENTAL PROCEDURE AND MATERIALS USED

Powder used

Höganäs sponge iron powder NC 100-24 was used throughout the tests reported here. This was blended with 0.7 per cent of zinc stearate for lubrication. Table 1 shows the characteristics of this iron powder.

Compaction machine

The compaction process was carried out with a 200 tonf (approximately 2000 KN) hydraulic testing machine, which provided a suitable means for applying and measuring the compaction load. The same machine was used for the homogeneous compression tests.

Tooling

A 20 mm bore diameter die was used for the production of cylindrical compacts. It was made of high-carbon steel and was heat-treated to a hardness of 60-62 Rc.

Sintering furnace

All sintering operations were carried out in a pusher type sintering furnace capable of providing sintering temperatures of up to 1300°C with an accuracy of ±5°C. It consisted of a preheating zone for de-waxing the components, a 50 cm long uniform heating zone and a water cooled jacket which was attached to one end of the furnace and had the same atmosphere. The atmosphere was cracked ammonia (25 per cent Nitrogen and 75 per cent Hydrogen).

TABLE 1 The characteristics of Höganas sponge iron powder NC 100-24

Screen analysis		Chemical composition	Apparent density	Flow rate
+ 65	0%	H ₂ -loss	max 0.3%	+ 0.10 max
- 65 + 100	max 1%	C	max 0.02%	2.4 g/cm ³
- 100 + 325	74-85%	SiO ₂	max 0.20%	- 0.05 32 s/50 g
- 325	15-25%	P	max 0.015%	
		S	max 0.015%	

Preform fabrication

Preliminary experiments were carried out to determine the compaction pressure-green density relationship (shown in figure 1). The load required to achieve

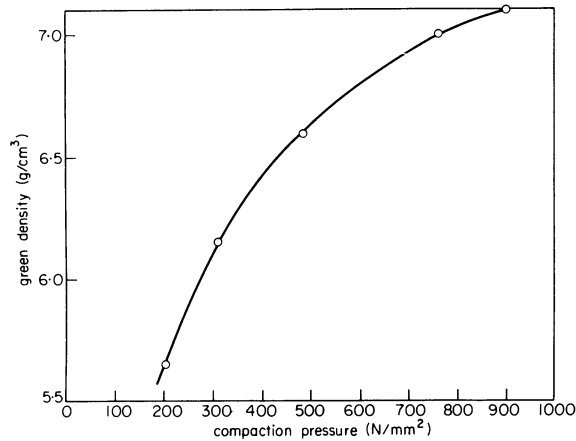


Figure 1. Variation of green density with the compaction pressure for NC 100-24+0.7 per cent zinc stearate.

a given density could then be determined. Appropriate compaction pressures were applied to different powder fills, in order to obtain cylindrical powder compacts all of the same height (20 mm) but having different densities. These compacts were in five groups with nominal densities of about 5.8, 6.25, 6.75, 6.99, 7.2 g/cm³ respectively (75, 80.5, 87, 90 and 93 per cent). The green compacts were then sintered for an hour at a temperature of 1150°C.

Homogeneous compression tests

Compression tests were carried out on samples by compressing them between polished, flat dies. Teflon sheets were used as the lubricant at both contact surfaces. This ensured the absolute minimum of friction, so that the resulting deformation behaviour could be considered homogeneous compression. Each group of compacts, all having the same density, were subjected to different loads in order to obtain different strains. After each experiment, both the height and the mean diameter of the sample were measured. From these measurements, the axial stress, axial strain, diametric strain and density were calculated (see figure 2). In addition, the deformed

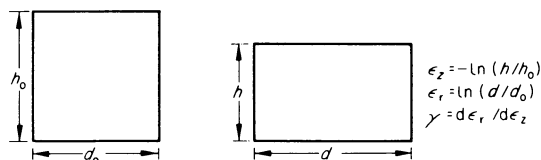


Figure 2. Schematic illustration of the homogeneous compression test.

compacts were examined visually to see whether there were any cracks.

RESULTS

Forging load

The variation of peak forging load with the percentage reduction in height is shown in figure 3, for

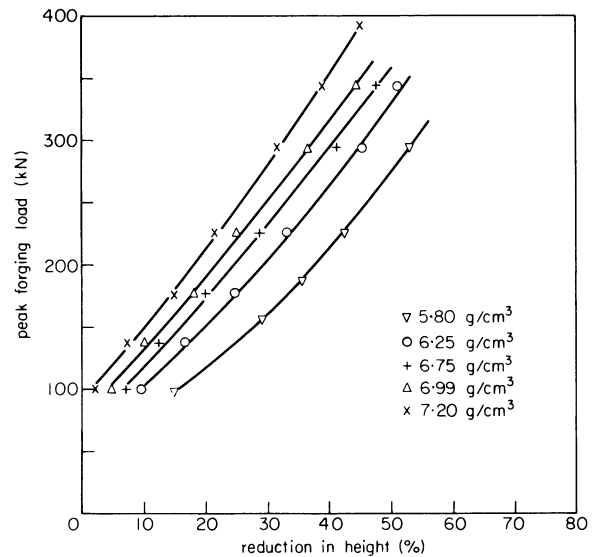


Figure 3. Variation of peak forging load with the percentage reduction in height for different initial preform densities.

different initial preform densities. As can be seen from the figure, the forging load is dependent not only on the percentage reduction in height, but also on the initial density of the preform. A slight increase in either the initial density or the percentage reduction in height will result in a considerable increase in the required forging load.

Densification

Figure 4 illustrates the relationship between the current forged density and the percentage reduction in height, for different initial preform densities. The results show that the densification rate is an inverse function of the initial as-sintered density of the preform. For low initial densities, densification begins after the yield stress of the porous body is reached and then rises rapidly, whereas the rate of densification for high-density compacts is very low at the beginning. Densification then slightly increases for higher values of percentage reduction in height and finally approaches a limit which is less than 100 per cent of the theoretical density. This limit appears to be dependent on the initial density of the preform.

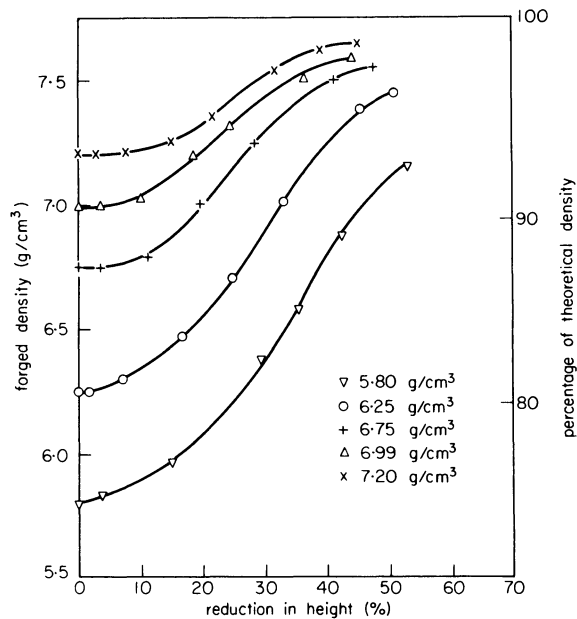


Figure 4. Variation of the current forged density with the percentage reduction in height for different initial preform densities.

Yield stress

When subjecting the preforms to homogenous compression tests, each compact showed a clear yield point at which the applied load appeared to be constant for some time. This yield point was found to be dependent upon the initial density of the preform. It increased with increasing preform density.

Failure of compacts

The Tested preforms exhibited a tendency to crack after uniaxial compression reducing them to about 50 per cent of their initial height. The failure was similar to that of wrought iron billets subjected to upsetting. Cracks appeared at an angle of about 45° with the axis of the preform.

True stress-true strain curves

The true stress-strain relationships are shown in figure 5 for different initial preform densities. It is

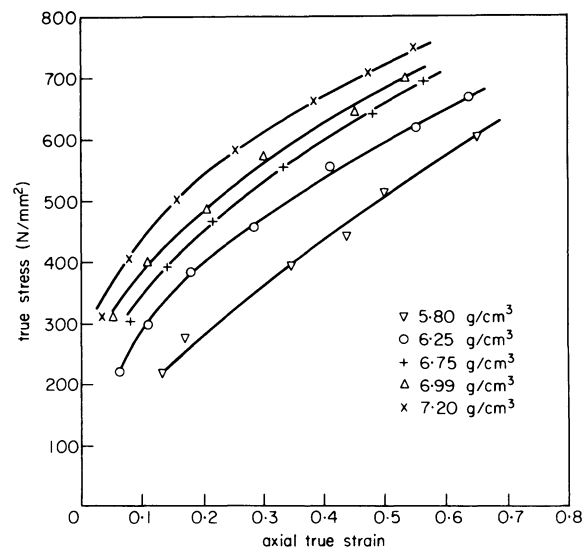


Figure 5. True stress-true strain curves for different initial densities.

well known that flow curves for many materials can be represented approximately by the following exponential relationship:

$$\text{true stress} = K (\text{true strain})^n$$

where both K and n are constants for a given material. It is clear that a log-log plot of true stress versus true strain, for this equation, yields a straight line.

Log-log plots were made for the various initial preform densities used in this study and are shown in figure 6. The values of K (the value of the stress corresponding to unit strain) and n (the work-hardening exponent) could then be calculated for each initial density.

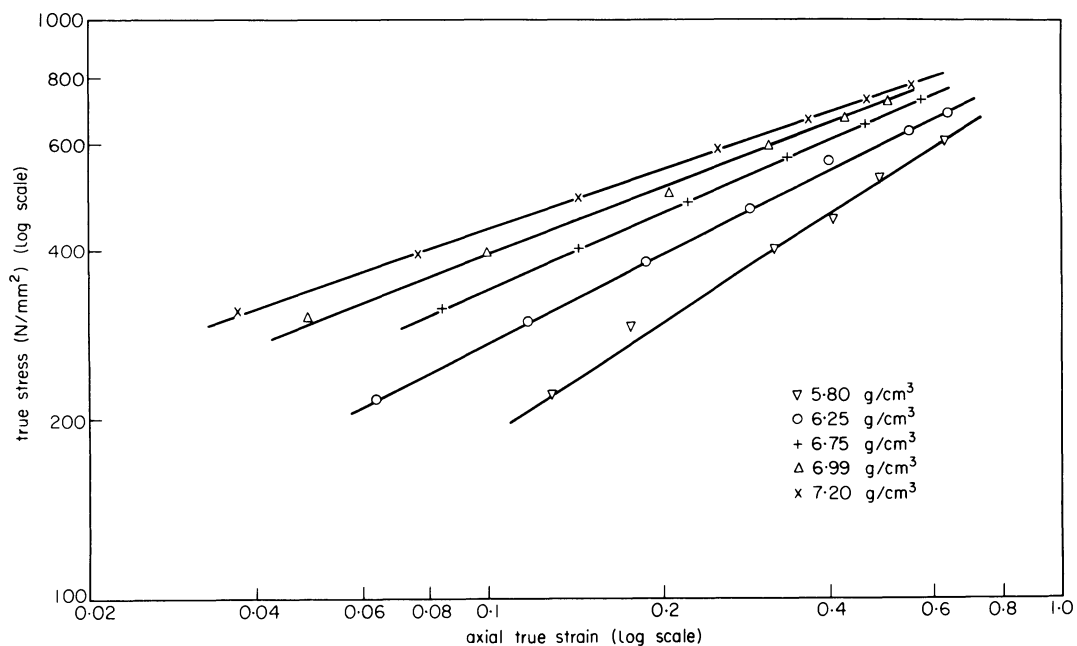


Figure 6. Log true stress versus log true strain for different initial densities.

The value of K

Figure 7 shows the relationship between the value of K and the initial density of the preform. It is clear

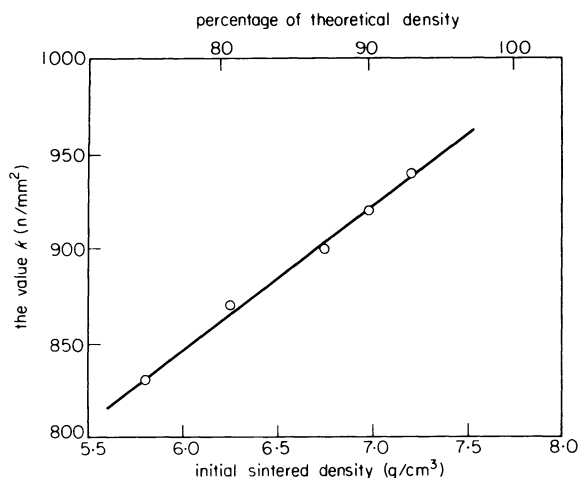


Figure 7. Variation of the value of K with the initial preform density.

from the figure that the value of K increases with the preform density.

Work-hardening exponent

Figure 8 illustrates the relationship between the work-hardening exponent n and the initial preform

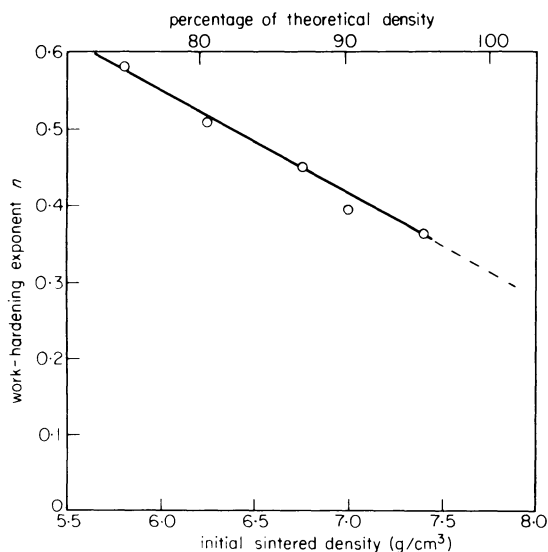


Figure 8. Variation of the work-hardening exponent n with the initial preform density.

density. As can be seen from the figure, the value of n decreases as the initial density of the preform increases. Extrapolating to the full theoretical density, the value of n becomes about 0.3 which is the value for wrought iron. The trend of the curve shown in figure 7 can be explained by considering the mechanism of work-hardening in both fully dense and porous materials. In the case of a fully dense material, work-hardening is the result of piling up of dislocation in the cold-worked grains, while work-hardening in a porous material is due to densification as well as cold-working of the individual powder particles. Accordingly, the work-hardening exponent n of a

porous material can be split into two parts. The first part, which represents work-hardening of the base material, is a constant value (for a given material) and can be determined by extrapolating the curve given in figure 8 to the full theoretical density. The second part, which is due to densification, decreases with increasing preform density. This is explained by the fact that the rate of densification is an inverse function of the initial density of the preform.

Poisson's ratio

Poisson's ratio was calculated by dividing the incremental diametric strain by the corresponding incremental axial strain. The results are plotted in figure 9 which shows the variation of Poisson's ratio

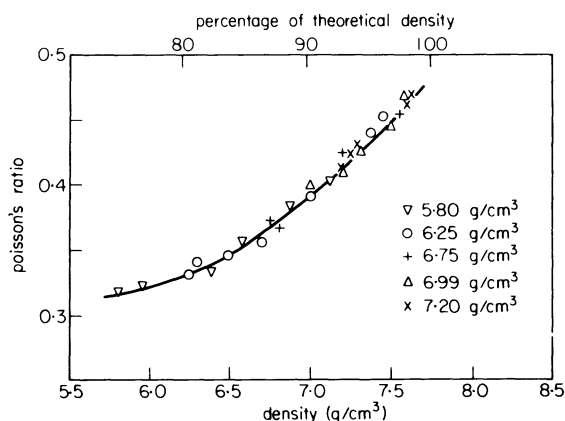


Figure 9. Variation of Poisson's ratio with the current density.

with the current density for all samples tested. As can be seen, Poisson's ratio is a function of the current density of the sample, irrespective of the method of achieving that density. In other words, Poisson's ratio is the same for compacts having the same density, whether that density is achieved by forging a sintered preform or by compaction and sintering. It can also be seen that Poisson's ratio increases with increasing density and extrapolates to 0.5 as full density is approached. It is well known that 0.5 is the value of Poisson's ratio for all conventional materials when plastically deformed. This results from the fact that their volume is constant during plastic deformation.

DISCUSSION OF RESULTS

The results obtained indicated the deformation and densification characteristics of sintered sponge iron powder compacts subjected to homogeneous compression tests at room temperature. A quantitative understanding of these characteristics is essential for the rational design of the powder preform as well as the die in the cold forming processes. It also provides a tentative method for the determination of the load requirements.

The initial density of the preform and the percentage reduction in height were found to be the two factors affecting the final density of the preform after upsetting. As can be seen from figure 4, a final forged density of 7.1 g/cm³ can be achieved by compressing either a compact with an initial density of 5.8 g/cm³ to 50 per cent of its initial height or a compact with

an initial density of 7.0 g/cm^3 by only 14 per cent of its initial height. The forging load in the first case will be about 80 per cent higher than that required in the second case (see figure 3). This indicates that a relatively high forged density can be achieved by compressing a low-density preform by a high percentage of its height. However, it must be borne in mind that there is a limiting value for the density obtained by this method. The lower the initial density of the preform, the lower will be this limiting value. In addition, compressing the height of the compact by more than 50 per cent is not recommended as cracking may take place. Accordingly, medium density preforms (density about 6.8 g/cm^3) should be used, if it is intended to obtain a high density after the cold forming process. This has the advantage of avoiding cracking of the preforms and only relatively low forging loads will be required.

Work-hardening of sintered powder compacts was found to be different from that of fully dense billets. The work-hardening exponent n is always constant for a given full dense material. For a sintered porous material, the work-hardening exponent was found to depend also on the initial density of the compact. This is explained by the fact that work-hardening in a porous material is due to densification as well as cold-working of the individual powder particles. The magnitude of each effect could be determined in the present work.

It was found from the experimental results that Poisson's ratio for a plastically deformed porous material is always less than 0.5 which is the value for full dense wrought or cast materials. This means that when compressing porous materials, they spread in the lateral direction to a lesser degree than would a dense material. It was also found that the value of Poisson's ratio for a porous compact is a function of its current density, no matter how this density was achieved. Accordingly, it can be concluded that there is a definite relationship between deformation and densification. In other words, if the initial density of the preform and the axial strain are known, the final density after upsetting can be calculated using the relationship between Poisson's ratio and the density. By obtaining the value of Poisson's ratio for the given density (from figure 9), the diametric strain, the change in volume and the final density can then be calculated. The relationship between Poisson's ratio and the preform density is also used in developing plasticity equations for sintered porous materials. It has been reported⁶ that the pressures in repressing sintered compacts could accurately be predicted using the Levy-Mises equations for plastic deformation. In that case, the value of Poisson's ratio was not taken as 0.5, but varied with the compact density according to a relationship similar to that shown in figure 9.

CONCLUSIONS

1. The deformation and densification characteristics of sintered sponge iron powder compacts, subjected to uniaxial compression were determined. These characteristics are essential in the

development of plasticity equations for sintered porous materials.

2. The value of Poisson's ratio for a sintered compact is a function of its current density, no matter how that density is achieved.
3. The work-hardening exponent n varies with the initial density of the compact. This is explained by the fact that work-hardening in a porous material is caused by densification as well as by cold-working of the individual powder particles.
4. During uniaxial compression, the rate of densification is an inverse function of the density of the compact.
5. The load required to achieve a given density increases with decreasing preform density.
6. Although a relatively high forged density can be achieved by compressing a low-density preform by a high percentage of its height, it is recommended to use medium density preforms ($6.7\text{--}6.8 \text{ g/cm}^3$), when a high final density is required. This has the advantages of relatively low load requirements and of avoiding the cracking of preforms during deformation.
7. The cold formability of sintered sponge iron powder compacts is surprisingly good. Compacts could be compressed to 50 per cent of their initial height without any cracking.

ACKNOWLEDGMENTS

The author would like to express his deep gratitude to Mr H. Khalifa, the technical manager of SABI Company for sintered components, for making available the facilities required to carry out the sintering process. The author is also grateful to Prof. A. Ata for his help.

REFERENCES

1. H. F. Fischmeister et al. Deformation and Densification of Porous Preforms in Hot Forging. *Powder Metallurgy* (1971) 14, No. 27, 144.
2. T. Nakagawa et al. Cold Forging of Ferrous M/P Billet 27th. *Symp. Japan Soc. Powder and Powder Metallurgy*, Tokyo, 1971.
3. T. Nakagawa et al. On the Cold Forging of Sintered Iron Powder Preforms. *Proc. 13th. Int. MTDR Conf.*, Birmingham, 1972.
4. A. Singh and R. Davies. Preliminary Investigations of the Cold Extrusion of Powder Preforms. *Proc. 13th. Int. MTDR Conf.*, Birmingham, 1972.
5. H. W. Antes. Cold Forging Iron and Steel Powder Preforms. *Modern Developments in Powder Metallurgy*, vol. 4—Processes, Plenum Press, 415, 1971.
6. H. A. Kuhn and C. L. Downey. Deformation Characteristics and Plasticity Theory of Sintered Powder Materials. *Internal J. Powder Met.* (1971) 7 (1), 15.
7. H. A. Kuhn et al. Deformation Characteristics of Iron-Powder Compacts. *Modern Developments in Powder Metallurgy*, vol. 4—Processes, Plenum Press, 463, 1971.

THE EFFECT OF PRESINTERING CONDITIONS ON THE COLD FORGEABILITY OF POWDER PREFORMS

by

M. A. RIFAI*

SUMMARY

The presintering conditions as well as the initial preform density affect the strength of preforms, which has to be sufficient to sustain the forging stresses during cold forging. It can be considered that the maximum requirement for preform strength before cold-forging is needed at the first stages of deformation when the porosity is at its maximum. As the preform deforms, sufficient strength is developed to withstand the forging stress. This increase in strength is due to the development of contact areas accompanied by an increase in cold welded portions.

It is well known that bonding takes place very early in the sintering process as the material heats up. Naturally, the greater the original density of the preforms, the greater is the amount of contact areas and potential grain boundary.

Therefore the possibility of minimizing the presintering conditions was investigated. The cold forgeability of preforms presintered at 1120°C and 900°C for 5, 10 and 20 min and the effect of initial preform density were studied.

It was found that presintering for 5 min at 900°C was sufficient for cold homogeneous uniaxial deformation. The required pressure for deformation was half that required for hot forging.

INTRODUCTION

The technique of cold forging of powder preforms becomes more attractive as a method for minimizing or eliminating the residual porosity. This is because of the complicated problems foreseen in using the hot forging process for powder preforms.

The heating process in the hot forging technique involves the risk of oxidation whatever common precautions may have been taken. This process is mainly used to increase the plasticity of powder particles for the severe deformation required to eliminate the residual porosity. The weldability of inter-particle contact areas may also be developed as a result of the hot forging. Aren¹ shows that there is no beneficial effect of presintering before hot forging. The forging temperature has a negligible influence on the densification produced by a given strain; the pressure to produce this strain, however, does depend on temperature. Presintering for 0, 4 and 24 hours at 1100°C gave the same density.

In the cold-forging technique the heating process is separately performed as a presintering process. Presintering in this case is to increase the strength of the preforms to sustain the deformation stresses during forging. Therefore it is a common practice^{2,3,4} to presinter iron and iron-based powder preforms to about 30–100 min at 1120°C before cold forging.

A substantial increase in the strength of the preforms before forging (by complete sintering) may hamper the densification process and increases the forging loads and process cost. Therefore the optimum conditions for presintering were studied.

EXPERIMENTAL PROCEDURE

Cylindrical test samples 24 mm in diameter by 26 mm high were formed from Höganäs NC 100-24 sponge iron powder of prepared sieve analysis. Compaction was carried out in a free-floating die with die-wall lubricant (zinc stearate in benzole). Two compaction pressures were used to get constant height for low and high-density specimens weighing 60 and 80 g. The resulting nominal densities were 5.4 and 6.5 g/cm³ respectively. Both samples of high and low initial densities were presintered at 1120°C and 900°C by introducing them into the hot zone of a tubular furnace with a dry hydrogen atmosphere, allowing 10 min for them to reach the presintering temperature and then holding them for periods of 5, 10 and 20 min, followed by cooling for 20 min to room temperature.

Forging was carried out on a 5000 kN hydraulic press fitted with plane-ground dies. The upper die was equipped with load and displacement transducers.

* Al-Azhar University, Cairo, presently NORAD fellow at Institutt for mekanisk teknologi, NTH, Trondheim

Signals from these were recorded on an x - y record.

Teflon sheets 0.1 mm thick were used as a lubricant at both contact surfaces and were replaced after each increment of deformation. This assured the absolute minimum of friction so that only the pure homogeneous uniaxial compression of the material resulted.

DEFORMATION BEHAVIOUR

Preform geometry

Figure 1(a) shows the effect of using teflon as a lubricant during the cold forging process. No barreling can be seen at any of the stages of deformation (30, 40, 60, 70 and 78 per cent reduction).

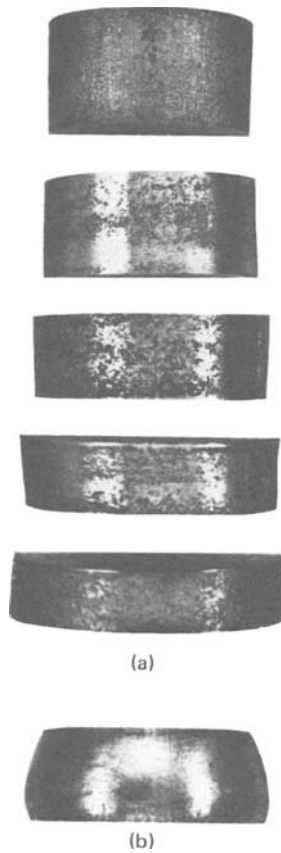


Figure 1. Cold-forged iron preforms at different steps of homogeneous uniaxial deformation. (a) Uniform initial density, (b) non-uniform initial density.

The small irregularity of the outer surface was a result of non-uniform initial density distribution within the preform. This irregularity in geometry was much less than warranted by the actual amount of non-uniformity in initial density distribution, the reason being the difference in the rates of deformation for different densities. However, the lateral flow of low density portions at the first stage of deformation was higher than the lateral flow of the high-density portions. Therefore in the case of single-acting compaction of powder, as the density decreased by the height the lateral flow was relatively increased, as shown in figure 1(b).

The uniformity of preform surface and geometry was better with preforms of high initial density (6.5 g/m^3) than ones of low initial density (5.4 g/m^3).

Preforms presintered at different conditions of temperature and time show no significant difference in geometry after cold forging. However, better results were obtained for preforms heated slowly (20 min) to reach the presintering temperature.

Pressure/strain relationship

Figure 2 shows the forging pressure required to produce a given strain. Pressure was calculated for five steps of deformation. The maximum force was recorded in the forging cycle at every step and the area of the deformed specimen was also measured.

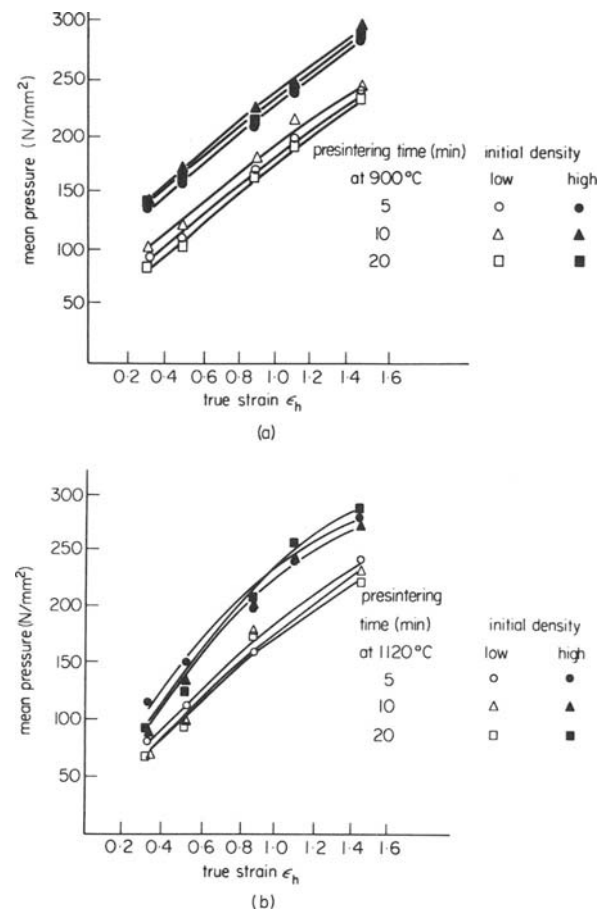


Figure 2. The relation between mean forging pressure and true axial strain of powder preforms.

The deformation behaviour of preforms of different initial densities was different for all the deformation steps. The preforms with low initial density deformed more easily than those with high initial density. This behaviour remained the same for different strains.

The pressure required to achieve a certain degree of deformation was very low compared with the results of hot forging⁵ and also cold forging⁶. The study of cold forging⁶ of sponge-iron powder preforms presintered for 1 hour showed that the pressure required to produce a true strain of 0.55 is about 500 N/mm² compared with 100–150 N/mm² as shown in figure 2.

This can be explained by low preform strength due to the presintering conditions studied. Low preform strength allows more deformation of the powder

TABLE 1 Work-hardening exponent n and k values from log-log plots of $p\sigma$ against ϵ

Initial density	time (min)	5		10		20	
	temp ($^{\circ}\text{C}$)	n	$K(\text{N/mm}^3)$	n	$K(\text{N/mm}^3)$	n	$K(\text{N/mm}^3)$
high	900	0.53	220	0.47	230	0.50	225
low	900	0.68	175	0.63	190	0.66	170
high	1120	0.72	210	0.67	215	0.69	225
low	1120	0.77	170	0.69	180	0.74	175

particles and therefore less pressure was required for a given deformation.

The difference in pressure required for cold forging powder preforms presintered at 900°C and 1120°C was relatively small. A small increase in the pressure required for a certain strain can be seen in the case of preforms presintered at 900°C over those at 1120°C , especially in the first stages of deformation.

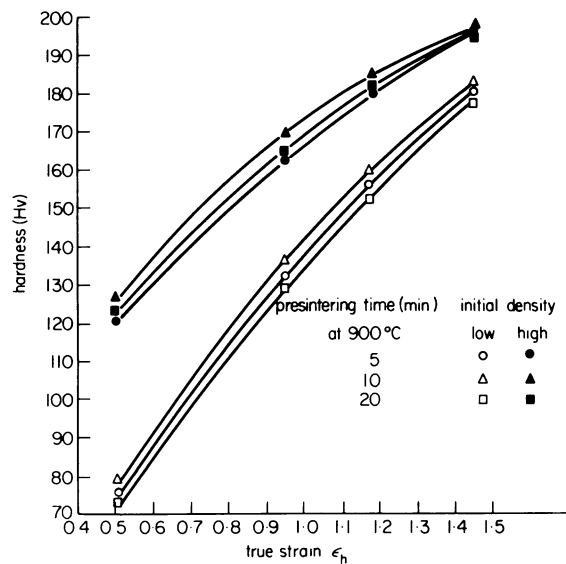
No significant effect on the required pressure due to the difference in presintering time could be seen. The small scatter of results for different presintering times may be due to the scatter in the initial density of preforms.

Flow curves for many materials approximate to the exponential relationship $\sigma = K\epsilon^n$. A log-log plot of σ against ϵ for this equation yields a straight line³. Log-log plots were made for the cases studied and approximate linear curves resulted. The values of n , the work hardening exponent, and K are listed in Table 1. This data indicates that as the presintering temperature increases from 900 to 1120°C , the value of n increases. This behaviour for preforms of high initial density is the same as for low initial density. Generally, the value of the work hardening exponent decreases as the time of sintering increases. The value of K , on the contrary, decreases as the presintering temperature increases.

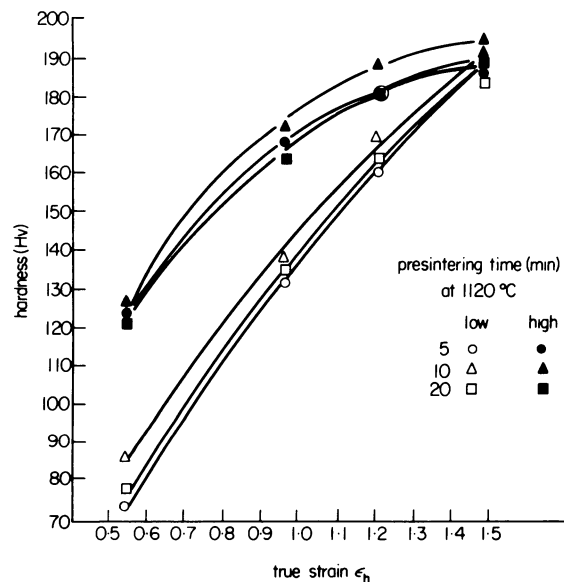
Hardness/strain relationship

Figure 3 shows the relation between hardness and strain for preforms presintered at 900 and 1120°C . The hardness of preforms of both low and high initial density increases at different rates for specimens presintered at 900°C and also for specimens presintered at 1120°C . For specimens presintered at 1120°C the hardness of preforms of high and low initial density was the same at axial strain $\epsilon = 1.45$. A presintering time of 10 min gave higher hardness values than 5 and 20 min for all values of height strain studied.

No cracks were seen at the different steps of deformation till the reduction ratio in height was 78 per cent. However the initial strength of preforms was low due to the limited presintering condition. This behaviour can be explained by the low stresses applied for deformation, especially at the first stages when the preforms are very weak. The improvement in the strength during deformation, by improving the areas of particle contact at higher reduction ratios, was sufficient to sustain the deformation stresses without the appearance of any surface cracks.



(a)



(b)

Figure 3. Relation between hardness and true axial strain of powder preforms.

Lateral flow

For a dense material which is subjected to the uniaxial compression, the change in axial true strain is equal to half the lateral true strain. This means that a decrease in height results in an increase in the diameter equal to half the decrease in height to satisfy the condition of constant volume which is imposed by the incompressibility of the material. The

deviation is mainly a function of the compressibility of the material or the volume change of porous preforms.

Figure 4 shows the relation between the axial true strain and lateral true strain for preforms homogeneously and uniaxially cold-forged.

All curves for the different conditions studied show significant deviation from incompressible behaviour, especially in the first stage of deformation. This indicates that the material at first flows mainly in the direction of punch movement, with little lateral flow. As the density increases, lateral flow

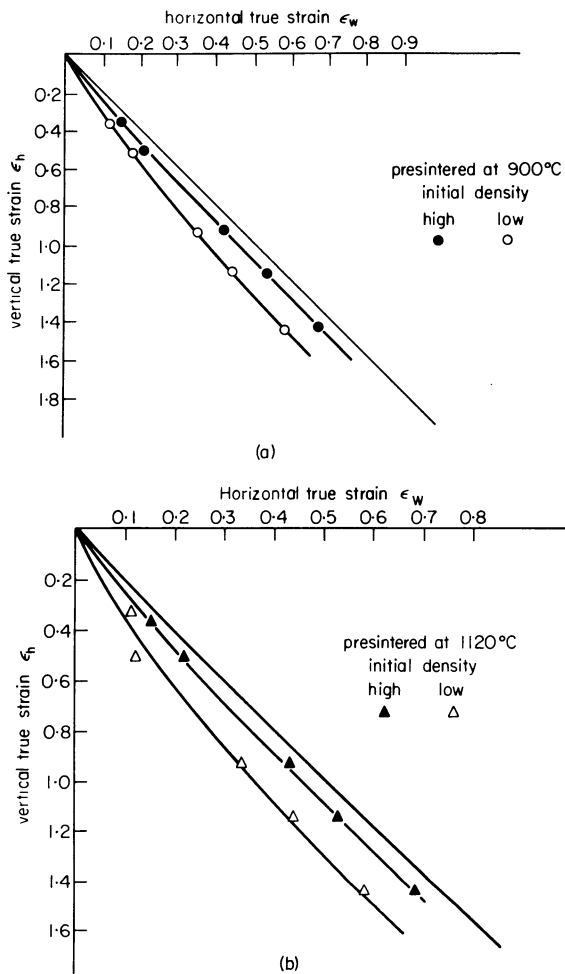


Figure 4. Lateral spreading (ϵ_w) during upsetting (ϵ_h) for preforms of different initial density.

increases. The rate of change of Poisson's ratio was changed continuously for all cases studied.

In the final stages of deformation the lateral flow approached the spreading behaviour of pore-free material at an axial true strain of 1.2–1.4.

Lateral spreading increased with increasing initial density of preforms for a given axial strain. This behaviour is clear for both the cases shown in figure 4.

Preforms presintered at 1120°C show lateral spreading somewhat closer to the pore-free material behaviour than those presintered at 900°C. This behaviour is more noticeable at higher strains.

The effect of presintering time on lateral spreading seems to be very limited. The reason may be that the

small time range in the cases studied (5, 10 and 20 min), was not sufficient to have a significant effect on lateral flow.

It is worth noting that the shortest presintering time, 5 min, was sufficient to allow extensive lateral flow, $\epsilon_w = 0.67$, which is desirable for good filling against the die walls, without any crack formation.

DENSIFICATION

Densification occurs during homogeneous uniaxial deformation. The micrographs in figure 5 show the pore form and distribution of presintered powder preforms at five different steps of deformation (30, 40, 60, 70 and 78 per cent reduction) and the initial condition of pores before deformation. There is a high rate of densification, especially in the first stages. In the final stage of deformation there is some residual porosity. The form of the pores in this stage is more oblong than spherical, which may be a disadvantage from the point of view of mechanical properties. Therefore exclusion of such severe deformation is recommended, especially if we consider that the increase in density by an axial strain such as 1.45 was not generally obtainable. (This will be discussed later.)

Figure 6 shows the change in the rate of densification during deformation. The densification rate was higher for preforms of low initial density than for ones of high initial density. This behaviour changes at different stages of deformation. However, the higher the initial density, the higher is the density after forging for all stages of deformation. Little increase in density was obtained by increasing the axial strain above 1.2 for preforms of high initial density, but it increases further for ones of low initial density, up to an axial strain of $\epsilon = -1.4$.

Figure 6 shows the effect of the presintering temperatures on the final density at different stages of deformation. Preforms presintered at 900°C densified somewhat less than those presintered at 1120°C. This difference can be seen only in the case of preforms of low initial density. In the case of high initial densities, the densities obtained for different strains were the same. At 900°C and with 10 min presintering, preforms reached densities 0.07–0.03 g/cm³ higher for all strains than did those presintered for 5 and 20 min at the same temperature. This difference can be neglected. For 1120°C the effect of presintering time on density was very small, and different for the different strains. The maximum density of 7.65 g/cm³ was obtained at $\epsilon_h = 1.2$ for preforms of initial density 6.6 g/cm³ presintered for 5 min at 900°C. These results agree with the results given by Antes³ for cold-forged sponge iron preforms uniaxially deformed and presintered at 1120°C for 30 min. It is also close to the results given by Fischmeister⁵ for hot forging of sponge iron preforms.

Effect of pressure on densification

Figure 7 shows the densification of the preforms with increasing pressure. The curves are similar to the compressibility curves obtained in normal cold compaction, especially in the first stages where the lateral

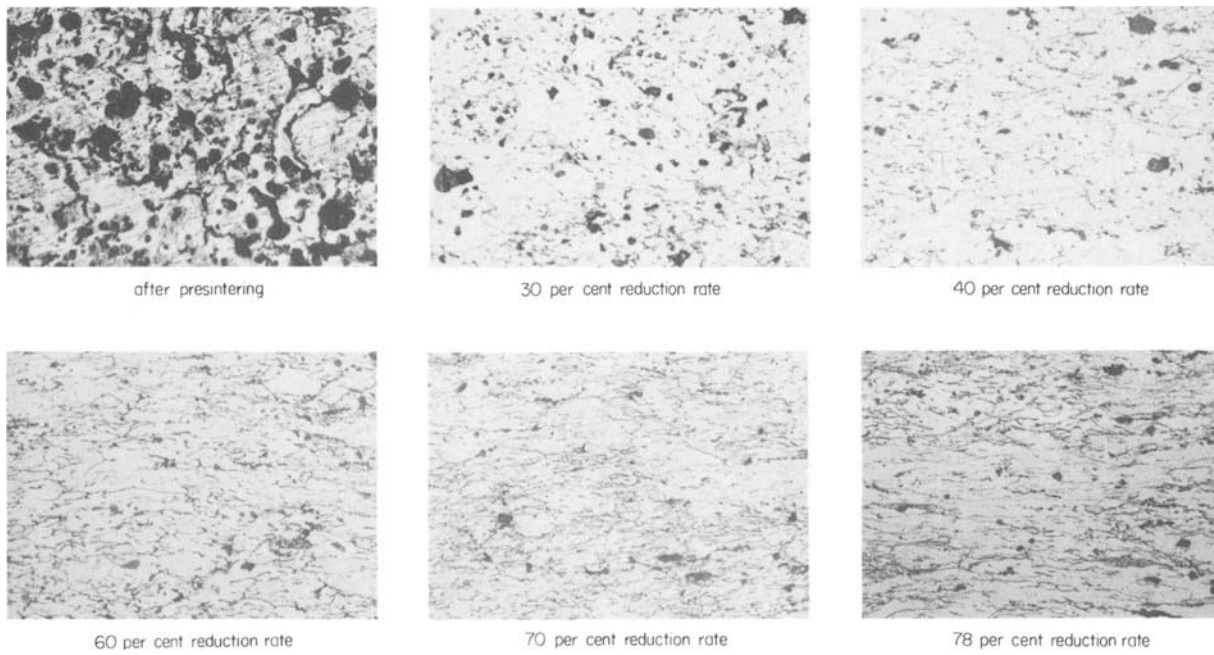


Figure 5. Microstructure of iron powder preforms (presintered at 900°C for 5 min) at various deformation ratios (0, 30, 40, 60, 70 and 78 per cent) (x 200).

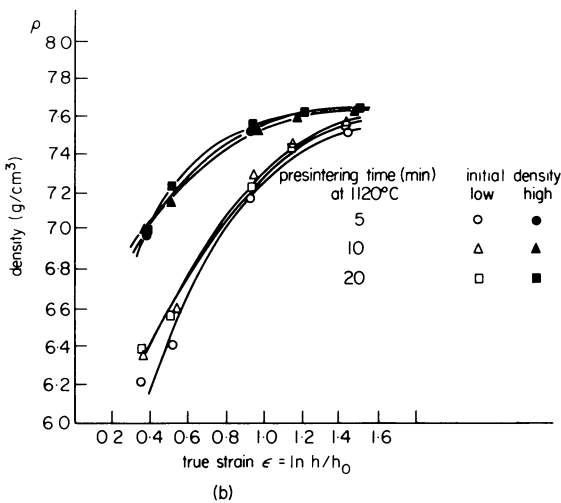
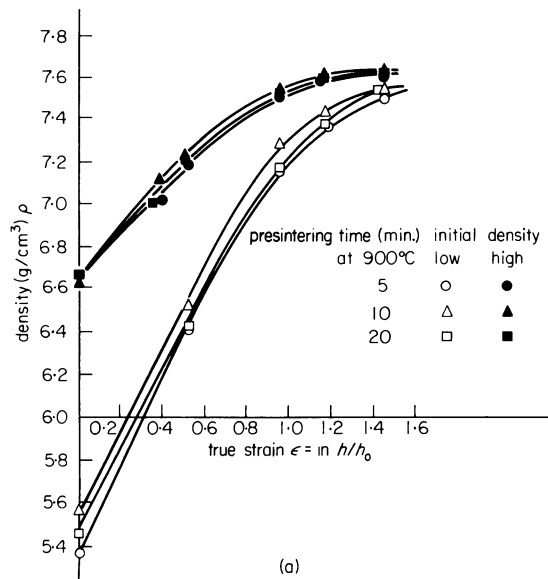


Figure 6 Relation between density and true axial strain.

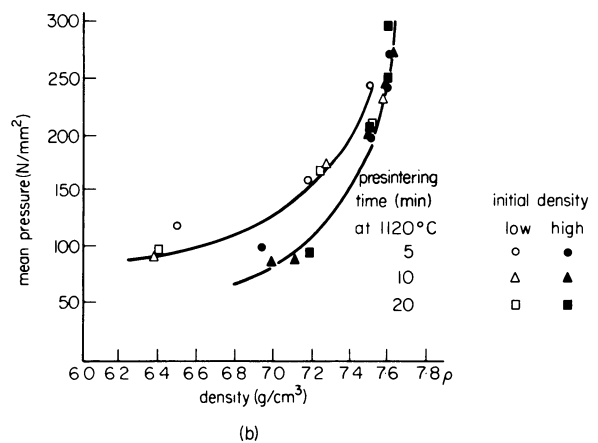
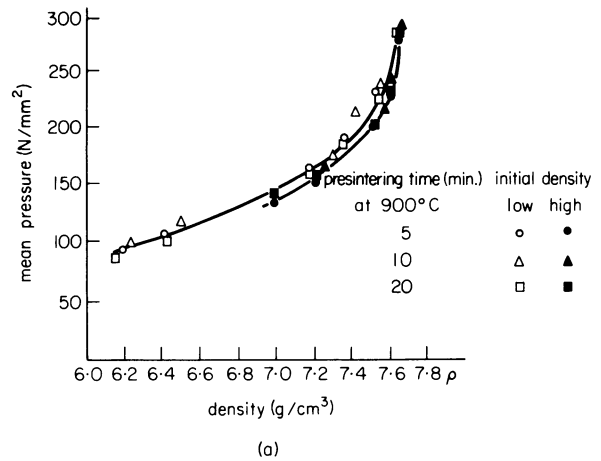


Figure 7. Relation between density and mean pressure.

flow was limited. The mechanism of densification by the cold-forging technique under the conditions studied is more similar to that of powder compacting if the effect of free lateral flow and absence of die wall friction is neglected.

The effect of preform density was different for

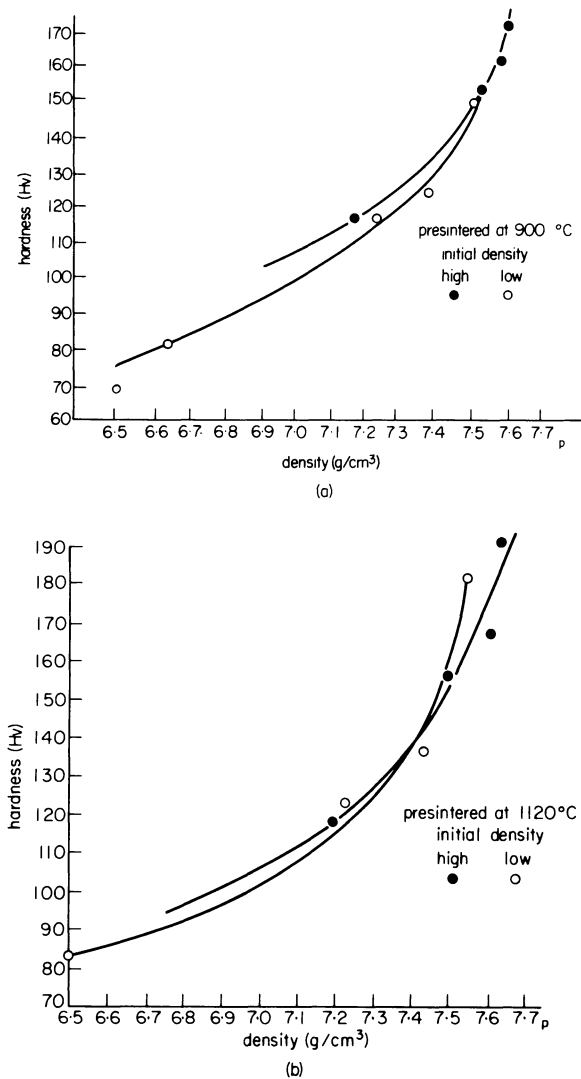


Figure 8. Relation between hardness and density.

different presintering temperatures. Preforms of low initial density required a somewhat higher pressure than ones of high initial density to get the same density. This was noticeable in the case of presintering at 1120°C. The difference in pressure required decreases as the density increases. For densities lower than 7.5 g/cm³, preforms presintered at 1120°C required less pressure to reach a particular density than preforms presintered at 900°C. For densities above 7.5 g/cm³ the difference in the pressure required was negligible.

The highest density obtained was 7.65 g/cm³ for preforms of initial density 6.6 g/cm³ presintered at 900°C for 5 min. This density was achieved at axial strain $\epsilon_h = -1.17$ and lateral strain $\epsilon_w = 1.966$. (The required pressure was 290 N/mm².) The same density was achieved at somewhat lower pressure (279 N/mm²) for preforms of initial density 6.55 g/cm³ presintered at 1120°C for 20 min.

Density/hardness relationship

The relation between density and hardness in figure 8 was established for specimens deformed to different reduction ratios (30, 40, 60, 70 and 78 per cent). Vickers hardness was measured and the mean value of thirty indentations was calculated. There was a scatter in the hardness values for the same specimen within a range of 10 per cent.

Figure 8 shows that hardness increases with increasing density. The rate of increase of hardness was lower at higher initial densities than at lower ones. Antes³ obtained completely different results by presintering the preforms for 30 min at 1120°C. However, for hot-forging⁵ the relationship between hardness and density agrees with the results given in figure 8. At low strains the higher the initial density, the higher the hardness value for preforms presintered

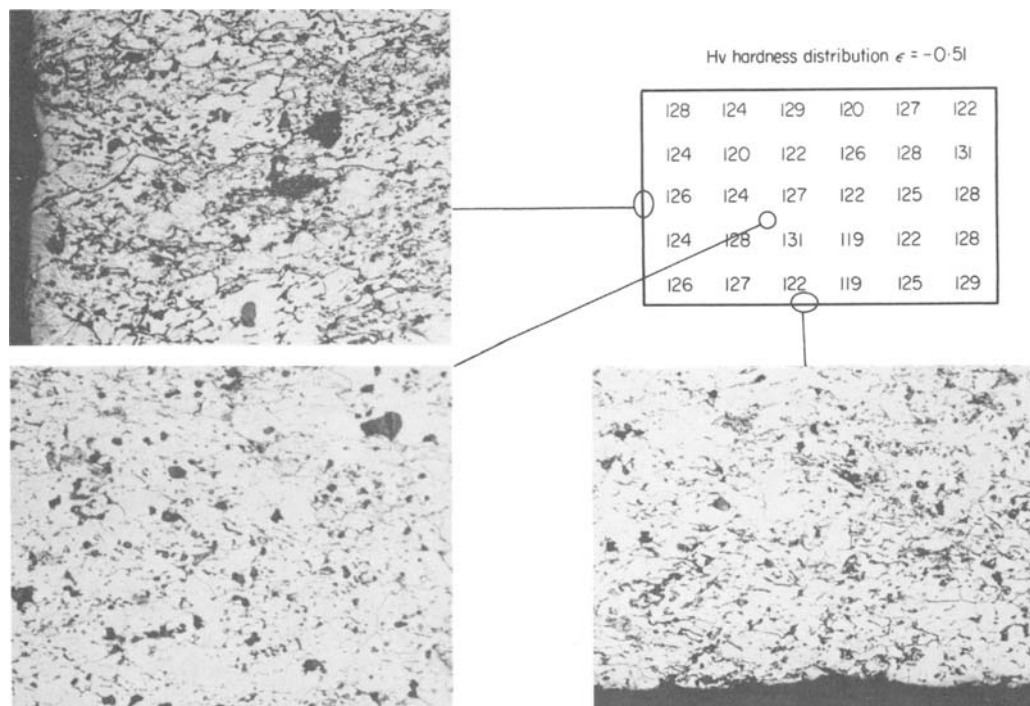


Figure 9. Hardness distribution and microstructures in centre cross-section of preform deformed to $\epsilon_h = 0.55$, presintered at 900°C for 5 min. (Micrographs x 200.)

at 900°C and 1120°C. At high strains, the higher the initial density, the lower the hardness value.

The same hardness was obtained for preforms of high and low initial density at a density of 7.40 g/cm³ in the case of presintering at 1120°C and 7.55 g/cm³ in the case of presintering at 900°C. Generally, preforms presintered at 1120°C had a higher hardness value than those presintered at 900°C.

Presintering time had very little effect on the hardness value at different height strains. Specimens presintered at 900°C for 5 min had somewhat less hardness than those presintered for 10 and 20 min.

Density distribution

The local densification effect was charted by means of a 2 x 2 mm grid of Vickers hardness impressions, made on the centre cross-section of the forged cylindrical preforms of high initial density presintered at 900°C for 5 min. Figure 9 shows the hardness distribution of a cold-forged preform deformed to a height strain $\epsilon_h = -0.51$, together with micrographs of the porosity in regions of particular interest. Figure 10 shows the porosity of corresponding regions at higher strain, $\epsilon_h = -1.4$. The density distribution was characteristically different from that obtained in cold compaction of loose powder in a die. In the case of cold forging by homogeneous uniaxial deformation, the hardness seems to be homogeneous all over the section though some scatter can be seen. Scatter in the hardness measurements can be explained by the irregular orientation of powder particles. Some of the powder particles may be oriented so that they carry severe stresses, while at the same time others may be oriented so that they are much less stressed. This can affect the hardness distribution at relatively high strain.

The uniformity in hardness distribution can be considered as the result of individual deformation and re-arrangement of powder particles during deformation due to the weak contact between them. Of course many of the weak contact points between powder particles could break, but the new contact areas formed during deformation under suitable cold welding conditions must be much stronger. That may also be the reason for higher lateral flow at higher strain.

Pores close to the surface of cold-forged preforms can be considered uniform with the core material, as can be seen from the micrographs in figures 9 and 10 which show the microstructure of different sections through the horizontal and vertical surfaces of a preform.

CONCLUSION

Bonding takes place very early in the sintering process as the material heats up. Much of this bonding process may take place while the sinter mass is being heated to a specific sintering temperature and during the early period of constant temperature. This stage of sintering does not lead to any dimensional change of the material. The bonding imparts a high degree of coherence and integrity to the material⁷; this is evidenced by rather large increases in strength and hardness even after relatively short exposures to an elevated temperature. Naturally the greater the original density of the material, the greater the amount of contact area and potential grain boundary.

Based on these facts the study of the effect of presintering conditions and initial density shows that:

- (1) 5 min presintering at 900°C was a sufficient condition for cold homogeneous uniaxial deformation of iron powder preforms, using teflon

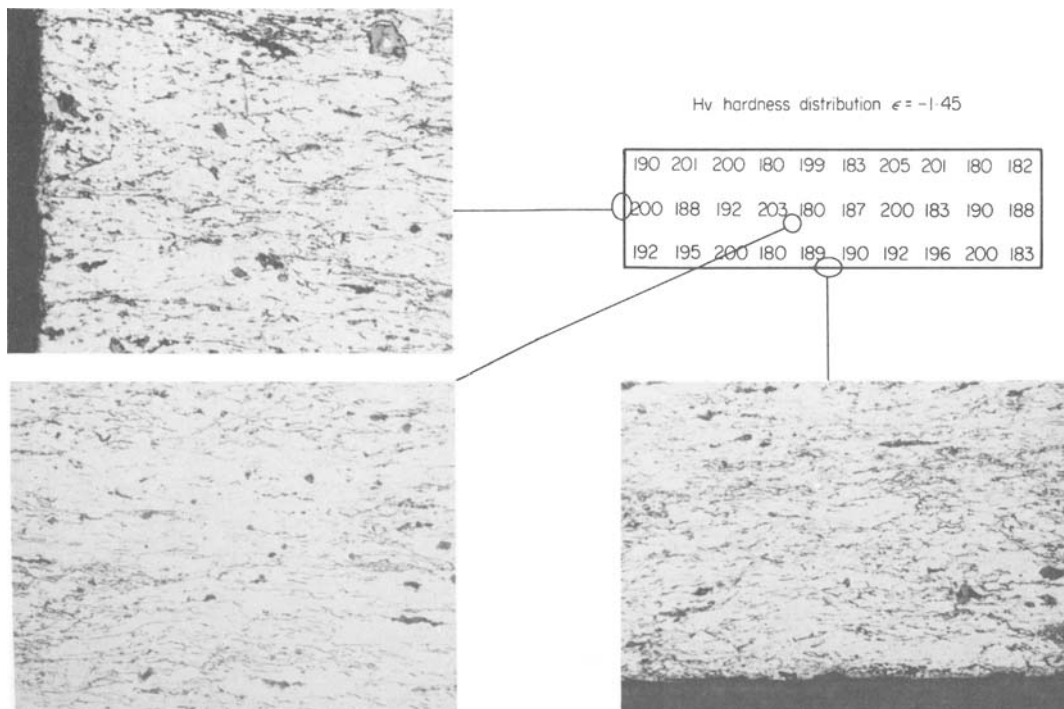


Figure 10. Hardness distribution and microstructures in centre cross-section of preform deformed to $\epsilon_h = 1.45$, presintered at 900°C for 5 min. (Micrographs x 200.)

sheets to minimize the friction between the die and the preform.

- (2) for these conditions the required pressure was less than half of that required for hot forging. No cracks appear until a strain of $\epsilon = -1.45$ and an extensive lateral flow $\epsilon_w = 0.67$ was allowed.
- (3) the higher the initial density, the higher the density after forging for a given degree of deformation, although at higher strains the rate of densification for preforms of low initial density was higher.
- (4) the homogeneity in density distribution in addition to the above results can be explained by the low preform strength which allows more free re-arrangement and deformation of powder particles.

ACKNOWLEDGMENTS

The author wishes to express his grateful appreciation to Professor Almar Almar-Næss for his encouragement and advice. He would also like to thank *siv. ing.* Sigurd Støren for his assistance.

This work was done in the Department of Metallurgy and Metal Forming, Institute of Mechanical

Technology, NTH Trondheim, and supported by NORAD fellowship.

REFERENCES

1. B. G. A. Aren, L. Olsson and H. F. Fischmeister. Powder forging. The influence of pre-sintering and forging temperature. Chalmers' University of technology, Gothenburg, Sweden, internal report (R106/71), 1971.
2. H. A. Kuhn, M. M. Hagerty, H. L. Gaigher and A. Lawley. Deformation characteristics of iron-powder compacts. *Modern developments in powder metallurgy*, volume 4, 1970.
3. H. W. Antes. Cold forging iron and steel powder preforms. *Modern developments in powder metallurgy*, volume 4, 1970.
4. Kunio Obara, Yasuo Nishino and Yuichi Saito. The cold forging of ferrous P/M preforms. *Proc. Powder Metallurgy Conference*, Toronto, 1973.
5. H. F. Fischmeister, B. G. A. Aren and K. E. Easterling. Deformation and densification of porous preforms in hot forging. *Powder Metallurgy* (1971) **14**, no. 27, p. 144.
6. M. A. Kuhn and C. L. Downey. Deformation characteristics and plasticity theory of sintered powder materials. *International Journal of Powder Metallurgy* (1971) **7**, no. 1.
7. S. A. Tsukerman. *Powder Metallurgy*, Pergamon Press, 1965.

THE PRODUCTION OF EXTRUDED MATERIAL FROM METAL POWDERS

by

T. SHEPPARD*

SUMMARY

One method of refining structure in metals is to ensure rapid solidification. Powders may be produced under such conditions by the atomization process and in the particular case of aluminium alloys, the oxide film surrounding such particles may be advantageous. The work reported uses an Al-Mn-Al₂O₃ alloy as an example and demonstrates how the extrusion load is reduced and work normally considered to be 'redundant' converted to useful application. The paper compares the properties with those of other powder materials and investigates the development of structure and sub-structure in the Al-Mn-Al₂O₃ alloy in some depth.

INTRODUCTION

The production of extruded sections from metal powders has been, and still is hindered to a considerable extent by the high cost of the raw material. However, it seems likely that the advent of powder forging will increase the tonnages of iron alloy powder produced, thus making a fabrication route from powder economically attractive in that field. The extension of such success into the light alloy field would probably first occur by exploitation of the special properties obtainable from aluminium powders. The research programme reported in this communication has produced extruded rods using copper alloy, steel, nickel rich alloy and aluminium alloy powders. For the sake of clarity, and because all the basic principles may thus be illustrated, this paper will concentrate on reporting the production of an Al-Mn-Al₂O₃ alloy, referring where necessary to the other materials produced.

In commercial practice the casting of a high weight percentage Mn content aluminium alloy is restricted to the production of a master alloy, which is then used to produce alloys such as Al-1.25 per cent Mn for wire and beer barrels, Al-2.5 per cent Mn-2.5 per cent Ni for gas burners and die casting and Al-4 per cent Mn for gas burners and anodising materials. The master alloy is usually an Al-10 per cent Mn and cannot be used directly because the extended freezing range gives feeding problems, leading to poor surface finish and hot tearing. The limited solid solubility also giving rise to porosity and segregation. When chill cast, coarse MnAl₆ is formed at grain boundaries causing poor ductility and rapid failure on application of a tensile stress. When hot working such an alloy, 'hot shortness' coupled with high yield stress at

working temperatures would result in both heavy investment in capital machinery and poor surface finish. It has been shown^{1,2} that because of the zones of massive shear prevailing in the extrusion process, metal powders may be formed into coherent material employing working loads which are much lower than those obtaining when cast material is processed. The powder metallurgical method of overcoming these problems is not new and many components are manufactured using accepted pressing and sintering techniques. It is therefore apparent that powder extrusion would be an economic method to produce those alloys which exhibit high temperature strength and/or casting problems. This paper reports the investigation of the densification process inherent in powder extrusion exemplified by Al-3.5 per cent Mn alloy. The properties and structure of the alloy are also reported in detail.

THEORY

Perhaps the most significant advantages of powder extrusion are that a lower pressure is required to implement the process and that the sintering usually associated with components fabricated via a powder metallurgical route is eliminated. Both of these phenomena may be explained by considering the mechanics of the extrusion process.

The pressure required for extrusion may be expressed as:

$$p = A\bar{Y}(a + b \ln R + c) \quad (1)$$

in which a represents redundant work which does not contribute directly to shape change, c is the work required to overcome friction and $b \ln R$ is the

* Department of Metallurgy & Materials Science, Imperial College, London

homogeneous work of deformation representing the least possible work required to effect the necessary shape change. Thus we may immediately recognize that the redundant part of this work should be available an energy source, should we desire to execute structural changes in the extrusion billet.

When considering a powder billet, it becomes clear that the homogeneous deformation term will no longer depend upon the properties of the extrude, but will be a function of the differing rheological properties obtaining in the quasi-static deformation zone. The effect may best be illustrated by reference to an idealized upper bound solution for plane-strain extrusion. Sheppard and Raybould³ have shown that whilst the same mathematical analysis does not apply in the axi-symmetric case, the solution may be modified to include this process because the same fundamental principles apply. Figure 1 shows an upper

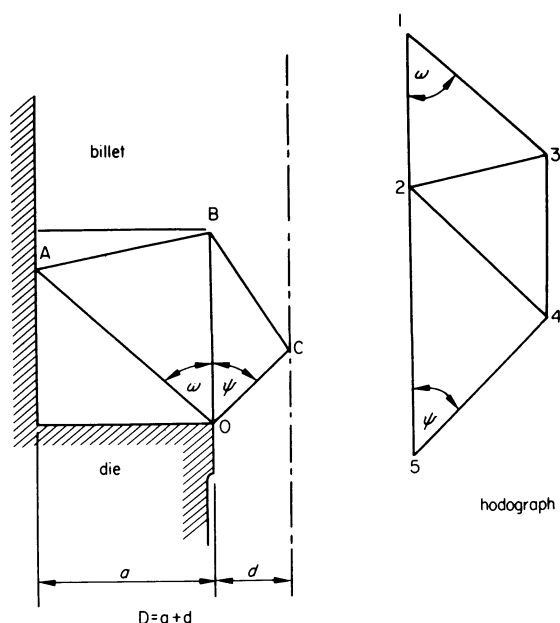


Figure 1. Upper bound solution for plane-strain extrusion.

bound solution of the type obtained using the concepts first proposed by Johnson⁴. In general we may write

$$dW = k \cdot u \cdot dS$$

where k = shear yield stress

u = magnitude of velocity discontinuity

S = length of the velocity discontinuity

W = work done

Applying this to the specific problem of the extrusion of a homogeneous material, we obtain:

$$\frac{p(a+d)}{2} \cdot l = k(AB\bar{2}3 + BC\bar{2}4 + BO\bar{3}4 + CO\bar{4}5 + AO\bar{1}3) \quad (2)$$

whilst for a non-homogeneous powder compact, we must write

$$\frac{p(a+d)}{2} \cdot l = k_1(AB\bar{2}3 + BC\bar{2}4) + k_2AO\bar{1}3 + k_3BO\bar{3}4 + k_4OC\bar{4}5 \quad (3)$$

Here k_1 represents the resistance to shear of the powder compact which is virtually zero and k_4 represents the resistance to shear of the final product which will be the same or greater than those of a conventionally processed product. All other shear terms in Equation (3) depend upon the average shear strength of material contained in the relevant rigid quasi-static triangle and their magnitudes lie somewhere between zero and final product values. Thus comparison of Equations (2) and (3) demonstrates that we may expect considerably reduced pressures when processing powder compacts than when forming conventionally cast materials.

EXPERIMENTAL WORK

Materials

The powder was supplied by the British Aluminium Company in the form of atomized and horizontally air blasted Al-Mn powder, nominally containing 3.5 per cent manganese. The size distribution was determined using standard 8 in (200 mm) BS sieves stacked on an 'Endrock' mechanical shaker and sieving for a specified time (1 hour for a 175 g sample). Particles below the 45 μm working limit of the conventional wire sieves were measured using a microscope and examining 75-100 particles. The results are shown in Table 1. Samples for analysis by

TABLE 1 Size distribution of Al-Mn Alloy Powder

Particle size μm	Weight (per cent)
+ 250	7.15
-250 + 150	35.8
-150 + 105	11.07
-105 + 75	10.12
- 75 + 63	14.6
- 63 + 45	9.15
- 45	12.3

the Imperial College Analytical Services Laboratory were taken by a coning and quartering technique. Their natural oxide contents were determined by a standard wet analysis method¹. The major and minor constituents of the powders were determined before oxide analysis and the results are shown in Table 2.

TABLE 2 Chemical analysis of powder

Particle size μm	Chemical analysis weight per cent					
	Al ₂ O ₃	Si	Mg	Fe	Cu	Mn
+ 250	0.68	0.13	tr	0.47	0.01	3.49
-105 + 75	0.75	0.19	tr	0.45	0.01	3.59
- 75 + 63	0.95	0.21	tr	0.49	0.01	3.57
Dust	0.95	0.09	tr	0.49	0.01	3.59

Powder compaction

It is not really essential that powders be compacted before extrusion, but handling at the press is facilitated if the billet is pre-compacted prior to heating. Accordingly, all billets were compacted to 85 per cent theoretical density before heating.

Extrusion

Extrusion was performed on an ENEFCO 5 MN hydraulic extrusion press having a container bore of 76 mm nominal diameter. The details of the equipment and instrumentation have been reported in a previous paper². The dies were square edged with reduction ratios of 5, 10, 20, 30, 40 and 50 and it was found that the easiest method of obtaining a satisfactory surface finish was to place a commercially pure aluminium pad in front of the billet and to perform the extrusion unlubricated. The billets were air heated in a furnace adjacent to the press to temperatures varying between 250 and 450°C. The temperatures were corrected for cooling between furnace and extrusion press, but in any case transfer was rapid, taking between 10 and 15 seconds. The temperature at the beginning of extrusion was also checked using a thermocouple threaded through the stem and dummy block located on the rear of the billet.

Preparation of specimens for mechanical testing and microscopy

The structure and properties of an extrusion vary along its length and to ensure some degree of uniformity during mechanical testing and microscopic investigation, it was necessary to extract specimens for hardness testing, electron microscopy and tensile testing from the same region in each extrusion.

For electron microscopy, longitudinal and transverse thin sheets were cut and thinned using a window technique by electrolytically polishing in a solution of 40 per cent acetic acid, 30 per cent orthophosphoric acid, 20 per cent nitric acid and 10 per cent water, at 20 V and 0°C. The specimens were examined in an AEI EM6G 100 kV electron microscope equipped with a tilting stage.

Tensile tests were carried out at room temperature and at 400°C on extruded rod machined to a Hounsfield specimen size (No. 14). The tests were performed on an Instron tensile testing machine equipped with a split furnace for the high temperature tests.

Microhardness tests were made on a standard Vickers machine using a 2 kg load. Specimens for hardness testing were prepared by electrolytic polishing and etching and measurements were made on transverse sections at intervals of 1 mm.

RESULTS AND DISCUSSION

Extrusion processes

The load-displacement diagrams for two different powder materials are compared with that of a commercial-purity cast billet in figure 2 which shows output traces from an X-Y recorder. It is clear that the powder billets exhibit a significant difference in shape than that of the cast material. The diagram for

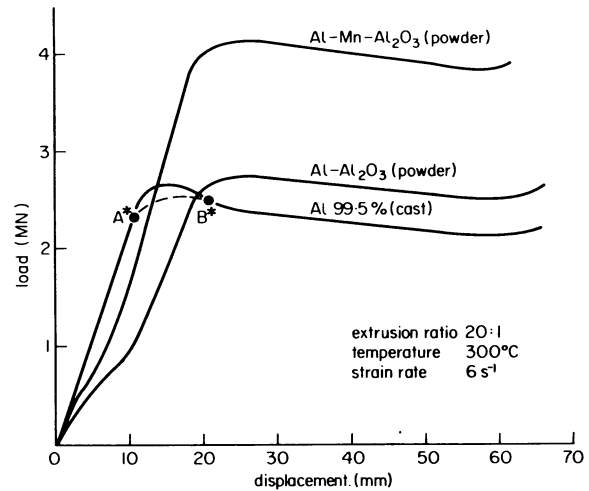


Figure 2. Load displacement diagrams.

these materials may be divided clearly into two regions: a compaction zone and a normal extrusion zone. In the compaction zone there is a non-linear increase in load until the billet attains 100 per cent density. This particular observation was checked by stopping several extrusions at this point and measuring the density. The extrusion zone is very similar to that of the cast material except for one important feature: the extrusion peak observed in the cast billet is not reproduced on the powder compact extrusions. This peak, represented by the curve A*B* in figure 2, is a characteristic of most extrusions and the reasons for its existence are not really understood. It has been suggested that the peak is a result of die-quenching of the billet, but little evidence has

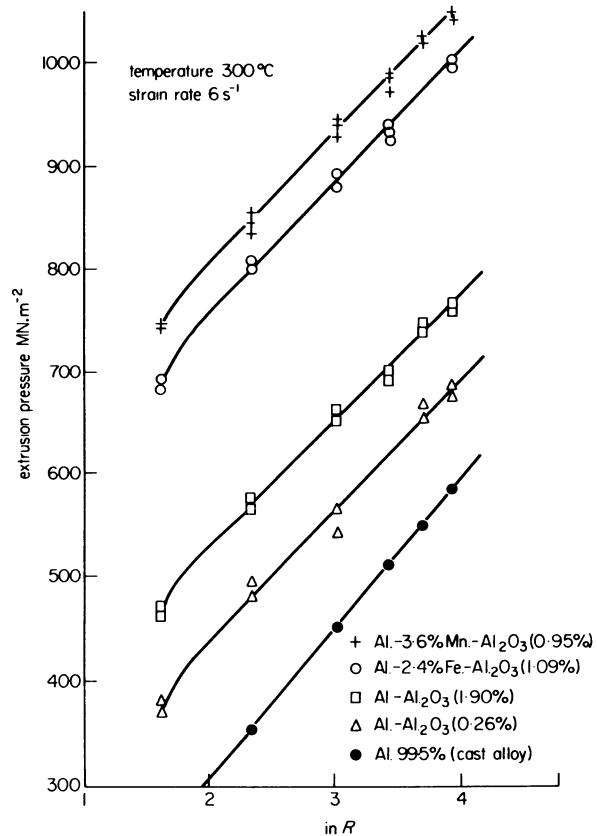


Figure 3. Extrusion pressure as a function of extrusion ratio.

TABLE 3 Pressure-extrusion ratio formulae

Powder	Extrusion Law
Al-3.6 per cent Mn-0.95 per cent Al ₂ O ₃	$p = 555 + 252.5 \ln R$
Al-2.4 per cent Fe-1.09 per cent Al ₂ O ₃	$p = 475.5 + 252.5 \ln R$
Al-1.9 per cent Al ₂ O ₃	$p = 277.5 + 252.5 \ln R$
Al-0.26 per cent Al ₂ O ₃	$p = 187.5 + 252.5 \ln R$
99.5 per cent Al (cast)	$p = 20 + 287.5 \ln R$

been presented and the effect is noticed in extrusions of low temperature where die and billet temperatures are almost identical. In the figure shown, for example, the billet, die and container were all at 300°C. The peak is certainly not purely a friction effect as it has been shown⁵ to occur in indirect extrusion where frictionless conditions exist. Hence, although the reasons for this load increase are not understood, it is known that pressures required for extrusion may be increased by 10-15 per cent because of this phenomenon. Thus, its elimination by the simple expedient of using a powder compact may be important.

Figure 3 shows a standard plot of extrusion pressure against $\ln R$ for a number of powder compacts and for a commercial-purity cast billet. It is not possible to make the vertical axis non-dimensional in the normal way by plotting pressure/flow stress because, as explained in the theory section, the rheological properties of the powder material must change throughout the deformation zone. It is clear, however, that the powder material obeys the same basic pressure-extrusion ratio laws as those for a cast material. Thus from Figure 3 we may write these laws as shown in Table 3.

These linear relationships do not hold at low extrusion ratios for the powder compacts because, as figure 3 indicates, there appears to be an extrusion ratio below which fully coherent material cannot be produced. This extrusion ratio lies between 5 and 10 and hence could not possibly be of industrial interest. The significant feature revealed by Table 3 is the relative magnitudes of the component terms of the equation. It is agreed that as far as the 99.5 per cent Al material is concerned, the first term will represent redundant work and the second term the homogeneous work of deformation. The implication, therefore, is that for the powder compact, the redundant work is a greater proportion of the total energy required for deformation. The theoretical section has in fact indicated that the work of homogeneous deformation will be lower and it is clear that in some way the material has to be built during extrusion. Thus, referring again to figure 1, the material approaching the first velocity discontinuity AB will be weak in shear and inter-particle bonds constructed during the compaction phase will easily be broken. The shearing action will, however, extend the area of metal/metal contact and re-welding will establish a slightly stronger weld neck. At the discontinuity OC, the powder has obtained final property values and normal shearing of the metal can occur. The *modus operandi* is of course much more complex and involves a continuous breaking and re-welding of powder contacts, gradually increasing the area of

contact until the material can shear along preferentially orientated planes as in conventional extrusion. There may also be a small 'internal friction' contribution to the 'redundant' portion due to abrasion of the powder particles but this can only be a minor effect. Thus we may suggest that this increase in the proportion of redundant work contributes significantly to the building of the powder material and is an essential feature of the powder extrusion process.

It is not clear from either figure 2 or figure 3 that the extrusion pressure is indeed lower for this process route, because it is not possible to produce finished material having the same properties. Even for nickel-rich alloys⁶, the final extrude properties are improved because material cooled by atomization, even under atmospheric conditions, is more homogeneous than its cast counterpart (nevertheless the actual extrusion pressures for these alloys are considerably reduced). Thus it is necessary to relate the final material properties to the extrusion pressure in order to obtain a true picture. Figure 4 shows the extrusion pressure divided by the 'as extruded' proof stress against the extrusion ratio and it is immediately apparent that the powder alloys require far less extrusion pressure

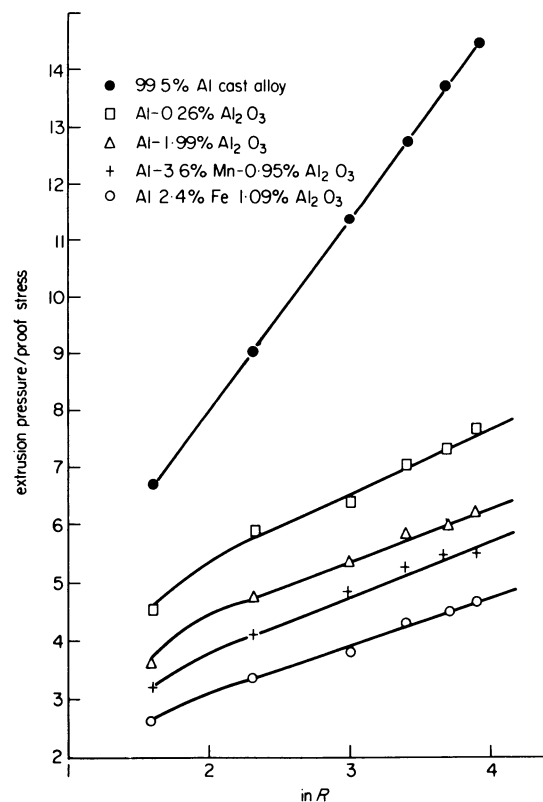


Figure 4. Extrusion pressure/proof stress as a function of extrusion ratio.

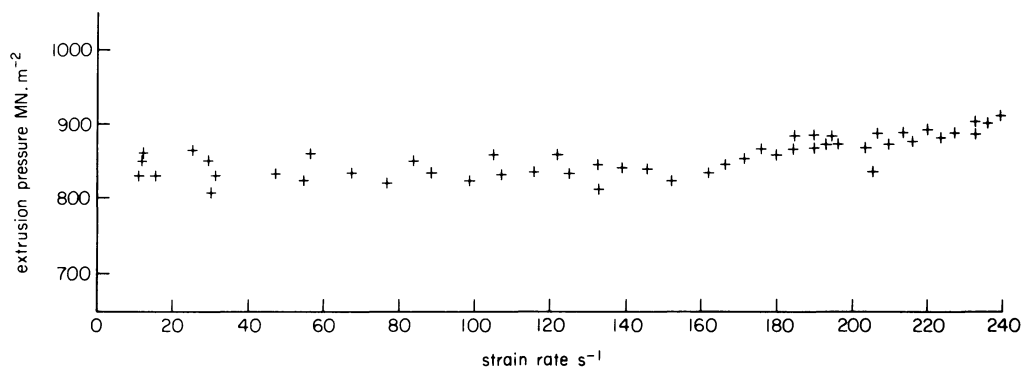


Figure 5. The effect of strain-rate upon extrusion pressure.

per unit of strength. It is also evident that powder products must not be produced below the threshold values of extrusion ratio previously produced.

An attempt was made, using the Al-Mn-Al₂O₃ powder, to investigate the effect of strain rate upon the extrusion pressure. Figure 5 shows fairly conclusively that all pressures recorded are within experimental error and that the powder extrusion process is essentially not strain rate sensitive. This is perhaps not surprising because it seems that the process is

dependent upon the strength of weld necks and local diffusion rates, rather than on finally developed flow stress. This must clearly be an advantage as production rates using such a process would appear to have no upper limit. However, it is equally clear that, initially at least, the process must be developed for special alloys having little use for high productivity.

Figure 6 shows that the extrusion pressure required is much higher for small particles than for large particles. This suggests that the 'redundant' work portion of the total pressure is higher. This would appear to be logical because the total surface area of the powder is greater and inevitably a greater number of weld necks must be formed on compaction, broken in shear and re-welded etc. However, there will clearly be an increase in the logarithmic term of the pressure equation, because greater surface area results in a higher oxide content, more dispersion hardening and higher proof stress in the final product.

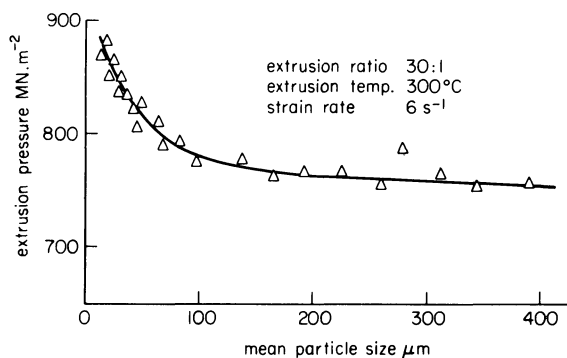


Figure 6. The influence of particle size on extrusion pressure.

Mechanical properties and structure

Figure 7 shows the variations in yield strength and elongation with mean particle diameter, both at room temperature, and at 400°C. The results are similar to those reported for aluminium powder, but the pro-

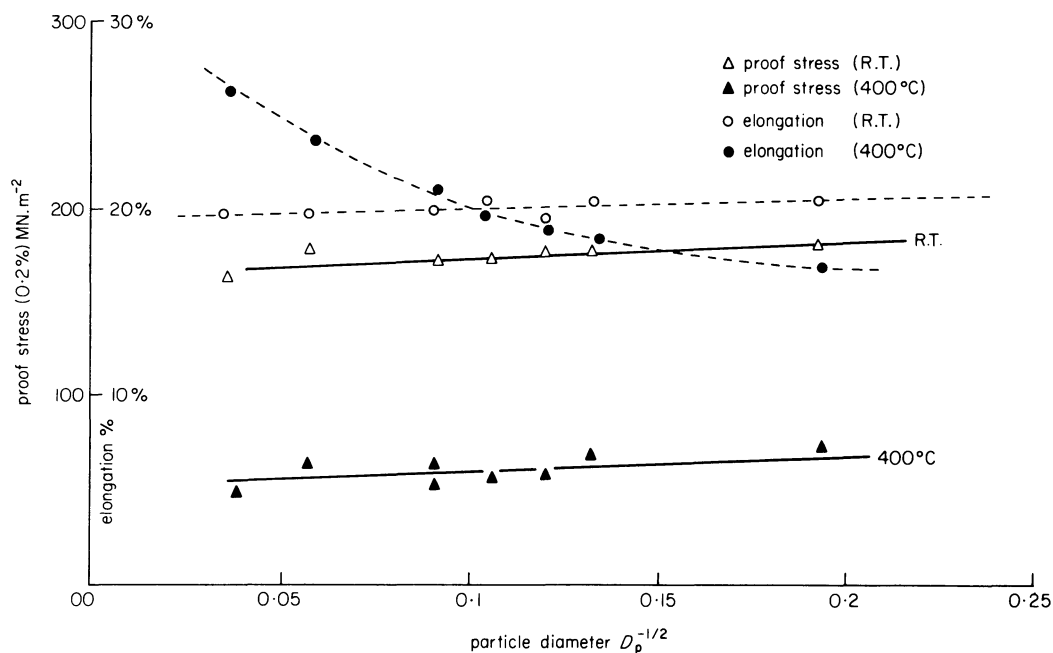


Figure 7. Variation of mechanical properties with particle size.



a



b

Figure 8. Structure of the initial powder: (a) $\times 90$ (b) $\times 800$.

properties are much less dependent upon particle size. The Al-Mn-Al₂O₃ alloy contains less oxide than present in the commercially pure Al powders and it is evident that final properties are much more dependent upon structure and shape of the particle, than on particle size. In other words, the dispersion strengthening effect is considerably reduced. This was a characteristic noted in all the alloy powders processed. The shape and structure of the powder is shown in figure 8, which indicates that powder produced in this way by horizontal air-blasting has a greater number of tails, than powder produced by a vertical blast; there also appears to be a larger proportion of powders having this 'tadpole' form. Figure 8(b) reveals the presence of a very fine structure, dendritic in nature, due to the high solidification rate in the atomizing process. There is also evidence of large primary precipitates and the size of these in relation to the fine unresolvable background indicates that in some cases the primary phase is precipitated out before the formation of the particles in the air stream. It is possible that small precipitates of MnAl₆, formed as the liquid temperature falls below 700°C, have acted as nuclei for primary precipitates in the large particles. Thus, the cooled material is a super-

saturated meta-stable alloy, containing the normally coarsely precipitated second phase, either in solid solution or in fine precipitate form. Thus we might expect the dispersion strengthening effect of any Al₂O₃ to be significantly less.

Figure 7 also shows that a Petch-type plot may be applied to powder material using the particle size as the grain size analogue. The equation may be written

$$\sigma = 147.6 + 152.5 D_p^{-1/2} \quad \text{at room temperature}$$

$$\sigma = 52.63 + 65.9 D_p^{-1/2} \quad \text{at } 400^\circ\text{C}$$

where σ is the proof stress and D_p is the particle size.

The initial billet temperature determines the strength ductility characteristics of the alloy as shown in figure 9. The room temperature proof stress falls considerably with increase in the extrusion temperature and the elongation is also controllable by the extrusion starting conditions. Thus a large range of material properties may be obtained by controlling the metallurgical process parameters.

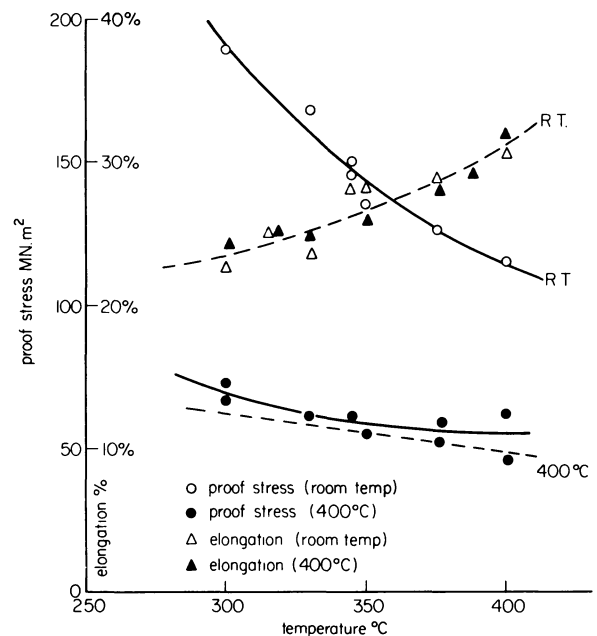


Figure 9. Variation of mechanical properties with extrusion temperature.

Table 4 shows a comparison of the properties of the powder materials produced with those of conventional alloys. It is clear that both high and low temperature tensile properties are far superior in the powder produced products and that quite simple alloys produced in this way could compete with the complex conventional alloys. Ductility although lower is still acceptable and much higher than in normal powder metallurgy parts. Further testing is required to determine the comparison of fatigue, impact and stress corrosion properties.

The electron microscope pictures of the Al-3.5 per cent Mn-1.09 per cent Al₂O₃ alloy (figure 10) show the presence of large primary precipitates at sub-grain or grain boundaries formed before the eutectic. The large precipitates, clearly seen in figure 10(a) are the nuclei for the sub-grains formed during the extrusion process and the small precipitates observable are

TABLE 4 Comparison of alloy properties

Material	At room temp.		At 400°C	
	UTS N/mm ²	Elong. per cent	UTS N/mm ²	Elong. per cent
Al-1.9 per cent Al ₂ O ₃	160	25	59	19
Al-0.36 per cent Al ₂ O ₃	122	36	29	29
Al-3.6 per cent Mn-0.74 per cent Al ₂ O ₃	208	27	64	45
Al-2.4 per cent Fe-0.51 per cent Al ₂ O ₃	276	16.6	92	24
99.5 per cent Al (conventional)	77	45	—	—
Al-1.25 per cent Mn (conventional)	108	40	18	70
Al-Cu-Mg-Si (conventional)	232	15	31	100

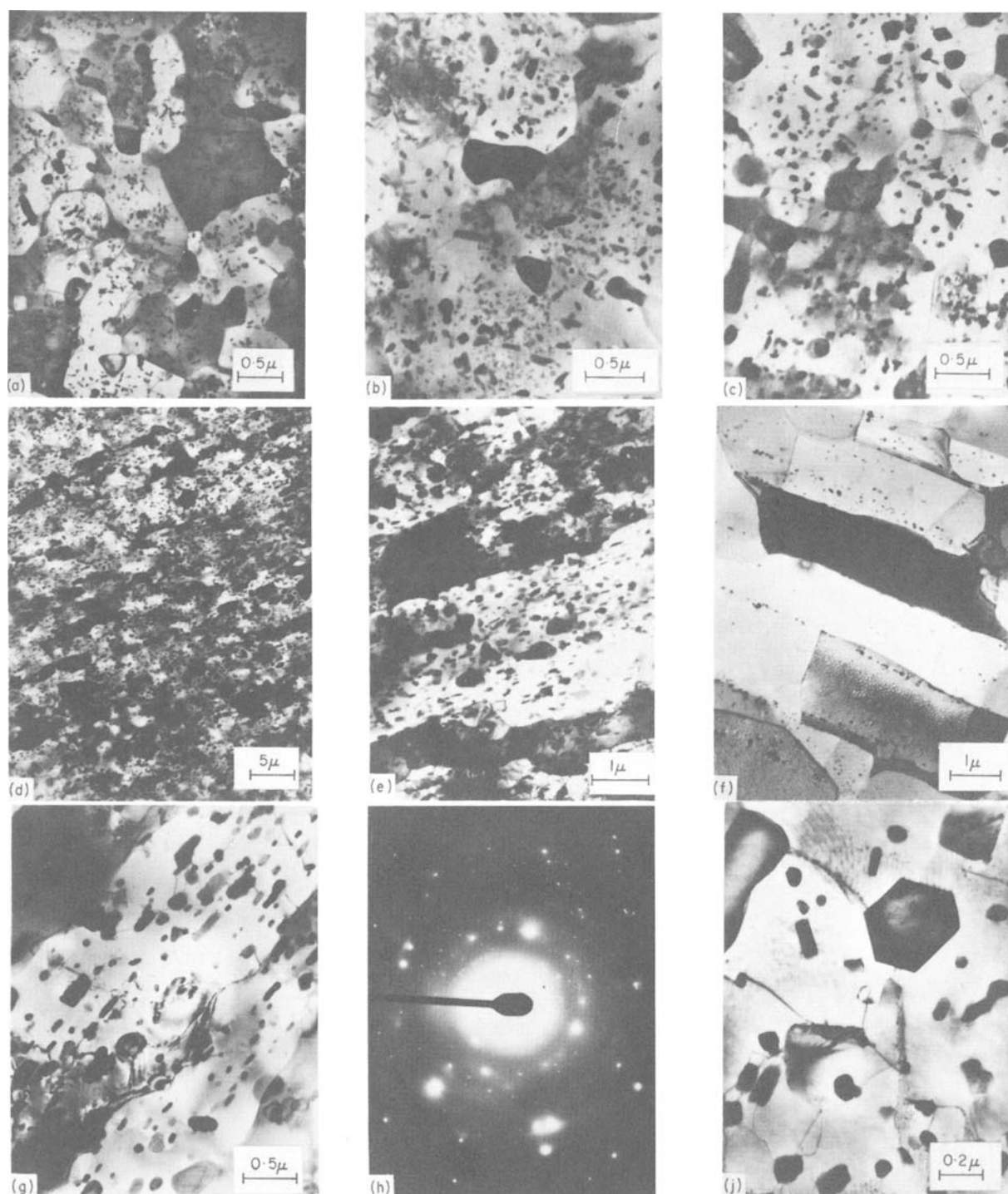


Figure 10. Electron micrographs of extruded sections (Al-3.6 per cent Mn-0.95 per cent Al₂O₃): (a) Transverse section, $R = 30 : 1$, $T = 345^\circ\text{C}$ (b) Transverse section, $R = 30 : 1$, $T = 360^\circ\text{C}$ (c) Transverse section, $R = 50 : 1$, $T = 405^\circ\text{C}$ (d) Longitudinal section, $R = 30 : 1$, $T = 340^\circ\text{C}$ (e) Longitudinal section, $R = 30 : 1$, $T = 350^\circ\text{C}$ (f) Al-1.9 per cent Al₂O₃ alloy transverse section, $R = 30 : 1$, $T = 250^\circ\text{C}$ (g) Longitudinal section, $R = 30 : 1$, $T = 350^\circ\text{C}$ (h) SADP of large precipitate (j) Dislocation network, $R = 30 : 1$, $T = 380^\circ\text{C}$.

those from the eutectic phase and precipitation from solid solution. During extrusion the super-saturated solid solution is relieved of the excess manganese by a precipitation similar to that occurring in an age-hardening system. The manganese of the super-saturated solid solution is present in small clusters that cause considerable lattice distortions and the process of plastic deformation causes the formation of a more stable phase by the actions of thermal energy and lattice strains. The small precipitates from the solid solution can be seen in figure 10(b) and (c) in conjunction with the large primary and equilibrium precipitates. The longitudinal micrographs (figure 10(d) and (e)) show that any large precipitates are broken up and subsequently elongated in the extrusion direction. The extremely large precipitates observed in the optical microscopy of the initial powders cannot be detected and must be assumed to have been disintegrated during extrusion.

Oxide particles which are easily identified in Al-Al₂O₃ alloys (figure 10(f)) cannot easily be detected and must be finely dispersed, as is the precipitated Mn phase from the solid solution. Figure 10(b) shows these smaller precipitates and oxide inclusions, together with the smaller Al-Mn precipitates; in the top left-hand corner can be seen a region of meta-stable precipitates from the super-saturated solution. There is no evidence of any interaction between the precipitate phase and the oxide dispersion and it would appear that the increased properties of the alloy are due to the separate actions of the oxide and precipitate dispersions. The better high temperature properties, compared with aluminium powders, are proof that the Mn precipitates have had insufficient time to coalesce. Figure 10(g) provides evidence of dislocation interaction with the primary precipitates; dislocations formed during the deformation process being arrested in motion by these obstacles and strengthening the final product. The structure of the large primary precipitates can be identified as MnAl₆ using the SADP shown in figure 10(h).

Dislocation networks, occasionally observed in the Al-Al₂O₃ system, are quite readily formed in the Al-Mn-Al₂O₃ alloy (figure 10(j)). There is, therefore, a strong possibility that these networks could have extended across the sub-grain during deformation and the microstructure shown would therefore be the remnants of the unpinned dislocation network.

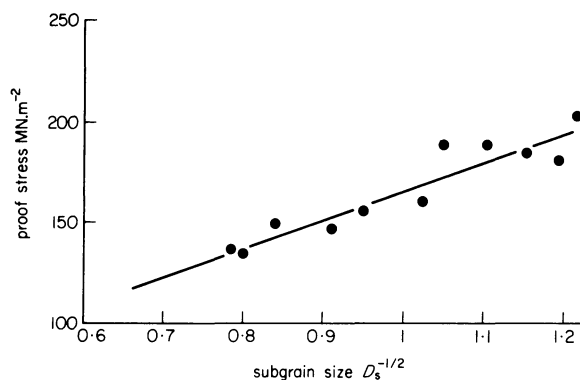


Figure 11.

In general, alloy systems will have a structure dependent upon the state of the solute present. In regions where primary precipitates exist, a region of recrystallized grains may be expected, whilst in areas of solid solution, recrystallization will be retarded. Thus, in figure 10(d), solid solution regions can clearly be identified and the presence of recrystallized grains associated with large particles present before the extrusion may be identified in figure 10(a).

It has been shown that a standard Hall-Petch plot may describe the relationship between the proof stress and the sub-grain size in cast and wrought materials. Figure 11 indicates that this also applies to materials produced from metal powders, such that we may write

$$\sigma = 32.5 + 121 D_s^{-1/2}$$

where σ is the proof stress and D_s is the sub-grain size.

CONCLUSIONS

- (1) The 'peak' or 'break-through' pressure normally required to initiate extrusion appears to be eliminated when extruding powdered material.
- (2) The extrusion pressures per unit of room temperature proof stress, required for powdered compacts are considerably lower than for cast materials.
- (3) The 'redundant' work term in the pressure equation is comparatively larger, but it is suggested that this energy is utilized to 'build' a coherent material.
- (4) There is a transition extrusion ratio below which coherent materials may not be produced.
- (5) The powder extrusion process is probably not as strain-rate sensitive as the normal extrusion process.
- (6) Material tensile properties are in general superior, both at room and high temperatures, to those of conventionally produced materials. However there is some deterioration in elongation figures.
- (7) The rapid cooling associated with powder atomization affects the precipitate.
- (8) Large precipitates are broken and elongated during extrusion.
- (9) There is no evidence of interaction between the Al₂O₃ dispersion and the precipitate in the Al-Mn-Al₂O₃ alloy.
- (10) The evidence presented suggests that standard Hall-Petch type plots may be applied to both grain and sub-grain plots.

REFERENCES

1. T. Sheppard and P. J. M. Chare. *Powder Metallurgy* (1972) 16, 17.
2. P. J. M. Chare and T. Sheppard. *Powder Metallurgy* (1973) 16, 437.

3. T. Sheppard and D. Raybould. *J. Inst. Metals* (1973) **101**, 33.
4. W. Johnson and H. Kudo. *The Mechanics of Extrusion* (1962) Manchester University Press.
5. W. Ziegler and K. Siegert. *Zeit Metallkde* (1973) **64**, 224.
6. T. Sheppard. *Proc. Int. Conf. Heat Resisting Steels* (1972) Dum Techniky CVTS Ostrava, p. 162.

COMPARISON BETWEEN UNIAXIAL AND ISOSTATIC DENSIFIED STAINLESS-STEEL POWDER PREFORMS FORGED IN A CLOSED DIE

by

H. ALTMANN*

SUMMARY

Isostatic compacts are known to have a more constant distribution of density and hardness than uniaxial densified compacts. In experimental comparison with uniaxial densified powder, compacts which were densified with the isostatic wet bag method did not show such constancy. In the compact both methods of powder densification generate fluctuations of density and hardness which could not be compensated by closed die forging. The inhomogeneities of the compacts together with the forging process determine the distribution of the mechanical properties in the final forging. Density and hardness distributions of preforms and forgings are represented over cross sections. The load and energy requirements necessary to produce a certain final density in the different forgings can be obtained from another graph. It is also shown that the tensile strength depends both on the direction in the forging and the length of the forged pin as well as the density and compacting method.

INTRODUCTION

The hot forging of powder compacts has developed from conventional sinter engineering. At first the compacts were repressed after sintering to lift the density. Nowadays the sintering is being replaced by a short period of heating to forging temperature, which is immediately followed by the forging process. This paper deals with the problems connected with this forging method.

Because the mechanical properties of P/M parts increase with greater density, the forging process took place at a high temperature in a closed die. The high generated tool pressure enabled relative densities up to $\rho = 0.98$ to be reached. There was only a small resemblance between the shape of preform and forgings, which means that new edging took place. In contrast to this method R. Davies and J. B. Marx¹ describe a procedure with a great similarity between preform and forging. The process carried out was applied to compacts which had been produced

- (1) by uniaxial densification;
- (2) by the isostatic wet bag method.

P. C. Eloff and L. E. Wilcox² produced pinions with a similar process, but their preforms were densified by the isostatic dry bag method.

EXPERIMENTAL PROCEDURE

Powder

The powder used is a water vaporized pre-alloyed stainless steel powder produced by Mannesmann-Werke. Its properties and chemical analysis can be seen in figure 1(a). On account of the high proportion of alloying metals of 18–20 per cent chromium and 9–11 per cent nickel it is an austenitic steel powder. A steel of this alloy shows great resistance to aggressive mediums at room temperature, together with good mechanical properties.

Both in forged austenitic chromium–nickel steels and in forged powder compacts of the same alloy it is possible after etching to recognize the typical metallographic twin formation. From this formation and the particle deformation one can conclude the flow behaviour of the material. In view of the short period of heating no lubricant was added to the powder. All samples weighed 400 g.

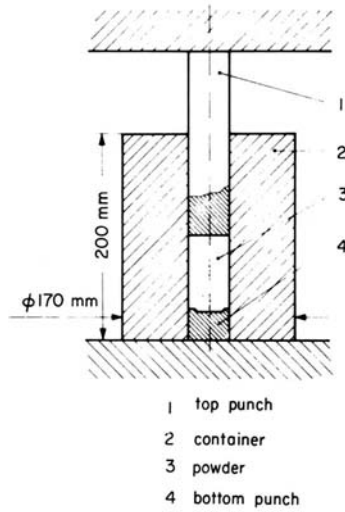
Uniaxial powder densification

In the well-known compacting method illustrated in figure 1(b), powder was compacted with the help of a top punch (1), a container (2) and a bottom punch (3). This uniaxial compaction is influenced by the friction at the inner tool walls, but demands little technical means. The tool can be put in nearly every

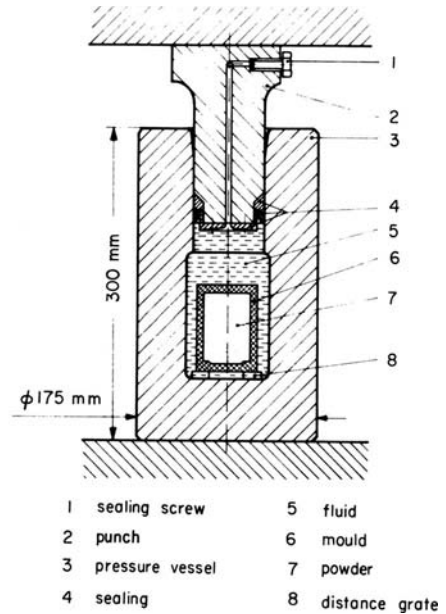
* Institut für Umformtechnik und Umformmaschinen, Technische Universität, Hannover

pre-alloyed stainless steel powder Fe-Cr-Ni	
	contents per cent
C	<0.08
Si	0.9-1.2
Mn	0.20
P	0.020
S	0.020
Cr	18-20
Ni	9-11
O	—
H ₂ -loss	—
O tot.	0.30
particle size	<150 μm
apparent density	270-2.95 [g/cm ³]

(a) chemical analysis



(b) uniaxial powder compaction



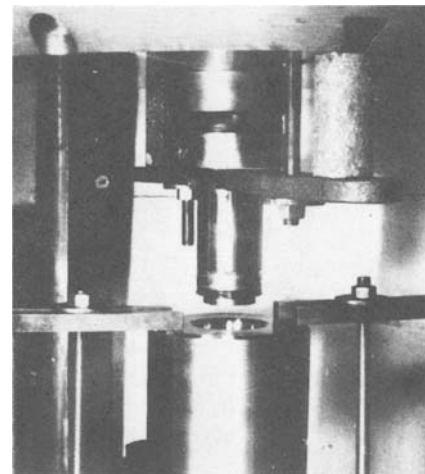
(c) isostatic wet bag powder compaction

Figure 1.

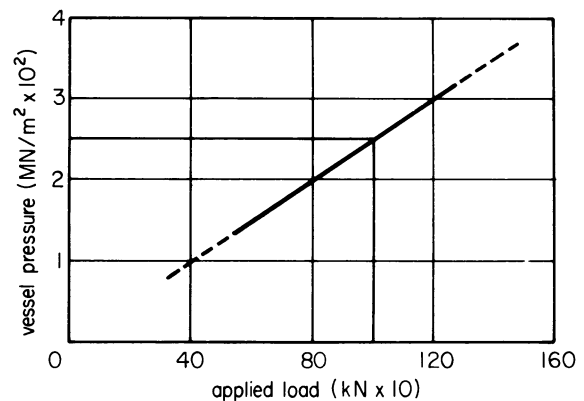
hydraulic press. The experiments were carried out using a Sack and Kiesselbach hydraulic press with a pressing speed of 0.04 m/s. The penetration depth of the top punch in the container determines the average density of the uniaxial compact. All compacts had the same relative density of $\rho = 0.75$.

Isostatic powder densification

S. Mokarski and P. C. Eloff³ describe the isostatic wet bag method in which the fluid pressure acts from all directions on the enclosed powder. Such a pressure can be generated by the isostatic appliance shown in figure 1(c). The pressure is not generated by a pump but by a punch. The advantage of this procedure is that the isostatic appliance may be put in the same press which was employed for uniaxial powder compaction. The prepared sample is immersed in the vessel and the punch is brought down. Air is allowed to escape from the vessel by opening the ventilating screw until fluid comes out. The sealing consists of an elastic seal for annular grooves for the low-pressure range, and a copper-beryllium ring with a triangular cross-section for high pressures. This metal ring expands and closes the clearance between punch and vessel cylinder. The rubber mould for the powder is cast from silicone rubber. The inner diameter and height of the mould are determined in experiments so that a compact is produced at a certain vessel pressure with the same geometry and density as a uniaxial densified compact. The relationship between vessel pressure and applied load is shown in figure 2(b). The photograph in figure 2(a) shows the isostatic appliance built into the hydraulic press.



(a)



(b)

Figure 2.

Forming geometry

Each of the preforms compacted in the two kinds of densification was forged to three pieces with different

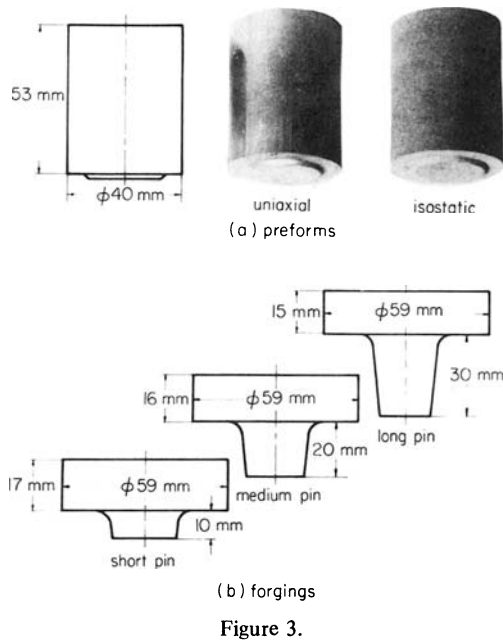


Figure 3.

pin lengths. Figure 3 shows the preform and forgings. Except for the rough surface of the isostatic preform in figure 3(a) there is no visible difference between it and the uniaxial preform.

The three forgings in figure 3(b) show that by reason of the edging the height of the head decreases with increasing pin length. Because of decreasing volume during forging, it is necessary to define a new ratio of deformation ϕ . In contrast to theoretical forming laws, the logarithmical ratio was based on the change in plane area during formation. The value of this ratio for all experiments was constant, $\phi = 0.8$. Every length of pin had been forged from both uniaxial and isostatic preforms; the two types will be called uniaxial and isostatic forgings respectively.

Inductive heating

After powder densification the compacts were given no special sintering treatment, but a short period of heating to forging temperature in an induction plant. This arrangement was built by the AEG-Elotherm and worked at 4000 Hz and 44 kVA. A compact to be heated was put in a wrought tube of non-scaling steel and inserted in the coil, as shown in figure 4(a).

The induction heating depends on the density of the inserted material in such a way that materials of greater density will be heated faster. On account of its greater density the wrought tube is raised quickly to higher temperatures and therefore accelerates the heating of the enclosed compact. The forging temperature was 1200°C for all experiments. The forging temperature was measured indirectly by gauging the heating time. Some uniaxial and isostatic compacts were drilled in the direction of the longitudinal axis as far as the centre for the insertion of thermocouples. At a given energy level of the induction plant the wrought iron tube with compact and thermocouple was inserted in the induction coil and the temperature was recorded at intervals. The curve of temperature against heating is shown in figure 4(b). It was found that for a given energy level there

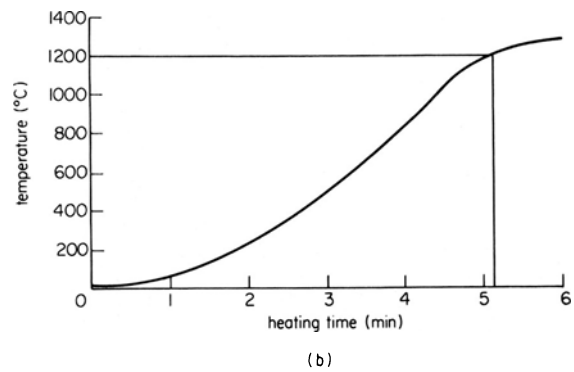
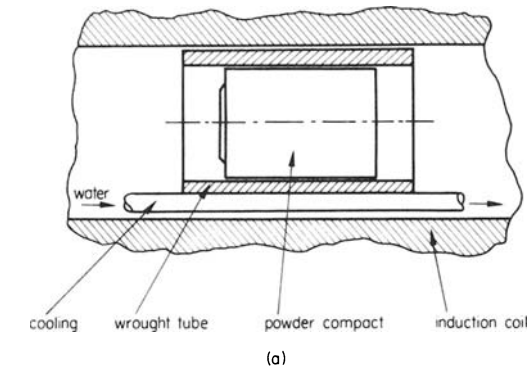


Figure 4.

was no difference between the two kinds of densification. The greatest difference in the time taken to reach 1200°C was ± 5 s. Therefore all compacts were heated at the same energy level of 55 per cent and taken from the coil after 5 min 10 s.

Forging process

In a closed die high pressures can be produced which are advantageous for the densification of compacts. The forming machine used was a mechanical crank press with a nominal load of 315.10 kN. The tooling design is illustrated in figure 5(a). The forging is squeezed out of the cavity by an ejector to minimize the contact time between forging and tool. Three ejectors with different cone lengths are provided. The pin length of the forgings (see figure 3(b)) is determined by the choice of ejector. By altering the table height, the cavity volume and therefore the forging density can be changed. The changes in forging load and displacement with forging time were recorded with the help of a load cell (wire strain gauge) and an inductive position sensor. The average forging time was 0.1 s and the punch speed was 0.5 m/s. The forgings were cooled in a container with an argon atmosphere for protection.

A cast alloy of the same material as the powder compacts was also forged in the tool. Figure 5(b) shows the load-displacement curves for forging processes which were carried out with the same forging load.

The solid curve is valid for both the uniaxial and isostatic compacts; the broken line belongs to the cast alloy.

The cast alloy preform had the same weight and forming ratio of $\phi = 0.8$ as the compacts but a smaller

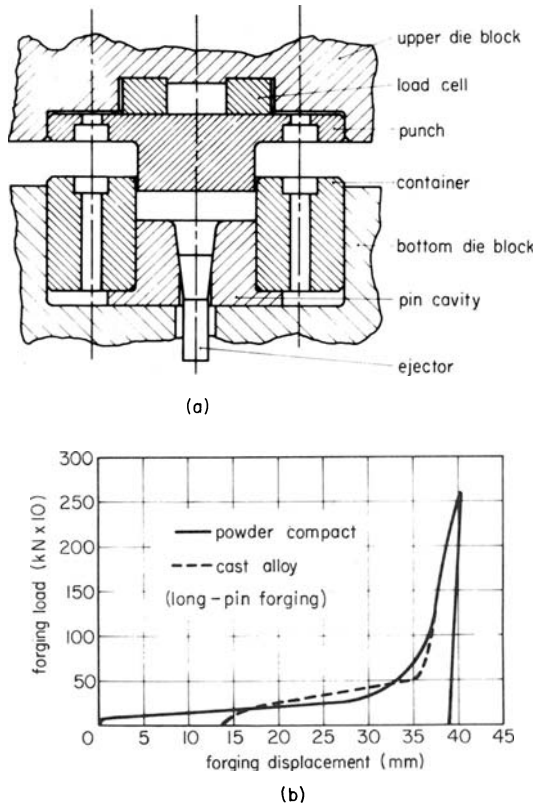


Figure 5.

initial height. This is the reason for the longer displacement of the compact.

The reason for the initial lower forging load in the case of the compacts in the first forging stage was the loosening of particle bonding which took place immediately after the punch had been lowered. In the draw-out stage there is no difference between the two curves. The energy requirement of the cast alloy was 10 per cent less than for the compacts. This is justified by the simultaneous occurrence of densification and forming of the compacts, resulting in great inner friction and slip-surfaces.

EXPERIMENTAL RESULTS

Forging load and forging energy

As shown earlier the uniaxial and isostatic compacts had the same form and the same average relative

density. Graphs were drawn for all forging operations to investigate the influence of forging load and forging energy on the achieved density of the forgings.

Figure 6(a) shows the average relative density for the three pin lengths as a function of the maximum forging load. With increasing forging load the density for all three pin lengths increases. Additionally it is clear that the greatest forging load is necessary to obtain a defined density in the long-pin forging. The curves for the uniaxial and isostatic compacts are similar. In the range of lower forgeability the requirement of forging load for isostatic compacts is lower for the medium and long-pin forgings.

Figure 6(b) shows the dependence of the average relative density on forging energy. For all lengths of pins the required energy increases with the density. According to figure 6(a) a smaller requirement of forging energy is necessary for short-pin forging than for the medium and long-pin forgings. Production of a forging with a long pin requires the greatest forging energy. Energy requirements for isostatic compacts are little smaller than for the corresponding uniaxial compacts. This means that less energy is necessary during forging of isostatic compacts than of uniaxial compacts to obtain a defined average density.

Density distribution

K. Morimoto et al.⁴ and T. W. Pietrocini⁵ point out that the mechanical properties of a forged compact depend very much on the distribution of density reached. In order to show a corresponding relationship for the pieces which were forged here in a closed die, an investigation into the density distribution must first be made.

For this reason some forgings were cut in pieces as shown by an example of a long-pin forging in figure 7. First the pin was sawn into discs; there were six discs in the long pin, four discs in the medium pin and two discs in the short pin. The head was cut by machining into $3 \times 4 = 12$ rings, which were concentric about the longitudinal axis. To determine their density the individual parts were weighed first in air and then, after paraffin impregnation, in water.

To determine the density of the compacts a sintering treatment was necessary because only in this way was it possible to divide a preform into

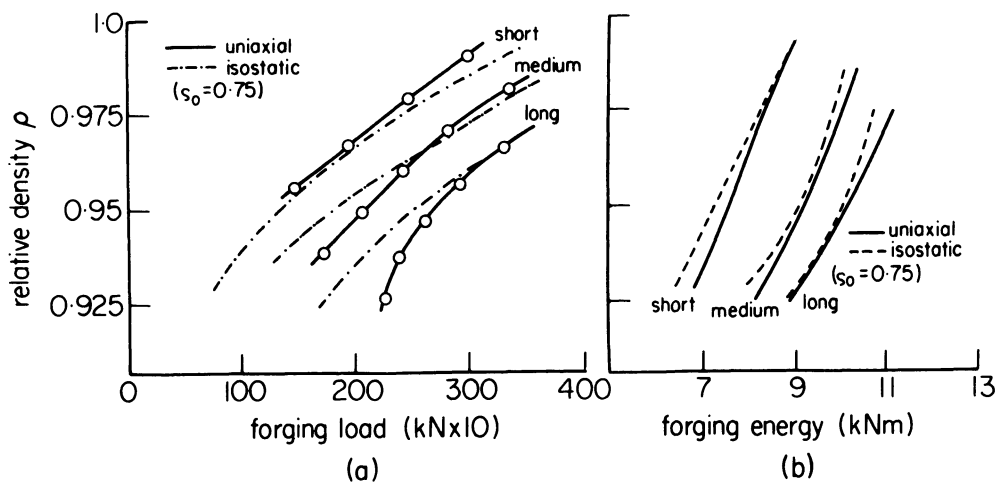


Figure 6.

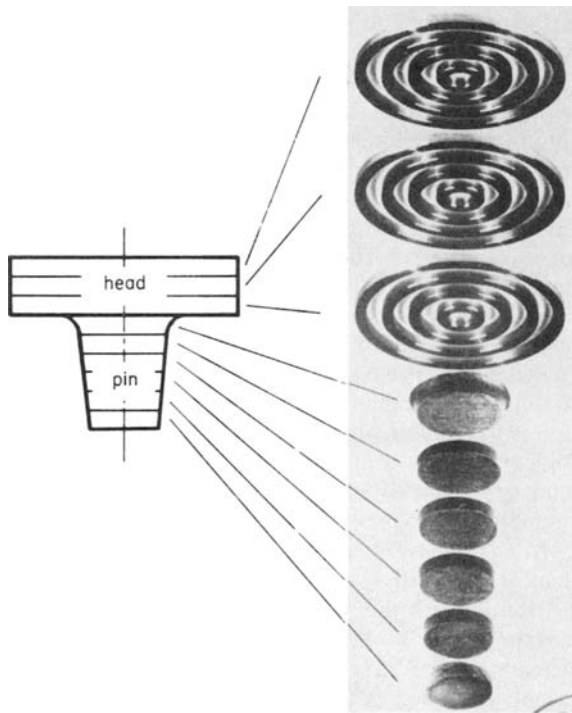


Figure 7.

6 × 3 = 18 rings. These rings were uniformly distributed in the cross section in the same way as for the head. For sintering treatment the compacts were enclosed in a metal tube which was evacuated before being closed by welding. The sintering temperature was 1200°C and the treatment operation took 90 min. During sintering treatment shrinkage occurs, resulting in enlargement of the compact's density.

The results of the density investigations are shown

in figure 8. The graphs show the distribution of relative density against radius of forging for (a) uniaxial compacts with the corresponding forgings and (b) isostatic samples. For the compacts and heads of the forgings one can see the corresponding scatter bands of measured density values; they are limited by straight lines.

No well-defined dependence of the density upon the sample height was proved and therefore only the distribution of density on the radius of the sample is shown.

Density values of the pins were measured by two discs situated next to each other, and are shown as broken horizontal lines across the ordinate. For this density distribution forgings were used which were forged by about the same maximum load of 330·10 kN.

The uniaxially-produced compact has a strong rise in density from the surface to the centre. This results from the wall friction which operates inversely to the uniform densification. The isostatic compact does not have this significant continuous rise. In the middle zone the density remains nearly constant in a wider range and decreases strongly only towards the edge.

For the short-pin forging made from a uniaxial compact, there is a steep slope of density distribution. In contrast to this the density of the short-pin forging made from isostatic compact increases towards the centre. It is remarkable that, in spite of inverse curves of density in the head body, both pins show similar values of density.

The medium-pin forging does not show such steep downward curves as short-pin forging. The uniaxial forging has an essentially lower slope of density in the

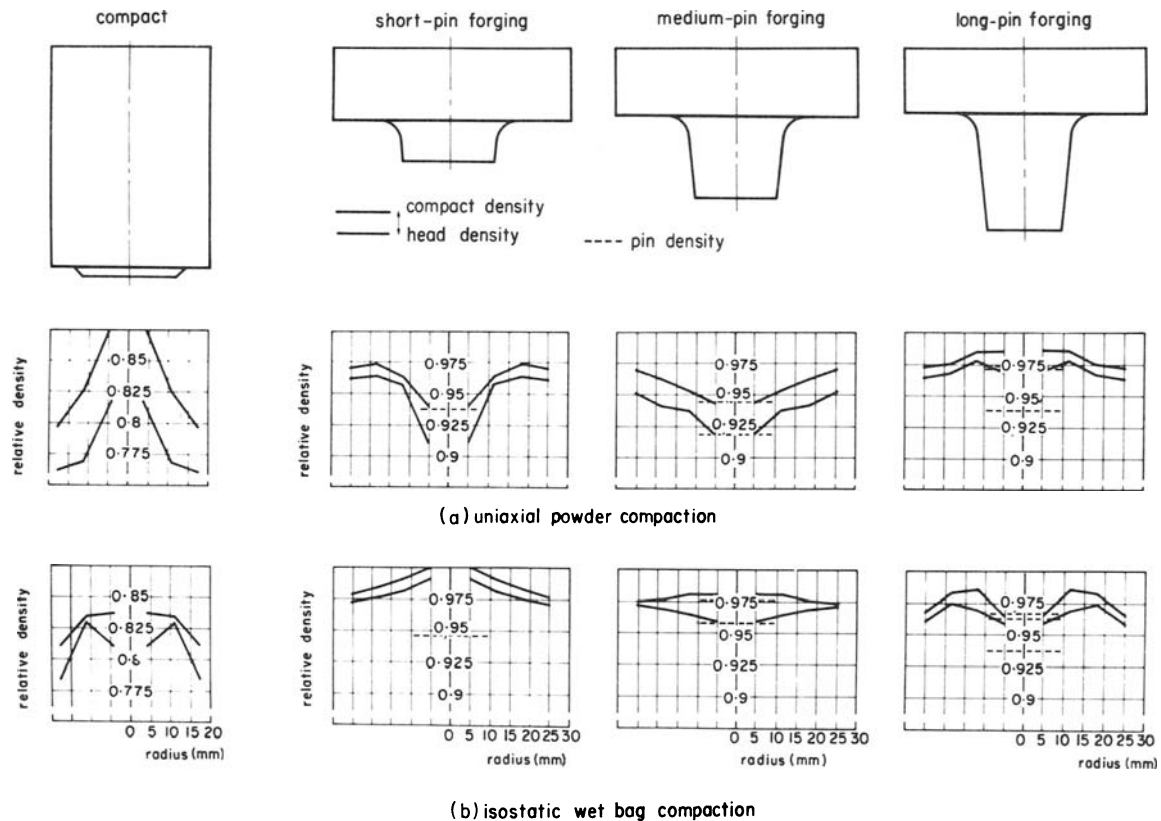


Figure 8.

centre of the sample. In the isostatic forging a rise in density can no longer be seen. Here we have a horizontal band.

The long-pin forging results in inverse curves of density compared with uniaxial and isostatic short-pin forgings. The density of the uniaxial long-pin forging increases slightly towards the centre of the sample. All values of density now lie in a narrow range, that is the curve of density against radius of forging is essentially more balanced than in the uniaxial short pin forging. In the case of isostatic compaction the long-pin forging first shows a rise of the relative density, as in the short-pin forging; in the centre, however, there is an inverse curve.

The investigation has shown that in uniaxial and isostatic forgings the direction of the curve of density against radius of forging is inverse to the increasing length of pin when the maximum forging load is nearly constant. In uniaxial forging the steep slope of density becomes a small rise towards the centre. In isostatic forging an initial rise in density is followed by a decrease in density in the centre zone. Distributions of density values of isostatic and uniaxial forgings are very similar in the case of long-pin forgings. This is reasonable because here the greatest motion and edging of material will occur. In the isostatic compact and the corresponding forgings there are not such great differences in density as in uniaxial samples.

Hardness distribution

The forgings and compacts used for investigations of hardness were cut by a diametrical longitudinal section and the sections were prepared for the Vickers hardness test by grinding and polishing. Investigations of hardness distribution of cold extruded powder preforms were made in the same way by T. Nakagawa et al.⁶.

A reference system was chosen in which, in an

analogous way to the procedure for density investigation, samples were forged by the same maximum forging load, in this case 260·10 kN. To include the powder compacts in this investigation a sintering treatment as described in the last section for the compacts was again necessary. The number of measuring points in the cross section of the sample was large enough to be able to establish iso-hardness contour levels. The results of the hardness measurements are given in figure 9. A definite value of hardness belongs to each contour; the difference between two contours next to each other is 10 Vickers units. The result of this examination is a field of contours in the cross section of the sample, which represents very clearly the distribution of hardness. For comparison contours of 60 HV-hardness (in the compact) and 190 HV and 230 HV-hardness (in the forging) are shown by the broken lines in figure 9.

The comparison of hardness distribution in the uniaxial and isostatic compacts shows a more uniform distribution in the uniaxial compact than in the isostatic one. In the uniaxial compact there is an increase of hardness towards the cylinder surface, except in zones near the face where contours of hardness are nearly constant in a radial direction. In the isostatic compact hardness remains constant in the middle zone over the cross section. Towards the surfaces, cone-shaped forms of greater hardness appear. This feature can be very clearly recognized in the upper zone of the compact. These zones must be regarded as typical for the isostatic wet bag compaction because all surfaces of the cylindrical compact were applied directly by compaction pressure. In dry bag compaction the round faces are applied by pressure indirectly, so one can expect a homogeneous hardness distribution over a bigger range of the compact height.

The figures of hardness distribution of the forgings show that new distributions of hardness are generated

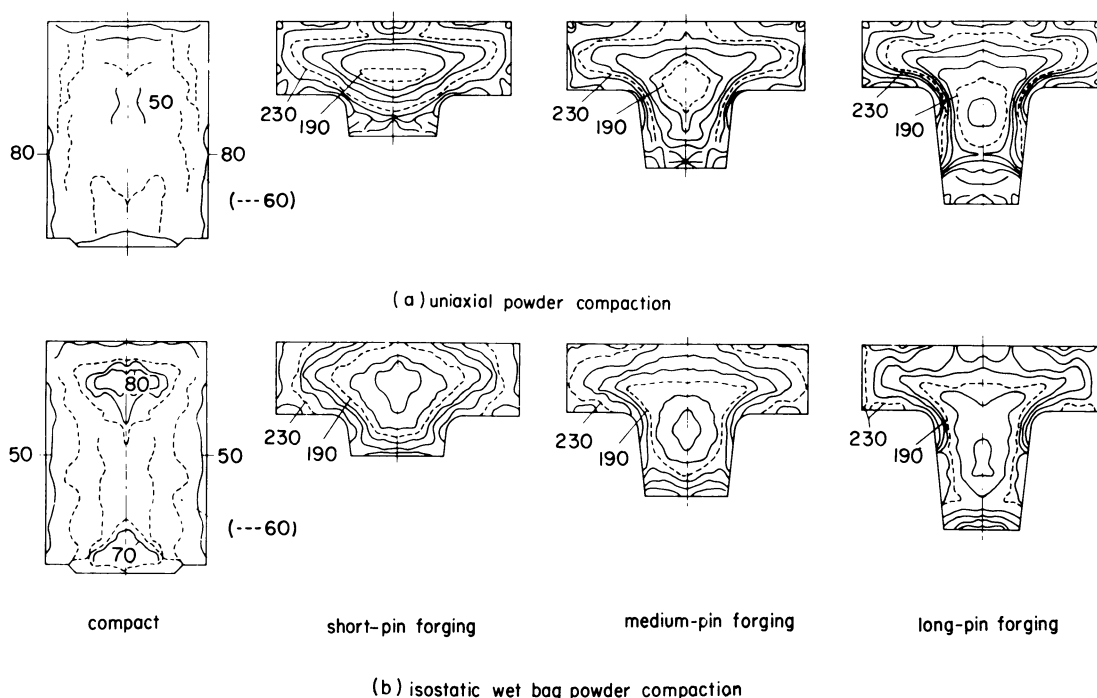


Figure 9.

by forging which is followed by slow cooling in an argon atmosphere. These distributions are uniform for all three pin lengths. In contrast to the compacts, the forgings show greater differences in hardness. No connection of density and hardness, as described in a report by H. Fischmeister et al.⁷ could be proved.

In the medium and long-pin forgings the less hard middle zone lies further in the pin of the forging than in the short-pin forging. This feature is clearer in the medium and long-pin isostatic forgings. In isostatic forgings hardness for all pin lengths is lower in the middle zone by 20 HV-units than in the uniaxial forgings. The smallest distances between hardness contours are to be found in the transition zone between the 'soft' middle zone and the transitional radius of the pin. This increase in hardness is most evident in the long pins. The greater part of the cross section is covered by closed contours of hardness. Towards the surface there are interrupted and irregular contours of hardness. This can be seen very clearly in the pin ends of the uniaxial forgings, where values of hardness vary very much. If one follows the contours of hardness in the longitudinal axis in the direction of the pin end, starting in the 'soft' middle zone, first the hardness increases in uniaxial as well as in isostatic forgings, and then decreases towards the pin end.

Tensile strength

K. H. Moyer⁸ and G. T. Brown⁹ point to the essential advantage of P/M pieces, which consists in the possibility of obtaining the same mechanical properties in directions which are perpendicular to each other. To examine the tensile strength of P/M-forgings in perpendicular directions, a small proportional test sample was designed with reference to DIN 50 125. The sample size was enough for a sample to be taken from the short-pin forging in the direction of the longitudinal axis. In this way it was possible to examine the tensile strength of the centre region of each type of forging in the directions of the longitudinal and transverse axes.

The measured values of the tensile strength for each type of forging are plotted in figure 10 against the average relative density. The longitudinal test samples of the medium and long-pin forgings were

taken from the punch side of the forgings. The transverse test samples were taken from the middle of the head.

Figure 10(a) shows the tensile strength values of the short-pin forgings. The values of tensile strength of the isostatic forgings are at a lower level for the longitudinal direction than in the uniaxial forgings. With increasing relative density the clearance becomes smaller. For the uniaxial medium-pin forging there is a great difference between the values of samples in the longitudinal and transverse direction. This difference is smaller in the case of the uniaxial medium-pin forging shown in figure 10(b). In the short-pin forging the measured tensile strength values of the isostatic transverse samples lie below the values of the uniaxial pieces. In figure 10(c) one can see that the tensile strength values are very similar in uniaxial forgings at a high level. The values of the isostatic longitudinal samples again lie far below the uniaxial comparison values.

With increasing pin length the values of tensile strength in the longitudinal direction reach higher levels in uniaxial and isostatic forgings. In contrast to this the values of tensile strength in the uniaxial transverse samples fell with increasing pin length. In the case of uniaxial forgings it is evident that with increasing pin length longitudinal and transverse values reach the same level. From this one can infer that a certain measure of flow is necessary to reach the same tensile strength in perpendicular directions. Furthermore it is striking that with increasing relative density the values of tensile strength dropped. One can see this effect in nearly all uniaxial samples and in the isostatic samples of the medium-pin forging. We can assume the reason for this is that with decreasing porosity the particle welding will be loosened. The required load and energy for higher relative density destroy the particle bonding. This means that in view of the tensile strength an optimum measure of forging load exists.

CONCLUSION

- (1) Density distributions of both compacts are changed by deformation. With increased length of pin a compensation of density takes place

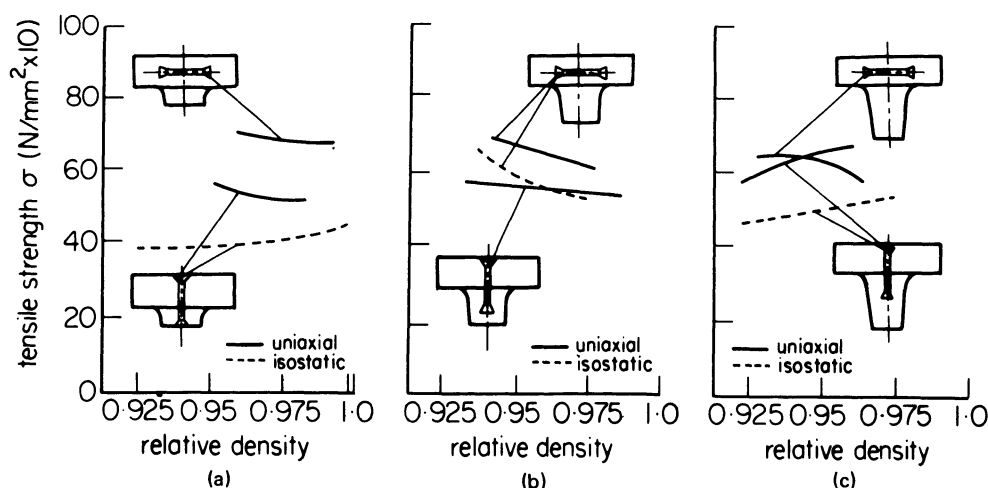


Figure 10.

which is inverse in uniaxial and isostatic forgings.

- (2) Hardness is not constant over the cross section of compacts pressed by uniaxial and isostatic wet bag compaction. After the forging operation there were similar hardness distributions for all pin lengths and for both compaction methods; there is an increase of hardness towards the surface of the forging (except the pin end). Isostatic forgings have a well-arranged distribution of hardness especially at the pin end.
- (3) It is not possible in all cases to find a correlation between distributions of hardness and density.
- (4) Edging of material in increasing pin lengths influences the forging as follows:
 - (a) Variability of density values decreases.
 - (b) In directions perpendicular to each other the tensile strengths are the same.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Professor H. Bühler for his support and encouragement throughout the work. He would also like to thank cand. mach. E. M. Offerhaus for his assistance. This investigation was sponsored by the 'Deutsche-Forschungs-Gemeinschaft'.

REFERENCES

1. R. Davies and J. B. Marx. The production of components by forging of powder preforms. *Report on First International Conference on the Compaction and Consolidation of Particulate Matter*, Brighton, 1972.
2. P. C. Eloff and L. E. Wilcox. Fatigue behaviour of hot formed powder differential pinions. *Report on Fourth International Powder Metallurgy Conference*, Toronto, 1973.
3. S. Mokarski and P. C. Eloff. Equipment considerations for forging powder preforms. *International Journal of Powder Metallurgy* (1973) 7 (2) 15-25.
4. K. Morimoto et al. Transmission spur gear by P/F and deformation characteristics of P/M preforms. *Report on Fourth International Powder Metallurgy Conference*, Toronto, 1973.
5. T. W. Pietrocini, Hot formed P/M applications. *Report on Fourth International Powder Metallurgy Conference*, Toronto, 1973.
6. T. Nakagawa et al. On the cold forging of sintered iron powder preforms. *Proc. 13th International MTDR Conference*, Birmingham, 1972.
7. H. F. Fischmeister et al. Deformation and densification of porous preforms in hot forging. *Powder Metallurgy* (1971) 14, 144-63.
8. K. H. Moyer. Comparison of the mechanical properties of deformed P/M iron-carbon preforms with commercial wrought iron-carbon alloys. *Report on Fourth International Powder Metallurgy Conference*, Toronto, 1973.
9. G. T. Brown and T. B. Smith. The relevance of traditional materials specifications to P/M products. *Report on Fourth International Powder Metallurgy Conference*, Toronto, 1973.

POWDER COMPACTION AT VERY HIGH PRESSURES

by

P. N. TOMLINSON,* R. L. HEWITT* and R. D. VENTER*

SUMMARY

Previous researchers in powder compaction have been limited to pressures of about 1500 MN/m², with 1000 MN/m² being the more usual limit. This paper describes some compaction work performed at much higher pressures.

Pressure-density results and hardness measurements are presented for a nickel-base superalloy compacted at pressures up to 4000 MN/m² together with scanning electron micrographs of the compacted particles. Even at 4000 MN/m² the powder reached only 94 per cent of its theoretical density, and was not sufficiently consolidated to permit any but the most delicate handling. The density and hardness values, however, showed no tendency to 'level-off' at this pressure.

The pressure-density results are further discussed in terms of a recent theory relating various stages of compaction to the extent of plastic deformation.

INTRODUCTION

Powder metallurgy is becoming increasingly important both as a means of reducing costs by minimizing machining operations and material losses and as a means of producing highly alloyed materials that cannot be successfully produced by conventional means. One aspect of powder metallurgy which has received considerable attention is that of powder compaction, but a recent review of the literature¹ showed that the majority of this work has been performed at pressures below 1500 MN/m², with 1000 MN/m² being the more usual upper limit.

Many of the new, high strength powders cannot be consolidated cold even at 1500 MN/m² and are generally formed using hot pressing or hot extrusion. This high temperature processing can be undesirable, however, and a means of cold compacting these powders would be useful. One possible way of achieving this is by using much higher pressures.

The only compaction study performed using very high pressures is that reported by Meerson, Islankina, Mel'nikov and Gol'man³. These authors gave a very brief report of the hydrostatic compaction of a stainless steel powder at pressures up to 2800 MN/m². Unfortunately, neither the powder properties nor particle shape were given in this report, nor was the degree of bonding between particles discussed. Thus it is very difficult to draw any conclusions with regard to the compactibility of other high strength powders from this work.

The purpose of this work, therefore, was to

investigate the possibility of using very high pressures (up to 4000 MN/m²) to consolidate high strength powders which cannot be consolidated at normal compacting pressures. An argon atomized nickel-base superalloy, INCO 713 LC, was used in the investigation since it is typical of many of the new high strength powders, having a regular particle shape and a yield strength of about 1000 MN/m². This powder cannot be consolidated by isostatic compaction at pressures up to at least 1400 MN/m².

EXPERIMENTAL METHOD

Compaction was performed in a single-acting piston-cylinder arrangement. The die bore was sprayed with PTFE to reduce die-wall friction and 55 g of powder were used for each compact which resulted in a diameter to height ratio of about three.

A continuous pressure-density curve was determined using the method proposed by Heckel⁴. A distance piece, 17.50 mm in length, was placed in the die and the piston advanced to meet this block. A dial gauge, calibrated in divisions of 0.1 mm and placed to measure the displacement of the head of the piston with respect to the die, was then set to zero. The powder was weighed, placed in the die and compressed to 4000 MN/m². The dial gauge was monitored continuously during the compression and the distance recorded corresponds to the height of the sample during compaction, less the compression of the piston. The dial gauge was similarly monitored

* De Beers Diamond Research Laboratory, Johannesburg, South Africa

during release of the pressure to give the appropriate corrections for piston compression and elastic recovery of the compact. The height of the compact was measured directly before ejecting from the die to provide a check on the final height.

Several other samples were similarly compacted to pressures between 0 and 4000 MN/m². None of the resulting compacts were sufficiently consolidated for any mechanical tests to be performed, but they were examined in a scanning electron microscope for particle deformation, and microhardness measurements were performed on the individual powder particles.

RESULTS

The loading and unloading curves for compaction up to 4000 MN/m² are shown in figure 1. In order to

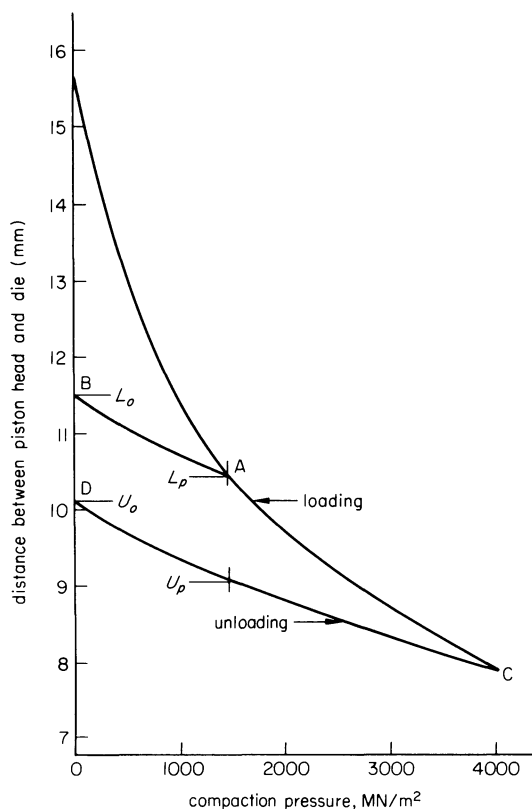


Figure 1. Piston displacement versus pressure for loading and unloading.

calculate the density of a sample compacted at pressure, P say, it is necessary to determine the height of a compact pressed to this pressure. This is achieved by assuming that if the sample were compressed to pressure, P , it would unload along curve AB in figure 1, where AB is such that, by vertical displacement, it can be brought into coincidence with curve CD . Then the height at zero pressure of a sample compacted to pressure P is given as

$$L_0 = L_p - (U_0 - U_p)$$

where L_p is the height at pressure P on loading, U_p is the height at pressure P during unloading and U_0 is the final unloaded height.

The assumed form of unloading curve has previously been shown to be correct by Heckel⁴, and was confirmed for this powder by obtaining unloading curves after pressing to several different end point pressures. All the loading curves were the same within the limits of experimental error, and the experimental unloading curves were coincident with the assumed unloading curves obtained from figure 1.

The accuracy of the distance calibration was checked by comparing the final height on the unloading curve with the height of the compact measured directly. Only the compact pressed to 4000 MN/m² was sufficiently consolidated to permit this direct measurement.

Heckel, who only compacted at pressures up to about 60 MN/m², was able to assume that the diameters of his compacts were equal to the diameter of the die. This is not valid when compacting to higher pressures, however, since the compact expands when ejected from the die because of elastic recovery. This was observed when compacting iron and copper in the present apparatus to 3500 MN/m². It was observed by direct measurement of the diameters of these materials, that there was a nearly linear variation of the diameters of these compacts with pressure up to a point where the hardness of the compact was no longer increasing. This is because as the pressure is raised, the yield strength of the powder increases as evidenced by the increase in hardness, and so the amount of elastic recovery increases. Since the superalloy compacts showed little levelling off in their hardness values, the diameter was assumed to vary linearly with pressure from the die diameter, to the diameter of compact pressed at

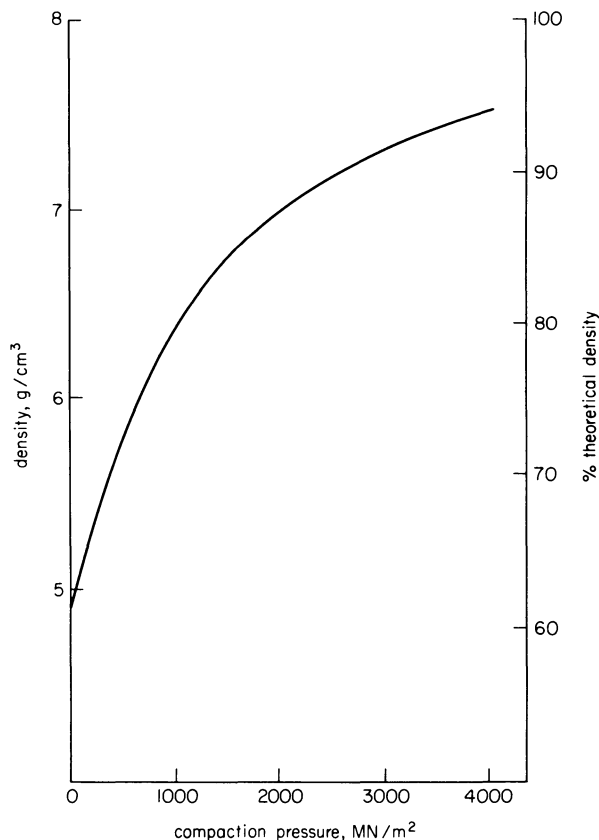


Figure 2. Density versus pressure.

4000 MN/m² which was sufficiently consolidated to permit a direct measurement.

The pressure-density curve obtained by the procedures outlined above is shown in figure 2 and the $\ln(1/(1 - D))$ versus pressure plot, where D is the relative density, is shown in figure 3. The theoretical

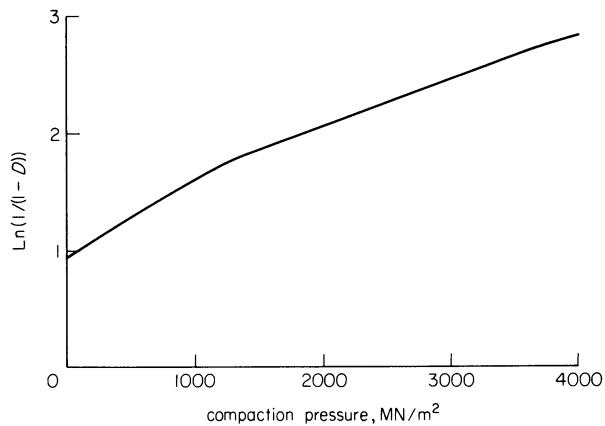


Figure 3. $\ln(1/(1 - D))$ versus pressure.

density was obtained from the International Nickel Company and was taken as 8.01 g/cm³.

The microhardness measurements obtained from individual powder particles are shown in figure 4 as a function of pressure, and scanning electron micrographs of individual particles after pressing to various pressures are shown in figures 5 to 8. Figure 9 shows the condition of compacts pressed to 3000 and 4000 MN/m² after removal from the die.

DISCUSSION

Macro examination

None of the compacts pressed at 3500 MN/m² or below were sufficiently well consolidated to enable them to be extracted from the die in one piece. At the lower pressures, the compacts disintegrated entirely, but as the pressure was raised, so the consolidation, or bonding, within the compact improved.

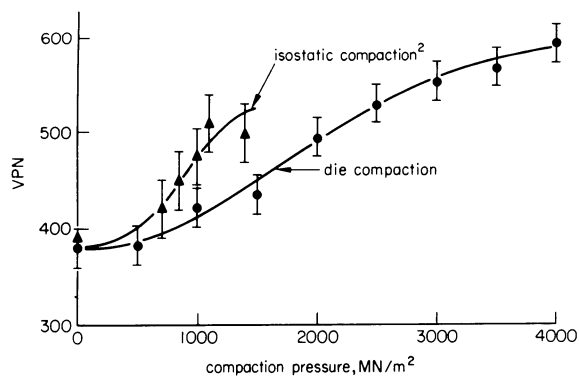


Figure 4. Hardness of individual powder particles versus pressure.

Figure 9 shows one compact pressed at 3000 MN/m² and one pressed at 4000 MN/m². The former shows regions of consolidation, these being

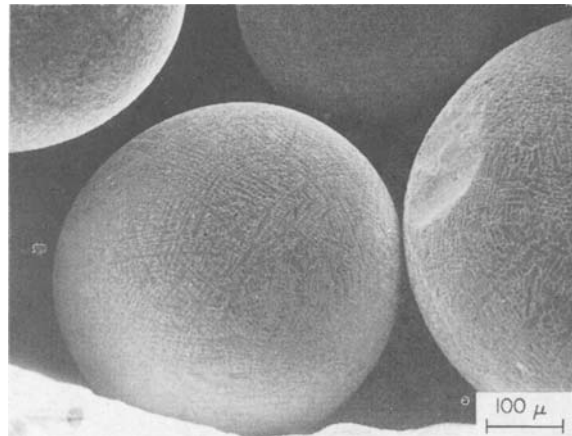


Figure 5. Scanning electron micrograph of as-received powder.

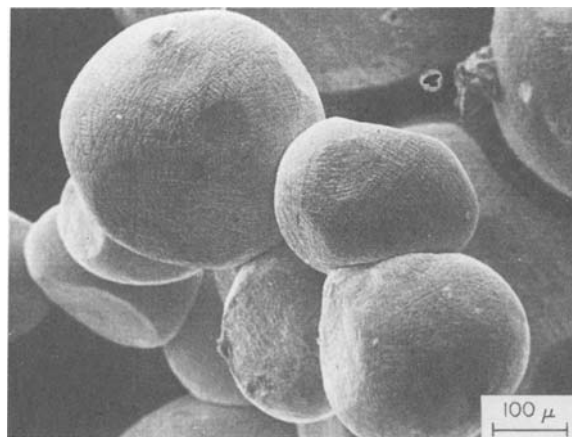


Figure 6. Scanning electron micrograph of powder compacted to 1000 MN/m².

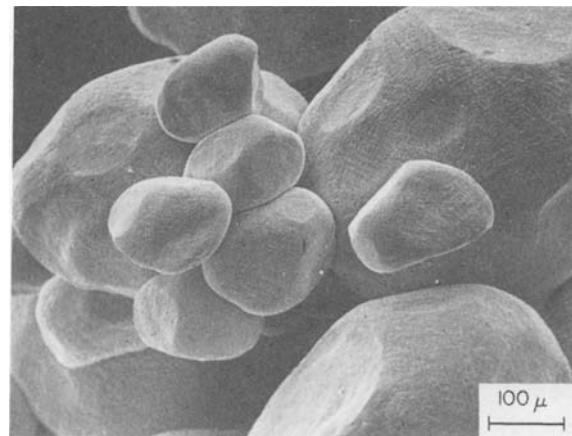


Figure 7. Scanning electron micrograph of powder compacted to 2500 MN/m².

principally around the outside of the compact. This may be due to an uneven pressure distribution within the sample during pressing. However, the cylindrical edges of the sample show particularly good bonding. This is due to the friction at the edges which promotes a shearing action. The presence of shear forces has been shown to be beneficial in improving bonding by Koerner and Quirus⁵. Thus the better consolidation around the edges of the samples is most

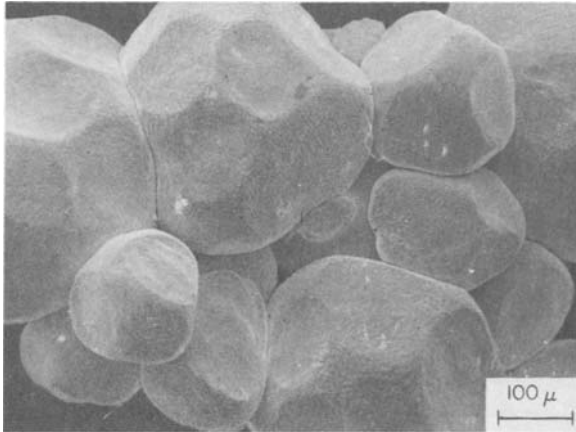


Figure 8. Scanning electron micrograph of powder compacted to 3500 MN/m^2 .

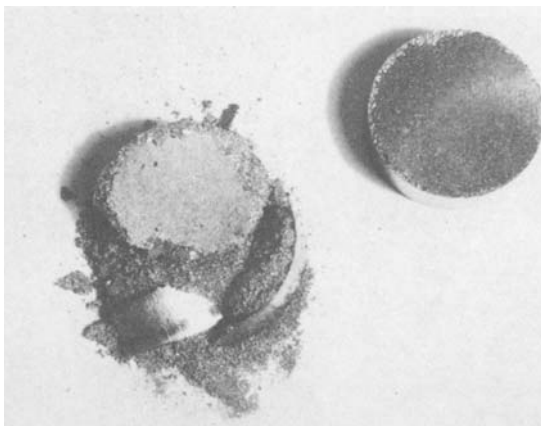


Figure 9. Compacts pressed at 3000 MN/m^2 (left) and 4000 MN/m^2 (right).

likely due to the shear forces acting in these regions.

It was possible to extract the sample compacted at 4000 MN/m^2 in one piece as shown in figure 9, and it could be handled with care. However, it was very fragile and no mechanical tests could be performed on it. On fracturing this sample, it was observed that there was little bonding within the centre of the compact.

Micro examination

Figure 5 shows a scanning electron micrograph of the as-received powder. It has a regular spherical shape and the surface shows evidence of the very fine dendritic structure. Figures 6 to 8 show samples of the powder after compaction to 1000 MN/m^2 , 2500 MN/m^2 and 3500 MN/m^2 respectively. The spherical particles are progressively deformed to yield irregular polyhedral particles, the degree of deformation clearly increasing with pressure. The particles compacted at 1000 MN/m^2 are essentially spherical with flat spots on them, while those compacted at 3500 MN/m^2 show little evidence of the original spherical surface and consist of flat or concave surfaces.

There is no evidence of any bonding between these particles which supports the conclusions of previous researches². In general, green strength in a compact is developed by a combination of mechanical interlocking and frictional welding. This welding only

occurs with irregular particle shapes, however, where the resultant asymmetric loading between particles creates shear stresses and interparticle sliding, with consequent frictional welding of freshly exposed surfaces. The powders used in this study, however, are of a very regular shape which leads to symmetrical point contact loading between particles and direct compressive stresses with little or no shearing (except around the edges) and hence no sliding or frictional welding effect. The point contacts merely deform into flat surfaces as is clearly shown in the figures. Thus any green strength which may be developed in these compacts is a result of shearing at the edges of the compact due to the motion within the die. This implies that it would be very difficult to produce a coherent compact from these powders using isostatic compaction.

One further point which is evident from the micrographs is that there is still considerable porosity within the compact even after pressing to 3500 MN/m^2 . This can be clearly seen in figure 8 from the size of the ridges between adjacent flat spots on the larger particles. The finite width of the ridges indicates that there was no interparticle contact in this region and hence there must have been a void.

Hardness measurements

The microhardness measurements shown in figure 4 indicate that the hardness of the individual particles increases with compaction pressure from an initial value of about 380 VPN to about 590 VPN at a pressure of 4000 MN/m^2 . It is generally observed^{2,6} that the hardness tends toward a constant value as the compact approaches its limiting value of density (which is, in general, less than the theoretical density). This set of data, however, shows little levelling off, which indicates that there is still considerable porosity within the compact which could be reduced by compacting at a higher pressure.

It is interesting to compare these results with those obtained previously² for isostatic compaction of the same powder. The hardness measurements for the isostatic compaction, also shown in figure 4, show a much more rapid increase with pressure at first, but then show a slower increase than for die compaction. At higher pressures than could be achieved with the available isostatic compaction equipment, it is probable that the curves would tend to coincide.

If it is assumed that compacts with the same hardness, though pressed by different methods, have the same density, then an isostatic compaction pressure of 1000 MN/m^2 , say, is equivalent to a die compaction pressure of about 1800 MN/m^2 . This confirms the view of Meerson et al.³ who suggested that isostatic compaction was 1.5 to 3 times more efficient than die compaction. At much higher pressures, however, it appears that little would be gained from isostatic compaction as suggested by Koerner and Quirus⁵.

Pressure-density data

The pressure-density curve shown in figure 2 confirms the observations made from the analysis of the hardness data and micrographs, i.e. there is still

considerable porosity (6 per cent) present after compacting to 4000 MN/m² and the porosity could be reduced by increasing the compaction pressure. Although this latter statement may seem self-evident, most pressure-density curves level off at high pressures at a limiting density, less than the theoretical density, and further increases in pressure producing little reduction in porosity.

Ln(1/(1 - D)) versus pressure

It is customary in the literature of powder compaction to represent pressure-density data in the form of ln(1/(1 - D)) versus pressure. This generally results in a straight line plot over part of the pressure range, stage 2, with curvature, or a line of different slope above and below this pressure range, stages 3 and 1 respectively. There has been much discussion concerning the significance of the slope of the straight line and the causes of curvature and these are summarized in reference 2. Most authors relate the slope to the strength of the powder but there is little agreement concerning the reasons for curvature. Hewitt et al.⁶ have recently shown that, for isostatic compaction, the linear portion can be related to bulk plastic flow of the powder particles. Using a model of a hollow, strain-hardening plastic sphere to represent the compaction behaviour of a porous body they predicted the compaction behaviour of iron and aluminium in the region where plastic flow is the predominant mechanism of compaction. They further showed² that the transition from stage 1 to stage 2 compaction occurred where the plastic flow in the powder became homogeneous instead of local, i.e. the compaction pressure exceeded the current flow stress of the material. The curvature at high pressures appeared to occur at about the same pressure as the hardness began levelling off.

The ln(1/(1 - D)) versus pressure plot for the present data is shown in figure 3. The first stage of compaction appears to extend to about 1200 MN/m², and the second to nearly 4000 MN/m². Although no absolute values of the yield strength of the powder are available, it is probably about 1000 MN/m² initially and it is quite possible that it would have increased to 1200 MN/m² by the time the pressure has reached 1200 MN/m². Thus the transition from stage 1 to stage 2 appears, at least qualitatively, to correspond to the change from local to homogeneous plastic flow. The hardness values are just beginning to level off at 4000 MN/m² as shown in figure 4, and so the beginning of stage 3 for this powder also seems to correspond with previous findings, i.e. this stage begins at about the same pressure as the hardness begins to level off.

CONCLUSIONS

The variation in density and hardness with compaction pressure has been obtained for INCO 713 LC powder die compacted at pressures up to 4000 MN/m². Even at the highest compaction pressure, the samples attained only 94 per cent of their theoretical density, but it appears that this could be reduced by increasing the compaction pressure.

Micro and macro examination of the compacts revealed little evidence of interparticle bonding, and what little green strength was developed was primarily due to shearing at the edges of the compact due to the motion with the die.

Comparing the hardness of the powder after compaction with those obtained by isostatic compaction, suggests that at low pressures (less than 1500 MN/m²), isostatic compaction may be 50 to 100 per cent more efficient than die compaction, but at much higher pressures (4000 MN/m²) the two methods are very similar.

Plotting the density data as ln(1/(1 - D)) versus pressure qualitatively confirmed a recent theory that the transition from stage 1 to stage 2 compaction occurred when the compaction pressure exceeded the current flow stress of the powder, i.e. when plastic deformation became homogeneous instead of local.

REFERENCES

1. R. L. Hewitt, R. D. Venter and M. C. deMalherbe. Theories Relating to the Compaction of Metal Powders. *South African Mechanical Engineer* (1973) 23, 214.
2. R. L. Hewitt, W. Wallace and M. C. deMalherbe. Plastic Deformation in Metal Powder Compaction. *Annals of CIRP* (1973) 22, 87.
3. G. A. Meerson, A. F. Islankina, V. N. Mel'nikov and L. D. Gol'man. Hydrostatic Pressing of Steel Powders. *Poroshkovaya Metallurgiya* (1969) 79, 13.
4. R. W. Heckel. Density-Pressure Relationships in Powder Compaction. *Trans. Metall. Soc. AIME* (1961) 221, 671.
5. R. M. Koerner and F. J. Quirus. High Density P/M Compacts Utilizing Shear Stresses. *Int. J. Powder Metallurgy* (1971) 7, 3.
6. R. L. Hewitt, W. Wallace and M. C. deMalherbe. The Effects of Strain Hardening in Powder Compaction. *Powder Metallurgy* (1973) 16, 88.
7. R. L. Hewitt, W. Wallace and M. C. deMalherbe. The Strain Hardening Hollow Sphere Model of Compaction Applied to Aluminium. *Int. J. Powder Metallurgy* (1973).

GENERAL ASPECTS OF FORMING

STRONG COMPOSITES BASED ON ALL-METAL SYSTEMS

by

D. V. WILSON*, W. T. ROBERTS*, J. J. MOORE†,
R. F. STOKES‡ and J. FARMER*

SUMMARY

High-strength composites based on two metallic components have received less attention than metal-non-metal combinations, although the former have special advantages in relation to economical manufacture and use. Strong metal-metal components have been manufactured by routes involving co-working of the components. Systems studied include aluminium alloy-steel, aluminium alloy-titanium and titanium-steel. Mechanical properties have been evaluated and the forming behaviour of some of the laminates has been assessed.

INTRODUCTION

Conventional light alloys are strengthened by three main mechanisms—work hardening, solute hardening and hardening by a dispersed phase. These mechanisms operate on the scale of the micro-structure and, at least to a first approximation, their strengthening effects are homogeneous on the macroscopic scale. An alternative way of strengthening is by the introduction of a strong second phase on a macroscopic scale, using a solid-phase bonding technique. The function of the second phase in this case is to sustain a greater proportion of the applied load in relation to its cross-section than does the matrix. An advantage of this method is that the strong phase can, in principle, be situated in regions where the service stress levels are highest so that the component is more closely matched with the requirements of the application. Despite the relatively brittle behaviour of many very strong materials it is possible, by incorporating them in a softer matrix, to produce a tough composite material. Possibilities for the reinforcing component range from single crystal whiskers of non-metallic material to strong metallic wires and sheets, the choice being determined by both economic and practical factors.

The present investigation was focused on the strengthening of aluminium and titanium alloys with steel wires or sheet. The only system in which steel was not used as reinforcement was the titanium-aluminium laminate, in which the possible utilization of the high biaxial strength of titanium sheets was investigated.

The strength levels of the steel required to produce

significant composite strengthening, and the range of heat treatment temperatures available for composite heat treatment are obviously different in the two matrix materials, aluminium and titanium. These considerations dictated the choice of the steel reinforcing material, which was to be incorporated in the composite by a route involving co-working of the two materials. Some emphasis has been given to the possibilities of working the composites after assembly because this is a potential advantage of using metallic reinforcement and it appears to have received relatively little attention up to the present.

In addition to achieving the highest strength levels in both components of the composite, several other factors are important in the production of a useful composite material. These are firstly, the production of a good bond between the two components, secondly, the achievement of a uniform or controlled distribution of reinforcement and finally, the control of diffusion and prevention of the formation of harmful brittle intermetallic compounds at the interface between the components.

For each system investigated, the first task was to select the best conditions for bonding and then to define the limits of further working and heat treatment to raise the level of the composite strength. This was followed by evaluation of the mechanical properties of the composite, including room temperature and elevated temperature tensile properties, impact strengths and behaviour under cyclic stress. A further important aspect in the case of laminate materials was an assessment of the formability and estimation of the strength of the formed parts.

* Department of Industrial Metallurgy, University of Birmingham

† now at West Bromwich College of Commerce and Technology

‡ now at J. Lucas Group Research Centre, Monkspath

ALUMINIUM-STEEL COMPOSITES

In order to achieve useful strengthening in aluminium alloys, and assuming that a Law of Mixtures strength relationship¹ is maintained, the steel reinforcement must attain a strength of about 220×10^3 psi. Such a strength level is readily attained in the stronger structural steels when they have undergone transformation during heat treatment. However, conventional steel heat treatments in the 800 to 900 °C temperature range clearly cannot be carried out after the steel is incorporated in a solid aluminium-base matrix, and if the steel is already in a strong condition before bonding, it is difficult to achieve the necessary deformation to produce a good bond by roll bonding techniques. It was therefore necessary to seek a relatively low-temperature strengthening mechanism for the steel which was controllable and which would preferably come into operation after bonding. The transformation from γ to α martensite which occurs during deformation at around room temperature in 18/8 austenitic stainless steels offers a useful hardening mechanism. The strength levels attained by the type 304 stainless steels can be surpassed in more complex alloys in which precipitation hardening can be utilized in addition to transformation hardening, and some of these alloys were used to attain greater strength levels in the steel.

Another group of steels capable of reaching high strength levels by work hardening and relatively low-temperature ageing treatments is the maraging group. The ageing temperature happens to be close to the temperature required for solution heat treatment

of heat-treatable aluminium alloys, and composite heat treatment at this temperature can thus serve a dual purpose.

It is clear that attention must also be paid to strengthening of the matrix, in this case aluminium alloy, by carrying out appropriate heat treatments and working operations on the composite. Two types of aluminium alloys were used—a heat-treatable alloy of the Duralumin type and aluminium-magnesium alloys strengthened mainly by solute hardening and work hardening.

Bonding

In general the strength of the bond depends on cleanliness of the two surfaces, the amount of reduction in the bonding pass and the temperature of the bonding pass. Whenever possible, the components to be bonded were degreased in a vapour bath of trichlorethylene and scratch-brushed. In order to achieve a good bond, it is usually necessary to obtain reductions of 40–50 per cent at moderately elevated temperatures². Systematic investigation of the bonding variables was undertaken to establish the best conditions for each system. When attempts are made to bond two metals that are sufficiently ductile to undergo the required deformation without fracture but which have widely different flow strengths and rates of work hardening, such as aluminium and steel, limits are imposed on the amount of deformation by the onset of plastic instability in the strong component. An example of strain concentration due to localized shear followed by necking of stainless steel in aluminium-steel laminates is shown in figure 1.

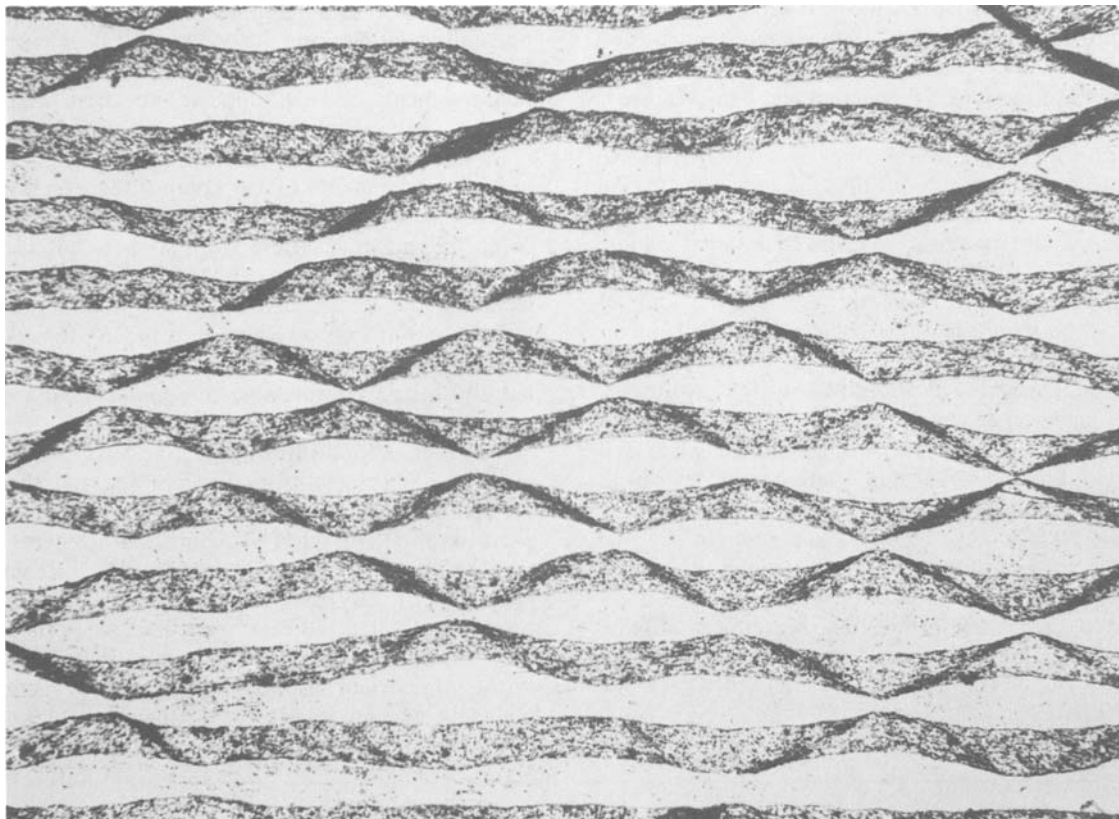


Figure 1. Longitudinal section of laminate consisting of 12 layers of 18/8 stainless steel with 13 layers of aluminium-3 per cent magnesium roll bonded at 300°C, and further cold rolled to give a total reduction in thickness of 80 per cent, showing co-operative shearing of stainless steel ($\times 40$).

In multi-layer laminates the necking is a co-operative phenomenon, necks forming on shear planes critically inclined at about 45° to the rolling plane. Bonding temperatures were therefore selected to limit the differences in flow strengths of the two components. In 18/8 stainless steel-aluminium laminates, low bonding temperatures during which the stress-induced martensitic transformation occurs were avoided while an upper limit to the bonding temperature was set by the very low strength of the aluminium component as its melting point was approached. For this particular system the optimum bonding conditions were defined as a reduction in thickness of about 45 per cent at temperatures of $300\text{--}350^\circ\text{C}$. Tests to examine the quality of the bond involved four-point bending and severe flange-shrinking in cup-drawing operations carried out on circular blanks cut from the composite sheets. These tests demonstrated that very good bonding between aluminium alloys and austenitic steels can be attained in roll bonding under the conditions used.

Working and heat treatment

The strength levels of both components of a metal-metal composite can often be raised by work hardening after bonding. The extent of deformation after bonding is again limited by instability or fracture of the strong component or, in some cases, by delamination. Generally in the aluminium-steel composites, warm working followed by a limited amount of cold working produced the highest strengths. The condition of the reinforcement and matrix prior to this final working operation governed the reductions and final properties of the composite. In figure 2 composite strength is plotted against rolling reduction for stainless steel wires in annealed or hard-drawn initial conditions rolled between sheets of Duralumin in either soft or aged initial conditions. Hard-drawn stainless steel wires in a soft matrix fracture after a small rolling reduction, but a harder matrix permits greater reduction before fracture occurs.

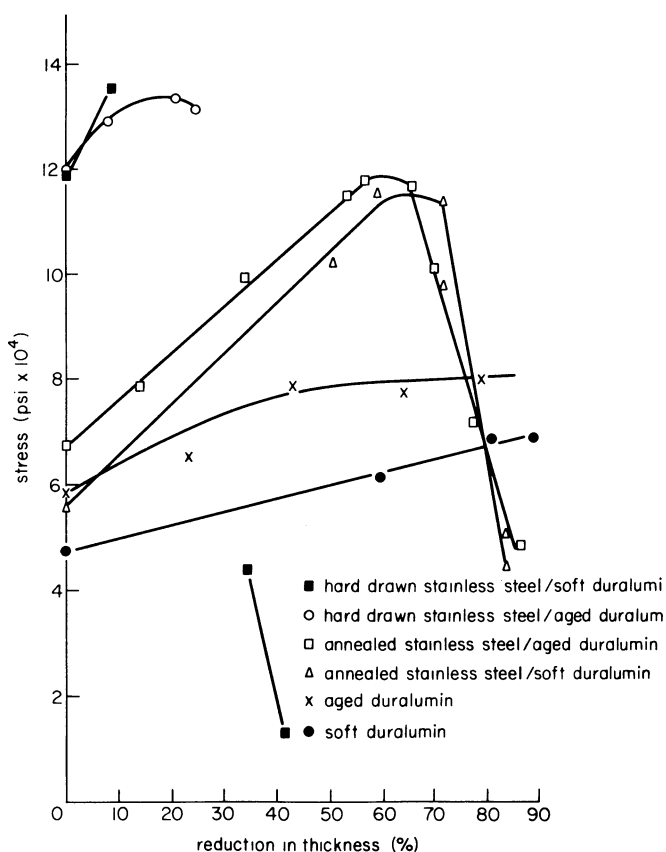


Figure 2. Strength measured in wire direction plotted against rolling reduction for hot-pressed composites of 18/8 stainless steel wires in Duralumin sheet.

Tensile properties

The room temperature tensile strengths and specific strengths achieved by various aluminium-steel laminates are shown in Table 1. Specific strength is defined here as the tensile strength divided by the specific gravity. The composites generally have higher specific strengths than the matrix materials. Slightly lower levels of specific strength in tests parallel to the

TABLE 1 Tensile properties of aluminium-steel laminates

Composition of components		Volume fraction of steel (per cent)	Angle between tensile axis and rolling direction	Tensile strength ($\text{psi} \times 10^3$)	Specific strength ($\text{psi} \times 10^3$)
Aluminium alloy	Steel				
Duralumin	18/8 stainless steel*	51.8	90°	119.1	21.7
Duralumin	18/8 stainless steel†	51.8	90°	134.3	24.5
Duralumin	18/8 stainless steel‡	46.7	90°	141.1	25.7
Duralumin (WP)	—	0		58	20.8
Al-5 per cent Mg	18/8 stainless steel	45.3	90°	141.2	28.0
Al-5 per cent Mg	18/8 stainless steel	45.3	0°	129.0	25.5
Al-5 per cent Mg ($\frac{1}{4}$ H)	—	0		67.2	25.4
Al-5 per cent Mg	11 R 51 §	44.4	90°	152.3	30.4
Al-5 per cent Mg	11 R 51 §	55.6	0°	164.0	29.4
Al-5 per cent Mg	Fortinox	35.2	90°	147.9	32.7
Al-5 per cent Mg	Fortinox	55.6	0°	159.0	28.4

Warm reduction after bonding pass (per cent) Final cold reduction (per cent)

*	55	0
†	55	32.4
‡	55	40

§ 1.15 Si 0.09 C 1.25 Mn 8.0 Ni 17.0 Cr 0.7 Mo
 || 0.65 Si 0.09 C 8.5 Mn 2.3 Ni 17.0 Cr

TABLE 2 Tensile properties of aluminium-maraging steel-hot pressed wires in sheets

Composition of components		Volume fraction steel	Tensile strength (psi x 10 ³)	Specific strength (psi x 10 ³)
Aluminium alloy	Steel			
Duralumin	(1)*	20	96	25.7
Duralumin	(2)*	9.6	88.5	27.4

* (1) 18.0 Ni 4.5 Mo 7.5 Co 0.21 Ti 0.005 S 0.003 P remainder Fe
 (2) 18.2 Ni 15.0 Mo 8.5 Co 0.54 Ti 0.11 Al 0.003 B 0.005 C 0.01 Zr 0.006 S 0.003 P remainder Fe

rolling direction compared with those in the transverse direction are probably due to the onset of necking or incipient failure of the steel. The highest specific strengths were obtained with a precipitation-hardened stainless steel (11 R51—supplied by Sandvik) and a high-strength steel (Fortinox—supplied by Samuel Fox and Co. Ltd.).

The laminates described above usually consisted of four layers of steel sandwiched between five layers of aluminium alloy, with a total final composite thickness of about 0.050 in. The composite tensile strengths achieved by hot-pressing maraging steel wires between sheets of Duralumin are given in Table 2, the strengths being measured at room temperature in the direction of the wire reinforcement.

Specific tensile strengths of some wire-in-sheet composites at temperatures between room temperature and 300°C are shown in figure 3. The composites

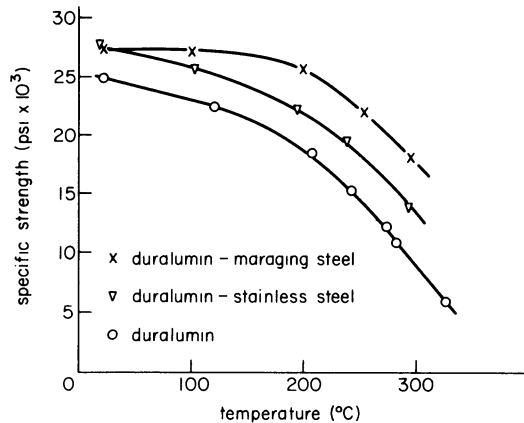


Figure 3. Variation of specific strength with testing temperature for Duralumin sheet reinforced with two types of steel wires.

are compared with results for Duralumin which was aged to peak hardness prior to testing. The maraging steel wires provide better reinforcement than 18/8 stainless steel, but at temperatures greater than 200°C both composite strengths start to fall rapidly.

Fatigue tests

Two types of fatigue test were carried out: tension-compression with zero mean stress and tension-tension with zero minimum stress.

Performance of the composites in the tension-compression tests was poor in relation to their static strengths. Typically their fatigue strength for 10⁶ cycles to failure was similar to that of the matrix alloy. In some cases the fatigue strength for the reinforced laminate was even lower than that of the aluminium alloy matrix, and higher volume fractions

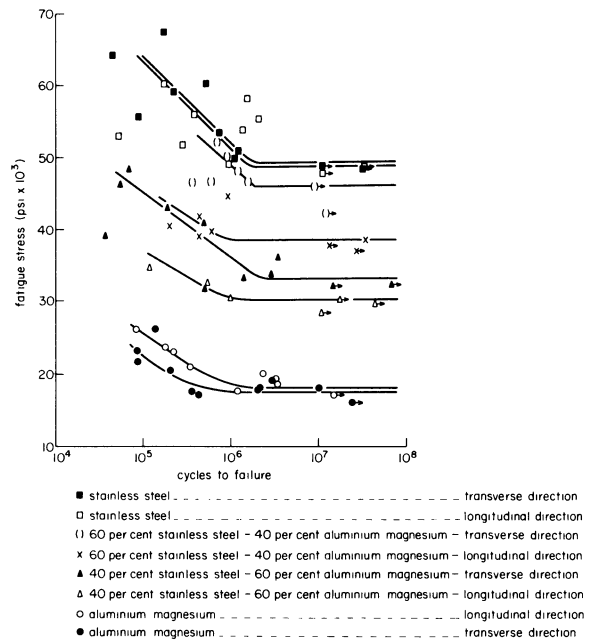


Figure 4. Tension-tension (zero minimum stress) fatigue results for 18/8 stainless steel-aluminium 7 per cent magnesium laminates.

of reinforcement led to poorer fatigue performance. Similar findings by Courtney and Wulff³ have been explained on the basis that the composite fatigue strength is controlled by matrix fatigue failure. This occurs when the reinforcement is deformed elastically whilst some plastic deformation occurs in the matrix. Reinforcement is ineffective in completely reversed stress cycling at relatively low plastic strain amplitudes. With increasing numbers of cycles, the effective stress range acting on the matrix increases as it work hardens until it fails. Since the plastic strain

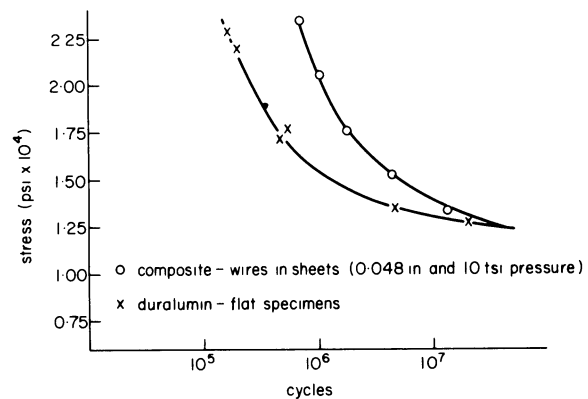


Figure 5. Tension-tension (zero minimum stress) fatigue results for Duralumin sheet reinforced with maraging steel wires.

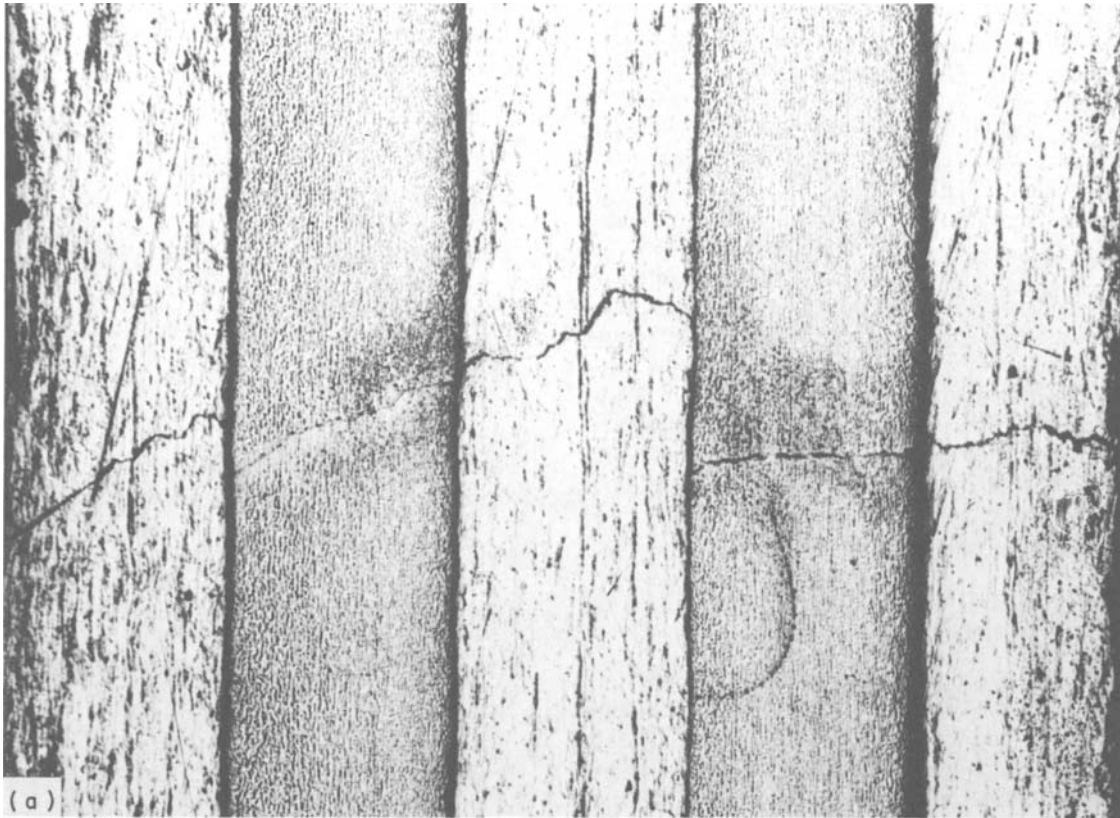


Figure 6. (a) Fracture path in 18/8 stainless steel–aluminium 7 per cent magnesium laminate with 3 layers of aluminium and 2 layers of steel after fatigue test to failure ($\times 55$).

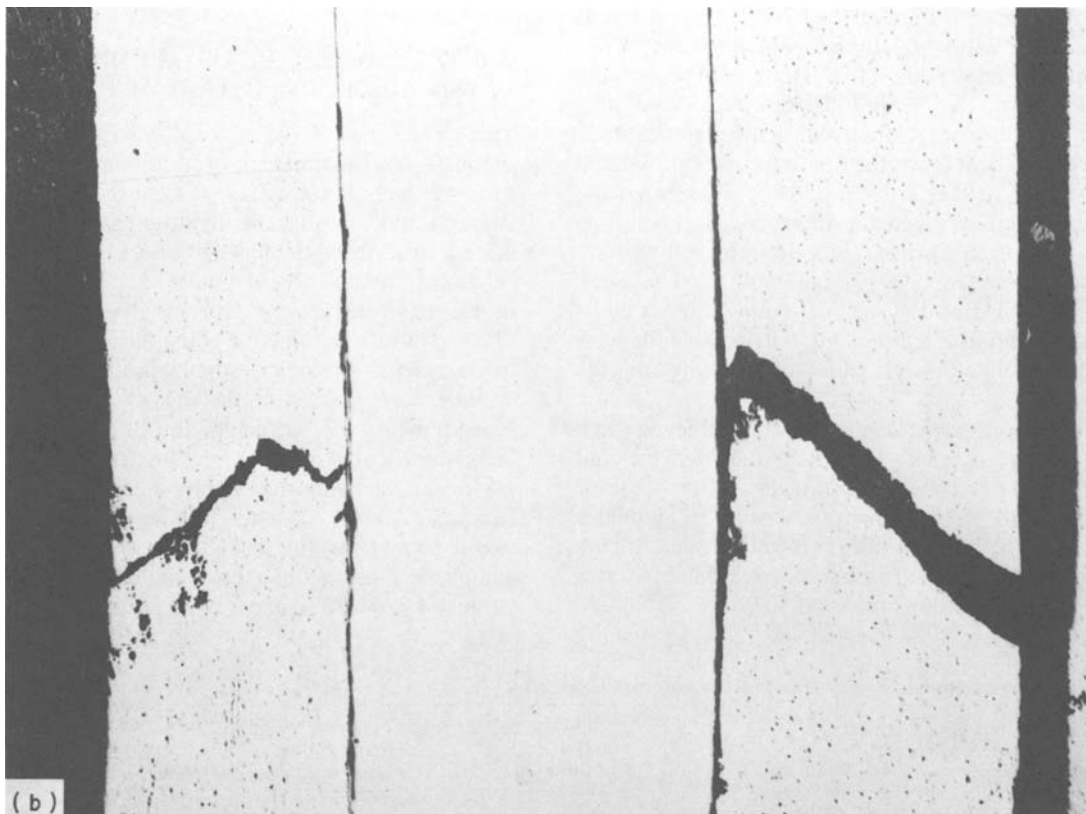


Figure 6. (b) Fatigue crack passing round a steel wire in Duralumin sheet reinforced with maraging steel wire ($\times 70$).

amplitude per cycle is appreciably greater for tension-compression than for tension-tension tests, the matrix will fail at a smaller number of cycles in the former case. A further factor which may lead to premature fatigue failure in tension-compression is the possibility of debonding or crack initiation at the reinforcement-matrix interface.

Under conditions of tension-tension, reinforced laminates showed considerable enhancement of fatigue strength (figure 4), whereas composites reinforced by wires showed fatigue strengthening only in the finite number of cycles to failure range (figure 5). The differences in performance associated with the two types of composite geometry may be due to the greater effectiveness of a planar interface in deflecting a propagating crack compared with the effectiveness of wire-matrix interfaces. Some evidence for this suggestion is provided by two photomicrographs showing fatigue cracks (figure 6).

ALUMINIUM-TITANIUM LAMINATES

The object of producing aluminium-titanium laminates was to exploit the high strength of titanium sheet in the through-thickness direction arising from the crystallographic texture developed in the titanium⁴. The component materials used for laminate manufacture were aluminium 5 per cent magnesium with IMI Titanium 130 or IMI Titanium 160, the Titanium 160 having a higher interstitial content than the Titanium 130 and being stronger but less ductile.

The temperature range at which bonding reductions could be made was limited by the formation of deleterious intermetallic compounds at the interface when the temperature exceeded 500°C. Good bonds were formed with reductions of 40 per cent in the temperature range 300–400°C. If the reductions were too large, necking of the titanium occurred in a similar fashion to that observed in the stainless steel.

The strength level of the bonded composites could be raised by further working. Figure 7(a) shows the composite tensile strength measured in the rolling direction plotted against reduction for rolling temperatures between 100°C and 400°C. The peak strength reached at each rolling temperature is about the same, and the subsequent fall in strength with increasing reductions is due to necking of the titanium.

Somewhat higher strength levels are achieved when warm rolling to peak strength is followed by cold rolling, the trend being shown in figure 7(b) for Titanium 130 reinforcement, whereas the laminates in figure 7(a) contain the stronger Titanium 160. Tensile strengths and specific tensile strengths achieved in this system are shown in Table 3.

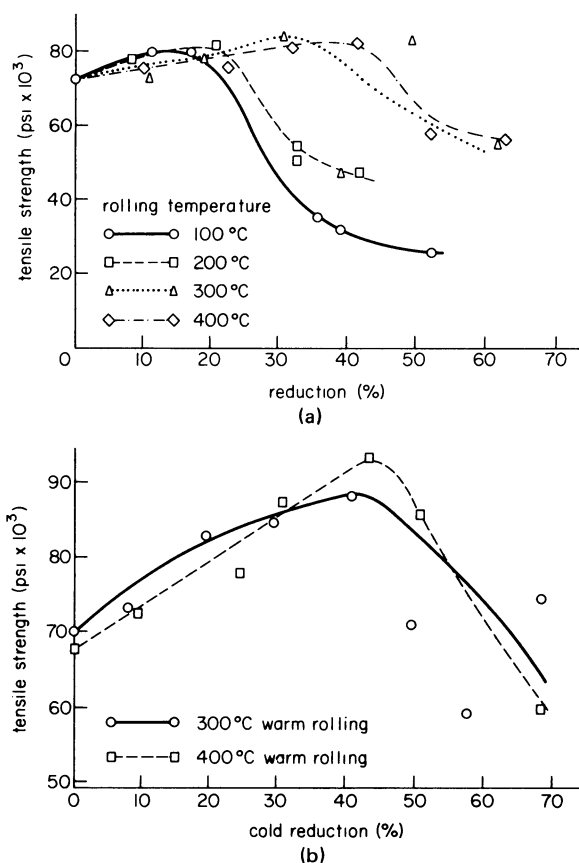


Figure 7. (a) Strength of Titanium 160-aluminium 5 per cent magnesium laminates after roll bonding at 300°C and further warm rolling at four different temperatures. (b) Strength of Titanium 130-aluminium 5 per cent magnesium laminates after roll bonding at 300°C, warm rolling with 45 per cent reduction at two different temperatures and further cold rolling.

SHEET FORMABILITY OF ALUMINIUM-STEEL AND ALUMINIUM-TITANIUM LAMINATES

These tests were made to explore the feasibility of shaping the laminates by ordinary sheet-metal forming techniques and to examine biaxial flow strength and ductility. As mentioned previously, the biaxial flow strength of titanium can be appreciably enhanced, relative to its uniaxial strength, by the development of a suitable crystallographic texture. The strain ratio r (where r is the ratio of width strain to thickness strain measured in uniaxial tension) provides an index of through-thickness strength. Measurements of r made on the titanium-reinforced laminates in the rolling direction were in the range 1.8 to 2.2, showing that plastic anisotropy favouring through-thickness strength had been developed to a useful extent in the titanium components of the laminates. The r values of the stainless steel reinforced laminates were of the order of 1, indicating that no

TABLE 3 Tensile properties of aluminium-titanium laminates

Composition of components		Volume fraction of titanium	Angle between tensile axis and rolling direction	Tensile strength (psi × 10 ³)	Specific strength (psi × 10 ³)
Aluminium alloy	Titanium				
Al-5 per cent Mg	Ti 130	44.4	90°	103.0	29.6
Al-5 per cent Mg	Ti 130	44.4	0°	92.7	27.3

significant enhancement of biaxial tensile strength due to textural effects could be expected in this case.

Five-layer laminates composed of sheets of the aluminium-5 per cent magnesium alloy with both CP Titanium 130 and 160 and with 18/8 stainless steel were examined in standardized deep-drawing and stretch-forming tests. Sheets of 0.05 and 0.08 in thickness were formed using a hemispherically nosed punch of 1.124 in diameter and a die of 1.394 in throat diameter with a 0.4 in profile radius. Oiled polythene lubrication was used in both tests.

STRETCH-FORMING TESTS ON TITANIUM REINFORCED LAMINATES

In Table 4 the true stresses and strains at the maximum loads attained in uniaxial and biaxial stretching of the titanium-reinforced laminates are compared. The values of the flow stress in biaxial tension are indirect estimates, derived by comparing the punch load-punch travel relationships with those of isotropic steel sheets of similar thickness and flow strength.

An outstanding feature of these tests is that the limiting strains in biaxial stretching were inferior to those in uniaxial tension, particularly with the Titanium 160 laminate. Increased flow strength due to plastic anisotropy in the titanium was almost certainly the main factor contributing to reduced ductility in biaxial stretching. However at failure in biaxial stretching, crack propagation involved locally high plastic strains in the aluminium alloy components; thus the work of fracture of the laminate is expected to have remained quite large. The results in Table 4 show that in uniaxial tension the specific strengths of the 'as bonded' titanium-reinforced laminates were inferior to those of fully heat-treated high-strength aluminium alloys, but in biaxial tension the specific strengths of the laminates, enhanced by texture-hardening of the titanium components, were comparable with those of high-strength aluminium alloys.

Stretch-formability of aluminium-5 per cent magnesium laminated with 18/8 stainless steel

The stainless steel components were only partially hardened after roll bonding at 300°C because this temperature is sufficiently high to inhibit transformation to martensite during deformation. The tensile strength values of the 'as bonded' laminates were in the range 70-80 × 10³ psi and uniaxial elongations in the range 25-30 per cent. Thus, despite the high reduction in bonding (about 40 per cent), the marten-

sitic transformation was still available for the enhancement of stable elongation and strength in subsequent straining at room temperature. The ductility of the aluminium alloy-stainless steel laminates was outstandingly good in biaxial stretching. Failure strains were generally in the range 0.6 to 0.7. This excellent ductility is clearly a result of enhanced work hardening in the steel components due to deformation-induced phase transformation, combined with an absence of brittleness at high flow strengths. Hardness values attained in the steel in regions of high strain were in the range 380-400 V.P.N. For laminates containing a 50 per cent volume fraction of the steel this hardness corresponds to a tensile strength of about 120 × 10³ psi and a specific strength of about 23 × 10³ psi.

Deep drawing behaviour

The limiting drawing ratios (LDR) of the four types of laminate tested were as follows:

Reinforcement in 5-layer laminates of aluminium-5 per cent magnesium alloy	LDR
2 layers of Ti-130	2.0
2 layers of Ti-160	1.9
2 layers of 18/8 stainless steel	2.2
3 layers of 18/8 stainless steel	2.2

The drawability of the steel-reinforced laminates was close to the value expected in a ductile material with $r \sim 1$. The drawability of the titanium-reinforced laminates was limited by the relatively low biaxial ductilities of the titanium components. However, with both types of reinforcement the drawability test provides a severe test of the integrity of the inter-layer bonds. The drawn cups show that in the absence of severe buckling, the composites were able to sustain severe compressive straining in the plane of the sheet without delamination.

TITANIUM-STEEL COMPOSITES

In order to compete with high-strength titanium alloys on a room temperature strength-to-weight basis, the strength requirement for a continuous steel reinforcement in titanium is of the order 300 × 10³ psi. One route by which strength levels of this order can be achieved in steels is by the process of ausforming⁵. The objective of this part of the work was to define procedures for the ausforming of steel *in situ* in a titanium-steel composite.

TABLE 4 True flow stress, σ_f , and limiting strain, ϵ_f , developed at maximum load in uniaxial and biaxial stretching of aluminium-5 per cent magnesium alloy titanium laminates

Reinforcement	Uniaxial Tests			Biaxial stretch-forming tests		
	σ_f (10 ³ psi)	ϵ_f	σ_f /specific gravity (10 ³ psi)	σ_f (10 ³ psi)	ϵ_f (thickness strain)	σ_f /specific gravity (10 ³ psi)
Ti-130 (33.8 per cent volume fraction)	63	0.17	19.4	~79.3	0.15	~24.3
Ti-160 (37.5 per cent volume fraction)	69.5	0.14	20.7	~71.0	0.05	~21.2

Composites were produced in the form of sheet laminates and rod, the latter consisting either of rods of the individual components (sometimes encased in tubes of titanium to produce a more uniform distribution of reinforcement) or of thin sheets of the two components in the form of a swiss roll.

Two grades of steel were used for most of the work, En30B (0.33 C, 4.16 Ni, 0.31 Mo, 0.17 Si, 0.51 Mn, 0.013 P, 0.009 S, remainder Fe) and Jessop-Saville H50 steel (0.4 C, 5.0 Cr, 1.35 Mo, 1.10 V, 1.0 Si, 0.45 Mn, 0.015 P, 0.015 S, remainder Fe), and three alloys of titanium were used, IMI Titanium 130 (commercial purity), IMI Titanium 230 (2.0 Cu, remainder Ti) and IMI Titanium 318 (6.0 Al, 4.0 V, remainder Ti).

Laminates

For the production of laminates, annealed sheets of En30B steel and Titanium 130 were enclosed in mild steel capsules and roll bonded with 60 per cent reduction at 650°C. Higher temperatures caused excessive oxidation while at lower bonding temperatures the steel hardened appreciably more than the titanium with an increased probability of necking of the steel. The titanium-steel interfaces were studied by electron microprobe analysis (figure 8(a)). At 650°C, there is little evidence of diffusion across the interface after bonding, but some diffusion has taken place after ausforming. After bonding at 900°C, nickel and chromium concentrate at the interface while the titanium concentration falls off gradually as the interface is approached from the titanium side.

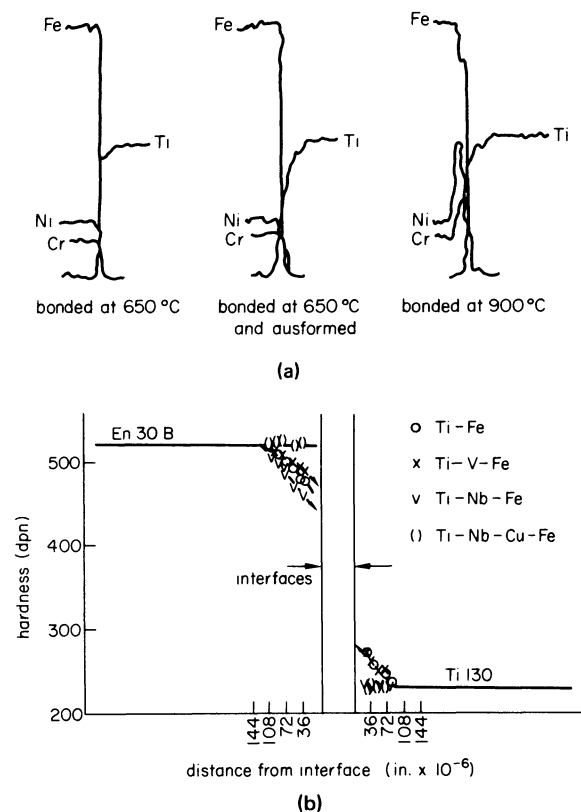


Figure 8. (a) X-ray microprobe analyser traces across the interface of Titanium 130-En30B steel laminates. (b) Variation of microhardness across the interface of Titanium 130-En30B steel laminates with and without barrier layers.

These observations are consistent with the characteristics of the various oxides⁶. Nickel and chromium oxides are anion conductors, so that nickel and chromium ions travel through the oxide layer towards the interface. On the other hand, titanium oxide conducts cations and negative oxygen ions pass through the oxide layers into the metal. It seems evident that extensive oxidation is occurring at bonding temperatures above 650°C, despite the efforts to exclude atmospheric contamination by encapsulation. Apart from oxides, brittle intermetallic compounds may be formed at the interface by interdiffusion, especially if prolonged treatments at high temperature are necessary during the austenitization and hot working sequences. In particular, it is possible for carbon to diffuse from the steel to form titanium carbide.

The introduction of thin barrier layers to reduce the effects of diffusion across the interface was briefly investigated. Laminates roll bonded with no barrier layer were compared with those in which layers of vanadium, niobium and a double layer of niobium and copper separated the steel from the titanium. Microhardness tests (figure 8(b)) indicate that the niobium-copper layer was effective in preventing carbon diffusion.

The presence of brittle compounds at the interface causes particular concern in the case of laminates with large areas of continuous interface. The problem may be less acute in rod composites, and in addition fabrication techniques for rods were devised which reduced the amount of oxidation occurring. The majority of the work on the titanium-steel system was therefore carried out on rod composites.

However, some tensile tests were performed on titanium-steel laminates and the results are given in Table 5. After roll bonding the laminate was heat treated at 850°C for a short time (10 min) in order to austenitize the steel, cooled to 600°C and rolled as much as possible (with reductions of about 70-80 per cent) in the temperature range 600-400°C.

Rod composites

Oxidation problems were lessened by comparison with laminates by the use of a more efficient encapsulation system involving sealing wires of titanium and steel in a welded mild steel cylinder, in one end of which were placed titanium turnings. These were used to getter the container by heating the end containing the turnings in an oxy-acetylene flame to about 900°C. Also, free surfaces were reduced by cold swaging assemblies of wires prior to the austenitization heat treatment and hot working. Initially, bundles of wires of titanium and steel were packed inside a larger titanium tube, but this led to composites with a poor distribution of components. The distribution was improved on a macroscopic scale by placing several cold-swaged composites into a larger titanium tube prior to ausforming. However, a uniform distribution of steel wires was obtained only after encasing each individual steel wire in a titanium tube. In this particular experiment H50 steel was used and the encapsulated composite was austenitized at 1050°C and 'hot-cold worked' at temperatures down to 400°C.

TABLE 5 Titanium-steel composites—tensile properties

Composition of components		Form of composite	Volume fraction of steel (per cent)	Tensile strength (psi $\times 10^3$)	Specific strength (psi $\times 10^3$)
Titanium	Steel				
Ti 130	En 30B	Laminate	50	216.5	34.3
—	En 30B	Sheet	100	304.6	38.5
Ti 130	—	Sheet	0	72.1	16.0
Ti 130B	En 30B	Bundled wires	38	180.6	30.7
—	En 30B	Rod	100	299.0	37.8
Ti 130	—	Rod	0	79.0	17.6
Ti 130	H50	Individual wires in tubes	9	102.4	21.3
Ti 130	H50	Individual wires in tubes	35	166.7	29.7
Ti 130	H50	Individual wires in tubes	41	168.9	28.8
—	H50	Rod	100	335.2	42.4
Ti 130	—	Rod	0	89.4	19.8
Ti 230	H50	Bundled wires	23	211	39.2
Ti 318	H50	Bundled wires	25	217	41.6
Ti 318	Alloy C*	Bundled wires	33	184	32.9
Ti 130	En 30B	Swiss roll	27	122.4	22.7
Ti 130	En 30B	Swiss roll	30	126.3	22.4
Ti 130	En 30B	Swiss roll	36	134.5	24.1
Ti 130	En 30B	Swiss roll	50	144.5	23.3

* 2.3 C, 0.35 Mn, 13.0 Cr, remainder Fe.

In order to raise the composite strength levels titanium alloys were used instead of commercial purity titanium, but since titanium alloy tubes were not available it was necessary to revert to the method of bundling wires into a large titanium tube.

Finally, a different form of rod composite was produced by rolling thin sheets (0.003 in.) of annealed En30B steel and Titanium 130 into the form of a 'swiss roll', which was placed inside a titanium tube, cold swaged, encapsulated and ausformed.

The tensile properties of these various rod composites are given in Table 5. The maximum room temperature specific strength achieved was 42×10^3 psi in a Titanium 318—25 per cent volume fraction H50 rod. The steel wire in this composite had been ausformed with a reduction in area of 85 per cent, giving a steel tensile strength of about 400×10^3 psi. The maximum strength quoted for ausformed H50 steel is 437×10^3 psi⁷ obtained after isothermal deformation with high deformation rates and ideal quenching conditions. Although the athermal treatment given to the steel in the composite is somewhat different, strengths approaching the maximum level have been achieved.

The variation of composite tensile strength with testing temperature in the range room temperature to 400°C is shown in figure 9. The fall in strength in this temperature range is relatively small, the Titanium 318-H50 steel composite having a specific strength at 400°C which is greater than the room temperature specific strength of Titanium 318. The strength of the H50 steel in the composite at 400°C is estimated to be 314×10^3 psi.

Fatigue properties for Titanium 130-En30B steel

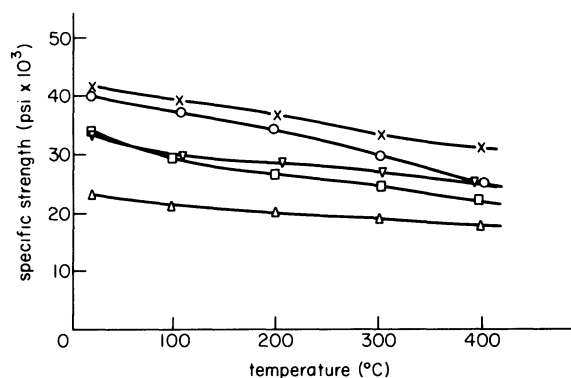


Figure 9. Variation of specific tensile strength with temperature for various titanium-steel composites.

x, v Ti 318-H50, o Ti 230-H50, o Ti 318, Δ Ti 230

rod composites are shown in figure 10, together with tension-compression curves for Titanium 130 in a condition similar to that of the titanium in the composite. At low alternating stresses around the endurance limit of the matrix, the plastic strain in the matrix is not sufficient to ensure that reinforcement by the steel is effective and failure occurs due to fatigue cracking in the matrix in the usual manner. At higher stress levels the steel wires are more effectively loaded and some reinforcement occurs. Some improvement in fatigue strength of the composites may also be due to deflection of the propagating fatigue crack at the interface between the titanium and steel.

One of the advantages generally associated with composite material is enhanced toughness. A comparison of impact resistance of a Titanium 130-En30B composite with Titanium 130 was made using the Charpy impact test over the temperature

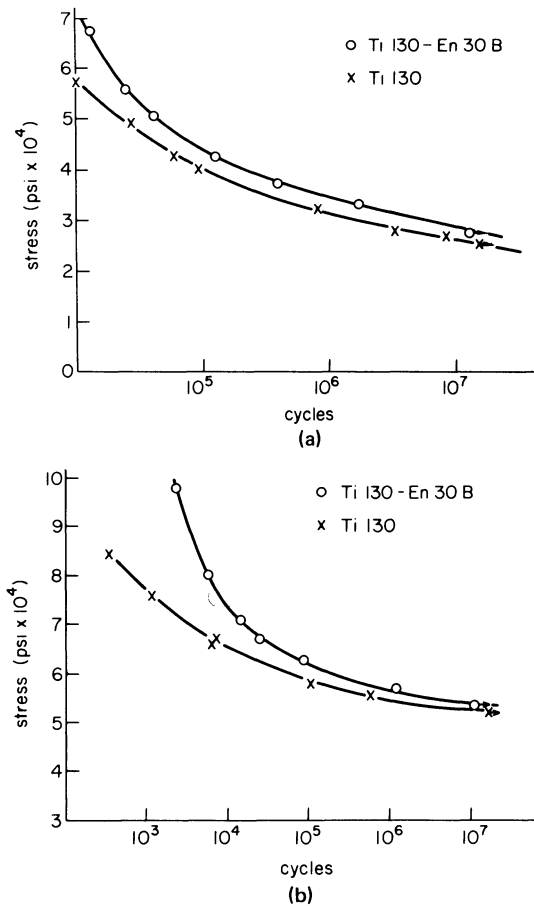


Figure 10. (a) Tension-compression (zero mean stress) fatigue results for Titanium 130-En30B steel rod composites. (b) Tension-tension (zero minimum stress) fatigue results for Titanium 130-En30B steel rod composites.

range $-196-600^{\circ}\text{C}$ (figure 11). Crack deflection at titanium-steel interfaces is effective in adding to the impact resistance of Titanium 130 in the temperature range below 150°C . At higher temperatures the impact resistance of titanium increases rapidly, but the steel passes through a relatively brittle condition at around 350°C ⁸. This probably results in fracture of the steel wires ahead of the propagating crack and

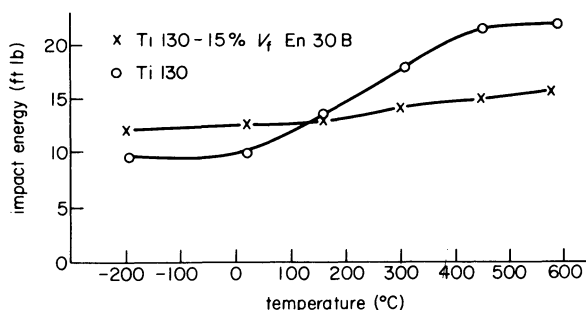


Figure 11. Charpy impact test results for Titanium 130-En30B steel rod composites.

leads to lower impact resistance of the composite compared with the matrix. However, the composite is not seriously embrittled at these higher temperatures and it provides reasonable fracture toughness over the complete temperature range.

CONCLUSIONS

It has been demonstrated that metal-metal composites with good combinations of mechanical properties can be made by co-working the components. When the flow strengths and work hardening rates of the constituents are widely disparate, the processing conditions for manufacture must be closely defined. Plastic instability of the strong component is a limiting factor both in manufacture and in subsequent working of these composites. High specific strengths have been achieved in combinations of aluminium alloys and steel. Under the conditions of alternating stress examined, main benefits accrued in laminated composites subjected to varying tensile stresses.

Sheet metal forming tests on aluminium alloy-titanium and laminates revealed texture strengthening under biaxial conditions, but rather low ductility. Aluminium alloy-stainless steel laminates had excellent biaxial ductility.

The feasibility of ausforming alloy steels *in situ* in a titanium matrix has been demonstrated. Problems of contamination at the interface suggest that the best configuration for this system is wire reinforcement.

ACKNOWLEDGMENT

This work was supported by the Ministry of Defence.

REFERENCES

1. A. Kelly and G. J. Davies. The principles of the fibre reinforcement of metals. *Metallurgical Reviews* (1965) **10**, 1.
2. D. R. Milner and G. W. Rowe. Fundamentals of solid-phase welding. *Metallurgical Reviews* (1962) **7**, 433.
3. T. H. Courtney and J. Wulff. *J. Materials Sci.* (1966) **1**, 383.
4. D. Lee and W. A. Backofen. *Trans. Met. Soc. AIME* (1966) **236**, 1077.
5. V. F. Zackay and W. M. Justusson. *ISI Special Report No. 76*, 1962.
6. O. Kubaschewski and B. E. Hopkins. *Oxidation of Metals and Alloys*. Butterworths, London, 1962.
7. J. J. Irani and P. R. Taylor. *ISI Special Report No. 108*, 1968.
8. J. Woolman and R. A. Mottram. *The Mechanical and Physical Properties of the British Standard En Steels*. Pergamon, Oxford, 1966.

CHARACTERISTIC FEATURES IN THE HOLE FLANGING AND PIERCING OF THIN AND THICK CIRCULAR PLATES USING CONICAL AND OGIVAL PUNCHES

by

W. JOHNSON*, N. R. CHITKARA* and P. A. BEX*

SUMMARY

The quasi-static piercing of mild steel, copper and brass plates having a thickness of $\frac{1}{8}$, $\frac{1}{4}$ and $\frac{1}{2}$ in, respectively was carried out using both conical and ogival punches $\frac{1}{2}$ in diameter; the semi-angle of the conical punches used was between 15° and 90° and the ogival punch ratio, that is, the height of the head to base radius, varied from 1 to 4. Autographic records of punch load against punch travel are presented and compared one with another and the deformation modes observed at different stages of punch penetration are discussed. Experimental results show that fracture directions at the tip of a lip are related to the planar anisotropy of the material.

INTRODUCTION

Most of the research into the punching, piercing, drifting or hole-flanging of plates has been carried out using conical punches and relatively thin material, as the fairly comprehensive review of the available literature shows¹. However, the important work carried out by Wilken² came only recently to the authors' attention. Published literature is available on the blanking of thick plates³⁻⁵, both at high and at low speed, using flat-ended punches and some work has been reported on the quasi-static perforation of relatively thick plates using conical or ogival punches⁶⁻⁹.

Quasi-static investigations are important in connection with armour penetration because they throw light on the behaviour of thick plates under impact conditions, suggest appropriate deformation modes, and indicate how they are related to punch load during penetration.

The present work adds to that previously reported¹ to include thin and thick plates penetrated by both conical and ogival punches. These punch shapes were chosen to facilitate broad experimental comparisons of their load-punch travel characteristics and, in particular, the maximum punch load.

EXPERIMENTAL WORK

Mild steel and copper plates of $2\frac{1}{2}$ in outside diameter, with thicknesses ranging from $\frac{1}{8}$ to $\frac{1}{2}$ in and central holes of zero, $\frac{1}{8}$ or $\frac{1}{4}$ in diameter, were used in the piercing experiments. The semi-angles of the conical punches were 15° , 30° , 45° , 60° , 75° and

90° . For the ogival punches, the height to base radius ratio β varied from 1 to 4 (see figure 1). Since almost all the work material used in this series of tests exhibited a small amount of anisotropy, experimental

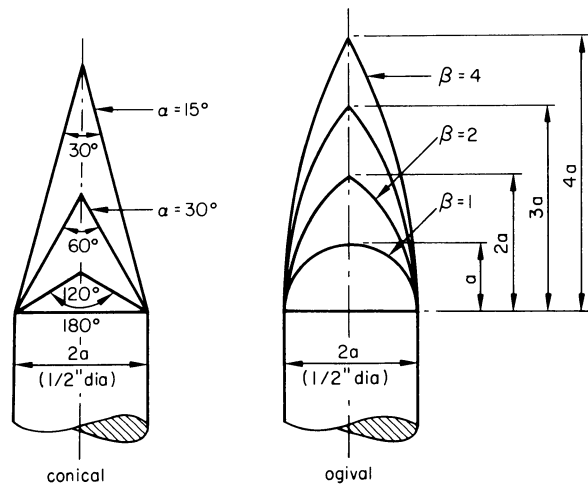


Figure 1. Shape of piercing punches.

work was also done on $\frac{1}{8}$ in thick brass specimens to correlate the number and the directions of lip fractures to the planar anisotropy of the material.

Preparation of the test specimens

Mild steel and copper were chosen as the two main test materials, but brass was chosen for tearing or lip fracture tests, since it was expected that it would show pronounced anisotropic behaviour.

* Division of Applied Mechanics, Department of Mechanical Engineering, the University of Manchester Institute of Science and Technology

The $\frac{1}{8}$ in thick specimens were prepared from plates cut as $2\frac{5}{8} \times 2\frac{5}{8}$ in squares and turned accurately to a diameter of $2\frac{1}{2}$ in. The $\frac{1}{4}$ and $\frac{1}{2}$ in thick specimens were machined to an accuracy of 0.003 in from bar stock supplied originally at $2\frac{1}{2}$ in outside diameter. Specimens of mild steel and copper circular plates were prepared either with no central holes or with holes of $\frac{1}{8}$ or $\frac{1}{2}$ in diameter, drilled and reamed.

Before machining, the $\frac{1}{8}$ in thick brass plates had two parallel lines marked on them in the direction of rolling.

Sub-press and die assembly: punches and lubrication

The general shape, construction and layout of the sub-press and die assembly was the same as that previously described¹. Punches were of two different shapes. One set had a conical head and a cylindrical shank, the other set had an ogival head with a cylindrical shank. All punches were made from S.23 high carbon steel, hardened, tempered and ground. The shape and dimensions of the punches are shown in figure 1.

Almost all the tests were carried out using a graphite based grease as lubricant. This was applied generously to both the specimen and the punch. To compare the effect of lubricants, that is the work done against friction, a series of tests was also performed using a thin film of polytetrafluorethylene (PTFE) as the lubricant medium.

Experimental procedure

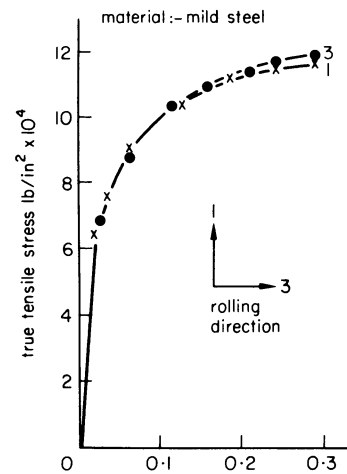
The procedure for the drifting and piercing tests was similar to that previously described¹. Specimens were tested to different stages of punch travel, as indicated by the load-punch travel graph. The load-punch travel history was recorded autographically on a testing machine.

Material properties: stress-strain curves

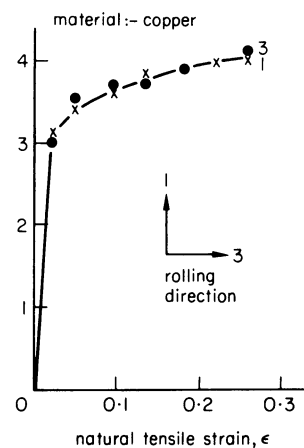
To determine the true stress-natural strain curves for the materials, suitable tensile specimens of 1 in gauge length were prepared. For material originally $\frac{1}{8}$ in thick, in the case of mild steel and copper, two specimens were prepared for each material, one in a direction parallel to the direction of rolling and the other perpendicular to it. Three specimens were however prepared from the brass plate, one in a parallel direction, one perpendicular to and the other at 45° to the direction of rolling. It was believed that these three chosen directions for brass would yield different flow stresses and therefore would easily demonstrate the anisotropic quality of the brass plates. In the case of the mild steel and copper, as originally supplied in the form of bar stock, two specimens of each were also prepared, the orientation of one being in the axial direction and the other in a transverse direction. The true stress-natural strain curves for these materials are shown in figures 2(a), (b) and (c).

Hardness

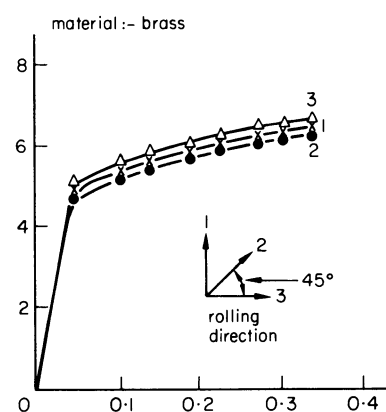
One specimen of each material was taken at random and the Vickers hardness value was measured at different points on the cross-section of the $\frac{1}{2}$ in thick specimens at the penultimate stage of deformation, as



(a)



(b)



(c)

Figure 2. Stress-strain curves of plate materials: (a) mild steel; (b) copper and (c) brass.

indicated on the load-punch travel graph. Typical hardness (number) distributions for copper specimens deformed by conical, ogival and flat punches are shown in figure 3(a). The hardness values of the copper plate before deformation are shown in figure 3(b).

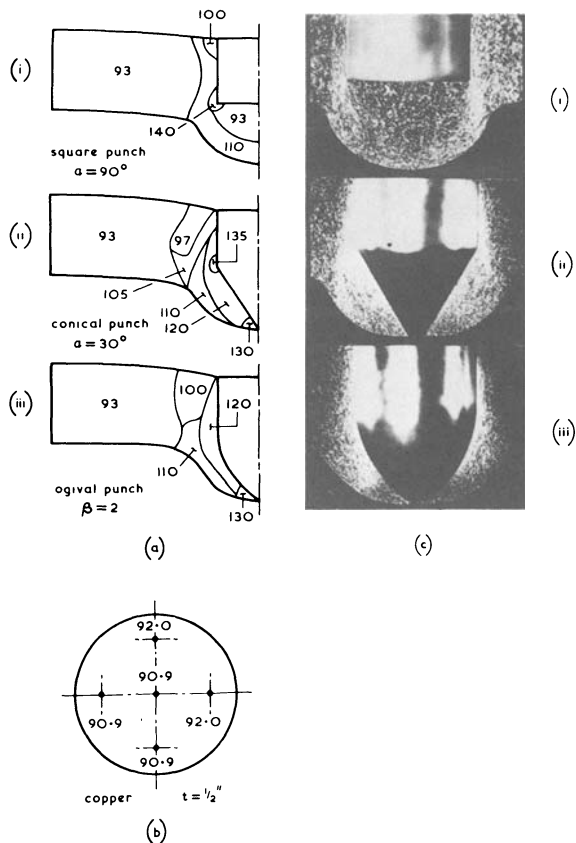


Figure 3. (a), (b) Hardness distribution (VHN) material: copper; $t_0 = \frac{1}{2}$ in. (a) Deformed specimen and (b) undeformed specimen.

RESULTS

Quasi-static drifting and piercing of plates

Typical experimental results obtained from the present series of tests on quasi-static drifting and piercing of plates¹⁰ are shown in figures 4 to 8.

Figures 4(a) and 4(b) show, respectively, typical autographic load-punch travel records when copper and mild steel plates of $\frac{1}{8}$ in thickness are drifted with conical punches of cone semi-angle $\alpha = 30^\circ$, the equivalent β value being 1.73. For comparison, load-punch travel results for drifting with ogival punches of height to base ratio $\beta = 2$ are shown in figures 4(c) and 4(d). Progressive changes in the shape of the cross-section of a copper plate during drifting with ogival punches for which $\beta = 2$, are also shown in figure 4(c), while sectional views of the plates at different depths of penetration appear in figure 5.

In figures 4(b) and 4(d), the effect of lubricating mild steel specimens with PTFE, which causes a significant drop in residual frictional load at the end of punch penetration, are shown marked as point 7 on the curve. Figures 6(a) and 6(b) show the punch load-punch travel curves for copper plates of $\frac{1}{4}$ in thickness when drifted with punches of cone semi-angle $\alpha = 30^\circ$ and ogival punches of $\beta = 2$, respectively, along with the progressive changes in the deformation modes of the plates due to gradually increasing penetration for an initial hole of $2b = \frac{1}{8}$ in. The stages are those marked on the load-punch travel curves. Sectional views of the plates drifted with

ogival punches are shown in figure 7 for $\frac{1}{4}$ in thick plates.

Variation of the punch load with travel for $\frac{1}{2}$ in thick mild steel specimens penetrated with conical punches of cone semi-angle $\alpha = 30^\circ$ are shown in figure 8(a), along with changes in deformation mode with increased penetration where the initial hole in the plate is $2b = \frac{1}{8}$ in; corresponding stages are marked on the diagram. In figure 8(b), the characteristic load-penetration curves obtained during the piercing of $\frac{1}{2}$ in thick copper plates by a flat punch and comparable conical and ogival punches are shown for comparative purposes. Figures 9 and 10 show the cross-section of deformed copper and mild steel specimens of, respectively, $\frac{1}{2}$ in thickness, at different stages during gradually increasing punch penetration, when piercing by ogival, conical and flat punches.

Figure 11 shows the top view of four of the brass plates $\frac{1}{8}$ in thick after piercing by ogival punches of $\beta = 1, 2$ and 3 and a conical punch of $\alpha = 30^\circ$. The views of the plates illustrated in figure 11 for ogival punches are those marked as point 6 on the load-punch travel curves in figure 12. The relationship of the fracture directions to the rolling direction, which is marked on the specimens, can be clearly seen. The load-penetration graphs for brass plates, when these are drifted by ogival punches of different shapes, are shown in figure 12.

OBSERVATIONS

(1) Solid plates: $2b = 0$

For both the conical ($\alpha = 30^\circ$) and ogival punches ($\beta = 2$) the plate remains flat as indentation proceeds from the first penetration; this is stage 1. In stage 2, the plate starts to collapse as a whole so that plate indentation and plate deflection (bending) thereafter proceed concurrently. The load-punch travel curve is then slightly concave upward and is reminiscent of typical curves of load against large deflection for plastically deformed plates under central transverse load. The maximum load is attained when the punch shank is about to pass through the plate and thereafter, stage 3, there is a rapid decrease in load (with rapid lip formation) until the punch head has passed wholly through the plate. The 'residual', more or less constant, load on the punch, stage 4, is due to friction between the shank and the metal plate and depends on the degree of springback or elasticity of the plate material. The figures show that using PTFE on the shank reduces the residual load.

Due to the limited ductility of real metals, radial fractures are created during the perforation process. These may occur in stage 2 and in that case the maximum punch load is less than what it would be had the fracture not occurred. Fractures often occur soon after maximum load is reached. There are usually four fractures and when fully developed they appear as 'petals', forming the lip.

The load-punch travel characteristics for comparable conical and ogival punches differ slightly in that for the latter the characteristic in stage 2 steepens more rapidly than is the case with the former, presumably because the effective punch angle increases as perforation proceeds. After maximum

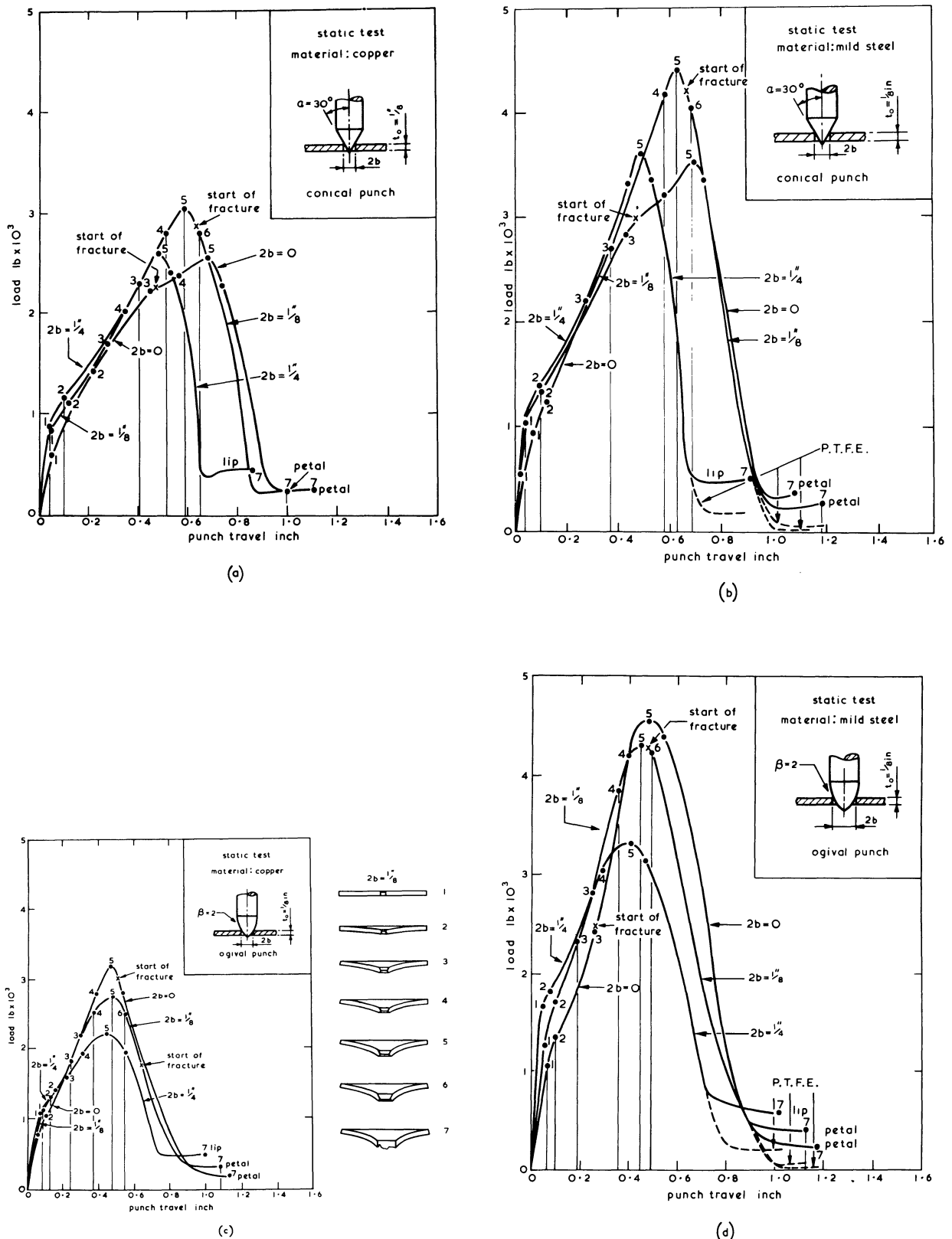


Figure 4. Load-punch travel characteristics and deflection patterns for copper and steel specimens when $t_0 = \frac{1}{8}$ in.

load in stage 3, the characteristic is straighter for the ogival than the conical punch.

Figures 4, 6 and 8 show that the thicker the plate penetrated, the less distinct is the change from stage 1 to stage 2.

Ogival punches give rise to a characteristic which, near to the maximum load, is more rounded than in

the case of the conical punch. This maximum also occurs sooner than in the case of the conical punch.

The total punch travel to the onset of stage 4 is very insensitive to plate thickness and is of about the same magnitude for plates for which $t_0 = \frac{1}{8}, \frac{1}{4}$ and $\frac{1}{2}$ in. Whether $\alpha = 30^\circ$ or $\beta = 2$, the maximum punch load is little affected.

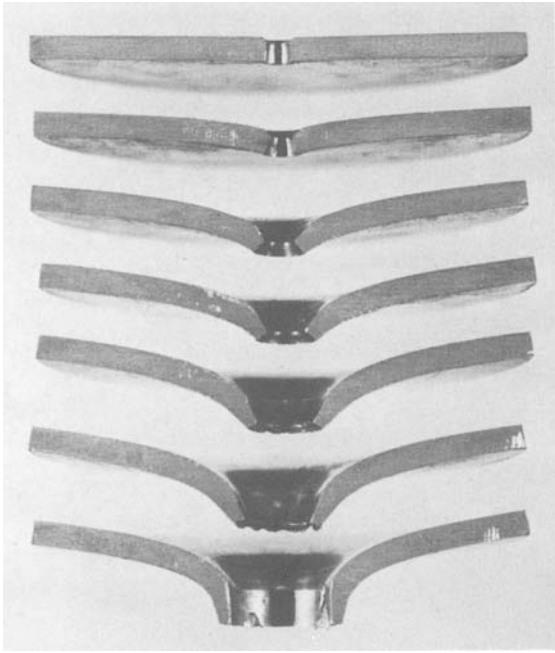


Figure 5. Sectional views of copper specimens when $\beta = 2$.

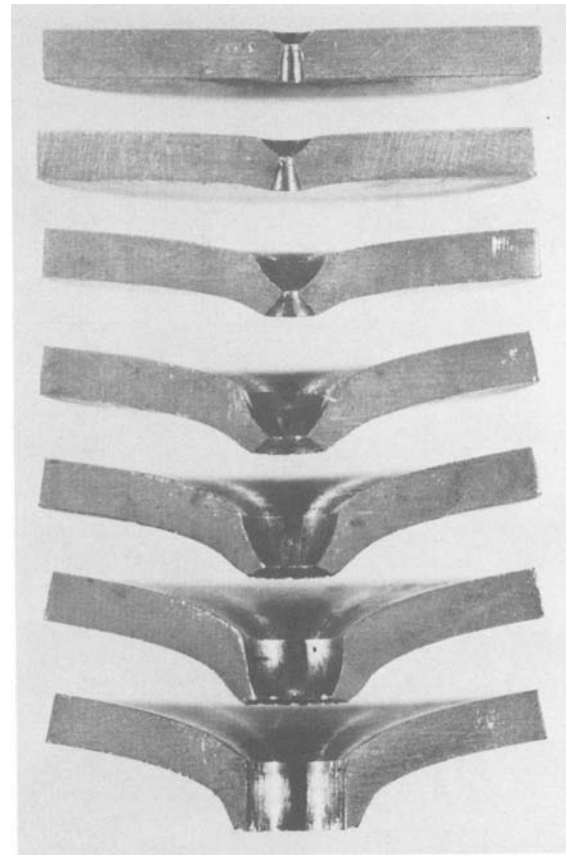
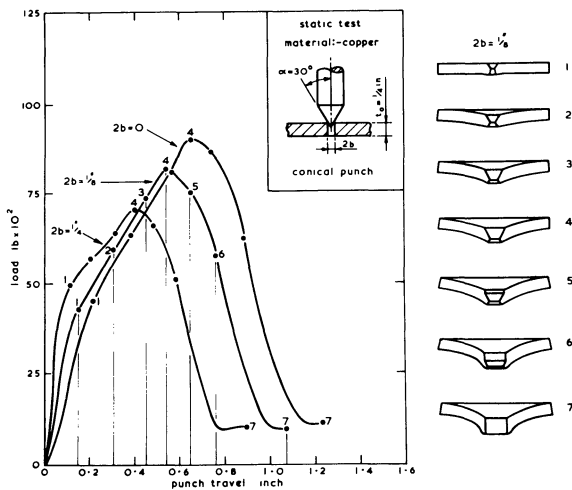
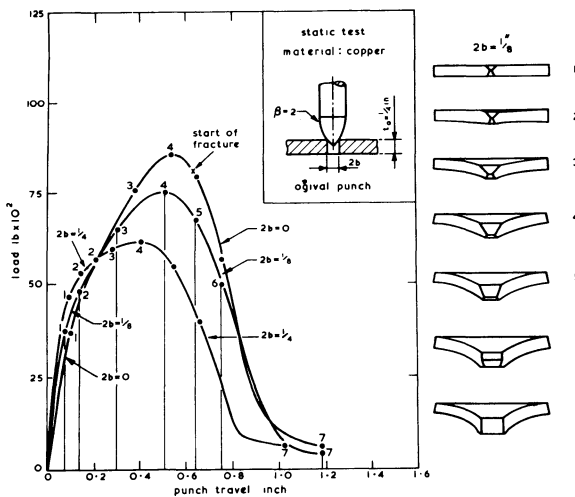


Figure 7. Sectional views of copper specimens, $\beta = 2$, $t_0 = \frac{1}{4}$ in, $2b = \frac{1}{8}$ in.



(a)



(b)

Figure 6. Load-punch travel characteristics and deflection patterns for copper specimens when $t_0 = \frac{1}{4}$ in.

(2) Plates containing holes: $2b \neq 0$

The same four stages are to be identified in the load-punch characteristic as above, though the change from stage 1 to stage 2 is the sharper the larger is b .

As would be expected, the maximum punch load is the larger, the smaller is b and the later in the process does fracturing occur, if it occurs at all. Maximum load occurs the sooner, the larger is b .

(3) An interesting observation from figure 8(b) is the fact that to penetrate a plate a flat punch ($\alpha = 90^\circ$) may require a lower load than a conical or ogival one; the required energy to fully perforate the plate may also be very much less, and this point is verified in figure 12. This feature is associated with plug formation in the one (shearing) and energy absorbing lip formation in the other two (i.e. mainly bending).

(4) The distribution of hardness in a formed lip (see figure 3) is entirely what would be expected. The hardness number will be the higher the more worked is the element of metal concerned, and the nearer the tip of the lip, the greater is the applied strain.

(5) Direction of the fractures for petal formation

The true stress-natural strain curves in figure 2(c) show in one way the anisotropic nature of the brass used, while the same is shown in another way in figure 11; the four splits in the brass plates when using different punches are clearly seen to be parallel with, or almost perpendicular to, the rolling direction of these specimens. The rolling direction was initially marked by two parallel lines. This indicates clearly

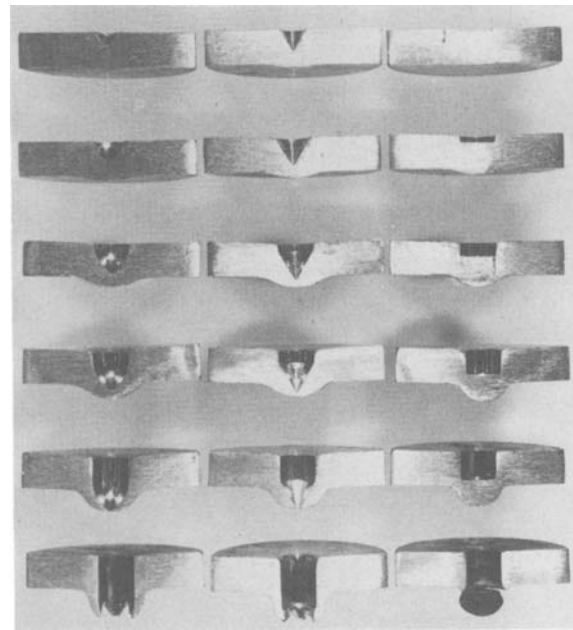
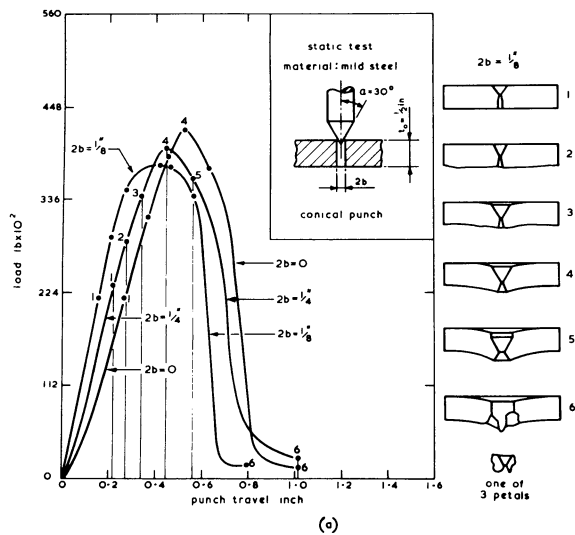


Figure 9. Sectional views showing progressive deformation in piercing of thick copper plates; $t_0 = \frac{1}{2}$ in. (i) Ogival punch, $\beta = 2$; (ii) conical punch, $\alpha = 30^\circ$ and (iii) flat punch.

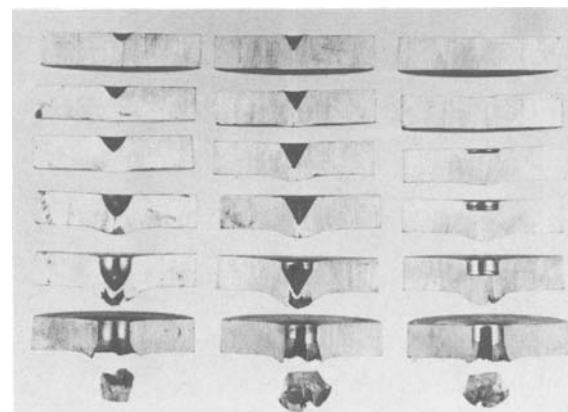
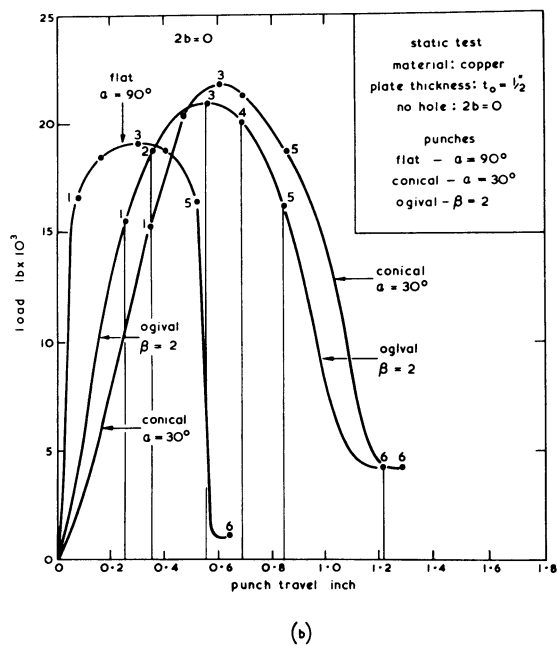


Figure 10. Sectional views showing progressive deformation in piercing of thick steel plates; $t_0 = \frac{1}{2}$ in. (i) Ogival punch, $\beta = 2$; (ii) conical punch, $\alpha = 30^\circ$ and (iii) flat punch.

Figure 8. (a), (b) Load-penetration characteristics for $t_0 = \frac{1}{2}$ in.

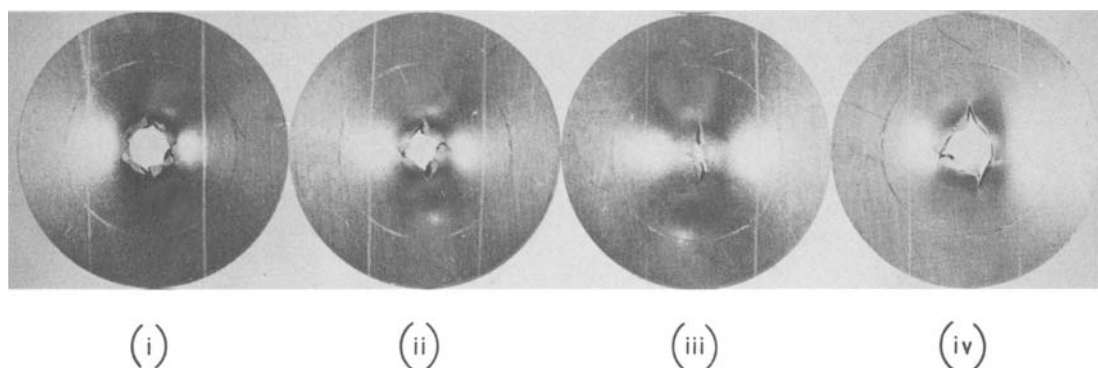


Figure 11. Top views of brass specimens showing lip fracture at stage 6 of the load-penetration curves; $t_0 = \frac{1}{8}$ in. (i) Ogival $\beta = 4$; (ii) ogival $\beta = 3$; (iii) ogival $\beta = 2$, and (iv) conical $\alpha = 30^\circ$.

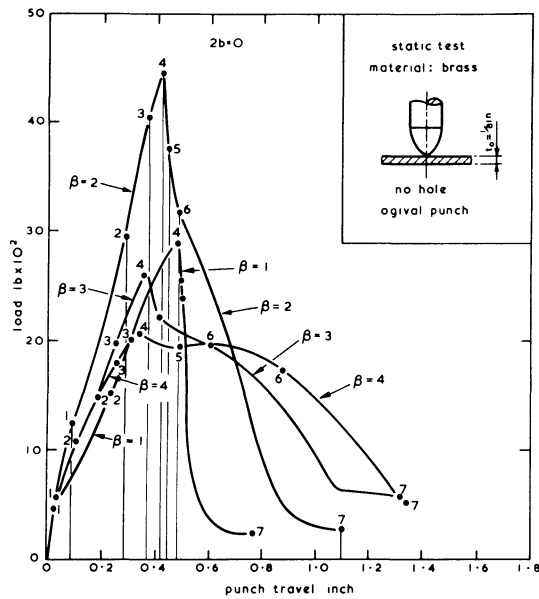


Figure 12. Load-punch travel curve for an ogival punch; brass; $t_0 = \frac{1}{8}$ in.

the link between the planar anisotropy of the sheet material and the number and direction of lip fractures in the drifting process.

ACKNOWLEDGMENTS

The authors wish to thank Mr G. Robinson and all other members of the Mechanical Engineering work-

shop staff for their assistance in the preparation of the equipment and the specimens. They would also like to express their thanks to Mr A. Bowers for his help with the drawings and diagrams. The authors are grateful to Sheet Metal Industries for their financial support of this work.

REFERENCES

1. W. Johnson, N. R. Chitkara, I. H. Ibrahim and A. K. Dasgupta. *J. of Strain Analysis*, Inst. of Mech. Engrs., London (1973) 8, 3, 228.
2. R. Wilken, *Das Biegen von Innenborden mit Stempeln*. Nr. 794, Westdeutscher Verlag, 1959.
3. R. A. C. Slater and W. Johnson. *Int. J. Mech. Sci.* (1967), 9, 271.
4. W. Johnson and P. B. Mellor. *Engineering Plasticity*. Van Nostrand, p. 624, 1973.
5. W. Johnson. *Impact Strength of Materials*. Ed. Arnold, 1972.
6. R. Hill *Mathematical Theory of Plasticity*. Oxford University Press, p. 106, 1950.
7. A. Brown. *Int. J. Mech. Sci.* (1964), 6, 257.
8. W. Goldsmith, T. W. Liu and S. Chulay. *Expt. Mech.* (1965), 5, 385.
9. N. Cristescu. *Dynamic Plasticity*. North Holland Publishing Co., Amsterdam, p. 614, 1967.
10. P. A. Bex. M.Sc. Dissertation, University of Manchester Institute of Science and Technology, 1972.

THE INFLUENCE OF HYDROSTATIC PRESSURE ON THE PLASTIC DEFORMATION OF METALLIC MATERIALS

by

WALTER LORREK* and OSKAR PAWELSKI*

SUMMARY

The pressure dependence of the flow and fracture behaviour of different metals and alloys has been studied at room temperature and constant strain rate. In a high-pressure apparatus designed to perform tensile and compression tests as well as some pattern-working processes, investigations were made under a hydrostatic pressure which could be continuously adjusted up to 15 kbars and held constant during deformation. Stress and strain could be obtained from direct measurement of specimen load and machine displacement during the tests. The materials investigated were 42 CrMo 4 steel, TiAl 6 V 4 titanium alloy, copper, zinc, cast iron (GG 25) and Fe₃Si single crystals.

For the ductile materials (steel, titanium alloy and copper), the ultimate deformability defined by the logarithm of the maximum amount of necking in the tensile test shows an increase with hydrostatic pressure which can be approximated by a linear relationship. The elongation until the beginning of necking remains unchanged under an environmental pressure. The flow stress was found to increase slightly with pressure up to about 1 per cent per kbar without any significant dependence upon material or degree of deformation. Using different fracture criteria the pressure-induced increase of ductility can be predicted from the mechanical data of an ordinary tensile test.

For the brittle materials investigated (zinc, cast iron, embrittled copper and Fe₃Si), the pressure dependence of the true strain at fracture is non-linear at first, but at higher strains it is linear as in the case of the ductile specimens. However, the increase in ductility per unit pressure for the brittle materials is not as large as for the ductile materials, and cannot be as uniquely described by criteria. The toughness of zinc shows a rather sudden transition from brittle to ductile behaviour. The brittle fracture in an ordinary tensile test yields a complete necking above a threshold pressure depending on the grain size.

NOTATION

		σ_m (N/mm ²)	mean normal stress (= $\frac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3)$)
D (mm)	grain size		
L (N)	load	σ_s (N/mm ²)	yield point
R (mm)	radius of curvature at neck	σ_B (N/mm ²)	tensile strength
a (mm)	outside radius of cross section of neck	σ_F (N/mm ²)	true flow stress at fracture
e	base of natural logarithm	σ_R (N/mm ²)	true fracture stress
p (kbar)	hydrostatic pressure ($p > 0$)	$\sigma_{R\text{ eff}}$ (N/mm ²)	effective true fracture stress (= $\sigma_R - p$)
r (mm)	current radius of cross section of neck ($0 \leq r \leq a$)	σ_{HI} (N/mm ²)	true stress at maximum load of tensile test
α (kbar ⁻¹)	= $\Delta\phi_{Br}/p$, increase in true strain at fracture per unit volume	ϕ	true (= logarithmic) strain
β (N/mm ²)	= $(d\sigma/d\phi)_{\phi \gg \phi_g}$, strain hardening coefficient assumed constant at large strains	ϕ_g	true strain until the beginning of necking
σ (N/mm ²)	(true) flow stress	ϕ_{Br}	true strain at fracture
σ_1 (N/mm ²)	} main normal stresses	$\Delta\phi_{Br}$	increase in ductility (= $\phi_{Br}^{(p)} - \phi_{Br}^{(0)}$)
σ_2 (N/mm ²)		Superscript (p)	means that an expression is associated with a hydrostatic pressure p
σ_3 (N/mm ²)		Superscript (0)	means that an expression is associated with atmospheric pressure

* Department of Metal Forming, Max-Planck-Institut für Eisenforschung, Düsseldorf, Germany

INTRODUCTION

The ultimate plastic deformability of a material is a function of stress configuration, strain rate and temperature. Among these, the stress configuration is the most significant. Its influence is characterized by σ_m , the mean normal stress. Compressive forces generating a low σ_m enhance ductility, while tensile ones with a high σ_m considerably reduce it. Thus in order to avoid an early failure of the material during deformation, a metal-forming process with a low σ_m may be preferred to one with a high σ_m . In a number of technical processes like rolling, forging or extrusion, the mean stress is less than zero and can be regarded as a hydrostatic pressure caused by that particular procedure.

When attempting to deform a brittle material, the mean stress of the procedure itself may not be low enough to avoid cracks or fracture. In this case σ_m can be deliberately lowered by superimposing a hydrostatic pressure. Using this additional technique, normally brittle materials have been successfully deformed and the ductility of normally ductile specimens has been extraordinarily enlarged. For example, on aluminium reductions in area up to 14000:1 have been reported (ASEA high pressure laboratories).

The object of the investigations discussed here was to study the influence of a widely variable hydrostatic pressure on the flow and fracture behaviour of different materials with special regard to forming technology. For metal-working the most important significance of pressure is, of course, the increase in ductility. This matter will be discussed in detail. However, pressure also influences the material itself. Changes in strain-hardening mechanism and fracture appearance become obvious. Experimental evidence resulting from those effects which are more important to metal physics than to metal forming will be mentioned briefly.

EXPERIMENTAL PROCEDURE

The tensile test with superimposed hydrostatic pressure meets the important requirements of a metal-working process for the study of pressure dependence on the deformation behaviour of a material. The basic state of stress is quite simple during uniform elongation and its calculation during necking is well established. Continuous data acquisition is possible to calculate stress-strain curves. In addition, compression tests under environmental pressure are also favourable for comparison with the results from tensile tests and for materials from which no tensile specimens could be machined (for example, single crystals with specific orientation).

A high-pressure apparatus was designed and built which enables a one-dimensional advance in a pressurized environment (figure 1). The flange side of the high-pressure cylinder (A) mounted on the base plate (B) is connected via three link rods (S) to a fixed plate (F). This comprises the local fixed part of the apparatus which is placed under a 1500 kN hydraulic press for pressurization during the test. The movable assembly consists of two plates (M) which are connected by three rods (R), and two load cells and high-pressure pistons (D and Z) which are attached to the plates.

Figure 2 shows the internal arrangement of the pressure chamber. Either a tensile specimen can be screwed into the front part of the split high-pressure piston (Z) and the specimen holder, or a compression specimen can be placed between two small plates which are connected by three springs around the workpiece in 120° symmetry, thus exerting a slight force on the specimen to keep it in place before deformation. After bringing the pistons D and Z into position, glycerine is poured through the vertical hole which is connected by two smaller holes with the two horizontal parts of the pressure chamber. Now the

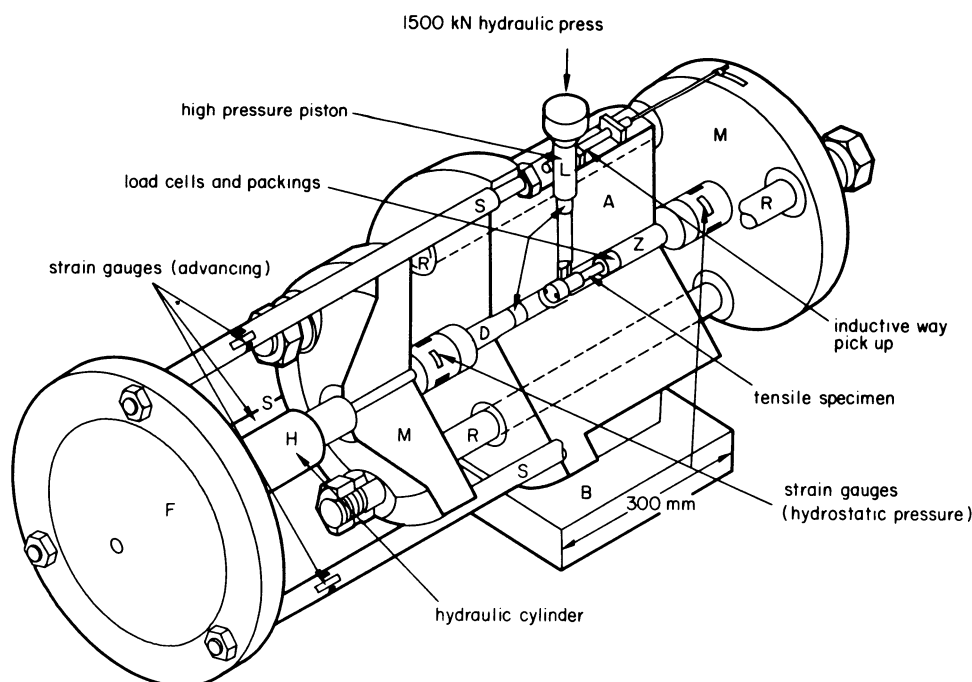


Figure 1. General arrangement of hydrostatic pressure apparatus up to 15 kbars.

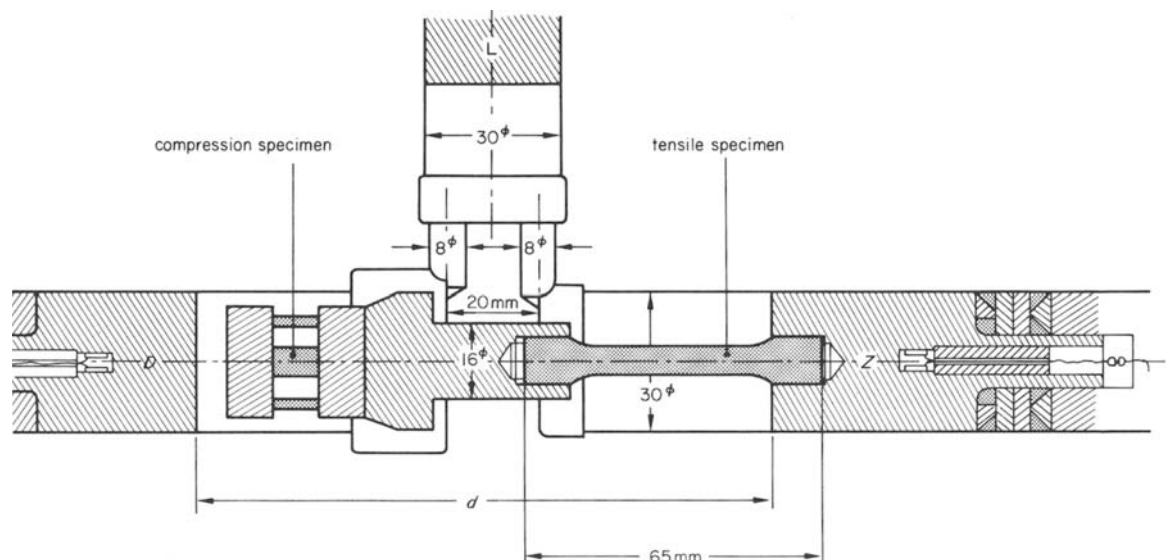


Figure 2. Internal arrangement of pressure chamber.

liquid is pressurized by the force provided by the external hydraulic press and transferred through the vertical high-pressure piston (L). The movable assembly is still at rest at this time because the forces exerted on the pistons D and Z, which have the same area, are identical. To start the deformation procedure the movable assembly is pushed by the hydraulic cylinder (H, figure 1) causing the two horizontal pistons (D and Z) to move to the right (figure 2) thus either elongating a tensile specimen or shortening a compressive one. Since the distance d between the two horizontal pistons is fixed, the high-pressure volume is not changed during deformation, thus guaranteeing an absolutely constant pressure throughout the test.

After testing different sealing systems, split high-pressure pistons were finally used (figure 3). They prevent leakage up to 15 kbars with perfect reliability. Their sealing technique is based on the principle of the 'unsupported area'¹ with a packing which consists of several rings of different materials. The strain gauges attached to the front parts of both horizontal pistons (figure 2) can be connected to the measuring amplifier in one of two different ways:

- (1) After pressurization but before deformation, either one of them can be connected as one branch of a Wheatstone bridge in order to measure the hydrostatic pressure as the force acting on an area of known value.
- (2) During the test the strain gauges of pistons Z (for 'tensile' side) and D (for 'compressive' side) are connected as opposite branches in a Wheatstone bridge, thus eliminating the signals due to hydrostatic pressure. The net signal then is a measure for the specimen load coming either from piston Z in a tensile test or from piston D in a compression test.

The advantage of this acquisition technique is the ability to measure both the load and the pressure free of error due to friction of the packing, nevertheless avoiding the need to put electrical wires into the high-pressure chamber.

The strain gauges attached to the link rods (S, figure 1) for measuring the load of advancing are only used in tests under atmospheric or low environmental pressures (up to 3 kbars) for the determination of the specimen load. In ordinary tensile tests the split pistons cannot, of course, be used and are replaced by unsplit ones. The constancy of the superimposed pressure during the test is checked by either one (depending on the kind of test) of the load cells attached to the moving plates (figure 1). The elongation or shortening of the specimen is obtained by the displacement of one of the movable plates with respect to the high-pressure cylinder using an inductive wave pick-up (LVDT). All tests were performed at room temperature and at an advance speed of 0.15 mm/s which equals a strain rate of 0.004/s at the beginning of a tensile test.

RESULTS AND DISCUSSION

Because of their different ductility behaviour it is convenient to distinguish between normally ductile and normally brittle materials. In this presentation a material will be called normally ductile when in an ordinary tensile test, that is without superimposed pressure, it shows a measurable amount of necking before failure. A material will be called normally brittle when it fails without necking in an ordinary tensile test.

Normally ductile materials

Tensile tests under different hydrostatic pressures were carried out with 42 CrMo 4 steel, TiAl 6 V 4 titanium alloy and copper. The specimens were 36 mm long and 6 mm in diameter. The materials were tested 'as received' and after various heat treatments. The results of these tests for the normalized steel are shown in figure 4, which is typical for all ductile materials investigated.

The true stress-true strain curve (figure 4(a)) of the ordinary tensile test is substantially enlarged by environmental pressure. The flow curves corres-

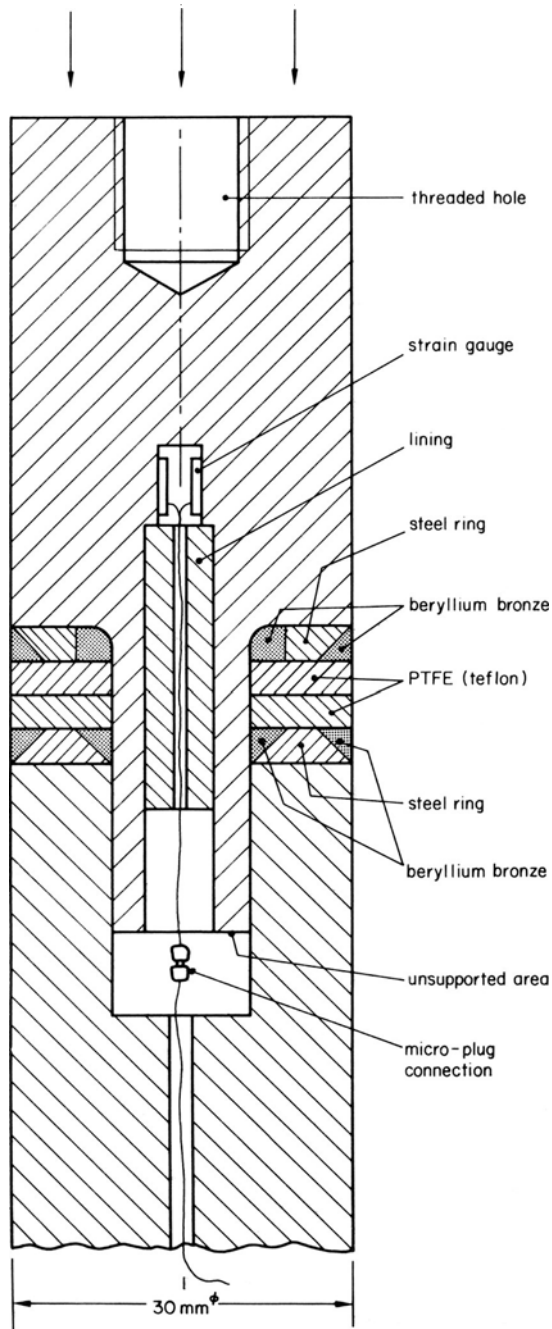


Figure 3. High-pressure piston with packing and measuring device for 'tensile' side.

ponding to pressures of 10 and 15 kbars run at higher flow stresses and to considerably higher strains at fracture. The solid line parts of the curves up to the end of uniform elongation (vertical broken line) are obtained from continuous measurements of load and elongation. The broken parts of the curves are defined by the slope at the beginning of necking, the end point indicating fracture, and by several points in between. The true stresses and true strains of these points are calculated from the measurements of the load, the radius of the curvature at the neck and the outside radius of the smallest cross section of the neck. The true flow stress was calculated from the axial stress using Siebel's formula² which is easier to handle than a similar one derived by P. W. Bridgman³. Within the range of the calculation, both

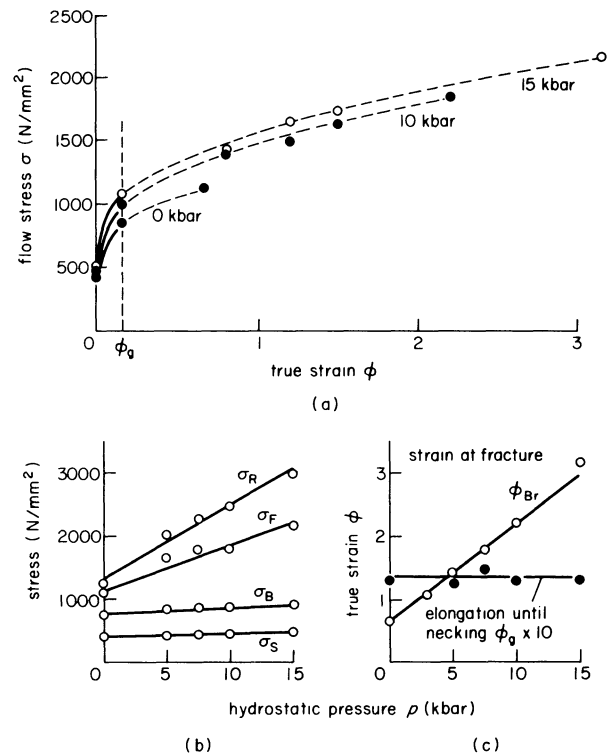


Figure 4. Stress-strain behaviour of steel (42 CrMo 4, normalized) in tensile tests under different environmental pressures.

formulas yield results of negligible difference. The diameter of the smallest cross section was measured like all other values of specimen geometry (original diameter, proof length and so on) by a travelling microscope with digital indication and an accuracy of $2\ \mu\text{m}$. The radius of curvature at the neck was measured with the help of a profile projector.

Figure 4(b) shows the pressure dependence of the characteristic stresses of the tensile test. For comparison with the ordinary tensile test all values presented reveal not the 'effective' stresses, that is the stresses actually present in the specimen, but those caused by the load after pressurization. This practice was, of course, also applied in figure 4(a) to the calculation of the flow stresses. The yield point σ_S and the tensile strength σ_B increase slightly with pressure as expected from the general rise of the flow stress for a given strain in figure 4(a). The true flow stress at fracture σ_F naturally increases with pressure which is mainly due to the additional strain hardening and also—by a much smaller amount—to the general rise described above. The increase of the true fracture stress σ_R with pressure is approximately the same as the rise of pressure itself. Thus one of the criteria to be tested later for a prediction of the pressure dependence of ductility will be that of a constant effective fracture stress:

$$\begin{aligned}\sigma_{R\text{ eff}} &= \sigma_R - p \\ &= \text{constant } (p > 0)\end{aligned}\quad (1)$$

The influence of pressure on the flow stress is small compared to its influence on ductility. In accordance with other authors the increase of the flow stress

amounts to about 1 per cent per kbar without any significant dependence on the strain. A survey⁴ on the pressure dependence of the flow stress reported in the literature yields the following conclusions. The increase in flow stress due to pressure for a given strain cannot be used for additional strain hardening because most of that increase is reversible. Two specimens pre-strained at atmospheric and at high hydrostatic pressure respectively show only a very small difference in their flow stresses when tested subsequently in an ordinary tensile test³. This difference, which is the irreversible part of the increase, can hardly be measured because it is of the same order of magnitude as the experimental error even if high pressures are used. It is due to structural changes during deformation under pressure⁵. One of the reasons for the larger, reversible part of the increase is the pressure dependence of the elastic constants. However, rises of the flow stress have frequently been found to be higher than would result from this dependence, so the change of the elastic values may be not the only reason for the reversible increase. An additional explanation may be a pressure dependence of the dislocation mobility as in the case of ionic crystals⁶. Decreasing mobility with increasing pressure causes a rise of the shear stress with pressure at constant strain rate. Although the pressure dependence on the dislocation speed is lower for a metal lattice than for an ionic crystal, a measurable influence on the flow stress cannot be excluded. Further investigations will be necessary to confirm this.

A measure of ductility is provided by the true strain at fracture ϕ_{Br} . In a tensile test it is defined by the logarithm of the ratio of the original cross section of the specimen and the smallest one of the neck at fracture. In a survey Martynov et al.⁷ present several different forms of pressure dependence of ductility observed so far. At first there is a linear relationship. In a second type the linear part is preceded by a non-linear beginning. In a third relationship the curve, which is linear at the beginning, turns towards the pressure axis at high pressures, indicating some sort of 'saturation'.

Figure 4(c) shows the pressure dependence of ductility for the steel investigated by the authors. Not only for steel but also for all other ductile materials tested after a number of heat treatments, a linear relationship between ϕ_{Br} and p was found to be the best approximation. From experimental evidence and from data of the literature (for example P. W. Bridgman³, H. L. D. Pugh⁸, M. Yajima et al.⁹) the ductility behaviour of a normally ductile material seems to be associated with a linear relationship between true strain at fracture and superimposed pressure.

The elongation up to the beginning of necking represented by the horizontal line in figure 4(c) is entirely unaffected by an environmental pressure.

The results on copper are shown in figure 5. They do not differ qualitatively from those on steel and titanium alloy. However, the amount of increase in ductility per unit pressure is substantially larger than that of the other materials (figure 5(a)). The ductility lines of annealed and cold worked ('as received') copper are parallel. The distance between them for a

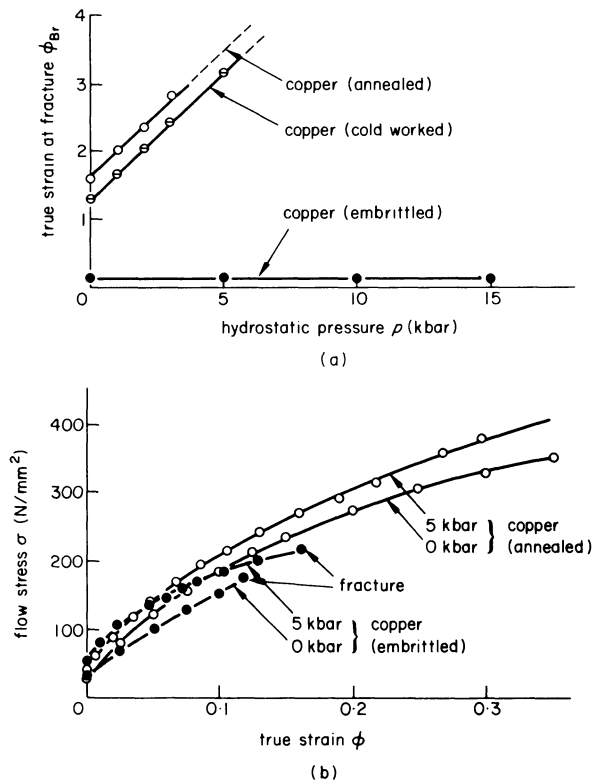


Figure 5. Ductility and true stress-true strain curves for copper under different hydrostatic pressures.

given pressure approximately equals their difference in true strain until necking ($= 0.35$). The investigation on copper becomes more interesting when viewing the results obtained from deliberately embrittled copper. Annealed specimens were exposed to nitric acid until small cracks appeared on the surface. In a subsequent ordinary tensile test the specimens failed at a true strain of only 0.12 (figure 5(b)) which is less than one-third of the strain until necking of the annealed material. Surprisingly, a superimposed pressure of 5 kbars caused hardly any rise in the ductility (0.16). Considering that the true fracture stress in this case of $p = 5$ kbars is about 210 N/mm², it is at first glance hard to understand that the specimen failed showing a macroscopically brittle fracture surface perpendicular to its axis, although the effective axial stress is a compressive and not a tensile stress. Even under 15 kbars the increase in ductility was so small that the graph of the data obtained (figure 5(a)) looks like a horizontal line when keeping the scale for the true strain at fracture. Without any significant pressure dependence, all specimens had a similar macroscopically brittle fracture appearance. A qualitative explanation for this behaviour will be given when discussing the results of the normally brittle materials.

As already mentioned, for metal-working the most interesting effect of a superimposed hydrostatic pressure is the increase in ductility of the material to be formed. So it would be useful to have a formula which allows a prediction of the rise in ductility to be expected for a given pressure. The first criterion to be checked for validity will be the one of constancy of the effective true fracture stress which has already been mentioned (Equation (1)). In a tensile test with

necking this stress is equal to the mean axial stress. Using Siebel's formula again the criterion yields

$$-p + \sigma_F^{(p)} \left(1 + \frac{1}{4} \cdot \frac{a^{(p)}}{R^{(p)}} \right) = \sigma_F^{(0)} \left(1 + \frac{1}{4} \cdot \frac{a^{(0)}}{R^{(0)}} \right) \quad (2)$$

where p is the hydrostatic pressure, σ_F is the true flow stress at fracture, a is the outside radius of the smallest cross section, and R is the radius of curvature at neck. The factors with superscripts (p) and (0) are associated with a hydrostatic pressure p or with atmospheric pressure, respectively.

The geometric factor a/R can be replaced by $\sqrt{(\phi_{Br} - \phi_g)^{1.0}}$ with ϕ_{Br} for the true strain at fracture and ϕ_g for the true strain until necking. After σ_F has been replaced by a suitable mathematical description for the true stress-true strain curve, Equation (2) represents a relationship between the true strains at fracture under atmospheric pressure $\phi_{Br}^{(0)}$ and hydrostatic pressure $\phi_{Br}^{(p)}$, respectively, and the superimposed pressure p in terms of the tensile strength σ_B and the true strain until necking ϕ_g only. Thus, with the knowledge of σ_B and ϕ_g the amount of increase in ductility $\Delta\phi_{Br}$ can be calculated for a given pressure.

The full line marked $\sigma_{R \text{ eff}} = \text{constant}$ in figure 6 shows the result when applying the criterion to 42 CrMo 4 steel ('as received'). Compared with the experimental evidence, the calculated improvement of ductility is too large. The mathematical flow curve used here was the power law

$$\sigma = \sigma_B \left(\frac{e}{\phi_g} \cdot \phi \right)^{\phi_g}$$

(where e is the base of natural logarithm) which was replaced by a relationship between stress and strain characterized by linear strain hardening at large strains. According to M. Yajima et al.⁹, for materials showing a constant strain hardening coefficient at large strains, the following relationship between the coefficient $(d\sigma/d\phi)_{\phi \gg \phi_g}$ and the true flow stress at maximum load σ_{HI} exists:

$$\frac{1}{\sigma_{HI}} \left(\frac{d\sigma}{d\phi} \right)_{\phi \gg \phi_g} = \text{constant} \approx \frac{1}{3} \quad (3)$$

Since at maximum load $d\sigma/d\phi = \sigma = \sigma_{HI}$ holds, Equation (3) states that the slope of the tangent drawn at the true stress-true strain curve for the maximum load point is three times larger than the slope of the curve itself at larger strains where linear work-hardening is assumed. These considerations proved valid for all ductile materials tested in the investigations presented here. So for the description of the stress-strain behaviour the power law was replaced by linear strain hardening when the slope of the curve became equal to $\frac{1}{3}\sigma_{HI}$ ($= \frac{1}{3}\sigma_B e^{\phi_g}$).

Considering that pressure also influences the flow stress, the calculation derived from the criterion of a constant effective fracture stress can be corrected assuming an amount of deviation of 1 per cent per kbar. The result is represented by the lowest line in figure 6, indicating an increase which is slightly too

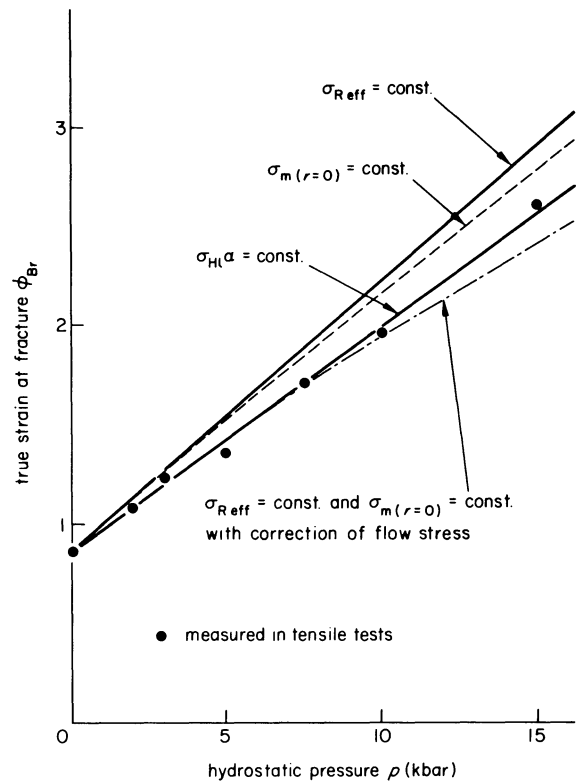


Figure 6. Calculated pressure dependence of ductility of steel (42 CrMo 4) for different criteria.

low at larger pressures compared to experimental evidence. Nevertheless, taking the correction of the flow stress to be slightly lower than 1 per cent per kbar and assuming that the strain hardening coefficient still decreases a little at larger strains instead of remaining constant (both assumptions would tend to raise the line) the criterion discussed describes the pressure dependence of ductility for ductile materials.

A second criterion to be tested claims the constancy of the mean normal stress in the centre of the smallest cross section of a necked specimen $\sigma_m(r=0)$ at the moment of fracture. This criterion yields the following mathematical expression:

$$-p + \sigma_F^{(p)} \left(\frac{1}{3} + \frac{1}{2} \frac{a^{(p)}}{R^{(p)}} \right) = \sigma_F^{(0)} \left(\frac{1}{3} + \frac{1}{2} \frac{a^{(0)}}{R^{(0)}} \right) \quad (4)$$

which is quite similar to Equation (2). Following the same considerations as in the case of the first criterion, the pressure dependence of ductility can be calculated. The result is represented by the broken line in figure 6. Including the flow stress correction the curve resembles that of the first criterion. Since figure 6 is typical for all results obtained, it cannot be decided from the tests on ductile materials which criterion describes the ductility behaviour better.

Since the results discussed so far in figure 6 could be derived from the criteria only by using several simplifying assumptions, and since their final mathematical expressions are rather difficult to deal with, a third criterion based on the empirical evidence alone will be introduced. Comparison of the ductility curves of steel (figure 4(c)) and copper (figure 5(a)) suggests that the increase in ductility per unit pressure may be inversely proportional to the strength of

the material. For all results on the ductile materials, including all heat treatment, the condition

$$\frac{\Delta\phi_{Br}}{p} \cdot \sigma_{HI} = \text{constant} \quad (5)$$

showed the best agreement with the experimental data, where $\Delta\phi_{Br}$ is the increase in ductility, p is the superimposed pressure and σ_{HI} is the true flow stress at maximum load. The value of the constant was determined as a mean value from all experiments performed; it is equal to 1.41. Figure 7 shows the deviation of the various materials and heat treatments from the generalized relationship. The deviation from the straight line remains within reasonable limits.

Because of its simplicity and good agreement with experimental results (see also the solid line $\sigma_{HI}\alpha = \text{constant}$ in figure 6) this third criterion can be used

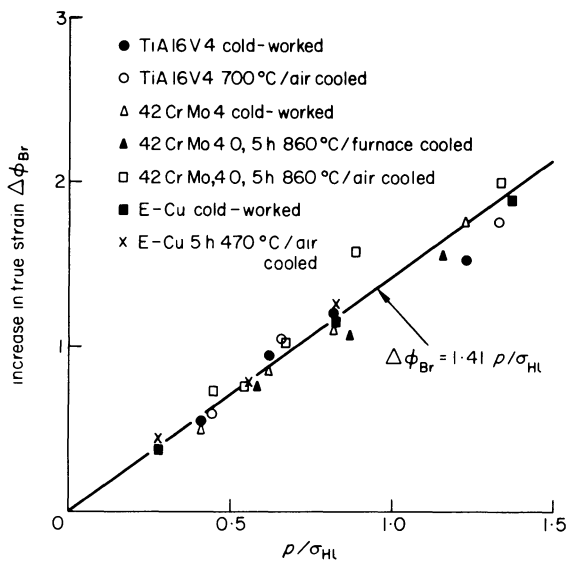


Figure 7. Increase in ductility against ratio of hydrostatic pressure and true stress at maximum load (for ductile materials). $\sigma_{HI} = \sigma_{Be}\phi_g$; for small elongations until necking $\sigma_{HI} \approx \sigma_B$.

to predict the pressure-induced increase of ductility for normally ductile materials. It can be interpreted in physical terms as follows (figure 7): In order to achieve an increase of true strain at fracture of one it is necessary to superimpose a hydrostatic pressure equal to about 70 per cent of the flow stress at maximum load σ_{HI} of the material to be formed. σ_{HI} is related to the tensile strength by $\sigma_{HI} = \sigma_{Be}\phi_g$.

Looking at the last criterion in physical terms it may not seem very reasonable to relate the increase in ductility to a value expressing the strength of a material. However, remembering Equation (3) which states that the true stress at maximum load is a measure of the constant strain hardening at large strains, condition (5) may be written

$$\frac{\Delta\phi_{Br}}{p} \cdot \left(\frac{d\sigma}{d\phi}\right)_{\phi \gg \phi_g} = \text{constant} \approx 0.47 \quad (6)$$

This equation is represented by the straight line in figure 8 when logarithmic scales are used. Figure 8 contains experimental evidence from tests and

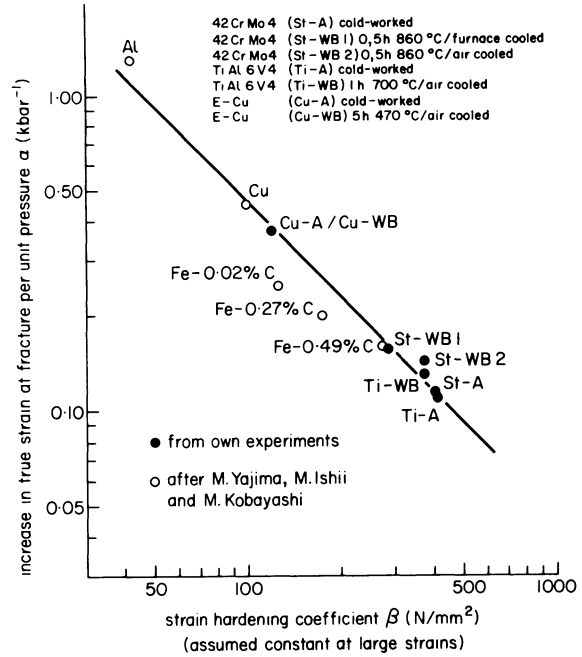


Figure 8. Increase in ductility per unit pressure against strain-hardening coefficient at large strains (for ductile materials).

from an investigation of M. Yajima et al.⁸. It can be generally stated that for materials which are already highly ductile (aluminium, copper) a superimposed pressure enhances ductility much more than for those with lower normal ductility (steel, titanium alloy).

The third criterion expressed in terms of Equation (6) is actually a greatly simplified version of the first criterion of a constant effective fracture stress. In physical terms it can be expressed as follows: For a

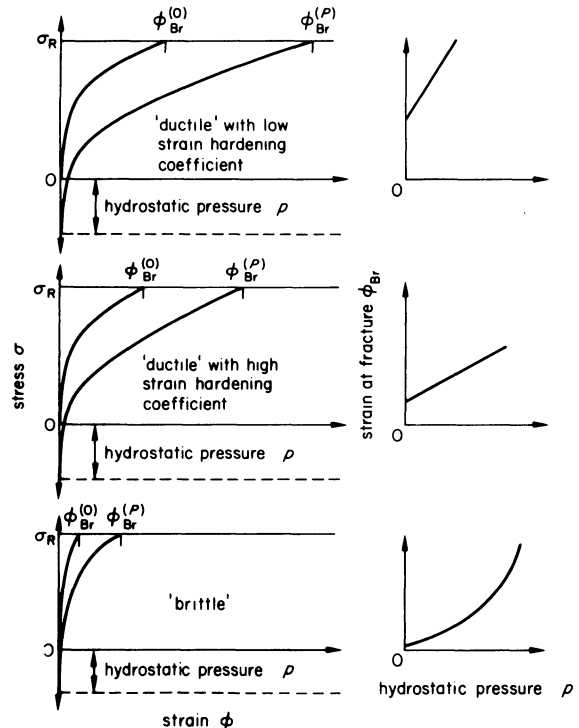


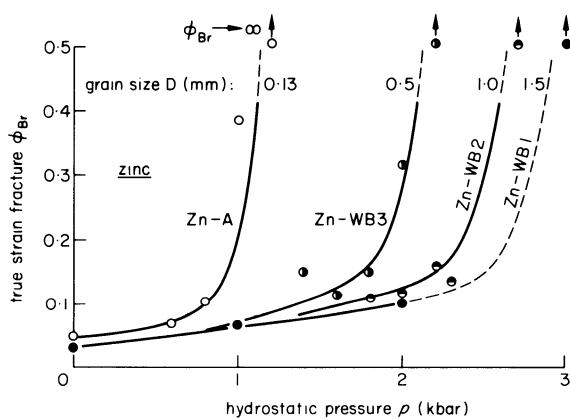
Figure 9. Schematic relationship between pressure and ductility assuming constant true fracture stress.

given pressure the strain at fracture increases by an amount which corresponds to the increase in strain due to strain hardening prescribed by the amount of the pressure.

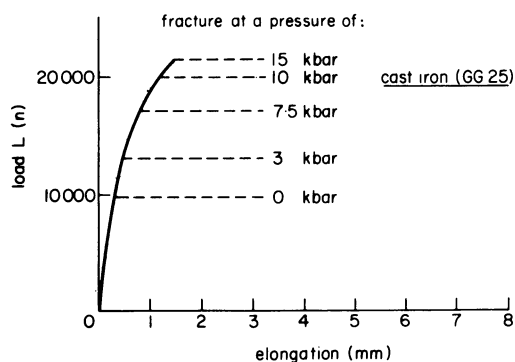
Using two examples this is demonstrated schematically in figure 9(a) and (b). In figure 9(a) a flow curve is assumed with a low strain-hardening coefficient at larger strains which is considered to be nearly constant when the flow curve without pressure meets the horizontal fracture stress line. With superimposed pressure the stress-strain curve is lowered while the fracture stress line stays in the same position. The increase in ductility is shown in the corresponding curve of ϕ_{Br} against p . In figure 9(b), with the same given pressure but a larger strain-hardening coefficient, the slope of the ductility curve (ϕ_{Br} against p) is lower than in the first case where the same strain hardening is achieved by a larger increase in strain.

Normally brittle materials

The brittle materials investigated were zinc, cast iron (GG 25) and Fe₃Si single crystals. Zinc was tested in tensile tests using specimens of various grain sizes. The dimensions of the specimens were the same as for the ductile materials. The resulting ductility curves are shown in Figure 10(a). The relationship between strain at fracture and pressure is no longer linear. The brittle fracture in a tensile test under atmospheric pressure yields a complete necking under only moderate amounts of superimposed hydrostatic



(a)



(b)

Figure 10. Pressure-induced ductility of zinc and cast iron in tensile tests.

pressure. From the point of view of the ductility criterion of a constant effective fracture stress, such a behaviour can at least be qualitatively explained using figure 9(c). When the fracture stress is so low that it 'cuts off' the normal stress-strain curve at its beginning—where the curvature is still high—or even before the yield point is reached, a superimposed pressure will raise ductility only slowly at first, as in figure 9(b). At higher strains when the work-hardening rate has become lower, the increase per unit pressure will be larger, as in figure 9(a). The result will be a relationship between strain at fracture and pressure which is non-linear at the beginning but may be linear later on. Thus the ductility and the flow curves are closely related.

However, the sudden transition from brittle to ductile behaviour cannot be quantitatively explained by the introduced criteria. While the true fracture stress slightly decreases with pressure, the criterion of a constant mean normal stress is entirely invalid. In an ordinary tensile test, fracture stress and fracture strain decrease with increasing grain diameter. In tests under pressure the greatest grain size needs the highest transition pressure to become ductile, as shown in figure 10(a).

The increase of ductility of cast iron obtained from tensile tests (specimen size as for zinc) is extremely small as presented by the load-elongation curves at different pressures in figure 10(b). As in the case of the deliberately embrittled copper, the specimens tested under high pressures fail although all stresses present at fracture are large compressive stresses. However, for a material containing holes or cracks the elastic stress concentration at the tip of such a crack may be three times as large as in homogeneous material. When taking this into account the criterion of the constant effective fracture stress again holds for cast iron up to 10 kbars. For 15 kbars, which is already associated with a certain amount of plastic strain, as well as generally for the embrittled copper, the accordingly expected increase in elongation was not achieved. This may be due to the plastic stress concentration in the vicinity of the cracks. A superimposed hydrostatic pressure does not close the holes completely; this is seen from microstructure photographs. So even under pressure the stress concentration between the cracks may be responsible for early failure.

The increase in ductility of cast iron depends upon the used procedure. Specimens extruded under an environmental pressure of 10 kbars with a reduction in area of 13 per cent showed no microscopic damage, while specimens conventionally extruded with the same amount of deformation broke into pieces. Figure 11(a) shows the ductility curve obtained from compression tests. The specimens used were small cylinders 10 mm long and 6 mm in diameter. Here the influence of the pressure seems to be much higher than in tensile tests, although the character of the relationship between strain and pressure is maintained.

Figure 11(b) shows the ductility behaviour of Fe₃Si single crystals as measured for two different orientations in compression tests. Specimens were 10 mm long and had a rectangular cross section of

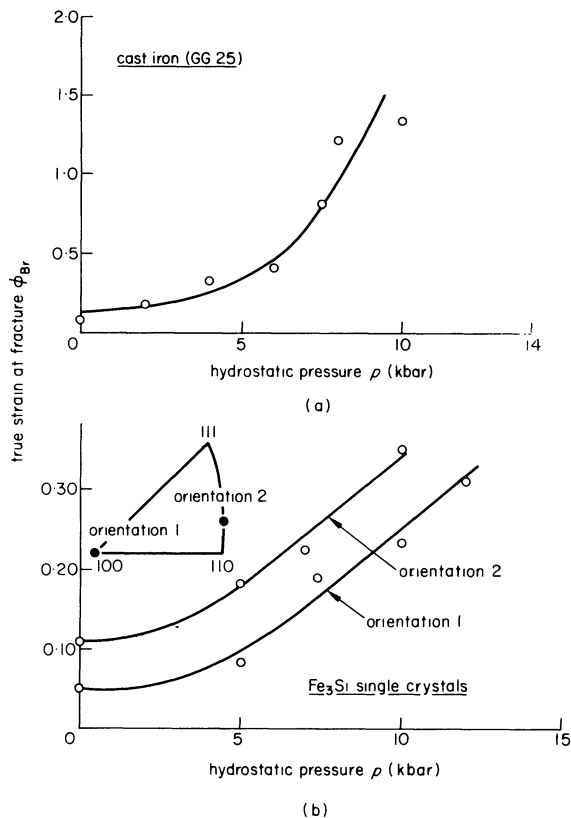


Figure 11. Increase in ductility due to hydrostatic pressure for cast iron (GG 25) and Fe_3Si single crystals in compression tests.

about $4 \times 5 \text{ mm}^2$. The ductilities of orientation I producing multiple glide and orientation II producing double glide are both affected by the superimposed pressure in the same way. The ductility curves again resemble the type received for zinc and cast iron. From this and further evidence from the literature, the ductility behaviour of normally brittle materials seems to be associated with a linear relationship between true strain at fracture and superimposed pressure which is preceded by a non-linear beginning.

CONCLUSIONS

For metal-forming the most interesting effect of an environmental hydrostatic pressure is the increase in ductility of the formed material. The gain of formability can be used to avoid early failure during deformation procedure. However, any rise of the strain at fracture produces additional work-hardening which can, according to the first ductility criterion introduced for ductile materials, be as large as the superimposed pressure. So an environmental pressure may be applied to a forming process in order to obtain special properties of the worked material, for example an extraordinary strength.

The pressure needed to realize a given increase in ductility can be easily estimated for normally ductile materials from the third criterion given in this presentation. For normally brittle materials no generally valid prediction of the pressure dependence on duc-

tility can be reported. However, the following features of the relationship between strain at fracture and superimposed pressure may be stated. The pressure needed for a given rise in ductility depends on the intended metal-working procedure and is considerably larger than for ductile materials of comparable strength. Furthermore, the linear relationship between strain and fracture is preceded by an initial non-linear part.

Other effects of pressure, such as slightly raising the flow stress or changing fracture appearances, may be less interesting to metal-forming than important improvements due to special processes, such as better lubrication and more homogeneity of properties over the cross section in hydrostatic extrusion. Considering the advantages of the presence of a hydrostatic pressure in a forming procedure, further applications in metal-working may be worthwhile.

DISCUSSION

During the presentation of this paper one of the authors, W. Lorrek, emphasized that the uniform elongation until the beginning of necking was found to be entirely unaffected by a superimposed hydrostatic pressure. So the beginning of plastic instability cannot be extended to higher values of strain by an all-around pressure. This is in complete agreement with the results of J. Chakrabarty¹¹ who elongated cylindrical bars under lateral fluid pressure with zero resultant longitudinal stress.

However, Lorrek also pointed out that the beginning of plastic instability can be influenced by using a one-dimensional lateral pressure instead of an all-around pressure. He referred to results from his experiments with rectangular bars which were tensile tested under lateral pressure provided by rigid dies (figure 12). The strain until the beginning of necking was considerably enlarged with increasing die pressure.

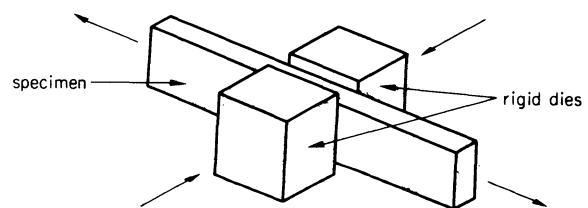


Figure 12.

REFERENCES

1. P. W. Bridgman. *The physics of high pressure*. G. Bell and Sons, Ltd., London, 1941.
2. E. Siebel, and S. Schwaigerer. *Zur Mechanik des Zugversuchs*. *Arch. Eisenhüttenwes.*, 19. S. (1948), 145-52.
3. P. W. Bridgman. *Studies in large plastic flow and fracture*. Harvard University Press, Cambridge, Mass, 1964.
4. W. Lorrek. Thesis, Technische Universität Clausthal, Germany, 1972.
5. P. Haasen, and A. W. Lawson, Jr. (1958) *Der Einfluß hydrostatischen Druckes auf die*

- Zugverformung von Einkristallen. *Z. Metallkde.* Bd. 49, Heft 6, S. (1958), 280-91.
6. F. E. Hanafee, and S. V. Radcliffe. Effect of hydrostatic pressure on dislocation mobility in lithium fluoride. *J. Appl. Phys.* (1967) **38**, no. 11, pp. 4284-94.
 7. Y. D. Martynov, B. J. Beresnev, K. P. Rodionov, Y. N. Ryabinin and M. J. Gavrilynk. The ductility limit of metals tensile-tested at high pressure. *Fiz. metal. metalloved.* **23** (1967) no. 6, pp. 1049-54.
 8. H. Ll. D. Pugh and D. Green. *The behaviour of metals under high hydrostatic pressure II-tensile torsion tests.* MERL plasticity Report Nr. 128, 1956.
 9. M. Yajima, M. Ishii and M. Kobayashi. The effects of hydrostatic pressure on the ductility of metals and alloys. *Int. Journ. of Fracture Mech.* (1970) **6**, 139-50.
 10. R. Hill. *The mathematical theory of plasticity.* Oxford University Press, 1964.
 11. J. Chakrabarty. The necking of cylindrical bars under lateral fluid pressure. *Proceedings of the 13th Int. M.T.D.R.* Macmillan, 1973.

THE TOUGHNESS OF M2 HIGH-SPEED STEEL

by

G. BERRY* and M. J. KADHIM AL-TORNACHI*

SUMMARY

The fracture toughness and the fatigue-crack growth characteristics of M2 high-speed steel have been measured for a wide range of hardening and tempering temperatures. The effect of differing hardening treatments on the occurrence of chipping of the cutting edge during a severe, single-point, surface-milling operation has been compared with these toughness results. A qualitative explanation for the occurrence of this edge chipping has been suggested in terms of the critical flaw size and the primary carbide size and distribution.

INTRODUCTION

The necessarily high hardness of tool materials, required for wear resistance, is concomitant with a relatively low toughness which can be sufficient to limit their cutting life and their application. Thus the restricted use of ceramic tool tips, for example, is a consequence of their low toughness. For any given cutting operation, it is important that the properties of hardness and toughness be balanced in order to achieve the optimum performance. It would be expected that the toughness of the material would indicate any tendency for either gross fracture of the tool tip or any micro-chipping along the cutting edge which can lead to a deterioration of tool performance. Alternatively, toughness is important in indicating resistance to the initiation and propagation of thermal or fatigue cracks. Thus the growth of thermal fatigue cracks in cemented carbides, for example, can be retarded by increasing the proportion of binder phase¹, i.e. by increasing the toughness. It has also been suggested² that inadequate toughness may accelerate the wear process.

Although the importance of toughness is apparent, its definition is less so. Intuitively one would expect it to describe a property of the material which combines a strength parameter³ with the ability to absorb energy without cracking, i.e. to resist crack initiation, and to resist propagation of this crack.

Most of the tool-material toughness tests used previously, measured a strength parameter such as the three- or four-point bend fracture stress or the ultimate torsion strength. These tests also provided an energy term, considered to be indicative of the toughness, which was derived from these stress-strain curves. The various impact tests have directly measured a fracture energy which combines both

energy for crack initiation from a notch or from a plain surface, with that for crack propagation.

The development of the fracture mechanics approach has made possible the measurement of a material's resistance to rapid crack extension. Using a stress function developed by Westergaard⁴, the stress distribution and its intensity, K , near a crack tip can be calculated. A critical value of this stress intensity factor (for plane strain), K_{Ic} , indicates the conditions close to the crack tip when rapid tensile (mode I) crack extension occurs. For a given material, K_{Ic} is unique and can be considered to be a material constant; it is often referred to as the fracture toughness. It is proportional to the critical applied stress, σ_c , and the square root of the crack length, a

$$K_{Ic} = M\sigma_c a^{1/2}$$

where M is a geometric constant accounting for specimen shape and size which can be calculated or determined experimentally. Once the fracture toughness has been measured, it is possible to predict⁵ the critical crack or flaw sizes a_{crit} , for a known working stress and similarly the critical failure stress can be calculated for a known flaw size

$$a_{crit} = \frac{K_{Ic}^2 Q}{1.21 \pi \sigma^2}$$

where Q is a function of the flaw shape and of the ratio of applied stress to 0.2 per cent yield strength.

The stress intensity approach can also be used for the analysis of fatigue-crack growth; crack velocity is found⁶ to be proportional to the n th power of the stress intensity factor range, ΔK

$$\frac{da}{dN} = C \Delta K^n$$

* Department of Mechanical Engineering, Sheffield University.

where n is normally between about 2 and 10 depending on the material.

Further useful information can also be derived from a combination of the fracture toughness and the fatigue-crack growth characteristics by the calculation⁷ of the total number of stress cycles, N , for a small flaw to grow to the critical size, i.e. the expected time to failure.

The few investigations^{8,9} of the fracture toughness of high-speed steel have only been made on the normally heat treated condition. In this work, both the fracture toughness and the fatigue-crack growth characteristics of M2 high-speed steel have been measured over a wide range of hardening and tempering temperatures and the results have been compared, qualitatively, with the results of tool performance tests in milling.

MEASUREMENT OF FRACTURE TOUGHNESS

Standard procedures¹⁰ have now been adopted for the measurement of fracture toughness. In these it is suggested that for plane strain conditions to be maintained at the crack tip, the thickness of the test piece, B , is greater than or equal to a function of the fracture toughness and the 0.2 per cent yield strength σ_{YS}

$$B \geq 2.5 \left[\frac{K_{Ic}}{\sigma_{YS}} \right]^2$$

The size of each bend test piece used in this investigation was 75 mm long by 15 mm wide (W) by 7.5 mm thick (B), the latter dimension fulfilling the above condition for all heat treatments.

Precracks were introduced into the spark-notched test pieces by fatigue, using a load range of 1960 N (in three-point bending) for crack initiation and a subsequent load range of 1420 N for propagation, in each case at a frequency of 52 Hz. Fatigue-crack growth characteristics were obtained using gauges attached to the test piece, having a grid spacing, i.e. a sensitivity, of 0.25 mm. Typically the total length of the precrack was 7.5 mm, which included a notch of length ~ 4 mm. Some fatigue-crack growth curves for several hardening and tempering temperatures are shown in figure 1(a, b) and their analysis in terms of $\log da/dN$ and $\log \Delta K_f$ is shown in figure 2 (a, b).

The fracture toughness was measured in three-point bending with a span to width ratio of 4 : 1 using an Instron tensile/compression testing machine. During each test, a crosshead speed of ~ 0.01 mm/sec was used and the load (P) and displacement were recorded simultaneously up to failure. Subsequently the precrack length was measured at equal intervals across the thickness and K_{Ic} was calculated

$$K_{Ic} = \frac{YP_c}{BW^{1/2}}$$

where Y , the compliance function is defined¹⁰

$$Y = 11.58 \left(\frac{a}{w} \right)^{1/2} - 18.42 \left(\frac{a}{w} \right)^{3/2} + 87.18 \left(\frac{a}{w} \right)^{5/2} - 150.66 \left(\frac{a}{w} \right)^{7/2} + 154.80 \left(\frac{a}{w} \right)^{9/2}$$

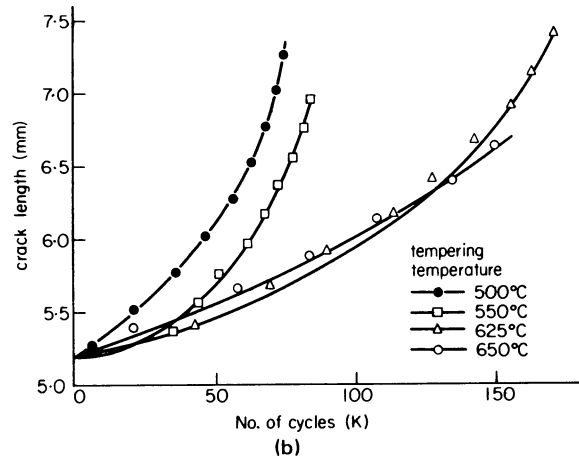
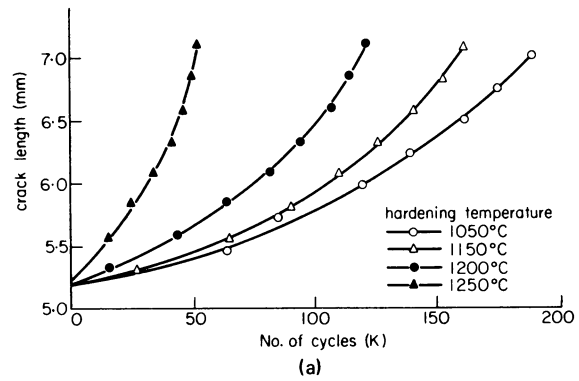


Figure 1. (a and b) Fatigue crack growth curves for several heat treatments; normalized to $N = 0$ for $a = 5.2$ mm.

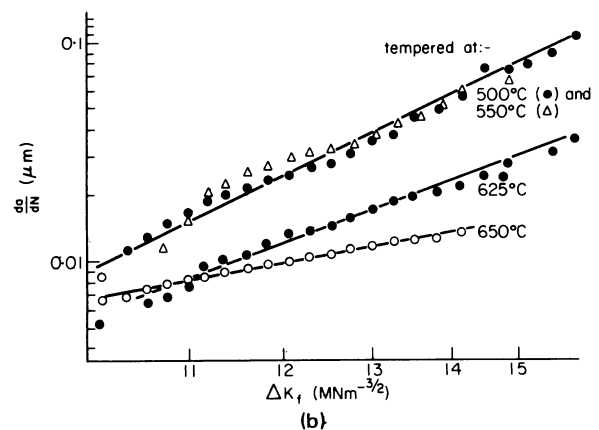
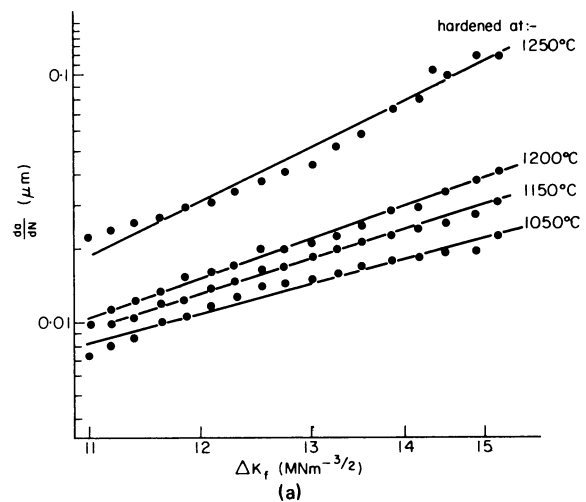


Figure 2. Variation of crack growth rate (da/dN) with alternating stress intensity (ΔK) for several (a) hardening temperatures, and (b) tempering temperatures.

AN ALTERNATIVE PRECRACKING TECHNIQUE

Precracking by fatigue is not always suitable, particularly for the more brittle tool materials which have a high value of the fatigue-crack growth exponent, n . For cemented carbides, for example¹¹, $n \approx 22$; such high values can be attributed¹² to the greater proportions of the tensile modes of failure which occur during fatigue-crack growth. In these circumstances, therefore, control of a fatigue precrack is difficult. An alternative precracking technique first suggested by Harris and Dunegan¹³ has been developed by Eriksson⁹. The stress intensity at the crack tip (introduced from a wedge-loaded notch) is reduced below the critical value by a sufficiently high stress, σ_T , applied through the thickness (figure 3). This technique has been used in this work for the precracking

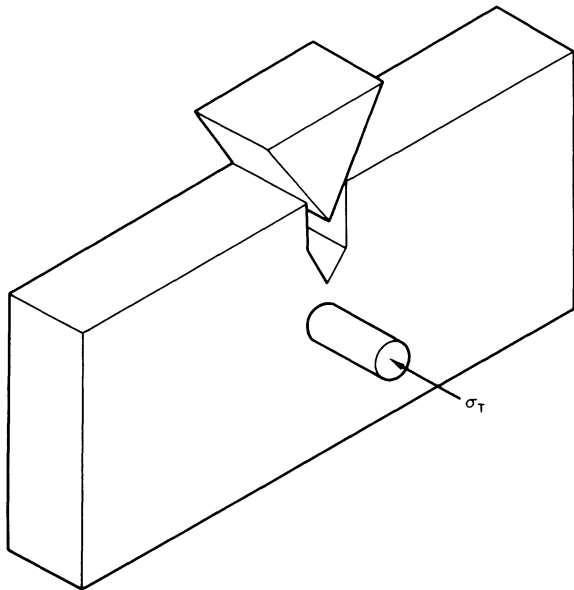


Figure 3. Schematic diagram of alternative precracking system using wedge loading.

of a high-speed steel in a relatively brittle condition. The values of K_{Ic} thus obtained, with different stresses, σ_T , are compared with K_{Ic} values obtained conventionally in Table 1. At low values of σ_T , the two sets of K_{Ic} values lie within the same scatter range; at higher values of σ_T ($> \sim \sigma_{YS}$), however, the K_{Ic} values are higher than the conventional values. Thus this technique is successful when low values of σ_T are used for crack arrest.

HEAT TREATMENT

The widely used M2 high-speed steel of nominal composition 0.85 per cent C, 6.3 per cent W, 5.0 per

cent Mo, 4.0 per cent Cr, 1.9 per cent V, which was chosen for this investigation was kindly supplied by Firth Brown Tools Ltd and by Osborn Steels Ltd. All test pieces were preheated for 6 minutes at 850°C; for the first series pieces were hardened for 5 minutes at temperatures ranging from 1050°C, quenched to 550°C, air cooled and double tempered for 1 hour (each) at 540°C; for the second series pieces were hardened at 1210°C for 5 minutes, quenched to 550°C, air cooled and double tempered for 1 hour at temperatures ranging from 400°C to 700°C. Subsequently all pieces were surface-ground to the final dimensions.

HARDENING TEMPERATURE EFFECTS

Both the fracture toughness and the hardness are plotted in figure 4 with respect to hardening temperature. Generally the toughness decreases as the hardening temperature increases; the slight increase between 1050°C and 1100°C is not significant. The increase in hardness over the same range is consistent

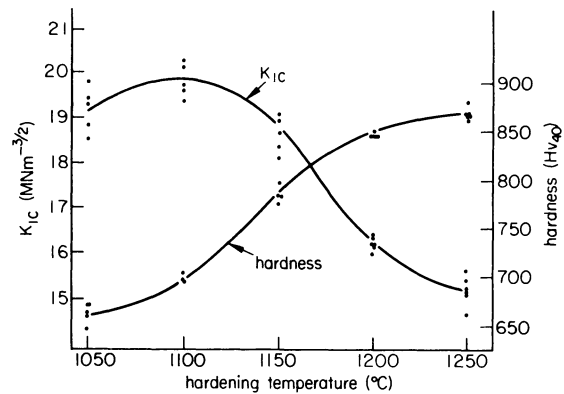


Figure 4. Variation of hardness and fracture toughness (K_{Ic}) with hardening temperature.

with the approximate inverse relationship between K_{Ic} and hardness also found for other tool materials^{9,14}; results from both the first and second series of tests are included in figure 5.

The changes in fracture toughness are accompanied by similar changes in the fatigue-crack growth characteristics, thus, n , the fatigue crack growth exponent almost doubles in value over this hardening temperature range (Table 2). This is in agreement with the general opinion that materials of low toughness also have poor fatigue resistance.

The rapid fracture mode consists essentially of regions of dimpled rupture, cleavage and quasi-cleavage, e.g. figure 6, their proportions depending on

TABLE 1 Comparison of fracture toughness data for alternative precracking techniques (high speed steel)

K_{Ic} (MNm ^{-3/2}) (fatigue precracking)	K_{Ic} (MNm ^{-3/2}) (wedge loading)	Side pressure, σ_T , (MNm ⁻²)
	15.9	1210
16.0	17.1	1380
15.8	18.2	2065
17.2	19.5	2420
16.4	20.2	2760
16.5	19.8	3450

TABLE 2 Hardening temperature series (mean values)

Hardening temp. (°C)	Hardness $H_v(40)$	Estimated σ_{YS} , (MNm ⁻²)	K_{Ic} (MNm ^{-3/2})	Critical flaw size (μm)		Fatigue props	
				$\sigma = \sigma_{YS}$	$\sigma = \sigma_{YS}/2$	$\sim n$	$\sim C$
1050	611	1725	19.2	35	168	3.3	3×10^{-12}
1100	697	1945	19.8	28	140	—	—
1150	784	2280	18.6	19	89	3.8	1×10^{-12}
1200	846	2720	16.2	11	53	4.4	3×10^{-13}
1250	869	2590	15.1	9	46	5.8	2×10^{-14}

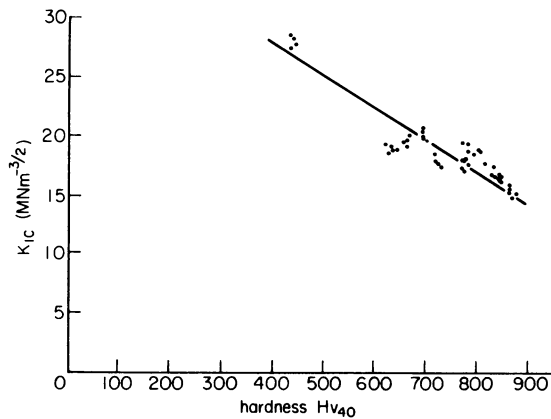


Figure 5. Relationship between hardness and fracture toughness for all heat treatments.

the hardening temperature. The fatigue fracture path is similar for all treatments, figure 7.

The changes that do occur in the toughness and fatigue properties cannot be directly attributed to any one specific cause because, by maintaining a constant holding time at each of the hardening temperatures, the carbide size and distribution, the grain size, and the solute content have all been altered.

Nevertheless it is possible to estimate from this data, the critical flaw sizes for each condition assuming a surface flaw shape having a value of 0.2 for the ratio of semi-minor axis : semi-major axis and

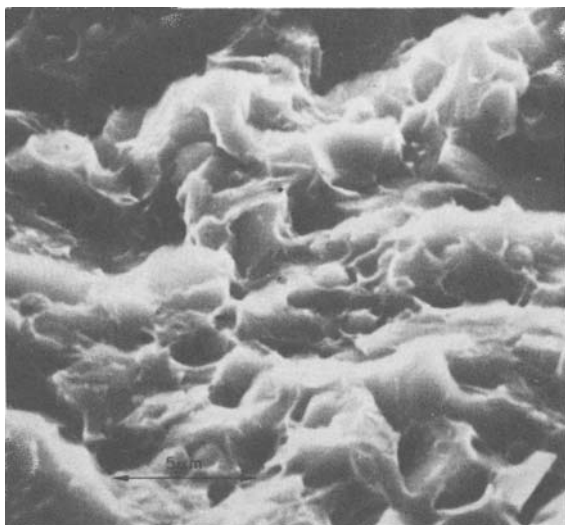


Figure 6. Typical rapid-fracture surface showing cleavage, quasi-cleavage, and dimpled-rupture failure modes (crack propagation from right to left).

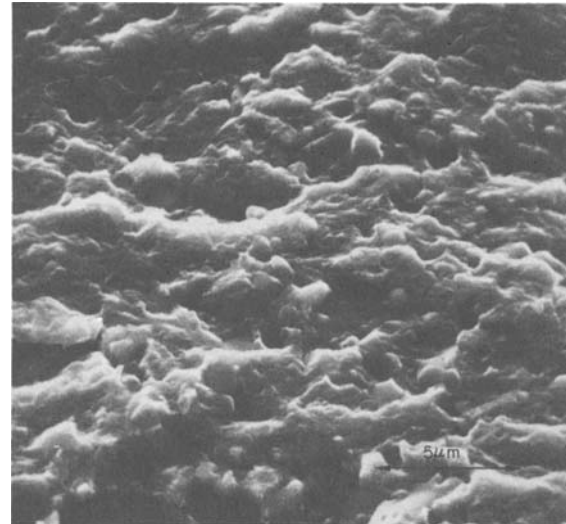


Figure 7. Typical fatigue crack path (crack propagation from right to left).

assuming working stresses of $\sigma_{YS}/2$ or σ_{YS} . The yield-strength values were determined by comparison with data for M2 high-speed steel obtained by Trent¹⁵. As indicated in Table 2, the critical flaw sizes decrease from 168 μm (at a hardening temperature of 1050°C) to 46 μm (at a hardening temperature of 1250°C) for a working stress of $\sigma_{YS}/2$ and from 35 μm to 9 μm over the same range for a working stress of σ_{YS} . Typically, carbide sizes range from about 1 μm to 8 μm , thus if carbides are considered to be flaws or flaw generators, then failure is likely to occur immediately if a group of the larger carbides are in close proximity or after some time if several smaller carbides are linked together by cracking. This approach confirms the requirements of specifications¹⁶ for high-speed steel which state that there should be no more than a certain number of large carbides of a given size in a standard area.

It would be expected that the in-service 'toughness' for this first series would depend on the critical flaw size, the carbide size and distribution (which varies in this series), and the distribution of surface scratches. It is interesting to note that the critical flaw size is a function of a strength parameter (σ_{YS}) and of the resistance to crack propagation (K_{Ic}); it is therefore approaching the intuitive definition of toughness.

TEMPERING TEMPERATURE EFFECTS

In contrast to the first series, the grain size, the primary carbide size and distribution remain similar in this second series of tests since each of the

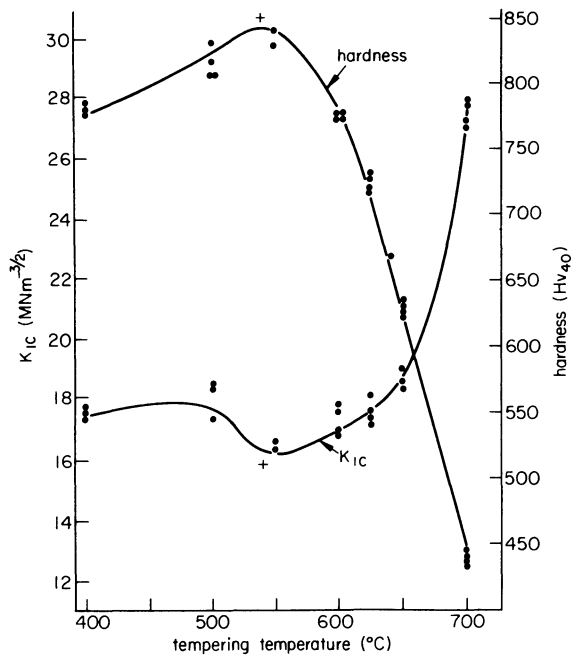


Figure 8. Variation of hardness and fracture toughness (K_{Ic}) with tempering temperature (+ indicates relevant value taken from figure 4).

tempering treatments is preceded by the same hardening treatment. The variation of toughness with tempering temperature (figure 8) shows little variation between 400°C and 500°C but at ~550°C there is a minimum which coincides with the secondary hardening peak. (The relevant results from the hardening series are also included in figure 8 and are indicated by +.) This minimum, which is often attributed¹⁷ to the phenomenon of tempered martensite embrittlement in other steels, occurs at tempering temperatures in the region of those used for the standard heat treatment. Torsion tests and C-notch Charpy¹⁸ tests are also sensitive to this minimum, but unnotched Izod tests have remained insensitive¹⁹. The critical flaw size (Table 3) also shows a minimum whose value is consistent with the values obtained from the first series for a similar tempering temperature (cf. Table 2).

At high tempering temperatures, the toughness increases slowly at first and then rapidly. There would appear to be little difference in properties close to either side of the minimum since the critical flaw size, the toughness, and the hardness are similar. The fatigue properties do not show a minimum but are similar between tempering temperatures of 500°C and 550°C (Table 3) and improve with increasing tempering temperature.

MILLING TEST RESULTS

Preliminary single-point milling tests were conducted on a Cincinnati horizontal milling machine using test pieces taken from the hardening series. Relatively severe cutting conditions were employed and tool lives were measured in terms of the tendency for chipping of the cutting edge.

The broken halves of the fracture toughness test pieces (~38 mm × 15 mm × 7.5 mm) were ground to the geometry described below and each piece was then mechanically clamped into the tool-holder for the single-point, centre-milling tests. The tool geometry was as follows: effective side rake 25°, back rake -15°, plan approach angle 15°, end and side clearances 10° and trail angle 10°. Each tip was given a nose radius of 1 mm using a specially constructed jig and a 0.1-mm land was diamond-honed on the cutting edge.

The dry cutting conditions were constant for all tests and were similar to those normally used for rough-face milling²⁰ with a cutting speed of 27.2 m/min, a feed of 1.35 mm/rev, a depth of cut of 3.75 mm, a width of cut of 46 mm and a milling cutter head of diameter 150 mm. The workpiece material had a Brinell hardness of 163 and a nominal composition of 0.2 per cent C, 0.7 per cent Mn, 0.04 per cent S, 0.5 per cent Ni, 0.5 per cent Cr, 0.5 per cent Mo. These fairly severe cutting conditions were chosen in order to minimize heat generation and to produce some sensitivity to the changes in toughness of the tool material. Each test was stopped at suitable intervals and the tip was examined for both wear and edge chipping, or the test was terminated immediately if a distinct failure occurred.

For tips hardened at 1250°C (tip A) and 1200°C (tip B), chipping of the cutting edge occurred abruptly and distinctly after 1000 and 700 cuts respectively. This point was recognized by the appearance of a dark line on the transient cutting surface and by the occurrence of sparking. Although the form of chipping (i.e. segmental) was similar in both tips (e.g. figure 9(a) the extent was different. The depth of the chipping in tip A on the rake and clearance face was 0.45 mm and 0.7 mm respectively and the extent of chipping along the cutting edge was 2.03 mm. The corresponding chip dimensions for tip B were 0.30 mm, 0.63 and 1.70 mm. Cracks were also observed on the rake face of tip B (figure 9(b)) which suggests that stable crack growth can occur in high-speed steel tips. More evidence of this effect was gained from the tip hardened at 1150°C (tip C) where

TABLE 3 Tempering temperature series (mean values)

Tempering temp. (°C)	Hardness $H_V(40)$	Estimated σ_{YS} (MNm^{-2})	K_{Ic} ($MNm^{-3/2}$)	Critical flaw size (μm)		Fatigue props	
				$\sigma = \sigma_{YS}$	$\sigma = \sigma_{YS}/2$	$\sim n$	$\sim C$
400	782	2520	17.6	14	66	—	—
500	820	2640	18.0	13	61	5.6	2×10^{-14}
550	835	2690	16.6	10	51	5.6	2×10^{-14}
600	776	2500	17.4	14	66	—	—
625	724	2330	17.6	16	76	4.2	3×10^{-13}
650	628	2020	18.7	24	114	2.1	5×10^{-11}
700	440	1420	27.6	107	514	—	—

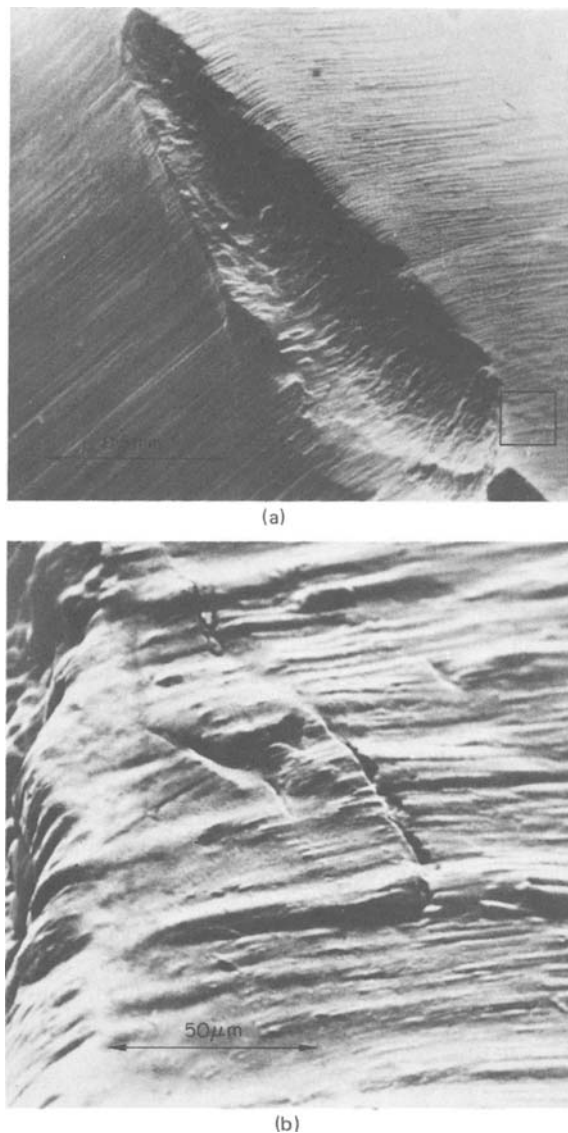


Figure 9. Tip hardened at 1200°C. (a) Example of segmental chipping of cutting edge (insert indicates position of figure 9(b)), and (b) cracking along rake face.

the characteristic dark line was observed on the transient cutting surface after ~ 900 cuts. When the tip was examined, chipping had (in contrast to tips A and B) occurred along a strip (~ 0.2 mm wide) on the cutting edge; a crack was also observed parallel to the cutting edge on the rake face. The test was continued for a few more cuts and further edge chipping occurred along the line of the previously observed crack.

The cutting edge of the tip hardened at 1050°C (tip E), deteriorated by uniform wear with no indication of chipping. The deterioration of the cutting edge of the tip hardened at 1100°C (tip D) appeared to occur by a process combining both wear and micro-chipping.

TOUGHNESS AND TOOL PERFORMANCE

From these preliminary milling tests, it is apparent that the room-temperature toughness properties are qualitatively consistent with the milling performance.

Thus the severity of chipping increases as the toughness decreases. It is surprising, however, that tip A (hardened at 1250°C) has a greater resistance to chipping (in terms of the number of cuts taken before chipping occurs) than tip B (hardened at 1200°C). A possible qualitative explanation of this effect can be found from their respective carbide sizes and distributions.

For this explanation it will be assumed²¹ that hardening at 1250°C produces a smaller volume of primary carbides than does hardening at 1200°C or other lower temperatures. Since the room-temperature critical flaw sizes for the two conditions are similar (Table 2), the probability of the occurrence of a group of carbides of critical or near critical size would be greater for tip B. Assuming an extrapolation of the room-temperature toughness properties, the probability of failure either directly from this group of carbides or by sub-critical crack growth from this group, would, therefore, be greater for tip B. For tip C (hardened at 1150°C), the critical flaw size and the fatigue properties have increased sufficiently to balance the effects of the greater volume of primary carbide present in the microstructure and therefore to make the probability of failure or chipping less than for tip B. For tips D and E, hardened at the lower temperatures, toughness properties are of less importance since wear becomes the dominant mechanism of cutting-edge deterioration.

The cracking observed during milling with tips A, B and C was parallel to the cutting edge, which suggest that it was of a mechanical rather than of a thermal origin (cf. thermal cracking of cemented carbides which occurs perpendicular to the cutting edge). Thus for these tips the resistance to stable crack growth under alternating loads, i.e. the fatigue properties, could also provide useful information.

Of course the severity of the cutting operation also defines the toughness required for a cutting material. In this work these particular cutting conditions were chosen because earlier milling tests at a higher cutting speed and at a smaller depth of cut and feed had produced severe wear rather than chipping failures. Thus, under cutting conditions which produced chipping of the cutting edge, a qualitative explanation of its occurrence has been obtained in terms of the fracture toughness.

CONCLUSIONS

From this investigation of M2 high-speed steel the following conclusions can be made.

- (1) For a standard tempering treatment, resistance to both stable and unstable crack propagation (i.e. the fatigue properties and the fracture toughness respectively) decrease with increasing hardening temperature.
- (2) For a standard hardening temperature a minimum in the fracture toughness occurs after tempering at $\sim 550^\circ\text{C}$.
- (3) The variation in room-temperature toughness for different hardening treatments is consistent with the occurrence of chipping during a severe milling operation.

- (4) For the tool tips hardened at 1150°C, 1200°C, 1250°C, there is some evidence of stable crack growth parallel to the cutting edge during this milling operation.

ACKNOWLEDGMENT

One of the authors (G.B.) acknowledges the financial support for this investigation given by the Science Research Council. The close cooperation of Professor D. S. Dugdale is also gratefully acknowledged.

REFERENCES

1. C. S. G. Ekemar, S. A. O. Iggstrom and G. K. A. Heden. Influence of some metallurgical parameters of cemented carbides on the sensitivity to thermal fatigue cracking at cutting edges, Iron and Steel Institute, Publ. 126, p. 133 (1970).
2. Stora Special Steels. *Stora A.S.P. high-speed steels*, Product information September 1972, p. 5, 1972.
3. E. Lardner. Review of current hardmetal technology, Iron and Steel Institute, Publ. 126, p. 122, 1970.
4. H. M. Westergaard. Bearing pressures and cracks, *J. Appl. Mech.* (1939) 6, A-49.
5. G. R. Irwin. Crack extension force for a part through crack in a plate, *J. Appl. Mech.* (1962) 84, E-651.
6. P. C. Paris and F. Erdogan. A critical analysis of crack propagation laws, *J. Basic Engineering, Trans. A.S.M.E.* (1963) 85, D-528.
7. E. T. Wessel and T. R. Mager. Fracture mechanics technology as applied to thick walled nuclear pressure vessels, from practical application of fracture mechanics to pressure vessel technology, *Proc. Inst. Mech. Engrs.*, London, p. 17, 1971.
8. M. A. Malling. Fracture toughness of high speed steel, Final year dissertation, Department of Mechanical Engineering, Sheffield University, 1971.
9. K. Eriksson. Fracture toughness of hard high speed steels, tool steels and white cast irons, *Scandinavian J. Metallurgy* (1973) 2, 197.
10. E. F. Walker. Test techniques, from fracture toughness, Iron and Steel Institute, Publ. 121, p. 49, 1968.
11. T. J. Davies and S. Barmana. Fatigue of a WC-25 per cent Co. alloy, *Planseeberichte fur Pulvermetallurgie* (1972) 20, 30.
12. R. O. Ritchie and J. F. Knott. Mechanisms of fatigue crack growth in low alloy steel, *Acta Met.* (1973) 21, 639.
13. D. O. Harris and M. L. Dunegan. Fracture toughness of beryllium, *J. Materials* (1968) 3, 59.
14. N. Ingelstrom and H. Nordberg. The fracture toughness of cemented tungsten carbides, Swedish Inst. for Metal Research, rep. IM-948, 1973.
15. E. M. Trent. Private Communication, 1973.
16. W. S. Caisley. Private Communication, 1973.
17. C. J. McMahon. Temper brittleness—an interpretive review, from Temper Embrittlement in Steel, *A.S.T.M. S.T.P. 407, American Soc. for Testing and Materials*, p. 127, 1968.
18. T. B. Smith and N. H. McBroom. High speed steel for cold forming, *Iron and Steel* (1968) 41, 56.
19. F. A. Kirk, M. C. Child, E. M. Lowe and T. J. Wilkins. High speed steel technology—the manufacturers' viewpoint, Iron and Steel Institute, Publ. 126, p. 67, 1970.
20. P. S. Houghton. *The Milling Machine*, Crosby and Lockwood, London, 1948.
21. G. Hoyle. Recent developments in high-speed steels, *Metallurgical Reviews* (1964) 9, 49.

ANALYSIS OF LARGE, NATURAL STRAIN INCREMENTS USING THE MOHR'S CIRCLE

by

J. WOLAK*

SUMMARY

An approximate relation is developed between the magnitude of direct strain induced in line elements and their mean angle of relative orientation during straining. This relation, shown to be representable by means of a Mohr's circle, leads to the evaluation of the principal strain components from the measurement of direct strain in three coplanar directions. Errors associated with such use are evaluated and plotted for various input parameters. The errors increase with the magnitude of the effective strain; for normal working conditions, effective strain increments of the order of 0.4 may be evaluated with an accuracy of about 1.3 per cent, or better. The proposed technique is simple, easily adaptable to computer work, and in usual practical applications yields results that differ insignificantly from those obtained by an 'exact' analysis.

	NOTATION		
\ln	Natural logarithm.	$\epsilon_{1c}, \epsilon_{2c}$	Principal strain increments calculated by means of equations (10) from data generated by equations (11) and (12) employing ϵ_{1k} and ϵ_{2k}
m	The number of small subdivisions (or steps) into which a large strain increment may be subdivided.	$\theta_A, \theta_B, \theta_C$	Orientation of elements A, B, C as in figure 1.
β	Rotation of a linear element, such as OE in figures 2 and 3, caused by the application of principal strain increments (ϵ_1, ϵ_2).	θ_{AB} , etc.	Angle between elements OA and OB , etc.
$(\epsilon_1, \epsilon_2, \epsilon_3)$	Simultaneous increments of principal strains associated with a large strain increment.	$\theta_{A_0}, \theta_{A_m}, \theta_{A_a}$, etc.	Initial, final, and average angles of orientation of a given lineal element OA subjected to a large strain increment ($\epsilon_1, \epsilon_2, \epsilon_3$); see figures 3 and 4.
$(\epsilon_x, \epsilon_y, \epsilon_z, \epsilon_{xy})$	Simultaneous increments of lineal and shear strain components, relative to an (x, y, z) co-ordinate system, direction z coinciding with one principal strain.		<i>Note:</i> $\theta_a = (\theta_0 + \theta_m)/2$ for a given line element.
$\left(\frac{\epsilon_1}{m}, \frac{\epsilon_2}{m}, \frac{\epsilon_3}{m}\right)$	A small sub-increment of a large strain increment ($\epsilon_1, \epsilon_2, \epsilon_3$), m being a large whole number.		
ϵ_E , etc.	Longitudinal (direct) strain induced in element OE , etc.		
$\epsilon_{1k}, \epsilon_{2k}$	Given values of principal strain increments.		

INTRODUCTION

The representation of the state of strain by means of the Mohr's circle¹ has been known for many years. Although attempts were made to use this construction for the description of the general case of the state of strain at a point in three dimensions^{1,2}, its most useful applications are to cases in which the direction (though not necessarily the magnitude) of one of the three principal strains is known to be perpendicular to a plane, which is parallel to the

* Department of Mechanical Engineering, University of Washington, Seattle, Washington U.S.A.

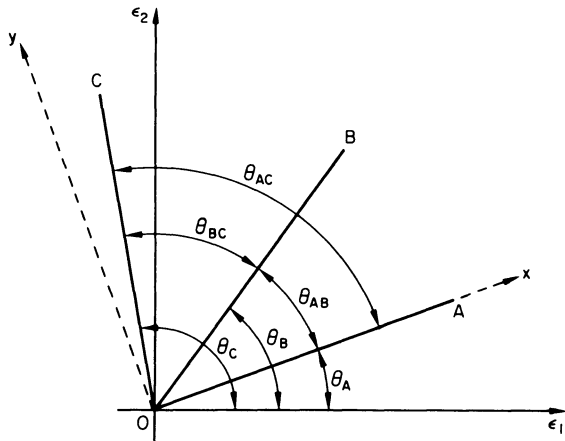


Figure 1. Three arbitrarily chosen orientations OA , OB , OC , along which lineal (direct) strain increments are usually measured for strain evaluation purposes.

remaining two principal strains both unknown in direction and magnitude. The solution of this problem for small strain increments is very well known¹⁻⁴. A choice is made of some convenient system of axes, xy , with respect to which the strain components may be denoted by ϵ_x , ϵ_y and ϵ_{xy} . In terms of these strain components the longitudinal strain ϵ_B of an element OB inclined at an angle θ_{AB} to the xx axis, as shown in figure 1, may be written

$$\epsilon_B = \frac{\epsilon_x + \epsilon_y}{2} + \frac{\epsilon_x - \epsilon_y}{2} \cos 2\theta_{AB} - \epsilon_{xy} \sin 2\theta_{AB} \quad (1)$$

If ϵ_B and θ_{AB} are measured Equation (1) would contain three unknowns, namely: ϵ_x , ϵ_y , and ϵ_{xy} . To solve for these, two more equations such as Equation (1) may be obtained by measuring the angle of inclination of, and the longitudinal strain induced in, any two other lines such as OA and OC in figure 1. This system of three simultaneous equations may be solved for ϵ_x , ϵ_y and ϵ_{xy} and these quantities may then be used to determine the corresponding principal strain increments, ϵ_1 and ϵ_2 , either by the Mohr's circle construction or by the use of the expression

$$\epsilon_1, \epsilon_2 = \frac{\epsilon_x + \epsilon_y}{2} \pm \left[\left(\frac{\epsilon_x - \epsilon_y}{2} \right)^2 + \epsilon_{xy}^2 \right]^{1/2} \quad (2)$$

Alternatively the three measured strains ϵ_A , ϵ_B , ϵ_C and the angles

$$\theta_{AB} = \theta_B - \theta_A \quad \text{and} \quad \theta_{BC} = \theta_C - \theta_B \quad (3)$$

may be used directly for the construction of the Mohr's circle, as proposed by Murphy³. Perry and Lissner⁴ give a detailed discussion of both these methods.

It is important to note that when the principal strain increments are small, the associated shear strain increments (ϵ_{xy}) must also be small and, consequently, the rotation due to the strain of any element, such as OA , in figure 1, will be small and its incorporation in equation (1) will have negligible effect on the result. When the deformation in any one step is relatively large, the resulting rotation of the

element being strained may be significant. For example if an element of material $OA_0B_0C_0$ is subjected to plane strain having principal components ϵ_1 † and ϵ_2 , it will be deformed into the shape $OA_mB_mC_m$ as shown in figure 2. A line OE_0 before deformation will occupy position OE_m after deformation. In this case it is evident that the Mohr's circle, based on Equation (1), is not suitable for the determination of strain along OE and somewhat more

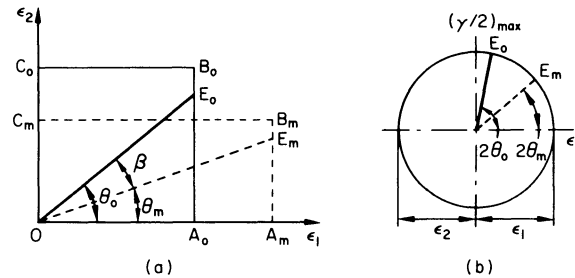


Figure 2. (a) Square element $OA_0B_0C_0$ deformed by pure shear ($\epsilon_1 = -\epsilon_2$) into element $OA_mB_mC_m$. (b) Inappropriateness of Mohr's circle, based on Equation (1), to describe large strain increments.

complex equations have to be used^{2,5} (see Appendix). However, if in a given strain field the initial and final angles of orientation of line elements are known relative to an arbitrarily chosen axis, a slightly modified form of equation (1) may be used for an approximate but, in most cases, sufficiently accurate analysis of large strain increments in the 'logarithmic description'. As shown below, this leads to a Mohr's circle that relates the longitudinal (direct) and shear strains induced in line elements with the mean (and thus with the initial and final) angles of orientation of these elements during straining.

THE CHOICE OF THE MEASURE OF STRAIN

When evaluating large, plastic strain increments it is often convenient to subdivide the straining process into a number of steps and to calculate the strain increments for each of these steps. For example, in metal cutting an element of material may be subjected to several unequal strain increments in its passage from the uncut specimen into the chip; these strain increments can be independently evaluated and then summed to give a measure of the total effective strain experienced by the element. The results of such evaluations should be unique and independent of both the number and the size of the 'sub-steps' into which the total process of straining is divided. As explained in greater detail elsewhere⁶, these requirements are fulfilled by the definition of the logarithmic strain. It should be noted that once the strain is evaluated in any one measure, it may easily be converted to other strain measures⁷.

† In all of this work the elastic strain components are considered very small by comparison with the plastic strain components and are, therefore, neglected.

ADAPTATION OF EQUATION (1) TO THE ANALYSIS OF LARGE STRAIN INCREMENTS

The element $OA_0B_0C_0$ in figure 2 is shown subjected to principal strain increments (ϵ_1, ϵ_2) acting parallel to the fixed directions OA_0^* and OC_0 . These principal strain increments, ϵ_1 and ϵ_2 , can be attained in m 'equal' steps, m being sufficiently large so that (ϵ_1/m) and (ϵ_2/m) are relatively small and, therefore, the resulting rotation of OE in any one step-increment is also small. For the first pair of strain increments $(\epsilon_1/m, \epsilon_2/m)$ a Mohr's circle can be drawn as shown in figure 3 and the corresponding strain increment

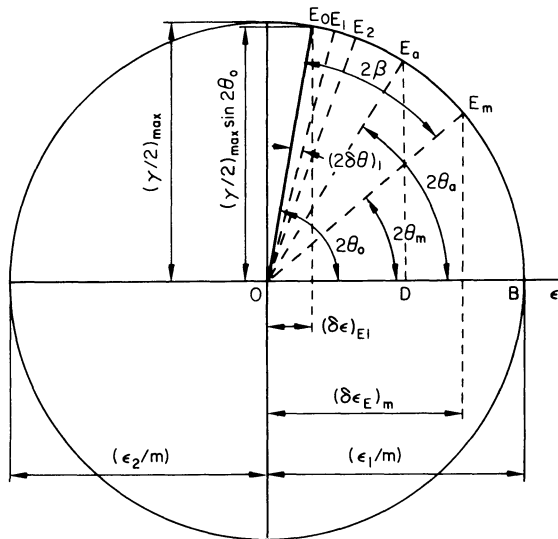


Figure 3. Superposition of m Mohr's circles for equal strain increments $(\epsilon_1/m, \epsilon_2/m)$.

$(\delta\epsilon)_{E_1}$ along OE , as well as the rotation $(\delta\theta)_1$ of OE in this first step can be evaluated⁸. The initial orientation of OE for the second stepwise strain increment is the algebraic sum of $\{\theta_0 + (\delta\theta)_1\}$, which could be employed in the Mohr's circle of figure 3 to obtain strain increment $(\delta\epsilon)_{E_2}$ and the corresponding 'increment' in rotation, $(\delta\theta)_2$.

Repeating this process for each of the m small strain increments $(\epsilon_1/m, \epsilon_2/m)$, all the components of the longitudinal strain in OE as it swings during deformation from position OE_0 to OE_m could be obtained from figure 3, and written in equation form as follows

$$\begin{aligned}
 (\delta\epsilon)_{E_1} &= \frac{\epsilon_1 + \epsilon_2}{2m} + \frac{\epsilon_1 - \epsilon_2}{2m} \cos 2\theta_0 \\
 (\delta\epsilon)_{E_2} &= \frac{\epsilon_1 + \epsilon_2}{2m} + \frac{\epsilon_1 - \epsilon_2}{2m} \cos 2(\theta_0 + \delta\theta_1) \\
 (\delta\epsilon)_{E_m} &= \frac{\epsilon_1 + \epsilon_2}{2m} \\
 &\quad + \frac{\epsilon_1 - \epsilon_2}{2m} \cos 2(\theta_0 + \delta\theta_1 + \dots + \delta\theta_{m-1})
 \end{aligned}$$

* The symbol OA denotes the length of the line OA in any position during straining and the subscript n signifies the n th sub-step of straining. The strain $\epsilon_A = \ln(OA_m/OA_0)$, where OA_0 and OA_m denote the initial and final lengths of OA .

The sum of these m components would give the total longitudinal strain experienced by OE , that is

$$\begin{aligned}
 \epsilon_E &= \frac{\epsilon_1 + \epsilon_2}{2} + \frac{\epsilon_1 - \epsilon_2}{2m} [\cos 2\theta_0 + \cos 2(\theta_0 + \delta\theta_1) \\
 &\quad + \dots + \cos 2(\theta_0 + \delta\theta_1 + \dots + \delta\theta_{m-1})] \\
 &= \frac{\epsilon_1 + \epsilon_2}{2} + \frac{\epsilon_1 - \epsilon_2}{2} \cos 2\theta_{ac}
 \end{aligned}$$

or

$$\left(\epsilon_E - \frac{\epsilon_1 + \epsilon_2}{2} \right) = \frac{\epsilon_1 - \epsilon_2}{2} \cos 2\theta_{ac} \quad (4)$$

where $\cos 2\theta_{ac}$ denotes the average value of the cosine as the element swings from θ_0 to θ_m (see Reference 6, wherein θ_{ac} is denoted by θ_a). Similarly, the rotation of the element OE relative to the principal direction ϵ_1 could be summed up as follows

$$\begin{aligned}
 \beta &= \delta\theta_1 + \delta\theta_2 + \dots + \delta\theta_{m-1} \\
 &= \left(\frac{\gamma_{max}}{2m} \right) [\sin 2\theta_0 + \sin 2(\theta_0 + \delta\theta_1) + \dots \\
 &\quad + \sin 2(\theta_0 + \delta\theta_1 + \delta\theta_2 + \dots + \delta\theta_{m-1})] \\
 &= \frac{\gamma_{max}}{2} [\sin 2\theta_{as}] = \frac{\epsilon_1 - \epsilon_2}{2} \sin 2\theta_{as} \quad (5)
 \end{aligned}$$

where $\sin 2\theta_{as}$ denotes the average value of the sine† as the element swings from θ_0 to θ_m and γ_{max} is the maximum shear strain associated with the principal strain components ϵ_1 and ϵ_2 . The angles θ_{ac} and θ_{as} are, in general, not the same; however, both of them are approximately equal to the average angle of orientation, θ_a , of the element OE during straining, that is

$$\theta_{ac} \simeq \theta_{as} \simeq \theta_a = (\theta_0 + \theta_m)/2 \quad (6)$$

Considering β as a measure of shear and replacing θ_{ac} and θ_{as} in Equations (4) and (5) respectively by θ_a and adding the squares of the corresponding sides of these equations, results in

$$[\epsilon_E - (\epsilon_1 + \epsilon_2)/2]^2 + \beta^2 = [(\epsilon_1 - \epsilon_2)/2]^2 \quad (7)$$

which is an equation of a Mohr's circle with ϵ_E and β as variables, with centre at $[(\epsilon_1 + \epsilon_2)/2, 0]$ and with a radius $[(\epsilon_1 - \epsilon_2)/2]$.

In other words figure 3 represents a series of m Mohr's circles one on top of another, and differing one from another only in the initial orientation of the same line element being strained through a given sub-step. If the scale of strain in figure 3 be changed so that OB represents ϵ_1 instead of (ϵ_1/m) , then figure 3 would be graphical representation of Equation (7) and the strain along OE would be represented on this circle by a point E_a such that the angle BOE_a would be approximately $2\theta_a$ as suggested by Equa-

† The summation for $2\theta_{as}$ can be done in the same manner as that for $2\theta_a$ in Reference 6.

tion (6), while the corresponding 'shear' would be represented by the ordinate DE_a .

Having established an approximate relationship suitable for a representation by the Mohr's circle, it is possible to use the properties of the latter to express the longitudinal strain in any direction in terms of longitudinal strains ϵ_x, ϵ_y , and shear strain ϵ_{xy} [$=\beta = (\epsilon_1 - \epsilon_2/2) \sin 2\theta_a$] associated with some 'reference set of axes, ox, oy ', which are at right angles to each other at some intermediate instant during the process of straining under consideration.* Expressed mathematically, this becomes

$$\epsilon_B = \frac{\epsilon_x + \epsilon_y}{2} + \frac{\epsilon_x - \epsilon_y}{2} \cos 2\theta_{AB_a} - \epsilon_{xy} \sin 2\theta_{AB_a} \quad (8)$$

where θ_{AB_a} is the average angle of orientation of element OB relative to the reference axis x (element OA in figures 1 and 4) during the process of straining. Equation (8) may be used for an admittedly approximate analysis of relatively large strains in the same manner as Equation (1) is used for small strains, but it should be emphasized that all input data to Equation (8) as well as the results are in terms of the mean angles of orientation of the elements during straining. In view of this it is desirable to express the mean angles between the elements during straining in terms of their initial and final angles of orientation since the latter are the quantities that will serve as data and can easily be measured during an experiment.

Referring to figure 4 consider elements OA, OB, OC , with initial orientation $\theta_{A_0}, \theta_{B_0}$ and θ_{C_0} ; due to straining these elements attain the final orientations $\theta_{A_m}, \theta_{B_m}$ and θ_{C_m} , respectively, relative to some fixed (i.e. non-rotating) fibre such as the one subjected to the principal strain, ϵ_{1k} . Let $\theta_{AB_0}, \theta_{AB_a}$ and θ_{AB_m} denote the initial, the average and the final angle, respectively, between OA and OB . Then, by Equation (6)

$$\begin{aligned} \theta_{AB_a} &= \frac{1}{2} \{ \theta_{AB_0} + \theta_{AB_m} \} \\ &= \frac{1}{2} \{ (\theta_{B_0} - \theta_{A_0}) + (\theta_{B_m} - \theta_{A_m}) \} \\ &= \frac{1}{2} \{ (\theta_{B_0} + \theta_{B_m}) - (\theta_{A_0} + \theta_{A_m}) \} = (\theta_{B_a} - \theta_{A_a}) \end{aligned} \quad (9)$$

Similarly

$$\theta_{AC_a} = \frac{1}{2} \{ \theta_{AC_0} + \theta_{AC_m} \} = (\theta_{C_a} - \theta_{A_a})$$

and

$$\theta_{BC_a} = \frac{1}{2} \{ \theta_{BC_0} + \theta_{BC_m} \} = (\theta_{B_a} - \theta_{C_a})$$

In a specific application the strains $\epsilon_A, \epsilon_B, \epsilon_C$ along OA, OB and OC † would be determined by

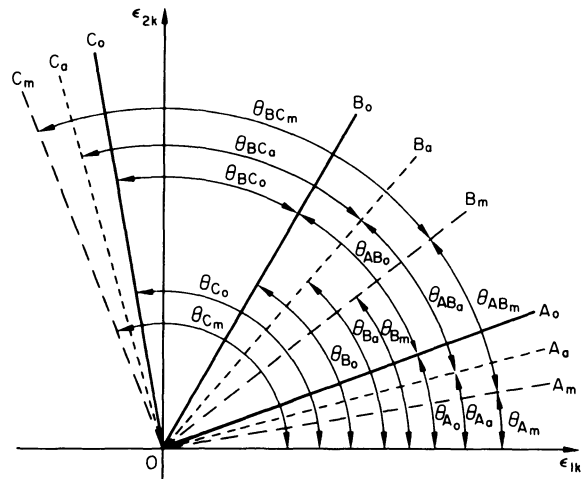


Figure 4. Relations between the initial, final and mean angles of orientation of the elements OA, OB and OC subscripted 0, m and a respectively.

measurement. Also the angles $\theta_{AB_0}, \theta_{AC_0}$, and $\theta_{AB_m}, \theta_{AC_m}$ would be measured before and after deformation respectively. Using the element OA in figure 4 as the OX axis of reference, Equation (8) written for the three directions OA, OB, OC would become

$$\begin{aligned} \epsilon_A &= \epsilon_x \\ \epsilon_B &= \frac{\epsilon_x + \epsilon_y}{2} + \frac{\epsilon_x - \epsilon_y}{2} \cos 2\theta_{AB_a} - \epsilon_{xy} \sin 2\theta_{AB_a} \\ \epsilon_C &= \frac{\epsilon_x + \epsilon_y}{2} + \frac{\epsilon_x - \epsilon_y}{2} \cos 2\theta_{AC_a} - \epsilon_{xy} \sin 2\theta_{AC_a} \end{aligned} \quad (10)$$

The solution of Equations (10) would yield ϵ_x, ϵ_y and ϵ_{xy} and hence the principal strains (ϵ_1, ϵ_2) could be calculated by means of Equation (2). Alternatively a solution might be obtained by the use of Murphy's construction (3).

ERRORS IN EQUATIONS (8) AND (10)

The errors in Equations (8) and (10) arise from the assumption that the angles θ_{ac}, θ_{as} and θ_a are equal to each other as stated in Equation (6).

Some aspects of these errors have already been discussed in conjunction with the evaluation of longitudinal strain in any direction for the case of known homogeneous principal strains, the latter acting on fibres fixed in direction during straining⁶. In this work the error will be considered from the point of view of using Equations (10) for the evaluation of the principal strains ($\epsilon_{1c}, \epsilon_{2c}$) in a given strain field with known principal strains ($\epsilon_{1k}, \epsilon_{2k}$). The difference between the calculated strains ($\epsilon_{1c}, \epsilon_{2c}$) and the known values ($\epsilon_{1k}, \epsilon_{2k}$) will be a measure of the error resulting from the use of Equations (8) and (10).

To accomplish this, 'exact data' should be used as input to equations (10) so that the results would

* It should be noted that both the axis ox and the element OB are subject to rotation relative to some 'fixed' direction, in this case the principal strain direction.

† The line elements OA, OB, OC might be thought of as elements of a 'usual strain rosette', with θ_{AB}, θ_{AC} as rosette angles, and θ_{A_0} as initial orientation of the rosette relative to the direction of the principal strain, ϵ_1 .

reflect the inherent errors of equations (8) and (10) and not of any experimental technique. Referring to figure 2(a) the exact expression for the lineal strain increment along any element such as OE brought about by known principal strain increments (ϵ_{1k} , ϵ_{2k}) is given by⁶

$$\epsilon_E = \frac{1}{2} \ln[\exp(2\epsilon_{1k}) \cos^2 \theta_0 + \exp(2\epsilon_{2k}) \sin^2 \theta_0] \quad (11)$$

while its final angle of orientation θ_m is

$$\theta_m = \tan^{-1}[\exp(\epsilon_{2k} - \epsilon_{1k}) \tan \theta_0] \quad (12)$$

Now, for any given values of the principal strain increments (ϵ_{1k} , ϵ_{2k}) and any set of initial orientations θ_{A0} , θ_{B0} and θ_{C0} of elements OA , OB and OC the exact values of the longitudinal strains ϵ_A , ϵ_B and ϵ_C as well as the final angles or orientation θ_{Am} , θ_{Bm} and θ_{Cm} may be determined using Equations

(11) and (12). These results may now be used as 'exact' data in equations (10) for the calculation of the principal strains, ϵ_{1c} and ϵ_{2c} . The differences ($\epsilon_{1c} - \epsilon_{1k}$) and ($\epsilon_{2c} - \epsilon_{2k}$) constitute the theoretical errors brought about by the approximations upon which equations (6) and (8) are based.

To reduce the labour of evaluating these errors for the infinite number of combinations of strain increments and of initial angular orientations it is desirable to consider those combinations of parameters for which the errors are largest. Firstly were there no shear strain, there would be no rotation and no error. Thus for a given value of effective strain the largest errors in the effective strain would occur for the largest possible shear strain, that is, for pure shear. Secondly it is evident from the geometry of the Mohr's circle that the errors are functions of the 'rosette' angles θ_{AB} and θ_{BC} as well as of the

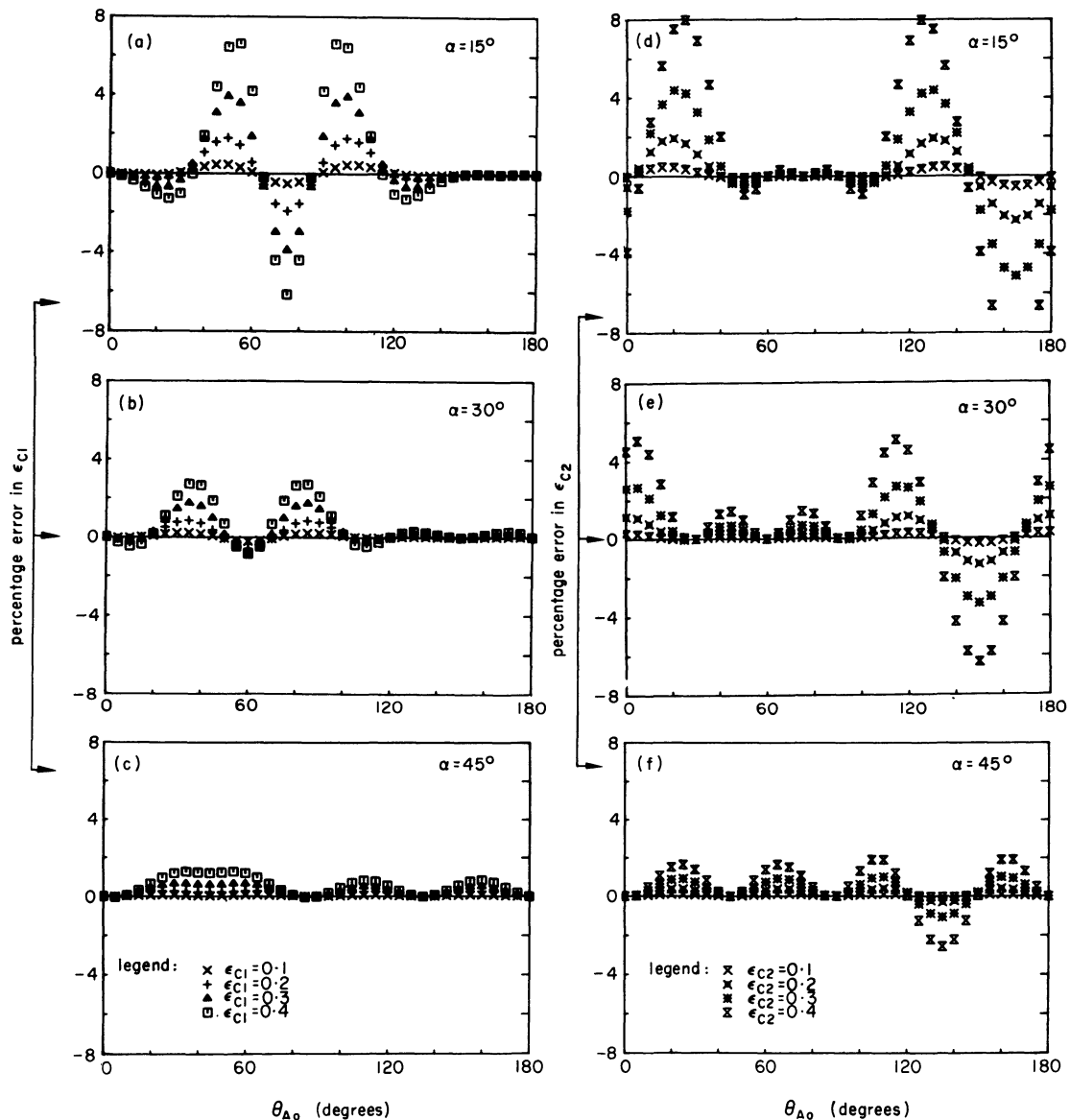


Figure 5. Percentage error in evaluating principal strain increments ϵ_{1c} and ϵ_{2c} by equations (10) plotted for selected values of $\theta_{AB} = \theta_{BC} = \alpha$ and of $\epsilon_{1k} = -\epsilon_{2k}$ versus the angle of rosette orientation θ_{A0} with respect to the direction of principal strain, ϵ_{1k} . The percentage error is defined as $100 (\epsilon_{1c} - \epsilon_{1k})/\epsilon_{1k}$ for ϵ_{1c} (plots a, b, c) and $100 (\epsilon_{2c} - \epsilon_{2k})/\epsilon_{2k}$ for ϵ_{2c} (plots d, e, f).

'orientation of the rosette', angle θ_{A_0} . Accordingly two computer programs were used: (A) to evaluate errors for the case when the angle θ_{AB} is not equal to θ_{BC} , and (B) where $\theta_{AB} = \theta_{BC}$; in both programs for a given value of known principal strain increments ($\epsilon_{1k}, \epsilon_{2k}$) the value of θ_{A_0} was varied to examine the magnitude of the error as a function of orientation of otherwise identical sets of measurements. For the case when the angles θ_{AB} and θ_{BC} are not equal the results show that for two sets of data-input, identical except for the angle θ_{BC} , the larger error is associated with the smaller value of θ_{BC} . It follows, therefore, that the results of computer program (B), in which $\theta_{AB} = \theta_{BC} = \alpha$ would serve as an envelope of maximum errors for given values of $\epsilon_1 = -\epsilon_2$, when the actual rosette angles are $\theta_{AB} \geq \alpha, \theta_{BC} \geq \alpha$. Sample results of this program are given in graphical form in figure 5, from which it is evident that the percentage error increases with the difference in principal strain increment, i.e. with the maximum shear strain, but decreases as the angles ($\theta_{AB} = \theta_{BC}$) between the elements increase. It should be noted in figure 5 that for the case of a 'rosette' with angles $\theta_{AB} = 45^\circ$ and $\theta_{AC} = 90^\circ$, the maximum error is less than 2.6 per cent for pure shear increment $\epsilon_{1k} = -\epsilon_{2k} = 0.4$, using principal strains as a basis of comparison. Calculated on the basis of effective strain this error is less than 1.3 per cent. For unusually small angles, such as $\theta_{AB} = \theta_{BC} = 15^\circ$, the maximum percentage error is about 9 per cent for $\epsilon_{1k} = -\epsilon_{2k} = 0.4$ and about 2.3 per cent for $\epsilon_{1k} = -\epsilon_{2k} = 0.2$, both evaluated on the basis of the given principal strains.

The *numerical error* is primarily a function of the maximum shear strain component considered as the difference between the principal strains; in other words, the numerical error for $\epsilon_1 = -\epsilon_2 = 0.3$ is the same as that for $\epsilon_1 = 0.4$ and $\epsilon_2 = -0.2$; the percentage error, however, depends on the magnitude of the basis of comparison. Thus, when the longitudinal strain ϵ_θ in any direction θ , is the basis of comparison, as was the case in Reference 6, the percentage error may become misleadingly large when, in fact, the actual error may have little significance on the results of the evaluation of the effective strain increment. For that reason the effective strain, ($\bar{\epsilon}$), or the maximum shear strain, or numerically maximum principal strain should be used as a basis of comparison when discussing errors.

The error in determining the direction of the principal strain increases with the magnitude of the shear strain, but decreases with the magnitude of the angles θ_{AB} and θ_{BC} . For the case of $\epsilon_{1k} = -\epsilon_{2k} = 0.4$ and for $\theta_{AB} = \theta_{BC} = 15^\circ$, the maximum error in question is about 1.3° and for the same strain increments, but for $\theta_{AB} = \theta_{BC} = 45^\circ$, the corresponding error is approximately 0.3° . For most practical purposes such errors would be considered negligible.

It should be emphasized at this point that the analysis of errors given above was undertaken not so much with the intent of correcting the results obtained by Equations (10), but rather to give the prospective user a means of estimating the upper bounds of the errors associated with Equations (10) and the specific conditions of taking data. Should his work demand a higher degree of accuracy than that attainable by the use of Equations (10), the 'exact' method of strain evaluation, outlined in the Appendix, should be employed.

CONCLUDING REMARKS

It has been shown that by introducing the mean angle of orientation of the element during straining into Equation (1), the latter may be employed for the analysis of large strain increments, such as those that arise in metal forming or in metal cutting. As an example of this consider figure 6, which is a Mohr's

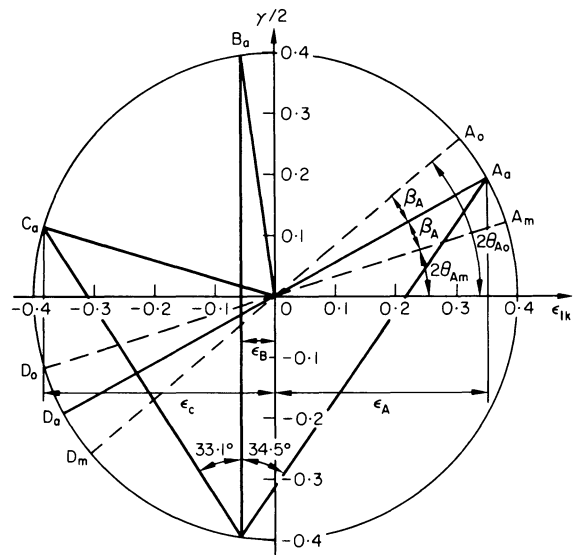


Figure 6. Murphy's construction of a Mohr's circle based on Equation (8) for data as follows:

$$\begin{aligned} \epsilon_{1k} &= -\epsilon_{2k} = 0.40 & \theta_{A_0} &= 20^\circ & \theta_{A_m} &= 9.29^\circ \\ \epsilon_A &= 0.351 & \theta_{B_0} &= 60^\circ & \theta_{B_m} &= 37.89^\circ \\ \epsilon_B &= -0.056 & \theta_{C_0} &= 85^\circ & \theta_{C_m} &= 78.98^\circ \\ \epsilon_C &= -0.385 & & & & \end{aligned}$$

circle constructed on the basis of Murphy's technique using data* calculated by Equations (11) and (12) for the case of known principal strains $\epsilon_{1k} = 0.4 = -\epsilon_{2k}$, for line elements initially oriented at $(\theta_{A_0}, \theta_{B_0}, \theta_{C_0}) = (20^\circ, 60^\circ, 85^\circ)$. It is seen that within the limits of drafting accuracy the principal strains given by the Mohr's circle correspond to the known values of ϵ_{1k} and ϵ_{2k} , which were used for the calculation of ϵ_A , ϵ_B , and ϵ_C as well as for the calculation of the final, and thus of the average, angular orientations of these elements. Taking for reference the direction of action of the principal strain, ϵ_{1k} , the rotation of the

* In practice data would be derived from experimental measurements.

element OA is $\beta_A = (\theta_{A_0} - \theta_{A_m}) =$ (shear strain ordinate corresponding to point A_a) radians $= (0.195 \times 180^\circ/\pi = 11.2^\circ)$. Constructing this angle on both sides of OA_a on the Mohr's circle in figure 6 gives the initial and the final position of OA (i.e. $2\theta_{A_0}$ and $2\theta_{A_m}$ respectively).

The concept of using the mean angle of orientation (θ_a) may at first appear to be somewhat inconvenient; however, it should be borne in mind that $2\theta_{AB_a}$ called for by Equation (8) is (approximately) equal to $(\theta_{AB_0} + \theta_{AB_m})$ and both of these values, in most practical cases, are known from experimental data, since they are the angles between line elements OA and OB before and after straining respectively. If such is the case, there is no need to find the initial or final orientation of any line element—what is needed is the evaluation of the principal strains in magnitude and direction. For this purpose the simple form of Equation (8) may be employed for both graphical or computerized analysis of large plastic strains in much the same manner as Equation (1) is employed for the analysis of small strain increments.

ACKNOWLEDGMENTS

The author is gratefully indebted to Professor P. L. Balise and to his son, Mr David Balise, who did all of the computer work associated with this study. Thanks are also due to Professor R. Taggart and Mr J. McCabe for their valuable comments on the manuscript. Finally, the author gratefully acknowledges that his appointment, in the winter of 1973, as a Visiting Fellow at the Battelle Seattle Research Centre, allowed him to complete this work.

APPENDIX

The 'exact' equation relating the logarithmic* longitudinal and the shear strains for the case of line elements in a plane containing the two principal strains ϵ_1 and ϵ_2 , may be obtained² by considering two elements OP and OQ , each of unit length, initially at right angles to each other as shown in figure A. If P' and Q' denote the positions of P and Q after deformation, describable by principal strain increments $(\epsilon_1, \epsilon_2, \epsilon_3)$, the co-ordinates of P' are $(e^{\epsilon_1} \cos \theta_0, e^{\epsilon_2} \sin \theta_0)$ and the strain along OP is

$$\epsilon_p = \ln \left(\frac{OP'}{OP} \right) = \ln [(\cos^2 \theta_0) e^{2\epsilon_1} + (\sin^2 \theta_0) e^{2\epsilon_2}]^{1/2}$$

or

$$e^{2\epsilon_p} = \frac{e^{2\epsilon_1} + e^{2\epsilon_2}}{2} + \frac{e^{2\epsilon_1} - e^{2\epsilon_2}}{2} \cos 2\theta_0 \quad (1A)$$

* Note: Exact strain analysis in the Lagrangian and Eulerian descriptions is discussed in 'Evaluation of Large Strains in Industrial Applications', by J. Wolak and V. J. Parks. *J. Testing and Evaluation* (1974) 2, 6.

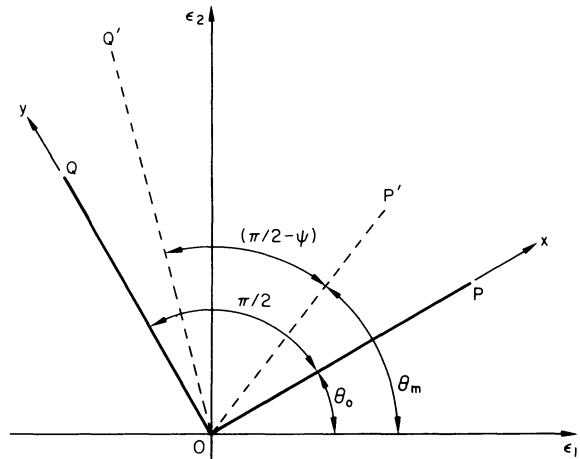


Figure A. Line elements OP, OQ take positions OP', OQ' due to principal strain increments (ϵ_1, ϵ_2) , the directions of which lie in the plane $POQ P'Q'$.

The direction cosines of OP' are

$$\left(\frac{e^{\epsilon_1} \cos \theta_0}{e^{\epsilon_p}}, \frac{e^{\epsilon_2} \sin \theta_0}{e^{\epsilon_p}} \right)$$

Similarly the direction cosines of OQ' are

$$\left(\frac{-e^{\epsilon_1} \sin \theta_0}{e^{\epsilon_q}}, \frac{e^{\epsilon_2} \cos \theta_0}{e^{\epsilon_q}} \right)$$

Associating the directions OP and OQ with OX and OY respectively, and denoting the change in angle POQ by ψ , we have

$$\begin{aligned} \cos \left(\frac{\pi}{2} - \psi \right) &= \sin \psi \\ &= \left[-\frac{e^{2\epsilon_1} \cos \theta_0 \sin \theta_0}{\exp(\epsilon_x + \epsilon_y)} + \frac{e^{2\epsilon_2} \cos \theta_0 \sin \theta_0}{\exp(\epsilon_x + \epsilon_y)} \right] \end{aligned}$$

or

$$(\sin \psi) \exp(\epsilon_x + \epsilon_y) = - \left[\frac{e^{2\epsilon_1} - e^{2\epsilon_2}}{2} \right] \sin 2\theta_0 \quad (2A)$$

Replacing ϵ_p with ϵ_x in Equation (1A), rearranging the latter, squaring and adding it to the square of (2A) results in

$$\begin{aligned} \left[e^{2\epsilon_x} - \frac{e^{2\epsilon_1} + e^{2\epsilon_2}}{2} \right]^2 + \exp[2(\epsilon_x + \epsilon_y)] \sin^2 \psi \\ = \left[\frac{e^{2\epsilon_1} - e^{2\epsilon_2}}{2} \right]^2 \end{aligned} \quad (3A)$$

Thus, the angular change ψ in the originally right angle between the OX and OY axes is related to the longitudinal strain increments ϵ_x and ϵ_y through a Mohr's circle expressed by Equation (3A).

Of the several possible definitions of large shear strain increment, γ_{xy} , two are frequently used

$$\gamma_{xy} = \psi \quad (4A)$$

and

$$\gamma_{xy} = \tan \psi \quad (5A)$$

For relatively small strain increments of the order of 0.1, both definitions yield much the same results. It should be noted, however, that according to definition (4A) the shear strain could not be greater than $(\pi/2)$. Hence, for very large shears, definition (5A) is more realistic than definition (4A).

REFERENCES

1. O. Mohr. *Abhandlungen aus dem Gebiete der technischen Mechanik*, 2nd ed., Wilhelm Ernst und Sohn, Berlin (1914).
2. A. Nadai. *Theory of Flow and Fracture of Solids*, vol. I, McGraw-Hill (1950).
3. G. Murphy. A graphical method for the evaluation of principal strains from normal strains, *Trans. A.S.M.E. (J. Appl. Mechanics)*, **67**, (A-209 to A-210), December 1945.
4. C. C. Perry and H. R. Lissner. *The Strain Gauge Primer*, McGraw-Hill, 1962.
5. V. J. Parks and A. J. Durelli. On the definitions of strain and their use in large strain analysis, *Proc. S.E.S.A. XXIV* (1967) **1**, 279-280.
6. J. Wolak, On the evaluation of large homogeneous plastic strains, *Experimental Mechanics* (1972) **12**, no. 10, 475-80.
7. V. J. Parks and A. J. Durelli. Various forms of the strain-displacement relations applied to experimental strain analysis, *Proc. S.E.S.A. XXI* (1964) **1**, 37-47.
8. G. Sines. *Elasticity and Strength*, Allyn and Bacon, Inc. 1969.

SMOOTH HOLE WALL PIERCING IN THE MANUFACTURE OF PRECISION MECHANISMS

by

R. JOHNSTON* and K. SWIFT†

SUMMARY

This paper describes an investigation into the feasibility of eliminating the bushes used in the workplates of a precision mechanism by the use of a smooth hole wall piercing method. The background to the process of smooth hole wall piercing or 'finish' piercing is discussed and initial experimental work is described in which a number of different possible punch profiles was assessed.

Pre-production testing on a basis of recommendations arising from the initial tests is also described. The Cost Savings resulting from the adoption of the preferred punch profile for workplate and mechanism manufacture are given.

INTRODUCTION

The mechanisms referred to in this paper are those used in pressure measurement instrumentation, a typical example of the mechanism is shown in figure 1. The top and bottom workplates, details of which are shown in figure 2, provide location and the bearing surfaces for the shafts carrying the pinion and the rack. The components of the class of mechanisms discussed in this paper are manufactured from Stainless Steel BSS EN58J in order to meet corrosion resistance requirements.

The mechanisms are manufactured in batches of about 20 000 at approximately three monthly intervals.

In the previous manufacturing method bearings for the pinion and rack shafts were provided by bushes which were pressed into holes in the stainless steel workplates pierced by a conventional piercing method. The original specification for the material of the bushes was sintered monel metal, but due to poor dimensional consistency and excessively long delivery periods of these, turned type BSS EN58J stainless steel bushes were specified as an alternative. The use of bushes pressed into conventionally pierced holes in this way suffered from the disadvantage that the alignment and positional accuracy of the bushes were subject to all the various inaccuracies associated with conventionally pierced holes.

The characteristic partly sheared, partly cracked, surface of the hole walls and the unevenness of the workplate surfaces due to local dishing and burr formation associated with the conventional piercing

process, all contributed to possible bush misalignment, hole pitch error, and variations between mechanisms in the amount of end float on the pinion shaft leading in some cases to binding due to the bushes not seating properly.

It was found necessary to introduce hand reaming of the bushes as a corrective measure during assembly and this was clearly unsatisfactory.

Consideration of possible methods of improving the quality of the pierced hole and reducing the effect of the piercing operation on the workplate material surrounding the pierced hole, in order to improve the alignment and positional accuracy of the bushes, eventually led to consideration of the possibility of eliminating the bushes altogether.

Initial consideration of the requirements of the piercing process if bushes were to be eliminated, i.e. a crack free hole wall and workpiece dishing and burr formation to be within acceptable limits, indicated that double or triple action tooling might be necessary.

It was thought, however, that the possibilities of essentially single action smooth hole wall piercing processes should be investigated at the outset.

PREVIOUS WORK

In piercing, using conventional square edged tooling, the process is generally characterized by cracking of the workpiece material at the wall of the pierced hole. The reasons for the presence of this cracking are discussed by Stromberger¹ in relation to a theory

* Department of Mechanical Engineering, University of Salford.

† University of Salford Industrial Centre, formerly with Budenberg Gauge Co., Altrincham, Cheshire.

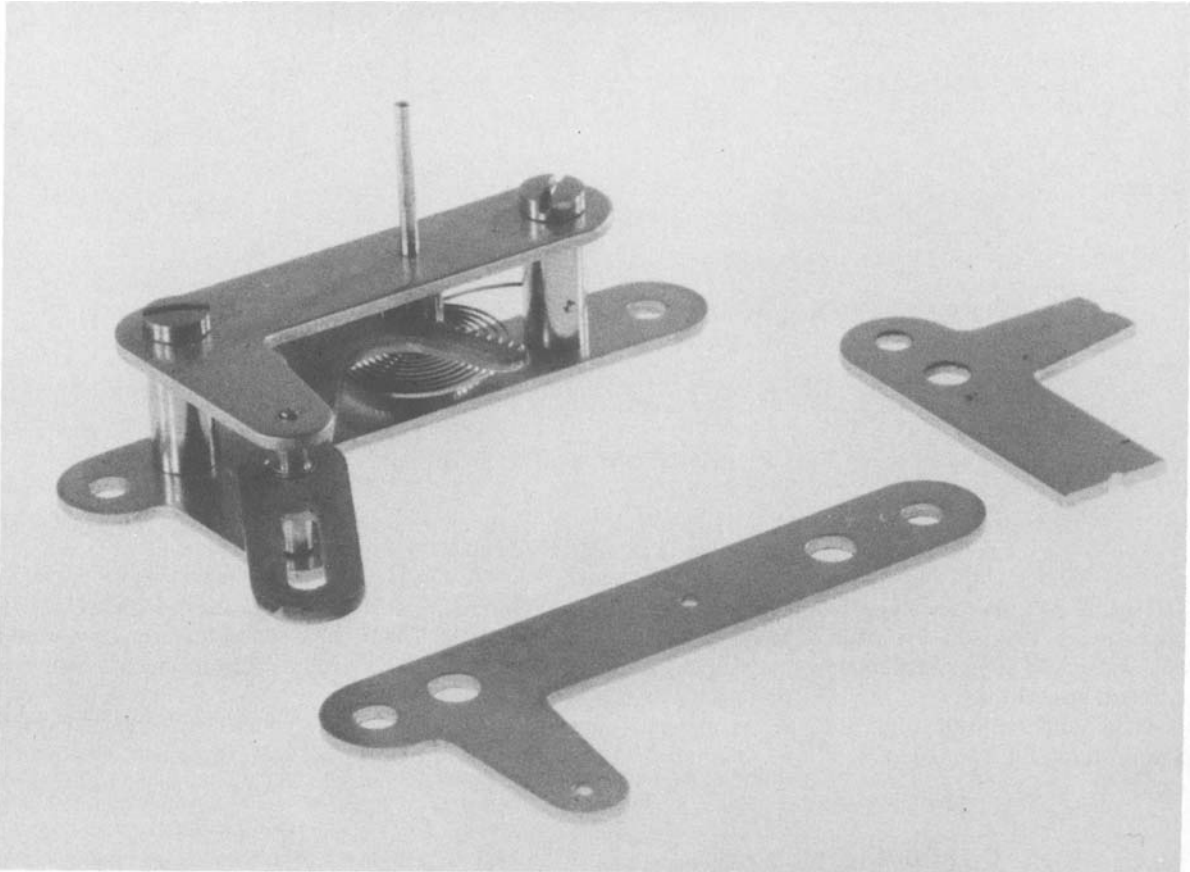


Figure 1. Photograph showing a typical pressure gauge movement and one of the workplates.

proposed by Siebel². This theory proposes that limits of yielding, i.e. smooth shear separation, and rupture (i.e. cracking), are defined in the Mohr Stress plane by intersecting straight lines, see figure 3, and that in conventional blanking, the level of hydrostatic stress

in the deformation zone is such that one principal stress is tensile, leading to separation of the workpiece material by cracking rather than by smooth shearing.

Stromberger goes on to suggest that in order to produce a pierced hole in which the hole wall is free from cracking, it is necessary that the stress state in the workpiece material in the deformation zone should be such that the hydrostatic stress is of sufficiently large compressive magnitude to ensure that the corresponding Mohr stress circle is bounded by the line defining the limit of yield stress in shear

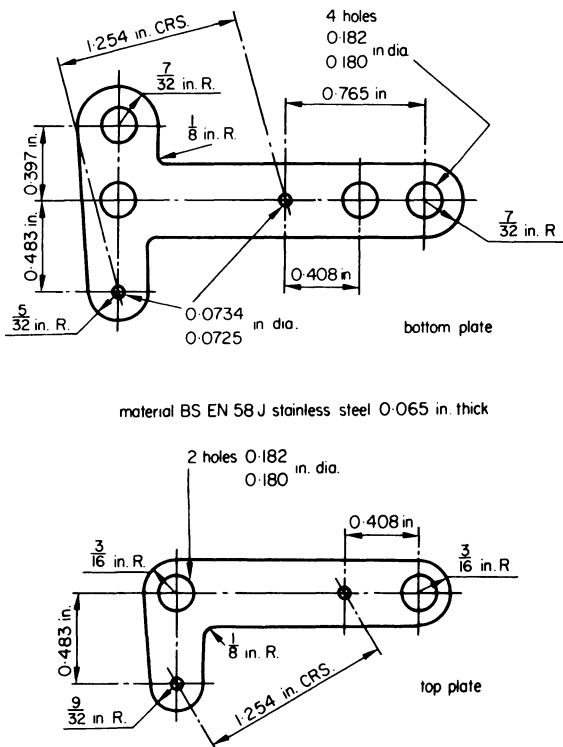


Figure 2. Details of the mechanism workplates.

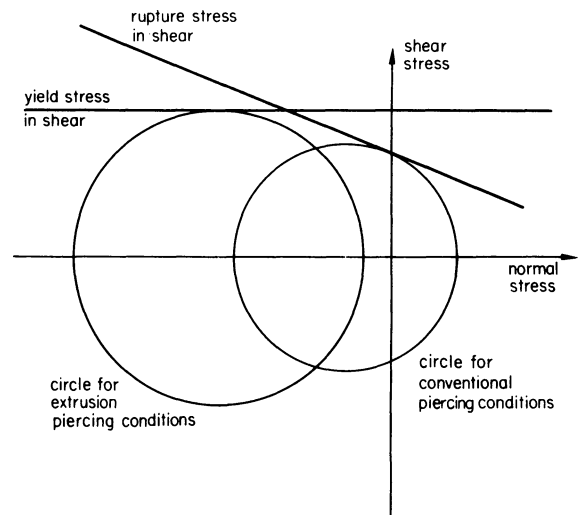


Figure 3. Mohr stress plane for piercing operations (after Stromberger).

rather than by the line defining the rupture stress in shear. Under these conditions all three principal stresses are compressive, and the material separates by smooth shearing.

The design of punch shown in figure 4 was proposed by Stromberger as being suitable for the fulfillment of the conditions that there should be no tensile stress in the deformation zone. The action of this punch is such that the punch face pre-pierces the workpiece material under conditions of large punch/die clearance, this produces a rough slug leaving a rough cracked hole wall. The shoulder part of the punch following up extrudes or shaves a ring of material under conditions of all-round compressive stress since the material being removed is trapped by the tapered part of the punch flank and the radius at the punch shoulder.

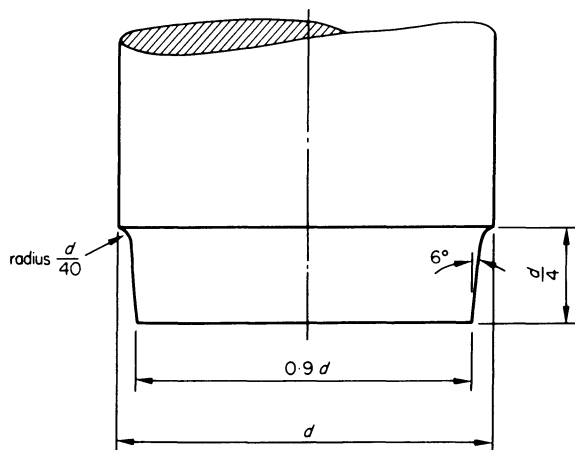


Figure 4. Proportions of the Stromberger extrusion piercing punch.

It has been shown by PERA*³ that under suitable conditions of punch/die clearance, a radius or chamfer on the punch corner can be effective in the prevention of cracking at the hole wall during piercing. The most suitable geometry of the radius or chamfer in a particular case is dependent upon the workpiece material properties and the proportions of hole size and material thickness. Microhardness surveys have been carried out by PERA⁴ on cross sections of partially pierced material which have shown that one of the effects of the introduction of a radius or chamfer at the punch corner is to cause a local broadening of the region of plastic deformation without any appreciable change in the levels of maximum hardness. The effect of this is to decrease the hardness gradient with respect to planes parallel to the surfaces of the workpiece.

Based on this experimental evidence Oxley⁵ has proposed a qualitative mechanism to account for the capability of the radiused or chamfered punch to prevent cracking. In this mechanism it is argued that the reduction in hardness gradient associated with the radiused or chamfered punch leads to an increase in

hydrostatic pressure in the deformation region particularly near the punch corner, and thus cracking in the region of the punch corner is inhibited.

The blanking process may be looked upon as the counterpart of the piercing process. From the point of view of fracture in the workpiece material, the significant factor is that in piercing the punch corner crack affects the stock material or the hole wall, whereas in blanking, the die corner crack affects the slug or blank edge. The mechanisms of crack prevention in the blanking and piercing processes clearly have much in common.

Johnston⁶ has reported evidence that in the fine blanking process the function of the tooling in prevention of cracking at the blank edge is to increase the magnitude of the hydrostatic stress in the region of plastic deformation in a compressive sense to such an extent that there is no tensile principal stress.

All of the available evidence therefore points to the need to provide a hydrostatic component of stress in the deformation zone of a magnitude of about the yield stress of the material in pure shear. The most promising methods of achieving this would appear to be either by the use of a Stromberger or PERA type punch profile or by all-round clamping of the material by a stripper plate and counterpunch. The former of these two possible methods is to be preferred on the grounds of simplicity.

EXPERIMENTAL WORK

The purpose of the experimental work was to examine the feasibility of producing pierced holes in the top and bottom workplates of sufficient quality for the pinion and rack shafts to be mounted directly in them thus eliminating the use of bushes.

It was decided to use a range of punches which were likely to give the desired result, i.e. Stromberger and PERA type punches, together with other types for comparison purposes.

A number of top and bottom workplates were obtained with the 0.180/0.182 in diameter holes pierced in them, see figure 2, but without the holes pierced to take the bushes.

An experimental press tool was constructed which was designed to accept either the top or bottom workplate, locating the plates by means of dowels, and to pierce the required 0.0734/0.0725 in diameter holes. The experimental tool was intended either for use on a fly press or on a small universal testing machine.

It was intended to examine the effect not only of punch profile but also the effect of variation in radial punch/die clearance within a limited range. It was originally intended to test clearances of about 0.5, 1.0 and 1.5 per cent.

However, starting with the 0.5 per cent clearance, it became clear during tests with 1 per cent clearance that the hole quality was so much worse than with the 0.5 per cent clearance that it was considered unnecessary to go on to investigate the effect of 1.5 per cent clearance.

Figure 5 shows details of all the punch profiles tested in the initial experimental work, the characteristics of the punches are shown in Table 1.

* Production Engineering Research Association of Great Britain.

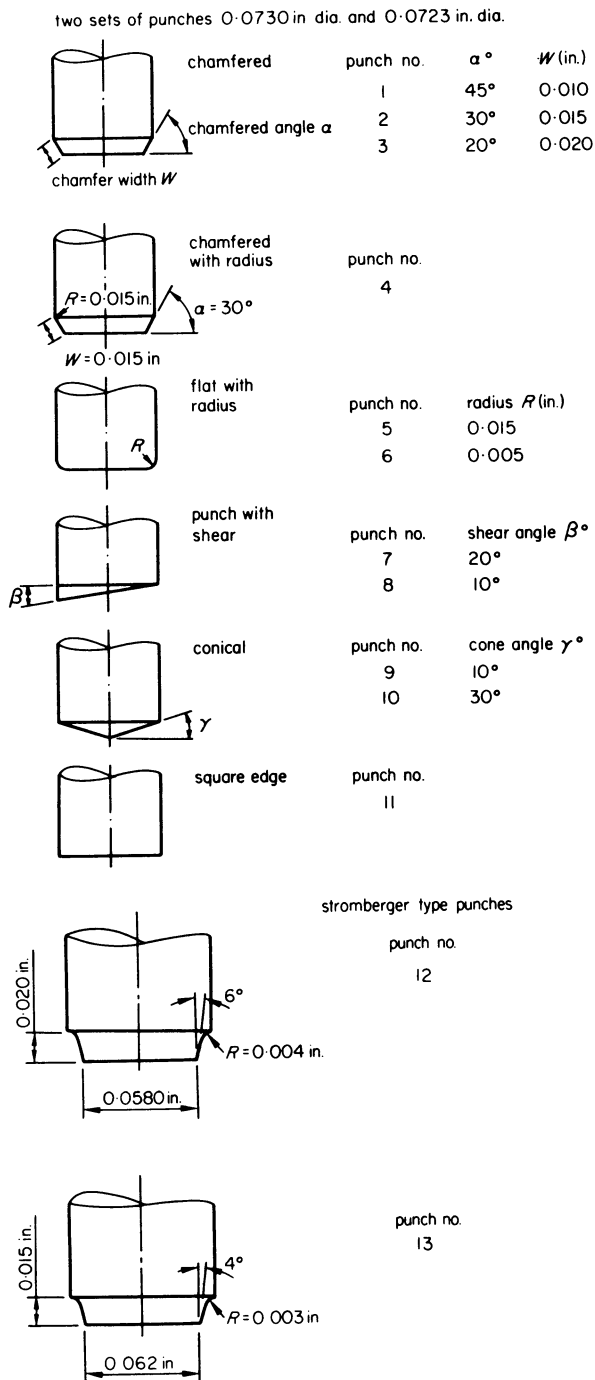


Figure 5. Details of the punch profiles used in the experimental work.

TABLE 1

Punch number	Description
1-4	PERA chamfered punch, variation of chamfer width and chamfer angle
5 and 6	PERA radiused punches different corner radii
7 and 8	Punches with 'shear' different 'shear' angles
9 and 10	PERA conical (semidome) different cone angles
11	Conventional sharp edged punch
12 and 13	Stromberger type punches

All of the workplates used in the experiments were of stainless steel BSS EN58J 0.0650 in thick. The die diameter used throughout was 0.0736 in. Two different sets of punch diameters were used, 0.073 in and 0.0723 in, giving corresponding radial punch/die clearances, expressed as percentages of the material thickness, of 0.46 per cent and 1.0 per cent. Both punch and die material were 18 per cent tungsten high speed steel.

The lubricant ESSO TSD 996, a chlorinated extreme pressure lubricant, was used throughout the tests.

The punch load and the height of the burr formed were measured for the plates using each of the punches tested. The results of these measurements are shown in Table 2. The amount of dishing was less easy to measure but was thought to be substantially consistent at about 0.004 in for all of the punches.

Pairs of workplates made using each of the punches were subsequently used in the assembly of a mechanism.

DISCUSSION

Examination of the hole wall conditions produced by each punch and radial punch/die clearance was carried out by visual inspection following sectioning of the plate across the centre of the hole. It was found necessary to pierce a number of plates at each condition in order to obtain a reasonably representative result. At the small punch diameter used in this work there were problems associated with maintaining concentricity of the punch and die. Any tendency for the punch to be displaced laterally was reflected in the occurrence of fracture at the hole wall. It was recognized that the experimental tool involved the use of punches which were rather longer than was desirable, but an attempt was made in the design of the stripper plate to provide as much punch guidance as possible.

The initial experimental work showed that three of the punches gave substantially better results (in terms of crack free holes) than the others. These three were:

- (1) Punch number 2 30° chamfer 0.015 in wide
- (2) Punch number 10 30° semidome
- (3) Punch number 12 Stromberger punch with larger taper angle.

The second phase of the experimental work was to carry out tests on punches using three preferred profiles on a production basis. All of the initial experimental work had been carried out under slow speed conditions either on a small universal testing machine or a fly press and it was recognized that the change to crank press speeds, and longer production runs, might reveal aspects of the process which had not been apparent in the initial tests.

PRE-PRODUCTION TESTS

In order to assess the suitability of the process under production conditions, the three punch profiles which had shown the best results in the initial

TABLE 2

Punch number†	Ratio $\frac{F_N}{F_{11}}$ *		Burr height $\times 10^{-3}$ in	
	0.46 per cent Radial clearance	1.0 per cent Radial clearance	0.46 per cent Radial clearance	1.0 per cent Radial clearance
1	1.17	1.17	3.5	3.5
2	1.15	1.16	3	3
3	1.14	1.16	3	3
4	1.17	1.20	4	4
5	1.18	1.21	4.2	4.2
6	1.10	1.16	4	4
7‡	1.28	—	4	—
8‡	1.30	—	3	—
9	1.13	1.18	2.5	3
10	1.12	1.17	3	2.5
11	1.0	1.00	2.5	2.5
12	1.30	1.30	2	2
13	1.31	1.31	2	2

* The ratio F_N/F_{11} is the ratio of the maximum punch force for a particular punch to the maximum punch force for the square edged punch (punch number 11).

† For identification of the punch numbers see figure 5.

‡ For punches 7 and 8 damage to the punch was experienced at the lower value of clearance and it was thought undesirable to carry out the tests for these punches at the higher value of clearance.

experimental work, i.e. punches numbered 2, 10 and 12 were subjected to a pre-production test.

A tool based on that used in the initial experimental work was designed and manufactured in order to carry out the pre-production and subsequent short term workplate production. The tool, which employed about 0.5 per cent punch/die radial clearance and incorporated a simple mechanized feed

arrangement, was used to pierce the workplates as a secondary operation. About 1000 workplates were produced using each punch. Of the three punches tested the Stromberger profile (figure 6) gave the most acceptable results based on the following criteria:

- (1) The surface finish of the pierced holes was superior in quality and consistently giving crack

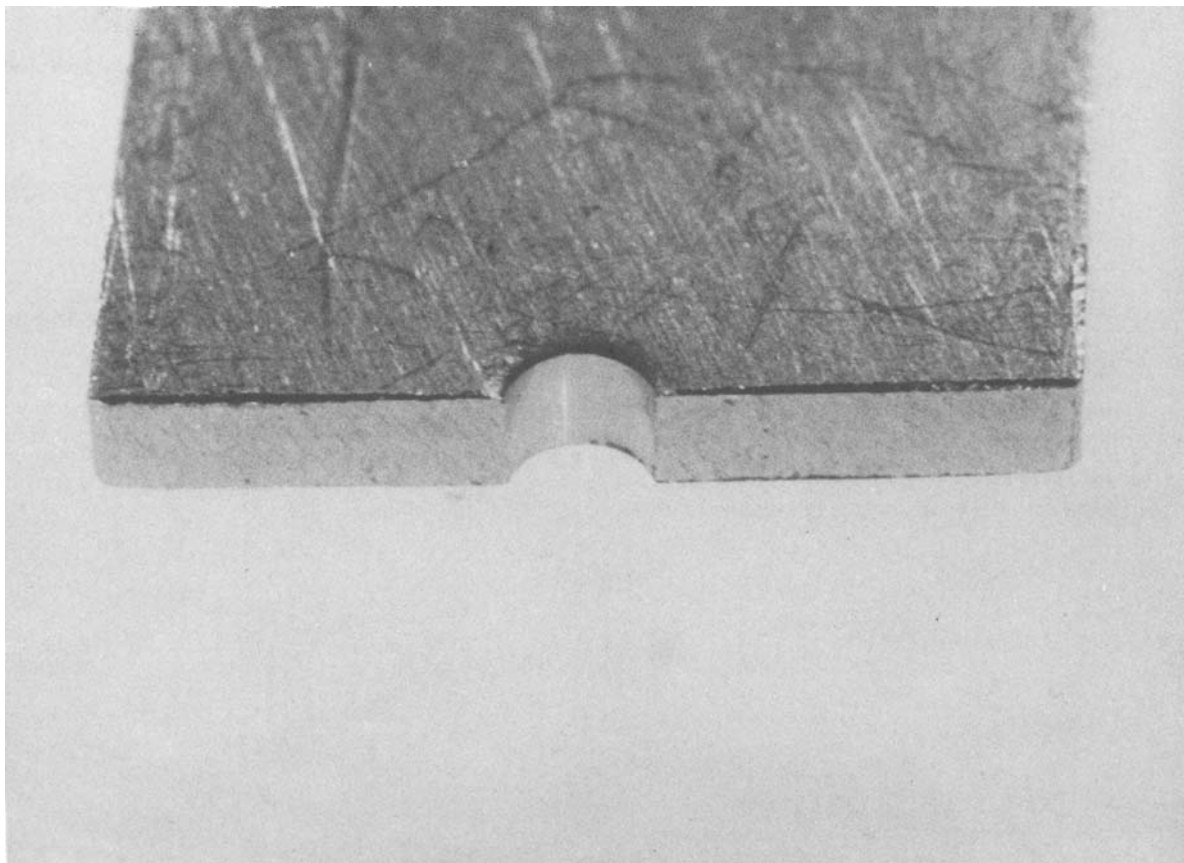


Figure 6. Photograph of a section of a workplate showing a hole pierced using a Stromberger type punch.

free hole walls, whereas the other two profiles tended to produce cracking at the lower edges of the holes.

- (2) The burr height and the amount of dishing of the component was at a minimum with the Stromberger profile.
- (3) The burr was removed from the workplates by a barrelling operation, and much more difficulty was experienced in deburring components produced using punches 2 and 10. In cases where the burr height was relatively large the barrelling operation resulted in a tendency for the burr to be turned back into the hole.

In order to provide further comparison of the workplate manufacture incorporating smooth hole wall piercing and the workplates incorporating bushes, accelerated life tests were carried out in which sets of mechanisms assembled from both types of workplates were subject to rapid cycling within the designed range of mechanism movement.

The results of this accelerated life testing showed that the mechanisms assembled from pierced workplates revealed rather less wear than corresponding mechanisms assembled from bushed workplates.

Sets of mechanisms were also evaluated in tests on high sensitivity instruments in which performance of the mechanism in providing displacement magnification free from sticking is critical. The assemblies incorporating pierced workplates showed a consistently superior performance to that of the bushed movement and enabled increased sensitivity to be achieved.

INDICATIONS OF COST SAVINGS

Comparison of the costs of plates, the mechanism assembly and the final product for both the plates with bushes and plates with pierced holes gave the following results

- (1) The saving in cost on a typical pair of workplates through the replacement of bushed plates by smooth hole wall piercing was about 60 per cent.
- (2) The consequent saving in cost of a mechanism assembly was about 30 per cent.
- (3) The consequent saving in cost of the final product was up to 7 per cent depending on the instrument type.

The savings indicated above are on the basis of the introduction of the smooth hole wall piercing method

as a secondary operation on plates which had been conventionally blanked.

It is clearly possible to achieve further significant cost savings by incorporating the piercing tool into a progressive blank and pierce tool enabling the finished workplates to be produced directly from strip material.

The major cost saving was achieved in the elimination of the bushes and their fitting into the plates. A further appreciable saving was made in the final mechanism assembly where the consistency of pierced hole diameter and pitch enabled the output of assemblies to be increased, completely eliminating the hand reaming operation necessary with the bushed assembly.

CONCLUSIONS

- (1) Tests carried out on experimental tooling have shown that the smooth hole wall piercing method is feasible and can be a viable alternative to the use of bushes in a precision mechanism movement in order to provide bearing surfaces for rotating shafts.
- (2) Initial experimental work and subsequent pre-production tests have shown that the preferred design of punch profile in order to produce smooth hole wall piercing with small burr height is that proposed by Stromberger.
- (3) Substantial cost savings were shown to result from the introduction of the smooth hole wall piercing method in the case of the application considered.

ACKNOWLEDGMENTS

The authors wish to acknowledge the work of John Morgan on the Initial Experimental Work, and the co-operation of Mr Richard Budenberg, Director of the Budenberg Gauge Co.

REFERENCES

1. C. Stromberger and T. Thomsen. *Werkstatt V Betrieb* (1965) **98**, 739.
2. E. Siebel. *Werkstoffmechanik* (1952) **14**, 465.
3. PERA Report, **70**.
4. PERA Report, **93**.
5. P. L. B. Oxley and C. F. Noble. *Research Note 139*, College of Aeronautics, Cranfield.
6. R. Johnston, B. Fogg and A. W. J. Chisholm. *Proceedings 9th MTDR Conference*, Birmingham (1968) **397**.

AUTHOR INDEX

- Abdul, N. A., 431
 Agullo, J. B., 349
 Akgerman, N., 57
 Alexander, J. M., 525
 Al-Hassani, S. T. S., 571
 Altmann, H., 669
 Al-Tornachi, M. J. K., 713
 Arndt, G., 203
 Arora, G. K., 155
 Athersmith, D., 45
 Attwell, J. W., 1
 Auramovic, D., 165
- Bedini, R., 209
 Berry, G., 713
 Besant, C. B., 115
 Bex, P. A., 695
 Biswas, S. K., 135, 445
 Blazynski, T. Z., 399, 407
 Bramley, A. N., 431, 437
 Bruce, J. W., 97
 Brunn, P., 19, 30
 Bueno, R. M., 175
 Burton, J. A., 227
- Carrie, A. S., 145
 Chia, H. T., 545
 Chitkara, N. R., 497, 695
 Chowdhury, M. I., 237
 Cole, B. N., 585
 Craven, F. W., 37
 Crookall, J. R., 45, 89, 373, 385
- Das, M. K., 617
 Datsko, J., 579
 Davies, B., 97
 Davies, R., 637
 Dean, T. A., 455, 479
 De Barr, A. E., 97
 Desai, I. M., 417
 Drapier, J.-P., 251
 Droubi, G., 259, 295
 Dugdale, D. S., 280
- Ebbesen, J., 473
 Edney, R. C., 115
 El-Sobky, H., 399
 El-Wakil, S. D., 645
 Etheridge, R. A., 195
 Eversheim, W., 7
- Farmer, J., 685
 Farmer, L. E., 597
 Farnworth, G. H., 227
 Ferenczy, J., 129
 Fort, P., 165
- Foster, T. G., 121
 Fukuoka, S., 423
- Gieseke, E., 71
- Hanna, N. H., 245
 Hawkyard, J. B., 507
 Hewitt, R. L., 677
- Jamil, A. T. M., 89
 Jebb, A., 115
 Johnson, W., 267, 497, 507, 695
 Johnston, R., 729
- Kaliszer, H., 311, 333
 Khor, B. C., 373
 Kiuchi, M., 423
 Knight, W. A., 135, 289
 Kobatake, K., 539
 Kobayashi, S., 561, 603
 Kwiatkowski, A. W., 245
- Langdale, D. J., 391
 Lee, P. W., 385
 Lee, S. H., 561
 Lehotzy, J., 129
 Leonard, R., 19, 30
 Lorrek, W., 703
 Lyon, J., 317
- McDermott, R. P., 437
 McNulty, G. J., 305
 Makino, H., 325
 Martin, P., 251
 Miese, M., 7
 Moore, J. J., 685
 Murthy, T. S. R., 155
- Nakagawa, T., 611
 Navaratne, F. R., 507
 Negm, M., 637
 Noble, C. F., 365
- Osakada, K., 539
 Osman, A. E. M., 289
- Pages-Fita, J., 349
 Parsons, B., 585
 Pattinson, E. J., 317
 Pawelski, O., 703
 Pinotti, P. C., 209
 Pirog, B. Z., 597
- Rathmill, K., 19, 30
 Raybould, D., 627
 Reid, S. R., 391

- Rifai, M. A., 651
Roberts, W. T., 685
Robertson, D. K., 515
Rooks, B. W., 445, 487
Rowe, G. W., 333, 417
- Sadek, M. M., 237, 259, 289, 295
Sætre, I., 473
Sansinenea, J. M., 175
Sansome, D. H., 515, 545, 551
Sarwar, M., 217
Scaife, T., 585
Scott, A. J. A., 195
Sekiguchi, H., 539
Sfantsikopoulos, M. M., 365
Shah, S. N., 603
Sheppard, T., 659
Shin, H. S., 417
Shiomori, K., 611
Silva, T. M., 479
Slater, R. A. C., 339
Slut&as, J., 473
Smart, E. F., 187
Spurgeon, D., 339
Stockmann, P., 107
- Stokes, R. F., 685
St&oren, S., 473
Sturgess, C. E. N., 455
Swift, K., 729
- Thomas, A., 467
Thompson, P. J., 217
Tobias, S. A., 237, 295
Tomlinson, P. N., 677
Townley, S., 407
Trent, E. M., 187
Trmal, G., 311, 333
Turner, T. W., 525
- Uttley, J. R. S., 497
- Venter, R. D., 677
- Wasiukiewicz, I., 79
Wilson, D. V., 685
Wolak, J., 721
Wollner, G., 107
Wong, V. G., 617
- Young, M. J. R., 551

SUBJECT INDEX

- CAD and CAM Session, 105**
automatic design, 121
axisymmetric hot forging dies, 135
combination of CAD and CAM, 107
generalized concept of production, 129
machine tools
 component design, 115
 manufacture, 107, 115
 production, 115
plant layout, 145
- CNC and DNC Session, 55**
adaptive control
 improvement of accuracy, 79
 turning operations, 71
closed loop positioning control, 63
computer NC system, 57
machining centre, 89
technological forecast for, 97
- Cold Forming, Blanking, Cropping Session, 577**
analysis of stresses, 617
clamp shearing, 612
cold extrusion, 597
cold formed parts
 high speed forming machine, 585
 strength of the material, 579
cold heading, 603
crankshafts, 612
cropping, 617
extrusion force, 597
- Compaction and Sinter Forging Session, 625**
compaction
 aluminium and iron powder, 627
 very high pressures, 677
metal powders, 659
plasticity, 645
powder pre-forms
 cold forgeability, 651
 hot extrusion, 637
 stainless steel, 669
production of extruded material, 659
sintered iron-powder, 675
- Die Life in Hot Forming Session, 465**
cracking and fracture, 467
die life estimation, 473
die wear
 drop forging dies, 479
 high speed forging, 487
effect of die temperature, 487
hot-work die steels, 467
metal flow, 487
- Drawing, Tube and Sheet Forming Session, 523**
anisotropic sheet metals, 561
bore expanding and flange drawing, 561
die-less drawing
 fundamental study, 539
 titanium and some steels, 525
drawing
 bi-metal rod and tube, 545
 oscillatory deep drawing, 551
tube inversion, 571
- ECM, EDM and Explosive Welding Session, 363**
bi-metallic tubes, cold drawing, 407
composite metal-foil cylinders, 399
diamond-grit cup-wheel, for ECM, 365
electrode wear and erosion, EDM, 385
machined surfaces, EDM, 373
modulation of interface waves, 391
- General Aspects of Forming Session, 683**
all-metal systems, 685
circular plates, 695
hole flanging and piercing, 695
hydrostatic pressure, 703
M2 high speed steel, 713
natural strain increments, 721
plastic deformation, 703
smooth hole wall piercing, precision mechanisms,
 729
strong composite, 685
toughness, 713
- Grinding and Surface Topography Session, 309**
criteria for wheel life, 311
criteria of redress life, chattering phenomena, 325
dressing treatment, 317
fibre-optics transducer, 341
optimization, 311
plain bearings, running-in, 333
stylus technique, 349
surface roughness, 341
- Group Technology Session, 5**
batch size selection, 29
developments and modes of application, 7
interaction of NC and GT, 37
organizational aspect, computer simulation, 45
total company appraisal, 19
- Hot Forming Session, 415**
alloys in mashy state, 423
closed die forging, 445
extrusion of fluted sections, 455
forging analysis, 437

- high speed steel, 455
 - load and energy estimation, 445
 - plastic working, 423
 - ring test, 431
 - section rolling and forging, 417
 - stress-strain curves, 431
 - transverse deformation, 417
 - upper-bound technique, 437
- Machine Tool Dynamics Session, 235**
- circular saws, 289
 - dynamic characteristics, 237
 - elastic wave transmission and reflection, 267
 - epoxy resin bonded structures, 237
 - instability in milling, 259
 - machine tool receptances, on line identification, 245
 - self excited vibrations, 289
 - tool wear and vibrations, 251
- Machine Tool Elements Session, 153**
- electro-hydraulic feed drives, 165
 - grinding wheel spindles, 155
 - hydrostatic journal bearings, 175
- Metal Cutting Session, 185**
- bending and torque in milling, 209
 - coolants and cutting temperature, 187
 - drill geometry, 227
 - orthopaedic surgery, 227
 - power hacksawing, 217
 - tool wear and tool design, 195
 - ultra-high-speed machining, 203
- Noise and its Reduction Session, 287**
- HERF cropping machine, 295
 - impact forming machine structure, 289
 - power hacksaws, 305
- Rolling Session, 495**
- ball rolling, 497
 - cold rolling, 507
 - ring gears, 507
 - S-mill rolling, 515