AD-785 968

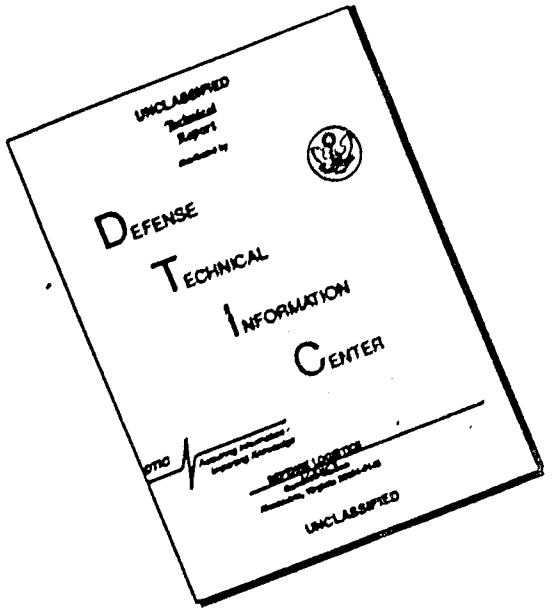
PROCEEDINGS OF THE CONFERENCE ON MATRIX METHODS IN STRUCTURAL MECHANICS (3rd) HELD AT WRIGHT-PATTERSON AIR FORCE BASE, OHIO ON 19-21 OCTOBER 1971

R. M. Bader, et al

Air Force Flight Dynamics Laboratory
Wright-Patterson Air Force Base, Ohio

December 1973

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

AFFDL-TR-71-160

PROCEEDINGS OF THE THIRD CONFERENCE ON MATRIX METHODS IN STRUCTURAL MECHANICS

PROCEEDINGS OF THE CONFERENCE HELD AT
WRIGHT-PATTERSON AIR FORCE BASE, OHIO
19-21 OCTOBER 1971

DECEMBER 1973

AIR FORCE FLIGHT DYNAMICS LABORATORY
AIR FORCE SYSTEMS COMMAND
AND
AIR FORCE INSTITUTE OF TECHNOLOGY
AIR UNIVERSITY

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD785 968

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFFDL-TR-71-160	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) PROCEEDINGS OF THE THIRD CONFERENCE ON MATRIX METHODS IN STRUCTURAL MECHANICS		5. TYPE OF REPORT & PERIOD COVERED Conference Proceedings 19-21 October 1971 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Flight Dynamics Laboratory/FBR Wright-Patterson Air Force Base, Ohio 45433		12. REPORT DATE December 1973 13. NUMBER OF PAGES ~~1062~~ 1021
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Structural Analysis Methods, Finite Element Methods, Discrete Element Methods, Matrix Structural Analysis, Structural Dynamics, Structural Optimization, Design Procedures, Computer Graphics

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Air Force Third Conference on Matrix Methods in Structural Mechanics, held at Wright-Patterson Air Force Base, Ohio on 19–21 October 1971, was sponsored jointly by the Air Force Flight Dynamics Laboratory, Air Force Systems Command, and the Air Force Institute of Technology, Air University. The purpose of the conference was to discuss the recent developments in the field of matrix methods of structural analysis and design of aerospace vehicles. The thirty-four papers presented were arranged into nine topic areas: Design Procedures, General Methods, Finite Elements, Dynamics,

1021

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

BLOCK 20 (CONT)

Optimization, Structural Applications, Computer Graphics, Non-Linear Effects, and Non-Structural Applications.

AFFDL-TR-71-160

PROCEEDINGS OF THE THIRD CONFERENCE ON MATRIX METHODS IN STRUCTURAL MECHANICS

TECHNICAL EDITORS:

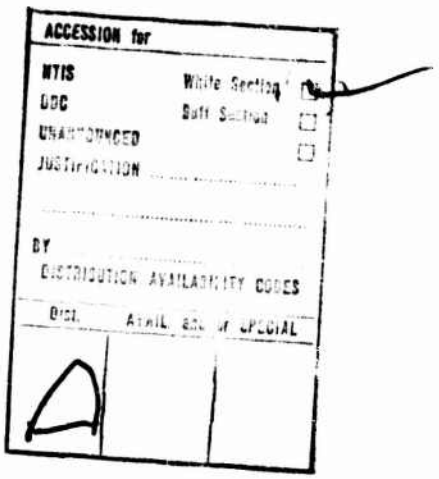
R. M. BADER
L. BERKE
R. O. MEITZ, LT COL, USAF
W. J. MYKYTOW
J. S. PRZEMIENIECKI
M. H. SHIRK

Approved for public release; distribution unlimited.

ib

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.



Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

AIR FORCE/56780/10 September 1974 — 1,100

ic

PREFACE

The Third Conference on Matrix Methods in Structural Mechanics, sponsored by the Air Force Institute of Technology (AFIT), Air University, and the Air Force Flight Dynamics Laboratory (AFFDL), Air Force Systems Command, was held on 19-21 October 1971 in the auditorium of the Air Force Museum at Wright-Patterson Air Force Base, Ohio. The purpose of the conference was to discuss the recent developments in matrix structural analysis and design of structural systems. This volume contains all the papers presented at the conference.

Three hundred and seventy-five scientists and engineers were in attendance including representation from the following foreign countries: Belgium, Canada, Germany, Holland, India, Ireland, Japan, Norway, Portugal, and the United Kingdom.

Thirty-four papers were presented at the conference in major topic areas entitled: Design Procedures, General Methods, Finite Elements, Dynamics, Structural Optimization, Structural Applications, Computer Graphics, Non-Linear Effects and Non-Structural Applications. The papers covered practically all major aspects of recent research and development work on matrix methods in structural mechanics.

The members of the Technical Committee for the conference, who also served as the Technical Editors, express their appreciation to the authors and session chairmen for their contributions toward the success of the conference, and to all those who attended and participated in the discussions. Special thanks are extended to the Director of the Air Force Museum, Colonel B. S. Bass and to the members of his staff. The Air Force Museum was a most appropriate setting for this conference and the staff's unselfish assistance in helping with the arrangements added greatly to the success of the conference. Special appreciation is extended to Lieutenant General James T. Stewart, Commander of the Aeronautical Systems Division for his excellent keynote address. Wholehearted thanks are offered to Mr. Martin Goland, President, Southwest Research Institute and President, American Institute of Aeronautics and Astronautics, for the outstanding address he presented at the conference banquet.

Grateful acknowledgement is made to Mr. John W. Thomas (now retired) and to Mrs. Helen Maxwell for their editorial efforts in preparing these proceedings for publication, and for the typing assistance of Miss Damaris Frantz and Miss Donna Schuh.

December 1973
Wright-Patterson Air Force Base
Dayton, Ohio 45433

Technical Editors

CONFERENCE COCHAIRMEN

Dr. J. S. Przemieniecki
Dean, School of Engineering
Air Force Institute of Technology

and Colonel C. A. Scolatti
Commander, Air Force Flight Dynamics Laboratory

TECHNICAL COMMITTEE

R. M. Bader (Chairman)
L. Berke
Lt Col R. O. Meitz
W. J. Mykytow
J. S. Przemieniecki
M. H. Shirk

CONTENTS

*now Lt Col

THE AIR FORCE THIRD CONFERENCE ON MATRIX METHODS IN STRUCTURAL MECHANICS

OPENING REMARKS: R. M. Bader
Air Force Flight Dynamics Laboratory

The conference was opened by Mr. R. M. Bader, Chairman of the Conference Technical Committee. Mr. Bader expressed his appreciation to the other members of the technical committee for their efforts in organizing the conference. The Cochairmen of the Conference were introduced: Colonel Charles A. Scolatti, Commander, Air Force Flight Dynamics Laboratory and Dr. J. S. Przemieniecki, Dean, School of Engineering, Air Force Institute of Technology.

WELCOME: Colonel B. S. Bass
Director, Air Force Museum

Colonel B. S. Bass, Director of the Air Force Museum was introduced by Mr. Bader. Colonel Bass welcomed the attendees as being members of the first technical conference to be held in the new Air Force Museum. He described the history of the Air Force Museum, the events leading to the construction of the new facility and the continuing search for items of historical interest to the Air Force. The museum had been recently dedicated in September 1971 with President Nixon as the speaker and honored guest. Colonel Bass then introduced Major General Pinson.

INTRODUCTION OF KEYNOTE SPEAKER: Major General E. A. Pinson
Commandant, Air Force Institute of Technology

The keynote address during the opening ceremony was delivered by Lieutenant General James T. Stewart, Commander of the Aeronautical Systems Division, Air Force Systems Command. This distinguished guest speaker was introduced by Major General E. A. Pinson, now retired, Commandant of the Air Force Institute of Technology. General Pinson was associated with research activities throughout his career and has participated in the three conferences on Matrix Methods in Structural Mechanics, held at Wright-Patterson Air Force Base.

KEYNOTE ADDRESS

Lieutenant General James T. Stewart
Commander, Aeronautical Systems Division
Air Force Systems Command

I am particularly pleased to be here. As Colonel Haviland, Director of Airframe Subsystems Engineering in the Aeronautical Systems Division (ASD), and his people, the SPO's in ASD, and the Air Force Flight Dynamics and Materials Laboratories well know, I am vitally interested in all facets of the structures discipline. My interest is shared by many people in the Air Force. We must follow the developments in this entire area including materials selection, design, analysis, test, manufacturing, fatigue life, and fracture characteristics. To do otherwise is to jeopardize the structural integrity of our future systems. Contrary to what some may think and contrary to some newspaper reports in the over-all, our track record is pretty good.

For example, over the last twenty years, the structural weight fraction of aeronautical systems has stayed fairly constant at about twenty-six percent. Uninformed critics would rush to point out that it should have decreased. That viewpoint is wrong. It is wrong because during the same period, the performance of systems has improved substantially. In effect, a measure of performance increases has been realized through optimization of the structure. How have we done this?

Speaking only to the subject of structural improvements, the last twenty years have brought out the use of higher strength materials. And although some of the lessons have been tough (to wit, the F-111), the pluses have far outweighed the temporary setbacks. And we have been working the materials harder. -- e.g., lower ultimate/working margins due largely to more sophisticated analytical methods. We could do this because our analytical techniques have generally become more precise and dependable. And here again, we have not always been perfect -- to wit, the C-5.

The advent of the large digital computer permitted us to exploit the finite element approach. Fine grid analyses can now be accomplished to a level of structural fidelity which was impossible only a few years ago. And of course you Gentlemen are more well aware than I of the advantages of the finite element approach.

The time required to respond to design changes can be reduced considerably with the use of these automated methods. Applying these methods earlier in the design cycle should not only result in a more efficient structure but, very probably, major design problems can be avoided.

This group generally represents two subsets of the technical community. Some of you are researchers -- from the Laboratory environment; others are engineers -- the users, the system people. I believe these groups must strive even harder to work closely together and build the strongest possible relationship between disciplines. It should be based on a reversible communication process in which the researcher develops more precise analytical techniques, passes them to the user and then the user develops a portfolio of applications experience for the researcher. This sort of dialogue, no one can fault. Everyone wins.

We have such a relationship here at Wright-Patterson to a degree, and we are doing everything we can to expand the information exchange. For example, we have physically located system engineers in local Air Force Laboratories, as well as at NASA Ames, and NASA Langley. Our motive is to expose them to the research environment for a period of about one year and then bring them back to the system acquisition world with fresh knowledge of current methods and techniques.

Also, the Laboratory people are assisting us solve aircraft development problems we have been experiencing, as well as actively working with us to apply their work more quickly to new systems.

Equally important is the need within the structures community for cross-talk between the other elements of the discipline. Designers should talk to the static and fatigue test people, the materials people should exchange information on NDI and manufacturing processes with everyone throughout the life cycle of the system. Fatigue failures and corrosion problems should be brought to the attention of all the engineers who worked on that system. That's the way it should work.

However, it doesn't work as well as it should. For example, in the past, some aircraft, or portions of it, have been designed by job shop engineers -- that is, people hired to do a specific task in a given time. They have no company affiliation or loyalty. When the job is over, they go somewhere else. The kind of feedback to the original designers I'm talking about is completely non-existent here.

Take another case, the company with an organic design group. On the average it gets a contract to design an airplane once every five years. After five years, where are the structural design people who worked on Airplane A? Most of the good ones have stayed with the company and moved up or into new projects. Few of them are willing to go back to the boards to work on Airplane B. So the wheel has to be reinvented again. It isn't practical to expect that the design supervisors will catch all the glitches. They can't unless they go into great detail and actually go back to the boards for a while.

So, the need for dedicated design groups working on new ideas all the time, not just every four or five years, has brought us to an appreciation of the experimental prototype concept about which you have heard much. Experimental prototypes can be used to explore promising theories or laboratory findings and their operational utility. They can bridge the gap between theory and application.

We have recently established a Prototype Program Office at ASD. It will establish technical goals, monitor the technical progress of contractors, provide technical assistance, and test and evaluate the end product. Industry, under this "adaptive management" approach will have the responsibility of developing the technical approach and establishing trade-offs, establishing management controls and design standards, and, most importantly, of demonstrating performance in flight.

The Prototype Program Office will have a Project Manager and Project Engineer for each prototype program. They will be supported as needed by designated lead engineers in the AFSC Laboratories and other ASD elements at Wright-Patterson. The key ingredients are small size and solid technical expertise from all appropriate sources.

The Air Force has identified four candidate projects for prototype development initiation in FY72: An advanced medium short take-off and landing (STOL) transport; a very low radar cross section test vehicle; a light-weight fighter aircraft; and a quiet aircraft.

It should be clearly understood that these projects are not aimed at either a pure research vehicle -- a la X-15 -- or an engineering development production prototype. They will be somewhere in between.

Their objective will be to provide options -- to demonstrate advanced technology and probable operational utility before proceeding into full-scale engineering development. The current A–X program is a conservative cut at prototyping.

All in all, we think prototyping is good. It's going to give a much larger degree of technical flexibility in the acquisition of hard data -- including structural performance -- before the big go/no-go decisions.

I think that you can see the dynamics of change in our ideas. For the future, we must avoid the past structural and management problems. One way to do this from the technical side is to increase cross talk between epople in the structural community. That's the name of the game at this conference.

Let me leave you with one thought. In the development of new techniques and exotic solutions to very sophisticated problems, keep in mind that solutions need to be practical (e.g., producible), cost-effective, and most of all, not create bigger problems than they solve..... In this latter regard, I could regale you for hours with almost unbelievable examples of problem solving via violation of structural fundamentals proven sound for at least thirty-five years.

CONFERENCE BANQUET

A conference banquet was held at the Officers Club, Wright-Patterson Air Force Base, on the evening of 19 October 1971. The guest speaker was Martin Goland, President, Southwest Research Institute, San Antonio, Texas, who at that time was the President of the American Institute of Aeronautics and Astronautics. Mr. Goland was a most appropriate speaker for the occasion because of his activity in several scientific advisory groups at the national level and his broad experience in aircraft design, applied mechanics and operations research. His address was entitled "The Engineering Challenge in the Post Industrial Era." He was introduced by the toastmaster for the evening, Dr. J. S. Przemieniecki, Dean, School of Engineering, Air Force Institute of Technology.

SESSION 1. DESIGN PROCEDURES AND GENERAL METHODS 1

Session Chairman

G. W. Brooks

NASA Langley Research Center
Hampton, Virginia

THE ROLE OF DESIGN ANALYSIS SYSTEMS FOR AEROSPACE STRUCTURES, AND FUTURE TRENDS

W. Lansing*

Grumman Aerospace Corporation

This paper attempts to lend some insight into the role of the design analyst during the evolution of an aerospace structural design, and examines the analysis tools available to him for doing his job. Of all the constraints to which he is subjected, the time schedule is likely to be dominant. This and other restrictions currently dictate that simplifying assumptions be made in carrying out his work. These are discussed and it is shown that the entire stream of analyses that must be employed, including those needed for obtaining applied loads, must be closely coordinated and integrated, in order to meet required schedules. Related comments are made also about structural optimization as it is currently being practiced in design, and trends for both it and integrated design analysis systems are indicated.

* Chief, Structural Mechanics Section

INTRODUCTION

We are all painfully aware of how the aerospace industry has fallen upon hard times. Employment is way down, the prices of our products are soaring, and our customers apparently are quite unsympathetic. The result is that our companies are asking their people, with much leaner technical staffs, to turn out better products than ever before, and at lower cost.

There are a number of areas related to the structures part of the business in which we can look for improvements. Apparently, there is a general consensus that the most promising of these areas are in the development and use of new materials, in better manufacturing methods, and in reduction of maintenance costs.(1)

But the design analysts can help too. They also play a vital role, specifically in their support of the designers during the evolution of a new vehicle. The key question is: are they able to perform their function properly, which is to provide structural evaluations sufficiently early that the right design decisions can be made? And an equally important, related question - do those of us who are developing methods for the design analysts to use really understand what their requirements are? One sometimes gets the depressing feeling that some of the best talent we have is misguidedly working on projects that are believed to be relevant to the design process, but really are not. Or if relevant, are peripheral rather than central to the main problem.

This is not to say that research and exploratory development work is not required. It most certainly is required. It adds needed depth to our understanding of basic phenomena, and in some cases obviously leads to methods that are useful in design. But the objectives of methods development work on the one hand, and research and exploratory development on the other, should not be confused. They are very different.

In an attempt to clarify this, I am going to try to lend some insight into the role of the design analyst during the evolution of an aerospace structural design. In so doing, naturally, we shall have to examine the tools that our design analyst has available to do his job, and how they might be improved.

In discussing the design process and the nature of the support that the design analyst provides, it is very difficult to generalize. Therefore, while attempting to get my message across, I shall have to refer extensively to examples. Since my experience is with Grumman hardware, this will be a major focal point. However, I believe that our experiences are broadly representative of those of the entire industry, and so should be of general interest.

A key point to bear in mind in reviewing a company's approach to the design process is that there are several external constraints that have a vital influence. One is the time available for the design work, i.e., the period from contract go ahead to first flight; this can vary greatly from one vehicle procurement to

the next. Another constraint is the design analysis philosophy of the customer. Our military services, of course, require final analyses of the airframes they buy, but they place greater emphasis upon satisfactory completion of static and fatigue tests of entire vehicles for proof of design adequacy. The airlines and FAA, on the other hand, rely much more heavily upon final detailed analyses for the large commercial transports, and less upon large scale testing. These factors are bound to make a difference in how a given manufacturer goes about his analytical work.

The design of an airframe structure is an exceedingly complex procedure. Early in the design, rather dramatic alternatives in the primary structure configuration, materials and methods of construction must be considered. Later on, the changes to be investigated are more those of detail. The earlier configuration changes, which are obviously of special importance to the analyst, are sometimes made solely for the purpose of increasing structural efficiency. More often, however, the motivation is likely to be a mixture of other things as well - an attempt to improve aerodynamic performance, a rearrangement of equipment and accessory locations for functional reasons, etc.

The applied loads acting upon the structure can also change significantly. In the case of a new vehicle whose external lines and performance represent real departures from anything already in existence, the preliminary applied loads which must be used initially are subject to sizable redistributions, as well as changes in magnitude, as more information (such as wind tunnel pressure distribution test data) becomes available. These changes can have a major impact upon the vehicle design.

During the evolution of such a structure, a fundamental responsibility of the analyst is to provide the designer with evaluations of his designs. Does the configuration currently being considered use the structural material efficiently? Are there undesirable discontinuities in strength or stiffness? Are there severe structural dynamics problems? This information must be available at the right time, so that the right decisions can be made as the design evolves. Since the major decisions that largely determine the success of the vehicle are made early in the design phase, this is where the analyst too must make a major contribution. Of course, he must also be responsible for the integrity of the structure as it is finally built.

It should be noted that in real life, the design analyst usually does more than just this. Ordinarily, he participates actively in the design function as well - resizing structure, making suggestions on changes in configuration, etc. Nevertheless, his basic responsibilites are primarily in performing the analysis tasks.

The aerospace industry is well aware of the nature of these tasks. And it is obviously looking for improvements, as evidenced by the activity in such fields as integrated design and analysis systems, structural optimization and computer graphics.

For its part, Grumman four years ago conducted a major study of the stream of analysis and optimization tools required to support its design activities, with primary attention being given to the time frame for their application. Deficiencies were identified, and a plan for remedial action was formulated. Initially, its implementation required a major effort to bring about certain basic changes in our procedures. Thereafter, our activities have been, and will continue to be, directed toward making evolutionary changes to the existing system, in the order in which improvements will be most cost effective. As might be expected, the Grumman approach is not necessarily identical in its philosophy to that of other members of the aerospace community.

This paper seeks first to describe the highlights of the study just mentioned and of the resulting analysis system as it now exists. Thereafter, it touches upon how the system has been used to contribute to the design of the new Navy fighter, the F-14. Comments are also made concerning its application to Space Shuttle design studies.

From this base we will take a brief look at other automated analysis methods development activities throughout the industry as they apply to design. Finally, I will attempt to identify the more critical needs of the next few years.

STUDY OF ANALYSIS NEEDS

Approximately eighteen months prior to the date that it was believed the Navy would be issuing an RFP for the F-14, Grumman management became convinced that the aerospace industry was headed for some major new challenges in the development of future military aircraft. In particular, for the F-14, these challenges were associated primarily with the drastically shortened schedules that were contemplated for its design, manufacture and flight testing. As a result, studies were made in all these areas to see what could be done to speed things up.

One of the studies concerned the analysis methods we used in determining the internal loads needed for sizing primary structure. This is obviously a crucial design requirement. The time period particularly emphasized was from the vehicles's contract go-ahead to its first flight. At the same time, we realized that methods suitable for this period would also be applicable for preliminary design, if they were streamlined and used properly.

As we examined our ability to quickly perform the required analyses, it immediately became clear that many technical groups and disciplines were involved. To obtain internal loads in primary structure, far more was involved than merely performing an adequate finite element analysis. If anything, the finite element work was the least of our troubles. The real pacing item turned out to be the determination of the applied loads acting upon the structure.

When one includes in the chain of required calculations all of the applied loads critical for design, he must draw upon many technical areas - aerodynamics, weights. dynamic analysis, loads and criteria, structural analysis, contour development, structural design, etc. All of these are needed to determine the loads due to flight maneuvers, gusts, landing, catapulting and taxiing. In addition, there are the closely related flutter analyses. In all cases, structural flexibility effects must be considered to some degree in the final determination of the applied loads for which the structure is designed. The whole process of ultimately determining internal loads is surprisingly complex, and there were actually very few people in the Grumman organization who understood the entire picture, except in a general way.

Personnel in each technical specialty had many analysis procedures, and many of these had been programmed for computer solution. The primary difficulty was that such computer programs had tended to grow in an uncoordinated fashion in each technical area. Some programs were not general enough and were limited to specific classes of airplanes. Some overall tasks within a given technical area had been programmed in pieces that were not properly combined for rapid execution. In general, output data from one technical group was not in a form suitable for use by some other technical group downstream. Instead, all too often, the "answers" from one group required a large amount of hand manipulation before they could truly be used as input data for the work of some other group.

In brief, the crucial problem was not one of lack of basic capability in any narrow technical area - it was the blending of all these skills into a properly balanced, coordinated systems approach. Frankly, this general situation also appeared to exist throughout the rest of the aerospace industry.

In the preceding discussion, no mention has been made of structural optimization. Instead, we have spoken only of "determining the internal loads needed for sizing primary structure." One might imply from this that structures engineers are unaware of the interaction which exists between changes in the distribution of material throughout the structure and the internal loads. Of course, this isn't true. Actually, as a matter of necessity, we have learned by experience how to make changes in material distribution in order to bring stresses reasonably into line with allowables, while at the same time making allowances for the accompanying changes in the internal loads. Whenever possible, revised analyses are then run as a check, but there is usually a limit to how much of this can be done before the drawings are signed and manufacture begins.

Under the preceding circumstances, it becomes quite clear why stress analysts are so much interested in automated methods for obtaining fully stressed designs. By this we mean idealized structures for which every element is either stressed to its allowable value for at least one design condition, or is at a specified minimum size. Historically, for strength critical structures, fully stressed designs are exactly what we have been striving for all along. They may or may not be truly minimum weight, but the capability for rapidly obtaining fully stressed designs of idealized airframe structures should be recognized as being a really significant step forward. We will have more to say about this later.

We were also aware of the necessity of using graphic display devices, wherever economically feasible, for visually checking input data for the computer, and displaying results. As for the interactive devices that have been emerging over the past few years, their potential was also very attractive, especially for smaller structural analysis applications. However, the benefits to be gained by their use appeared to be less crucial than the alleviation of our more basic problems. Accordingly, less attention has been given to them than some of the other matters to be discussed next.

INTEGRATED DESIGN AND ANALYSIS SYSTEM (IDEAS)

We began to put our house in order. Approximately 50 men from the applicable engineering groups worked as a team for about a year to establish the present analysis system. It was a high priority effort in regard to selection of the best qualified personnel and in regard to computer availability. The result is called IDEAS. The system has been made compatible with Grumman's computing facilities which embody the IBM 360/75 computer.

Basically, IDEAS[2] is an integrated collection of many different kinds of computer program modules and formalized hand calculations. The sequencing and execution of these program modules allow one to plan, schedule, coordinate and control all of the analytical efforts of the several sections that make direct contributions to the determination of internal loads for the sizing of primary structure for an aerospace vehicle. It is the "systems approach" for the problem at hand.

System capability was planned from the start to be sufficiently broad to provide state of the art analysis procedures for both variable-sweep and fixed-wing subsonic and supersonic aircraft. As such, it treats all applicable flight maneuvers, landing and catapulting conditions, taxiing conditions, and gust and flutter analyses. A suitable starting point for application of the system can occur with a definition of the loading condition requirements and knowledge of external vehicle contours, including control surfaces and stores. Also needed are weight estimates of concentrated masses, fuel, and major structural components such as wing, fuselage, etc.; a configuration for the primary structure and landing and arresting gear; and initial estimates for the sizing of primary structure. Early in the design, such information will necessarily be very sketchy. For the later IDEAS cycles, the data will naturally become more definitive and detailed.

The technical groups associated with IDEAS use various mathematical models to perform their calculations. The establishment and standardization of these models in regard to sign conventions and coordinate systems is another key feature of the IDEAS system. With standardized models, it is easy to pass data between the various technical groups by means of suitable transformation matrices.

A main feature of the IDEAS effort is the establishment of a "Data Bank" which provides a central storage facility for all calculated data. This manually controlled Data Bank consists of stored magnetic tapes, cataloged listings, CALCOMP plots, Orthomat drawings, CDC-DD80 microfilm, punched cards and engineering drawings of mathematical models. Any given program draws its input data from the Data Bank and returns its output to it. There are approximately 50 major computer programs currently stored on the accompanying "IDEAS disc." Included among them are the Grumman developed programs for fully stressed design, of which "ASOP" (Automated Structural Optimization Program) is the most recent addition.[3]

Another feature of IDEAS is that it is open-ended and lends itself well to the incorporation of new analysis methods. For example, recently, because of the Space Shuttle and its thermal stress problems, a new structural temperature analysis capability has been formulated, and is being integrated into the system. It provides for coordination of all of the necessary calculations, beginning with trajectories and the accompanying aerodynamic heat transfer coefficients, and ending with transient temperature distributions in appropriate structural models. In combination with the proper coefficients of thermal expansion, these temperatures yield initial strains that are treated thereafter as load cases, along with the mechanically applied loads.

Interactive computer graphics is not yet a part of IDEAS. As has already been indicated, this is the result of deliberate management decision - an attempt to put first things first. We are now investigating ways to incorporate these tools into the system in a cost effective manner.

APPLICATION OF IDEAS TO F-14

As I have indicated, the first project to which we have applied the IDEAS system is the F-14. I shall try to give some of the highlights of the effort. In scope, it should be fairly representative of the analytical work required in support of any advanced airframe design.

As shown in Figures 1 and 2, the F-14 is a supersonic, carrier-based fighter with variable sweep wings, all movable stabilizer and twin vertical tails. It flew for the first time in December 1970, and is now engaged in its flight test program. Its most significant structural innovations are the wing center box, made of titanium, and the stabilizer, with its skins of boron/epoxy composite. The fuselage is also quite complex, being very broad and rather shallow. The way in which it carries overall bending is quite complicated, the bending structure being concentrated in the centerbody in the forward region, and in the widely spread engine nacelles aft, with a transition region in between.

Mathematical Models

The key to developing a feel for the scope of work covered is to look at the various mathematical models that have been employed. There are essentially four different types.

Air Loads
Weights
Dynamics
Structures

Actually, the dynamics and structures models have been changed a number of times during the course of the design. For simplicity, in all cases I shall concentrate primarily upon the models we used for release of structural drawings.

Let's start out by examining the air loads model for the fuselage, Figure 3. This shows the resultant air load distribution for a typical supersonic symmetrical pullout. For the final design iteration, the source of such distributions is wind tunnel tests performed on models with many pressure taps. Earlier in the design, loads are approximated by traditional estimating procedures. Shown is the form in which the results are summarized for application to the structure. Essentially, we have here three types of running loads, one applied down the centerline of the fuselage, one applied along the centerline of each engine inlet, and the third along a curved line, representing the loading on the fairing between the fuselage and the wing. Bear in mind that these running loads must ultimately be applied to a finite element model of the fuselage.

The wing air loads distribution model is shown in Figure 4. The unswept case is illustrated. There are four other models for various sweep positions. Notice that the model goes all the way into the airplane centerline. For the early design iterations, we treat the wing as an isolated surface, and calcu-

Figure 1

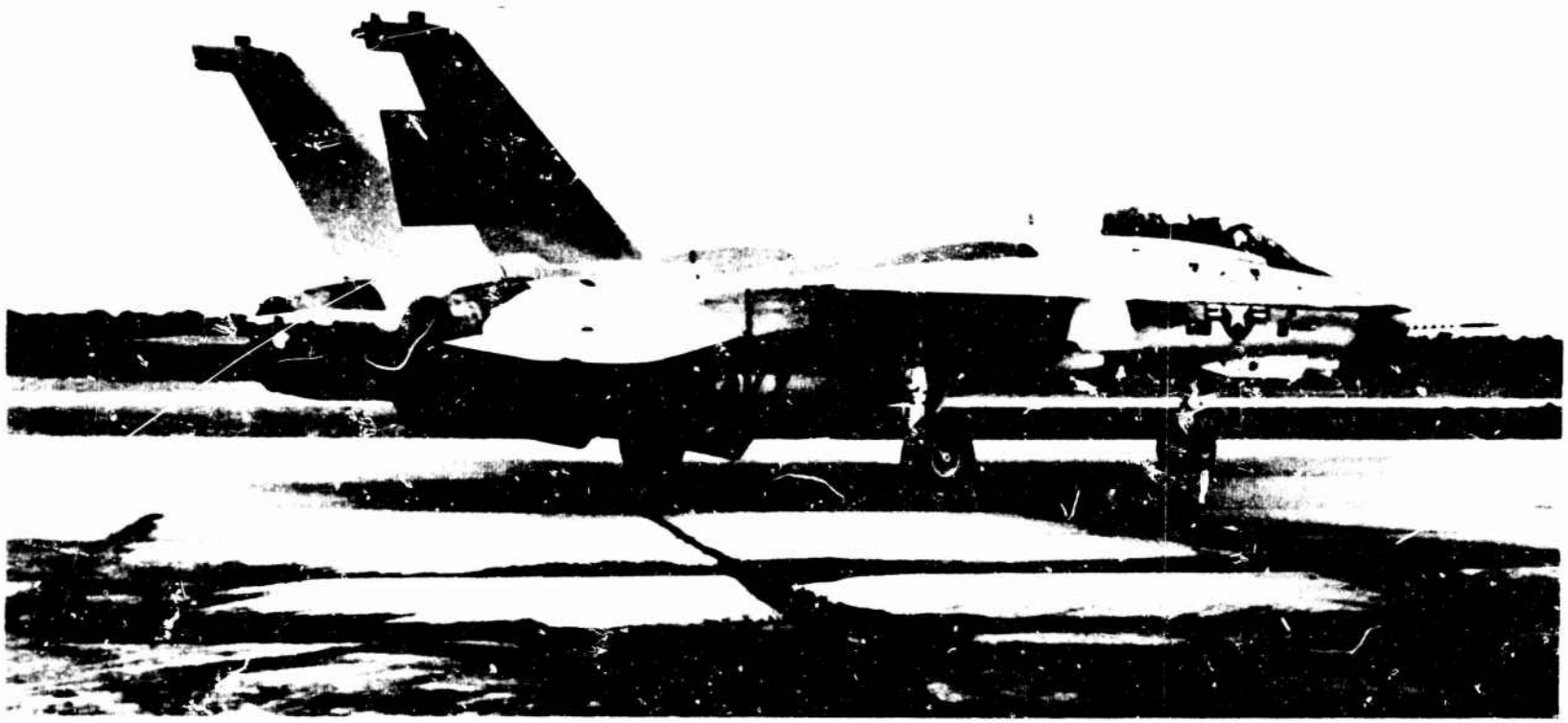


Figure 2

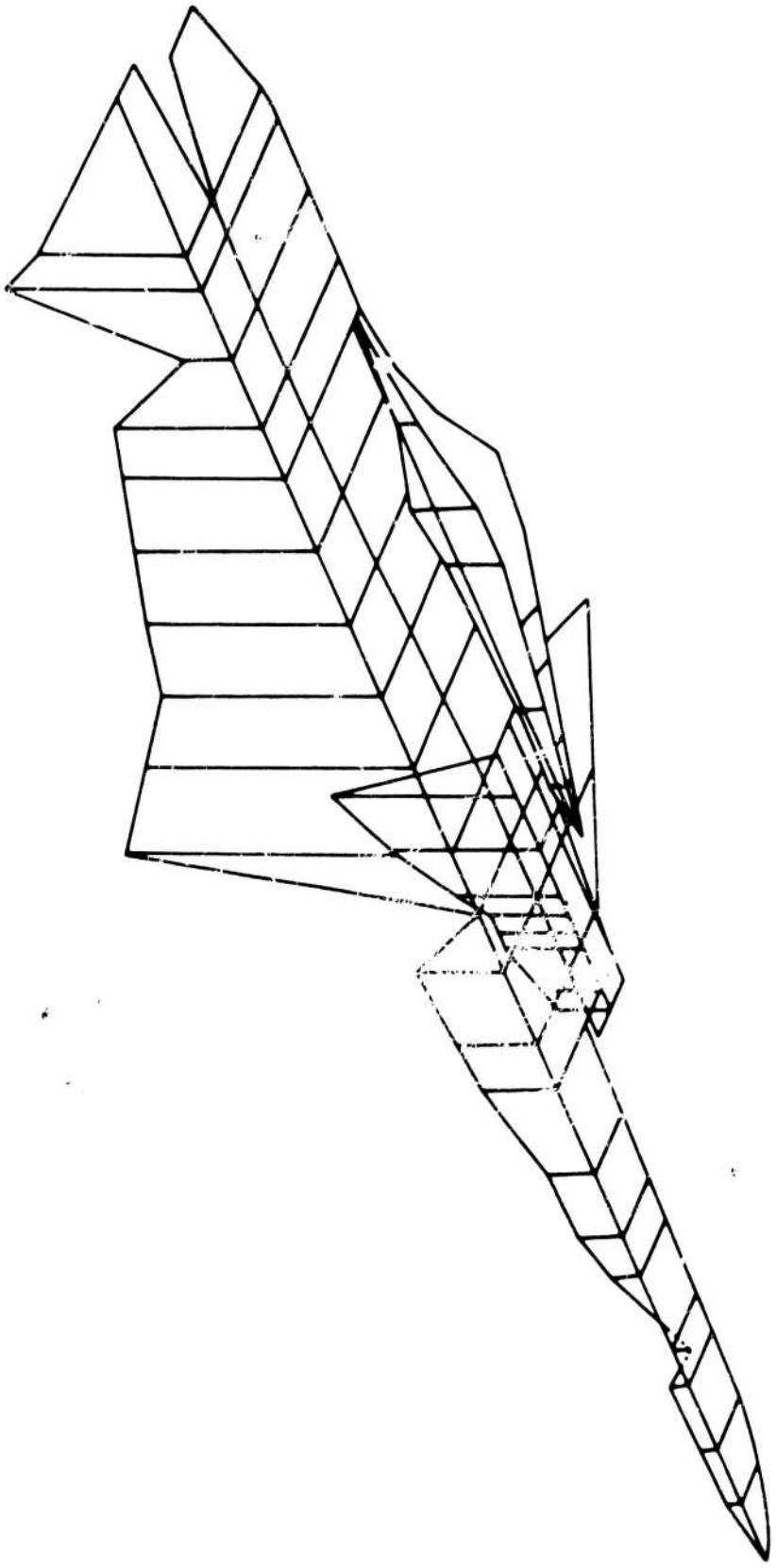


Figure 3 Fuselage Air Loads Model

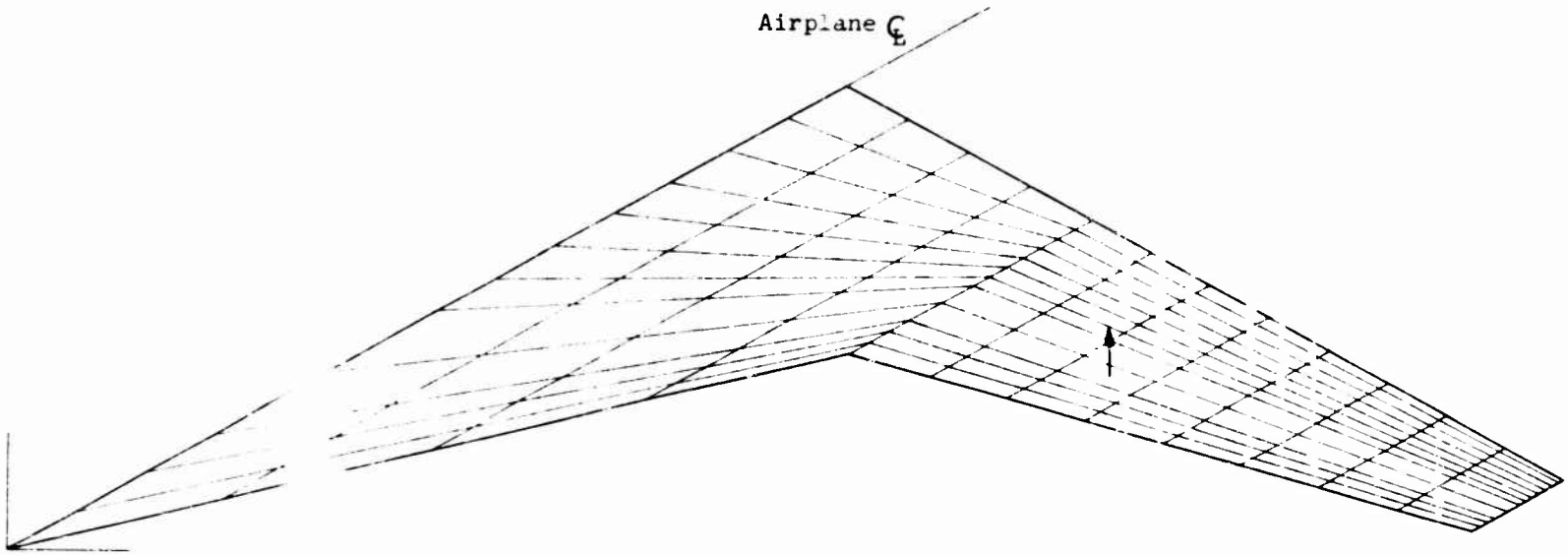


Figure 4 Wing Air Loads Model

late distributions of loads acting on all of the 144 quadrilateral areas shown. We then correct for the effects of the fuselage aerodynamics on a semi-empirical basis using overall experimental aerodynamic derivatives wherever possible. For the final iteration, wind tunnel pressure distribution data is used to obtain forces acting upon the portion of the quadrilaterals that represents exposed wing. Air load models for the stabilizer and the combined fin and rudder are similar in nature to the wing.

As for the weights models, the fuselage is shown in Figure 5, again in a configuration suitable for a flight maneuver. Note that two types of masses and moments of inertia are considered, one representative of distributed, fixed structure and fuel, and the other representative of large, concentrated items such as engines and external stores. Some of these items will of course vary from one design condition to the next. As in the case of the fuselage air loads, the inertia loads that go with these masses must be redistributed to the finite element model of the fuselage.

In the case of the wing and empennage structures, the weights people have used the same grids as employed for the air loads.

The third class of models is dynamics. Figure 6 shows the flutter model. The dots indicate lumped masses, and the arrows show the degrees of freedom considered. In the case of the curved arrows, for rotational degrees of freedom, moments of inertia and mass first moments are included as well. (Existence of the mass first moments means that, for the wing and for the fin and rudder, the mass cg's actually do not fall on straight lines, as seems to be indicated in the figure.) Some degrees of freedom are not shown; they include those to represent various combinations of external stores. When all of the stores are included, the number of degrees of freedom for the half airplane is 222 symmetric, and 253 antisymmetric. There is also a dynamics model for use in calculating landing and catapulting loads. This model is similar to that for flutter, except that certain simplifications have been made, as for instance, in the wing and empennage structures. On the other hand, the fuselage portion of the airframe is now modelled in more detail, and there are roughly twice as many degrees of freedom for this region as in the flutter model.

Of course, structural flexibilities are also required. These are essentially provided by the structures models to be described next.

As for the structures models, the most complex is that for the fuselage. Figure 7 shows one half of the principal fuselage structural model, including a wing center section which has been detailed sufficiently to account for its interaction with the fuselage. The model as shown employs about 3000 structural elements, including bars, beams, warped (non-planar) shear panels, warped membrane panels, and a few triangles. Figure 8 shows a portion of the idealization in greater detail. As in most analyses of this sort, the applied loads enter at the structural node points. Their values are obtained by assuming that the air loads and the inertia loads associated with the models described previously are distributed to the structural nodes by simple beaming action. There are some obvious gross assumptions involved in such procedures, but the process is consistent with the accuracy with which the applied loads themselves are known.

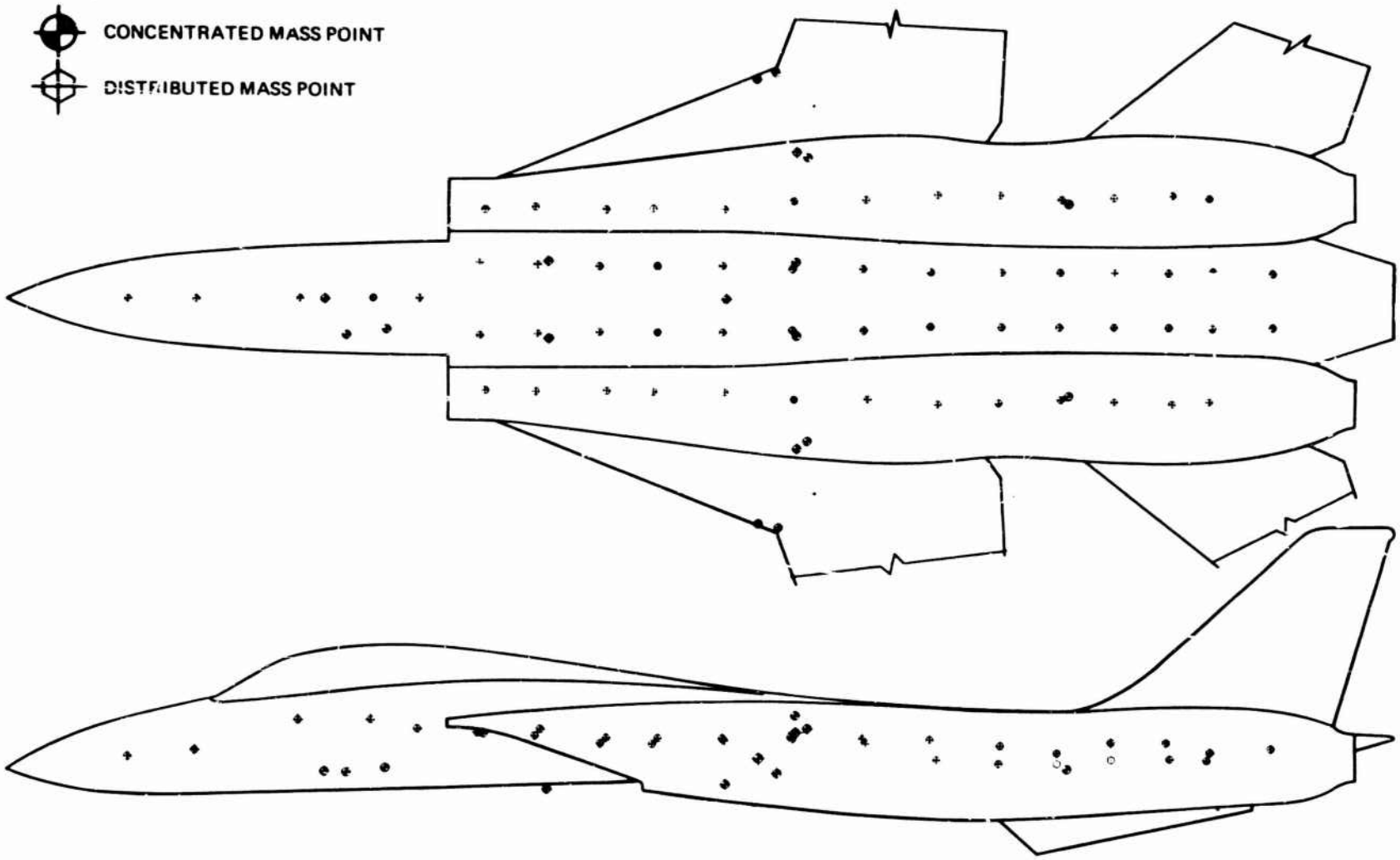


Figure 5 Fuselage Weights Model

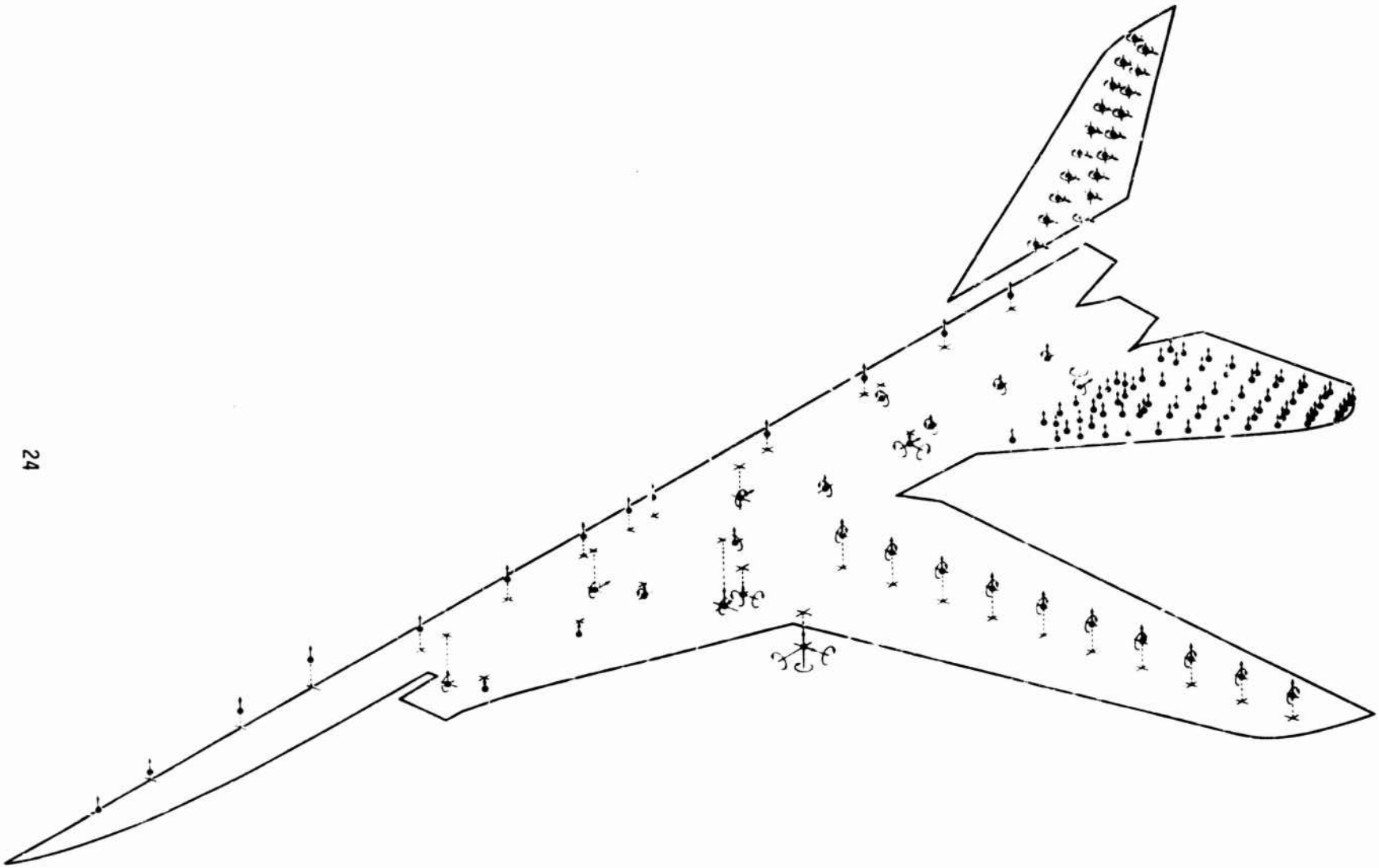


Figure 6 Flutter Model

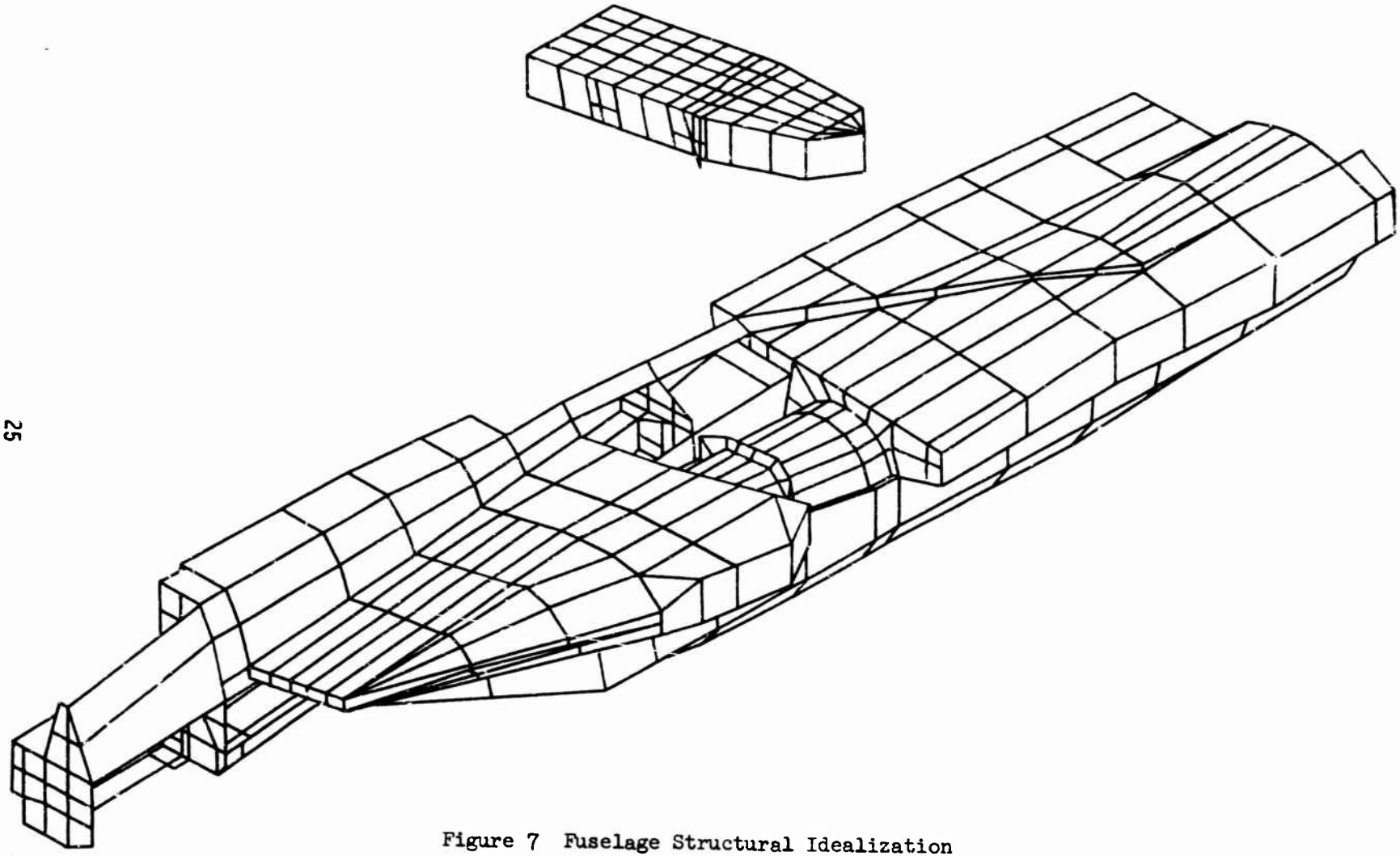


Figure 7 Fuselage Structural Idealization

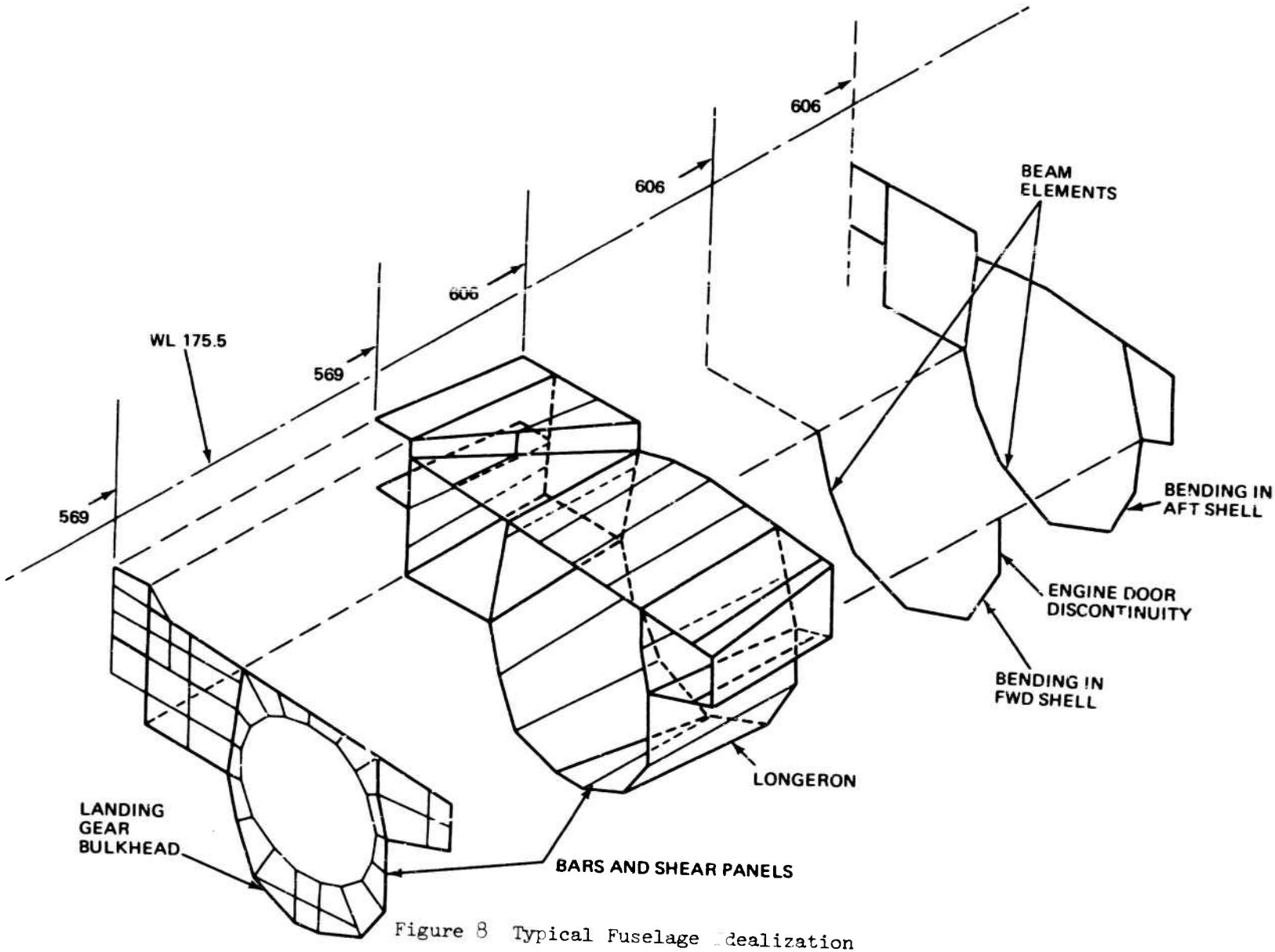


Figure 8 Typical Fuselage Idealization

The wing center section has also been idealized somewhat more elaborately in another model for its detail design. The wing outer panel is of high aspect ratio, and is conventional in its design except for the pivot region. Figure 9 shows its planform, and the region of the primary structure idealized by a finite element model. The remainder, including leading edge, tip and moveable surfaces, is analyzed by conventional methods based upon engineering beam theory.

The fin and rudder are also conventional, and are treated similarly to the wing outboard of the pivot. Again, a finite element model is used for only a portion of the structure, in this case, that adjacent to the fuselage attachments, because of departures from engineering beam theory in this region.

The stabilizer is a low aspect ratio surface which pivots about a shaft fixed to the fuselage. Beam theory is marginal in this case, and a finite element model for the entire primary structure would thus be preferred even if it were not for the boron fiber covers. The stabilizer structural arrangement and its corresponding finite element model are shown in Figures 10 and 11.

These models have been supplemented by many other finite element models which represent smaller regions of structure previously described but in greater detail. The wing pivot region is an obvious example. The principal fuselage bulkheads comprise additional cases.

Loads Calculations

The design conditions that are critical for the F-14 are primarily the flight maneuver conditions, and arrested landing and catapulting. The flight maneuver calculations are made originally with simplified airframe flexibility effects included. In later iterations, more refined aeroelastic corrections are introduced. From the calculated flight maneuver time histories, critical values of gross air and inertia loads are determined. As mentioned previously, five different wing sweep positions are covered in the loads work. Thus, the number of maneuver conditions to be considered is increased approximately by a factor of five, as compared to a fixed wing aircraft. These gross loads are then distributed to the structure in accordance with the air loads and weights models previously described by means of appropriate transformation matrices.

As for the arrested landing and catapulting conditions, here the critical loads are induced in the structure through the arresting hook and landing gear forces. (The catapulting loads in all of our modern Naval aircraft are applied through the nose gear.) Consequently, in carrying out the necessary time history calculations, the main and nose gears must be carefully modelled, because their nonlinear load-stroke and inertial characteristics enter into the determination of the loads in a fundamental way. Airframe flexibility effects are also significant, and are included as well.

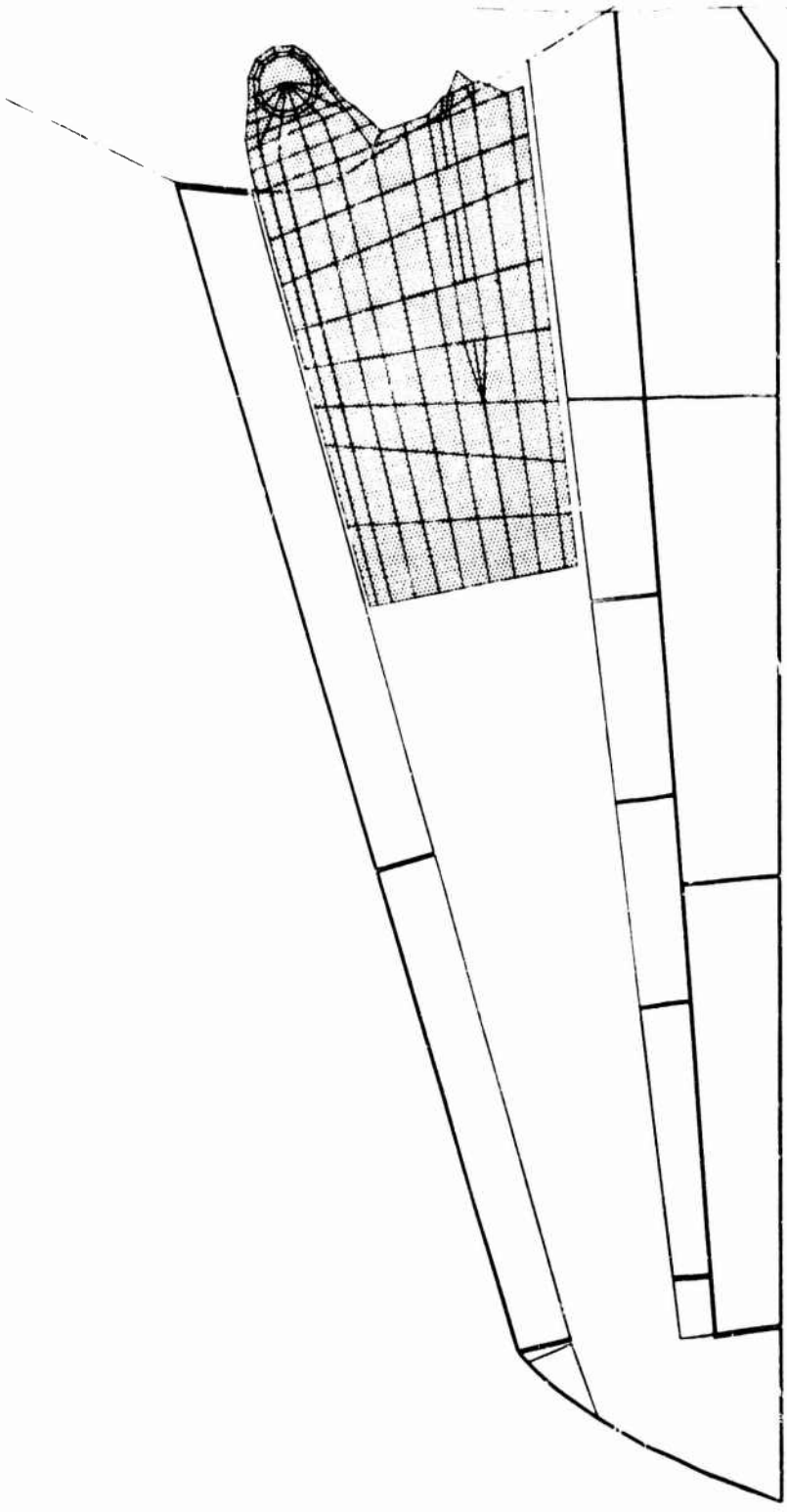


Figure 9 Idealized Wing Structure

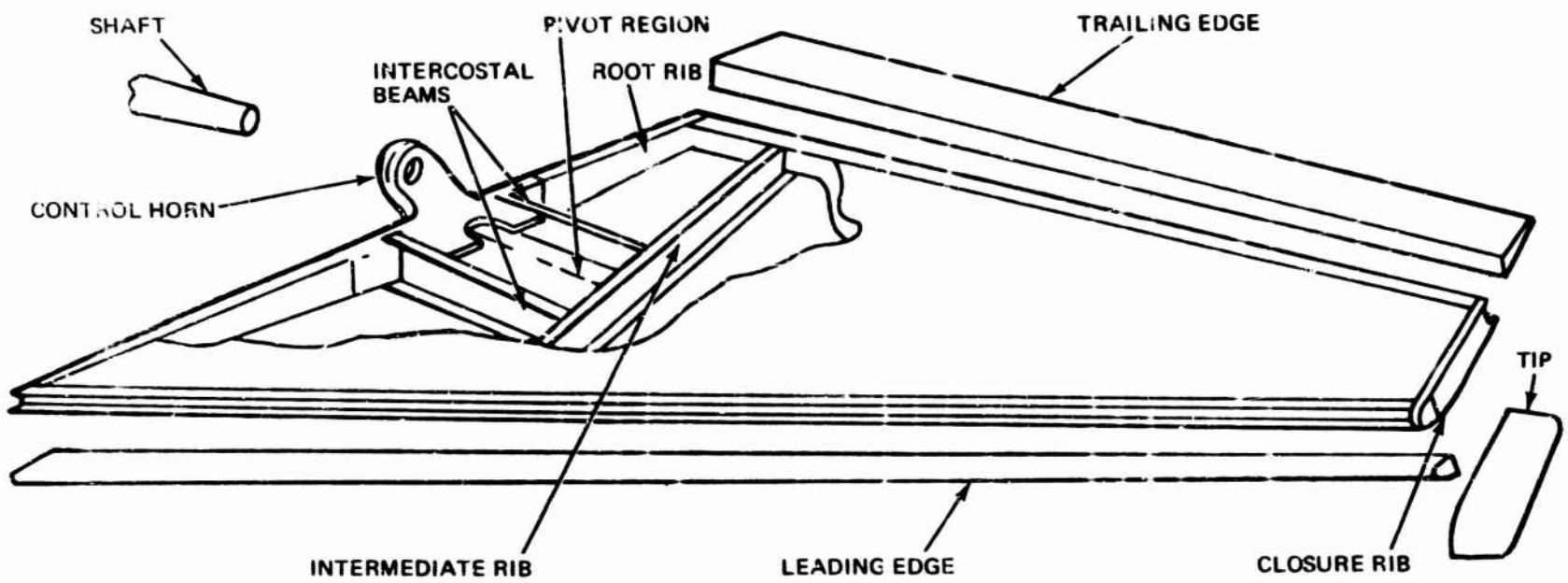


Figure 10 Exploded View of Stabilizer

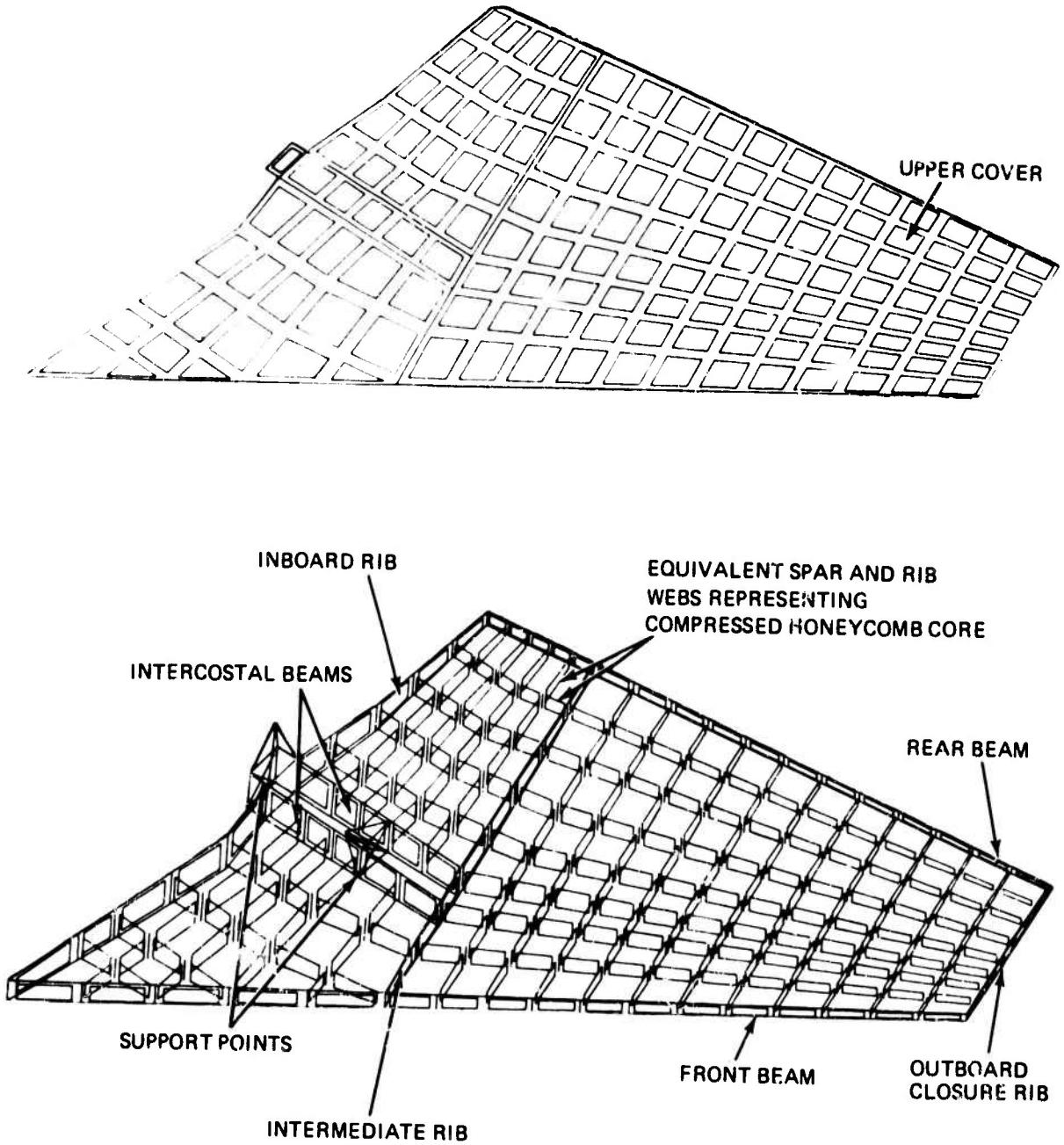


Figure 11 Stabilizer Idealized Structure

Analysis Iterations

As I have emphasized previously, structural design is an iterative procedure in which one of the analyst's tasks is to provide timely evaluations as the design progresses. In the case of the F-14, a review of the timing of the various analysis iterations that were done should be of interest. I have summarized these for the fuselage in Table 1.

As shown in this table, the time between the F-14 contract go-ahead and first flight was twenty three months, a very short period for such a sophisticated design. During this period, four major fuselage analyses were conducted. The first one was pretty crude, and actually began before contract go ahead. The next two were progressively more complex, and were completed prior to release of major structural drawings. The fourth was completed prior to first flight. All but the first of these analyses involved the 4 types of models I've been talking about -- loads, weights, dynamics, and structures -- and included all the steps that were needed to go from basic design requirements to internal structural loads.

Similar analysis iterations took place for the wing and vertical tail. As for the stabilizer, its design was fixed much earlier than would normally be the case, because of the use of the new material - the boron fiber/epoxy composite - in the covers. This was done in order to permit additional lead time for early completion of its static and fatigue testing, well before the aircraft's first flight.

The flutter analyses that were conducted proceeded along conventional lines. As is customary, the wing and empennage structures were first designed for strength and then checked for flutter. The checks made followed the structural design iterations, and increased in complexity until finally, the complete airplane model of Figure 6 was used. The result was a very small amount of weight added to the stabilizer early in the program for increased torsional stiffness. Very small weight increments were also required in the wing outer panel and in the upper region of the fin, and thus the F-14 has an almost negligible weight penalty for flutter.

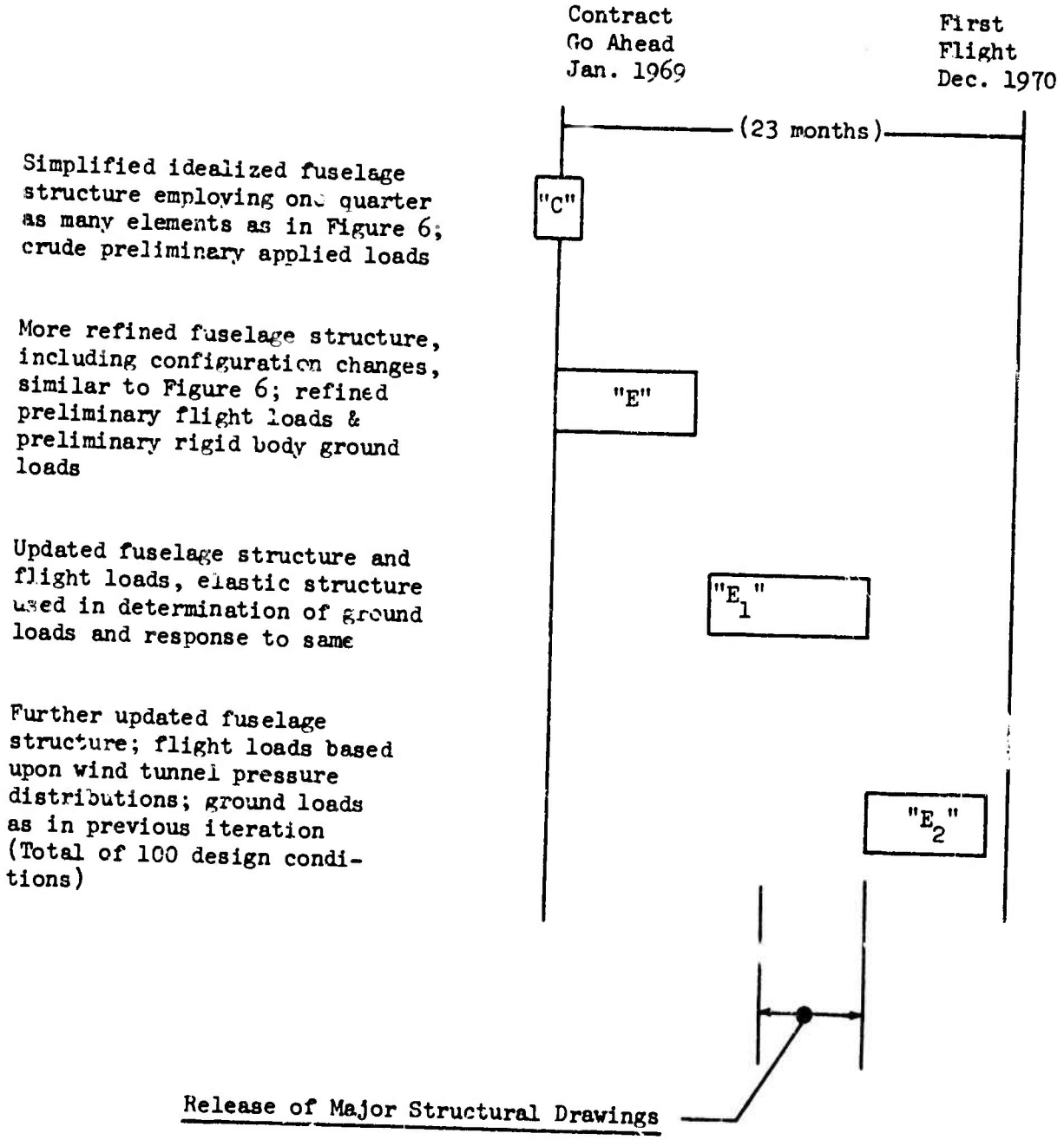
Use of Internal Loads Rather than Stresses

It may have been noticed that in discussing the role of the automated methods used in supporting the F-14 structural design, I have always referred to the goal as calculating the internal loads needed for sizing primary structure. I have specifically avoided any mention of stress distributions as such. There is a very important reason for this.

I have tried to emphasize, all along, how often the design will change during its more creative stages. It is therefore inevitable that the analyses will always lag behind; we just can't keep changing them instantaneously to keep up with the changes in design. Also, because of necessary simplifications, a given idealization may not have represented the structure precisely, even before any changes in design took place. Under these circumstances, the stress analyst frequently has to "make do" with somewhat obsolete analysis results. What he needs to do is put a portion of a large redundant structure,

TABLE 1

F-14A FUSELAGE ANALYSIS TIME LINE



Simplified idealized fuselage structure employing one quarter as many elements as in Figure 6; crude preliminary applied loads

More refined fuselage structure, including configuration changes, similar to Figure 6; refined preliminary flight loads & preliminary rigid body ground loads

Updated fuselage structure and flight loads, elastic structure used in determination of ground loads and response to same

Further updated fuselage structure; flight loads based upon wind tunnel pressure distributions; ground loads as in previous iteration (Total of 100 design conditions)

say a bulkhead, in equilibrium, and size it for the loads acting upon and within it. If he has internal loads, then even if the real structure is somewhat different from that for which the redundant analysis was performed, he can still make assumptions and go ahead with his work. However, if all he has are stresses, he's in deep trouble.

This reality of the design procedure is so important that at Grumman, we convert the results of our standard displacement analyses into the older force method form for use by our stress analysts. By this I mean results in the form of bars with axial force varying along the length, and shear panels subjected to shear flows only.[4] If we had not hit upon this expedient, Grumman might still be using the force method in support of structural design.

Preparation of Structural Models

No doubt the most creative task associated with any redundant structure analysis is the selection of the structural model. Unfortunately, it can also be one of the most laborious.

Early in a design, before the structural details have become established, one has available only the external lines of the vehicle. At this stage, the designer requests of the contour development people that appropriate sections through the fuselage or wing, or whatever, be drawn for his use. (Most of the aerospace companies, including Grumman, now have fairly extensive automated drafting systems for such purposes.) The analyst and the designer then rough out a primary structure configuration within these external contours. As for the member sizes, their selection can be a major task. They are chosen based upon whatever information is available or can be conveniently obtained - results of previous analyses, engineering beam theory calculations, pure estimates based upon past experience, etc. In this work, estimates of the applied loads for the more critical conditions obviously are necessary. And if one has an automated fully stressed design capability available, it can be tremendously helpful in sizing the members. One can simply start the fully stressed design calculations with unit areas or gages for all members.

Later on in the design, when structural layout drawings become available, new structural idealizations are developed. Even at this time, however, some departures from the actual structural configuration are usually made. Examples of this are the simplification of complex fastening details between major parts, "lumping" of stringers and frames, etc. Regardless of the degree of refinement of the model, it must be translated into nodal geometry, topology (i.e., how the members are interconnected), and member sizes, for submittal to the computer. On the F-14, this translation was done very laboriously by hand. The idealized structure in each case was then plotted by the "Orthomat," an automated drafting machine, to provide a visual check on the input data. This is certainly an area where further automation would appear to be most cost effective.

Graphical Display of Output

One of the F-14 fuselage analyses mentioned earlier consists of approximately 3000 elements for the half structure, and is subjected to 100 design conditions. When one considers that there are two bar end loads for each bar, four shear flows for each panel and up to twelve end loads and moments for each bending element, the sheer volume of the printed results is quite disheartening to the person who has to use them. To make it as easy as possible to use results such as these, we have done two things. First, we have listed the output so that it is as convenient as we can make it. Figure 12 shows the portion of fuselage between stations 569 and 606. Figure 13 shows the corresponding member loads by condition, while in Figure 14, the member loads are ranked by condition according to maximum positive and negative values. The other thing that we have done is to display the results on a DD-80, and make hard copies from its microfilm output. Figure 15 represents a fuselage station cut just aft of station 668. The trailers leading to a number identify cap loads, while the numbers written along the members represent shear flows. Sketches like Figure 15 were produced for all loading conditions and for cuts forward and aft of each bulkhead, as well as for the bulkheads themselves. As for displacements and distortions, Figure 16 shows these for the bulkhead at station 668. Similar DD-80 plots were made for all bulkheads and all design conditions. They were a big help in understanding the way in which the structure carries its loads. The graphical display of outputs such as these is very cost effective, and seems certain to be pushed everywhere.

Another application of computer graphics consisted of partial use of the Orthomat as an aid in plotting vibration mode shapes of the previously described dynamics models. As an example, Figure 17 shows a symmetric fuselage-bending mode. In this application, the Orthomat plotted only the original undeflected grid and the corresponding displacement vectors. As an expediency, the deflected grid was then drawn in by hand. Even so, it was a tremendous aid in plotting the 480 mode shapes that were required because of the variety of sweep angles and fuel conditions to be considered.

Structural Optimization of the F-14

The design iterations of which we have spoken, including the final sizing of primary structure, really refer to an attempt at structural optimization. It is the time honored engineering approach. How near the structure comes to optimum depends upon many things - first and foremost, the ingenuity of the engineers doing the work, but also upon the tools available to them and the time allotted for their application.

We have seen that in the case of the F-14 fuselage, both the idealized configuration and the applied loads varied from one model to the next. Nevertheless, to the best of their ability, the analysts made use of the information gained from each model to size the structural members of the next model downstream. The fully stressed design program that we had operational at this time did not include a bending element, which was essential in idealizing the

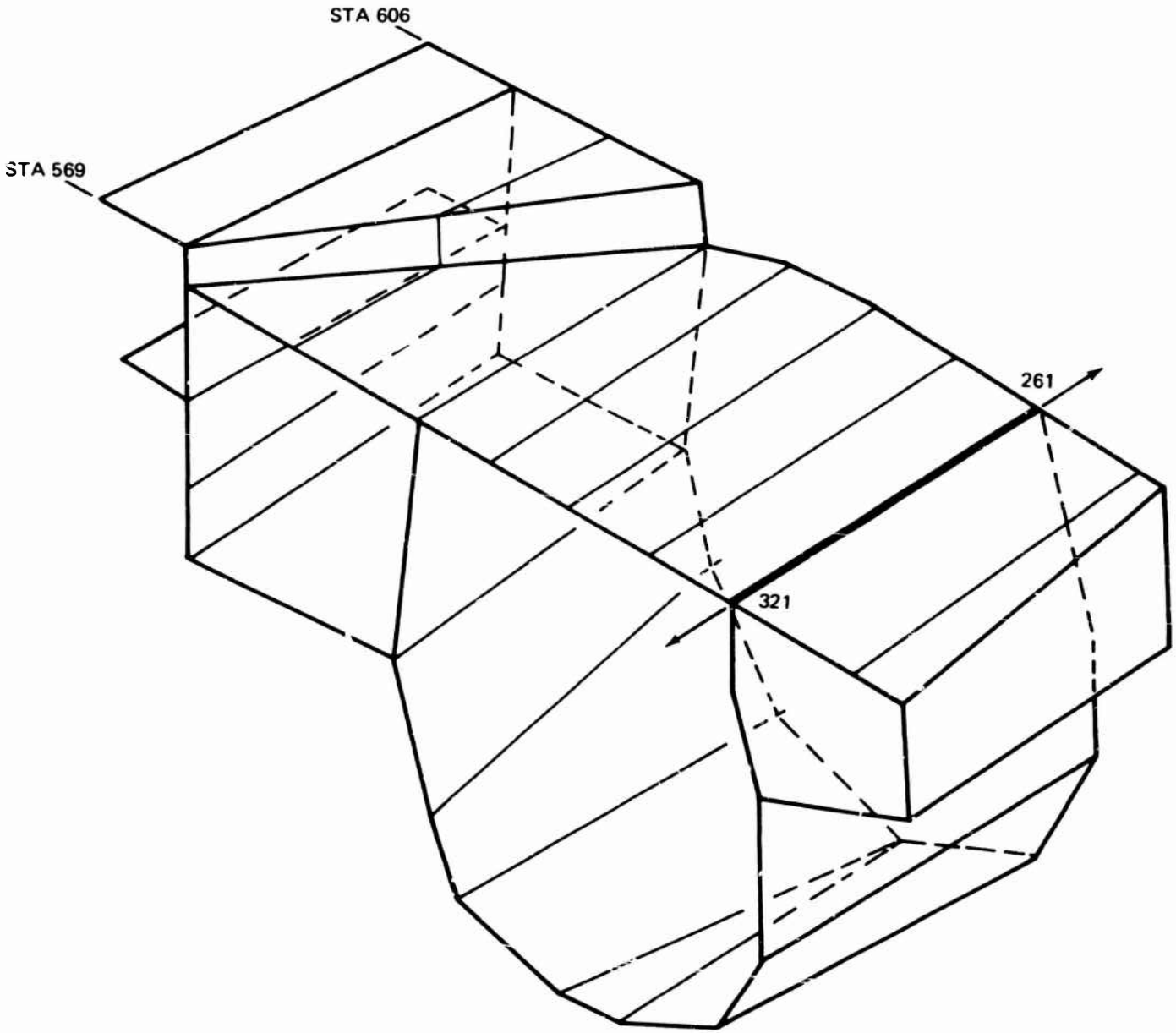
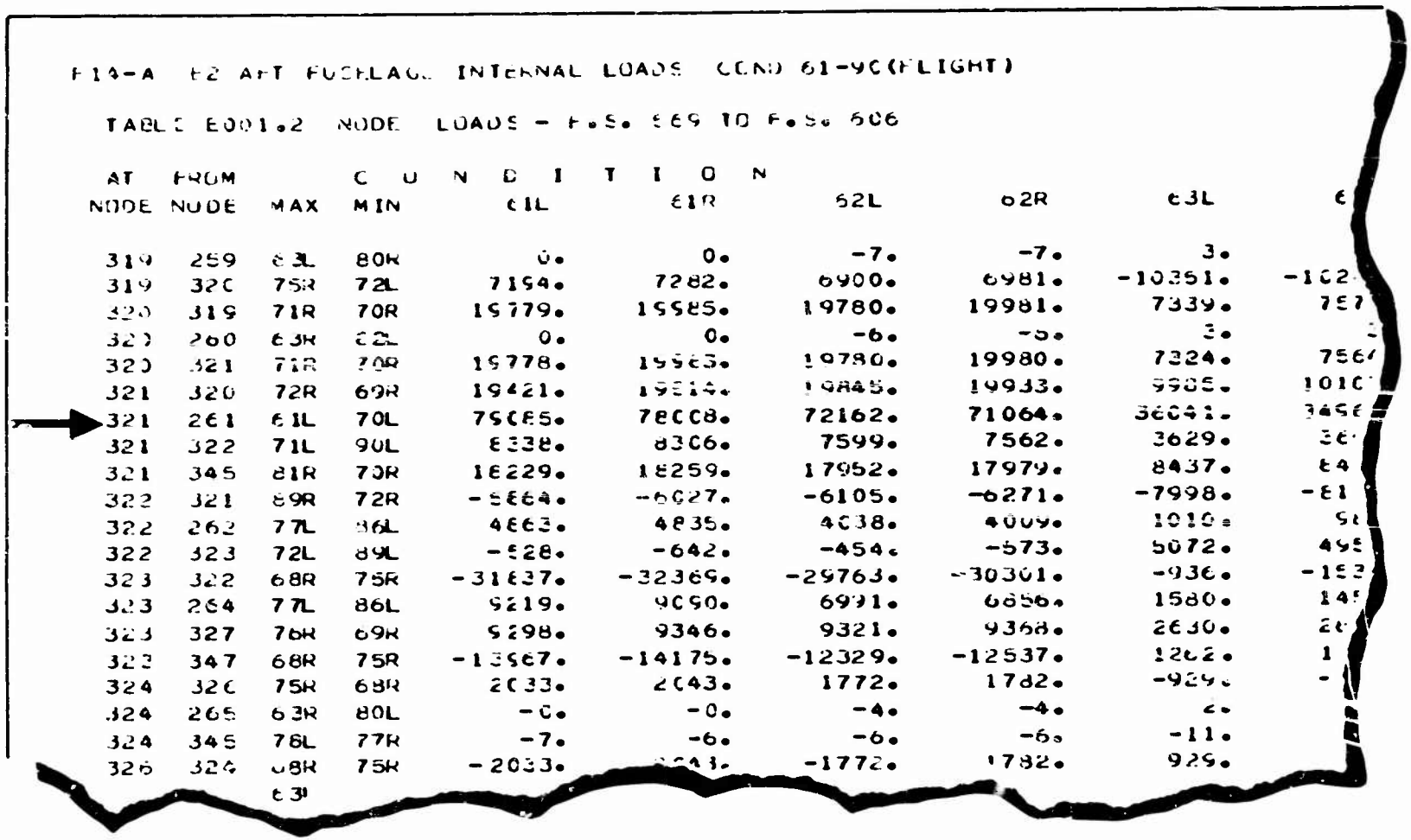


Figure 12 Typical Fuselage Section

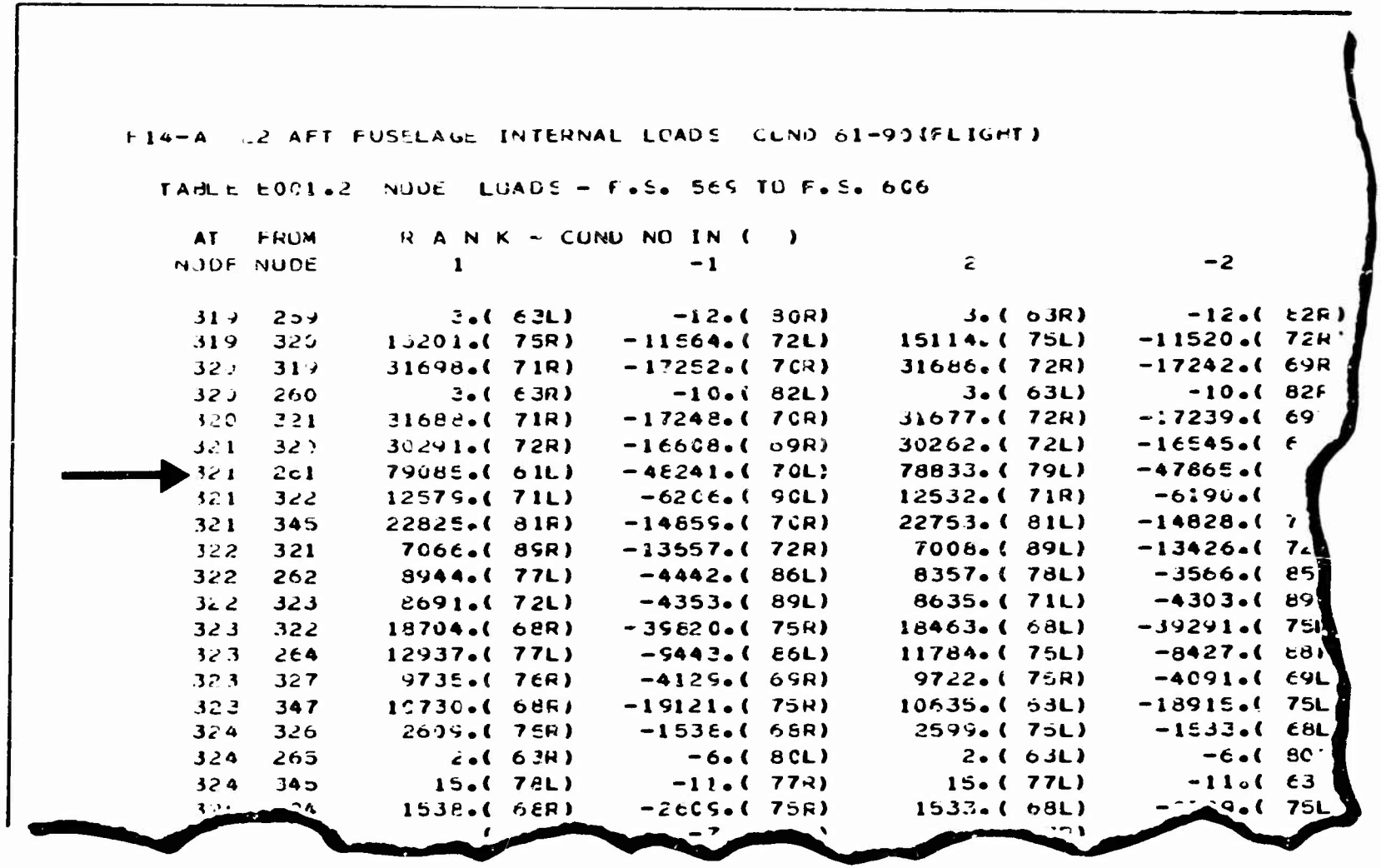


F14-A E2 AFT FUSELAGE INTERNAL LOADS COND 61-9C(FLIGHT)

TABLE E001.2 NODE LOADS - F.S. 569 TO F.S. 606

AT NODE	FROM NODE	MAX	CONDITION MIN	61L	61R	62L	62R	63L	6
319	259	63L	80R	0.	0.	-7.	-7.	3.	
319	320	75R	72L	7194.	7282.	6900.	6981.	-10351.	-102
320	319	71R	70R	19779.	19585.	19780.	19981.	7339.	757
320	260	63R	62L	0.	0.	-6.	-5.	3.	
320	321	71R	70R	19778.	19583.	19780.	19980.	7324.	756
321	320	72R	69R	19421.	19214.	19845.	19933.	9985.	1010
321	261	61L	70L	75085.	78008.	72162.	71064.	36041.	3456
321	322	71L	90L	8338.	8306.	7599.	7562.	3629.	36
321	345	81R	70R	18229.	18259.	17952.	17979.	8437.	84
322	321	89R	72R	-5864.	-6027.	-6105.	-6271.	-7998.	-81
322	262	77L	86L	4863.	4835.	4038.	4009.	1010.	9
322	323	72L	89L	-528.	-642.	-454.	-573.	5072.	495
323	322	68R	75R	-31837.	-32369.	-29763.	-30301.	-936.	-152
323	264	77L	86L	9219.	9090.	6991.	6856.	1580.	145
323	327	76R	69R	9298.	9346.	9321.	9368.	2630.	26
323	347	68R	75R	-13967.	-14175.	-12329.	-12537.	1262.	1
324	326	75R	68R	2033.	2043.	1772.	1782.	-929.	-
324	265	63R	80L	-0.	-0.	-4.	-4.	2.	
324	345	78L	77R	-7.	-6.	-6.	-6.	-11.	
326	324	68R	75R	-2033.	-2043.	-1772.	1782.	929.	
		63							

Figure 13



F14-A .2 AFT FUSELAGE INTERNAL LOADS COND 61-90(FLIGHT)

TABLE E001.2 NODE LOADS - F.S. 565 TO F.S. 606

AT NODE	FROM NODE	RANK - COND NO IN ( ) 1	-1	2	-2
319	259	3.(63L)	-12.(80R)	3.(63R)	-12.(82R)
319	320	15201.(75R)	-11564.(72L)	15114.(75L)	-11520.(72R)
320	319	31698.(71R)	-17252.(70R)	31686.(72R)	-17242.(69R)
320	260	3.(63R)	-10.(82L)	3.(63L)	-10.(82R)
320	321	31688.(71R)	-17248.(70R)	31677.(72R)	-17239.(69
321	320	30291.(72R)	-16608.(69R)	30262.(72L)	-16545.(6
321	261	79085.(61L)	-48241.(70L)	78833.(79L)	-47865.(
321	322	12579.(71L)	-6206.(90L)	12532.(71R)	-6190.(
321	345	22825.(81R)	-14859.(70R)	22753.(81L)	-14828.(7
322	321	7066.(89R)	-13557.(72R)	7008.(89L)	-13426.(7
322	262	8944.(77L)	-4442.(86L)	8357.(78L)	-3566.(85
322	323	8691.(72L)	-4353.(89L)	8635.(71L)	-4303.(89
323	322	18704.(68R)	-39820.(75R)	18463.(68L)	-39291.(75
323	264	12937.(77L)	-9443.(86L)	11784.(75L)	-8427.(88
323	327	9735.(76R)	-4129.(69R)	9722.(75R)	-4091.(69L
323	347	10730.(68R)	-19121.(75R)	10635.(68L)	-18915.(75L
324	326	2609.(75R)	-1538.(68R)	2599.(75L)	-1533.(68L
324	265	2.(63R)	-6.(80L)	2.(63L)	-6.(80
324	345	15.(78L)	-11.(77R)	15.(77L)	-11.(63
32[illegible]	[illegible]24	1538.(68R)	-2609.(75R)	1533.(68L)	-[illegible]9.(75L

Figure 14

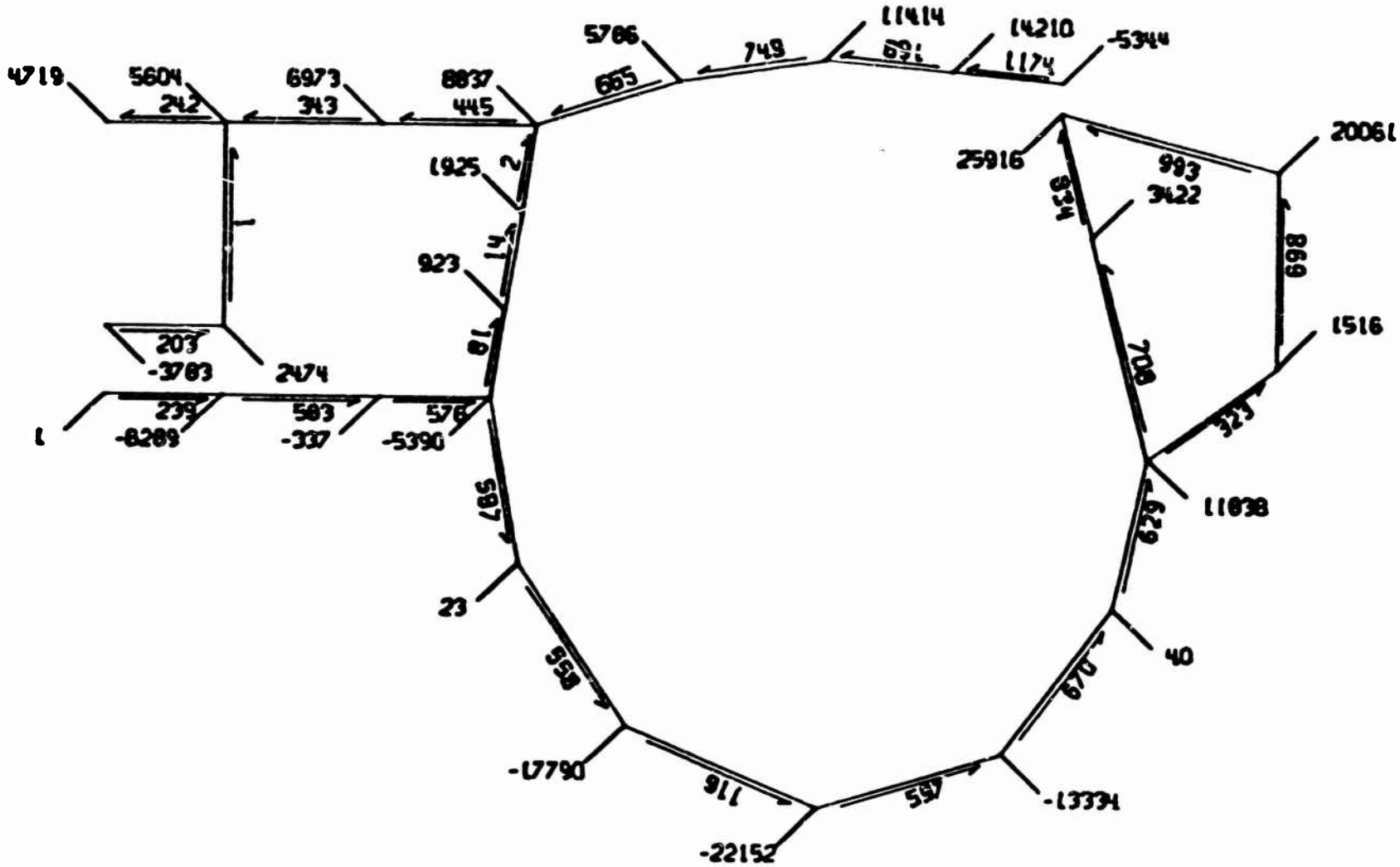


Figure 15 Internal Loads at Station Cut Aft of FS 668

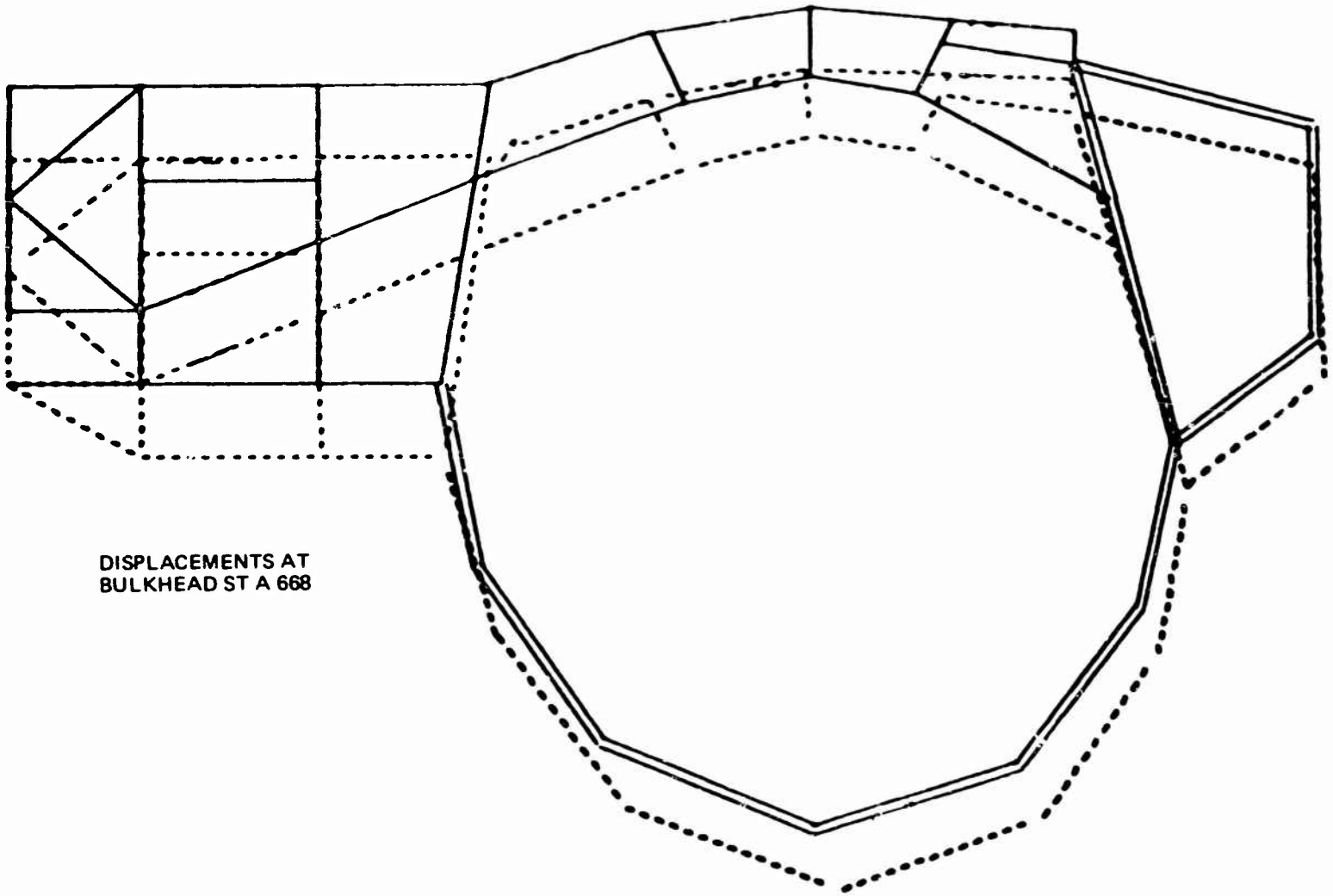


Figure 16 Displacements at Bulkhead ST 668

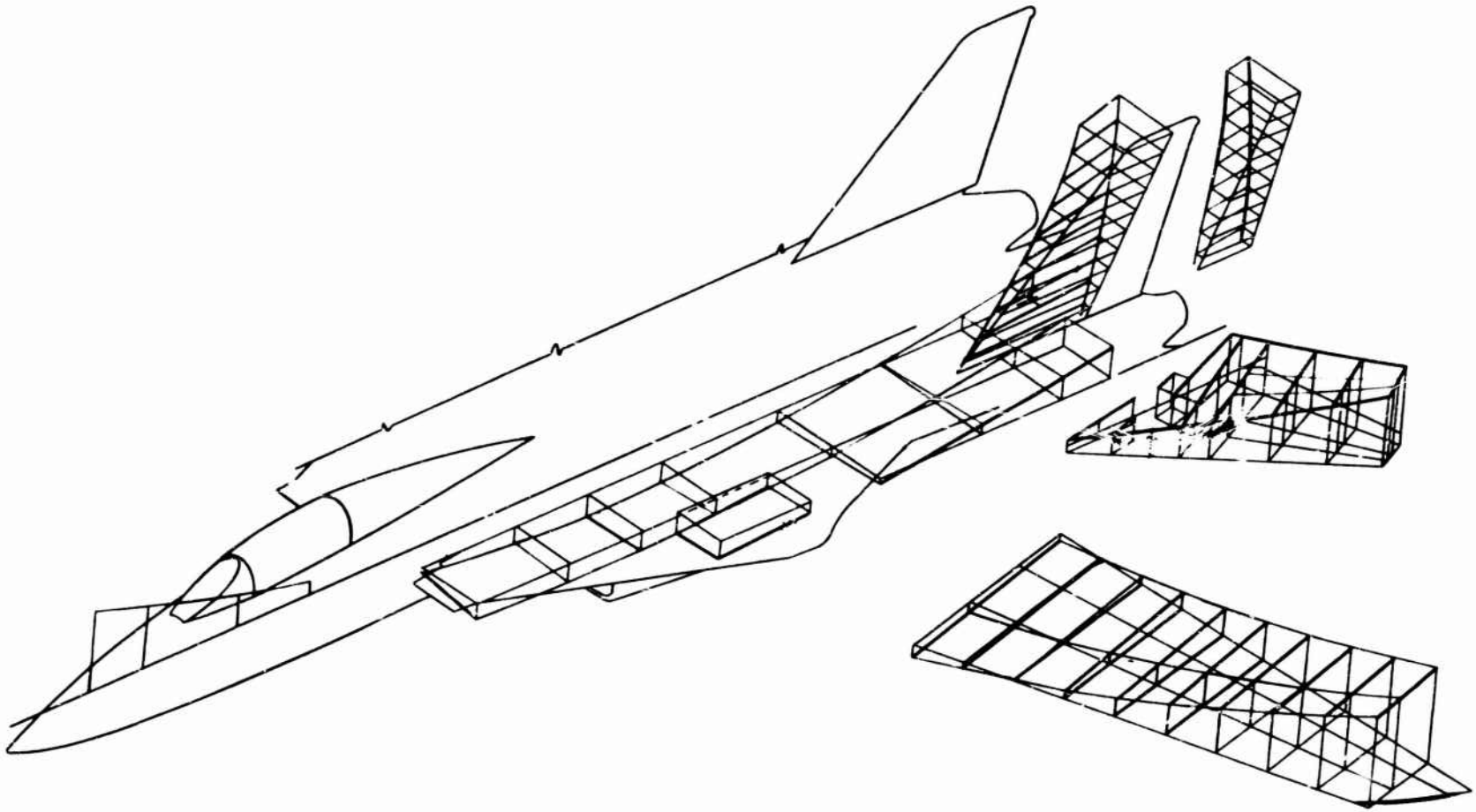


Figure 17 Second Airfame Vibration Mode

fuselage. Thus the use of fully stressed design was not even considered. This element deficiency has since been remedied in ASOP (Automated Structural Optimization Program). However, there are many other practical problems that remain in the optimization of fuselage structures, such as introducing the appropriate frame design criteria. It may still be a while before we are really ready to approach this task in any automated fashion.

In the case of the stabilizer, the situation was somewhat better. Because of the nature of the structure, it was logical to idealize it primarily by bars and anisotropic membrane elements. (See Figure 11). Thus it could be handled by an existing fully-stressed design computer program developed for wings and empennages. It was only necessary to add to the program a capability for selecting optimized boron epoxy layups according to a suitable allowable strength criteria. This work has already been discussed[5] and so will not be repeated here. However, one observation should be made. Even after attempting to introduce realistic minimum sizes into the fully stressed stabilizer design, the weight of the actual structure as built was far above that of the idealized structure. Specifically, for the main box the ratio of weights was 1.51, while for the covers alone, it was 1.28. The differences can all be accounted for in joints, splices, fasteners, adhesives, bearings, fittings, etc. The point is that a surprisingly large proportion of actual primary structure weight is not included in its corresponding finite element idealization.

The fully stressed design program also was used effectively on the titanium wing center box. It was not used on the wing outer panel, largely again because of the unavailability of a bending element in the program. The latter was needed in idealizing the wing covers adjacent to the pivot region.

Concluding Remarks on the F-14 Design Analysis

It would be an exaggeration to say that the design of the F-14 is representative of all modern American fighters. After all, with the exception of the F-111, the F-14 is the first of the only two new fighters we've had in the last fifteen years. (The F-15, now being designed by McDonnell Douglas, is the other.) Nevertheless, one can conjecture that the F-14 program contains elements of what lies ahead for most new aircraft of its type.

By far the most important F-14 design constraint was the short time available from contract go ahead to first flight. This, in combination with the requirement to do the applied loads work effectively for five different airplanes because of the variable sweep feature, absolutely dictated that the design analysts employ the systems approach in the determination of internal loads for sizing structure. By use of the IDEAS system, it was possible for them to make their inputs sufficiently early to provide a sound basis for changes in design during its more creative stages.

USE OF IDEAS IN PRELIMINARY DESIGN

As discussed previously, the first IDEAS application has been to the F-14, during the period from contract go ahead to first flight. More recently, we have used the system for preliminary design purposes on the Earth Orbital Shuttle. It would be inappropriate to go into detail on this work at the present time; however, some of the highlights can be mentioned.

Grumman, in partnership with Boeing, currently has a contract with NASA titled "An Alternate Space Shuttle Concepts Study." During our investigations last winter, we became convinced that the orbiter's main propellant liquid hydrogen tanks should be mounted externally to the fuselage. See Figure 18. The tanks could then be jettisoned while the vehicle is still in orbit, thus reducing the latter's overall size and weight for reentry. NASA was interested in the concept, but was concerned about the possibility of undesirable dynamic effects. To establish credibility for the design, it became essential to evaluate its dynamic behavior.

The specific problem of most concern was the response of the vehicle to sinusoidal variation of the thrust load during first stage of boosted flight. As in all "piggy back" arrangements, there is a primary interaction between longitudinal excitation, such as variation in engine thrust, and transverse response. (This effect is almost completely absent in the Saturn/Apollo vehicle.) In the case of our shuttle configuration, the long, flexible, end supported, LH_2 tanks were especially suspect.

In order to carry out the investigation, we first had to come up with a sufficiently realistic dynamic model of the coupled orbiter/booster vehicle. The booster was relatively easy to model, because of its extremely simple structural arrangement and load paths. The orbiter was something else. Because of the provisions necessary for payload, various internal "floating" tanks for liquid propellants, and transfer of thrust load from booster to orbiter during first stage boost - as well as the external hydrogen tanks - the orbiter structural arrangement was much more complex. The idealized structure correspondingly had to be more detailed.

Based upon the preliminary layouts available for our H3T configuration, (Figure 18), we roughed out node locations and topology for a complete, idealized orbiter/booster structure. The total number of structural elements employed (bars, beams, shear panels, etc.) was 1763, of which 92% were in the orbiter. Preliminary applied loads for various design conditions were obtained, and distributed to the idealized structural components. In the case of the orbiter wing and fin, beginning with completely arbitrary distributions of material, fully stressed designs were obtained by use of the ASOP program. For the fuselage, member sizes were first selected by the stress analysts using conventional design procedures. An analysis was then run, and based upon the results, the structure was resized. The resized fuselage structure was then used to compute flexibilities in the subsequent dynamics calculations.

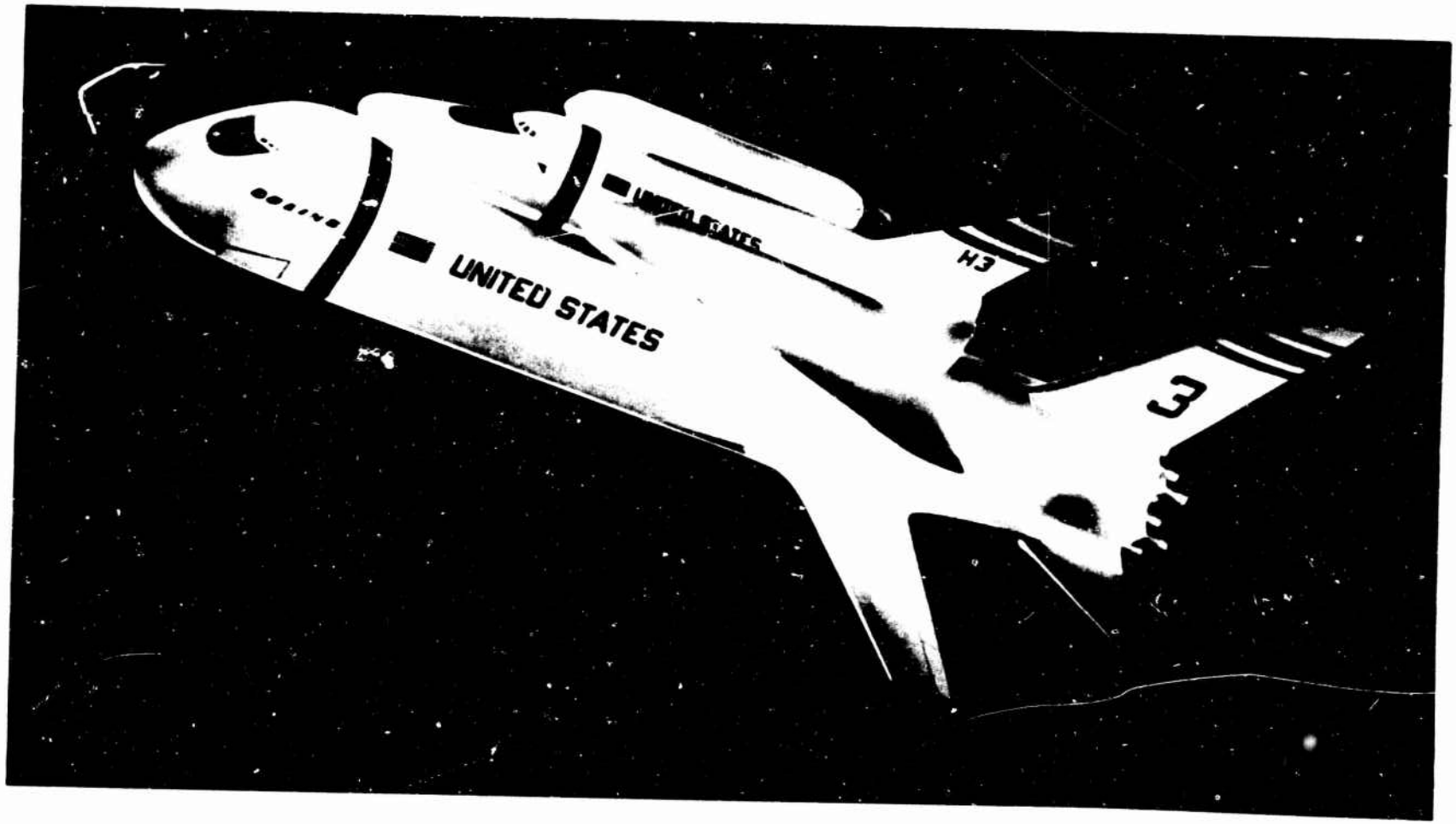


Figure 18 Grumman/Boeing External Hydrogen Tank Space Shuttle

Mass and inertia idealizations had been obtained concurrently, together with the necessary transformation matrices. It was now possible, using selected portions of IDEAS, to carry out all of the matrix manipulations required. These included coupling together the booster and the fuselage, wing and fin of the orbiter, transforming the structural influence coefficient matrices to conform to the dynamics model, obtaining the normal mode characteristics, and finally obtaining the desired response to sinusoidal variations in booster engine thrust.

The results of the investigation were very gratifying. They were completed on schedule and indicated that the external hydrogen tanks presented no additional dynamic problems. Of course the analyses also yielded all of the usual static structure information - internal loads, deformations, etc.

Once the preceding objectives had been met and the pressure was off, we went back and applied the ASOP program to the orbiter fuselage idealized structure. This time we started out with completely arbitrary member areas, skin gages, etc., i.e., set equal to unity, together with specified minimum sizes and simplified strength allowables. Because the latest applied loads, as used in this calculation, were different from those employed earlier, a direct comparison with the earlier, hand resized results was not possible. However, there was now a drastic reduction in the new weight of the frames relative to the external shell structure. Before being able to place any confidence in such results, we shall have to examine the failure criteria employed for the frames much more closely.

CURRENT DEFICIENCIES AND FUTURE TRENDS

Structural Optimization

Earlier in this paper, I have tried to show that, from a practical design point of view, it is still a struggle merely to obtain internal loads in primary structure on a time schedule that is attuned to the needs of the designer. Traditionally, once the internal loads are known, the analyst and the designer use them to resize the structural members as required. An essential ingredient in the redesign is a good knowledge of the failure criteria and the accompanying allowable stresses to be employed; of course these will depend upon the nature of the various structural elements and their loadings. Member resizing is customarily based upon the ratios of the predicted stresses to these allowables. However, the method is not applied rigidly. Also involved are the judgements of the analyst and designer as to what constitutes a well-balanced design, taking into account good manufacturing practices, etc. This is followed up by a reanalysis check whenever possible.

A very normal, evolutionary improvement in this procedure is to automate the resizing, based upon the stress ratios just mentioned. Of course, this is the motivation that has led to the inclusion of fully stressed design features in a number of computer systems recently reported in the literature - a recent version of W. D. Whetstone's original SNAP program[6] and NASA Langley's DAWNS program[7], to name a few in addition to the Grumman references cited earlier. All these programs have the characteristic that between analysis cycles, the structural elements are either resized to carry some prescribed allowable stress in at least one design condition, or are at their specified minimum sizes. In most cases, the stress ratio is based upon some average stress that can be identified for each element. The Grumman programs are different in that for the skin portion of a semi-monocoque structure, the individual element stresses are combined into "nodal stresses."[3,4,5] In our experience, this leads to smoother distributions of material than the average stress approach. In either event, after a small number of analysis and redesign cycles, say five, the structural weight and material distribution usually settle down, so that further recycling causes negligible improvements. When one considers that the idealized structure weight may only represent on the order of two thirds of the equivalent actual primary structure weight (as in the case of the F-14 stabilizer), the last percent or so of idealized structure weight saved may be relatively meaningless.

Theoretically, of course, one can argue that fully-stressed designs are not necessarily optimum even for strength critical structures. Perhaps as we gain more experience in the use of these procedures, we will encounter practical airframe design cases where fully-stressed designs are in fact significantly non-optimum. The writer doubts that this will be the case for wings, but fuselages could be another story. We just don't know. However, this possibility should not deter us from using the tool until we have a more optimum one to replace it. The latter could be a long time in coming; current projections of computer running times for reasonably large structures optimized by mathematical search techniques are still quite high.

Fox and Schmit[8] have suggested that the mathematical search approach could be made workable for such large idealized structures by abandoning the one to one relationship between the number of optimization variables and the number of finite elements employed. Instead, they propose that only certain key elements be optimized, and that one interpolate between them for the sizes of the other elements. Tocher and Karnes[9] have incorporated this interpolation concept into a mathematical search program which can handle structures up to 500 members in size, of which less than twenty percent of the members are actively optimized. They have applied it successfully to several relatively small structural components. It will be interesting to see whether this approach can be extended to larger structures in a practical way.

No matter what resizing algorithm eventually turns out to be best, it is only part of the story for strength critical structures. From a practical design point of view, a very significant part of the total methods development effort will have to go into delineation of all of the strength criteria and resulting allowable stresses that will be required. I referred to one especially sticky case of this previously, that of fuselage frames. Here stiffness rather than strength may be critical, and the situation will become further complicated if the fuselage skins buckle and go into partial diagonal tension. Compression allowables for stiffened panels such as are used in wing skins is another example. Fibrous composites and all of their applications is a whole new field. Fatigue allowables different from static strength allowables must be accommodated. Until features such as these are incorporated into the programs, their usefulness in practical applications will be severely limited. The allowables aspect of the work apparently has been given very little attention so far, at least in the literature.

Another matter of extreme practical importance is that of cost. More and more, in the current environment, weight penalties are being taken in order to reduce the overall price of a design. One illustration of this is simplified machining. By reducing the number of cuts in an integrally stiffened wing skin, for example, one can save on machining costs, and perhaps at a fairly small weight penalty. We ought to be thinking in terms of evaluating such alternatives by means of our optimization procedures.[10]

Up to this point, the discussion of structural optimization has dealt exclusively with strength as the constraint. This is as it should be - the first requirement for any aerospace vehicle is that it be able to carry the "static" loads applied to it. However, there are other constraints over and above strength that sometimes become critical.

Displacements as a constraint have been included practically from the beginning in the mathematical search work that has been done[11]. In the original work by Schmit, they have been treated exactly as the strength constraints, and travel in design space has been either normal to or parallel to the equal weight surfaces until the constraint surfaces have been encountered.

More recently, Venkayya et. al.[12] have suggested that travel normal to a deflection constraint surface - combined with travel associated with "scaling" the sizes of all structural members proportionately - would be preferable. In the examples examined to date, in combination with a suitable stress ratio algorithm for strength, the method appears to lead to designs that are generally competitive with those obtained by traditional mathematical search, and in significantly less computer time. It has been incorporated into and is operational in the ASOP program. Berke[13] has another interesting intuitive approach to the constrained displacement problem. The resizing algorithm is a very simple one which is theoretically correct only for statically determinate structures. It is somewhat analogous to the stress ratio algorithm of fully stressed designs, and is used iteratively in a similar fashion. Exploratory investigations conducted so far indicate that it may prove to be even more desirable than Venkayya's approach, from the point of view of computer running time.[14]

Vibration frequencies of aerospace structures can also impose design constraints. There have been a number of papers on this recently, of which the one by Rubin[15] is especially attractive. Here, one of the travel directions is normal to the frequency constraint surface - the other along the frequency constraint and directed so as to reduce weight. In order to be applicable in airframe structural design, strength constraints would have to be introduced as well.

Flutter is still another constraint. There have been a number of papers already in this area, of which Turner's[16] was the first. One of the most recent, by Rudisill and Bhatia[17], is very promising in that it provides the direct means for obtaining gradients to the flutter velocity constraint surface. Programs for optimizing directly for both strength and flutter could be of great practical assistance on many aircraft, particularly those designed for low maneuver load factors. These can be significantly flutter critical.

A constraint of special interest for the Space Shuttle is thermal stress. This one is especially tricky, because it could involve configuration changes. A common design approach to the thermal stress problem is to reconfigure the structure if possible, so that the thermal effects will be minimized. Perhaps, automated redesign procedures can help in this area too.

Integrated Design and Analysis Systems

Of course, other organizations in addition to Grumman are hard at work on automated design systems. We shall be hearing about the Boeing experiences on the 747 and the SST at this meeting. In addition, there is the NASA DAWNS program,[7] which is part of an exploratory development project currently under way at Langley. DAWNS integrates the structures and aerodynamics disciplines into an automated design operation, and it can be used in an interactive mode with a cathode ray tube display. Primarily for wing type structures, it could be the beginning of a very attractive tool for use in preliminary design.

Langley has other goals that extend far beyond DAWNS. Over the next five to ten years, in partnership with industry, they contemplate the development of a system called IPAD, Integrated Programs for Aerospace Vehicle Design.[7] It is intended to embrace all of the disciplines that we have discussed previously as being required in aerospace structural design. Perhaps the most significant improvement that they visualize is in the complete automation of the interfaces that exist between the disciplines, as well as automated access to all stored data. The goal thereby is to reduce the total cycle time for review of a given design from a period of months to one of weeks. In achieving this objective, an avowed guideline is that the system be sufficiently versatile and flexible that industry will want to use it. They are soliciting industry's participation in developing it, so that this can in fact take place. The plan that Langley proposes appears to be a desirable and necessary one. For any one company to develop such a system using only its own resources appears to be very unrealistic under current conditions.

Computer Graphics

Large scale batch mode analyses of the type discussed for the F-14 have already benefited substantially from the use of graphic aids. Thus far, most aerospace companies have concentrated their efforts for these large analyses in two areas:

Graphic Input Data Checking

Graphic Display of Results

In the first of these areas, commonly used techniques involve the use of automated drafting equipment, such as the Orthomat mentioned earlier, for display purposes. In addition, several companies have techniques for interactively displaying and correcting the input data for structural models. Lockheed in particular has done very interesting work along these lines, and will be reporting upon some of it at this meeting.

The second big payoff area is the graphic display of structural analysis output data. In past years the results of an analysis were generally in the form of huge computer listings containing internal load distributions. These were handed over to the structural analysis personnel for their use. The information usually was hand transcribed onto model drawings so that examination and use of the results could proceed. The time consuming nature of the work, of the order of many man months, dictated a better approach. Initially, automated drafting or plotting equipment was tried, but this was still very slow. More recently, devices such as the DD-80 and other similar machines have made an appearance. They are based upon display of results on a cathode ray tube plus microfilm recording, and are much speedier. A recent advance in high resolution microfilm plotting which enables reproduction up to 30" x 40" (the FR80) has made possible the display of results in larger segments of the model. In conjunction with computer programs that can search for critical conditions, display of results ready for input into final stress reports is now a normal mode of operation in several companies.

While all these methods have led to great cost and time savings in the area of large scale structural analysis, a major bottleneck still remains, namely in the initial preparation of the input data. New hardware configurations enabling the marriage of low cost graphics terminals, x-y digitizers, and hard copy devices now permit the generation of input data in the engineering areas, interactively, without resorting to former hand transcription, card punching, verifying and associated time consuming processes. By using these devices in conjunction with a time share computer one can generate input data tapes for future runs on a batch computer. Work in this area is currently under way and we believe there will be a great payoff in the next few years.

The large scale batch mode analyses discussed earlier basically require for their staffing specialists having rather distinctive aptitudes, training and experience. The smaller, more detailed analyses that accompany and compliment the large scale ones are less specialized. Here, the design analyst himself should be in complete charge from beginning to end. (Bulkheads and fittings are good candidate components.) In this situation the interactive programs centered about machines like the IBM 2250 are very attractive. By their use, the engineer at the console can first enter the input data describing the structure either by light pen or appropriate data cards. He then calls for a solution, and the results are displayed on the cathode ray tube. While the computer is still "on line," he makes any changes to the structure he desires and reanalyses. Lockheed has had such a system operational for several years for two dimensional structures.[18] Other companies have similar capabilities in various stages of development, and this tool should come into widespread usage in the next few years.

Another interactive computer graphics area under development in a number of organizations is the tie in of the preceding efforts with a master dimensions library. These systems, usually referred to as design drafting systems, generally start with a stored geometric definition of the external contours of the vehicle. A designer sitting at a console can call up desired sections of the aircraft, perform a modelling function, and exercise various engineering application programs to arrive at final detail designs. Drawings, including all details, may also be created at the scope. Following the creation of the drawing, hard copy can be obtained via a flat bed plotter, high resolution computer output microfilmer or high speed laser-type plotting equipment. The full exploitation of such techniques appears years away, but several aerospace companies already have made sufficient progress to be marketing systems of this type.

CONCLUDING REMARKS

As indicated at the outset, my basic goal has been to portray the role of the design analyst during the evolution of an aerospace structural design, and the analysis tools that he needs to do his work. There is one principal idea that I have tried to develop and that I hope I can leave with you. It is that the design process and the analyses that must accompany it cover a wide range of technical specialties. Each specialty plays a vital role, and if any one of them falters, either schedules will not be met, or what is more likely, design decisions will be made on schedule, but they will be based upon incomplete information. We must be realistic. Our methods development work must focus, more effectively than heretofore, upon those areas where we are most deficient. If our industry is to become more competitive in the world marketplace, these are the areas where we should be expending a major portion of whatever resources are available for analytical procedures.

REFERENCES

1. Hedrick, I. G., "Future Trends of Structures and Materials for Spacecraft, Aircraft and High Speed Surface Vehicles," presented at AIAA 7th Annual Meeting in Houston, Texas - 19 October 1970.

2. Wennagel, G. J., Mason, P. W., and Rosenbaum, J. D., "IDEAS, Integrated Design and Analysis System," Paper No. 680728, presented at Society of Automotive Engineers Aeronautic and Space Engineering and Manufacturing Meeting, Los Angeles, Calif., October 7-11, 1968.

3. Dwyer, W. J., Emerton, R. K. and Ojalvo, I. U., "An Automated Procedure for the Optimization of Practical Aerospace Structures," AFFDL-TR-70-113, Vol. I and II, April 1971.

4. Dwyer, W., Rosenbaum, J., Shulman, M., and Pardo, H., "Fully Stressed Design of Airframe Redundant Structures," Proceedings of the Second Conference on Matrix Methods in Structural Mechanics, AFFDL TR-68-150, December 1969.

5. Lansing, W., Dwyer, W., Emerton, R., and Ranalli, E., "Application of Fully Stressed Design Procedures to Wing and Empennage Structures," AIAA Journal of Aircraft, Vol. 8, No. 9, September, 1971, pp. 683-688.

6. Whetstone, W. D., "Computer Analysis of Large Linear Frames," J. Struct. Div., ASCE, October, 1969.

7. Fulton, R. E. and McComb, H. G., "Automated Design of Aerospace Structures," presented at the ASME International Conference on Design Automation, Toronto, Canada, September 1971.

8. Fox, R. L. and Schmit, L. A., "Advances in the Integrated Approach to Structural Synthesis," Journal of Spacecraft and Rockets, Vol. 3, No. 6, June 1966, pp. 858-866.

9. Tocher, J. L. and Karnes, R. N., "The Impact of Automated Structural Optimization on Actual Design," AIAA Paper No. 71-361, presented at AIAA/ ASME 12th Structures, Structural Dynamics and Materials Conference, Anaheim, California, April 19-21, 1971.

10. Taig, I. C., "Automated Analysis and Efficient Design of Real Structures," AGARD Conference Proceedings No. 36 on Symposium on Structural Optimization, Istanbul, Turkey, October 1969.

11. Schmit, L. A., "Structural Design by Systematic Synthesis," Proceedings of the ASCE Second Conference on Electronic Computation, Pittsburgh, Pennsylvania, September 1960, p. 105.

12. Venkayya, V. B., Khot, N. S., and Reddy, V. S., "Optimization of Structures Based On The Study of Energy Distribution," Proceedings of the Second Conference on Matrix Methods in Structural Mechanics, AFFDL TR-68-150, December 1969.

13. Berke, L., "An Efficient Approach to the Minimum Weight Design of Deflection Limited Structures," AFFDL TM 70-4-FDTR, July 1970.

14. Gellatly, R., and Berke, L., "Optimal Structural Design," AFFDL TR 70-165, February 1971.

15. Rubin, C. P., "Minimum Weight Design of Complex Structures Subjected to a Frequency Constraint," AIAA Journal, Vol. 8, No. 5, May 1970, pp. 923-927.

16. Turner, M. J., "Optimization of Structures to Satisfy Flutter Requirements," AIAA Journal, Vol. 7, No. 5, May 1969, pp. 945-951.

17. Rudisill, C. S., and Bhatia, K. G., "Optimization of Complex Structures to Satisfy Flutter Requirements," AIAA Journal, Vol. 9, No. 8, August 1971, pp. 1487-1491.

18. Batdorf, W. J., Kapur, S. S., and Sayer, R. B., "The Role of Computer Graphics in the Structural Design Process," Proceedings of the Second Conference on Matrix Methods in Structural Mechanics, AFFDL TR 68-150, December 1969.

THE BOEING SST PROTOTYPE

INTERNAL LOADS ANALYSIS SYSTEM AND PROCEDURES

ANDREW F. GRISHAM*

THE BOEING COMPANY

This paper covers the management and organization required to implement the Boeing Supersonic Transport Internal Loads Program, the level of technical detail attained via the analysis, the quality control of the final results, and the tools used by engineering to formulate, solve, and display the total program.

* Structures Specialist Engineer

Aerospace Division
Seattle, Washington

1.0 INTRODUCTION

Probably the single most significant change in structural analysis methods in the past 25 years has been the application of matrix structural analysis methods to aircraft and aerospace structures. This change has been possible through the development of the high speed digital computer with large data storage devices, and has been encouraged by the high cost effectiveness of removing unneeded weight from modern commercial aircraft transports. (The Boeing Company considered an expenditure of $150 to $200 per pound of weight saved per airplane to be cost effective during the development of the SST.)

A major consequence of the Boeing Saturn, 747, and SST programs is a significant advancement in the capabilities of structural analysis finite element methods. The structural technical requirements of these programs, culminating with the SST, were such that many millions of dollars and approximately 100 man years of effort were spent in the development of a fully interfaced system of finite element programs to make them handle all types of structural problems, and function efficiently in the process. Furthermore, Boeing developed highly skilled individuals who are very versatile in the use of these tools. More than 40 such specialists were employed within the SST stress group alone during the drawing release phase of the prototypes. The capabilities of these computer programs in type and size of problem treatable; in detail, accuracy and quality control of results; and in applicability to the hardware development, design, and release phases has been well demonstrated in the Boeing 747 and Super Sonic Transport programs.

The reasons for extensive development of a fully interfaced finite element computer program on the SST were: 1) a six-year development period prior to drawing release of the prototypes, 2) a ready access to SST monies, 3) a third-generation computer facility, and 4) a complex environment and structural system. For these reasons advances in the state-of-the-art in finite element analysis were made in: 1) finite element program development, 2) size and scope of problems treated, 3) run time improvement, 4) quality control of solutions, 5) technical procedures, and 6) management of the development and application of finite element methods.

The balance of this paper presents in four sections the background, finite element tools developed, management of finite element program development and use, and examples of the application of these programs to the SST prototype aircraft.

2.0 BACKGROUND

The advent of the supersonic and hypersonic aerospace vehicle together with the continuing demand for higher structural performance, including both weight efficiency and service life, have created the environment for extensive application of the computer to solving structural problems. This situation existed to a high level on the Super Sonic Transport program where the complexity of the structural system, the wide range of load and thermal conditions, and the demands for maximum structural efficiency combined to present a formidable problem to the structures engineers.

To handle a structural system the size of the Super Sonic Transport, in the level of detail required to insure maximum structural efficiencies, required development of a large size computer structural analysis program with a set of fully interfaced peripheral programs that can handle the multitude of supportive tasks that are part of any detail structural analysis.

The requirements set for the structural analysis were stringent, but were management decisions based on extensive experience with such aircraft as the B-70 and the X-20. Specifically, it was decided to insure that there was sufficient size capability that very large problems could be run both as discreet sub-structures and if necessary through an interact system. All six degrees of freedom for each node of a structural element would be included (secondary stresses are very important in fatigue). Structural modeling will include all structural elements (minimize lumping and de-lumping), and must be able to interface all disciplines supporting a structural analysis within the computer structural analysis program.

In addition to these technical requirements there existed one other requirement that may be considered to be the most important. That was a management dedicated to insuring the computer programs were developed, computers were available, and that the engineering drawing release schedules reflected the time it takes to achieve computer developed structural solutions. This management commitment did exist within the Structures Staff on the SST. The result was development of the system of fully interfaced computer programs to be discussed, and perhaps just as important, the tools and techniques used to manage and schedule the data used to support the computer programs and the programs use. The programs to be discussed were applied to the initial release of the Super Sonic Transport structural drawings. The extensive computer use resulted in a very significant weight savings and would have been the major factor in insuring a long service life for the SST structural system.

3.0 FINITE ELEMENT ANALYSIS CAPABILITIES

The primary objective of the SST internal loads program development activity was to produce a set of fully interfaced programs which would, in a timely and cost effective manner, produce a complete set of quality structural element internal loads for each structural element of the SST airframe. Structural geometry, sizing, redundancies, buckling, large deflection effects, loadings (mechanical and thermal), and criteria are accounted for and evaluated in a finite element closed solution. The programs are equally applicable to large and small problems of varying degrees of complexity without imposing significant penalties on either the small or large size problem. Internal loads on each structural element are obtained to avoid time consuming engineering effort and loss of accuracy associated with lumping and de-lumping of mathematical models.

The programs are designed to be fully interfaced: since for all structural analyses it is important that the various computational and display modules at certain points in the analysis procedure report back to the engineering staff before further computation is performed, because the engineering requirements downstream in an analysis change frequently based upon the information learned from the earlier data processing, and because of the large quantity of data which must flow across interfaces without incurring any error in the process. Because of this interfacing it is possible to exercise quality control of the

data processing, and also to conduct engineering reviews of technical and structural aspects of the solutions routinely and as the solutions progress. Built into the interfaced system is a capability for the engineer to bypass any or all pauses in the data processing. These options may be selected where problems are sufficiently small or routine to preclude, in the judgment of those managing the analysis, the necessity of reviewing interface reports prior to the completion of a larger segment of the analysis.

Major objectives in using finite element analyses in support of the SST prototype were:

1) to produce a complete set of quality structural element internal loads for each structural element of the SST airframe,
2) to allow treatment of both thermal induced and mechanically induced loads for all load conditions of interest, including those used for strength design, fatigue, aerodynamic shape and performance,
3) to allow large scale finite element analysis of the entire structural assembly, primary and secondary structure included,
4) to allow fine-grid analyses for stress and deflection prediction within the large scale finite element analysis, using substructuring concepts,
5) to allow three complete internal loads cycles of each structural element prior to 25% drawing release,
6) to produce rapidly a fully interacted solution in a local region where numerous stiffness, grid, and/or connectivity variations are being studied,
7) to allow the inclusion of large deflection and skin post-buckling effects as an integral part of the analysis rather than after the fact,
8) to take advantage of the repeatability and consistency of solution results, which a closed form finite element solution offers, in treating both a single concept and in comparing concepts.
9) to take advantage of a relatively economical tool to treat the large variety of concepts, details, and conditions expected in a technically optimum manner,
10) to support structural development of components, subassemblies, and assemblies by analysis allowing optimization of structural parameters,
11) to allow as detailed an evaluation of the state of stress and deflection in a local region as is required,
12) to produce compatible deflected shapes for components throughout the airframe for load cases of interest,
13) to facilitate definition of those items requiring further structural development,
14) to provide a firm technical base for program decisions on weight, cost, materials, scheduling, and functional reliability,
15) to search for and flag critical stresses for strength, fracture, and fatigue analysis and
16) to support test planning by analysis before-the-fact of test loads and support conditions.

Of fundamental importance to the success of a finite element internal loads analysis is the finite element program which generates and solves the stiffness (or flexibility) equations. With an efficient general purpose program of this type the user greatly extends the scope of the problem he can treat. It is necessary to surround such a finite element program with a complete system of fully interfaced peripheral programs if the system is to be fully capable of practically dealing with the full range of complex problems associated with an SST airframe. This finite element system in turn may become a fully interfaced subsystem of the technology staff programs (which include programs used

by aerodynamics, performance, stability and control, flutter, and dynamic and static external loads groups).

The SAMECS (Structural Analysis Method for Evaluation of Complex Structures) finite element program system now includes SAMECS and the following fully interfaced peripheral programs: 1) the automatic transformation of loads to the structural grid program, 2) the structural analysis input language program (SAIL), 3) the iterative large displacement program, 4) a library of master dimensions programs of both general and special purpose types which generate geometric data for use by the system as structural geometry becomes more completely defined, 5) a freebody program for forming freebodies within SAMECS models and also for further substructing or updating these models, 6) a substructure interaction capability, 7) automatic input data sorting and updating, 8) an input data format and card sequence logic check program, 9) input and output data display programs, and 10) the post-buckling and diagonal tension effects program.

3.1 SAMECS

The SAMECS program is a powerful tool commonly used to analyze complex structures which may be idealized as an assemblage of plate and beam elements. Other types of elements may be added as necessary via an external stiffness and loads matrix pre-merge option (also used for interaction of substructures). The sub-paragraphs of this section describe the capabilities of SAMECS.

3.1.1 SAMECS Maximum Problem Size

There is no limit on the maximum size problem treatable by SAMECS except, of course, available computer budget. Any structure may be subdivided into any number of substructures. A substructure may contain as many as 2000 six-degree-of-freedom nodes, 10,000 plate elements, and 10,000 beam elements. Loads may be applied in any combination of thermal and mechanical environments up to 40 load cases per pass.

3.1.2 SAMECS Run Time and Computer Costs

Because of the run time improvements incorporated in SAMECS and the large data storage capabilities of the Boeing computing facilities, SAMECS is currently considered to be capable of processing large structural systems and with less run time and computer cost than any other known finite element program. For a structure which fits within the size restrictions of the SAMECS substructure (and therefore does not require interaction to complete the analysis) the computer cost varies between 1/2 cent and $2\frac{1}{4}$ cents per node freedom per load case with the unit cost decreasing as the complexity of the problem decreases and as the number of load cases processed increases. Data processing costs are slightly less when SAMECS runs are processed using the full core of the computer in a single job rather than multi-job processing mode. However, because full core jobs must be scheduled, smaller problems are usually processed in the multi-job processing mode in order to obtain fast turn-around.

Structures which do not fit within the size restrictions of a SAMECS substructure and must be interacted cost between 1/2 cent and 6 cents per node freedom per load case with the unit cost decreasing as the complexity of the problem decreases, as the number of load cases processed increases, and where

3.1.2 (Continued)

more than one interaction run is required. These cost figures are based on a computing facility cost of $2.75 per Computer Resource Unit.

Computer data processing run time can be approximately computed from run cost since the Boeing CDC 6600 facility costs approximately 250 computer resource units or $687.50 per hour for SAMECS full core (stand alone) processing. No prediction of run time is possible in the multi-processing mode since the occupancy on the computer is then a function of the number and types of other jobs processing simultaneously, relative job priority in the system, and utilization of disk, tape, printers, core, card readers, card punches, and other components of the computer by the multi-processing jobs. However, a typical statistic on a 240-node six-degree-of-freedom job having six load cases would be 55 computer-resource-units and an occupancy on the computer in the multi-processing mode of 18 minutes.

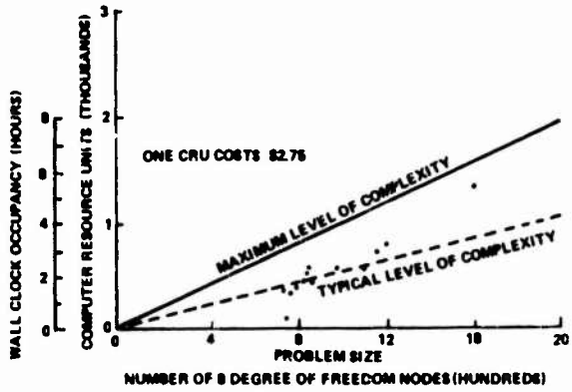


Figure 3.1.2: SAMECS INTERNAL LOADS RUN TIME AND COST

On the SST program 800 to 1000 node runs processed routinely overnight. The cost of processing a 1000-node (6000-degree-of-freedom) run with 20 load cases is typically 500 computer-resource-units and the occupancy full core is typically 2 hours. A graph showing typical computer data processing statistics on run time and cost is shown in Figure 3.1.2.

3.1.3 SAMECS Finite Element Capabilities

Nodes, plates, and beams may be randomly numbered in SAMECS. Node freedoms may be fixed, free, specified or sprung.

Plate elements are isotropic, orthotropic, isotropic stiffened, or orthotropic stiffened; may be pre-strained, heated through their thickness and across their surface linearly, may be offset, and allow heating of offsets. The quadrilateral plate element is always in precise force and moment equilibrium, even though warped, allowing moment and force equilibrium checks for all structural assemblies of elements without exception. All varieties of plate elements posses both inplane shear and stretch, and out-of-plane shear and bending. The quadrilateral plate element is formed through the use of a centrally located fifth node and four triangular plate elements. An inplane rotational stiffness for the plate element is generated to assist the user in generating three independent rotational stiffnesses at all plate nodes. The effect of this inplane rotational stiffness on plate element inplane translational stiffness is not allowed to exceed 1% for a plate element regardless of aspect ratio.

Beam elements are curved or straight, may be heated axially and through their depths along both major and minor axes of bending, may be pinned torsionally

and about both bending axes at each end of the beam, may be offset in any orientation relative to its end nodes (same as the plate element), allows heating of offsets, has constant, linearly varying and piecewise linearly varying section properties, and may be loaded between nodes with uniform, linearly varying, piecewise linearly varying, and concentrated loads in both local and general coordinate systems. Section properties and loads may be defined at as many as 50 stations between the nodes of a beam element.

In addition to plate and beam elements, super elements (stiffness matrices or reduced stiffness matrices for other structures) and solid element stiffness and loads matrices (full or reduced) may be included in a SAMECS analysis via a pre-merge feature.

3.1.4 SAMECS Solution Capabilities

Both the multi-processing and the full core version of SAMECS will produce: 1) internal loads, stresses, deflections, and reactions, 2) reduced stiffness matrices up to 1000 x 1000 for a substructure or interacted group of substructures, 3) reduced flexibility matrices up to 1000 x 1000 for a substructure or interacted group of substructures, 4) interaction of substructures, and 5) pre-merge, which allows substructure interaction within SAMECS.

Equilibrium checks comparing the equilibrium of input loads and resulting reactions, conditioning checks on every freedom of the stiffness matrix, stiffness matrix maps, run diagnostics, and timing information for each phase of the SAMECS solution are reported automatically for all runs.

SAMECS uses the SOLPAC (SOLution PACkage) solution program based on Choleski decomposition.

3.1.5 SAMECS Output Communication

SAMECS outputs all of the data described in Section 3.1.4 in both printed form and on multi-file output tapes. In addition to this data, all input data, element transformation, stiffness, and loads matrices, and gross stiffness matrices are available on request. The multi-file output tapes are used for communication with the SAMECS peripheral programs.

3.2 SAMECS PERIPHERAL PROGRAMS

A system of peripheral programs was developed to interface with SAMECS. The purpose of these programs is to facilitate the preparation of mathematical models, the interpretation and application of the results, to add certain analysis capabilities of an iterative nature, and to permit interaction analysis using substructures. The SAMECS peripheral programs are therefore capable of dealing with any problem which the SAMECS program can treat.

Interfaced programs were developed for use on the SST for several reasons; 1) there is no overhead associated with an executive system, 2) the CDC 6600 system has in effect already built in an executive control system which allows the stringing together through the use of control cards one or more of the SAMECS system programs, 3) checkpoint restart options are automatically available when interfaced programs are used, 4) problems which may develop in one program do not effect any of the other programs, 5) a string of activities may be conveniently grouped together to make a computer run of reasonable run time

depending upon the urgency of the problem, availability of computer resources, and confidence of the analyst, 6) it is logical to break programs where engineering checks and/or decisions routinely occur, 7) new programs having increased capabilities are used in conjunction with SAMECS and all of the SAMECS peripherals by simply complying with system interface requirements. It is pointed out that some of the above points may be equally applicable to integrated (or executive system controlled) programs. A chart showing the flow of data through the SAMECS system of programs is shown in Figure 3.2.

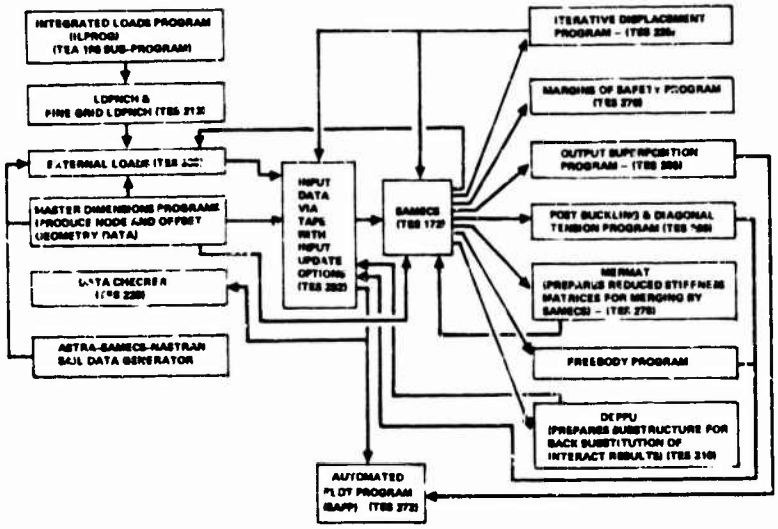


Figure 3.2: THE FULLY INTERFACED SYSTEM OF SAMECS FINITE ELEMENT PERIPHERAL PROGRAMS USED BY THE SST STRESS GROUP

3.2.1 SAMECS External Loads Transformation Peripheral Program

The SAMECS External Loads Transformation Program completely eliminates the need for the engineer to manually prepare nodal loads data for SAMEC runs. It furthermore provides a means of positively guaranteeing the quality of the nodal loads data to management prior to a SAMECS run, and it provides complete reports on the formulation of the loads data so that future questions regarding the source of loads data are easily and completely answered.

The SAMECS external loads program performs a multiplicity of tasks. These tasks are accomplished in four distinct sections of the program. These four sections were used to prepare virtually all of the loads data for SST mathematical models. The sections: 1) transform local loads of all types to the structural grid (including mechanical, pressure, and inertia type loads), 2) distribute loads outside of the domain of mathematical models to the nodes on the boundaries of the models using engineering theory, 3) perform various types of shear and moment calculations using the node loads computed in 1) and 2) as the source of data for the integrations, and 4) scale unit load conditions to form actual load conditions, convert loads into symmetric and anti-symmetric loads where structural symmetry is taken advantage of, combine, scale, and/or correct unit load conditions, and merge reduced loads from substructures for substructure interaction with SAMECS. This program is fully interfaced with loads group programs for automated data transfer of actual loads scalers. A more detailed description of these capabilities is given in Section 3.2.1.1 through 3.2.1.4.

3.2.1.1 External Loads Peripheral Program Transformation of Local Loads Option

The transformation of local loads option is of primary importance in reducing the engineering effort required to prepare loads input to SAMECS. Input consists of the SAMECS node coordinates and a list of nodes, beams, and triangular and/or quadrilateral plates which are loaded. Loads may be applied as forces (force vectors or moment vectors in space), pressure acting on panels, pressure acting on projected areas of panels, and loads acting on beams. The program: 1) equates a uniform pressure load (pressure may be specified as other than normal to the surface and to act on the true or projected area of the panel), on a triangular or quadrilateral (may be warped without loss of equilibrium) panel to an equivalent system of forces at its corners, or at any other group of points (not necessarily the corners of the panels); 2) equates beam loads and point loads and moments (vector direction of loads may or may not be required depending upon option selected) to an equivalent system of forces at the ends of the loaded elements or at any other group of points; 3) performs the same operations as described in 1) and 2) except computes weights of elements from weights data supplied by the weights group and then computes for each of the rigid body unit accelerations the equivalent inertia forces induced on the corner (or other) node points; 4) sums all element corner (or other) forces at nodes for each load or unit inertia conditions automatically and stores results with an identifying unit load case number, 5) performs equilibrium checks of all load and unit acceleration conditions formed (including necessary summations for calculation of centers of gravity of mass groups), 6) calculates for the air loads group the projected areas and centroids of projected areas of air loads group panels on the three-dimensional surface.

At early stages of structural development, finite element analysis air loads are usually received from the air loads group as generated on a two-dimensional aero-elastic grid balanced by vertical and pitching accelerations. At more advanced stages forces and accelerations in all directions are of interest, including the three-dimensional slope of all aerodynamic surfaces. Achieving a balance for the two-dimensional problem is straight forward since the two-dimensional aero-elastic panels are overlaid on the structural panels and using the "LINES AND PLANES" program precise transfer of the air paneling systems pressure load to the structural paneling system and integration is accomplished. However, when a three-dimensional balance is undertaken, then the air loads group must be given the precise centroids and projected area of the structural panels which are equivalent to the aeroelastic paneling system. In this way drag and lateral forces and the effect of surface slope is taken into account in the balance, and the accuracy of the total effort is assured. The transformation of local loads option produces the required information on punched cards for the air loads group in a format compatible with the balance programs.

The transformation of local loads capability is the key to balanced load conditions, since it permits exact transformation of loads without incurring any of the engineering effort that would otherwise be associated with such precision.

3.2.1.2 External Loads Peripheral Program Transformation of Loads to Model Boundaries Option

The boundary loads option is used to apply loads to mathematical model boundaries prior to the time when an interact is accomplished which encompasses that boundary in a larger solution, or for problems where the inaccuracies of load distribution induced by assuming simple engineering theory are considered to be small either because of the nature of the structural system, or because

the boundaries are sufficiently removed from the region of interest within the model. Loads which are applied to any reference point in space are in turn applied equivalently to all of the nodes in a defined section made up of plates and beams (which may or may not be offset). In addition to calculating the equivalent node loads on the sect! and performing appropriate equilibrium checks on the resulting set for comparisons with the orginally applied loads, this option also prints out section properties and resulting stress distributions in the boundary elements.

Elements need not be oriented perpendicularly to the section or boundary plane since in aircraft structure such an assumption will in general cause an ovalization of the section and different frame bending moments and skin shear flows.

3.2.1.3 External Loads Peripheral Program Shear and Moment Option

In addition to the equilibrium checks which are performed routinely on all load conditions formed, it is sometimes desirable to perform a stepwise integration through one or more load cases, transforming the resulting forces and moments to specific points of interest. A series of integration intervals may be defined using ranges of coordinates, a list of nodes, or a combination, and integration may be obtained approaching the SAMECS model array of nodes from different directions. It is the purpose of the Shear and Moment Option to provide a general tool for obtaining the specific information needed to produce checks and plots of data not fully covered by the automatic equilibrium checks.

3.2.1.4 External Loads Peripheral Program Unit - Actual Loads Options

The heart of the external loads program is the unit actual loads routine. This routine accommodates new load conditions and changes to old load conditions on short notice, and gives the user the check information required to verify the quality of all loads data. This module takes the unit conditions formed by the previous options, modifies them as required, forms new unit conditions from existing unit conditions, adds any special unit load conditions, receives unit-actual load condition scalars on punched cards from the loads groups and scales and adds the unit load conditions to form actual load conditions, breaks the load conditions into symmetric and anti-symmetric components, performs equilibrium checks of actual load conditions about selected points of interest, and performs the merging of reduced loads from several substructures into a single set of loads for an interact run. The resulting node loads card images are sorted in a proper sequence for use by SAMECS and written on tape ready for updating onto the SAMECS input data tape. When interaction between substructures is being accomplished and the node and/or load case numbering differs from that being used in the interaction solution, correspondence tables are read by this option and the revised numbering automatically assigned to all nodes and load conditions.

3.2.2 SAMECS SAIL (Structural Analysis Input Language) Peripheral Program

The SAIL program is a structural analysis input language which allows the user to take advantage of repetition in the numbering, sizing, and geometry in the preparation of the structural elements comprising a mathematical model. Node, plate, beam, and some loads data may be generated using SAIL. Equations may be written for, or tables input directly in, the SAIL program to assign values to the various required input parameters of SAMECS. The SAIL program is particularly useful for generating super refined grids where repetition in the mathematical

modeling is prevalent. Models constructed in part or in total with SAIL may be modified in ways that result in completely new data sets but require only small changes to the original SAIL data. For example, quadrupling the refinement of a mathematical model grid may require only a one card change in SAIL, but it may result in hundreds of new card images being produced on tape for SAMECS input. The SAIL program is also capable of outputting data decks for ASTRA and NASTRAN finite element programs using a SAIL input deck that orginally produced a SAMECS data set.

3.2.3 SAMECS Iterative Displacement Peripheral Program

The iterative displacement program is used in conjunction with SAMECS to correct equilibrium equations of any finite element mathematical model for the effects of large displacements. The program has been used successfully to treat a large number of structures for which the rotations of the elements due to deflections induced by loads and temperatures were important in increasing or decreasing the stresses in the finite elements. Included in the list of analyses performed recently using the method are: 1) the 747 fiberglass honeycomb wing leading edge panels, 2) the SST cab acrylic window analysis, 3) several analyses of skin-stringer, integrally stiffened, and sandwich panels, and 4) a beam-column type improved stress analysis in the aft body of the SST immediately aft of the rear spar. Analysis results have been substantiated by test for the 747 fiberglass honeycomb wing leading edge panels and the SST cab acrylic window.

3.2.4 SAMECS Node and Offset Data Generation Programs

The geometry of SAMECS mathematical models is entirely defined by node coordinates and element offsets from the nodes. This data is generated in many ways, occasionally manually, but more frequently by SAIL, one of the special purpose geometry generators (such as that for windows), or finally and most frequently by a library of general purpose geometry programs developed and used by a project support group called Master Dimensions.

SAIL is used where the shape of the model being treated may be adequately described using mathematical equations for the surfaces, planes, lines, and intersection points involved. The engineer codes these equations directly in SAIL language and can have the resulting nodes and offsets back the next day. Time-sharing teletype terminals are frequently used for small programs to allow immediate results in support of a modeling effort.

Specialized programs have been written by Master Dimensions and turned over to the structures groups for their use. The Master Dimensions "window" program is a good example. The window program allows the user to stipulate: the thickness and corner radii of the several panes of glass, the thickness of the rebate, the thickness of the seal material around the edge of the glass, post and sill geometry (where the post and sill is modeled as a built up section of plates and beams), hard points such as intersection points of posts with sills and stiffeners with posts and sills, and finally the fineness of the grid into which the window panes are to be divided. From this information the Window Master Dimensions program generates node coordinates for the glass, seal, posts and sills assembly. All coordinates may be skewed to conform precisely to the true geometry of the vehicle and facilitate interaction of window with surround and/or imposition of surround displacements on the post and sill structure. Because a Master Dimensions program is used to generate this node data, it is known: that the seal springs which transmit shear and moment to the edge of the

glass panes are precisely normal to the plane of the glass, that the glass pane is precisely flat, and that the loading which is dependent upon the node coordinates is accurately and precisely applied. Small inaccuracies in the node coordinates of the window analysis could result in large errors in internal loads and displacements. Similar programs are used to obtain refined grids in the corners of cutouts.

As the project definition of hardware to be analyzed evolves then the Master Dimensions group geometry definition which is used for preliminary lofting, firm lofting, drawing release and manufacturing can be referenced to obtain node and offset data for the structural analysis. Since the primary concern of the finite element structures analyst is not a tight tolerance on the accuracy of the geometric definition (a .1" difference in the radius of the fuselage monocoque will cause no significant change in internal loads) but a precisely smooth definition (a .1" joggle in the surface of the fuselage monocoque over a .5" arc length can ruin the local results, is difficult to detect, and costly to correct). Master Dimensions is used almost exclusively by structure analysts performing medium and large size analyses using SAMECS. Data supplied by Master Dimensions to the analyst is node coordinates on magnetic tape or punched cards, surface normals on magnetic tape, and direction cosines of the lines of intersection of structural element intersections such as frame web with stringer web, again on magnetic tape.

Offset data is generated by an offsets program using SAMECS structural data, section y bars, and the magnetic tapes supplied by Master Dimensions. Some additional comments on the function of Master Dimensions are made in Section 4.3.1 dealing with quality control of the mathematical analyses.

3.2.5 SAMECS Freebody Peripheral Program

The Freebody Program acts as a data extractor in that it extracts from any SAMECS solution a selected subset of the data which constitutes a complete SAMECS data set within itself. Deflections solved for in the previous SAMECS solution may be automatically imposed on the boundaries of the subset structure. Reactions obtained from the subset SAMECS run are the freebody loads. Furthermore, the new input data tape formed by the Freebody program may now be altered in sizing, refinement, heating, geometry, and/or loading and with the reactions applied as loads the subset may be reprocessed for a SAMECS solution, or the altered input data tape may be used to update the original SAMECS input data tape so that the most current information in that particular region of the structural system will be included in the next run of the original model.

3.2.6 SAMECS Substructure Interaction Programs, Including DEFPU and MERMAT

A capability to substructure (or partition) the mathematical model of a complex structural system was of key importance to the SST structures staff in obtaining adequate internal loads analyses simultaneously with the development and drawing release of the SST prototypes. Substructuring and substructure interaction was developed for several reasons:

1) Use of substructures to idealize structural systems permits a detailed idealization within each substructure at the same time a comprehensive analysis of the total structure is obtained.
2) Structure not well defined may be modeled, using a coarse grid idealization, and used to accomplish, as desired, interaction with fine grid substructures for purposes of improving the boundary conditions of the fine grid models.

3) All information required to idealize a large section of structure may not be available at the time an analysis of some portion of that section is required. Use of substructures permits setup of each portion as design information comes available and also allows interaction at a later date with the whole structural system.
4) Substructure analysis reduces the waiting period for large size idealizations, setup, and execution, and spreads out the model preparation and data reduction effort more uniformly over the design and drawing release periods.
5) Failsafe analysis is economically accomplished using the substructure interaction approach.
6) Computer program modules may be designed to handle a certain fixed size substructure which is primarily a function of the computer system used, and run times may be appropriately optimized around that size without affecting the generality of the type, size, and treatment of structures problems by the analyst.
7) Run times on computer equipment may be broken into several segments of controllable size, and with engineering check points for quality control between the segments: to allow verification of analysis progress and quality, to eliminate the possibility that a very large block of computer processing time may be lost due to computer equipment malfunction, and to allow engineering access to valuable internal loads and deflection data prior to obtaining the fully interacted final result.
8) The flexibility of an interact system, properly designed, is such that literally hundreds of alternative routes through the interaction procedure are available. This provides an analysis system which has the best opportunity of satisfying each technical and scheduling problem in an optimum manner. For example, options are available as to: sequence of data processing of the substructures, whether only stiffness or stiffness and load interaction is required, which of the substructures will be back-substituted from boundary deflections to obtain internal loads and deflections and for what load cases, which substructures will not require reduced stiffness matrix updating in a given cycle, whether a full or a partial interaction is required, whether a pure interaction or a hybrid interaction is required (a hybrid interaction is one in which certain elements are included immediately in the interaction run and internal loads and deflections are obtained directly from the interact run for these elements), whether fail safe or cracked structure considerations should alter the interaction procedure, and whether first, second, third, etc. level interactions are to be employed.

Interaction of substructures is accomplished through the use of four key programs: SAMECS, External Loads Peripheral, MERMAT, and DEFPU.

SAMECS pre-processes each substructure to obtain a reduced stiffness and a reduced loads matrix. MERMAT merges the reduced stiffness matrices from the several substructures using a correspondence table for nodes. The External Loads Peripheral Program merges the reduced loads matrices from the several substructures using the same node correspondence table and a load case correspondence table. SAMECS reads the merged matrices and solves for the interacted boundary displacements. DEFPU imposes the interact displacements on the boundaries of each substructure using the same correspondence tables used above, and SAMECS is used to calculate internal loads and deflections within each substructure.

The MERMAT program has the capability of transforming stiffness matrices to

simulate the effects of pivoting and/or translation of a substructure relative to other substructures, or to accommodate interaction of substructures having different coordinate systems.

3.2.7 SAMECS Input Data Sort and UPDATE Peripheral Program

The UPDATE program eliminates the need for all deck sorting, sequencing, and stacking for SAMECS runs. All cards or card images on tape produced by any of the peripheral programs to SAMECS have entered in card columns 73 through 80 a uniquely defined sequence number. These numbers are also carried on all of the standard input forms used in preparing SAMECS input data manually. This system together with the total elimination of node loads greatly reduces the chance of engineering error in forming, making additions, corrections, or deletions to SAMECS data sets on tape.

In addition to sorting, adding, correcting, and deleting SAMECS data sets, the UPDATE program automatically provides and inserts all blank cards, data delimiters, and standard header cards with sequencing information entered for the user.

3.2.8 SAMECS Data Checker Peripheral Program

The Data Checker Program is used to check the input data before a SAMECS run is attempted. If the SAMECS data set is accepted by the data checker, then the successful execution of that data set using SAMECS is guaranteed, and it is almost certain that scheduled computer time for the run will be efficiently and profitably used. It checks for the logical sequence of SAMECS input data sections (including nodes, plates, beams, and loads) and for the logical sequence within each section. The sequence numbering in columns 73 through 80 is checked along with the format of each field of each data card. The program will then simulate the steps which SAMECS performs in reading in the data, screening it, and computing element matrices. It produces the identical diagnostics that SAMECS would produce, but does not generate the matrices. This requires 20 to 25% of the residency processing time normally required by SAMECS in the SAMECS generate phase saving a small amount of computer time, but more important, because of its small core requirements, rapid turn-around is assured.

3.2.9 SAMECS Automatic Plot Peripheral Program

A major contributor to the quality of SAMECS analyses is the automatic plot program. Although the Master Dimensions group performs numerous plots of the geometric data which that group generates for SAMECS, the final check on node geometry and the element orientation and connectivity to that node geometry must be made using exclusively the SAMECS input data deck. Equally important is the capability to display the solution, observing trends in load paths, insuring that there are no undetected "soft spots" in the analysis, and finally for use by the strength, fatigue, fracture, mechanical systems, loads, performance, and design groups for displaying specific information in support of their respective problems.

The SAMECS Automatic Plot program will produce any number of two-dimensional and three-dimensional views constructed from SAMECS input data or output results, or from SAMECS External Loads Transformation Program Elements. The whole structure or any parts thereof may be plotted. Two-dimensional and three-dimensional views of models are plotted using the Cal Comp Plotter. The program also plots internal loads, stresses, and node deflections in any of the standard

or three-dimensional views desired.

Features available are: arbitrary angular orientation and scale of system axes in the displays, automatic compression of plot size to fit a page dimension specified, labeling of nodes, plates, beams, element local coordinate system orientations, arbitrary lettering sizes, labeling of any and all internal loads data produced as output by SAMECS, and construction of three-dimensional deflection plots. Data produced for the SAMECS Loads Program, for and by SAMECS, and by the SAMECS Output Superposition Program may be plotted.

3.2.10 SAMECS Output Superposition Peripheral Program

The output superposition program allows superposition of unit or actual load cases processed by SAMECS using appropriate scalars. Plate and beam element stresses and internal loads, and node deflections and reactions may all be superposed in the same run as required. Different loadcases in the same SAMECS run, or same or different loadcases in different SAMECS runs of the same structure may be superposed. The program is commonly used to combine critical design flight conditions with appropriate pressures and thermal environments, to form left-and right-side results from symmetric and anti-symmetric results, to combine load conditions appropriately for use with the post-buckling peripheral program, to ratio loads from limit to ultimate, and to perform max-min critical condition searches of the superposed data. Max-min data is printed at the bottom of each element's superposed load conditions, and also by itself as a max-min envelope package. Output data on tape and in printout is in precisely the same format as SAMECS output data, so that peripheral programs may operate with the superposed data as well as the original SAMECS data.

The Output Superposition Program is frequently used to prepare a special set of data required by a specific group, eliminating those conditions not of particular interest. In this way the user is able to focus quickly on the pertinent data for his problem, and the amount of printout data retained in each area is minimized. The SAMECS Margins of Safety Peripheral Program is even more selective in the superposition and display of results.

3.2.11 SAMECS Margins of Safety Peripheral Program

The Margins of Safety program is used for detecting and displaying critical stresses in SAMECS beam and plate elements. This program is capable of: 1) extracting internal loads and section properties for selected beam and plate elements from one or more SAMECS output tapes dealing with the same structure, 2) preparing design loads by superposing the SAMECS internal loads using specified scalars, 3) selecting the proper allowable stress for each design case corresponding to the thermal condition included with that design case, 4) calculating the fiber stresses and margins of safety in the extreme fibers of the selected beams, 5) calculating plate element principal stresses and margins of safety for selected conditions, 6) analysis of beam elements with section properties different from those used in the last SAMECS solution, 7) presenting the input and output data with various printout options in an easy-to-read format, 8) x-y plotting of beam unit moments, axial loads and shears as ordinates, and beam numbers as abscissas, 9) x-y plotting of beam section properties, design loads, and stresses, and 10) execution of several sets of beams and plates in one run.

3.2.12 SAMECS Post-Buckling (Including Diagonal Tension) Peripheral Program

For structural systems where the buckling of the structure under load causes

major redistributions of internal loads the SAMECS Post-Buckling program is most important. There is little point in producing internal loads and deflections for a complex structural system by a detailed finite element analysis procedure and then leaving the engineer to his own devices to correct the solution for buckling effects which when incorporated alter the solution or even change the signs of the internal loads in many of the structural elements.

The Post-Buckling program treats post-buckling effects (including diagonal tension) in SAMECS plate elements, using the constant initial strain ($\epsilon_x, \epsilon_y, \gamma_{xy}$) inputs available for SAMECS plate elements, in an iterative procedure. The program evaluates each plate element's degree of buckling for each load case and determines the corresponding initial strains for input to SAMECS. It also prepares symmetric and anti-symmetric pre-strains where principles of structural symmetry are taken advantage of in the modeling. Following are the important aspects of the program:

1) Post-buckling can be handled by a series of SAMECS iterative runs using the initial strain capability for each run. Iteration gives the model opportunity to redetermine load paths due to compression buckling and diagonal tension effects, and to re-evaluate the buckled state of plates in as many cycles as are required to achieve the desired degree of convergence. A closed solution is obtained which exactly meets force equilibrium and deformation compatibility in the model since the SAMECS program is used to process the pre-strain plate data.
2) Compression buckling in one direction is permitted without affecting the stiffness in the other direction.
3) Compression buckling in both directions is permitted with different stiffnesses, respectively, which correspond to the degree of buckling.
4) Shear buckling is permitted, which produces diagonal tension in plates.
5) Combination of compression and shear buckling with tension present is permitted.
6) No change in the stiffness matrix of the structure is required in the SAMECS program. Consequently, all load cases may be treated simultaneously in a single series of iterations.

3.3 PROGRAMS INTERFACING WITH THE SAMECS SYSTEM OF PROGRAMS

A large number of programs too numerous to list or discuss here, are interfaced with the SAMECS system of finite element programs. They encompass all of the technology areas outside of the Stress group, and include static and dynamic loads, flutter, stability and control, performance, and aerodynamics groups. Some of the programs provide data directly to the SAMECS system. Such data might be aeroelastic panel geometry, panel pressures, load factors, powerplant thrust, etc. Other of the programs may require data from the SAMECS system such as deflections, internal loads, reduced flexibility matrices, and reduced stiffness matrices. Several eigenvalue programs, dynamic response, transient response, aeroelastic, and stability and control programs may at the option of the analyst use SAMECS originated data. The SAMECS system multi-file concept embraces a generality which encompasses or can be easily made to encompass all of the technical programs which would have reason to use SAMECS originated data. Furthermore, as programs are developed or improved within the various technology areas, there is sufficient documentation available on the SAMECS system to allow these programs to maintain a fully interfaced condition with the SAMECS system where applicable. Figure 3.3 shows technology group computer programs and the flow of data between groups across program interfaces.

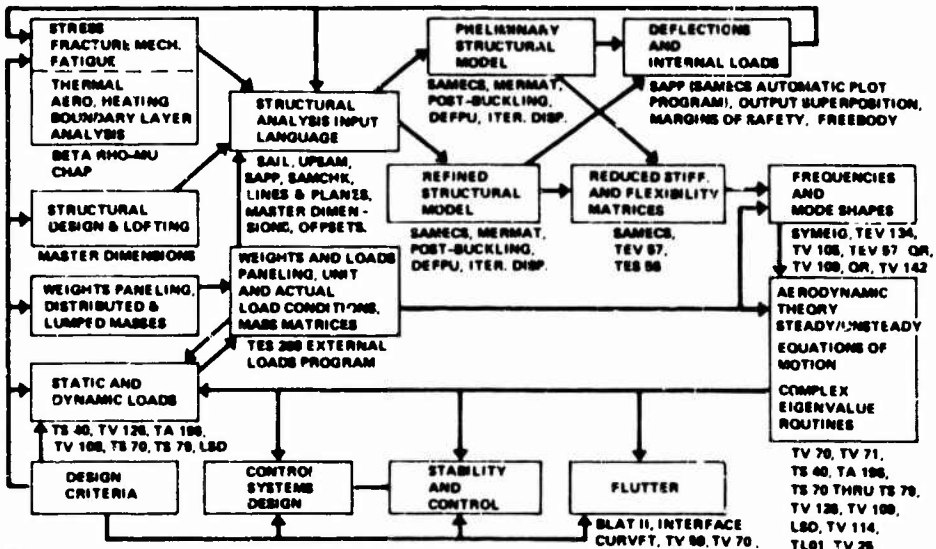


Figure 3.3: TECHNOLOGY GROUPS AND INTERFACES, WITH CORRESPONDING COMPUTER PROGRAMS (PROGRAMS LISTED BELOW BOXES ARE USED TO ACCOMPLISH ACTIVITIES INSIDE BOXES)

4.0 MANAGEMENT OF FINITE ELEMENT PROGRAM USE

The attention to detail required of management and engineering to produce analyses of the behavior of complex structural systems has become more demanding as the scope and level of detail of the analyses has been extended. Members of Boeing Structures Management recognize that the solution of the algebraic equations for the structural internal loads system and the peripheral processing associated with preparation and reduction of the analysis is only one part of the total task. Important to the success of the analysis effort is the involvement of engineering managers who have a thorough understanding of the tools being used and their application to the design task, and who recognize that only with proper emphasis on the disciplined development and management of the total problem are we able to use the computer to the extent that the most up-to-date computing facilities allow. Without a well coordinated and thoroughly integrated effort with a specified end product deliverable on a committed and realistic schedule, the results obtained may lose much, if not all, of their value.

A set of management and technical procedures were developed at Boeing for the SST program and routinely employed in treating the SST structure. Standardized procedures for construction of the finite element mathematical models and for operation of the finite element programs were adopted. Documentation and standardization of all aspects of the operating procedures for the SAMECS system was developed and completed. By following these procedures Boeing was able to assure consistency throughout all phases of solution, proper record keeping of all activities preparing finite element models, timely and efficient coordination of model construction and associated data transfer, quality control of solution results, and complete documentation of analysis results to give a complete technical historical perspective to design decisions.

Management procedures adopted relate to: 1) preparation prior to undertaking the finite element analysis effort, 2) organization required to implement modeling activity, 3) execution of the solutions with thorough quality control of all

facets of the analysis effort and 4) data transfer of the results to the technology groups relying on the finite element information. Each of these aspects of the management and technical procedures are discussed in Section 4.1 through 4.4. A brief discussion of the scope and cost effectiveness of finite element methods for the SST is presented in Sections 4.5 and 4.6.

4.1 PREPARATION PRIOR TO UNDERTAKING THE FINITE ELEMENT ANALYSIS EFFORT

Management preparation includes:
1) Adoption of a Finite Element modeling philosophy,
2) Establishment of a mode of management control,
3) Standardization of technical procedures, and
4) Adoption of modes of information transfer between Technology Computer Programs.

4.1.1 The Finite Element Modeling Philosophy Used on the SST

Management adopted an overall finite element modeling philosophy for use on the SST program. This general philosophy was adopted to provide guidelines for construction of models, and was modified as necessary to fit the objectives and timing of the various phases of the SST prototype development and drawing release activities. The modeling philosophy adopted stated that:
1) Substructuring of the structure would be required,
2) Six-degrees-of-freedom would be used at all nodes in the finite element grids,
3) Lumping or delumping of primary structural elements would be avoided, every physical structural element would be modeled with corresponding finite elements adequately representing the discreteness, stiffness, and connectivity of the actual structure.

This philosophy ground rule regarding lumping and delumping merits further comment. Lumping and delumping of structural elements were avoided where possible because: such activity is time consuming, it acts to degrade the quality of the mathematical model, and it makes it more difficult to later use the plate element post-buckling and diagonal tension program. (This is so because the buckling loads are a function of the geometry of the plate elements or element groups and if the finite element plate isn't a reasonably good representation of the shape of the actual platethen special parameters have to be prepared for the post-buckling program furnishing this information for each plate which is different. If there is good correspondence, then the handling of the plate elements to be buckled is automatically accomplished by the computer program from the basic input data.)

It is important to note that only because of the large problem size capabilities and improved efficiencies of the SAMECS program were we able to avoid lumping. SAMECS was applied using this philosophy on virtually all of the SST wing and body primary structure.
4) Key section properties of stiffener elements will be modeled whether it appears that they are significant to the analysis or not, i.e., A, I_x and I_y, and j, will be included in beam elements.
5) Additional subelement refinements will be used in areas of rapidly varying stress gradients to produce the required results, for stress, fracture and fatigue groups.
6) Only balanced load conditions will be analyzed.
7) Master Dimensions will be the source for geometric data for fine and superfine grid finite element models.

8) Engineering will follow a strict set of technical and management procedures. Specific quality control checks will be made on all analysis activity, specific minimum standards of quality will be met, and summary reports on quality will be published for all phases of modeling activities.
10) Non-linear, large displacement, and post-buckling and diagonal tension capabilities will be used at the finite element level to answer as required the questions of effectivity and overall stiffness of structural elements and models.
11) Joint stress severity factor analysis will require the modeling of discrete fasteners in interacted or partially interacted analyses to allow optimum joint design.

4.1.2 Established Manner and Modes of Management Control of Finite Element Modeling

Management decisions were made at the inception of all internal loads modeling activities in support of SST structural analysis efforts. These decisions related to the following aspects of the problem: 1) technical interfaces, 2) sign conventions, 3) coordinate systems, 4) adoption of standard input forms, 5) adoption of quality control measure, 6) adoption of procedures, 7) implementation of the system by which management would review modeling activities to insure compliance with their requirements and procedures, 8) definition of items needing management approval before committing runs, 9) definition of the level of detail, type of forms, and cataloging system to be used for recording all math model inputs inlcuding geometry, sizing, loads, and stress calculations in support of model development, 10) assignment of math models to specific stress group engineers who would be responsible for the structural input to the models, assembly, loading, quality assurance, selection of output display, overall coordination of interactions of substructures, overall coordination of data transfer between technology groups, direction of computer programming, and consultation on finite element modeling in total, and 11) assignments for scheduling of the major milestones finite element schedule, determination of manpower requirements, and determination of computer budget and monthly use required to support the major milestones schedules.

Where management procedures adopted for the SST called for specific management approval, the SAMECS system was designed to offer management optimum visibility of key steps taken by the computer, and quality control check sheets were designed to make it impossible for engineering to continue beyond certain points without recognizing the management requirement for approval.

Finally, SST management provided data storage and display areas to provide maximum management and technical staff visibility of the analysis efforts at all stages of development. In these areas were kept math model node and element diagrams, interact, weights, and air loads paneling diagrams, major milestones schedules, computer data management charts, input and output computer data logs, tapes, cards, printouts, plots, and geometry, stiffness, external loads, and structural sizing references.

4.1.3 Standardized Technical Procedures for Construction of Finite Element Models

Standardized procedures for engineering to follow in the construction of SAMECS models and for the operation of SAMECS and the peripheral programs to SAMECS were developed for use on the SST program. These procedures were adopted to assure that all associated with the definition, scheduling, preparation, and application of mathematical models were fully aware of the system needs in insuring quality,

compatibility, proper utilization, and adequate documentation. Such procedures were used to acquaint new employees and subcontractor employees with the requirements of using the SAMECS system. Topics discussed in the procedures include ground rules, peripheral programs, technology interface requirements, input requirements, substructure interface requirements, descriptions of specific coordination which must be accomplished on interfaces to assure timely and proper support, descriptions of the modeling philosophy to follow, general comments (including a list of common pitfalls to be avoided), and an item by item description of each of the major milestones on the major milestones schedule.

4.1.4 Information Transfer Between Technologies and Computer Program Interfaces Which Correspond

From a managerial and technical point of view, the technology staffs supporting the design and release of a specific structural system are linked to each other and to the design project in a complex and highly interdependent manner. The various programs which make up a system of fully interfaced programs used to perform the technical analyses of a structural system must similarly be linked to each other and so organized internally and externally that data can be passed through the interfaces with minimum effort, and maximum visibility, control, and responsiveness.

Technology interfaces occur between Project, Master Dimensions, Weights, Static Air Loads, Dynamics, Flutter, Thermal, Aerodynamics and Performance, Stress, Fracture Mechanics, Fatigue, and Materials and Processes. Major quantities of information must be transferred between certain of these technology groups. These include geometry, section properties, weights data, air loads data, stiffness matrices, flexibility matrices, deflections, internal loads, and stresses.

4.2 ORGANIZATION REQUIRED TO IMPLEMENT MODELING ACTIVITY

The organization required to initiate a modeling activity focuses primarily on: 1) the scheduling of the support required, be it configuration, component development and test, layout review or drawing release, 2) the scheduling of the supporting finite element analysis major milestones, and 3) the commitment of computer facilities.

4.2.1 SST Drawing Release Schedules

The SST stress groups used the network of SAMECS and SAMECS peripheral programs to support layout review and drawing release of all of the body and wing structure. The goal in using finite element analyses was to obtain runs early enough in the design process for the results to have a solid impact on weight and fatigue lift. A minimum of three cycles of internal loads runs prior to 25% drawing release was the goal set for the mathematical model analysis group. The sequence of events by which release of structure was accomplished, focusing on the stress group activities primarily, was as follows, using the wing-body intersection structure as an example:

1) Obtain spar, rib, stiffener centerlines; upper and lower surface definition, frame, stringer, and cutout locations, construct additional finite element grid lines, and request from project loft Master Dimensions group the resulting node coordinates, offsets, and related geometry in such format that it is directly usable by the SAMECS system,

2) Obtain surface temperatures, structural element gradients throughout the substructure from the thermal group; external pressure distributions and associated load factors on punched cards from the loads group, unit and actual weights data from the weights group,
3) Add sufficient information to the data received in (2) to build and insure balance of the critical design load and nominal flight conditions of interest,
4) Calculate internal loads, deflections, stresses, and stiffnesses of structure using SAMECS,
5) Perform a preliminary sizing of the structure from the layouts using the results of this coarse grid finite element solution, and stiffness requirements,
6) Provide nominal flight condition deflections to project for preliminary deflected shape information to be used for jigging the vehicle,
7) Compare deflections and stiffness of coarse grid finite element models with those of dynamics, flutter, and loads groups to verify compatibility of stiffness representations arrived at using different fineness SAMECS models,
8) Supply reduced stiffness matrices to flutter, dynamics, and load groups for use in the next design data cycle in support of final drawing release,
9) Supply sizing of layouts to project for update of drawings, and to weights for update of weights data,
10) Refine the finite element grid to obtain better internal loads distributions in areas of high stress concentration and including the effects of post-buckling and diagonal tension directly in the finite element analysis. Use the SAMECS substructuring and freebody options to treat specific components, joints, etc. in greater detail as required.
11) Obtain new centerlines, node coordinates, and geometry corresponding to the refined finite element grid from the project loft Master Dimensions group,
12) Obtain updated thermal, weights, loads, and stiffness data from the respective staff and project groups,
13) Add sufficient information to the data received in (12) to build the load conditions of interest for final drawing release, jigging, and stiffness comparisons,
14) Calculate internal loads, deflections, stresses, and reactions of the structure using the refined finite element mathematical model,
15) Perform, in support of final drawing release, necessary stress, fatigue, safe life, fail safe, and stability analysis using the results of the refined model and recycles as required. Release drawings. Verify deflected shape for jigging with project, and stiffnesses with flutter, dynamics, and loads,
16) After final drawing release, initiate an additional pass through Master Dimensions, weights, loads, thermal, and flutter groups to obtain a final analysis for the structure as released. Steps taken are the same as (10) through (14).
17) Use the results of (16) to verify that all structure as released has positive or zero margins, to identify potential areas of weight saving, to verify stiffnesses used by flutter, dynamics and loads groups, to assist in location of strain gages and correlating with them in test, and to assist in definition of test loads.

4.0.2 Scheduling of Finite Element Analysis Major Milestones

All SST structural modeling activities of medium and large size were scheduled against a MMDMM (Major Milestones for Development of Mathematical Models) schedule modified to suit the requirements of each structural problem as necessary. The MMDMM schedule contains all of the major technical activities,

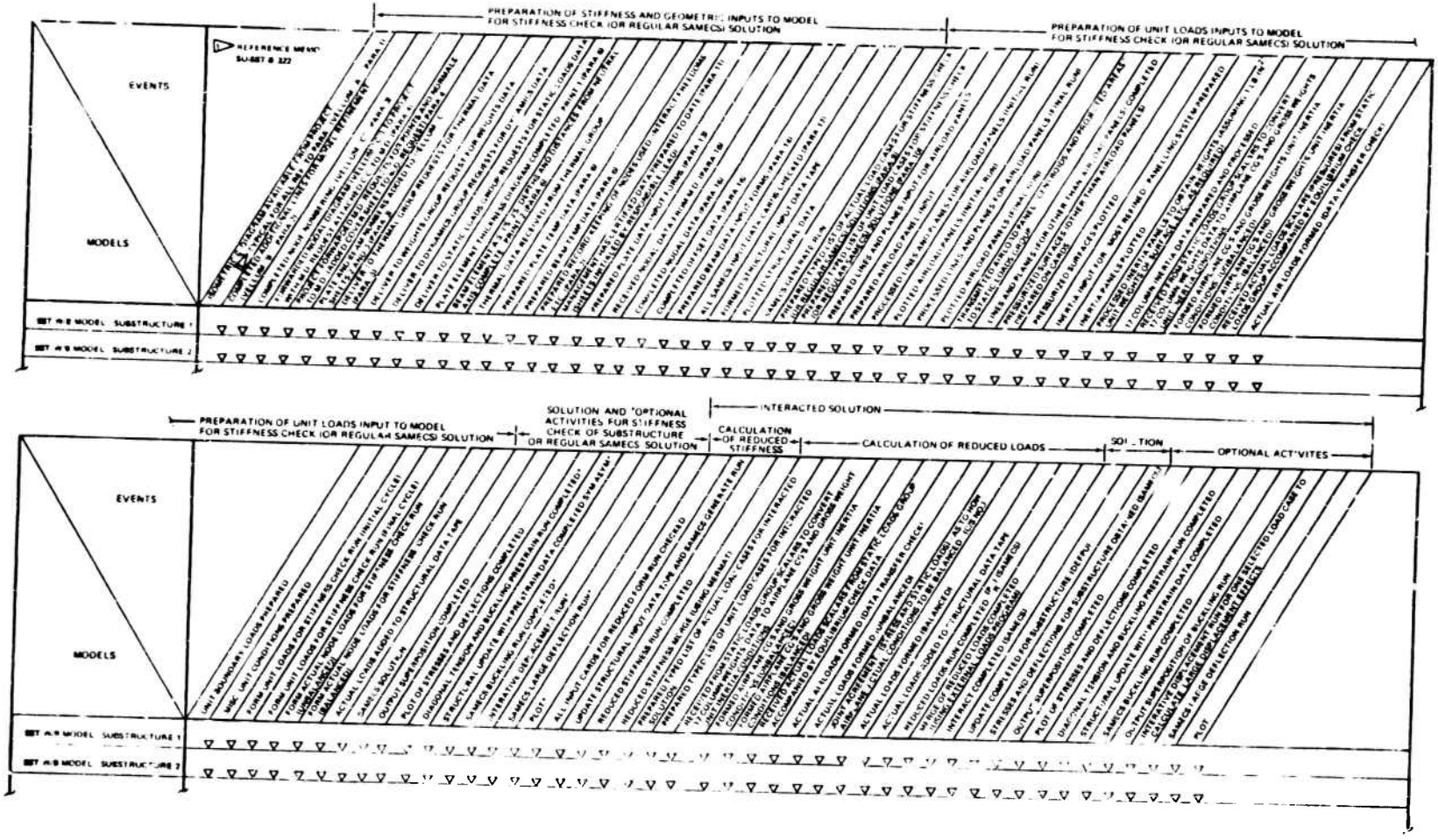


Figure 4.2.2: MAJOR MILESTONES FOR DEVELOPMENT OF MATH MODELS

arranged approximately in the order in which they most commonly occur, which must be accomplished in the development of finite element models. Each milestone in the MMDMM schedule is described in detail as part of the Boeing SAMECS technical procedures. The MMDMM schedule is shown in Figure 4.2.2. The schedule is subdivided into major phases which deal with: 1) preparation of the stiffness and geometric inputs to the model, 2) preparation of the unit loads inputs to the model, 3) preparation of the actual loads conditions for the stiffness check solution on substructures, 4) solution of stiffness check runs on substructures, 5) calculation of reduced stiffnesses, 6) calculation of reduced loads, 7) interact solution, and 8) post-processing activities including superposition, plotting, post-buckling, large displacement analysis, and iterations. There are 92 major milestones listed on the MMDMM schedule for each substructure including the interaction of that substructure.

4.2.3 Commitment of Technical Skills

The SST program retained an average of 6 engineers in finite element research and development for a six year period. Approximately 15 programmers performed finite element coding, consultation, and logistical support over the same time period. At the peak of the SST prototype drawing release phase Boeing employed 45 engineers who were experts in finite element methods and specifically familiar with SAMECS and the peripheral system of programs. This team of engineers produced multiple finite analysis solutions for over 30 different key substructures used to model the entire primary structure airframe. These structures averaged 800 nodes in complexity, and varied from 500 nodes to 2000 nodes. All of the SAMECS peripheral programs were used on these substructures depending upon the type of problems being dealt with. Most of the engineers involved with finite element modeling were also performing strength check and drawing release work. A complex analysis group within the stress group served as consultants in all facets of the analysis planning and execution. The average educational background of the 45 engineers involved was a Masters Degree, and many were equipped with their Doctorates. A definite management commitment to the use of finite element analysis procedures on the SST program was evidenced by the level at which technical skills were committed to the mathematical modeling activities. It was necessary that those used in modeling be well educated, trained, experienced, and familiar with the drawing release cycle in order to gain maximum utilization of the solutions results. On the SST program they were assigned in sufficient strength to allow the careful planning, preparation, checking and cross-checking required to produce quality results.

4.2.4 Commitment of Computer Facilities for Support of Analysis

The computing facility on which SAMECS and the SAMECS peripheral programs operated during the SST prototype layout review and drawing release was a back-to-back installation of two CDC 6600 computers, each with 131,000 core storage and 60 bit words. Three 6638 physical disks containing six logical disks were attached and switchable from computer to computer as required. The facility was equipped with high speed printers, card readers, tape drives, on-line card punch, on-line graphics display, teletypes, and highspeed remote terminals. Off-line data display capability included Cal Comp, Gerber, and SC 4020 plotters.

For approximately two years prior to the SST prototype cancellation the SST technology staffs required a dedicated CDC 6600 facility. Only because engineering management took the care to repeatedly forecast the required levels of computer support and present this picture to upper management was the staff adequately equipped and able to function without severe computing delays.

Without any question, the single most difficult task which faced engineering management was the procurement of adequate computer time, and in blocks of time sufficiently concentrated to correspond to and support periods of intensive technical effort.

4.3 QUALITY CONTROL OF THE MATHEMATICAL ANALYSES

Quality Control of the finite element analysis system is the key to the performance of the system. As larger problems are solved and the associated output data from these analyses increases it is necessary to carefully check all along the solution path to insure quality. Such control also provides a means for quickly detecting, locating, and resolving problems with a minimum expenditure of engineering and computer resources.

Many of the quality control checks in finite element analyses are exactly the same as would be used in any other analysis approach. The computer, however, opens up a large number of additional methods for exercising quality control. In most cases these new methods hinge upon the precision of the computer, which is an inherent by-product of the properly functioning machine. If engineers are precise in the way in which they feed information to the machine, then a lack of precision in the data returned indicates a malfunction or input error, and potential loss in accuracy. Though load conditions may be known to be somewhat inaccurate, they are always processed as balanced load conditions. The smoothness of node and offset geometry produced by a Master Dimensions program may be unnecessary for one load case and vital for another, unnecessary for one grid refinement and vital for a finer grid refinement of the same problem, particularly where the analysis being performed is using linear rather than large deflection analysis procedures. A very common cause of oscillating frame moments is the practice of scaling rather than the precise calculation of the coordinates.

Quality control of the analysis from generation of element sizing and mathematical geometry to final output display of results is necessary. In Sections 4.3.1 through 4.3.4 quality control brought about through the use of: 1) Master Dimensions, 2) equilibrium checks, 3) plotting, and 4) other means will be discussed.

4.3.1 Quality Control Through the Use of Master Dimensions

Procedures for precisely defining the geometry of the structural system through the use of the computer and a numerically controlled drafting machine were used on the SST to guarantee smooth and accurate geometry definition. The group which accomplished this activity in support of project is known as Master Dimensions. Master Dimensions also assists project with the production of drawings and provides manufacturing with jigging coordinates and magnetic tapes for operation of numerically controlled manufacturing equipment. It is possible to produce finite element analyses from the same geometry definitions that are used to produce the hardware, and proper geometric representation as well as smoothness of the structure geometry is assured. All Master Dimensions data is developed through project using project definitions for outside mold lines, centerlines, stringer and stiffener locations, etc. Any changes made by project in the lines are readily picked up by the math models. Quality control checks performed by Master Dimensions include multi-view plotting and comparisons with loft group equations and drawings.

4.3.2 Quality Control Through Equilibrium Checks

Following is the equilibrium check data that is automatically reported by the

SAMECS and SAMECS peripheral programs for use by the analyst in constructing finite element models:

1) unit air loads paneling projected areas and centroids,
2) unit weights paneling weights and c.g.'s,
3) airplane total weights and c.g.'s,
4) individual substructure total weights and c.g.'s,
5) air loads paneling equilibrium for each actual load case,
6) net inertia and air loads for each actual load case about the c.g. corresponding to each actual load case,
7) equilibrium of load cases output from loads transformation program versus equilibrium of load cases input to finite element program,
8) equilibrium of load cases input to finite element program against equilibrium of reactions output from finite element program,
9) force check equilibrium (force check involves applying a unit force or moment for every freedom in the mathematical model and then determining if every freedom load is precisely reacted by the reactions to eight significant digits),
10) thermal load case equilibrium is a computed zero.

4.3.3 Quality Control Through Plots of Mathematical Model Inputs and Outputs

Plots of the mathematical model are made throughout the analysis activity. Typical of the types of plots made are the following:

1) unit air loads panels
2) unit weights panels
3) node and element diagrams
4) element local coordinate systems
5) element connectivity
6) outside contours viewed from several angles
7) deflections of the mathematical model for load cases of interest to examine for soft spots in the idealization and reasonableness of deformations,
8) internal loads and stresses of the math model to determine if load paths are reasonable, to locate and understand stress concentrations, and to determine if the structure as sized is capable of carrying the loads with near zero margins.

4.3.4 Other Quality Control Checks

A uniform thermal expansion condition was processed as one of the load conditions for all SST mathematical models to verify that the elements are stress free (a computed zero). This condition gives the user an idea as to how accurate the internal loads are for the thermal conditions.

The conditioning number (number of significant digits lost in the solution for each freedom) is printed out for every freedom. This data is scanned and if more than four digits of the 14 available are lost, the idealization is scrutinized to determine the reason for the loss and the structure stiffness is corrected to reduce the digits lost to a maximum of four. This was a firm requirement for all interaction analyses.

All nodal diagrams, actual load conditions to be processed, initial reference data for stiffness and geometry definition, skin thicknesses, frame and stringer properties and other structural element properties are signed off by the responsible supervisors or their designees. All input forms and listings are checked. Boundary conditions and retained freedom data are double checked. (At no time is data checked by the same person who prepared the data). A

management review of all models to verify that checks are complete, and to insure that management concurs with the structural idealization and loads is routinely held before data processing of a SAMECS run is initiated. Check sum reports are used to verify that appropriate tapes containing reduced stiffness matrices have been merged correctly. The criteria for superposing load conditions is reviewed and approved by management, as is the grouping of these load conditions for max-min purposes, and of course a check is made to see that the superposition was accomplished as intended.

Each mathematical model has its own peculiar check list selected from a list of over 100 categorical checks for each model or substructure. The wing-body joint analysis of the SST had 98 checks of the type described in Section 4.3 for each of the substructures.

4.4 TRANSFER OF THE OUTPUT FROM SAMECS TO THE TECHNOLOGY GROUPS

Several tools have been described earlier which are used for producing only significant output in a convenient form for use by the various engineering groups. They are the SAMECS program, the output superposition program, the output two-dimensional and three-dimensional plot program, the freebody program, and the margins of safety program. These programs are described in Section 3.0.

The various groups requiring information prepared cards requesting of the various programs the specific information and format required. Depending upon the priority of the request the results of such runs were obtained in from 2 hours to 24 hours on a normal backlog data processing day. This turn-around was in part realized because the SST program had available to it upon demand a computer fulltime. As an additional service to the engineers requiring data, the complex analysis subgroup frequently collected lists of requirements and produced the required results at the earliest possible time following completion of an analysis. The initial data produced by such runs was always processed, bound, and catalogued by that group.

4.5 SCOPE OF FINITE ELEMENT ANALYSES ON THE SST PROGRAM

The scope of the stress group mathematical modeling on the SST program was such that all of the primary structure and much of the secondary structure of that aircraft was treated using SAMECS. The mathematical model philosophy avoiding lumping of structural elements was adopted in virtually all of the body structure and in most areas of the empennage and wing structure. Figure 4.5 shows the major substructures treated on the SST and their approximate sizes. In most of the mathematical models the refinement of the grid was greater than that of the actual structure in order to obtain specific local stress concentration, and to include local large displacement effects. The total number of nodes used in the substructures of the SST airplane numbered approximately 30,000.

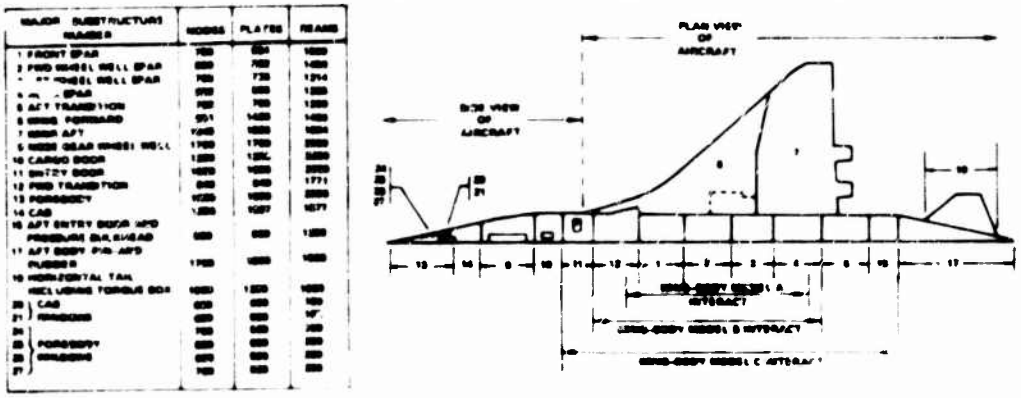


Figure 4.5 SCOPE AND SIZE OF TYPICAL SST SAMECS FINITE ELEMENT MODELS

4.6 COST EFFECTIVENESS OF USING THE FINITE ELEMENT METHODS FOR STRUCTURAL ANALYSIS OF THE SST

The question as to how cost effective the use of the computer is in obtaining internal loads and stress for aircraft hardware was clearly indicated in the course of the SST program. In many areas the benefits of finite element analyses were realized. These areas included the elimination or reduction of the scope of test programs (including quarter or half scale load distribution models and component testing that would have cost millions of dollars and added months of calendar time to schedules). The complex thermal environment, mechanical load and structural system defied effective treatment by any means other than the computer as mentioned in Section 2.0. Major emphasis was placed by management on the reduction of vehicle weight, and on the importance of structural reliability and life.

The question of cost effectiveness may be conservatively answered using only a portion of one of these realized benefits, namely that of weight reduction in the fuselage forward of the passenger compartment aft pressure bulkhead. Six key areas in which weight was saved are:

1)	Sections 43 and 45 frames	285 pounds per airplane
2)	Sections 46 and 47 frames	547 pounds per airplane
3)	Entry doors (6) and cargo door cutout	560 pounds per airplane
4)	Redistribution of load with wing in Section 46	150 pounds per airplane
5)	Redistribution of load with strake in Section 45 lower lobe.	100 pounds per airplane
6)	Reduction in Section 41 cab weight	300 pounds per airplane
		1942 pounds per airplane

Several other areas of weight saving in the body and all areas of weight savings in the empennage and wing are neglected. Boeing management determined that $150 to $200 per pound saved per airplane was a cost effective expenditure. Assuming a 200 airplane breakeven point in development and production costs, this meant that a maximum of $30,000 to $40,000 could be spent to remove a pound of recurring weight in the design of the SST. Based on these figures, and acceptable total cost of eliminating the 1942 pounds saved could have been as much as 58 to 77 million dollars.

The actual computing costs for all data processing by the SST stress group including the body, wing, and empennage sections was less than $1,250,000. If it were necessary to justify the entire stress group computer budget expended on the basis of the weight saved in the six key areas listed, the computing cost per pound of weight saved per airplane would be $3.21 which compares most favorably with the allowable $150 to $200 figures. The direct cost of computer time used by the stress group to save weight in the Sections 43, 45, 46, and 47 frames actually ranged from 15 cents to $2.75 per pound per airplane depending upon whether the frame in question was part of the major wing/body interaction or not.

The idea that the computer which consistently stays busy is cost effective is erroneous. Achievement of this goal also implies infinite turn-around for the unanticipated job. A cost effective use of the computer in solving structural problems requires that the computer system stand idle at times in order to accommodate the more active data processing periods of engineering schedules. In other words, engineering should be ready to pay a premium for guaranteed computing capacity and turn-around.

Computing facilities which are used by engineering to support drawing release schedules should be backed up by an equivalent facility to be used when machine breakdown occurs. It is not reasonable that a large engineering program on a strict schedule and costing hundreds of millions of dollars should count upon the proper functioning of a computer facility without a backup to support engineering and manufacturing schedules.

5.0 APPLICATION OF THE SAMECS FINITE ELEMENT ANALYSIS SYSTEM TO THE SST PROTOTYPE AIRCRAFT

Several mathematical models processed on the SST program serve to illustrate the application of finite element methods on the SST program, see Figure 4.5. Four models are presented: 1) the SST swing-wing wing-body joint analysis accomplished in 1969 before the switch to a delta wing, 2) the SST forward entry door model showing the level of detail treated around cutouts, 3) the SST crew compartment, and 4) the SST wing-body model B interact analysis completed in June of 1970.

5.1 THE SST SWING-WING/BODY JOINT ANALYSIS

A finite element model of the SST swing-wing/body joint was processed in 1967. Every stringer, frame, spar chord, rib chord, and stiffener was included in its offset location and with all section properties assigned. The entire analysis was accomplished within the confines of a single substructure, illustrating the capability of the SAMECS analysis system at the substructure level. The model contained 1899 nodes, 2191 plates, and 2978 beams in the half structure. Forty load conditions were processed including thermal and mechanical loads. The processing time for SAMECS in 1967 was 17 hours. The 1971 version of SAMECS would process the same problem in 7 hours. Figure 5.1 shows the model element diagram.

5.2 THE SST FORWARD ENTRY DOOR ANALYSIS

The forward entry door model illustrates the detail used around cutouts. This 1200 node, 1200 plate, 2000 beam substructure included every structural element discretely, and used still greater refinement in and around the corners of the entry cutout. Skin plate elements in the corners were $1\frac{1}{4}$" by $1\frac{1}{4}$" in size. Figures 5.2.1 and 5.2.2 show the model monocoque and local door cutout detail.

5.3 THE SST CREW COMPARTMENT ANALYSIS

The SST cab analysis included the body shell, window posts and sills, forward pressure bulkhead and floor. A portion of the node-element diagram is shown in Figure 5.3.1. Three-hundred unit-load and 80 actual load conditions were processed. The load cases covered critical design conditions including fatigue.

5.4 THE SST WING/BODY MODEL B INTERACT ANALYSIS

In June of 1970 the B model of the SST delta wing airplane was processed. The [illegible] involved 5939 nodes, 7484 plates, 10,565 beams, and approximately 2500 [illegible] conditions which were superposed to form 80 actual symmetric and anti-[illegible] load cases. There were 2560 interact freedoms on the boundaries of the seven substructures. Loads included ascent, cruise, and descent thermal conditions, model quality check conditions, design and fatigue conditions, and a passenger compartment pressurization condition. Run time for this wing/body interact was [illegible] hours occupancy which included 50 hours central processing time. The SAMECS interact run processed for $3\frac{1}{2}$ hours occupancy. Thirty-thousand simultaneous algebraic equations were solved in the wing/body B analysis. Figure 5.4 is an element diagram for a typical substructure.

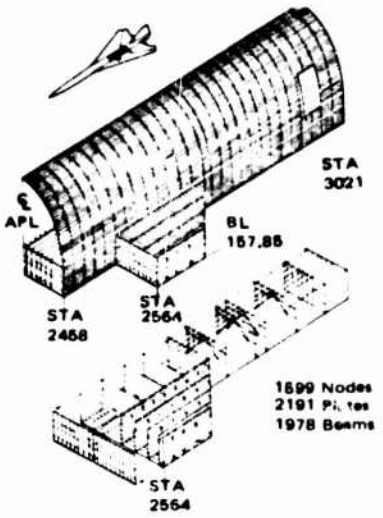


Figure 5.1: SST SWING-WING/BODY JOINT ELEMENT DIAGRAM

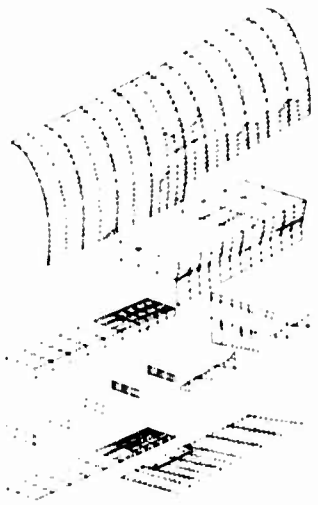


Figure 5.4: SST WING/BODY MODEL B SUBSTRUCTURE NO. 4 ELEMENT DIAGRAM

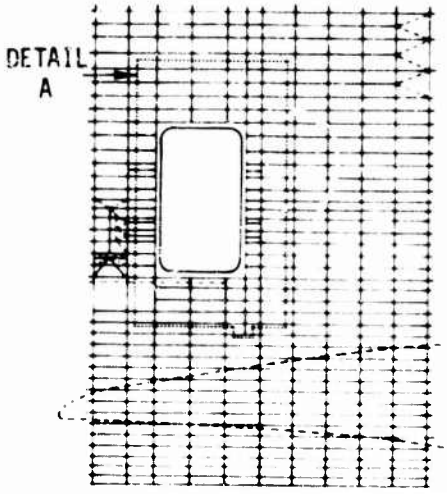


Figure 5.2.1: SST FORWARD ENTRY DOOR ELEMENT DIAGRAM

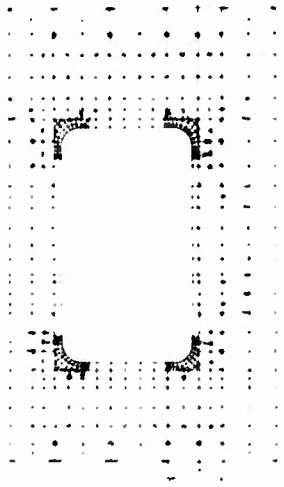


Figure 5.2.2: DETAIL A

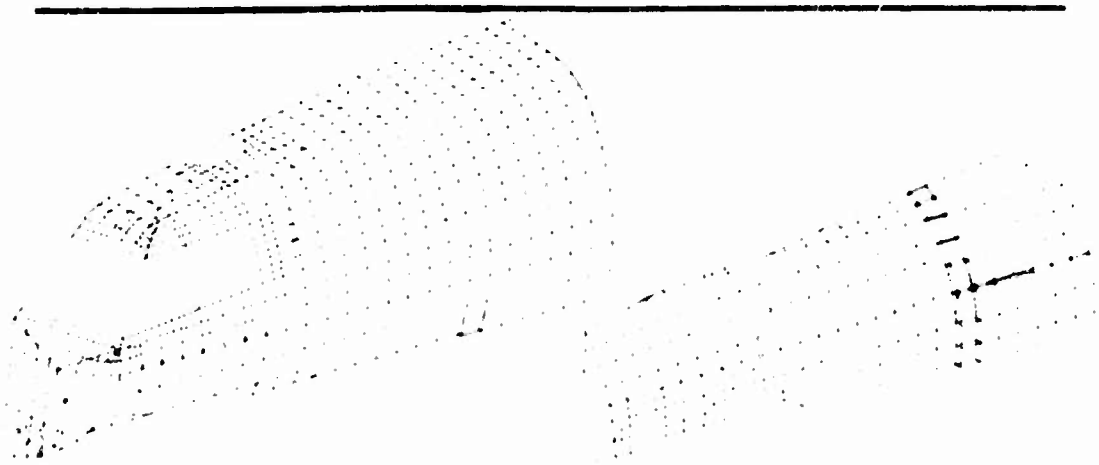


Figure 5.3: SST CREW COMPARTMENT UPPER AND LOWER LOBES ELEMENT DIAGRAMS

ACKNOWLEDGMENT

The author wishes to recognize Melvin A. Nelson and Andrew K. Hepler, members of SST Structures Staff Management, for their dedication to the improvement and extensive application of finite element programs and procedures for the Super Sonic Transport Program. Their commitment was indispensable to the existence of an internal loads analysis system of the quality described in this paper, and to the application of SAMECS to the SST and 747 aircraft.

THE MATRIX DEFORMATION METHOD OF STRUCTURAL ANALYSIS

Paul H. Denke* and Gordon R. Eide**
Douglas Aircraft Company, McDonnell Douglas Corporation

Three new matrix methods of structural analysis are presented: (1) Matrix Deformation, (2) Unified Force Deformation, and (3) Static Force. These approaches are related to the matrix force method in that they require the selection of redundants.

The distinguishing feature of the matrix deformation method is that deformations of elements in the continuous structure, which are initially designated as "statically determinate," are computed before displacements and redundants. A comparison of the matrix deformation method with the matrix force approach shows that the matrix force method is most efficient when the ratio of number of redundants to number of statically determinate forces is small, while the matrix deformation method is most efficient when this ratio is greater than unity.

The order of the matrix to be solved in computing static deformations is the same as the order of the K matrix in the displacement method. Compared to the displacement method the matrix deformation approach offers the advantage of improved matrix conditioning through proper selection of static deformations.

The unified force deformation method provides a theoretical basis from which the redundant force and matrix deformation methods can be derived.

The static force method was investigated and found to be inefficient.

SECTION I

INTRODUCTION

The matrix displacement method of structural analysis requires less computation than the force method when the number of redundants exceeds the number of statically determinate forces. [illegible] the force method has a tendency to yield better conditioned equations and, consequently, [illegible]. The reason for this advantage is that redundants in the force approach can be selected to minimize computational errors. The displacement method has no equivalent operation (Reference 1). Consequently, the force method is a good choice when the structural configuration requires additional accuracy and the number of redundants is less than the number of statically determinate forces.

*[illegible]
**[illegible]

Unfortunately, the extra computations required by the force method limits its feasibility when the number of redundants exceeds the number of statically determinate forces. Therefore, a need exists for a method that combines the speed of the displacement method and the accuracy of the force method in this regime. Such a method would be particularly useful in the analysis of structural models composed of finite elements.

The matrix deformation method provides the required capability. The equations of this method are similar to the equations of the displacement method. The major item of computation is the solution of a set of equations having a coefficient matrix K_{QQ}, which is the same size as the structural stiffness matrix K in the displacement method. The unknowns in the matrix deformation approach are deformations in the continuous structure initially designated as "statically determinate," which are analogous to the displacements of the displacement method. The difference is that statically determinate deformations can be selected to minimize computational errors in a manner analogous to the selection of redundants in the force method.

During study of the deformation approach, two other methods were discovered: (1) Unified Force Deformation, and (2) Static Force. The unified force deformation method is analogous to the unified method of structural analysis and serves as a convenient basis for deriving the matrix deformation equations. The static force method appears to have no utility.

This paper also describes the matrix deformation equations and their applications. The derivations of the methods presented in Appendix I represent an extension and refinement of work reported in Reference 2.

SECTION II

DISCUSSION

The equations of the matrix deformation method can be summarized as follows:

$$K_Q = k\, q_{FI}^I \tag{1}$$

$$K_{QQ} = q_{FI}\, K_Q \tag{2}$$

$$F_I = k\left(q_{RI}^I\, \Delta_R - e_I\right) \tag{3}$$

$$K_{QQ}\, e_Q = q_{F\phi}\, \phi + q_{FI}\, F_I \tag{4}$$

$$F = K_Q e_Q + F_I \tag{5}$$

$$Q_R = q_{RI}\, F + q_{R\phi}\, \phi \tag{6}$$

$$\Delta = q_{F\Delta}^I\, e_Q - q_{R\Delta}^I\, \Delta_R \tag{7}$$

where

k = a square matrix of element stiffnesses (nonsingular, referred to local coordinates)

q_{Ff}, q_{Rf} = matrices of statically determinate element forces and reactions resulting from unit element forces applied as external loads to the statically determinate structure

$q_{F\phi}, q_{R\phi}$ = matrices of statically determinate element forces and reactions resulting from external loads on the statically determinate structure

K_{QQ} = a square matrix of stiffness coefficients

ϕ, e_f, Δ_s = column matrices of loads, unassembled element deformations, and support displacements

e_Q, F, Δ, Q_R = column matrices of statically determinate element deformations (in the continuous structure), element forces, joint displacements, and reactions

$q_{F\Delta}, q_{R\Delta}$ = matrices of statically determinate element forces and reactions resulting from unit displacement loads on the statically determinate structure.

As a simple example, consider the structure shown in Figure 1. The structure has four element forces numbered as shown. Element force No. 1 is designated as statically determinate. This force is assigned the additional symbol Q_1, while the corresponding element deformation is e_{Q1}. No unassembled deformations or support displacements are introduced; consequently $e_f = \Delta_R = 0$. The unassembled stiffness matrix is

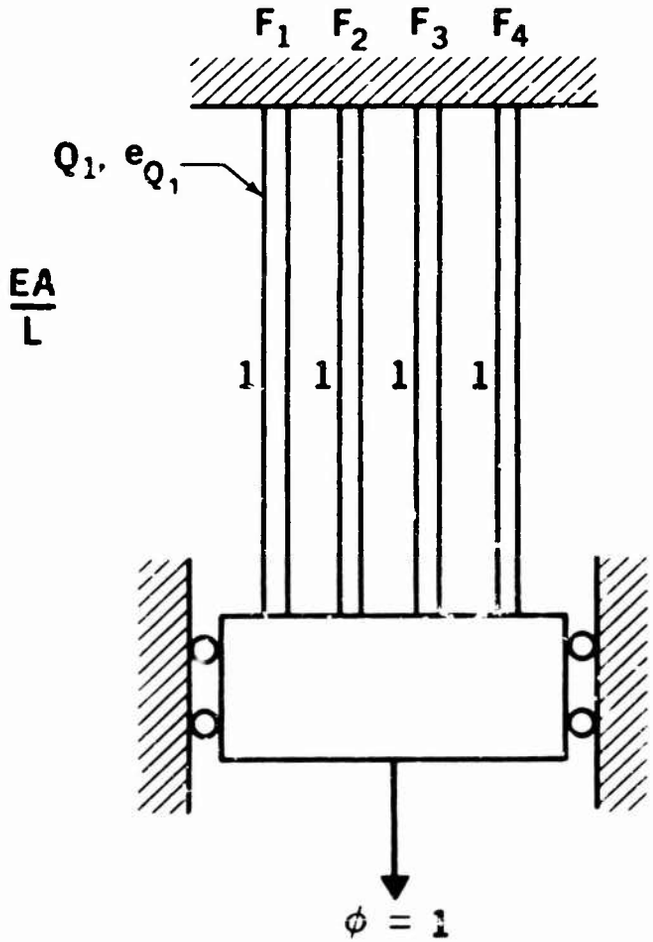
$$k = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \tag{8}$$

The symbol q_{ff} denotes a matrix of statically determinate forces resulting from unit element forces applied to the statically determinate structure as external loads. This matrix is

$$q_{ff} = [\ 1\ |\ -1\ |\ -1\ |\ -1\ ] \tag{9}$$

The matrix is derived as follows: When a column of q_{ff} corresponds to a statically determinate element force, a -1 is inserted in the row corresponding to the force. Thus, the -1 in column 1 can be interpreted as the force that would exist in the first element if unit tensile forces produced by the element on adjacent joints were placed upon the statically determinate structure as external loads. The remaining columns are statically determinate forces resulting from unit forces in the cut members.

FIGURE 1. A PARALLEL STRUCTURE



The symbol $q_{F\phi}$ denotes a matrix of statically determinate element forces resulting from unit loads on the cut structure. This matrix is

$$q_{F\phi} = I = q_{F\Delta} \tag{10}$$

From Equations 1 and 2

$$K_Q = \{-1 \mid -1 \mid -1 \mid -1\}, \; K_{QQ} = 4 \tag{11}$$

where the symbols $\{\ \}$ denote a column matrix.

The matrix of external loads is

$$\phi = 1 \tag{12}$$

From Equation 4

$$e_Q = 1/4(1)(1) = 1/4 \tag{13}$$

From Equation 5

$$F = \{1/4 \mid 1/4 \mid 1/4 \mid 1/4\} \tag{14}$$

From Equation 7

$$\Delta = 1/4 \tag{15}$$

Analysis of this structure by the matrix force method requires the solution of three simultaneous equations. This example demonstrates that the deformation method requires less calculation than the force method when the number of redundants exceeds the number of statically determinate forces.

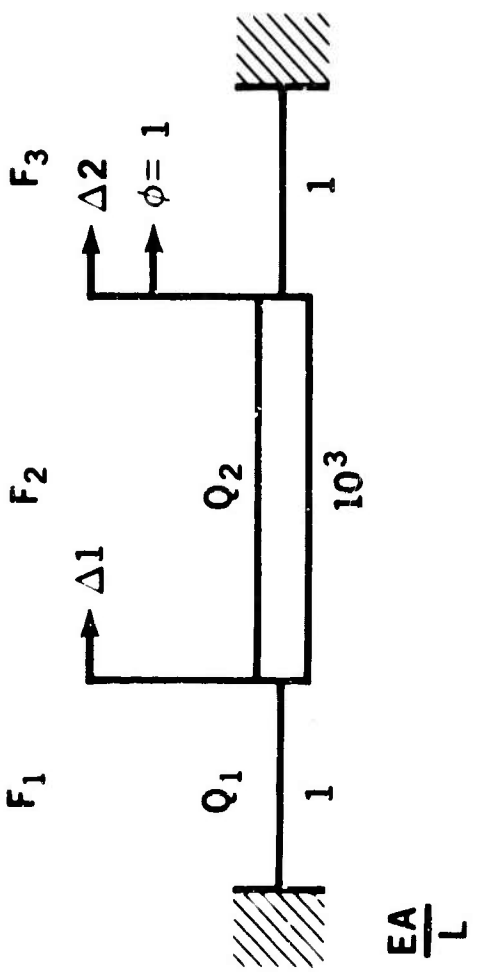
As a second example, consider the structure shown in Figure 2.

Designate the forces in the two left-hand members as statically determinate. Now

$$k = \left[\begin{array}{c|c|c} 1 & & \\ & 10^3 & \\ & & 1 \end{array}\right] \qquad \phi = 1 \tag{16}$$

$$q_{FF} = \left[\begin{array}{c|c|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right] \qquad q_{F\phi} = \left[\begin{array}{c} 1 \\ 1 \end{array}\right] \qquad q_{F\Delta} = \left[\begin{array}{c|c} 1 & 1 \\ 0 & 1 \end{array}\right] \tag{17}$$

FIGURE 2. STRUCTURE WITH A STIFF MEMBER



From Equations 1 to 7

$$K_Q = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & -1000 \\ 1 & 1 \end{array}\right] \qquad K_{QQ} = \left[\begin{array}{c|c} 2 & 1 \\ \hline 1 & 1001 \end{array}\right] \tag{18}$$

$$e_Q = \left[\begin{array}{c|c} 2 & 1 \\ \hline 1 & 1001 \end{array}\right]^{-1} \left[\begin{array}{c} 1 \\ 1 \end{array}\right] \quad (1) = \left\{\begin{array}{c} 0.50 \\ 0.00050 \end{array}\right\} \tag{19}$$

$$F = \left\{\begin{array}{c} 0.50 \\ 0.50 \\ 0.50 \end{array}\right\} \qquad \Delta = \left\{\begin{array}{c} 0.50 \\ 0.50 \end{array}\right\} \tag{20}$$

The results shown for e_Q are obtained by solving the simultaneous equations involved by Gaussian elimination, keeping two significant figures in the calculations. The results are correct to two significant figures.

The displacement method equation for the structure shown in Figure 2 is:

$$K\Delta = P \tag{21}$$

where

$$K = \left[\begin{array}{c|c} 1001 & -1000 \\ \hline 1000 & 1001 \end{array}\right] \qquad P = \left\{\begin{array}{c} 0 \\ 1 \end{array}\right\} \tag{22}$$

and where K is the structural stiffness matrix and P is a matrix of loads in the joint degrees of freedom.

Equation 21 cannot be solved by any procedure keeping only two significant figures because all accuracy is lost. Consequently, the deformation equations are considered better conditioned.

The improved conditioning of the deformation equations results from the proper choice of a statically determinate structure. If the stiff member in the center is selected as a redundant, the deformation equations are also poorly conditioned.

This example illustrates the ability of the deformation method to provide well conditioned equations after proper choice of statically determinate forces.

SELECTION OF STATICALLY DETERMINATE ELEMENT FORCES AND CALCULATION OF q_{FF}, $q_{F\phi}$, q_{RF} AND $q_{R\phi}$

The scope of this paper allows only a brief discussion of these subjects. The statically determinate element forces, and consequently the redundants, can be selected by a pivoting technique applied to the rectangular coefficient matrix of unknown element forces in the equilibrium equations for structural joints. This technique, known as the "Structure Cutter," is described in Reference 3. Member stiffnesses are accounted for by the introduction of weighting factors. Presently, the Structure Cutter outputs are more suitable for the force method, but necessary modifications can be easily and efficiently introduced. As the modified Structure Cutter selects statically determinate forces, it can simultaneously calculate the matrices q_{FF}, $q_{F\phi}$, q_{RF} and $q_{R\phi}$. The matrices $q_{F\Delta}$ and $q_{R\Delta}$ can be extracted from $q_{F\phi}$ and $q_{R\phi}$.

In the generation and solution of the equilibrium equations in the deformation approach, the following advantages of the force method can be preserved: (1) joint reactions can be applied in any direction; (2) the number of calculated displacements can be less than the number of unconstrained degrees of freedom; and (3) joints can have any number of degrees of freedom from 1 to 6, depending on available internal constraints.

PARALLEL ELEMENT FORCES

In the direct stiffness version of the displacement method, considerable simplification results from the use of element forces parallel to the axes of a global coordinate system. Some of this advantage can be secured for the deformation method without sacrificing the option of considering only the minimum number of joint degrees of freedom. Thus, if element forces of finite elements are selected parallel and perpendicular to element edges, then element forces of adjacent elements tend to be parallel. Columns of the matrix of coefficients in the equilibrium equations corresponding to parallel element forces are identical; consequently, duplicate columns can be deleted with significant savings in the equilibrium solution. This consideration can eliminate as much as 60 percent of the effort involved in solving equilibrium equations.

COMPARISON WITH THE DISPLACEMENT EQUATIONS

The equations of the matrix displacement method can be defined:

$$K_U = k\, P_{UF}^{1} \tag{23}$$

$$K = P_{UF}\, K_U \tag{24}$$

$$F_T' = k\left(P_{CF}^{1}\, \Delta_R + e_T\right) \tag{25}$$

$$K\Delta = P_{U\phi}\, \phi + P_{UF}\, F_T' \tag{26}$$

$$F = K_U \Delta + F_T' \tag{27}$$

$$Q_R = P_{CF}\, F - P_{C\phi}\, \phi \tag{28}$$

where

P_{UF} and $P_{U\phi}$ = matrices of components in the unconstrained joint degrees of freedom of unit values of element forces and external loads

P_{CF} and $P_{C\phi}$ = matrices of components in the constrained joint degrees of freedom of unit values of the element forces and external loads.

Equations 23 through 28 are the displacement equations as they are used essentially in practice although practices may vary. In the direct stiffness method, the calculations represented by Equations 23 and 24 are performed in scalar rather than matric form, but the same work must be accomplished. Reactions (Equation 28) can be computed from joint displacements and unassembled deformations, but the calculation is no more concise.

A comparison of Equations 1 through 7 with Equations 23 through 28 immediately shows a strong similarity. In fact, a one-to-one correspondence exists except for Equation 7. There are two major differences between the methods: (1) the deformation method requires calculation of q_{FF}, $q_{F\phi}$, q_{RF}, and $q_{R\phi}$ from P_{UF}, $P_{U\phi}$, P_{CF}, and $P_{C\phi}$, and (2) no calculation in the displacement method, analogous to Equation 7, is required since displacements are already given by Equation 26. The extra calculation required by the deformation approach can be minimized by taking advantage of the extreme sparsity of P_{UF}, $P_{U\phi}$, P_{CF}, and $P_{C\phi}$, the existence of parallel element forces, and the fact that the complete set of joint displacements is seldom required and, therefore, need not be computed.

The remaining extra calculations in the deformation method represent the minimum additional effort required to develop a well-conditioned set of matrix equations. Thus, a trade-off can exist between computing time expended to develop a well-conditioned set of equations, and computing time and reliability saved by not having to exercise extra precision in subsequent calculations.

In Appendix I, the static deformation stiffness matrix is shown to be related to the structural stiffness matrix by a simple transformation. As an example of this transformation, again consider Figure 2. If external loads are assumed to act in both joint degrees of freedom, then

$$q_{F\phi} = \left[\begin{array}{c|c} 1 & 1 \\ \hline 0 & 1 \end{array}\right] \tag{29}$$

Now, from Equations 22, 29, and A-57

$$K_{QQ} = \left[\begin{array}{c|c} 1 & 1 \\ \hline 0 & 1 \end{array}\right] \left[\begin{array}{c|c} 1001 & 1000 \\ \hline 1000 & 1001 \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 1 & 1 \end{array}\right] = \left[\begin{array}{c|c} 2 & 1 \\ \hline 1 & 1001 \end{array}\right] \tag{30}$$

Thus, the static deformation stiffness matrix is correctly computed from the structural stiffness matrix, if four significant figures are saved during the calculation.

ACCURACY OF THE DEFORMATION METHOD

The preceding section reflects a belief that the deformation method is able to provide a better conditioned set of equations in a significant number of practical cases. This belief is based on three considerations: (1) the deformation equations can be shown to be better conditioned for simple structures containing stiff members, (2) the deformation method offers a means of improving conditioning through selection of statically determinate forces (the displacement method offers no corresponding means), and (3) the example of a fuselage shell analysis (discussed in subsequent paragraphs) shows that the deformation method provides a well-conditioned set of equations with accuracy comparable to the force method.

COMPARISON OF THE FORCE AND DEFORMATION METHODS

Appendix II compares the amount of computation required by the force and deformation methods for each of two procedures: (1) Force and Deflection Analysis, in which element forces and deflections are calculated for a limited number of loading conditions, and (2) Deflection Influence Coefficients, in which only the deflection influence matrix is computed. The basis of comparison is the number of multiplications involved in the computations.

Figure 3 summarizes results of the comparison for the force and deflection procedure. The curves are based on Equations B-6 and B-13. The figure shows that the force method is most efficient when the structure has fewer redundants than statically determinate forces. Otherwise, the deformation method is most efficient.

The dotted vertical lines show the value of the ratio r for two types of idealizations. Type 1 is the lumped parameter idealization commonly used in conjunction with the force method. Type 2 represents an idealization composed of rectangular plate elements subjected to in-plane forces of the kind often used with the displacement method. These values of r were established by counting the numbers of statically determinate and redundant forces in a typical wing-box structure for the two idealizations. Figure 3 shows that the deformation method is more efficient than the force method for the type 2 idealization.

Figure 4 shows a similar comparison for the deflection influence computation. The conclusions are similar, but the advantage of the deformation method is more pronounced.

DEMONSTRATION PROBLEM

A section of stiffened cylindrical shell (Figure 5) was analyzed to demonstrate the capability of the deformation method. Both structure and applied loads were symmetric about the Y – Z plane; therefore, the idealized model was reduced to a half shell with symmetric boundary constraints on the plane of symmetry. The shell was rigidly supported at one end and was loaded by internal pressure and shear and moment applied at the free end.

The structure was modeled in terms of lumped parameter elements in a manner that produced a large ratio of redundants to statically determinate forces, simulating a model composed of finite elements. The problem size is summarized in Table 1. Element properties were based on materials and sizing of a typical airframe structure.

The analysis was performed on the IBM 360-85 computer in double precision using up to 300K of core and the FORMAT Structural Analysis System (Reference 3). FORMAT Phase 1 was used to generate the required matrices. The Phase 2 matrix abstraction capability was employed to

FIGURE 3. FORCE AND DEFLECTION ANALYSIS

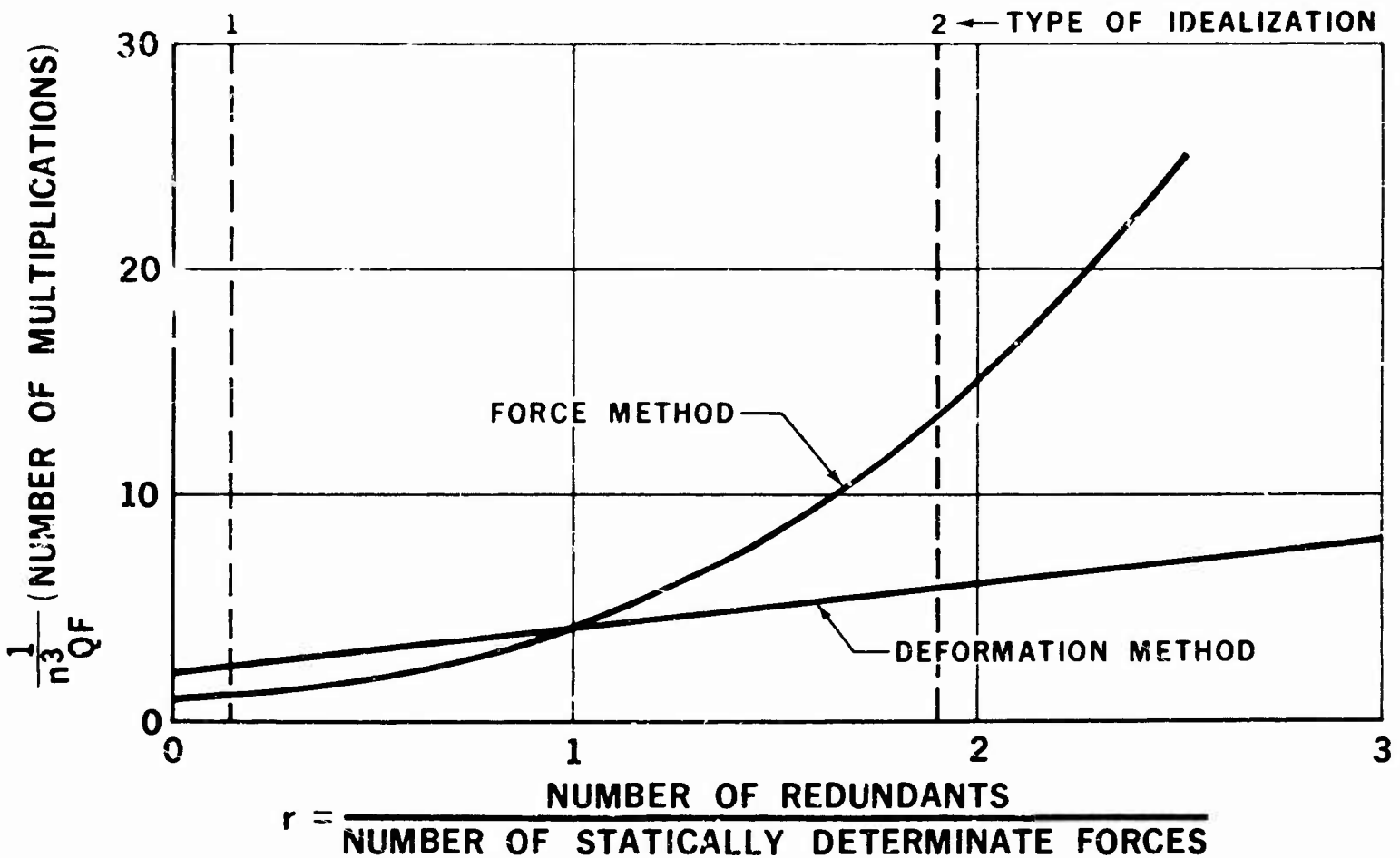


FIGURE 4. DEFLECTION INFLUENCE ANALYSIS

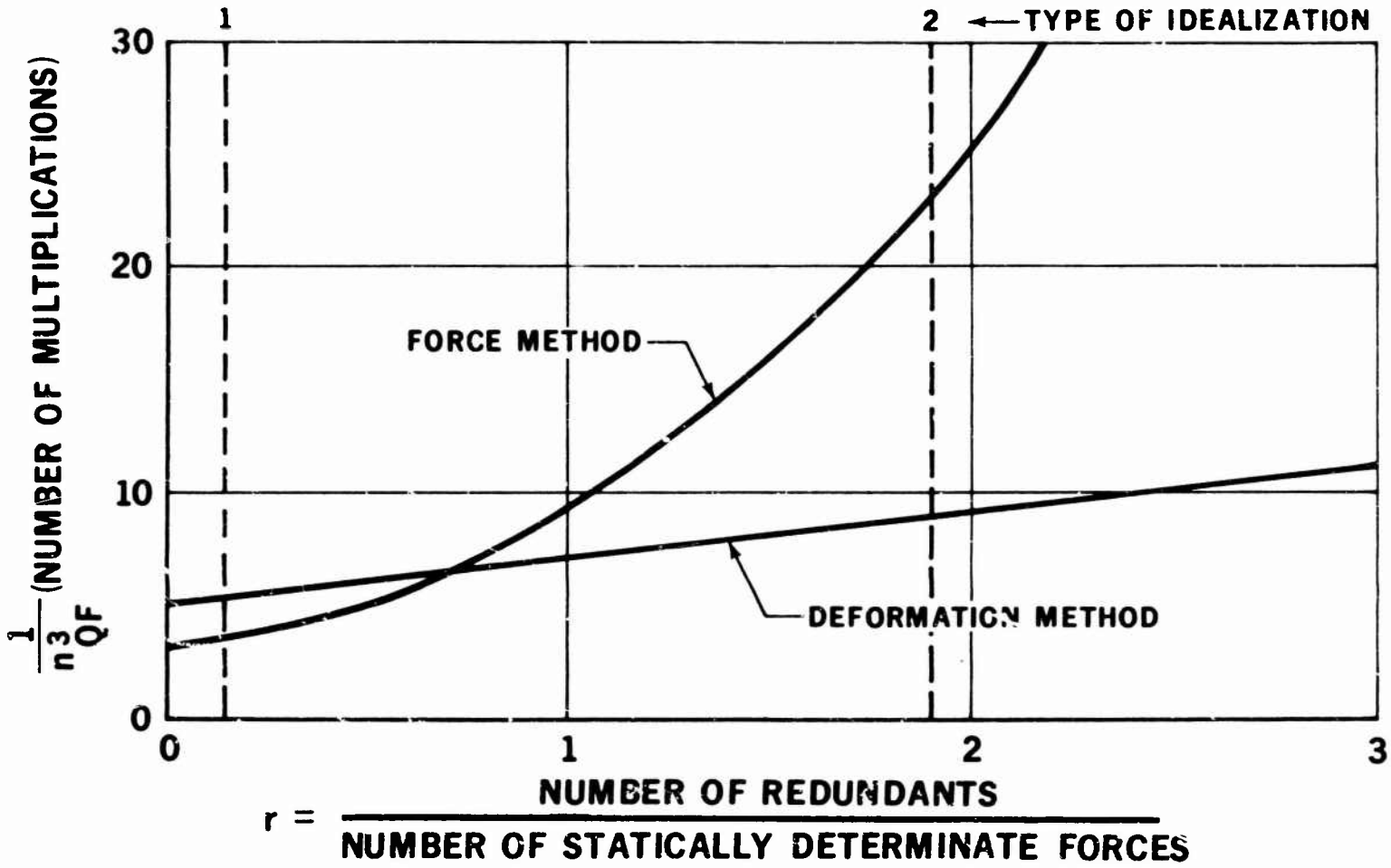


FIGURE 5. DEMONSTRATION STRUCTURE

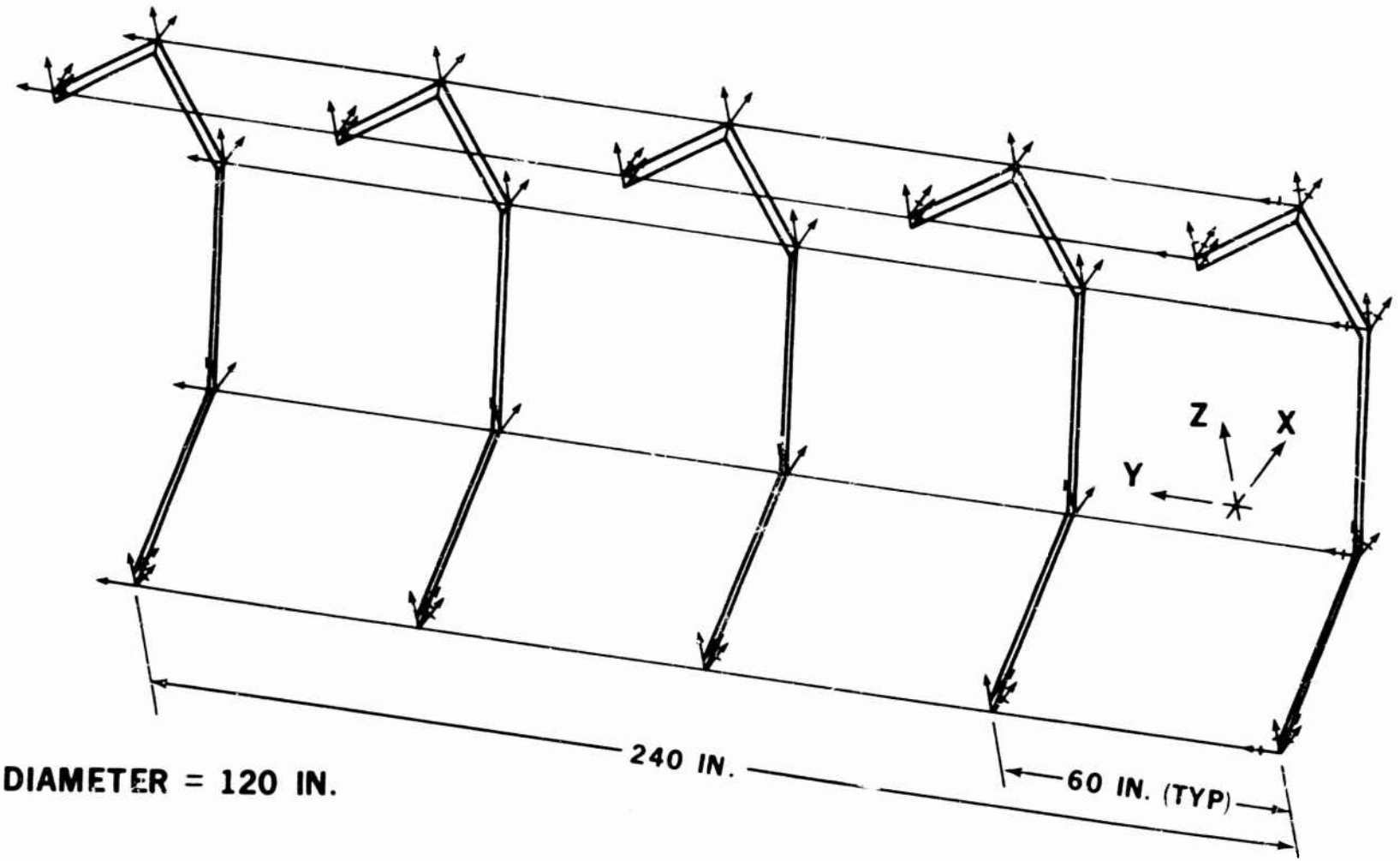


TABLE 1
PROBLEM SIZE

• UNCONSTRAINED DEGREES OF FREEDOM	177
• CONSTRAINED DEGREES OF FREEDOM	33
• REACTION FORCES	33
• ELEMENT FORCES	534
• STATICALLY DETERMINATE FORCES	177
• REDUNDANT FORCES	357

TABLE 2
MACHINE TIME COMPARISON (IBM 360 - 65)

NO.	OPERATION		STATIC DEFORMATION METHOD T_{SDM} (MIN)	FORCE METHOD T_{FM} (MIN)	RATIO $\frac{T_{SDM}}{T_{FM}}$
①	MATRIX GENERATION		1.18	1.18	1.0
②	EQUILIBRIUM SCLUTION		1.427	1.427	1.0
③	CONTINUITY SOLUTION FOR DEFLECTION INFLUENCE COEFFICIENTS	DM	2.647		0.212
		FM		12.472	
④	CONTINUITY SOLUTION FOR FOUR APPLIED LOAD CONDITIONS	DM	2.685		0.257
		FM		10.436	
⑤	COMPLETE ANALYSIS	DEFL INFL COEF'S ① + ② + ③	5.254	15.079	0.348
		APPLIED LOADS ① + ② + ③	5.292	13.043	0.406

solve the matrix equations of the Force and Deformation Methods. A final operation in the Phase 2 step subtracted one solution from the other. The result showed that the solutions were identical to 10 or 12 significant figures. Computer time comparisons are shown in Table 2. The time ratios of 0.406 and 0.348 shown in Table 2 compare favorably with values of 0.393 and 0.354 computed from Equations B-6 and B-13 (see Figures 3 and 4).

UNIFIED FORCE DEFORMATION METHOD

A derivation of this method is presented in Appendix I. The approach is similar to the unified method of structural analysis, except that the unknowns in the force deformation method are redundants and static deformations, whereas the unknowns in the unified method are element forces and displacements. The force deformation method does not lend itself to practical calculation. It is included as a convenient basis for deriving the deformation method.

STATIC FORCE METHOD

Appendix I also contains a derivation of the static force method. The distinguishing feature of this method is that statically determinate element forces are computed before other unknowns. A brief examination of the equations shows that the method is not feasible because the calculations required are excessive.

SECTION III

CONCLUSIONS

The matrix deformation method is a practical approach to the analysis of complex statically indeterminate structures. This method requires less calculation than the force method when the ratio of redundants to statically determinate forces is on the order of 1 or greater. When deflection influence coefficients only are computed, the "break even" point is less than 1.

The conditioning of the deformation method equations can be improved by proper choice of statically determinate deformations. The displacement method has no corresponding capability; consequently, the deformation method can produce better conditioned equations than the displacement method for certain classes of structures.

SECTION IV

REFERENCES

1. Melosh, R.J., and Palacol, E.L., "Manipulation Errors in Finite Element Analysis of Structures," NASA CR-3875, August 1969.

2. Denke, P.H., "Some New Methods of Structural Analysis," Douglas Aircraft Company Report No. DAC-68425, August 1969.

3. Pickard, J., "FORMAT Fortran Matrix Abstraction Technique," AFFDL-TR-66-207, Volume V, October 1968.

APPENDIX I

DERIVATION OF EQUATIONS

The equations of the unified force method, the static deformation method, and the static force method are derived in the following paragraphs.

GOVERNING EQUATIONS

Classify element forces and reactions as statically determinate or redundant. Statically determinate element forces and reactions can be expressed as follows:

$$Q_F = q_{FXF} X_F + q_{FXR} X_R + q_{F\phi} \phi \qquad \text{(A-1)}$$

$$Q_R = q_{RXF} X_F + q_{RXR} X_R + q_{R\phi} \phi \qquad \text{(A-2)}$$

where

Q_F and Q_R = column matrices of statically determinate element forces and reactions

X_F, X_R and ϕ = column matrices of redundant element forces, redundant reactions, and loads

q_{FXF}, q_{FXR} and $q_{F\phi}$ = statically determinate element forces resulting from unit values of redundant element forces, redundant reactions, and loads

q_{RXF}, q_{RXR} and $q_{R\phi}$ = statically determinate reactions resulting from unit values of redundant element forces, redundant reactions, and loads.

The matrices q_{FXF}, q_{FXR}, $q_{F\phi}$, q_{RXF}, q_{RXR} and $q_{R\phi}$ can be computed from statics.

Consider the joints of the structure as free bodies. Apply a unit value of the ith redundant element force $\bar{F}_{X_i}$ to the joints. Balance with the $q_{FXF_{ji}}$ and the $q_{RXF_{ki}}$ which are statically determinate element forces and reactions resulting from the unit $\bar{F}_{X_i}$. Produce a virtual displacement. The virtual work is zero since the joints are rigid bodies in equilibrium.

$$\therefore -1 \cdot e_{X_i} - \sum_j q_{FXF_{ji}} e_{Q_j} + \sum_k q_{RXF_{ki}} \Delta_{QR_k} = 0 \qquad i = 1, 2 \cdots \qquad \text{(A-3)}$$

where the summations extend over the total numbers of statically determinate element forces and reactions, and

e_{X_i} = the ith redundant element deformation

e_{Q_j} = the jth statically determinate element deformation

Δ_{QR_k} = the displacement of the kth statically determinate reaction

From Equation A-3

$$e_X = q^I_{FXF}\ e_Q + q^T_{RXF}\ \Delta_{QR} \tag{A-4}$$

where e_X, e_Q and Δ_{QR} are column matrices of redundant element deformations, statically determinate element deformations, and displacements of statically determinate reactions.

Similarly

$$\Delta_{XR} = q^T_{FXR}\ e_Q \quad q^T_{RXR}\ \Delta_{QR} \tag{A-5}$$

where Δ_{XR} = a column matrix of displacements of redundant reactions.

Also

$$e_Q = e_{QF} + e_{QT} \tag{A-6}$$

$$e_X = e_{XF} + e_{XT} \tag{A-7}$$

where

e_{QF} and e_{QT} = column matrices of statically determinate elastic and unassembled deformations

e_{XF} and e_{XT} = column matrices of redundant elastic and unassembled deformations

Finally

$$\begin{Bmatrix} e_{QF} \\ e_{XF} \end{Bmatrix} = \left[\begin{array}{c|c} D_{QQ} & D_{QX} \\ \hline D_{XQ} & D_{XX} \end{array}\right] \begin{Bmatrix} Q_F \\ X_F \end{Bmatrix} \tag{A-8}$$

$$\begin{Bmatrix} Q_F \\ X_F \end{Bmatrix} = \left[\begin{array}{c|c} k_{QQ} & k_{QX} \\ \hline k_{XQ} & k_{XX} \end{array}\right] \begin{Bmatrix} e_{QF} \\ e_{XF} \end{Bmatrix} \tag{A-9}$$

where D_{QQ}, D_{QX}, D_{XQ}, and D_{XX} are partitions of the element flexibility matrix corresponding to statically determinate and redundant element forces and deformations, and k_{QQ}, k_{QX}, k_{XQ}, and k_{XX} are similar partitions of the element stiffness matrix.

UNIFIED FORCE DEFORMATION EQUATIONS

Eliminating e_Q from Equations A-5 and A-6 gives

$$q_{FXR}^T\ e_{QE} = -q_{FXR}^T\ e_{QT} + q_{RXR}^T\ \Delta_{QR} + \Delta_{XR} \tag{A-10}$$

Now,

$$Q_F = D_{QQ}^{-1}\left(D_{QQ}\ Q_F + D_{QX}\ X_F - D_{QX}\ X_F\right) \tag{A-11}$$

From Equations A-8 and A-11

$$Q_F = D_{QQ}^{-1}\left(e_{QE} - D_{QX}\ X_F\right) \tag{A-12}$$

From Equations A-1 and A-12

$$\left(q_{FXF} + D_{QQ}^{-1}\ D_{QX}\right) X_F + q_{FXR}\ X_R - D_{QQ}^{-1}\ e_{QE} = -q_{F\phi}\ \phi \tag{A-13}$$

It can be shown from Equations A-8 and A-9 that

$$k_{XX}^{-1}\ k_{XQ} + D_{XQ}\ D_{QQ}^{-1} = 0 \tag{A-14}$$

$$e_{XF} = e_{XF} + \left(k_{XX}^{-1}\ k_{XQ} + D_{XQ}\ D_{QQ}^{-1}\right) e_{QF} \tag{A-15}$$

$$= k_{XX}^{-1}\left(k_{XQ}\ e_{QE} + k_{XX}\ e_{XF}\right) + D_{XQ}\ D_{QQ}^{-1}\ e_{QE} \tag{A-16}$$

From Equations A-9 and A-16

$$e_{XE} = k_{XX}^{-1}\ X_F + D_{XQ}\ D_{QQ}^{-1}\ e_{QE} \tag{A-17}$$

Eliminating e_X, e_Q, and e_{XE} from Equations A-4, A-6, A-7 and A-17 gives

$$k_{XX}^{-1}\ X_F + \left(q_{FXF}^T + D_{XQ}\ D_{QQ}^{-1}\right) e_{QE} = -q_{FXF}^T\ e_{QT} - e_{XT} + q_{RXF}^T\ \Delta_{QR} \tag{A-18}$$

From Equations A-10, A-13, and A-18

$$\begin{bmatrix} k_{XX}^{-1} & 0 & q_{FXF}^T + D_{XQ}\ D_{QQ}^{-1} \\ 0 & 0 & q_{FXR}^T \\ q_{FXF} + D_{QQ}^{-1}\ D_{QX} & q_{FXR} & -D_{QQ}^{-1} \end{bmatrix} \begin{Bmatrix} X_F \\ X_R \\ e_{QE} \end{Bmatrix} = \begin{Bmatrix} -q_{FXF}^T\ e_{QT} - e_{XT} + q_{RXF}^T\ \Delta_{QR} \\ -q_{FXR}^T\ e_{QT} + q_{RXR}^T\ \Delta_{QR} + \Delta_{XR} \\ -q_{F\phi}\ \phi \end{Bmatrix} \tag{A-19}$$

Equation A-19 is the unified force deformation equation. The unknowns are redundant element forces, redundant reactions, and statically determinate element deformations. Note the symmetry of the coefficient matrix of unknowns. The redundant force equations can be derived from Equation A-19 by eliminating e_{QE}. The static deformation equations are derived by eliminating X_F and X_R.

MATRIX DEFORMATION EQUATIONS

The characteristic feature of the matrix deformation method is that e_{QE} is the first matrix of unknowns to be computed. However, a study of Equation A-19 shows that e_{QE} cannot be computed before X_R. The redundant reactions are, therefore, eliminated by establishing the convention that reactions are never selected as redundants. This requirement does not compromise the method since the selection of reactions as redundants is never a good choice. The following simplified notation is introduced:

$$\begin{aligned} X &= X_F & q_{FX} &= q_{FXF} \\ \Delta_R &= \Delta_{QR} & q_{RX} &= q_{RXF} \end{aligned} \qquad \text{(A-20)}$$

From Equation A-19, deleting X_R:

$$k_{XX}^{-1} X + \left(q_{FX}^T + D_{XQ} D_{QQ}^{-1}\right) e_{QE} = -q_{FX}^T e_{QT} - e_{XT} + q_{RX}^T \Delta_R \qquad \text{(A-21)}$$

$$\left(q_{FX} + D_{QQ}^{-1} D_{QX}\right) X - D_{QQ}^{-1} e_{QE} = -q_{FQ} \phi \qquad \text{(A-22)}$$

Eliminate X from Equations A-21 and A-22:

$$\left[\left(q_{FX} + D_{QQ}^{-1} D_{QX}\right) k_{XX} \left(q_{FX}^T + D_{XQ} D_{QQ}^{-1}\right) + D_{QQ}^{-1}\right] e_{QE}$$

$$= q_{FQ} \phi - \left(q_{FX} + D_{QX}^{-1} D_{QX}\right) k_{XX} \left(q_{FX}^T e_{QT} + e_{XT} - q_{RX}^T \Delta_R\right) \qquad \text{(A-23)}$$

From Equations A-8 and A-9 it can be shown that

$$D_{QQ}^{-1} = k_{QQ} - k_{QX} k_{XX}^{-1} k_{XQ} \qquad \text{(A-24)}$$

Combining Equations A-14, A-23, and A-24 gives

$$\left(k_{QQ} - k_{QX} q_{FX}^T - q_{FX} k_{XQ} + q_{FX} k_{XX} q_{FX}^T\right) e_{QE}$$

$$= q_{FQ} \phi + \left(k_{QX} - q_{FX} k_{XX}\right)\left(q_{FX}^T e_{QT} + e_{XT} - q_{RX}^T \Delta_R\right) \qquad \text{(A-25)}$$

Equation A-25 can be written

$$\left[\begin{array}{c|c} I & q_{FX} \end{array}\right] \left[\begin{array}{c|c} k_{QQ} & k_{QX} \\ \hline k_{XQ} & k_{XX} \end{array}\right] \left[\begin{array}{c} -I \\ \hline q_{FX}^T \end{array}\right] e_{QE}$$

$$= q_{F\phi}\,\phi - \left[\begin{array}{c|c} I & q_{FX} \end{array}\right] \left[\begin{array}{c|c} k_{QQ} & k_{QX} \\ \hline k_{XQ} & k_{XX} \end{array}\right] \left(\left[\begin{array}{c} -I \\ \hline q_{FX}^T \end{array}\right] e_{QT} - \left[\begin{array}{c} 0 \\ \hline q_{RX}^T \end{array}\right] \Delta_R + \left\{\begin{array}{c} e_{QT} \\ \hline e_{XT} \end{array}\right\} \right) \qquad \text{(A-26)}$$

$$\therefore\ q_{F\bar{F}}\,\bar{k}\,q_{F\bar{F}}^T\,e_{QE} = q_{F\phi}\,\phi - q_{F\bar{F}}\,\bar{k}\left(q_{F\bar{F}}^T\,e_{QT} - q_{R\bar{F}}^T\,\Delta_R + \bar{e}_T\right) \qquad \text{(A-27)}$$

where

$$q_{F\bar{F}} = \left[\begin{array}{c|c} -I & q_{FX} \end{array}\right] \qquad \text{(A-28)}$$

$$q_{R\bar{F}} = \left[\begin{array}{c|c} 0 & q_{RX} \end{array}\right] \qquad \text{(A-29)}$$

$$\bar{k} = \left[\begin{array}{c|c} k_{QQ} & k_{QX} \\ \hline k_{XQ} & k_{XX} \end{array}\right] \qquad \text{(A-30)}$$

$$\bar{e}_T = \left\{\begin{array}{c} e_{QT} \\ \hline e_{XT} \end{array}\right\} \qquad \text{(A-31)}$$

From Equations A-6 and A-27

$$q_{F\bar{F}}\,\bar{k}\,q_{F\bar{F}}^T\,e_Q = q_{F\phi}\,\phi + q_{F\bar{F}}\,\bar{k}\left(q_{R\bar{F}}^T\,\Delta_R - \bar{e}_T\right) \qquad \text{(A-32)}$$

Equation A-32 is the basic deformation equation. Solving for e_Q gives statically determinate deformations as functions of external loads, support displacements, and unassembled deformations.

The matrix $q_{F\bar{F}}$ can be given a physical interpretation. The rows of $q_{F\bar{F}}$ and the columns of the first partition (Equation A-28) correspond to statically determinate element forces. The columns of the second partition correspond to redundants, thus, the columns of $q_{F\bar{F}}$ correspond to element forces. The matrix can be interpreted as a matrix of statically determinate element forces resulting from unit element forces applied as external loads to the statically determinate structure. Thus, a statically determinate element carrying a unit tensile force will carry a unit compressive force if the tension is applied to the appropriate joints as an external load. In this manner the partition $-I$ is interpreted. The partition q_{FX} is simply a matrix of statically determinate forces resulting from unit redundants.

Similarly, $q_{R\bar{F}}$ is interpreted as a matrix of reactions resulting from unit element forces applied as external loads to the statically determinate structure.

From Equation A-9

$$\bar{F} = \bar{k}\,\bar{e}_F \tag{A-33}$$

where

$$\bar{F} = \begin{Bmatrix} Q_F \\ X \end{Bmatrix} \qquad \bar{e}_F = \begin{Bmatrix} e_{QF} \\ e_{XF} \end{Bmatrix} \tag{A-34}$$

From Equations A-6, A-7, A-31, and A-34

$$\bar{e} = \bar{e}_F + \bar{e}_T \tag{A-35}$$

where

$$\bar{e} = \begin{Bmatrix} e_Q \\ e_X \end{Bmatrix} \tag{A-36}$$

From Equations A-4, A-35, and A-36

$$\bar{e}_F = \begin{Bmatrix} e_Q \\ q_{FX}^T e_Q + q_{RX}^T \Delta_R \end{Bmatrix} - \bar{e}_T \tag{A-37}$$

$$\therefore \bar{e}_F = \begin{bmatrix} I \\ q_{FX}^T \end{bmatrix} e_Q - \bar{e}_T + \begin{bmatrix} 0 \\ q_{RX}^T \end{bmatrix} \Delta_R \tag{A-38}$$

From Equations A-28, A-29, and A-38

$$\bar{e}_F = q_{FF}^T e_Q - \bar{e}_T + q_{RF}^T \Delta_R \tag{A-39}$$

From Equations A-33 and A-39

$$\bar{F} = \bar{k}\,q_{FF}^T e_Q + \bar{k}\left(q_{RF}^T \Delta_R - \bar{e}_T\right) \tag{A-40}$$

From Equation A-2, deleting X_R and introducing the simplified notation

$$Q_R = \left[0 \mid q_{RX}\right] \begin{Bmatrix} Q_F \\ X \end{Bmatrix} + q_{R\phi}\,\phi \tag{A-41}$$

From Equations A-29, A-34, and A-41

$$Q_R = q_{RF}\,\bar{F} + q_{R\phi}\,\phi \tag{A-42}$$

The following equation can be derived from virtual work:

$$\Delta = q_{F\Delta}^{T} e_Q \quad q_{R\Delta}^{T} \Delta_R \tag{A-43}$$

where

Δ = a column matrix of joint displacements

$q_{F\Delta}, q_{R\Delta}$ = matrices of statically determinate element forces resulting from unit loads coinciding in position and direction with the desired joint displacements

Equations A-32, A-40, A-42, and A-43 give statically determinate deformations, element forces, reactions, and displacements

Transformation of the Element Stiffness Matrix

The matrix k (Equation A-30) is divided into partitions corresponding to statically determinate and redundant element forces. This partitioning can lead to inefficiency in subsequent operations performed on the stiffness matrix. A transformation that produces the usual and more efficient banded form of the unassembled stiffness matrix is presented.

Let F = a column matrix of element forces in which the forces corresponding to any particular element appear consecutively. Then

$$F = T \bar{F} \tag{A-44}$$

where T is a Boolean transformation that reorders the rows of $\bar{F}$. This transformation has the property

$$T T^{T} = T^{T} T = I \tag{A-45}$$

From Equations A-32 and A-45

$$q_{F\bar{F}} T^{T} T \bar{k} T^{T} T q_{F\bar{F}}^{T} e_Q = q_{F\phi} \phi + q_{F\bar{F}} T^{T} T \bar{k} T^{T} T \left(q_{R\bar{F}}^{T} \Delta_R - \bar{e}_T \right) \tag{A-46}$$

$$q_{FF} k q_{FF}^{T} e_Q = q_{F\phi} \phi + q_{FF} k \left(q_{RF}^{T} \Delta_R - e_T \right) \tag{A-47}$$

where

$$\begin{aligned} q_{FF} &= q_{F\bar{F}} T^{T} & q_{RF} &= q_{R\bar{F}} T^{T} \\ k &= T \bar{k} T^{T} & e_T &= T \bar{e}_T \end{aligned} \tag{A-48}$$

The transformed element stiffness matrix k has the desired banded form.

Let

$$K_Q = k\, q_{FF}^T$$

$$K_{QQ} = q_{FF}\, K_Q \qquad \text{(A-49)}$$

$$F_T = k\left(q_{RF}^T\, \Delta_R - e_T\right)$$

From Equations A-47 and A-49

$$K_{QQ}\, e_Q = q_{F\phi}\, \phi + q_{FF}\, F_T \qquad \text{(A-50)}$$

Let

$$e = T\,\bar{e} \qquad e_E = T\,\bar{e}_E \qquad \text{(A-51)}$$

Note that e, e_E, and e_T are column matrices of element deformations, elastic element deformations, and unassembled element deformations arranged in the same order as the element forces in F. From Equations A-40 and A-45

$$T\,\bar{F} = -T\,\bar{k}\,T^T\,T\,q_{F\bar{F}}^T\,e_Q + T\,\bar{k}\,T^T\,T\left(q_{R\bar{F}}^T\,\Delta_R - \bar{e}_T\right) \qquad \text{(A-52)}$$

From Equations A-44, A-48, A-49 and A-52

$$F = -K_Q\, e_Q + F_T \qquad \text{(A-53)}$$

From Equations A-42, A-44, A-45, and A-48

$$Q_R = q_{RF}\, F + q_{R\phi}\, \phi \qquad \text{(A-54)}$$

Statically determinate deformations, element forces, reactions, and displacements are computed from Equations A-50, A-53, A-54, and A-43.

An equation for deflection influence coefficients can be obtained by substituting e_Q from Equation A-50 into Equation A-43, and setting $\phi = I$, $F_T = 0$, $\Delta_R = 0$, and $q_{F\Delta} = q_{F\phi}$.

$$\therefore \delta = q_{F\phi}^T\, K_{QQ}^{-1}\, q_{F\phi} \qquad \text{(A-55)}$$

where δ is a deflection influence matrix. If $q_{F\phi}$ is a matrix of statically determinate element forces resulting from unit loads in all unconstrained joint degrees of freedom, then $q_{F\phi}$ is square and nonsingular and δ is a nonsingular matrix of deflection influence coefficients for all unconstrained degrees of freedom. In this case

$$\delta = K^{-1} \qquad \text{(A-56)}$$

where K is the structural stiffness matrix. From Equations A-55 and A-56

$$K_{QQ} = q_{F\phi}\, K\, q_{F\phi}^T \qquad \text{(A-57)}$$

The static stiffness matrix is related to the structural stiffness matrix by a simple transformation.

STATIC FORCE METHOD

The possibility of computing statically determinate element forces before other unknowns is explored.

Reactions are not selected as redundants. From Equations A-6 and A-8, introducing the simplified notation of Equation A-20

$$e_Q = D_{QQ}\, Q_F + D_{QX}\, X + e_{QT} \tag{A-58}$$

From Equations A-7 and A-8

$$e_X = D_{XQ}\, Q_F + D_{XX}\, X + e_{XT} \tag{A-59}$$

Eliminating e_Q and e_X from Equations A-4, A-58, and A-59 gives

$$\left(q_{FX}^T\, D_{QQ} + D_{XQ}\right) Q_F + \left(q_{FX}^T\, D_{QX} + D_{XX}\right) X = -q_{FX}^T\, e_{QT} - e_{XT} + q_{RX}^T\, \Delta_R \tag{A-60}$$

$$\therefore\ X = -AQ_F + B \tag{A-61}$$

where

$$\begin{aligned} A &= \left(q_{FX}^T\, D_{QX} + D_{XX}\right)^{-1} \left(q_{FX}^T\, D_{QQ} + D_{XQ}\right) \\ B &= \left(q_{FX}^T\, D_{QX} + D_{XX}\right)^{-1} \left(-q_{FX}^T\, e_{QT} - e_{XT} + q_{RX}^T\, \Delta_R\right) \end{aligned} \tag{A-62}$$

Eliminating X from Equations A-1 and A-61 and solving for Q_F gives

$$Q_F = \left(I + q_{FX}\, A\right)^{-1} \left(q_{FX}\, B + q_\phi\, \phi\right) \tag{A-63}$$

A study of Equations A-62 and A-63 shows that the computations involved in this approach are excessive.

APPENDIX II

COMPARISON OF COMPUTATIONAL EFFORT FOR FORCE AND DEFORMATION METHODS

The comparison is accomplished by counting the approximate number of calculations required by each method to solve a typical aircraft structural analysis problem on the following basis:

- Only multiplications are counted.
- Element stiffness and flexibility matrices are assumed to be banded and very sparse.
- The number of loading conditions is negligible compared to the number of element forces.

This basis is considered adequate to establish the "break even" point between the force and deformation methods. The following computational procedures are considered separately:

- Force and Deflection Analysis — element forces and deflections are calculated for a limited number of loading conditions.
- Deflection Influence Coefficients — only the deflections influence matrix is computed.

Force and Deflection Analysis

The required force method equations are

$$\delta_{XX} = f_X^T \, Df_X \tag{B-1}$$

$$X = -\delta_{XX}^{-1} f_X^T \, Df_\phi \, \phi \tag{B-2}$$

$$F = f_X X + f_\phi \, \phi \tag{B-3}$$

$$\Delta = f_\Delta^T \, DF \tag{B-4}$$

where f_X, f_ϕ and f_Δ are matrices of element forces in the statically determinate structure resulting from unit redundants, external loads, and displacement loads.

The number of multiplications involved in Equation B-1 is approximately $n_X^2 \, n_{QF}$, where n_X = the number of redundants and n_{QF} = the number of statically determinate forces. This estimate accounts for the sparse banded nature of D, and the existence in f_X of a partition comprising a unit matrix. The number of multiplications in Equation B-2 is approximately n_X^3. The solution of the equilibrium equations for f_X and f_ϕ is an additional source of effort. The number of multiplications involved in the solution of these equations is estimated to be less than or equal to $(n_{QF} + n_X) \, n_{QF}^2$. The remaining equations involve negligible calculation.

$$\therefore n_{REM} = n_{QF}^3 + n_{QF}^2 \, n_X + n_{QF} \, n_X^2 + n_X^3 \tag{B-5}$$

where n_{REM} = the number of multiplications involved in the force method.

$$\therefore \frac{n_{REM}}{n_{QF}^3} = 1 + r + r^2 + r^3 \tag{B-6}$$

where

$$r = \frac{n_X}{n_{QF}} \tag{B-7}$$

The required deformation method equations are

$$K_{QQ} = q_{FF} \; k \; q_{FF}^T \tag{B-8}$$

$$e_Q = K_{QQ}^{-1} \; q_{F\phi} \; \phi \tag{B-9}$$

$$F = -K_Q \; e_Q \tag{B-10}$$

$$\Delta = q_{F\Delta}^T \; e_Q \tag{B-11}$$

The number of multiplications involved in Equation B-8 is approximately $n_{QF}^2 \; n_X$. This estimate accounts for the sparse banded nature of k and the existence in q_{FF} of a partition comprising a unit matrix. The number of multiplications in Equation B-9 is approximately n_{QF}^3. The number of multiplications involved in solving the equilibrium equations for q_{FF} and $q_{F\phi}$ is estimated to be less than or equal to $(n_{QF} + n_X) \; n_{QF}^2$. The remaining equations involve negligible calculation

$$\therefore n_{SDF} = 2n_{QF}^3 + 2n_{QF}^2 \; n_X \tag{B-12}$$

where N_{SDF} = the number of multiplications involved in the static deformation method.

$$\therefore \frac{n_{SDF}}{n_{QF}^3} = 2 + 2r \tag{B-13}$$

Deflection Influence Coefficients

The required force method equations are again Equations B-1 to B-5 inclusive, and the equilibrium equations. Take $\phi = I$, a unit matrix, and $q_{F\Delta} = q_{F\phi}$. In this case, the effort involved in Equations B-3 and B-4 is not negligible. The estimated number of multiplications is given by

$$\frac{n_{REM}}{n_{QF}^3} = 3 + 3r + 2r^2 + r^3 \tag{B-14}$$

The deformation equations are

$$K_{QQ} = q_{FF}\ k\ q_{FF}^{T} \tag{B-15}$$

$$\Delta = q_{F\phi}^{T}\ K_{QQ}^{-1}\ q_{F\phi} \tag{B-16}$$

plus the equilibrium equations. The estimated number of multiplications is

$$\frac{n_{SDM}}{n_{QF}^{3}} = 5 + 2r \tag{B-17}$$

FINITE ELEMENT ANALYSIS OF THIN-WALLED STRUCTURES BASED ON THE MODERN ENGINEERING THEORY OF BEAMS

By Tadahiko Kawai* and Toyohiko Muraki**
Institute of Industrial Science,
University of Tokyo, Tokyo, Japan

Noboru Tanaka*** and Takashi Iwaki****
Mitsui Shipbuilding and Engineering
Company, Ltd. Tamano, Okayama, Japan

A practical method of overall finite element analysis of thin-walled beam-type structures such as aircrafts, ships, bridges etc. is proposed in this paper.
The analysis consists of two steps, the first of which is evaluation of various sectional properties of a beam with arbitrary cross section, especially St. Venant's torsional stiffness and warping torsional stiffness. The second step is solution as well as formulation of overall stiffness equation of a given structure by using general stiffness matrix of a beam element which is defined with respect to an arbitrary point on the cross section. For justification of the proposed method, results of shear flow analysis of some beam cross sections and overall torsional analysis of a container model ship as well as an actual container ship with their experimental data are shown.

* Professor
** Graduate Student
*** Chief Research Engineer
**** Senior Research Engineer

SECTION I

INTRODUCTION

In the field of structural engineering, it is a common practice to idealize slender plate structures as thin-walled beams with rigid cross sections and make stress analysis by using the modern engineering theory of beams. Ships, bridges, aircrafts, railroad trucks, large machine bases etc. are typical examples of such structures.
Design philosophy of these structures has been based on the bending-torsion theory of thin-walled beams proposed by many outstanding scholars such as S. P. Timoshenko, F. Bleich, V. Z. Vlasov and so on before advent of electronic digital computers. With the progress of high speed digital computers of large capacity, however, finite element method has been regarded as the most reliable tool on the stress analysis of thin-walled structures since the first celebrated paper on the displacement method of analysis proposed by Mr. M. J. Turner, Professor R. W. Clough, late Professor H. C. Martin, and Mr. L. J. Topp. In the conventional finite element method, such structures are divided into large number of beam or plate elements and overall stiffness equation is formulated by calculating a stiffness matrix of each element and assembling them after their transformation from the local to global coordinates. And finally unusually large number of linear simultaneous equations should be handled.
Today, however, no body doubts that elaborate analysis of such structures can be made only by solving these king size overall stiffness equations. In the initial design stage, use of such a general purpose computer program of the finite element structural analysis may not be economically feasible, instead it would be the most powerful and effective if they are used in the final stage of structural design. It is the purpose of the present paper to propose a new finite element method of analysis of slender plate structures based on the modern engineering theory of beams, by which overall structural analysis of thin-walled structures can be made easily and quickly with minimum labour and cost for the required accuracy.

SECTION II

THEORETICAL BASIS FOR FINITE ELEMENT ANALYSIS OF THIN-WALLED STRUCTURES WITH A VARIABLE CROSS SECTION

In order to formulate the theory of finite element analysis of thin-walled structures with a variable cross section, it is essential to derive a general stiffness matrix of a beam element which is defined with respect to an arbitrary nodal point on a given cross section. In this section, therefore, theoretical basis for finite element analysis of a thin-walled structure will be briefly described.

2.1 Recapitulation of Modern Engineering Theory of a Beam

Modern engineering theory of a beam is based on the following assumed displacement functions: (See Fig. 1)

$$\left.\begin{aligned} U(x,y,z) &= u(z) - y\theta(z) \\ V(x,y,z) &= v(z) + x\theta(z) \\ W(x,y,z) &= w(z) - x u'(z) - y v'(z) + \theta'(z)\,\omega_n(x,y) \end{aligned}\right\} \quad (1)$$

where (U, V, W) represents the displacement vector of a general nodal point, while $(u(z), v(z))$ is the lateral displacement vector of the centroid, $w(z)$ is the average axial displacement of the given section, and $\theta(z)$ is the angle of rotation as shown in Fig. 1.
$\omega_n(x,y)$ is the so-called normalized warping function of St. Venant's torsion problem and it is defined as follows:

$$\left.\begin{aligned} \omega_n(x,y) &= \omega(x,y) - \frac{S_\omega}{A} \\ S_\omega &= \iint \omega(x,y)\,dxdy \\ A &= \iint dxdy \end{aligned}\right\} \quad (2)$$

According to the well-known St. Venant's torsion theory, $\omega(x,y)$ is defined as a solution of the following boundary value problem:

$$\Delta\omega(x,y) = 0 \quad (3)$$

with the following boundary condition:

$$\frac{\partial\omega}{\partial n} = y\cos(n,x) - x\cos(n,y) \quad (4)$$

on the boundary curve C for a simply connected region. (See Fig. 2)
It is easily shown that the corresponding variational problem of St. Venant's torsion problem is given as follows:

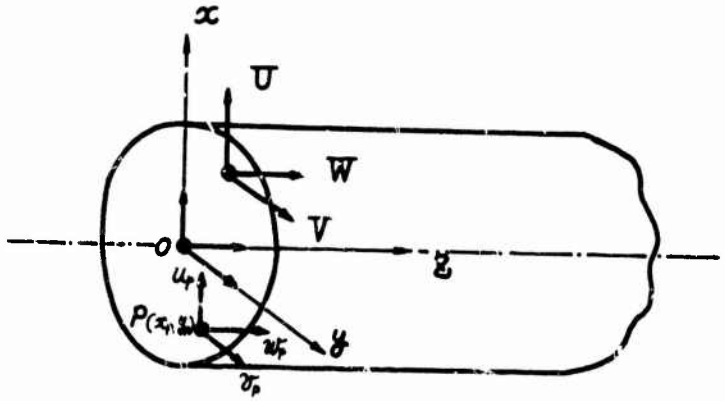


Fig. 1 COORDINATES AND NOTATIONS EMPLOYED

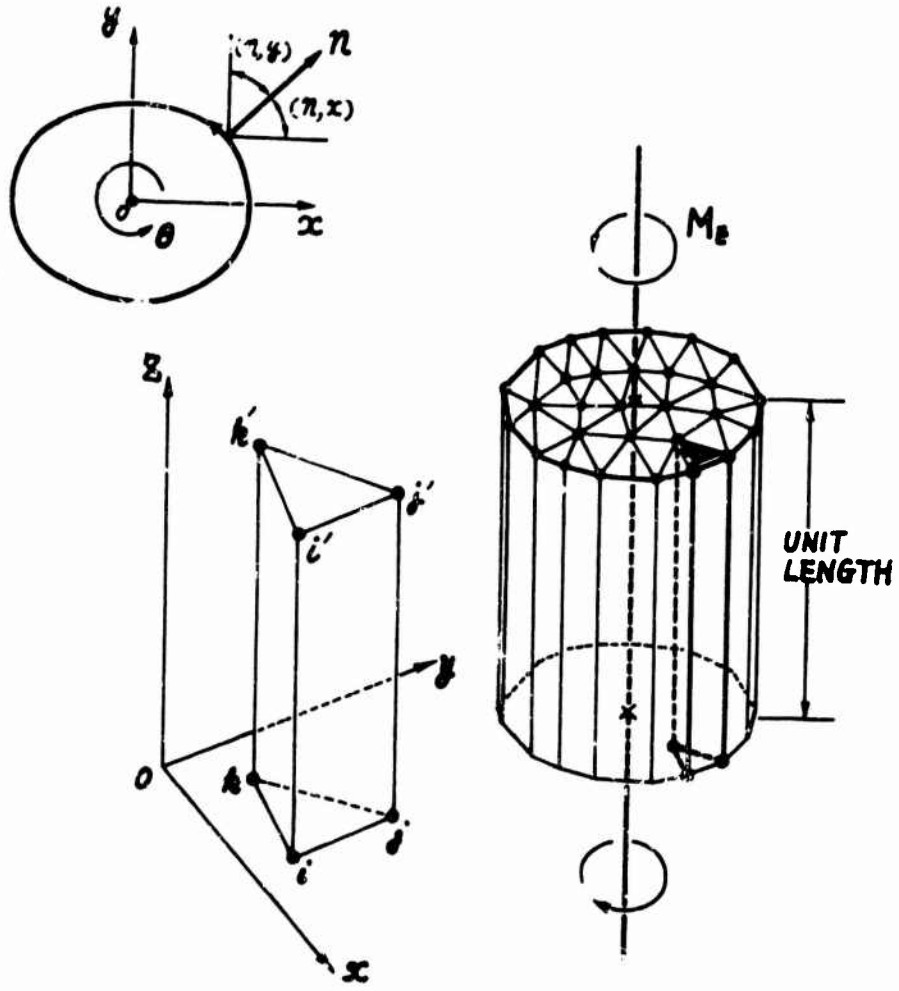


Fig. 2 FEM ANALYSIS OF ST. VENANT'S TORSION PROBLEM BY A TRIANGULAR PRISM ELEMENT

$$\delta \iint_{\omega} \frac{1}{2} G\alpha^2 \left\{ \left(\frac{\partial \omega}{\partial x} - y\right)^2 + \left(\frac{\partial \omega}{\partial y} + x\right)^2 \right\} dx\,dy$$

$$- M_z \delta\alpha = 0 \qquad (5)$$

in which variation should be taken with respect to the warping function $\omega(x, y)$ as well as the rate of twist α.
The following two variational equations can be derived from Eq.(5):

$$(M_z - GK\alpha)\delta\alpha = 0 \qquad (6)$$

and

$$\delta \iint \frac{1}{2} \left\{ \left(\frac{\partial \omega}{\partial x} - y\right)^2 + \left(\frac{\partial \omega}{\partial y} + x\right)^2 \right\} dx\,dy = 0 \qquad (7)$$

where K is the St. Venant's torsional rigidity and it is defined by the following equation:

$$K = \iint \left\{ \left(\frac{\partial \omega}{\partial x} - y\right)^2 + \left(\frac{\partial \omega}{\partial y} + x\right)^2 \right\} dx\,dy \qquad (8)$$

It can be easily seen that Eq.(6) gives the well-known St. Venant's torsion formula in strength of materials, while Eq. (7) is mathematically equivalent to Eqs.(3) and (4).
The first displacement analysis of St. Venant's torsion problem should be credit to Professor L. R. Herrmann.[1] Independently T. Kawai proposed the method which is exactly the same as Prof. L. Herrmann and applied it to the torsional analysis of a shaft with a key hole in May 1968.[2]
Once the warping function $\omega(x, y)$ is determined, various sectional properties of warping-torsion can be calculated by the following formulae.

(i) St. Venant's torsional constant K

$$K = \iint \left\{ \left(\frac{\partial \omega}{\partial x} - y\right)^2 + \left(\frac{\partial \omega}{\partial y} + x\right)^2 \right\} dx\,dy \qquad (9)$$

(ii) Warping torsional rigidity EI_ω with respect to the centroid

$$I_\omega = \iint \left\{ \omega_n(x, y) \right\}^2 dx\,dy \qquad (10)$$

(iii) Warping cross products of inertia $I_{\omega x}$, $I_{\omega y}$ with respect to the centroid

$$I_{\omega x} = \iint x\, \omega_n(x, y)\, dx\,dy$$

$$I_{\omega y} = \iint y\, \omega_n(x, y)\, dx\,dy \qquad (11)$$

(iv) Location of shear center (x_s, y_s)

$$x_s = -\frac{I_{\omega y}}{I_{yy}}, \qquad y_s = \frac{I_{\omega x}}{I_{xx}} \tag{12}$$

where

$$I_{xx} = \iint x^2 dx dy, \quad I_{yy} = \iint y^2 dx dy$$

Practical computation of various section properties can be made by using Simpson's Rule and computed results of $\omega_n(x, y)$.
The three different finite elements for St. Venant's torsional analysis are prepared and they are shown in Table 1.
The triangular prism element is the most general element, especially suitable for torsional analysis of irregular shapes, while in case of thin-walled structures, the rectangular prism element is very practical and high accuracy of computation can be maintained with the coarse mesh division. As a matter of fact, it is interesting to note that the finite element analysis with the plate strip element is equivalent in accuracy to the so-called shear flow analysis of the given thin-walled structures.
Some numerical examples of St. Venant's torsion problems are shown in Fig.3.

2.2 Derivation of a New Stiffness Matrix of a Beam Element for Warping Torsion

According to the modern warping-torsion theory of a beam, the governing differential equations can be given as follows:

$$E I_\omega^{(s)} \theta'''(z) - GK\theta''(z) = m_z(z) \tag{13-a}$$

with the following boundary conditions at $z = 0$ or ℓ

$$(GK\theta'(z) - E I_\omega^{(s)} \theta'''(z))\delta\theta = 0 \tag{13-b}$$

$$E I_\omega^{(s)} \theta''(z)\delta\theta' = 0 \tag{13-c}$$

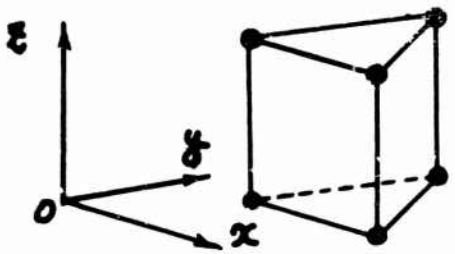
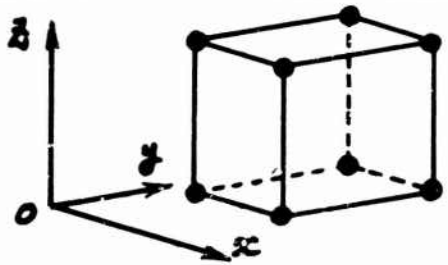
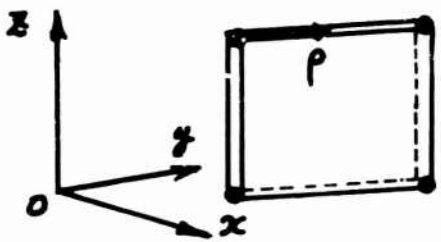
where $E I_\omega^{(s)}$ is the warping torsional rigidity of a given beam with respect to the shear center.
It is easy to show that solution of Eq. (13-a) together with the boundary conditions Eqs. (13-b) and (13-c) is equivalent to obtain the approximate solution to satisfy the following variational equation:

$$\delta \int_0^\ell \left(\frac{1}{2} E I_\omega^{(s)} \theta''^2 + \frac{1}{2} GK\theta'^2\right) dz - \int_0^\ell m_z \delta\theta \, dz = 0 \tag{14}$$

To derive the stiffness equation for warping-torsion of a beam, the most of investigators use the following displacement function:

$$\theta(z) = C_0 + C_1 z + C_2 \cosh\frac{kz}{\ell} + C_3 \sinh\frac{kz}{\ell}. \tag{15}$$

Table 1. THREE DIFFERENT FINITE ELEMENTS FOR ST. VENANT'S TORSIONAL ANALYSIS

ELEMENT		DISPLACEMENT FUNCTION ASSUMED
	TRIANGULAR PRISM ELEMENT	$w = C_0 + C_1 x + C_2 y$
	RECTANGULAR PRISM ELEMENT	$w = C_0 + C_1 x + C_2 y + C_3 xy$
	PLATE STRIP ELEMENT	$w = C_0 + C_1 \rho$

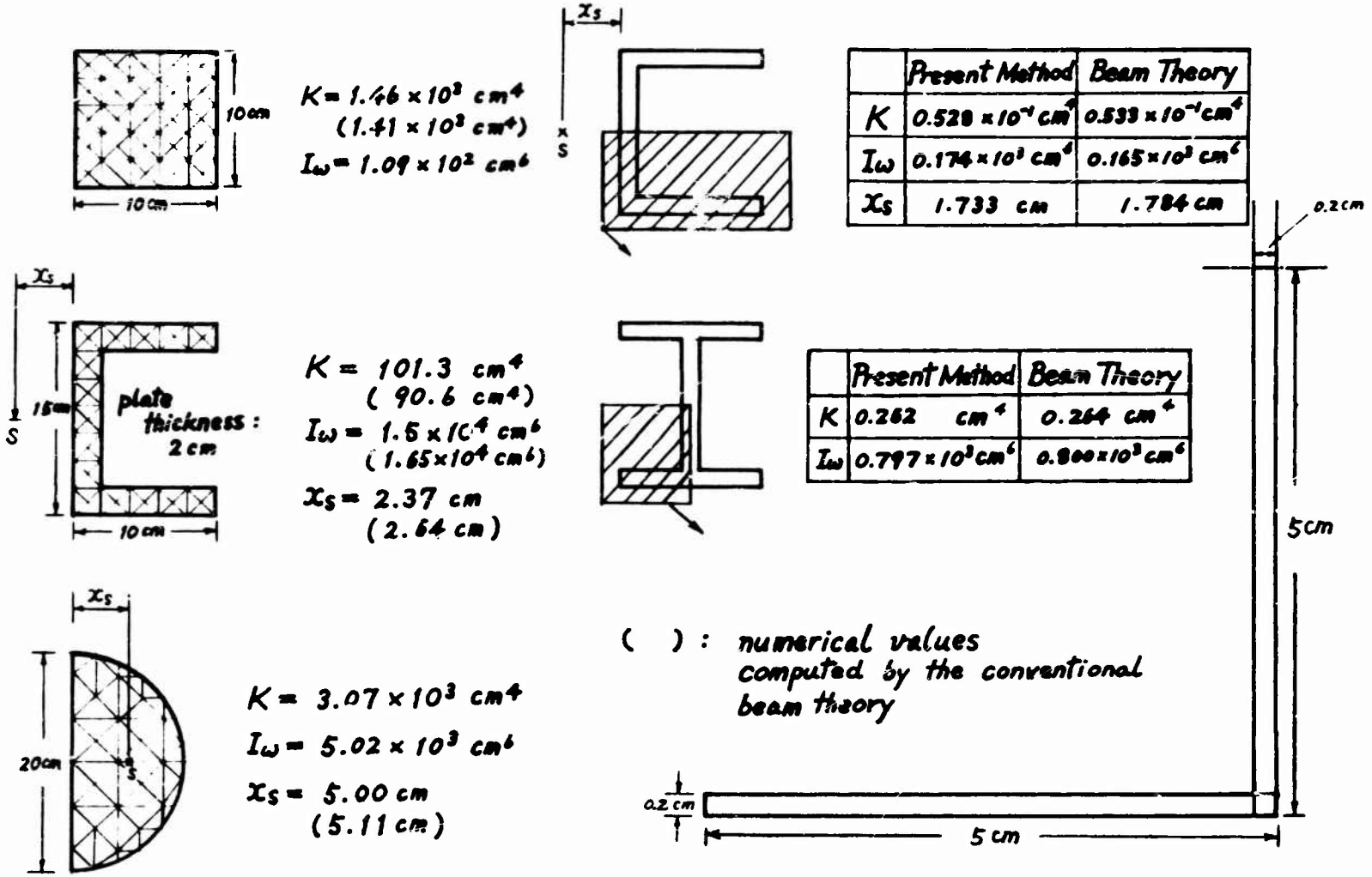


Fig. 3 SOME NUMERICAL EXAMPLES ON THE TORSIONAL RIGIDITIES OF BEAM CROSS SECTIONS WITH MESH DIVISION

where

$$k^2 = \frac{GK\ell^2}{EI_\omega^{(S)}} \tag{16}$$

This is the general solution of the following homogeneous differential equation:

$$EI_\omega^{(S)}\theta'''(z) - GK\theta'(z) = 0 \tag{17}$$

Professor J. [illegible]vris has shown in his recent paper[3] that the following stiffness matrix can be derived by using Eq. (15):

$$\begin{bmatrix} M_{EL} \\ T_{EL} \\ M_{ER} \\ T_{ER} \end{bmatrix} = -\frac{EI_\omega^{(S)}}{D} \begin{bmatrix} \left(\frac{k}{\ell}\right)^3 \sinh k & & \text{SYM.} & \\ \left(\frac{k}{\ell}\right)^2(\cosh k - 1) & \left(\frac{k}{\ell}\right)(k\cosh k - \sinh k) & & \\ -\left(\frac{k}{\ell}\right)^3 \sinh k & -\left(\frac{k}{\ell}\right)^2(\cosh k - 1) & \left(\frac{k}{\ell}\right)^3 \sinh k & \\ \left(\frac{k}{\ell}\right)^2(\cosh k - 1) & -\left(\frac{k}{\ell}\right)(k - \sinh k) & -\left(\frac{k}{\ell}\right)^2(\cosh k - 1) & \left(\frac{k}{\ell}\right)(k\cosh k - \sinh k) \end{bmatrix} \begin{bmatrix} \theta_L \\ \theta_L' \\ \theta_R \\ \theta_R' \end{bmatrix} \tag{18}$$

where $D = k\sinh k - 2(\cosh k - 1)$.

Undoubtedly this is an elaborate stiffness matrix of a beam element for warping-torsion, and indeed it is frequently used by many researchers in warping torsional analysis of beams.

From the viewpoint of numerical computation, however, it is rather laborious to manipulate such matrices which are composed of hyperbolic functions. In this paper, T. Kawai proposed a new stiffness matrix of a beam element for warping-torsion by using a polynomial of the 3rd order as follows:[4]

First of all, $\theta(z)$ is assumed in the following form:

$$\theta(z) = [H(z)]\{C\} \tag{19-a}$$

where

$$\left.\begin{aligned} [H(z)] &= [1,\ z,\ z^2,\ z^3] \\ \{C\}^T &= [C_0,\ C_1,\ C_2,\ C_3] \end{aligned}\right\} \tag{19-b}$$

and the nodal displacement $\{\theta\}$ is defined by the following equation:

$$\{\theta\}^T = [\theta_L,\ \theta_L',\ \theta_R,\ \theta_R'] \tag{20}$$

Denoting the length of a beam element by ℓ and taking the origin of the coordinates at the left end of a given element, the following equation can be easily derived:

$$\theta(z) = [H(z)][A^{-1}]\{\theta\} \tag{21-a}$$

where

$$[A^{-1}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -\frac{3}{\ell^2} & -\frac{2}{\ell} & \frac{3}{\ell^2} & -\frac{1}{\ell} \\ \frac{2}{\ell^3} & \frac{1}{\ell^2} & -\frac{2}{\ell^3} & \frac{1}{\ell^2} \end{bmatrix} \tag{21-b}$$

Following the standard method of deriving the stiffness matrix of a beam element for bending, the following equation of strain energy variation can be easily obtained:

$$\delta V = \int_0^\ell E I_\omega^{(s)} \theta'' \delta\theta'' dz + \int_0^\ell GK\theta'\delta\theta' dz$$
$$= \delta\{\theta\}^T [K_\omega]\{\theta\} + \delta\{\theta\}^T [K_{st}]\{\theta\} \tag{22-a}$$

where

$$\left.\begin{aligned} [K_\omega] &= [A^{-1}]^T [\bar{K}_\omega][A^{-1}] \\ [K_{st}] &= [A^{-1}]^T [\bar{K}_{st}][A^{-1}] \end{aligned}\right\} \tag{22-b}$$

and

$$\left.\begin{aligned} [\bar{K}_\omega] &= \int_0^\ell [H''(z)]^T E I_\omega^{(s)} [H''(z)] dz \\ [\bar{K}_{st}] &= \int_0^\ell [H'(z)]^T GK [H'(z)] dz \end{aligned}\right\} \tag{22-C}$$

Carrying out the calculation, the following matrices can be obtained:

or

$$\begin{bmatrix} M_{zL} \\ T_{zL} \\ M_{zR} \\ T_{zR} \end{bmatrix} = \begin{bmatrix} \frac{12EI_\omega^{(s)}}{\ell^3} & & & \text{SYM.} \\ \frac{6EI_\omega^{(s)}}{\ell^2} & \frac{4EI_\omega^{(s)}}{\ell} & & \\ -\frac{12EI_\omega^{(s)}}{\ell^3} & -\frac{6EI_\omega^{(s)}}{\ell^2} & \frac{12EI_\omega^{(s)}}{\ell^3} & \\ \frac{6EI_\omega^{(s)}}{\ell^2} & \frac{2EI_\omega^{(s)}}{\ell} & -\frac{6EI_\omega^{(s)}}{\ell^2} & \frac{4EI_\omega^{(s)}}{\ell} \end{bmatrix} \begin{bmatrix} \theta_L \\ \theta_L' \\ \theta_R \\ \theta_R' \end{bmatrix} \tag{23-a}$$

$$\{F_\omega\} = [K_\omega]\{\theta\} \tag{23-b}$$

$$\begin{bmatrix} M_{zL} \\ T_{zL} \\ M_{zR} \\ T_{zR} \end{bmatrix} = \begin{bmatrix} \frac{6GK}{5\ell} & & \text{SYM.} & \\ \frac{GK}{10} & \frac{2GK\ell}{15} & & \\ -\frac{6GK}{5\ell} & \frac{GK}{10} & \frac{6GK}{5\ell} & \\ \frac{GK}{10} & -\frac{GK\ell}{30} & -\frac{GK}{10} & \frac{2GK\ell}{15} \end{bmatrix} \begin{bmatrix} \theta_L \\ \theta_L' \\ \theta_R \\ \theta_R' \end{bmatrix} \qquad (24\text{-}a)$$

or

$$\{F_{st}\} = [K_{st}]\{\Theta\} \qquad (24\text{-}b)$$

Denoting the nodal force vector and stiffness matrix of a beam element for warping-torsion as $\{F_\theta\}$ and $[K_\theta]$ respectively the following stiffness equation for warping-torsion can be obtained as follows:

$$\{F_\theta\} = [K_\theta]\{\Theta\} \qquad (25\text{-}a)$$

in which

$$\{F_\theta\} = \{F_\omega\} + \{F_{st}\} \qquad (25\text{-}b)$$

$$[K_\theta] = [K_\omega] + [K_{st}] \qquad (25\text{-}c)$$

and the complete form of $[K_\theta]$ will be seen in Appendix (I). Application of the newly derived stiffness matrix $[K_\theta]$ for warping-torsion to some simple beam torsion problem will show superiority of this matrix to the stiffness matrix defined by Eq. (18).

2.3 General Stiffness Matrix of a Straight Beam Element with a Constant Cross Section which is Defined with Respect to an Arbitrary Point $P(x_p, y_p)$ [5]

In order to obtain the general stiffness matrix, the stiffness matrix of a beam element can be formulated by simply assembling bending stiffness matrices with respect to principal axes passing through centroid $[K_u]$, $[K_v]$, the warping torsional stiffness $[K_\theta]$ and the stiffness matrix of axial deformation $[K_w]$ as follows:

$$\left[\begin{array}{c} f_u \\ \hline f_v \\ \hline f_w \\ \hline f_\theta \end{array}\right] = \left[\begin{array}{c|c|c|c} [K_u] & 0 & 0 & 0 \\ \hline 0 & [K_v] & 0 & 0 \\ \hline 0 & 0 & [K_w] & 0 \\ \hline 0 & 0 & 0 & [K_\theta] \end{array}\right] \left[\begin{array}{c} u_s \\ \hline v_s \\ \hline w_s \\ \hline \theta \end{array}\right] \qquad (26)$$

It should be mentioned that the lateral displacements (u_s , v_s) of a beam element are defined with respect to the shear center, while displacement w is the average axial displacement of a given section and that it is not necessarily equal to w_s.
$[K_u]$, $[K_v]$, $[K_w]$ and $[K_\theta]$ are given in Appendix (I).
Now the general stiffness matrix which is defined with respect to an arbitrary point P can be obtained by a standard transformation procedure of stiffness matrices.
In order to obtain the transformation matrix $[T]$, Eq. (1) is slightly modified and given by the following equations:

$$\left.\begin{aligned} U(x,y,z) &= u_s(z) - (y-y_s)\theta(z) \\ V(x,y,z) &= v_s(z) + (x-x_s)\theta(z) \\ W(x,y,z) &= w(z) - x u_s'(z) - y v_s'(z) + \theta'(z)(\omega_n(x,y) \\ &\quad - x y_s + x_s y) \end{aligned}\right\} \qquad (27)$$

where ($u_s(z)$, $v_s(z)$) is the lateral displacement vector of the shear center.
From Eq. (27) the following equations can be easily derived:

$$\left.\begin{aligned} u_s(z) &= u_p(z) + y_{ps}\,\theta(z) \\ v_s(z) &= v_p(z) - x_{ps}\,\theta(z) \\ w(z) &= w_p(z) + x_p u_p'(z) + y_p v_p'(z) - \theta'(z)\,\omega_{np} \\ u_s'(z) &= u_p'(z) - y_{ps}\,\theta'(z) \\ v_s'(z) &= v_p'(z) - x_{ps}\,\theta'(z) \end{aligned}\right\} \qquad (28\text{-}a)$$

where

$$\left.\begin{aligned} x_{ps} &= x_p - x_s \;, \quad y_{ps} = y_p - y_s \\ \omega_{np} &= \omega_n(x_p, y_p) \end{aligned}\right\} \qquad (28\text{-}b)$$

($u_p(z)$, $v_p(z)$, $w_p(z)$) is the displacement vector of an arbitrary point P on a given cross section.
Using Eq. (28), the following transformation matrix can be easily obtained:

$$\begin{bmatrix} u_s \\ v_s \\ w \\ \theta \end{bmatrix} = \left[\begin{array}{c|c|c|c} [\mathbf{I}] & 0 & 0 & y_{ps}[\mathbf{I}] \\ \hline 0 & [\mathbb{I}] & 0 & -x_{ps}[\mathbf{I}] \\ \hline [t_x] & [t_y] & [\mathbb{I}] & [t_\omega] \\ \hline 0 & 0 & 0 & [\mathbf{I}] \end{array}\right] \begin{bmatrix} u_p \\ v_p \\ w_p \\ \theta \end{bmatrix} \tag{29}$$

where $[\mathbf{I}]$ is an unit matrix of (4 x 4) or (2 x 2) and rectangular matrices $[t_x]$, $[t_y]$ and $[t_\omega]$ are defined by the following equations:

$$\left.\begin{aligned} [t_x] &= \left[\begin{array}{c|c|c|c} 0 & 0 & x_p & 0 \\ \hline 0 & 0 & 0 & x_p \end{array}\right] \\ [t_y] &= \left[\begin{array}{c|c|c|c} 0 & 0 & y_p & 0 \\ \hline 0 & 0 & 0 & y_p \end{array}\right] \\ [t_\omega] &= \left[\begin{array}{c|c|c|c} 0 & 0 & -\omega_{xp} & 0 \\ \hline 0 & 0 & 0 & -\omega_{xp} \end{array}\right] \end{aligned}\right\} \tag{30}$$

Writing Eqs.(26) and (29) in the following form:

$$\{f\} = [\overline{\mathbb{K}}]\{d\} \tag{31}$$

$$\{d\} = [\mathbb{T}]\{d_p\} \tag{32}$$

where

$$\left.\begin{aligned} \{f\}^T &= [\, f_u,\ f_v,\ f_w,\ f_\theta \,] \\ \{d\}^T &= [\, u_s,\ v_s,\ w,\ \theta \,],\ \{d_p\}^T = [\, u_p,\ v_p,\ w_p,\ \theta \,] \end{aligned}\right\} \tag{33-a}$$

$$[\overline{\mathbb{K}}] = \begin{bmatrix} [\mathbb{K}_u] & 0 & 0 & 0 \\ 0 & [\mathbb{K}_v] & 0 & 0 \\ 0 & 0 & [\mathbb{K}_w] & 0 \\ 0 & 0 & 0 & [\mathbb{K}_\theta] \end{bmatrix}, \quad [\mathbb{T}] = \begin{bmatrix} [\mathbf{I}] & 0 & 0 & y_{ps}[\mathbb{I}] \\ 0 & [\mathbf{I}] & 0 & -x_{ps}[\mathbb{I}] \\ [t_x] & [t_y] & [\mathbf{I}] & [t_\omega] \\ 0 & 0 & 0 & [\mathbf{I}] \end{bmatrix} \tag{33-b}$$

the general stiffness matrix $[K]$ can be finally obtained as follows:

$$[K] = [T]^T[\bar{K}][T]$$

$$= \left[\begin{array}{c|c|c|c} [k_{11}] & [k_{12}] & [k_{13}] & [k_{14}] \\ \hline [k_{21}] & [k_{22}] & [k_{23}] & [k_{24}] \\ \hline [k_{31}] & [k_{32}] & [k_{33}] & [k_{34}] \\ \hline [k_{41}] & [k_{42}] & [k_{43}] & [k_{44}] \end{array}\right] \tag{34}$$

where $[k_{ij}]$ (i, j = 1,2,3,4) are defined by the following equations:

$$\left.\begin{array}{ll}
[k_{11}] = [K_u] + [t_x]^T[K_w][t_x], & [k_{12}] = [t_x]^T[K_w][t_y] \\
[k_{13}] = [t_x]^T[K_w], & [k_{14}] = y_{ps}[K_u] + [t_x]^T[K_w][t_\omega] \\
[k_{21}] = [t_y]^T[K_w][t_x], & [k_{22}] = [K_v] + [t_y]^T[K_w][t_y] \\
[k_{23}] = [t_y]^T[K_w], & [k_{24}] = -x_{ps}[K_v] + [t_y]^T[K_w][t_\omega] \\
[k_{31}] = [K_w][t_x], & [k_{32}] = [K_w][t_y] \\
[k_{33}] = [K_w], & [k_{34}] = [K_w][t_\omega] \\
[k_{41}] = y_{ps}[K_u] + [t_\omega]^T[K_w][t_x], & \\
[k_{42}] = -x_{ps}[K_v] + [t_\omega]^T[K_w][t_y], & [k_{43}] = [t_\omega]^T[K_w] \\
[k_{44}] = [K_\theta] + y_{ps}^2[K_u] + x_{ps}^2[K_v] + [t_\omega]^T[K_w][t_\omega] &
\end{array}\right\} \tag{35}$$

The final explicit form of the general stiffness matrix $[K]$ is given in Appendix (II).

2.4 Evaluation of Additional Stiffness of Transverse Stiffening Members such as Bulkheads, Ring Frames or Cross Decks

In actual ship structures it is a standard practice to put transverse stiffening members such as transverse rings, bulkheads or cross decks in order to increase their stiffness. Additional stiffness matrix of these stiffeners $[K_s]$ can be computed by the following standard formulae:

(see Fig. 4)
$$[K_s] = [\Phi]^T [\bar{K}_s] [\Phi] \tag{36}$$

in which $[\bar{K}_s]$ is the stiffness matrix with respect to the local coordinates taken on the stiffeners and they are evaluated by basing on the simple beam theory. The transformation matrix $[\Phi]$ is also given by the following expression:

$$[\Phi] = [\Lambda] [\Omega] \tag{37}$$

in which $[\Lambda]$ is the angular transformation matrix between coordinates of a stiffener and coordinates of a given beam element, and it is given as follows:

$$[\Lambda] = \left[\begin{array}{c|c} \lambda_\ell & 0 \\ \hline 0 & \lambda_r \end{array}\right] \tag{38}$$

$[\lambda_\ell]$, $[\lambda_r]$ are the angular transformation matrices at the point Q and R respectively. While $[\Omega]$ is the point transformation matrix between the points Q or R and the beam nodal points. It is generally given by the following equation:

$$[\Omega] = \left[\begin{array}{c} \omega_\ell \\ \hline \omega_r \end{array}\right] \tag{39}$$

in which $[\omega_\ell]$ is defined by the following matrix equation:

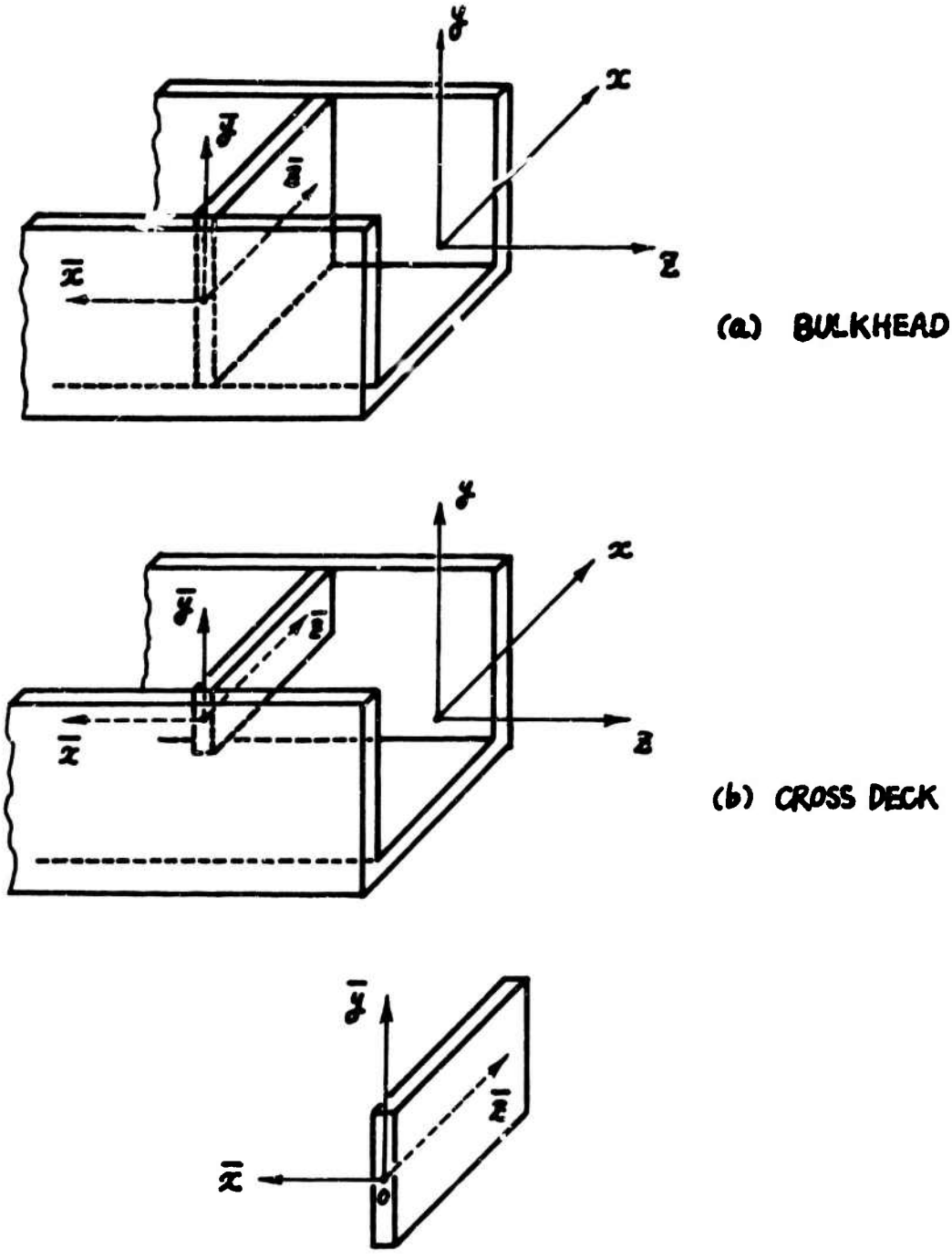


Fig. 4. CALCULATION OF ADDITIONAL STIFFNESS MATRICES OF CROSS DECKS OR BULKHEADS

$$\begin{Bmatrix} U_q \\ V_q \\ W_q \\ \theta_{xq} \\ \theta_{yq} \\ \theta_{zq} \end{Bmatrix} = \left[\begin{array}{c|c|c|c} N(z) & 0 & 0 & -y_{qs}N(z) \\ \hline 0 & N(z) & 0 & x_{qs}N(z) \\ \hline -x_q N(z) & -y_q N(z) & n(z) & \omega_{nq} N(z) \\ \hline 0 & -N'(z) & 0 & x_s + \frac{1}{2}\Gamma_{xq} \\ \hline N'(z) & 0 & 0 & y_s - \frac{1}{2}\Gamma_{yq} \\ \hline 0 & 0 & 0 & N(z) \end{array}\right] \begin{Bmatrix} u_s \\ \hline v_s \\ \hline w \\ \hline \theta \end{Bmatrix} \qquad (40)$$

$(\theta_{xq}, \theta_{yq}, \theta_{zq})$ is the angular displacement vector of a point Q and it is defined by the following equation:

$$\left.\begin{aligned} \theta_x &= \frac{1}{2}\left(\frac{\partial W}{\partial y} - \frac{\partial V}{\partial z}\right) \\ \theta_y &= \frac{1}{2}\left(\frac{\partial U}{\partial z} - \frac{\partial W}{\partial x}\right) \\ \theta_z &= \frac{1}{2}\left(\frac{\partial V}{\partial x} - \frac{\partial U}{\partial y}\right) \end{aligned}\right\} \qquad (41)$$

where

$$\left.\begin{aligned} [N(z)] &= [(1 - 3\xi^2 + 2\xi^3),\ \ell(\xi - 2\xi^2 + 3\xi^3),\ (3\xi^2 - 2\xi^3),\ \ell(-\xi^2 + \xi^3)] \\ [n(z)] &= [(1-\xi),\ \xi] \qquad \xi = z/\ell \end{aligned}\right\} \qquad (41\text{-}a)$$

$$\left.\begin{aligned} x_{qs} &= x_q - x_s \quad , \quad y_{qs} = y_q - y_s \\ \omega_{nq} &= \omega_n(x_q, y_q) - x_q y + x y_q \\ \Gamma_{xq} &= \left(\frac{\partial \omega_n}{\partial x}\right)_q - x_q \quad , \quad \Gamma_{yq} = \left(\frac{\partial \omega_n}{\partial y}\right)_q + y_q \end{aligned}\right\} \qquad (41\text{-}b)$$

It should be mentioned that stiffness matrices of stiffeners thus calculated $[K_s]$ are defined with respect to the shear center of a given beam element. Therefore, $[K_s]$ should be added to the stiffness matrix of a given beam element before applying the point transformation matrix $[T]$ between the shear center and an arbitrary point P. At this point, it is interesting to note that reduction of the element stiffness due to cutouts of openings can be estimated exactly in the same manner, but in this case, $[K_s]$ should be considered the negative stiffness.

2.5 Finite Element Analysis of a Beam with Variable Cross Section

Once the stiffness matrices which are defined with respect to an arbitrary point P of a given section are obtained, it is a standard practice in the displacement method of analysis to obtain the following overall stiffness equation:

$$[K]\{d\} = \{F\} \tag{42}$$

That is, formulation of overall stiffness matrix can be made after transformation of element stiffness matrices from local to global coordinates. It should be worthwhile to mention the method how to restrain the rigid body movement in case of unrestrained structures such as ships, aircrafts and so on. In case of statical analysis, such a rigid body movement can be easily eliminated only by suppressing the displacement vector of a certain nodal point, i.e. $u_p = v_p = w_p = u_p' = v_p' = \theta = 0$, and it is clear that such operation does not produce any stress in a given structure. In the following Section, torsional analysis of a container ship model and actual ship will be described. In case of ship structures, the cross section is symmetric with respect to y axis and only distributed torque will be considered as external forces in these cases. Therefore the overall equation can be written in the following form:

$$\begin{bmatrix} [K_{uu}] & 0 & 0 & [K_{u\theta}^{(P)}] \\ 0 & [K_{vv}] & [K_{vw}] & 0 \\ 0 & [K_{wv}] & [K_{ww}] & 0 \\ [K_{\theta u}^{(P)}] & & 0 & [K_{\theta\theta}^{(P)}] \end{bmatrix} \begin{bmatrix} u_p \\ v_p \\ w_p \\ \theta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ F \end{bmatrix} \tag{43}$$

From this matrix equation, the following equations are finally derived:

$$\left.\begin{aligned} [K_{uu}]\{u_p\} + [K_{u\theta}^{(P)}]\{\theta\} &= 0 \\ [K_{\theta u}^{(P)}]\{u_p\} + [K_{\theta\theta}^{(P)}]\{\theta\} &= \{F\} \end{aligned}\right\} \tag{44}$$

with $u_p = w_p = 0$

Eliminating $\{\ddot{u}_p\}$ from both equations of Eq. (44), the following equation for warping torsion will be derived:

$$([K_{\theta\theta}^{(P)}] - [K_{\theta u}^{(P)}][K_{uu}^{-1}][K_{u\theta}^{(P)}])\{\theta\} = \{F\} \tag{45}$$

Solving Eq. (45), torsional deformation of a given structure can be determined. And the associated horizontal movement can be also determined by the following equation:

$$\{u_p\} = -[K_{uu}^{-1}][K_{u\theta}^{(P)}]\{\theta\} \tag{46}$$

SECTION III

APPLICATION OF THE PRESENT METHOD TO THIN-WALLED STRUCTURES

In order to investigate the effectiveness and applicability of the present method to stress analysis of actual thin-walled structures, the finite element analysis on the torsional strength of a container ship model was made and a fairly good agreement between results of the present finite element analysis and experiment was confirmed.
Then overall torsional analysis of an actual container ship was carried out and again the result of analysis based on the present method was found consistent with experimental data obtained by the full-scale testing of the same ship.

3.1 Statical Analysis of a Container Ship Model

Container ships would be typical examples of thin-walled structures where the torsional strength is considered one of important design parameters. K. Enomoto and others previously conducted experimental and theoretical studies on this problem by using a ship model as shown in Fig. 5.[6]
Fig. 5 shows the mesh division of the model structure.
Furthermore, effects of cross beams or transverse members on the overall stiffness of a given structure were considered in this analysis and comparative study between theory and experiment was conducted for the ship model with and without eleven bulkheads. The detail of the bulkhead structure is also shown in Fig. 5.

3.1.1. Evaluation of Torsional Stiffness of a given Ship Model

This model ship has four different types of cross section shown in Fig. 5. Evaluation of these section properties was made by using the computing program for analysis of St. Venant's torsion problem as indicated in Section II. It is found that evaluation of torsional stiffness can be made easily by using very coarse mesh division with rectangular prism elements shown in Table 1.

3.1.2. Experiment

The ship model was fixed to a rigid base at the point (1) in Fig. 5 and concentrated twisting moment was applied at the point (23).
The applied torque was measured by a load cell and the deformation of the ship model was measured by dial gages placed at several different locations and from the results of which the angles of rotation were calculated.

3.1.3. Comparison of Analytical and Experimental Results

In Fig. 6 the result of analysis is shown in comparison with the experimental data, and result of the convertional finite element analysis is also shown in Fig. 6.
In view of a fairly good agreement between theory and test results it may be concluded that the finite element analysis based on the present method could be a simple and yet reliable method as far as the overall structural analysis is concerned.

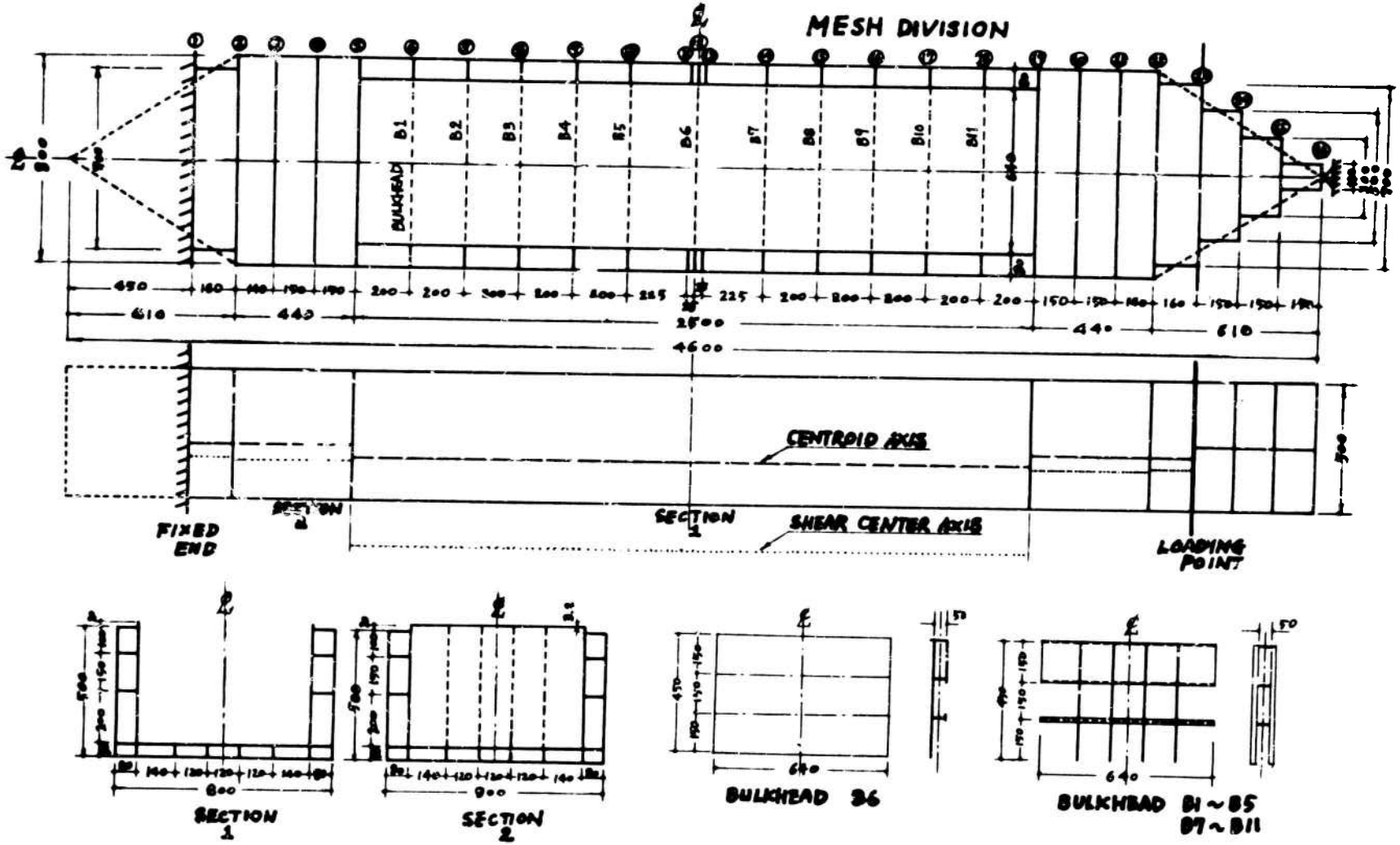


Fig. 5. CONTAINER SHIP MODEL (TAMANO RESEARCH INSTITUTE, MITSUI SHIP BUILDING CO.) IN MM SCALE 1/10

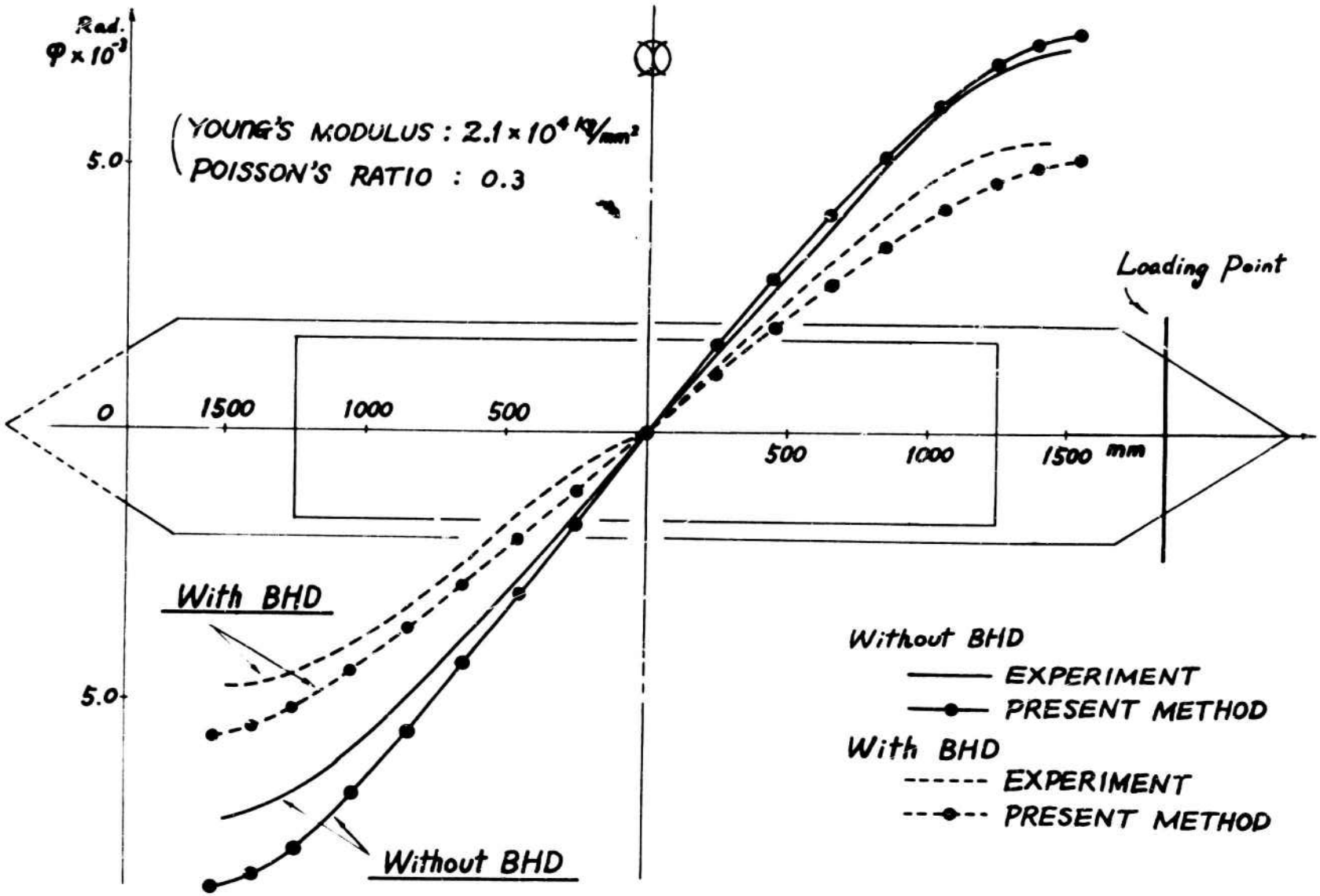


Fig. 6. TORSION TESTING OF A CONTAINER SHIP MODEL

3.2 Application of the Present Method to Analysis of an Actual Ship Test

Full-scale torsion test of a newly built container ship, whose name is "Australia-Maru" (total length 212,992 MM, breadth 29,000 MM, depth 16,3000 MM, draft 10,500 MM) was previously conducted at the Tamano Shipyard of the Mitsui Shipbuilding Company and at the same time conventional finite element analysis was carried out to check consistency between the theory and experiment.[47] In order to confirm reliability of the calculated results when the present method is applied to analysis of actual structures, the same test result was analyzed again. (See Fig. 7-a,b,c)

3.2.1. Full-scale Testing of the "Australia-Maru"

Torsional loading was applied to the "Australia-Maru" by filling water in the ballast tanks. More precisely, two antisymmetric but equal loading conditions with respect to the vertical plane passing through the centerline of the ship were considered. In each loading condition, water was filled into the ballast tanks such that she may be not only in self-equilibrium but also under combind bending and torsional loading. Therefore pure torsional deformation can be obtained by simply taking the difference of measurement data corresponding to each loading condition. Fig. 8 shows the distribution of torsional moment.

3.2.2. Measurement of Torsional Deformation and Horizontal Displacement

Horizontal displacement of the ship hull was measured by using laser beam on the starboard deck and differences of vertical displacements on both side of the ship were also measured by using the manometers at eleven stations. Using measurement data by the manometers the rotation angle of each cross section φ can be computed by the following equation:

$$\varphi = \sin^{-1}\left(\frac{\Delta}{b}\right) \doteq \frac{\Delta}{b} \qquad (47)$$

where Δ : difference of water height on port and starboard
b : distance between manometers placed on port and starboard

In order to conduct the full-scale testing under the same ambient condition, sea, air, and hull temperature histories were continuously measured by Cu-Con thermocouples. Measurements of wind and draft were also made. During the test effects of rolling and heeling motion of the ship were carefully controlled.

3.2.3. Finite Element Torsional Analysis of the "Australia-Maru" based on the Present Method

Basing on the present method, torsional analysis of the "Australia-Maru" was carried out by dividing the whole structure into 25 beam elements as shown in Fig. 7 and evaluation of sectional properties of each element was made by using the computer program for shear flow analysis of St. Venant's torsion problem. Fig. 9 shows two typical examples of these section properties evaluation. The calculated distribution of warping-torsional as well as St. Venant's torsional rigidities are shown in Fig. 10.

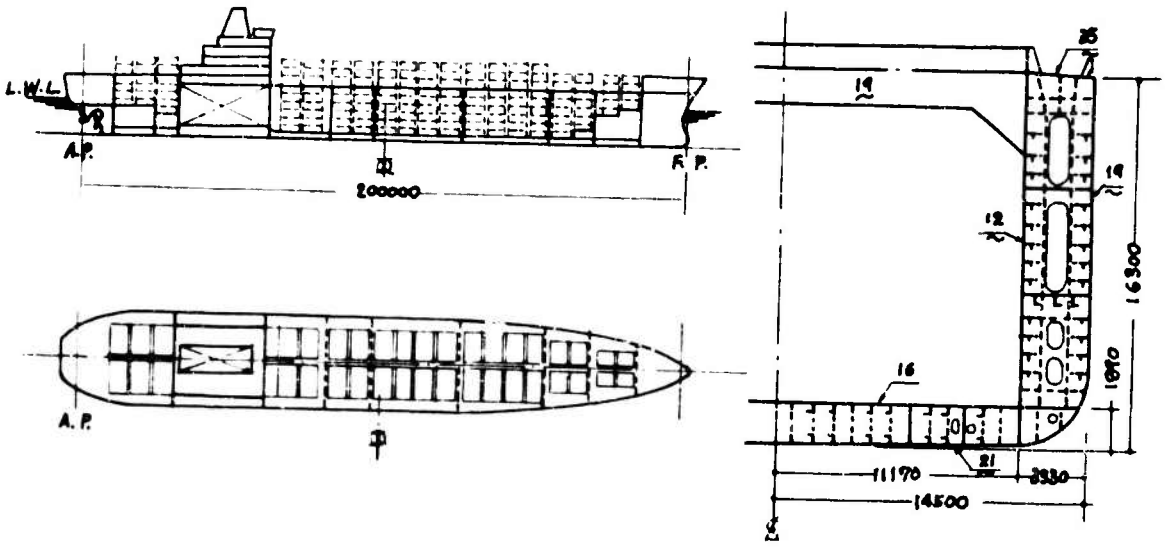


Fig. 7-a GENERAL ARRANGEMENT PLAN AND MIDSHIP SECTION OF THE "AUSTRALIA-MARU"

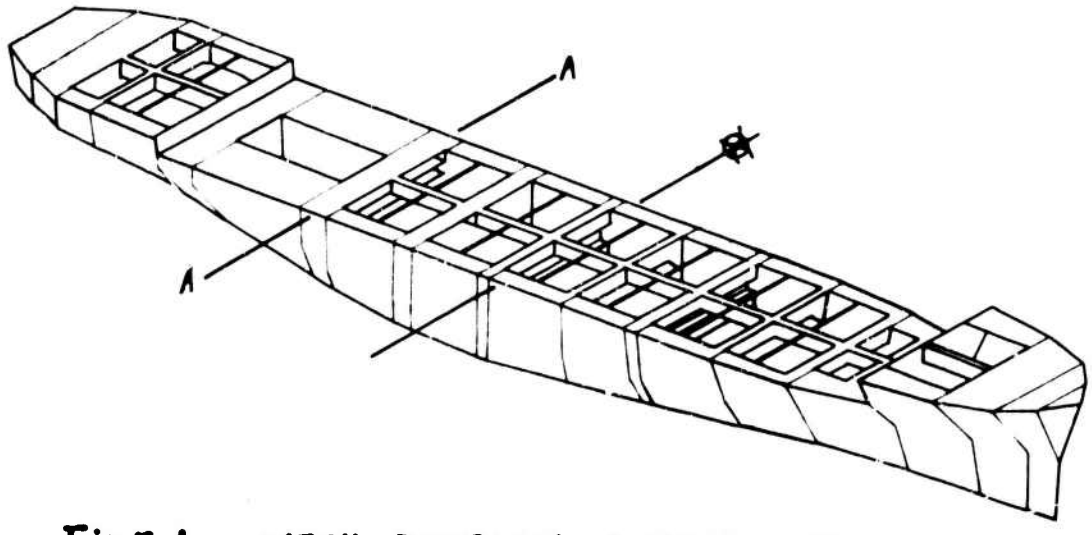
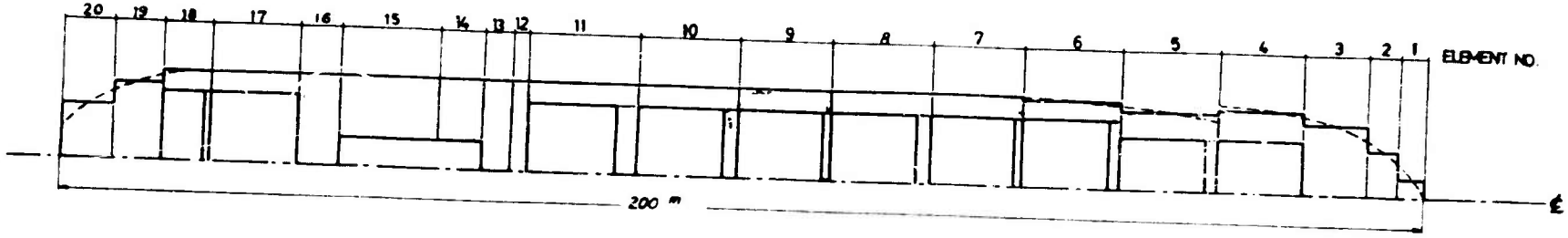
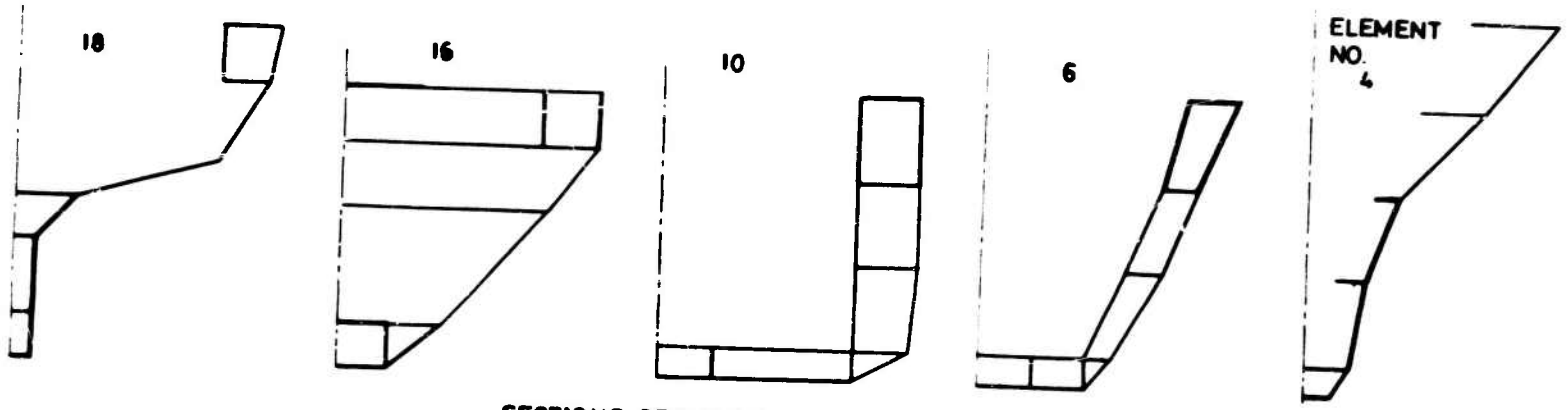


Fig. 7-b MESH DIVISION SCHEME OF THE "AUSTRALIA-MARU"



PLAN



SECTIONS OF THE ELEMENT

FIG. 7-C STRUCTURAL IDEALIZATION OF THE "AUSTRALIA-MARU"

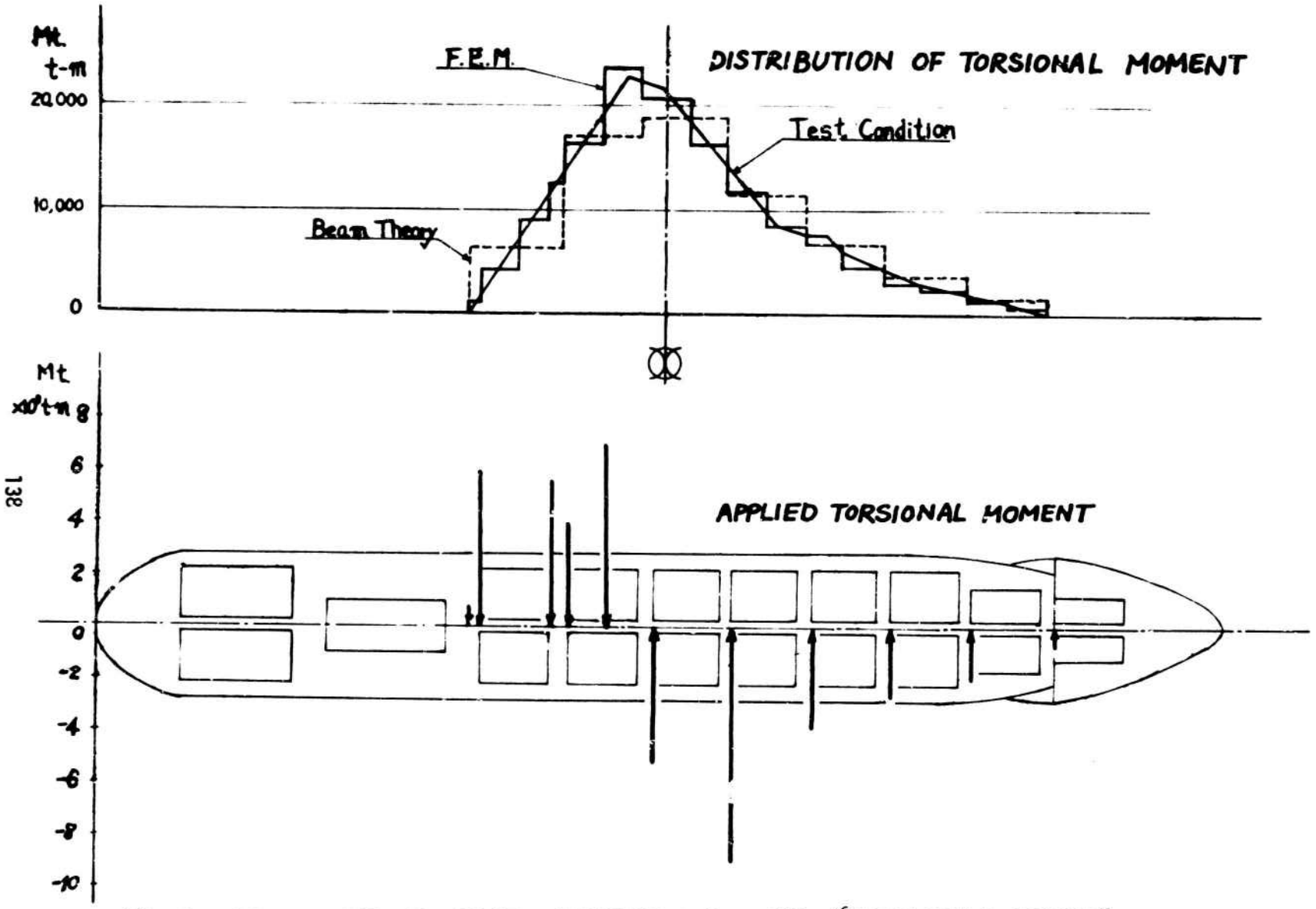


Fig. 8 TORSIONAL MOMENT APPLIED TO THE "AUSTRALIA-MARU" BY MEANS OF WATER BALLAST LOADING

A	0.395×10^{7}	mm^2
Ixx	0.348×10^{15}	mm^4
Iyy	0.141×10^{15}	mm^4
Iω	0.725×10^{21}	mm^6
Iωx	-0.316×10^{12}	mm^6
Iωy	0.0	mm^6
K	0.277×10^{15}	mm^4

NO. of Elements : 56
NO. of Nodes : 100

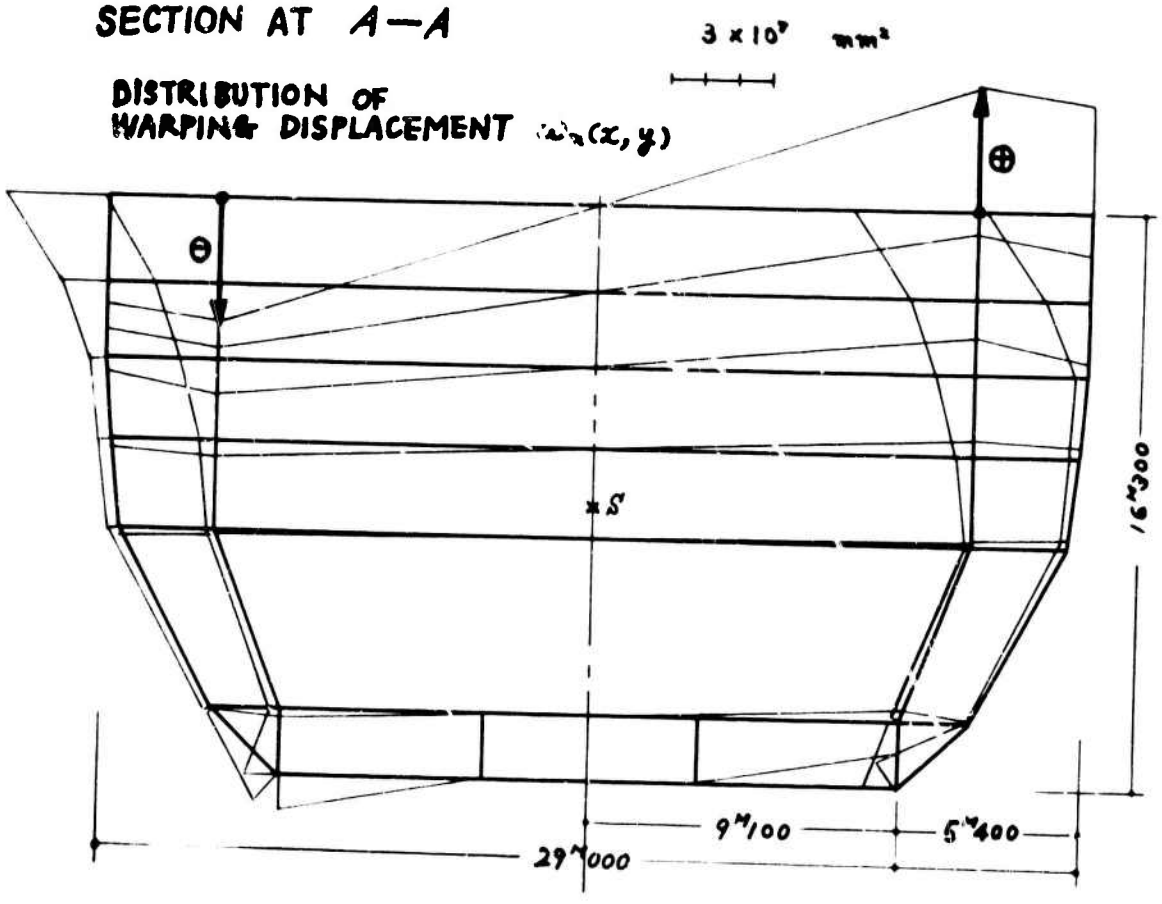


Fig. 9-a EXAMPLES OF SECTION PROPERTIES AT TYPICAL CROSS SECTIONS OF THE "AUSTRALIA-MARU"

A	0.266×10^{7}	mm^2
Ixx	0.337×10^{16}	mm^4
Iyy	0.965×10^{14}	mm^4
Iω	0.712×10^{22}	mm^6
Iωx	-0.445×10^{19}	mm^5
Iωy	0.0	mm^5
K	0.142×10^{14}	mm^4

NO. of Elements : 31
NO. of Nodes : 54

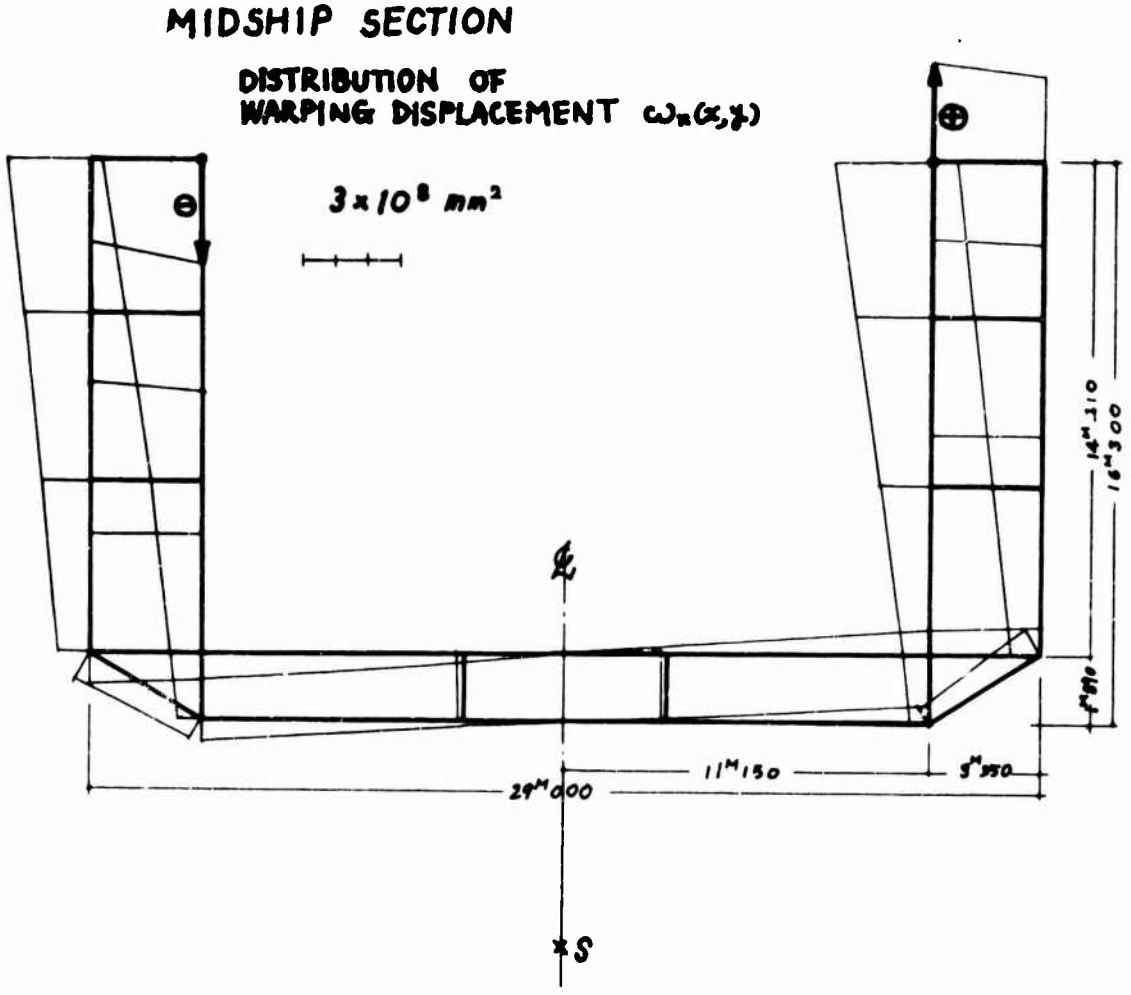


Fig. 9-b EXAMPLES OF SECTION PROPERTIES AT TYPICAL CROSS SECTIONS OF THE "AUSTRALIA-MARU"

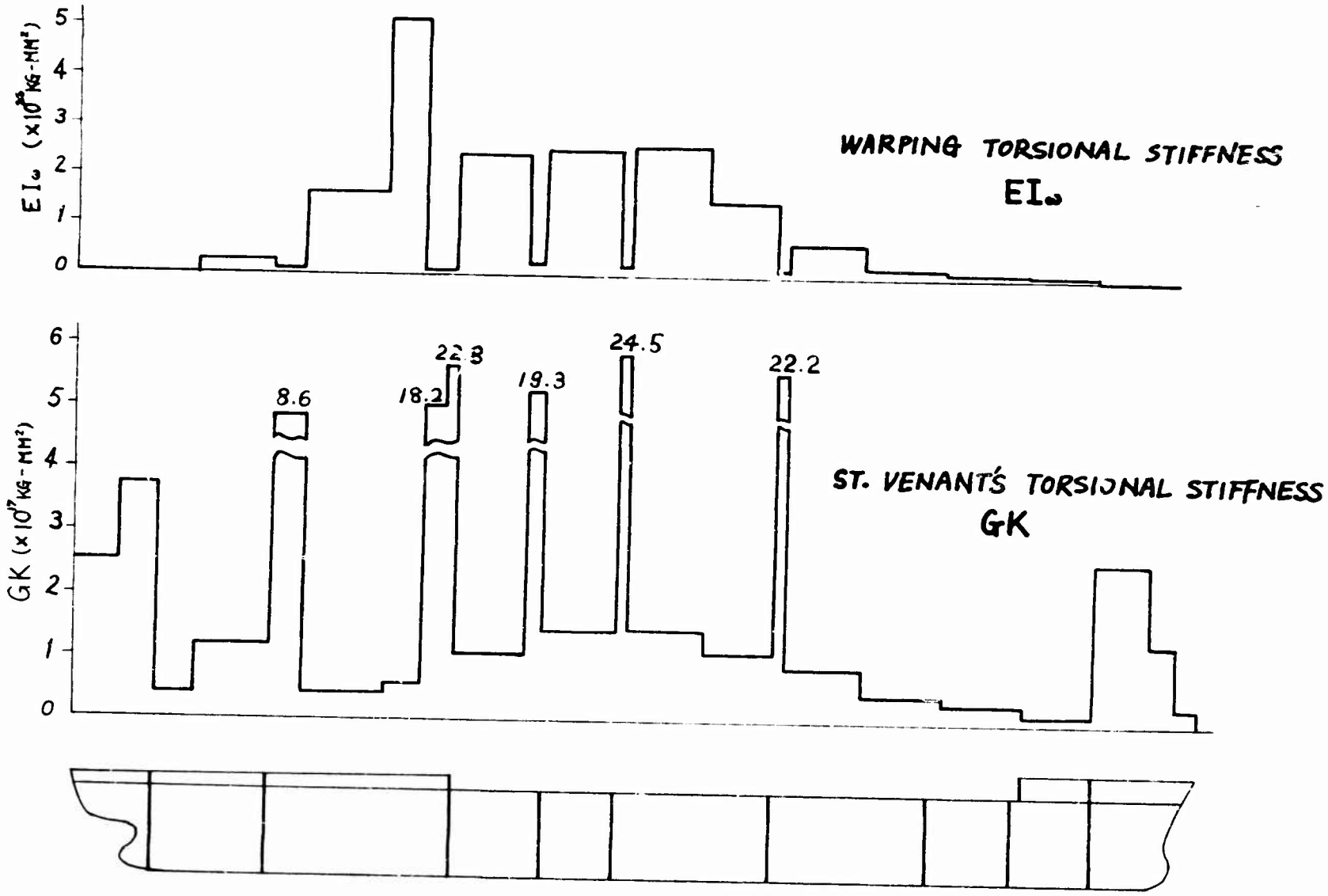


Fig. 10 CALCULATED DISTRIBUTION OF TORSIONAL STIFFNESS OF THE 'AUSTRALIA-MARU"

And distributed twisting moment m_x for one loading condition is approximated by eleven concentrated torque as shown in Fig. 8.
Formulating the overall stiffness equation of 101 x 101, it was solved by the HITAC 5020E (65 KW, 32 bits/W) installed at Main Campus of the University of Tokyo.
Computer-used time for evaluation of 20 section properties and overall torsional analysis are 2 min. 06 sec. and 1 min. 37 sec. respectively.
All input data preparation was made by a design engineer and it took him about 7 days.

3.2.4. Comparison between Result of the Present Analysis and Experimental Data

Calculated distributions of rotation angle and horizontal displacement are shown in Figs. 11 and 12, respectively.
In these figures not only measured data but also results of two theoretical analyses are plotted together in comparison with the results of the present analysis. The one of which is a theoretical analysis made by K. Enomoto, N. Tanaka and others, and it may correspond to the solution of $[K_\phi]\{\theta\} = \{F_\phi\}$ i.e. Eq. (25-a)[9]. In this analysis they obtained the numerical solution of the differential equation governing a given warping-torsion problem by using a digital computer. Another analysis is the conventional finite element analysis made by S. Sanbongi, N. Tanaka and others.[7]
From these figures it can be seen that the result obtained by the present method is in good agreement with measured data of the full-scale testing and the present authors believe high potentiality of the proposed method in stress analysis of actual thin-walled structures.
It should be especially mentioned that the present method of analysis can be made even by using small computers with reasonably low cost and labour for preparation of input data to compare with the conventional finite element method of analysis.

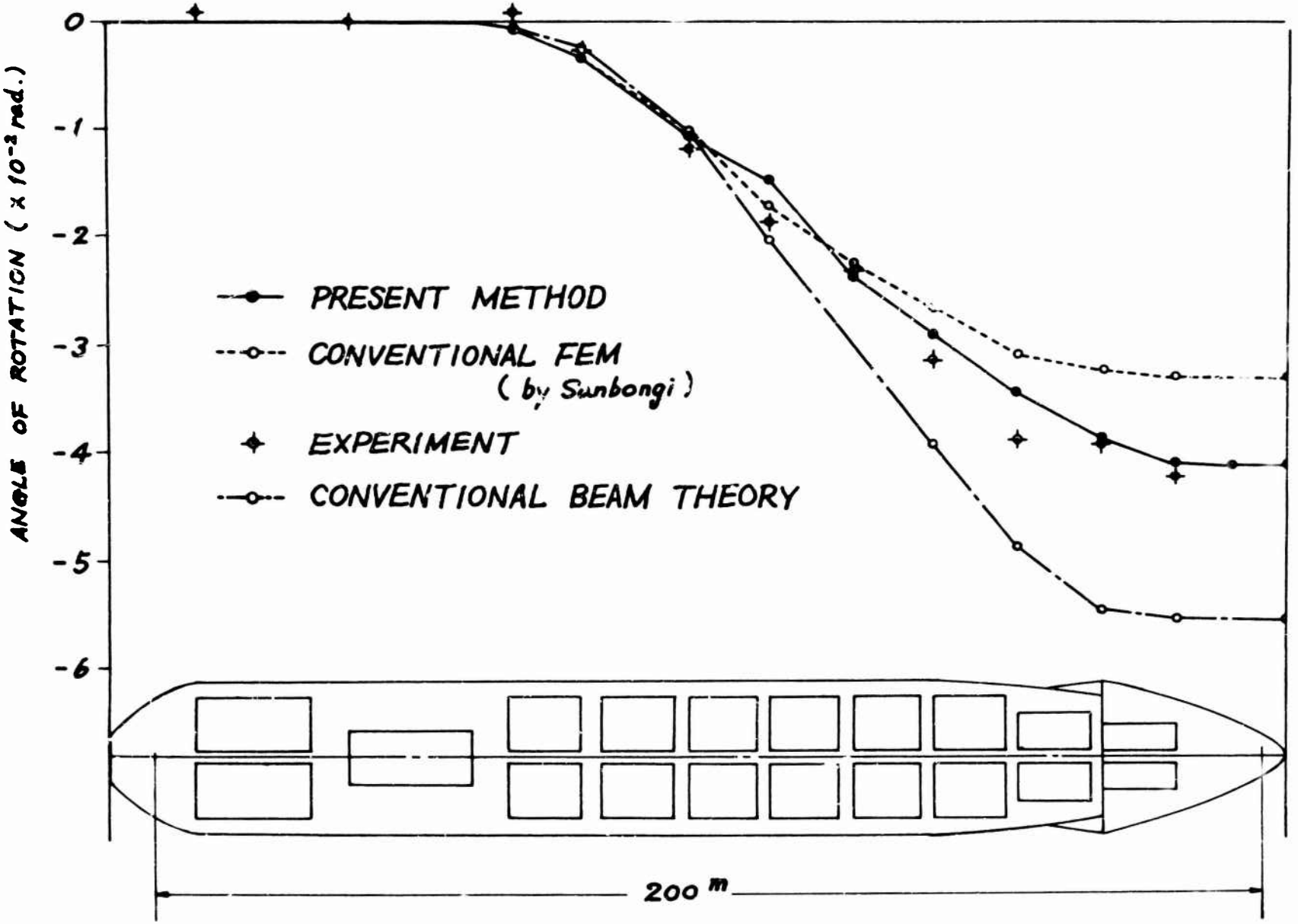


Fig. 11 TORSION TESTING RESULT OF THE "AUSTRALIA-MARU"
(DISTRIBUTION OF ANGLE OF ROTATION)

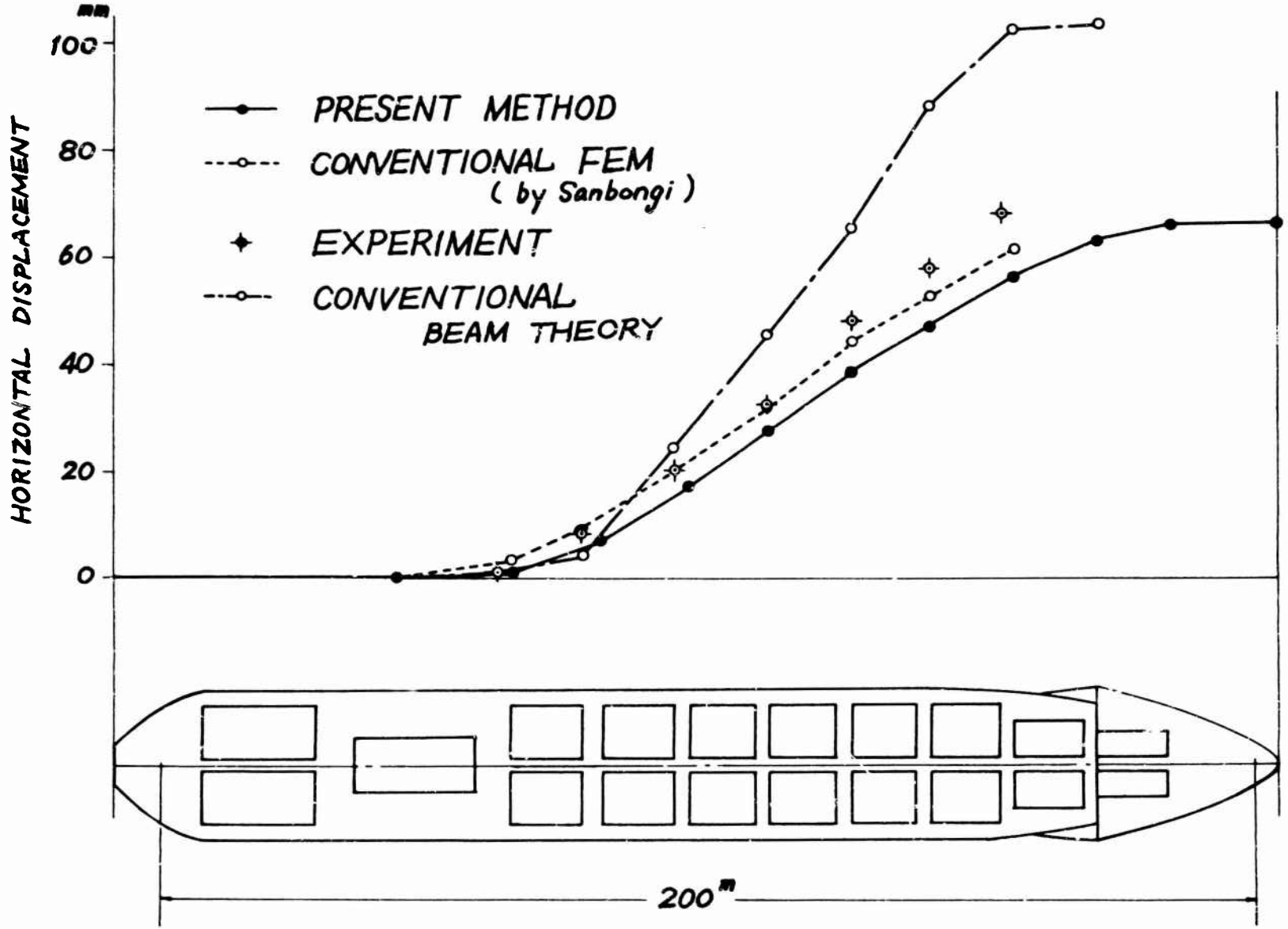


Fig. 12 TORSION TESTING RESULT OF THE "AUSTRALIA-MARU"
(DISTRIBUTION OF HORIZONTAL DISPLACEMENT)

SECTION IV

CONCLUSION

Basing on the modern beam theory, a new finite element method of analysis of thin-walled structures was proposed in this paper.
Applying the present method to stress analysis of a container ship model as well as a real container ship "Australia-Maru", fairly good agreement between theory and experiment was confirmed.
Important feature of the present method may be stated as follows:

(a) Finite element analysis on the St. Venant's torsion problem initiated by Prof. L. R. Hermann is extended to evaluate various torsional constants such as, I_ω , $I_{\omega x}$, $I_{\omega y}$ and location of shear center.

(b) A new stiffness matrix of the beam element for warping-torsion is developed by using the 3rd order polynomial.

(c) Finite element method of analysis of a beam with variable cross section is proposed. This method may be in principle the same as the method proposed by Professor J. H. Argyris, but it may be superior in simplicity of theoretical formulation.

(d) Practical method of estimating effects of cut-out or transverse stiffnening members on the section properties of a given beam structure is proposed.

(e) Size of stiffness equation may seldom exceed several hundreds in the present method so that small computing facilities will be sufficient for the overall analysis of structures at initial design stage.

ACKNOWLEDGEMENT

The authors would like to express their acknowledgement to Mr. M. Irie, the Manager and Mr. M. Nishikido, Assistant to Manager of Technical Development Headquarters, Mitsui Shipbuilding and Engineering Company Ltd. for their constant encouragement and support.
The present authors also wish to express their hearty thanks to many people for their kind co-operation and sacrificing efforts shown in the course of the present study, especially Mr. H. Takeda, Itochu Electronic Computing Center for his kind co-operation, Mr. Tsujioka for his effort in preparation of input data of the "Australia-Maru" and Mr. Y. Fujitani, Mr. S. Shiina and Miss S. Suzuki of Kawai's Laboratory at the Institute of Industrial Science, University of Tokyo for their kind assistance in this manuscript preparation.

SECTION V

REFERENCES

(1) L. R. Herrmann, "Elastic Torsional Analysis of Irregular Shapes" ASCE, Engineering Mechanics Division, Dec. 1965.

(2) T. Kawai and N. Yoshimura, "Finite Element Analysis on the Torsion of a Bar with Uniform Cross Section", Seisan Kenkyu, Vol.20, No.5, May 1968.

(3) J. H. Argyris and D. Radaj, "Steifigkeitsmatrizen dünnwandiger Stäbe und Stabsysteme", Ingenieur-Archiev 40, 1971, S.198-210.

(4) T. Kawai, "Matrix Methods in Vibration and Dynamic Response Analysis of Structures", Series Textbooks on Computerized Structural Engineering I - (4) Baifukan Book Co., Tokyo, Feb. 1971.

(5) T. Kawai and others
   (i) "Warping Torsional Analysis of a Beam with Variable Cross Section"
   (ii) "Vibration Analysis of a Beam with Variable Cross Section Based on the Finite Element Method"
   (iii) "Finite Element Analysis of an Aircraft Wing Structure Based on the Modern Beam Theory"
   (iv) "Shear Deformation Analysis of a Beam based on the Finite Element Method"

   The Proceeding of the J.S.S.C. Symposium on Matrix Method of Structural Analysis and Design, June 1971.

(6) K. Enomoto, "On the Effectiveness of Bulkhead Structure to the Torsional Rigidity of the Container Ship", Jour. Soc. Naval Architects Japan, No.125, 1969.

(7) N. Tanaka, S. Sambongi and T. Kawai, "A Torsional Strength Analysis on the Container Ship by Means of the Finite Element Procedure and Full-scale Testing", International Shipbuilding Progress, Shipbuilding and Marine Engineering Monthly, Vol.18, No.201, May 1971.

APPENDIX (I) STIFFNESS MATRICES OF BEAM ELEMENTS FOR BENDING, TORSIONAL AND AXIAL DEFORMATION

$$[K_u]=\begin{bmatrix} \frac{12EI_{xx}}{l^3} & -\frac{12EI_{xx}}{l^3} & \frac{6EI_{xx}}{l^2} & \frac{6EI_{xx}}{l^2} \\ -\frac{12EI_{xx}}{l^3} & \frac{12EI_{xx}}{l^3} & -\frac{6EI_{xx}}{l^2} & -\frac{6EI_{xx}}{l^2} \\ \frac{6EI_{xx}}{l^2} & -\frac{6EI_{xx}}{l^2} & \frac{4EI_{xx}}{l} & \frac{2EI_{xx}}{l} \\ \frac{6EI_{xx}}{l^2} & -\frac{6EI_{xx}}{l^2} & \frac{2EI_{xx}}{l} & \frac{4EI_{xx}}{l} \end{bmatrix}$$

$$[K_v]=\begin{bmatrix} \frac{12EI_{yy}}{l^3} & -\frac{12EI_{yy}}{l^3} & \frac{6EI_{yy}}{l^2} & \frac{6EI_{yy}}{l^2} \\ -\frac{12EI_{yy}}{l^3} & \frac{12EI_{yy}}{l^3} & -\frac{6EI_{yy}}{l^2} & -\frac{6EI_{yy}}{l^2} \\ \frac{6EI_{yy}}{l^2} & -\frac{6EI_{yy}}{l^2} & \frac{4EI_{yy}}{l} & \frac{2EI_{yy}}{l} \\ \frac{6EI_{yy}}{l^2} & -\frac{6EI_{yy}}{l^2} & \frac{2EI_{yy}}{l} & \frac{4EI_{yy}}{l} \end{bmatrix}$$

$$[K_w]=\begin{bmatrix} \frac{EA}{l} & -\frac{EA}{l} \\ -\frac{EA}{l} & \frac{EA}{l} \end{bmatrix}$$

$$[K_\theta]=\begin{bmatrix} \frac{12EI_{zz}}{l^3}+\frac{6GK}{5l} & -\frac{12EI_{zz}}{l^3}-\frac{6GK}{5l} & \frac{6EI_{zz}}{l^2}+\frac{GK}{10} & \frac{6EI_{zz}}{l^2}+\frac{GK}{10} \\ -\frac{12EI_{zz}}{l^3}-\frac{6GK}{5l} & \frac{12EI_{zz}}{l^3}+\frac{6GK}{5l} & -\frac{6EI_{zz}}{l^2}-\frac{GK}{10} & -\frac{6EI_{zz}}{l^2}-\frac{GK}{10} \\ \frac{6EI_{zz}}{l^2}+\frac{GK}{10} & -\frac{6EI_{zz}}{l^2}-\frac{GK}{10} & \frac{4EI_{zz}}{l}-\frac{2GKl}{15} & \frac{2EI_{zz}}{l}-\frac{GKl}{30} \\ \frac{6EI_{zz}}{l^2}+\frac{GK}{10} & -\frac{6EI_{zz}}{l^2}-\frac{GK}{10} & \frac{2EI_{zz}}{l}-\frac{GKl}{30} & \frac{4EI_{zz}}{l}-\frac{2GKl}{15} \end{bmatrix}$$

APPENDIX (II) GENERAL STIFFNESS MATRIX OF A BEAM ELEMENT DEFINED WITH RESPECT TO AN ARBITRARY NODAL POINT P ON THE BEAM CROSS SECTION

$I_{xy} = y_p I_{xx} - I_{\omega x}$, $I_{yy} = -x_p I_{yy} - I_{\omega y}$, $I_{\omega p} = I_\omega + y_p^2 I_{xx} + x_p^2 I_{yy} - 2y_p I_{\omega x} + 2x_p I_{\omega y}$

	u_L	u_R	u_L'	u_R'	v_L	v_R	v_L'	v_R'	w_L	w_R	θ_L	θ_R	θ_L'	θ_R'
V_{xL}	$\frac{12EI_{xx}}{l^3}$													
V_{xR}	$-\frac{12EI_{xy}}{l^3}$	$\frac{12EI_{xx}}{l^3}$										SYM.		
M_{yL}	$\frac{6EI_{xx}}{l^2}$	$-\frac{6EI_{xx}}{l^2}$	$\frac{4EI_{xy}}{l} + \frac{EAy_p^2}{l}$											
M_{yR}	$\frac{6EI_{xy}}{l^2}$	$-\frac{6EI_{xx}}{l^2}$	$\frac{2EI_{xy}}{l} - \frac{EAy_p^2}{l}$	$\frac{4EI_{xy}}{l} + \frac{EAy_p^2}{l}$										
V_{yL}	0	0	0	0	$\frac{12EI_{yy}}{l^3}$									
V_{yR}	0	0	0	0	$-\frac{12EI_{xy}}{l^3}$	$\frac{12EI_{xy}}{l^3}$								
M_{xL}	0	0	$\frac{EAx_p y_p}{l}$	$-\frac{EAx_p y_p}{l}$	$\frac{6EI_{xx}}{l^2}$	$-\frac{6EI_{xy}}{l^2}$	$\frac{4EI_{xy}}{l} + \frac{EAy_p^2}{l}$							
M_{xR}	0	0	$-\frac{EAx_p y_p}{l}$	$\frac{EAx_p y_p}{l}$	$\frac{6EI_{xy}}{l^2}$	$-\frac{6EI_{yy}}{l^2}$	$\frac{2EI_{xy}}{l} - \frac{EAy_p^2}{l}$	$\frac{4EI_{xx}}{l} + \frac{EAy_p^2}{l}$						
V_{zL}	0	0	$\frac{EAx_p}{l}$	$-\frac{EAx_p}{l}$	0	0	$\frac{EAy_p}{l}$	$-\frac{EAy_p}{l}$	$\frac{EA}{l}$					
V_{zR}	0	0	$-\frac{EAx_p}{l}$	$\frac{EAx_p}{l}$	0	0	$-\frac{EAy_p}{l}$	$\frac{EAy_p}{l}$	$-\frac{EA}{l}$	$\frac{EA}{l}$				
M_{zL}	$\frac{12EI_{xy}}{l^3}$	$-\frac{12EI_{xy}}{l^3}$	$\frac{6EI_{xy}}{l^2}$	$\frac{6EI_{xy}}{l^2}$	$\frac{12EI_{yy}}{l^3}$	$-\frac{12EI_{xy}}{l^3}$	$\frac{6EI_{yy}}{l^2}$	$\frac{6EI_{xy}}{l^2}$	0	0	$\frac{12EI_{yy}}{l^3} + \frac{6GK}{5l}$			
M_{zR}	$-\frac{12EI_{xy}}{l^3}$	$\frac{12EI_{xy}}{l^3}$	$-\frac{6EI_{xy}}{l^2}$	$-\frac{6EI_{xy}}{l^2}$	$-\frac{12EI_{xy}}{l^3}$	$\frac{12EI_{xx}}{l^3}$	$-\frac{6EI_{xy}}{l^2}$	$-\frac{6EI_{xx}}{l^2}$	0	0	$-\frac{12EI_{xy}}{l^3} - \frac{6GK}{5l}$	$\frac{12EI_{xy}}{l^3} + \frac{6GK}{5l}$		
T_L	$\frac{6EI_{xy}}{l^2}$	$-\frac{6EI_{xy}}{l^2}$	$\frac{2EI_{xy}}{l} - \frac{EAx_p\omega_p}{l}$	$\frac{2EI_{xy}}{l} + \frac{EAx_p\omega_p}{l}$	$\frac{6EI_{yy}}{l^2}$	$-\frac{6EI_{xy}}{l^2}$	$\frac{2EI_{xy}}{l} - \frac{EAy_p\omega_p}{l}$	$\frac{2EI_{xx}}{l} + \frac{EAy_p\omega_p}{l}$	$-\frac{EA\omega_p}{l}$	$\frac{EA\omega_p}{l}$	$\frac{6EI_{yy}}{l^2} + \frac{GK}{10}$	$-\frac{6EI_{xy}}{l^2} - \frac{GK}{10}$	$\frac{4EI_{xx}}{l} + \frac{2GK}{15} + \frac{EA\omega_p^2}{l}$	
T_R	$\frac{6EI_{xy}}{l^2}$	$-\frac{6EI_{xy}}{l^2}$	$\frac{2EI_{xy}}{l} + \frac{EAx_p\omega_p}{l}$	$\frac{4EI_{xy}}{l} - \frac{EAx_p\omega_p}{l}$	$\frac{6EI_{xy}}{l^2}$	$-\frac{6EI_{xy}}{l^2}$	$\frac{2EI_{xy}}{l} + \frac{EAy_p\omega_p}{l}$	$\frac{2EI_{xy}}{l} - \frac{EAy_p\omega_p}{l}$	$\frac{EA\omega_p}{l}$	$-\frac{EA\omega_p}{l}$	$\frac{6EI_{yy}}{l^2} + \frac{GK}{10}$	$-\frac{6EI_{xy}}{l^2} - \frac{GK}{10}$	$\frac{2EI_{xy}}{l} - \frac{GK}{30} - \frac{EA\omega_p^2}{l}$	$\frac{4EI_{xx}}{l} + \frac{2GK}{15} + \frac{EA\omega_p^2}{l}$

SESSION 2. GENERAL METHODS 2

Session Chairman

Major R. O. Meitz*

Air Force Institute of Technology
Wright-Patterson Air Force Base, Ohio

*now Lt Col

FINITE ELEMENT MODELS OF NONLINEAR OPERATOR EQUATIONS

J. T. Oden*

The University of Alabama in Huntsville

The formulation of finite-element models of nonlinear potential operators is presented in this paper. The variational theorems of Mikhlin are extended to nonlinear and nonpositive operators using the theory of Vainberg on potential operators. Completeness and convergence criteria for certain nonlinear boundary-value problems are examined. A method for constructing variational principles for nonlinear operators that do not satisfy the potentiality conditions is presented. Examples are given which involve nonlinear partial differential equations and integral equations.

---

*Professor and Chairman of Engineering Mechanics

I. INTRODUCTION

We shall be concerned with the approximate solution of certain nonlinear boundary-value problems characterized by equations of the form

$$\theta(u) = 0 \tag{1}$$

where θ is a nonlinear operators and $u = u(\underset{\sim}{x})$ is a function whose domain is a closed region $\mathcal{R}$ of k-dimensional euclidean space $\mathcal{E}^k$. Here $\underset{\sim}{x} = (x_1, x_2, \ldots, x_k)$ is a point in $\mathcal{R}$ and u must also satisfy certain boundary conditions on the boundary $\partial\mathcal{R}$. We shall emphasize finite-element approximations of (1) based on equivalent variational statements of the boundary-value problem, and, consequently, θ shall be considered to be a potential operator. Further restrictions on the class of operators and the function u are to be stated in subsequent sections.

Some Definitions and Theorems. To set the stage for the approximate solution of (1) by the finite-element method, we present here a number of definitions and theorems basic to the study of variational methods for nonlinear operators. For additional details, the work of Vainberg [1] can be consulted.

We shall be primarily concerned with approximations of elements of real Hilbert spaces. The inner product of two elements u and v is denoted $\langle u,v\rangle$ and the natural norm $\|u\|$ of an element is $\langle u,u\rangle^{\frac{1}{2}}$. The norm of the difference of two elements, $\|u - v\|$, defines a metric or distance between the elements u and v. If $\mathcal{H}$ is a real Hilbert space, sequences $u^1, u^2, \ldots, u^n, \ldots$ in $\mathcal{H}$ are denoted $\{u^n\}$. If v and u_o are fixed elements of $\mathcal{H}$, a sequence $\{u^n\}$ converges weakly to u_o if $\lim_{n\to\infty} \langle v,u^n\rangle = \langle v,u_o\rangle$; otherwise, convergence of a sequence $\{u^n\}$ to u_o implies that $\lim_{n\to\infty} \|u^n - u_o\| = 0$. $\mathcal{H}$ is complete if every Cauchy sequence (i.e., every sequence $\{u^n\}$ such that $\lim_{n,n\to\infty} \|u^n - u^n\| = 0$) converges in $\mathcal{H}$ (i.e. has a limit u_o that belongs to $\mathcal{H}$).

The construction of variational principles related to (1) is based on the concepts of differentiation in normed spaces and the idea of potential operators. Let θ denote a mapping of some subset $\mathcal{A} \subset \mathcal{H}$ into a normed space $\mathcal{Y}$ and let R denote the real numbers. Then

• θ is Gateaux differentiable at the point $u \in \text{int}(\mathcal{A})$ if there exists an operator $D_\theta(u,h)$ such that for any $h \in \mathcal{H}$

$$\lim_{\alpha\to 0} \frac{1}{\alpha} \|\theta(u + \alpha h) - \theta(u) - D_\theta(u,h)\| = 0 \tag{2}$$

Here $\alpha \in R$ and $D_\theta(u,h)$ is called the Gateaux (or weak) differential of θ at u.

· Let Θ have a Gateaux differential $D_\Theta(u,h)$ at u linear in h; then the linear operator $\dot{\Theta}(u)(\ )$, where $\dot{\Theta}(u)(h) = D_\Theta(u,h)$, is called the Gateaux derivative of Θ at u.

· Θ is Frechet differentiable at $u \in \text{int}(\mathfrak{D})$ if there exists a linear operator $\delta\Theta(u,h)$ such that for any $h \in \mathcal{H}$,

$$\lim_{\|h\| \to 0} \|\Theta(u + h) - \Theta(u) - \delta\Theta(u,h)\| = 0 \qquad (3)$$

$\delta\Theta(u,h)$ is called the Frechet differential of Θ and the collection of terms $\Theta(u + h) - \Theta(u) - \delta\Theta(u,h) = \omega(u,h)$ is called the remainder; likewise, the linear operator $\Theta'(u)$, where $\Theta'(u)(h) = \delta\Theta(u,h)$, is called the Frechet derivative of Θ at u.

· Theorem 1. (Vainberg [1]). If the Gateaux derivative exists and is continuous in some neighborhood of u, then $D_\Theta(u,h)$ exists and equals $\delta\Theta(u,h)$.

We remark that second-and higher-order Gateaux and Frechet derivatives can be constructed by applying (2) or (3) to the first derivatives $\Theta(u)$ or $\Theta'(u)$. For example, if

$$\lim_{\|k\| \to 0} \|\delta\Theta(u + k,h) - \delta\Theta(u,h) - \delta^2\Theta(u,h,k)\| = 0$$

then the bilinear operator $\Theta''(u)$, where $\Theta''(u)(h,k) = \delta^2\Theta(u,h,k)$, is the second Frechet derivative at u. If $\delta^2\Theta(u,h,k) = \delta^2\Theta(u,k,h)$, then Θ is symmetric at the point u. To continue:

· Let $K: \mathcal{H} \to R$ denote a functional which has a linear Gateaux differential on some set $\mathfrak{D} \subset \mathcal{H}$. Then the operator Θ defined by the formula

$$\langle \Theta(u),h \rangle = \lim_{\alpha \to 0} \frac{1}{\alpha} [K(u + \alpha h) - K(u)] \qquad (4)$$

for any $h \in \mathcal{H}$ is called the (weak) gradient of K at u and is denoted $\Theta = \text{grad } K$. (If K is Frechet differentiable, grad K is called the strong gradient at u).

· An operator $\Theta: \mathcal{H} \to \mathcal{V}$ is potential on a set $\mathfrak{D} \subset \mathcal{H}$ if there exists a functional K(u) such that for every $u \in \mathfrak{D}$, $\Theta(u) = \text{grad } K(u)$.

· Theorem 2 (Vainberg [1]). Let Θ be a continuous operator with a linear Gateaux differential at every point u in the set $\mathfrak{D} = \{u: \|u - u_0\| < r\}$; then a necessary and sufficient condition that Θ be potential on $\mathfrak{D}$ is that for every $h,k \in \mathfrak{D}$,

$$\langle D_\Theta(u,h),k \rangle = \langle D_\Theta(u,k),h \rangle \qquad (5)$$

This theorem provides a test for the potentiality of any given operator $\mathcal{O}$. If the operator in (1) satisfies (5), then we can proceed directly to the construction of a variational principle corresponding to (1) by applying the following fundamental theorem.

Theorem 3 (Vainberg [1]). Let $\mathcal{O}$ be a potential operator satisfying (5). Then there exists a unique functional K(u), whose value at u_o is K_o, such that grad K = $\mathcal{O}$. Furthermore, K(u) is given by

$$K(u) = \int_0^1 \langle \mathcal{O}(u_o + su - su_o),\ u - u_o \rangle ds + K_o \qquad (6)$$

Here s is a real parameter.

A point $u_o \in \mathcal{S}$ is a critical point of a functional K(u) defined on $\mathcal{S}$ if grad $K(u_o) = 0$, 0 being the zero element in $\mathcal{S} \subset \mathcal{H}$. Moreover, if $K(u) \geq K(u_o)$ for all $u \in \mathcal{S}_o \subset \mathcal{S}$, then u_o is called a global minimizer of K on $\mathcal{S}_o$ (or, if $K(u) \geq K(u_o)$ for all u in $\mathcal{S} \cap \mathcal{S}$, $\mathcal{S}$ being an open neighborhood of u_o, then u_o is a local minimizer of K). As expected, if u_o is a local minimizer of a Gateaux-differentiable functional, then grad $K(u_o) = 0$; that is, u_o is a critical point of K(u). If the second Gateaux differential $D^2K(u,h,h)$ exists and is positive definite (i.e. $D^2K(u,h,h) > 0$ for all $h \in \mathcal{S}_o$) at a critical point u_o of K(u), then u_o is a proper local minimizer of K(u)(i.e. $K(u) > K(u_o)$ for all $u \in \mathcal{S}_o$). If $D^2K(u_o,h,h)$ is positive-semidefinite at u_o, u_o is a local minimizer of K(u). Finally, if u_o is a critical point of K(u) and $D^2K(u_o,h,h)$ is negative definite or negative semidefinite, then u_o is a proper maximizer or a local maximizer of K(u), respectively.

II. GALERKIN'S METHOD; LINEAR OPERATORS; NONPOTENTIAL OPERATORS

Galerkin's Method. It is interesting to note that the idea of potential operators can be used to motivate the approximate solution of (1) by Galerkin's method. Let $\mathcal{O}$ be a potential operator defined on a set $\mathcal{S}$ of a Hilbert space $\mathcal{H}$ and let $\bar{\mathcal{H}}$ denote a G-dimensional subspace of $\mathcal{H}$ spanned by a set of G linearly independent functions $\varphi_1(\underset{\sim}{x}), \varphi_2(x), \ldots, \varphi_G(\underset{\sim}{x}) \in \mathcal{H}$. We wish to determine coefficients a^α so that the function

$$\bar{u} = a^\alpha \varphi_\alpha(\underset{\sim}{x}) \qquad (7)$$

represents an approximate solution to (1) in $\bar{\mathcal{H}}$. Since $\mathcal{O}$ is a potential operator, there exists a functional K(u) such that grad K(u) = $\mathcal{O}$(u): indeed, if u* is the solution of (1),

$$\text{grad } K(u^*) = \mathcal{O}(u^*) = 0 \qquad (8)$$

that is, the solution is a critical point of K(u). In view of (4), (8) implies that

$$\langle \theta(u^*),\ h\rangle = 0 \tag{9}$$

In other words, $\theta(u^*) = 0$ by virtue of the fact that it is orthogonal to every h in a set $\mathfrak{D}_o$ assumed to be dense in $\mathcal{H}$.

Galerkin's method amounts to forcing $\theta(\bar{u})$ to satisfy (9) in $\bar{\mathcal{H}}$. If $\bar{h}$ is an arbitrary element of $\bar{\mathcal{H}}$, we set

$$\langle \theta(\bar{u}),\bar{h}\rangle = \langle \theta(\bar{u}),h^{\beta}\varphi_{\beta}\rangle = h^{\beta}\langle \theta(\bar{u}),\varphi_{\beta}\rangle = 0 \tag{10}$$

and since (10) must hold for arbitrary coefficients h^{β}, we have

$$\langle \theta(a^{\alpha}\varphi_{\alpha}),\ \varphi_{\beta}\rangle = 0 \tag{11}$$

Equations(11) represent G independent (generally nonlinear) equations in a^{α} which, when solved, determine the approximate solution (7). We also observe that (11) represents the requirement that the residual

$$r_o \equiv \theta(a^{\alpha}\varphi_{\alpha}) \neq 0$$

be orthogonal to $\bar{\mathcal{H}}$. In fact, (11) must be satisfied if $\bar{u}$ of (7) is to be the "best" approximation of u_o in $\mathcal{H}$. To prove this assertion, assume the contrary; i.e. $\langle -r_o,h\rangle = \alpha \neq 0$, and, without loss in generality, set $\|h\| = 1$. Then, if $r = r_o + \alpha h \in \bar{\mathcal{H}}$, consider

$$\|\theta(u^*) - r\|^2 = \langle \theta(u^*) - r_o - \alpha h,\ \theta(u^*) - r_o - \alpha h\rangle$$

$$= \|\theta(u^*) - r_o\|^2 - \alpha\langle h,\theta(u^*) - r_o\rangle - \alpha\langle \theta(u^*) - r_o,h\rangle + \alpha^2$$

Since $\theta(u^*) = 0$, $\langle h,\theta(u^*) - r_o\rangle = \langle h,- r_o\rangle = \alpha$ and

$$\|\theta(u^*) - r\|^2 = \|\theta(u^*) - r_o\|^2 - \alpha^2 < \|\theta(u^*) - r_o\|^2$$

Consequently, the measure $\|\theta(u^*) - \theta(\bar{u})\|$ is a minimum when $\alpha = 0$. Then $r = r_o$ and $\langle r_o,h\rangle = 0$, as indicated in (11)

Linear Operators. Mikhlin [2] presented a detailed investigation of variational principles associated with linear positive-definite operators encountered in boundary-value problems characterized by equations of the form

$$\mathfrak{L}(u) = f \tag{12}$$

$\mathcal{L}$ being the linear operator. If a solution to (12) exists, then it can be shown to minimize the functional

$$J(u) = \langle \mathcal{L}u,u\rangle - 2\langle f,u\rangle \tag{13}$$

It is interesting to note that (13) can be obtained as a special case of the theory outlined in the previous section. Consider the nonlinear operator

$$\mathcal{O}(u) = \mathcal{L}(u) - f \tag{14}$$

Setting $u_o = 0$, $K_o = 0$ and introducing (14) into (6), we get

$$K(u) = \int_0^1 \langle \mathcal{L}(su) - f,u\rangle ds = \int_0^1 [s\langle \mathcal{L}u,u\rangle - \langle f,u\rangle]ds = \frac{1}{2}\langle \mathcal{L}u,u\rangle - \langle f,u\rangle$$
$$= \frac{1}{2} J(u) \tag{15}$$

Thus (12) is the Euler equation of $2K(u)$.

Nonpotential Operators. The restriction that $\mathcal{O}(u)$ of (1) must be potential in order to develop an associated variational statement of the problem appears to be very severe, for it is not difficult to cite important examples in which (5) is not satisfied. Yet, in these examples, application of (11) may lead to meaningful approximate solutions. The question arises as to whether or not it is possible to subject a given operator equation, involving a nonpotential operator, to some type of transformation which will render it potential and still preserve the properties of the solution of the original equation. We present here an affirmative answer to this question based on the procedure often employed in the solution of systems of nonlinear algebraic equations of constructing a functional which takes on local minimum values at the points satisfying the given system of equations [3].

Let $\mathcal{O}$ be a nonlinear operator from $\mathcal{H}$ into $\mathcal{Y}$, not necessarily potential, and consider the problem of solving the equation $\mathcal{O}(u) = \theta$, where θ is the zero element of $\mathcal{Y}$ (we use θ instead of 0 temporarily to distinguish elements of $\mathcal{Y}$ from the real number 0). Let $g:\mathcal{Y} \to R$ be a functional with the following properties:

(i) $g(\theta) < g(v)$ for every $v \in \mathcal{Y}$, $v \neq \theta$

(ii) $g(\theta) = 0$

Set

$$j(u) = g(\mathcal{O}(u)) \tag{16}$$

Then the solution u* of $\mathcal{O}(u) = \theta$ is a (unique) global minimizer of j(u). Furthermore, if

(iii) g(u) is Frechet differentiable and has a unique critical point at $\theta \in \mathcal{V}$

then

$$\text{grad } j(u^*) = 0 \quad \mathcal{O}(u^*) = \theta \tag{17}$$

For example, suppose that $\|v\|$ denotes the norm in $\mathcal{V}$. If we set

$$g(v) = \|v\|^2 = \langle v, v \rangle \tag{18}$$

then

$$j(u + \alpha h) - j(u) = \langle \mathcal{O}(u + \alpha h), \mathcal{O}(u + \alpha h) \rangle - \langle \mathcal{O}(u), \mathcal{O}(u) \rangle$$

$$= \|\delta\mathcal{O}(u,\alpha h)\|^2 + 2\langle \delta\mathcal{O}(u,\alpha h), \mathcal{O}(u) \rangle$$

$$+ 2\langle \omega(u,\alpha h), \mathcal{O}(u) \rangle + \langle \mathcal{O}(u + \alpha h), \omega(u,\alpha h) \rangle$$

and

$$\lim_{\alpha \to 0} \frac{1}{\alpha} [j(u + \alpha h) - j(u)] = 2\langle \mathcal{O}'(u)(h), \mathcal{O}(u) \rangle = g'(\mathcal{O}(u)).\mathcal{O}'(u)(h) \tag{19}$$

Hence, if grad j(u*) = 0, g'(θ) = 0, which implies that v = θ is the critical point of g. Likewise, if $\mathcal{O}(u^*) = \theta$, then grad j(u*) = 0.

Obviously, the choice (18) leads to least-squares approximations. Still other choices are possible. For example, if $\mathcal{L}$ is a positive-definite, symmetric, linear operator on $\mathcal{V}$, the procedure leading to (19) yields

$$\langle \text{grad } j(u), h \rangle = 2\langle \mathcal{L}(\mathcal{O}(u)), \mathcal{O}'(u)(h) \rangle \tag{20}$$

Since h can assumedly be taken from outside the null space of the linear operator $\mathcal{O}'(u)$, vanishing of (20) implies that $\mathcal{L}(\mathcal{O}(u^*)) = 0$. But since $\mathcal{L}$ is positive definite, the unique solution of $\mathcal{L}(\mathcal{O}(u^*)) = 0$ is $\mathcal{O}(u^*) = \theta$, which is the solution of the original problem.

III. FINITE ELEMENT APPROXIMATIONS

In this section we shall briefly review certain properties of finite element approximations viewed as a special method of interpolation. Specifically, if U(x) is a given function ($x \in \mathcal{R} \subset \mathcal{E}^k$, $U(x) \in C^p(\mathcal{R})$), what are the properties of a finite-element approximation $\tilde{U}(x)$ of U(x) which coincides with U(x) and certain of its partial derivatives at a finite number G of nodal points $\tilde{x}^\Delta \in \mathcal{R}$, $\Delta = 1,2,\ldots,G$? We first record a number of results and definitions given in [3]:

· To construct a finite element model of $U(x)$, we partition $\mathcal{R}$ into E subregions $r_1, r_2, \ldots, r_E$ (finite elements) connected continuously together along interelement boundaries. While x^Δ denote global nodal points in the "connected model" $\bar{\mathcal{R}}$ of $\mathcal{R}$, locally we denote nodal points in element r_e by x_e^N, $N = 1, 2, \ldots, N_e$; N_e being the number of nodal points of r_e. The decomposition of $\bar{\mathcal{R}}$ into elements is then implied by

$$x_e^N = {}^{(e)}\Omega_\Delta^N x^\Delta \tag{21}$$

where the Boolean matrix ${}^{(e)}\Omega_\Delta^N = 1$ if node N of r_e is coincident with node Δ in the connected model and ${}^{(e)}\Omega_\Delta^N = 0$ if otherwise. Here and henceforth, repeated nodal indices $\Delta(N)$ are summed from 1 to $G(N_e)$. Also, for simplicity, we assume that the local and global coordinate systems x_e, x coincide and we affix an element label (e) on various local quantities only when their local character is of special importance or is not clear in the context.

· Let $u_e(x_e)$ denote the restriction of $U(x)$ to r_e and let u_e^N, $u_{e,i}^N$, $\ldots, u_{e,i_1\ldots i_r}^N$ denote the values of $u_e(x)$ and its partial derivatives up to order r at node N or r_e. The local finite-element representation of $u_e(x)$ is of the form

$$u_e(x) \approx \bar{u}_e(x) = u_e^i \psi_N^{0}(x) + u_{e,i}^N \psi_N^{i}(x) + \ldots + u_{e,i_1\ldots i_r}^N \psi_N^{i_1\cdots i_r}(x) \tag{22}$$

where N is summed from 1 to N_e and the interpolation functions $\psi_N^0(x), \ldots, \psi_N^{i_1\cdots i_r}(x)$ $(i_1 \le i_2 \le \ldots \le i_r = 1, 2, \ldots k)$ have the following properties:

$$\text{(i)} \quad \frac{\partial^r \psi_N^{i_1 i_2 \cdots i_s}(x_e^M)}{\partial x_{j_1} \partial x_{j_2} \cdots \partial x_{j_r}} = \begin{cases} 0 & r \neq s \\ \delta_N^M \delta_{j_1}^{i_1} \cdots \delta_{j_s}^{i_s} & r = s \end{cases} \tag{23}$$

(ii) If ℓ_e is a characteristic length of r_e, there exist bounded dimensionless functions $\varphi_N^0(\xi)$, where $\xi_1 = x_1/\ell_e$, such that

$$\psi_N^0(x) = \varphi_N^0(\xi), \quad \psi_N^i(x) = \ell_e \varphi_e^i(\xi), \ldots, \psi_N^{i_1\cdots i_r}(x) = \ell_e^r \varphi_N^{i_1\cdots i_r}(\xi) \tag{24}$$

and for every $x_e = x \in r_e$,

$$|\psi_N^0(x)| \le K_0, |\psi_N^i(x)| \le \ell_e K_1, \ldots, |\psi_N^{i_1\cdots i_r}(x)| \le \ell_e^r K_r \tag{25}$$

where $K_0, K_1, \ldots, K_r$ are constants $< \infty$ and independent of ℓ_e

$$\text{(iii)} \quad \frac{\partial^s \psi_N^{i_1\cdots i_s}(x)}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_s}} \le \ell_e^{n-s} K_n \tag{26}$$

$m \leq r$, $s = 0,1,\ldots,r$.

A finite element approximation of the form (22) is said to belong to a family $\mathfrak{F}^r$ of finite element approximations if at least one partial derivative of order r is specified at a node. All elements of the family are generated by allowing the coefficients $u_e^N,\ldots,u_{e,i_1\cdots i_r}^N$ to take on all real values. For example, $u_e = u_e^N \psi_N^0(\underset{\sim}{x}) + u_{e,11}^N \psi_N^{11}(\underset{\sim}{x})$ and $u_e = u_{e,12}^N \psi_N^{12}(\underset{\sim}{x})$ define two different $\mathfrak{F}^2$ families of finite-element approximations.

· The global approximation of $U(\underset{\sim}{x})$ is of the form

$$U(\underset{\sim}{x}) \approx \bar{U}(\underset{\sim}{x}) = U^{\Delta}\Phi_{\Delta}^{0}(\underset{\sim}{x}) + U_{,1}^{\Delta}\Phi_{\Delta}^{1}(\underset{\sim}{x}) + \ldots + U_{,i_1\ldots i_r}^{\Delta}\Phi_{\Delta}^{i_1\cdots i_r}(\underset{\sim}{x}) \tag{27}$$

where U^{Δ}, $U_{,1}^{\Delta},\ldots,U_{,i_1\cdots i_r}^{\Delta}$ are the values of U and its partial derivatives at global node $\underset{\sim}{x}^{\Delta} \in \bar{\mathfrak{R}}$ and $\Phi_{\Delta}^{0}(\underset{\sim}{x}),\ldots,\Phi_{\Delta}^{i_1\cdots i_r}(\underset{\sim}{x})$ are the global approximation functions; that is, $\psi_N^{i_1\cdots i_s}(\underset{\sim}{x})$ is the restriction of $\Phi_{\Delta}^{i_1\cdots i_s}(\underset{\sim}{x})$ to $\mathfrak{r}_e$.

Refinements, Diameters, and r-Equivalence. We can generate sequences of approximations of the form (27) by constructing refinements of $\bar{\mathfrak{R}}$; that is, we refine $\bar{\mathfrak{R}}$ by identifying a new finite-element model $\bar{\mathfrak{R}}'$ obtained by increasing the number of elements and nodes in $\bar{\mathfrak{R}}$. We consider only uniform, regular refinements; i.e. refinements in which the same types of elements and interpolation functions defined on $\bar{\mathfrak{R}}$ are used for $\bar{\mathfrak{R}}'$ and every node and interelement boundary in $\bar{\mathfrak{R}}$ is a node and interelement boundary in $\bar{\mathfrak{R}}'$. The diameter δ_e of an element $\mathfrak{r}_e$ is the maximum distance between two points $\underset{\sim}{x}_e, \underset{\sim}{y}_e \in \mathfrak{r}_e$ ($\delta_e = \max \|\underset{\sim}{x}_e - \underset{\sim}{y}_e\|$), and the mesh δ of a model $\bar{\mathfrak{R}}$ is the maximum diameter of all elements in $\bar{\mathfrak{R}}$ ($\delta = \max\{\delta_1,\delta_2,\ldots,\delta_E\}$). Obviously, for a sequence of refinements $\{\bar{\mathfrak{R}}_\Gamma\}$ we require that the meshes $\delta_\Gamma \to 0$ as $\Gamma \to \infty$.

· Let $f(\underset{\sim}{x}_e)$ and $g(\underset{\sim}{x}_e)$ be the restrictions of two functions $f(\underset{\sim}{x})$ and $g(\underset{\sim}{x})$ to an element $\mathfrak{r}_e$ of diameter δ_e. Then $f(\underset{\sim}{x})$ and $g(\underset{\sim}{x})$ are r-equivalent on $\mathfrak{r}_e$ with respect to the norm $\|\cdot\|$ if

$$\|f(\underset{\sim}{x}_e) - g(\underset{\sim}{x}_e)\| \leq K\delta_e^r \tag{28}$$

where K is a constant independent of δ_e. Unless noted otherwise, by equivalence we shall henceforth mean uniform equivalence of functions, i.e. $f(\underset{\sim}{x})$ and $g(\underset{\sim}{x})$ are uniformly r-equivalent iff $|f(\underset{\sim}{x}_e) - g(\underset{\sim}{x}_e)| \leq K\delta_e^r$, $\underset{\sim}{x}_e \in \mathfrak{r}_e$.

· Let $U(\underset{\sim}{x})$ be a function defined over a finite element model $\bar{\mathfrak{R}}$; $U(\underset{\sim}{x})$ is r-comformable if $U(\underset{\sim}{x})$ and its partial derivatives up to and including those of order r are continuous at every interelement boundary in $\bar{\mathfrak{R}}$. Obviously, if $U(\underset{\sim}{x}) \in C^r(\bar{\mathfrak{R}})$, then $U(\underset{\sim}{x})$ is r-conformable.

Theorem 4. [4] Let $u(\underset{\sim}{x})$, $\underset{\sim}{x} \in \bar{\mathfrak{R}}$, be continuous, together with its partial derivatives of order p, and let the p + 1 st partial derivatives

be piecewise continuous and bounded. Let $\hat{u}(\underset{\sim}{x}) \in \mathfrak{F}^r$ be an r-conformable finite-element approximation of $u(\underset{\sim}{x})$ which, together with its partial derivatives up to order r, coincides with $u(\underset{\sim}{x})$ at each node $x^{\Delta} \in \bar{R}$. Moreover, let $\mathfrak{F}^r$ contain a subfamily $\mathcal{P}$ of complete polynomials of degree p and let r = p - 1. Then, at every point $\underset{\sim}{x} \in r_e \subset \bar{R}$,

$$|u(\underset{\sim}{x})_{,i_1 \cdots i_j} - \hat{u}(\underset{\sim}{x})_{,i_1 \cdots i_j}| \leq C\delta_e^{p+1-j} \tag{29}$$

$j = 0,1,2,\ldots,p$, where C is a constant independent of δ_e.

Proof: A similar theorem was discussed by Arantes e Oliveira [4]. We shall reproduce briefly the basic ideas of the proof following the plan in [3].

Since $u(\underset{\sim}{x}) \in C^p(\bar{R})$, we may write

$$u_e(\underset{\sim}{x}) = u^N + u^N_{,i}(x_i - x_i^N) + \ldots + \frac{1}{(p+1)!}\, u(\underset{\sim}{x}^*)_{,i_1 \cdots i_p}(x_{i_1} - x^N_{i_1}) \times (x_{i_{p+1}} - x^N_{i_{p+1}})$$

where $\underset{\sim}{x}, \underset{\sim}{x}^N \in r_e$ and $\underset{\sim}{x}^* \in \{\underset{\sim}{y}: \theta\underset{\sim}{x} + (1 - \theta)\underset{\sim}{x}^N = \underset{\sim}{y},\ 0 \leq \theta \leq 1\}$. If $p(\underset{\sim}{x}) \in \mathcal{P}$ is a complete polynomial of degree p, we can adjust its coefficients so that

$$|u_e - p| \leq C_0\delta^{p+1}, \ldots, \ |u_{e,i_1 \cdots i_j} - p_{,i_1 \cdots i_j}| \leq C_j\delta^{p+1-j} \tag{30}$$

Here C_j depends on the bound of the p+1 derivative of u_e and we have used the fact that $|x_i - x_i^N| \leq \delta_e$ for every i. Since $\hat{u}$ and $p \in \mathfrak{F}^r$,

$$p(\underset{\sim}{x}) - \hat{u}(x) = (p^N - \hat{u}^N)\psi_N^0(\underset{\sim}{x}) + \ldots + (p^N_{,i_1 \cdots i_r} - \hat{u}^N_{,i_1 \cdots i_r})\psi_N^{i_1 \cdots i_r}(\underset{\sim}{x}) \tag{31}$$

where, in view of the fact that, by definition, $\hat{u}$ coincides with u at its derivatives of order p-1 = r at the nodes and p can be made to coincide with the first p terms of an expansion of u_e about $\underset{\sim}{x}^N$,

$$|p^N_{,i_1 \cdots i_s} - \hat{u}^N_{,i_1 \cdots i_s}| \leq N\delta_e^{p+1-s} \tag{32}$$

Introducing (31) into (30) and making use of (22) - (25), we get

$$|p_{,i_1 \cdots i_j} - \hat{u}_{,i_1 \cdots i_j}| \leq M_j\delta_e^{p+1-j} \tag{33}$$

Thus, from (a),

$$|u_{,i_1\cdots i_j} - \hat{u}_{,i_1\cdots i_j}| = |u_{,i_1\cdots i_j} - p_{,i_1\cdots i_j} - (\hat{u}_{,i_1\cdots i_j} - p_{,i_1\cdots i_j})|$$

$$\leq |(u - p)_{,i_1\cdots i_j}| + |(\hat{u} - p)_{,i_1\cdots i_j}|$$

$$\leq (C_j + M_j)\delta_e^{p+1-j} = C\delta_e^{p+1-j}$$

which proves the theorem.

It is not necessary to limit the finite element approximations in $\mathfrak{F}^r$ to polynomials. We can generalize Theorem 4 as follows:

Theorem 5 Let $u(\underset{\sim}{x})$ be as in Theorem 4 (i.e. $u(\underset{\sim}{x}) \in C^p(\bar{R})$) and $\hat{u}(\underset{\sim}{x}) \in \tilde{\mathfrak{F}}^r$ be an r-conformable finite element approximation of $u(\underset{\sim}{x})$ at which, together with its partial derivatives of order r, coincides with $u(\underset{\sim}{x})$ at each node $\underset{\sim}{x}^\Delta \in \bar{R}$. Let $\mathfrak{F}^r$ contain a subfamily $\tilde{\mathfrak{F}}$ of functions whose j^{th} partial derivatives are p+1-j-equivalent to the corresponding derivatives of a complete polynomial of degree p; j=0,1,2,...,p. Then (29) holds.

Proof: Let $\tilde{p}(\underset{\sim}{x}) \in \tilde{\mathfrak{F}}$. Then, by hypothesis, the following hold:

$$|(p - \tilde{p})_{,i_1\cdots i_j}| \leq \tilde{M}_j \delta_e^{p+1-j} \tag{34}$$

$$\hat{u}(\underset{\sim}{x}) - \tilde{p}(\underset{\sim}{x}) = (\hat{u}^N - \tilde{p}^N)\psi_N^0(\underset{\sim}{x}) + \ldots + (\hat{u}^N_{,i_1\cdots i_r} - \tilde{p}^N_{,i_1\cdots i_r})\psi_N^{i_1\cdots i_r}(\underset{\sim}{x}) \tag{35}$$

However, since

$$|\hat{u}^N_{,i_1\cdots i_j} - \tilde{p}^N_{,i_1\cdots i_j}| \leq |u^N_{,i_1\cdots i_j} - p^N_{,i_1\cdots i_j}| + |p^N_{,i_1\cdots i_j} - \tilde{p}^N_{,i_1\cdots i_j}|$$

we have from (30) and (34),

$$|(\hat{u} - \tilde{p})_{,i_1\cdots i_j}| \leq \tilde{N}_j \delta_e^{p+1-j} \tag{36}$$

where $\tilde{N}_j$ are constants. Equations (30), (34), and (36) now lead directly to the conclusion of the theorem: note that

$$(u - \hat{u})_{,i_1\cdots i_j} = (u - p + p - \tilde{p} + \tilde{p} - \hat{u})_{,i_1\cdots i_j}$$

Thus

$$|(u - \hat{u})_{,i_1\cdots i_j}| \leq |(u - p)_{,i_1\cdots i_j}| + |(p - \tilde{p})_{,i_1\cdots i_j}| + |(\tilde{p} - \hat{u})_{,i_1\cdots i_j}|$$

$$\leq (C_j + \tilde{M}_j + \tilde{N}_j)\delta_e^{p+1-j} = \bar{C}\delta_e^{p+1-j}$$

Theorem 5 is thus proved.

We remark that, because of the assumption of r-conformity, we can go directly from (28) to the global estimate

$$|(u - \hat{u})_{,i_1 \cdots i_j}| \leq C\delta^{p+1-j} \tag{37}$$

where δ is the mesh of the model $\bar{\mathcal{R}}$. Obviously, if $j \leq p$, $\lim_{\delta \to 0} |(u - \hat{u})_{,i_1 \cdots i_j}| = 0$.

IV. CONVERGENCE

We shall now demonstrate that it is relatively simple to prove convergence of finite-element approximations of (1) for the case in which $\mathcal{P}$ is potential and the solution u^* is a local minimizer of the functional $K(u)$ corresponding to $\mathcal{P}$ (i.e. $\mathcal{P}(u) = \text{grad } K(u)$). Indeed, the proof is essentially the same as that for linear positive definite operators because, under the stated conditions, the quadratic terms in h in $K(u + h) - K(u)$ dominate those of higher order in the neighborhood of the critical point u^*.

Let $\bar{U}$ denote the solution of the finite-element approximation of the nonlinear equation (1). That is, among all functions of the form (27), that which minimizes the functional (15) in the subspace $\bar{\mathcal{F}}$ described by the approximating functions $\phi_\Delta(x),\ldots$ in (27), is the function $\bar{U}$. Further, let $\hat{U}$ denote another member of $\bar{\mathcal{F}}$ which, together with its derivatives of order r, coincides with the exact solution u^* at the nodal points of a finite-element model of the domain $\mathcal{R}$ of u^*. Then

$$|K(\bar{U}) - K(u^*)| \leq |K(\hat{U}) - K(u^*)| \tag{38}$$

Thus, we need only show that there exists a constant C and an interger $m > 0$ such that

$$|K(\hat{U}) - K(u^*)| \leq C\delta^m \tag{39}$$

where δ is the mesh of $\bar{\mathcal{R}}$. Then, in the limit as $\delta \to 0$, $|K(\hat{U}) - K(u^*)| \to 0$, which, according to (38), implies that also $|K(\bar{U}) - K(u^*)| \to 0$.

Let

$$h = u^* - \hat{U} \tag{40}$$

and consider the functional

$$K(u) = \int_{\mathcal{R}} F(\underline{x}, u)\, d\mathcal{R} \tag{41}$$

where, for brevity, $F(\underset{\sim}{x},u)$ denotes a function

$$F(x,u) = F(x,u;u_{,i};u_{,i_1};\ldots,u_{,i_1\ldots i_p}) \tag{42}$$

We assume that $F(\underset{\sim}{x},u)$ is continuous on $\mathcal{R}$ and $\mathcal{S}$ and that all of its partial derivatives with respect to the arguments indicated in (42) are continuous and bounded for every $\underset{\sim}{x} \in \mathcal{R}$.

Subject to all the conditions stated thusfar,

$$\Delta K(u,h) \equiv K(u^* + h) - K(u^*) = \int_{\mathcal{R}} \mathcal{E}(u^*)h\,d\mathcal{R} + \int_{\mathcal{R}} F^*(\underset{\sim}{x},u,h)\,d\mathcal{R} \tag{43}$$

where, for every h,

$$\mathcal{E}(u^*) = 0 = \frac{\partial F}{\partial u} - \frac{\partial}{\partial x_i}\left(\frac{\partial F}{\partial u_{,i}}\right) + \ldots + (-1)^p \frac{1}{p!} \frac{\partial^k}{\partial x_{i_1}\ldots\partial x_{i_p}}\left(\frac{\partial F}{\partial u_{,i_1}\ldots\partial u_{,i_p}}\right) \tag{44a}$$

and

$$F^*(\underset{\sim}{x},u,h) = \sum_{n=2}^{\infty} \frac{1}{n!}\left[\frac{\partial^n F}{\partial u^n} h^n + \frac{\partial^n F}{\partial u_{,i_1}\ldots\partial u_{,i_n}} h_{,i_1}\ldots h_{,i_n} + \ldots + \frac{\partial^n F}{\partial u_{,i_1\ldots i_{p_1}}\ldots\partial u_{,i_n\ldots i_{p_n}}} h_{,i_1\ldots i_{p_1}}\ldots h_{,i_n\ldots i_{p_n}}\right] \tag{44b}$$

Equation (44a) is merely a restatement of the assumed fact that $\mathcal{E}(u) = 0$ is the Euler equation of $K(u)$.

It follows from (43) and (44) and the assumed properties of $F(\underset{\sim}{x},u)$ that

$$|\Delta K(u,h)| \le \mathcal{V}\bar{C} \sum_{n=2}^{\infty} \frac{1}{n!}\left[\max\left|\frac{\partial^n F}{\partial u^n}\right| |h|^n + \max\left|\frac{\partial^n F}{\partial u_{,i_1}\ldots\partial u_{,i_n}}\right| |h_{,i_1\ldots i_n}| + \ldots\right]$$

wherein $\mathcal{V}$ is the volume of $\mathcal{R}$ and $\bar{C}$ is a constant. Therefore, there exist positive constants such that

$$|\Delta K(u,h)| \le \tilde{C} \sum_{n} \frac{1}{n!}\left[k_{0n} \max_{\underset{\sim}{x}} |h|^n + K_{1n} \max_{\underset{\sim}{x}} |h_{,i}|^n + \ldots + K_{p_n} \max_{\underset{\sim}{x}} |h_{,i_1\ldots i_p}|^n\right] \tag{45}$$

We now assume that the conditions of Theorem 5 (or Theorem 6) hold so that (37) applies. Then, from (40) and (45), it follows that

$$\left|K(\hat{U}) - K(u^*)\right| \le \tilde{C}^* \sum_{n=2}^{\infty} \frac{1}{n!} \left[K_{o\,n} C_o^n (\delta^{p+1})^n + \ldots + K_{p\,n} C_p^n \delta^n\right]$$

or, as $\delta \to 0$,

$$\left|K(\hat{U}) - K(u^*)\right| \le C\delta^2 \tag{46}$$

Therefore, m of (39) equals 2 and we have proven convergence of $K(\bar{U})$ to $K(u^*)$.

V. APPLICATIONS

A Nonlinear Partial Differential Equation. To demonstrate representative applications of the theory presented thusfar, we consider the nonlinear boundary-value problem

$$-\mathcal{\theta}(u) = 2(u_x^2 + u_y^2)\nabla^2 u + 4(u_x^2 u_{xx} + 2u_x u_y u_{xy} + u_y^2 u_{yy}) - u + f = 0 \tag{47}$$

where $u = u(x,y)$, $u_x = \partial u/\partial x$, $u_{xy} = \partial^2 u/\partial x \partial y$, etc. and $(x,y) \in \mathcal{R}$; on the boundary $\partial\mathcal{R}$ we assume, for simplicity, that $u = u_x = u_y = 0$. Here $f = f(x,y)$ is a prescribed twice continuously differentiable function.

It is natural to first question whether or not the operator $\mathcal{\theta}(u)$ is potential. Introducing (47) into (3) (or (2)) and performing the indicated manipulations, we obtain for the first Frechet differential of $\mathcal{\theta}$

$$\delta\mathcal{\theta}(u,h) = hP^o(u) + h_x P^1(u) + h_y P^2(u) + h_{xx} P^3(u) + h_{xy} P^4(u) + h_{yy} P^5(u) \tag{48}$$

where, for simplicity in notation, we have used

$$P^o(u) = 1 \quad ; \quad P^1(u) = -4(u_x \nabla^2 u + 2u_x u_{xx} + 2u_y u_{xy})$$

$$P^2(u) = -4(u_y \nabla^2 u + 2u_x u_{xy} + 2u_y u_{yy}); \quad P^3(u) = -2(3u_x^2 + u_y^2)$$

$$P^4(u) = -8u_x u_y \qquad P^5(u) = -2(u_x^2 + 3u_y^2) \tag{49}$$

Thus, $\mathcal{\theta}$ is potential if $\delta\mathcal{\theta}(u,h)$ satisfies (5). If k is an arbitrary function from the class of functions satisfying the boundary conditions, we find

$$\langle \delta\mathcal{P}(u,h),k\rangle - \langle \delta\mathcal{P}(u,k),h\rangle = \int_{\mathcal{R}} h(kQ_o + k_x Q_1 + k_y Q_2 + k_{xx} Q_3 + k_{xy} Q_4 + k_{yy} Q_5)dxdy \tag{50}$$

where

$$\begin{aligned} Q_o &= P^3_{xx} - P^1_x - P^2_y + P^4_{xy} + P^5_{yy} \\ Q_1 &= P^4_y - 2P^1 + 2P^3_x \\ Q_2 &= 2P^5_y - 2P^2 + P^4_x \end{aligned} \tag{51}$$

and, in our particular example, $Q_3 = Q_4 = Q_5 = 0$. The vanishing of the functions Q_o, Q_1, and Q_2 are clearly sufficient for $\mathcal{P}$ to be potential. Upon introducing (49) into (51) we indeed find that all Q_i in (50) vanish. Hence (5) holds and $\mathcal{P}$ is therefore a potential operator.

It follows that there exists a functional K(u) whose gradient equals $\mathcal{P}$. Introducing (47) into (6), with $u_o = 0$, we find after some lengthy algebra that

$$K(u) = \frac{1}{2}\int_{\mathcal{R}} [(u_x^2 + u_y^2)^2 + u^2 - 2uf]dxdy \tag{52}$$

It is easily verified that (47) is the Euler-equation for K(u).

Observing that

$$\begin{aligned} \Delta K(u,h) &= K(u+h) - K(u) \\ &= \int_{\mathcal{R}} h\mathcal{P}(u)dxdy - \frac{1}{2}\int_{\mathcal{R}} [P^3(u)h_x^2 + \frac{3}{4}P^4(u)h_x h_y + P^5(u)h_y^2]dxdy \\ &\quad + 2\int_{\mathcal{R}} (u_x h_x^3 + u_y h_y h_x^2 + u_x h_x h_y^2 + u_y h_y^3)dxdy + \frac{1}{2}\int_{\mathcal{R}} (h_x^2 + h_y^2)^2 dxdy \end{aligned}$$

Hence, for h, h_x, and h_y sufficiently small, the solution u* of (47) is a global minimizer of K(u) if, for $u^* \neq 0$,

$$-P^3(u^*) > 0 \qquad P^3(u^*)P^5(u^*) - \frac{4}{3}P^4{}^2(u^*) \geq 0 \tag{53}$$

In view of (49), these conditions are satisfied for real-valued functions u. Thus, we may impose the convergence criteria discussed earlier.

To construct a finite-element approximation of the solution of (47), we first observe that p of (37) is, in this case, unity. Therefore, we need only use piecewise linear approximations of u. For example, locally we have

$$u_e \approx (a_N + b_{N\alpha}x^\alpha)u^N \tag{54}$$

where, for triangular elements, N = 1,2,3 and α = 1,2; $x^1 = x$, $x^2 = y$. Hence, $u_x = b_{N1}$, $u_y = b_{N2}$. Locally, the functional K(u) of (52) becomes

$$K_e(u) = \frac{1}{2}\,(u^N u^M u^R u^S A_{NMRS} + B_{MN}u^M u^N - 2u^N f_N) \tag{55}$$

where

$$A_{NMRS} = a_e b_{N\alpha} b_{M\alpha} b_{R\beta} b_{S\beta} \qquad B_{MN} = \int_{r_e} \psi_M \psi_N dxdy \qquad f_N = \int_{r_e} f\psi_N dxdy \tag{56}$$

Here a_e is the area of the element, $\psi_N(x,y) = a_N + b_{N\alpha}x^\alpha$, and α, β = 1, 2. Thus, minimizing $K_e(u)$ among all functions of the form (54), we obtain the system of cubic equations,

$$u^N u^M u^S (A_{NMKS} + A_{KNMS}) + u^M B_{MK} = f_K \tag{57}$$

Global forms are obtained by connected elements together in the usual fashion.

A Nonlinear Integral Equation. As a final example, we consider the nonlinear integral equation

$$-\theta(u(\underline{x},t)) = \int_0^t G(t-\tau)\left(\frac{\partial u(\underline{x},\tau)}{\partial x}\right)^2 \frac{\partial^2 u(\underline{x},\tau)}{\partial x^2}\,d\tau - f(\underline{x},t) = 0 \tag{58}$$

where $\underline{x} = (x,y) \in \mathcal{R} \subset \mathcal{E}^2$ and, for simplicity, we assume that $u_x(\underline{x},t) = u_y(\underline{x},t) = u(\underline{x},t) = 0$ on $\partial\mathcal{R}$ and $u(\underline{x}0) = 0$. We also loosen the conditions on $\langle u,v\rangle$, requiring only symmetry and linearity in u and v and that $\langle u,v\rangle = 0 \Rightarrow v$ or $u = 0$. In the present case, we define [5]

$$\langle u,v\rangle = \int_{\mathcal{R}} u*v\,d\mathcal{R} \tag{59}$$

where u*v denotes the convolution

$$u*v = \int_0^t u(t-\tau)v(\tau)d\tau \tag{60}$$

Alternately, to account for possible discontinuities in u, we could use the Stieltjes' integral u*dv; but we assume sufficient continuity in t in the present example so that (59) is adequate.

It follows that

$$\langle \theta(su),u\rangle = \int_{\mathcal{R}} [s^3 G*u_x^2 u_{xx}*u - f*u]dxdy \tag{61}$$

Thus, introducing (61) into (6), integrating by parts, and making use of the homogeneous boundary conditions and the commutativity and associativity of the convolution operator, we obtain the functional

$$K(u) = \frac{1}{12} \int_{\mathcal{R}} [G*u_x^3*u_x - 12f*u]d\mathcal{R} \tag{62}$$

Hence, we have derived a variational principle associated with the nonlinear integral equation (58); that is, K(u) of (62) assumes a stationary value when u satisfies (58). It is not difficult to verify that (58) is, indeed, the Euler equation of (62), for

$$\lim_{\alpha\to 0} \frac{1}{\alpha} [K(u + \alpha h) - K(u)] = \lim_{\alpha\to 0} \frac{\partial}{\partial\alpha} K(u + \alpha h) = \int_{\mathcal{R}} (G*u_x^2 u_{xx} - f)*h d\mathcal{R} = \langle \theta(u),h\rangle$$

That is, $\theta(u)$ = grad K(u).

The construction of finite-element approximations of (58) is now straightforward. Locally, if we use the linear approximation in (54) with $u^N = u^N(t)$, we find that

$$\lim_{\alpha\to 0} \frac{\partial}{\partial\alpha} K(\psi_N u^N + \alpha\psi_N h^N(t)) = h^N*[\frac{1}{4} C_{NRSM} u^R u^S (G*u^M) + \frac{1}{2} C_{SRNM}(G*u^S u^R u^M) - f_N] = 0 \tag{63}$$

where

$$C_{NRSM} = ab_{N1}b_{R1}b_{S1}b_{M1} \qquad f_N = \int_{r_e} f(\underline{x},t)\psi_N(x)dxdy \tag{64}$$

Since $h^N(t)$ and their histories are arbitrary, the term in brackets in (63) must vanish, and we obtain as a discrete model of (58) a system of nonlinear integral equations in $u^N(t)$.

Acknowledgement: The support of this work by the U. S. Air Force Office of Scientific Research under Contract F44620-69-C0124 is gratefully acknowledged.

VI. REFERENCES

1. M. M. Vainberg, Variational Methods for the Study of Nonlinear Operators, Translated from the 1956 Russian monograph by A. Feinstein, Holden-Day, San Francisco, 1964.

2. S. G. Mikhlin, Variational Methods in Mathematical Physics, Translated from the 1957 Russian edition by T. Boddington, Pergamon Press, Oxford, 1964.

3. J. T. Oden, Finite Elements of Nonlinear Continua, McGraw-Hill, New York, 1971.

4. E. R. Arantes e Oliveira, "Theoretical Foundations of the Finite Element Method," International Journal of Solids and Structures, Vol. 4, pp. 929-952, 1968.

5. M. E. Gurtin, "Variational Principles in the Linear Theory of Viscoelasticity", Archives for Rational Mechanics and Analysis, Vol. 13, pp. 179-191, 1963.

A FINITE ELEMENT METHOD FOR VARIOUS KINDS OF INITIAL VALUE PROBLEMS

Fumio Kikuchi*

Yoshio Ando**

University of Tokyo

Tokyo, Japan

This paper presents some basic considerations to a most fundamental finite element scheme for initial value problems. The stability and the convergence of the approximate solution are obtained theoretically under some fundamental assumptions on the spatial operator, and it is shown that this scheme is applicable to wide and important classes of evolution equations. Numerical experiments are also performed to various kinds of evolution equations in order to demonstrate the validity of the present method.

SYMBOLS

H	: a real or complex Hilbert space (usually L_2)
X	: a Hilbert space in which the problem is considered
A	: an operator in X; A:X X
$D(A)$	: domain of A; D(A) X
$R(A)$	: range of A; R(A) X
u, v etc.	: elements of X; if u is a vector with m components, it is designated as follows; $u = (u^1, \ldots\ldots, u^m)^T$ (T: transpose)
$(\ ,\ )$	: inner product of X; sometimes $(\ ,\ )_X$ is also employed.
$\|\ \|$	: norm of X; sometimes $\|\ \|_X$ is also employed.
B	: a self-adjoint operator in H; it is often positive bounded below.
H_B	: derived energy space by a self adjoint positive bounded below operator B.
$[\ ,\ ]$	: inner product of H_B (energy product)
$\|\ \|$	: norm of H_B (energy norm)
S_n	: subspace of X in which approximate solutions are sought (n=1,2,. . .)
u_n, v_n, etc.	: elements of S_n
A_n	: continuous approximation operator of A; A_n: S_n S_n
t	: time
Δt, Δx	: time and spatial mesh sizes
θ	: parameter; $0.0 \leq \theta \leq 1.0$
t_θ	: $t + \theta\ \Delta\ t$
u_θ	: $\theta u(t+\Delta t) + (1-\theta)\, u(t)$
u_n^*	: reference element in S_n for the check of consistency

---

*Graduate Student, Graduate School of Engineering

**Professor, Department of Nuclear Engineering, Faculty of Engineering

I. INTRODUCTION

Recently, finite element methods have been employed not only in structural problems but also in non-structural ones, and a range of their applicability is enlarged from linear boundary value problems and eigenvalue ones to nonlinear problems and initial ones. In fact, equilibrium or steady states in nature can exist only approximately and linearity is also idealization of nonlinearity. Therefore, these effects should be taken together and analyzed at the same time in rigorous treatment, and as a matter of course much consideration should be now given to nonlinear and dynamic problems in the finite element methods (References 1 and 2).

However, the foundation of such analysis does not seem to be sufficiently developed, though some basic investigations have already been made mainly in wave and heat equations (References 3–11). For example, the instability of numerical solutions which one encounters in the analysis of dynamic problems by the finite element methods is an important thing that must be solved, and the situation will become more complicated if nonlinearity is introduced. It is also to be noted that the finite element methods should be valid irrespective of types of differential equations such as hyperbolic and parabolic, because different kinds of phenomena must be treated at the same time in the analysis of general complex systems.

This paper is intended to give some basic considerations and to present a unified treatment to a certain extent to a fundamental class of linear and/or nonlinear initial value problems from the above-mentioned standpoint.

The finite element method for initial value problems heretofore developed can be probably classified as follows, where the space is discretized by the usual finite element technique in all cases;

(1) finite difference methods: time is discretized by suitable finite difference schemes. (References 3–7, 9–11).

(2) direct generalization of the finite element methods in space: time and space is simultaneously discretized by similar methods to the usual finite element methods (Reference 12).

(3) some special methods based on the linearity or other character of the considered systems; mode-superposition methods and Laplace transform are examples of such methods.

(4) other methods: for example, combination of (1) and (2) may be sometimes possible. (Reference 13).

Though the methods belonging to (3) are very convenient for special classes of problems and have been often employed, they have common fatal defect of difficult application to nonlinear and time-varying systems. On the other hand the methods belonging to (1) and (2) can be used to nonlinear problems at least formally.

This paper first presents some considerations to the methods belonging to (2), and then discusses mainly a special fundamental type of method belonging to (1) both theoretically and experimentally. That is, the popular method in which the first order finite difference quotient used in the discretization of the time is exclusively discussed. This method has often been employed in special classes of

problems, but some extension is made so that it can deal with wider classes of problems by imposing some fundamental properties on the spatial operator independent of linearity and nonlinearity of the problem. As a result, it is shown that the present method is applicable to fairly wide classes of evolution equations such as heat, wave, and Schrodinger equations, and unified treatment becomes possible to a certain extent. Furthermore, it is also shown that the class of problems treated in this formulation nearly corresponds to the class to which the theory of nonlinear semigroups in Hilbert space can be applied.

Numerical experiments are also conducted to some evolution equations in order to show the validity of the present theory, and it is demonstrated that the present method is actually applicable to various kinds of initial value problems.

In the course of the present theory, the elemental part of the functional analysis is used, and the concepts and techniques of the finite difference methods are also introduced if they are considered to be necessary.

II. PRELIMINARIES

In this section, some preliminaries are given briefly in preparation for the subsequent sections.

II – 1. General

Let X (or sometimes H) be a real or complex separable Hilbert space and the inner product and the norm of X be denoted by (,) and ‖ ‖ respectively (Reference 14).

Let A be an operator with both its domain and range in X;

$$A : X \quad X \tag{1}$$

A is generally unbounded and nonlinear. The domain and the range of A are designated by D(A) and R(A) respectively. Boundary conditions associated with A are assumed to be homogeneous in the theoretical treatment of this paper. In the present theory, A is required to satisfy the following condition;

$$\text{(CA)} \quad \mathrm{Re}(Au - Av, u - v) > \quad \|u - v\|^2 \quad \text{for} \quad u, v \in D(A) \tag{2}$$

where is a real constant independent of u and v. If ≥ 0 (>0), then A is called accretive (strictly accretive). In such cases, $-A$, is called dissipative (strictly dissipative) (References 14–17).

The condition (CA) can be regarded as a kind of stability condition. That is, the system is stable if A is strictly accretive, and, even if ≤ 0, it can be stabilized by adding λI to A, where I is the identity operator in X and $\lambda >$ – . Therefore, this condition may not be sufficiently general but can be regarded as fundamental.

If X is a general Banach space in which the inner product is no longer available, (CA) is replaced by the next condition;

$$(CA')\quad \| u - v + \lambda (Au - Av) \| \geqq (1 + \lambda) \| u - v \| \quad \text{for} \quad u, v \epsilon D(A) \qquad (3)$$

where λ is an arbitrary positive number that satisfies $1 + \lambda > 0$.

Let S_n be a finite-dimensional subspace of X (n = 1,2, ...). It is not necessarily required that $S_n \quad S_m$ if $n < m$, but the next condition is always required;

for $u \epsilon X$, at least one $u_n \epsilon S_n$ can be chosen for each n(= 1,2, . . .) in such a way that

$$\lim_{n \to \infty} \| u_n - u \| = 0 \qquad (4)$$

Let us consider a continuous approximation operator of A;

$$A_n : S_n \qquad S_n, \quad D(A_n) = S_n \qquad (5)$$

It is required that

for $u \epsilon D(A)$, at least one $u_n \epsilon S_n$ can be chosen for each n such that

$$\lim_{n \to \infty} \| A_n u_n - Au \| = 0, \qquad \lim_{n \to \infty} \| u_n - u \| = 0 \qquad (6)$$

Moreover, A_n conserves (CA) in S_n;

$$\mathrm{Re}(A_n u_n - A_n v_n, u_n - v_n) \geqq \| u_n - v_n \|^2 \quad \text{for} \quad u_n, v_n \epsilon S_n \qquad (7)$$

where is not necessarily equal to , but it is independent of n, and $\geqq 0$ (> 0) if $\geqq 0$ (> 0).

The existence of such subspaces and approximation operators is assumed in this paper, but the condition that X is a separable Hilbert space is usually necessary.

II – 2. Approximate method for boundary value problems

Let us consider the following boundary value problem:

$$(BVP) \qquad Au = f \qquad \text{for } f \epsilon X \qquad (8)$$

where A is the operator given in II – 1 and it is assumed to be strictly accretive in this case. The uniqueness and existence of the solution are assumed here, but Galerkin's method is often employed as a mean of constructive proofs (Reference 15).

The approximate equation to (BVP) is given as follows by use of A_n and S_n introduced in II – 1;

$$A_n u_n = f_n \qquad \text{for} \quad f_n \in S_n \tag{9}$$

where f_n satisfies the following condition;

$$\lim_{n\to\infty} \| f_n - f \| = 0 \tag{10}$$

The uniqueness of u_n in (9) is apparent because >0, and its existence can be proved by use of Brouwer's fixed point theorem because A_n is continuous in S_n (Reference 15). In linear problems, the existence is derived from the uniqueness.

A simple error estimation of u_n in (9) is given as follows;

$$\| u_n - u \| = \min_{u_n^* \in S_n} (\| u_n^* - u \| + {}^{1} \| A_n u_n^* - f_n \|) \tag{11}$$

This estimation is exact because u_n^* can be taken as u_n and the second term in the right side of (11) vanishes for u_n. The convergence is assured if the conditions (6), (7) and (10) hold, where the next relation is used;

$$\| A_n u_n^* - f_n \| \leqq \| A_n u_n^* - f \| + \| f - f_n \| \tag{12}$$

The formula (11) is useful if a simple u_n^* can be easily found. In the finite difference methods, the restriction of u of (8) to nodal points is usually used as u_n^*, but its analog does not seem to be always effective in the finite element method.

In the finite element method, S_n is usually constructed as follows;

(a) S_n $D(A)$

(b) for $u \epsilon D(A)$, at least one $u_n \epsilon S_n$ can be chosen for each n such that

$$\lim_{n\to\infty} \| Au_n - Au \| = 0, \qquad \lim_{n\to\infty} \| u_n - u \| = 0 \tag{13}$$

Then A_n and f_n are decided as follows;

$$A_n = P_n A, \qquad f_n = P_n f \tag{14}$$

where P_n is the projection operator from X into S_n. In this case, the conditions (6) and (10) are assured because $\| P_n \| \leqq 1$ and (13) holds, and in (7) can be taken as because

$$\mathrm{Re}(A_n u_n - A_n v_n, u_n - v_n) = \mathrm{Re}(Au_n - Av_n, u_n - v_n) \geqq \| u_n - v_n \|^2 \tag{15}$$

for $u_n, v_n \epsilon S_n$. The continuity of A_n is assured if A is hemi-continuous (Reference 15).

Sometimes, the present conditions for S_n and A_n are too strong for practical uses. If A is a self-adjoint positive bounded below operator in X (or H), then the theory of energy space can be employed and the conditions (13) can be replaced by the following ones;

(a') $S_n \quad H_A$

(b') for $u \in H_A$, at least one $u_n \in S_n$ can be chosen for each n such that

$$\lim_{n \to \infty} |u_n - u| = 0 \tag{16}$$

where H_A is the energy space derived by A, and $|\ \ |$ is the energy norm. A_n is defined as follows;

$$(A_n u_n, v_n) = [u_n, v_n] \qquad \text{for} \quad u_n, v_n \in S_n \tag{17}$$

where $[\ ,\ ]$ is the energy product (References 18 and 19). In this case can be taken as again. A simple error estimation is, as is well known, given as follows;

$$|u_n - u| = \min_{u_n^* \in S_n} |u_n^* - u| \qquad (u, u_n = \text{solutions of (8) and (9) respectively})$$

$$\| u_n - u \| \leq |u_n - u| / \tag{18}$$

Therefore, u_n itself can be used as u_n^* in (6), though it seems meaningless for the present problem (BVP).

The theory presented in this section is an abstract theory of approximate methods for (BVP) in Hilbert spaces, and it is effective not only to the finite element methods but also to some other methods such as the finite difference methods. It is also to be noted that the non-conforming solutions in the (pseudo-) finite element methods are still vali ! if the corresponding A_n satisfies the above-mentioned conditions.

III. INITIAL VALUE PROBLEMS AND SOME APPROXIMATE METHODS

In this section, the evolution equation to be treated in this paper is introduced and some approximate methods for it are given.

III – 1. Initial value problem

In this paper, initial value problems are treated in the form of the following evolution equation;

$$\text{(IVP)} \qquad \frac{du}{dt} + A(t)\, u = f(t) \quad , \qquad 0 \leq t \leq T$$
$$u(0) = u_0 \tag{19}$$

where t indicates time and T is a positive constant which can be usually taken arbitrarily. It is assumed that $u_0 \in D(A)$ and $f(t) \in X$ at each t. A(t) is an operator

in X at each t and is admitted to depend on t smoothly. Furthermore, A(t) satisfies (CA) at each t, and is dependent only on T. If (IVP) is considered in $0 \leq t < \infty$, then is assumed to be constant. The term f(t) is separated from A(t)u for convenience sake.

If some subsidiary conditions are imposed on A, f and u_0, then the well-posedness of (IVP) is assured by the theory of nonlinear semi-groups (References 16 and 17). The unified treatment of nonlinear evolution equations is available to the extent of the above-mentioned classes, as far as the authors know. It is also to be noted that representative initial value problems in mathematical physics can be included to the present classes as will be seen in Section V.

In order to solve (IVP) approximately, the following discretizations are generally needed;

(a) discretization with respect to t

(b) discretization with respect to X

In the subsequent sub-sections, some approximate methods will be shown briefly.

III – 2. Boundary value techniques for initial value problems

Generally speaking, initial value problems can be interpreted as a kind of generalized boundary value problems if the initial conditions are regarded as boundary values. Therefore, the method given in III – 2 can be employed to them if they satisfy the conditions given there. In such treatment, the Hilbert space should be constructed on $X \times [0,T]$.

Of course, the popular step by step methods are usually more convenient because less numbers of unknowns are needed at each step. However, the boundary value techniques may be effective to periodic and short transient problems, in which long-time calculation is not necessary. Moreover, this approach can be regarded as a direct extension of the usual finite element methods based on Rayleigh–Ritz–Galerkin's method. This type of method is proposed by Zienkiewicz and Parekh (Reference 12), and a finite difference method using a similar (a little different, though) approach is discussed by Caraso and Parter (Reference 20). Some results will also be given by the present authors (Reference 21).

In order to treat (IVP) as (BVP), the next inner product and norm are introduced;

$$((u, v)) = \int_0^T (u, v)\, dt \qquad \| u \| = \sqrt{((u, u))} \tag{20}$$

Thus a Hilbert space can be constructed on $X \times [0,T]$.

In the case that A is strictly accretive, it can be proved that the next operator is strictly accretive in this new Hilbert space;

$$\tilde{A} = \frac{d}{dt} + A \tag{21}$$

The proof is;

$$\mathrm{Re}((\dot{A}u - \dot{A}v,\ u - v)) = \tfrac{1}{2}\,\|u(T) - v(T)\|^2 + \mathrm{Re}\int_0^T (Au - Av,\ u - v)\,dt$$

$$\geq \int_0^T \|u - v\|^2\,dt = |||\,u - v\,|||^2 \qquad (22)$$

where the following assumption is made;

$$u = v \qquad \text{at} \quad t = 0 \qquad (23)$$

In the case $\lambda \leq 0$, the transformation $u = \exp(\lambda t)\,v$ gives the next problem;

$$\frac{dv}{dt} + \exp(-\lambda t)\,A\,\exp(\lambda t)\,v + \lambda v = \exp(-\lambda t)\,f(t), \quad v(0) = u(0) \qquad (24)$$

and $\dot{A} = d/dt + \exp(-\lambda t)\,A\,\exp(\lambda t) + \lambda I$ is strictly accretive if $\lambda > -$.

Of course some consideration is necessary with respect to the domain of $A\exp(\lambda t)$ as was pointed by Kato (Reference 17), but this method is applicable at least formally and is certainly valid if A is linear.

III – 3. Semi-discrete methods

In this sub-section, two types of semi-discrete methods (Reference 22) are given

(a) discretization only to t; in this case, d/dt is replaced by its finite difference analog;

$$\frac{u(t + \Delta t) - u(t)}{\Delta t} + A(\tilde{t})\,\tilde{u}(t) = f(\tilde{t}), \quad u(0) = u_0 \qquad (25)$$

where Δt : time mesh size ($t = m\,\Delta t$, $m = 0, 1, 2, \ldots$)

$\tilde{t} = t + \theta\,\Delta t$, (θ : parameter, $0 \leq \theta \leq 1.0$)

$\tilde{u}(t) = \theta u(t + \Delta t) + (1 - \theta)\,u(t)$ (Fig. 1)

(b) discretization only to X : in this case, X and A are replaced by S_n and A_n;

$$\frac{du_n}{dt} + A_n u_n = f_n(t), \qquad u_n(0) = u_{n0} \qquad (26)$$

where f_n and u_{n0} are suitable approximations of f and u_0. This method is called Faedo–Galerkin's method if A_n and f_n defined in (14) or (17) are used (Reference 15)

The discussion of these methods are omitted here, but it is to be noted that they sometimes give useful information as the foundation of the perfectly discrete

method given in the next section. Especially, the scheme (25) is valid if $0.5 \leqq \theta \leqq 1.0$. Furthermore, it is valid even in general Banach space if $\theta = 1.0$ (Reference 14).

IV. A FINITE ELEMENT METHOD FOR INITIAL VALUE PROBLEMS

In this section, a finite element scheme belonging to step by step methods is given and its validity is discussed.

IV – 1. Formulation

The equation (19) is now discretized with respect to both time and space;

$$\frac{u_n(t+\Delta t) - u_n(t)}{\Delta t} + A_n(\tilde{t})\,\tilde{u}_n(t) = f_n(\tilde{t}), \qquad u_n(0) = u_{n0} \tag{27}$$

where $\tilde{t}$, $\tilde{u}_n(t)$, A_n, f_n and u_{n0} are perfectly the same as those of (25) and (26). In order to solve (27), one begins from u_{n0} and proceeds step by step. The scheme for $\theta = 0.5$ is known as Crank–Nicolson's one in the finite difference methods (Reference 23). Practically, the schemes for $\theta = 0.0$, 0.5 and 1.0 are mainly employed. Of course, it is assumed that A_n satisfies (7) at each t.

Some similar schemes are also available;

$$\frac{u_n(t+\Delta t) - u_n(t)}{\Delta t} + \theta A_n(\tilde{t})\,u_n(t+\Delta t) + (1-\theta)A_n(\tilde{t})u_n(t) = f_n(\tilde{t}) \tag{28}$$

$$\frac{u_n(t+\Delta t) - u_n(t)}{\Delta t} + \theta A_n(t+\Delta t)u_n(t+\Delta t) + (1-\theta)A_n(t)u_n(t) = f_n(\tilde{t}) \tag{29}$$

where $f_n(\tilde{t})$ can be replaced by $\tilde{f}_n(t)$ if it is preferred. These schemes coincide with (27) if A_n is linear and independent of t, but the results obtained for (27) do not seem to hold generally except for $\theta = 0.0$ and 1.0.

IV – 2. On the validity of the present method

In order to give rigorous treatment to the approximate system the next three concepts that have been mainly used in the finite difference methods are employed in a little modified forms (References 23, 24 and 25).

(a) consistency this condition implies that the approximate system actually approximates the original system (19), and is defined as follows,

for the solution u of (19), at least one $u_n^* \in S_n$ can be chosen for each n in such a way that

$$\lim_{\substack{n\to\infty \\ \Delta t \to 0}} \tau_n^1 = 0, \qquad \lim_{n\to\infty} \tau_n^2 = 0 \tag{30}$$

where
$$\tau_n^1 = \frac{u_n^*(t+\Delta t) - u_n^*(t)}{\Delta t} + A_n(\tilde{t})\tilde{u}_n^*(t) - f_n(\tilde{t})$$
$$\tau_n^2 = u_n^* - u \tag{31}$$

The convergence in (30) is assumed to be uniform in $[0, T]$. If some relations are necessary between n and Δt to achieve consistency, then the scheme is called conditionally consistent. Otherwise, it is called unconditionally consistent.

(b) convergence : this means that the approximate solution u_n converges to the exact solution of (19) in the metric of X as $n \to \infty$ and $\Delta t \to 0$. The convergence is pointwise in $[0, T]$ and not necessarily uniform.

(c) stability : this means that the growth of the approximate solution is bounded to a certain extent. In the present case, only stability with respect to initial value and external source term is considered and it is given as follows;

$$\| u_n(t) \| \leq C_1(T) \| u_n(0) \| + C_2(T) \| \| f_n \| \|_\infty \quad 0 \leq t \leq T \tag{32}$$

where C_1 and C_2 are nonnegative constants dependent only on A and T, and

$$\| \| f_n \| \|_\infty = \sup_{0 \leq t \leq T} \| f_n(t) \| \tag{33}$$

Usually $u_n(0)$ and f_n in (32) can be replaced by u(0) and f. Sometimes a little stronger condition is preferred;

$$\| u_n(t) - v_n(t) \| \leq C_1(T) \| u_n(0) - v_n(0) \| + C_2(T) \| \| f_n - g_n \| \|_\infty \tag{34}$$

where $u_n(t)$ and $v_n(t)$ are solutions of (27) corresponding to the initial values $u_n(0)$ and $v_n(0)$ and the force terms f_n and g_n respectively. Equations (32) and (34) are identical in linear problems but seriously different in general nonlinear problems.

The existence and uniqueness of the approximate solution are also essential, though they can be derived from stability in linear problems. In general nonlinear problems, uniqueness can be obtained from the stability of (34) but existence is not assured from the above-mentioned three conditions. Therefore, a proof of the existence is given in the next section in the case of general approximate methods. It is also to be noted that convergence is assured if the stability of the type (34) and consistency hold. This is known as Lax-Kreiss' equivalence theorem in the case of linear problems. On the other hand, uniqueness and convergence are not generally obtained from the stability of (32).

The definition of consistency presented here is a little modified so that it is available to general purposes. This is mainly due to the fact that the restriction of the exact solution u to nodal points, which is usually employed in the finite difference methods, is no longer effective in general approximate methods. Furthermore, this consistency condition is necessary for convergence by the same reason in II – 2.

In the case of the finite element method, the consistency can be assured if S_n is chosen as follows (cf. II – 2);

(i) S_n $D(A)$

(ii) for u of (19), there exists at least one u_n^* in each S_n such that

$$\text{(ii – 1)} \quad \lim_{n\to\infty} \| u_n^*(t) - u(t) \| = 0$$

$$\text{(ii – 2)} \quad \lim_{\substack{n\to\infty \\ \Delta t \ 0}} \left\| \frac{u_n^*(t+\Delta t) - u_n^*(t)}{\Delta t} - \frac{du}{dt}\Big|_t \right\| = 0$$

$$\text{(ii – 3)} \quad \lim_{\substack{n\to\infty \\ \Delta t \ 0}} \| A(\tilde{t})\tilde{u}_n^*(t) - A(\tilde{t})\tilde{u}(t) \| = 0 \qquad (35)$$

$$\text{(ii – 4)} \quad \lim_{\Delta t \ 0} \| A(\tilde{t})\tilde{u}(t) - A(\tilde{t})u(t) \| = 0$$

where the convergence is assumed to be uniform in $[0,T]$. (ii – 4) is a smoothness condition for A and u. Generally it can be expected that the value of (ii – 2) is smallest for $\theta = 0.5$ if u is sufficiently smooth.

As for the stability, it can be easily shown that the next conditions are sufficient ones;

$$\| u_n(t+\Delta t) \| \leq (1 + C_3\Delta t) \left(\| u_n(t) \| + C_4\Delta t \| f_n \|_\infty \right) \text{ for (32)} \qquad (36)$$

$$\| w_n(t+\Delta t) \| \leq (1 + C_3\Delta t) \left(\| w_n(t) \| + C_4\Delta t \| h_n \|_\infty \right) \text{ for (34)} \qquad (37)$$

where C_3 and C_4 are nonnegative constants independent of n, t and Δt, and dependent only on ', and $w_n = u_n - v_n$, $h_n = f_n - g_n$.

By using (7), the next main result of this paper with respect to existence and stability of approximate solutions can be obtained;

> In the case $0.5 \leq \theta \leq 1.0$, the approximate solution of (27) exists uniquely and the stability of (34) holds if $\Delta t < C_5$ ('), where C_5 is a positive constant dependent only on '. (if ' ≥ 0, then C_5 can be taken as ∞.)

Of course, the convergence is also assured if the consistency is guaranteed.

IV – 3. Sketches of proofs and some remarks

In this sub-section, the outline of the proofs of existence and stability is given.

(a) existence and uniqueness of the approximate solution ($0 \leq \theta \leq 1.0$)

For θ = 0.0, the uniqueness and existence are apparent. In other cases, the procedure in each step is essentially to solve the next boundary value problem;

$$\tilde{u}_n + \theta \Delta t A_n \tilde{u}_n = e_n \tag{38}$$

where e_n is a known element in S_n. Because $\theta\Delta t$ is positive, the existence and uniqueness can be derived by use of the results in II – 2 if ≥ 0. Same conclustion is derived if

$$\Delta t \leq \frac{-1}{\theta}, \qquad \text{for} \quad ' < 0 \tag{39}$$

(b) stability ($0.5 \leq \theta \leq 1.0$)

The inequality (37) is to be derived. From the definitions of u_n, v_n, w_n and h_n, the next equation is obtained;

$$\frac{w_n(t+\Delta t) - w_n(t)}{\Delta t} + A_n(\tilde{t})\,\tilde{u}_n(t) - A_n(\tilde{t})\,\tilde{v}_n(t) = h_n(\tilde{t}) \tag{40}$$

The following inequalities can be derived by use of such fundamental inequalities as the triangle inequality and Cauchy–Schwartz' one;

$$\mathrm{Re}(w_n(t+\Delta t) - w_n(t),\ \tilde{w}_n(t)) \geq \tfrac{1}{2}\left(\| w_n(t+\Delta t)\|^2 - \| w_n(t)\|^2\right) \qquad \text{for } 0.5 \leq \theta \leq 1.0 \tag{41}$$

$$\| \tilde{w}_n(t)\| \leq \theta \| w_n(t+\Delta t)\| + (1-\theta)\| w_n(t)\| \qquad \text{for } 0.0 \leq \theta \leq 1.0 \tag{42}$$

The next inequality is derived from (7);

$$\mathrm{Re}(A_n\tilde{u}_n(t) - A_n\tilde{v}_n(t),\ \tilde{w}_n(t)) \geq \ '\| \tilde{w}_n(t)\|^2 \tag{43}$$

By multiplying (40) by $\tilde{w}_n(t)$ and taking the real part of their inner product, the following inequality can be obtained with the aid of (41) and (43);

$$\begin{aligned} &\tfrac{1}{2}\left(\| w_n(t+\Delta t)\|^2 - \| w_n(t)\|^2\right) + \ '\Delta t \| \tilde{w}_n(t)\|^2 \\ &\leq \Delta t \,\| h_n \|_\infty \| \tilde{w}_n(t)\| \end{aligned} \tag{44}$$

In the case ≤ 0 and $\| w_n(t+\Delta t)\| \geq \| w_n(t)\|$, the next inequality is derived by use of (42);

$$\begin{aligned} \| w_n(t+\Delta t)\|^2 \leq \| w_n(t)\|^2 &- 2\ '\Delta t \| w_n(t+\Delta t)\|^2 \\ &+ 2\Delta t \,\| h_n\|_\infty \| w_n(t+\Delta t)\| \end{aligned} \tag{45}$$

Therefore,

$$\| w_n(t+\Delta t) \| \leq \frac{1}{1+2\ '\Delta t}(\| w_n(t) \| + 2\Delta t \| h_n \|_\infty) \tag{46}$$

if $\Delta t < \frac{-1}{2\ '}$. In the case $\| w_n(t) \| \geq \| w_n(t+\Delta t) \|$, (46) is apparent.

In a similar way, the next inequality can be derived in the case $' \geq 0$;

$$\| w_n(t+\Delta t) \| \leq \| w_n(t) \| + 2\Delta t \| h_n \|_\infty \tag{47}$$

In any case, (37) is valid because $1/(1+2\ '\Delta t) \leq 1-2\ '\Delta t$ ($'<0$) for sufficiently small Δt.

(c) Remarks

(c–1) As is seen from the above proof, C_3 in (37) can be taken to be 0 if $' \geq 0$. In such cases, the calculation by the present scheme is very stable and the accumulation of round-off error, which is not considered in this paper, is not so serious.

(c–2) In the case $h_n = 0$, the next estimation is obtained for $0.5 \leq \theta \leq 1.0$ by use of some inequalities;

$$\| w_n(t+\Delta t) \|^2 \leq \frac{1-2\ '(2\theta-1)(1-\theta)\Delta t}{1+2\ '\theta(2\theta-1)\Delta t} \| w_n(t) \|^2 \quad \text{for} \quad ' \geq 0$$

$$\| w_n(t+\Delta t) \|^2 \leq \frac{1-2\ '(1-\theta)\Delta t}{1+2\ '\theta\Delta t} \| w_n(t) \|^2 \quad \text{for} \quad ' \leq 0 \tag{48}$$

Therefore, w_n grows at most exponentially with respect to t.

(c–3) The stability of type (32) can be derived in a similar way under the next assumption of A_n;

$$\mathrm{Re}(A_n u_n, u_n) \geq \| u_n \|^2 \quad \text{for} \quad u_n \in S_n \tag{49}$$

The result is obtained by Kreiss in his study of linear finite difference equations (Reference 26).

(c–4) In general Banach space, in which scalar product is generally no longer available, stability can be obtained in a similar way if $\theta = 1.0$ and A_n satisfies (3) in S_n. Therefore, the present theory is applicable even in general Banach space if A_n is linear because the existence can be derived from stability. As for nonlinear problems, the existence is not assured by the method given in this sub-section. Such consideration is not so useful for the finite element methods, which are usually constructed in Hilbert spaces, but may be effective in general approximate methods.

(c–5) In the case of schemes (28) and (29), the uniqueness and existence of u_n can be shown similarly. The stability is assured if

$$\Delta t \leqq \frac{2(\quad' + C)}{(1-\theta)^2 M_n^2} \tag{50}$$

where C is an arbitrary nonnegative constant and $C > -\quad'$ if $\quad' \leqq 0$. M_n is the Lipshitz constant in S_n (if it exists);

$$\| A_n u_n - A_n v_n \| \leqq M_n \| u_n - v_n \| \qquad \text{for} \quad u_n,\ v_n \in S_n \tag{51}$$

M_n generally increases as $n \to \infty$, and its existence is assured if A_n is linear because A_n is continuous in S_n. However, its existence is not guaranteed in nonlinear problems except special cases.

V. SOME FURTHER CONSIDERATIONS ON THE PRESENT METHOD

V – 1. Examples of evolution equations to which the present method is applicable

In this sub-section, some examples of evolution equations to which the present method can be employed are shown with brief explanations. The condition (CA) can be proved by use of symmetric properties of the system in the case of (b), (c), (d), and (e).

(a) heat equation : $\dfrac{du}{dt} + B\,u = f(t), \quad f \in X$ (external source term) (52)

B is a self-adjoint positive bounded below operator in X, the examples of which are certain classes of elliptic operators (Reference 18). In this case, (CA) is apparent because B is positive. This result can be extended to wider classes of elliptic operators such as strongly elliptic ones (Reference 14).

(b) wave equation : $\dfrac{d^2u}{dt^2} + B\,u = f(t), \quad f \in H$ (external force term) (53)

B is same as that of (a) and is independent of t. In order to transform (53) into the form of (19), the following two methods are available.

(b–1) decomposition of B :

Because B is self-adjoint and positive, it can be decomposed as follows;

$$B = T^*T, \quad T : H \quad H', \quad T^* : H' \quad H \tag{54}$$

where H' is another Hilbert space and T^* is the adjoint operator of T (Reference 27). Then the equation (53) can be transformed as follows;

$$\frac{du^1}{dt} + T^*u^2 = f(t), \quad \frac{du^2}{dt} - T\,u^1 = 0 \quad \left(u^1 = \frac{du}{dt},\ u^2 = T\,u\right) \tag{55}$$

where X is taken as HxH' and its scalar product and norm is defined by

$$(u, v)_X = (u^1, v^1)_H + (u^2, v^2)_{H'}, \quad \| u \|_X = \sqrt{(u, u)_X} \tag{56}$$

The norm can be regarded as square root of the total energy (usually ½ is multiplied). The condition (CA) can be shown by use of the relation

$$(T^* u^2, u^1)_H = (u^2, Tu^1)_{H'} \tag{57}$$

(b–2) use of energy space :

In this case, the equation (53) is transformed into the next form;

$$\frac{du^1}{dt} - u^2 = 0, \quad \frac{du^2}{dt} + B\, u^1 = f(t) \quad (u^1 = u,\ u^2 = \frac{du}{dt}) \tag{58}$$

X should be taken as H_BxH, where H_B is the energy space derived by B, and its inner product and norm are defined as follows;

$$(u, v)_X = [u^1, v^1]_{H_B} + (u^2, v^2), \quad \| u \|_X = \sqrt{(u, u)_X} \tag{59}$$

Again, this norm can be regarded as the square root of the total energy, and (CA) can be shown by use of the following fact;

$$[u^2, u^1]_{H_B} = (u^2, B\, u^1)_H \tag{60}$$

(c) first order wave equation : $\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = f(t), \quad f \in X = L_2$ (61)

It is assumed that u and f have compact support. The condition (CA) can be shown because

$$\mathrm{Re}\,(\frac{\partial u}{\partial x}, u) = -\,\mathrm{Re}\,(u, \frac{\partial u}{\partial x}) = 0 \tag{62}$$

This result can be extended to general symmetric hyperbolic systems and the above-mentioned support condition can be moderated to a certain extent (Reference 28). A majority of differential equations describing reversible phenomena can be expressed in this form.

(d) Schrodinger's equation : $\frac{du}{dt} + i\, B\, u = 0$ (i = imaginary unit) (63)

In this case, the source term is usually absent and B is self-adjoint. (CA) can be derived by use of the fact that B is self-adjoint.

(e) equation of coupled sound and heat flow:

$$\frac{\partial u^1}{\partial t} - c \frac{\partial}{\partial x}(u^2 - (\gamma - 1) u^3) = 0, \quad \frac{\partial u^2}{\partial t} - c \frac{\partial u^1}{\partial x} = 0$$

$$\frac{\partial u^3}{\partial t} - \sigma \frac{\partial^2 u^3}{\partial x^2} + c \frac{\partial u^1}{\partial x} = 0 \tag{64}$$

where c, γ (>1) and σ are positive constants. The physical meanings of (64) are given in Reference 23, and a similar equation is treated by Oden and Kross (Reference 29). This is a coupled type of equation of parabolic and hyperbolic, and (CA) can be shown by imposing suitable boundary conditions (e.g. $u^1 = 0$, $u^3 = 0$ on the boundary) and replacing u^3 by $u^3/\sqrt{\gamma - 1}$.

(f) semilinear heat equation; $$\frac{du}{dt} + B u + h(u) = 0 \tag{65}$$

where B is same as that of (a) and h (nonlinear source term) is assumed to be smooth and non-decreasing to assure (CA). Sometimes certain classes of perfectly nonlinear heat equations can be treated similarly (Reference 15).

The present method is valid irrespective of linearity and non-linearity of A and many other examples can be found. However, it is seen from these examples that representative linear evolution equations in mathematical physics can be analyzed by the present method.

V – 2. Consistency in the case that the theory of energy space is applicable.

In the preceding sections, a general theory for initial value problems is presented. However, the conditions (35) are sometimes too severe in the finite element methods. In this sub-section, it will be shown that such a situation can be moderated to a certain extent in the case that the theory of energy space is applicable.

The equations treated here are : heat equation (a) (in V – 1), wave equation (b – 2) and Schrodinger equation (d). In all cases, B is assumed to be self-adjoint, positive bounded below and independent of t. S_n is so constructed as (a') and (b') in II – 2 and A_n is constructed as (17).

Then, the best approximation of u(t) (exact solution) in H_B can be used as u_n^* in (31) if u is sufficiently smooth. In this way, the error estimation of $B_n u_n^*$ becomes very easy.

It is also to be noted that $d^m u_n^*/dt^m$ ($m = 1, 2, \ldots$) is the best approximation of $d^m u/dt^m$ in H_B if it belongs to H_B. Furthermore, stability and convergence in H_B are assured in the case of heat equation, and stability in H_B is assured in the case of Schrodinger's equation.

V – 3. Conservation of norms of approximate solutions

Heretofore, the convergence and stability of the approximate solution have been mainly discussed. However, there is an important character for certain classes of initial value problems that a certain norm is conserved. Therefore, it is desirable

that not only the approximate solution converges to the exact solution but also the corresponding norms of the approximate solution are conserved. Especially such a situation is favorable for long time calculation because an important physical value is conserved even if coarse mesh division is employed. In practical calculations, it is often observed that the error in the norm (not error of the norm) becomes large as time passes because of the deviation of the phase. Even in such cases, approximate solutions still can be regarded as good ones if such valuables as amplitude, period and norm are in good agreement with the exact values and the wave of decay is small (Fig. 2).

Such a situation can be realized in the following equations if θ is taken as 0.5. In the following examples, the source terms are absent in all cases.

(a) wave equation : (b−1) and (b−2) in V − 1. The corresponding norm is square root of total energy.

(b) first order wave equation : (c) in V − 1. The corresponding norm is L_2 − norm. This result can be extended to general symmetric hyperbolic systems with constant coefficients.

(c) Schrodinger's equation : (d) in V − 1. The corresponding norm is the norm in X (or H) (Reference 30). If B is positive bounded below, the norm of H_B is also conserved.

The proofs are omitted here, but they can be shown in a similar way to that of stability (section IV).

V − 4. Some considerations on the lumping of the mass matrix

As was discussed in the preceding sections, it is shown that the present method is valid at least in the case that $0.5 \leqq \theta \leqq 1.0$. However, the scheme is perfectly implicit in such cases, and, moreover, a new set of simultaneous equation must be solved at each step if A changes from step to step. This situation especially causes great computational difficulty if A is nonlinear.

To avoid this difficulty, the scheme for $\theta = 0.0$ can be used, though stability is not assured unless Δt is sufficiently small (see section V). In this case, a common set of linear simultaneous equations is required to be solved at all steps. However, the scheme thus obtained is still implicit, and use of the lumping the matrix derived from I (identity operator in X) is often made to make it explicit. This modified matrix is usually called lumped mass matrix in structural analysis, because it appears as the term of inertia.

In this sub-section, it will be shown in simple examples that more stable schemes can be sometimes obtained if lumping is suitably made. In the following examples, the piecewise linear shape function is employed and spatial mesh size Δx is assumed to be uniform. Furthermore, u is used instead of u_n.

(a) heat equation : $$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0 \qquad (66)$$

The finite element scheme for θ = 0.0 at interior mesh points is given as follows:

$$\frac{u(t+\Delta t,x+\Delta x)+4u(t+\Delta t,x)+u(t+\Delta t,x-\Delta x)-u(t,x+\Delta x)-4u(t,x)-u(t,x-\Delta x)}{6\,\Delta\, t}$$

$$-\frac{u(t,x+\Delta x)-2u(t,x)+u(t,x-\Delta x)}{\Delta\, x^2} = 0 \qquad (67)$$

This scheme is stable in L_2 if $\Delta t/\Delta x^2 \leqq 1/6$ (see V – 5). If lumping is ma made with respect to both t and t+Δt, the next scheme is obtained;

$$\frac{u(t+\Delta t,x)-u(t,x)}{\Delta t} - \frac{u(t,x+\Delta x)-2u(t,x)+u(t,x-\Delta x)}{\Delta x^2} = 0 \qquad (68)$$

As is well known, this scheme is stable not only in L_2 but also in the uniform norm if $\Delta t/\Delta x^2 \leqq 0.5$ (Reference 23).

(b) first order wave equation: $$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = C \qquad (69)$$

The finite element scheme for θ = 0.0 at interior mesh points is given as follows;

$$\frac{u(t+\Delta,x+\Delta x)+4u(t+\Delta t,x)+u(t+\Delta t,x-\Delta x)-u(t,x+\Delta x)-4u(t,x)-u(t,x-\Delta x)}{6\,\Delta\, t},$$

$$+\frac{u(t,x+\Delta x)-u(t,x-\Delta x)}{2\,\Delta\, x} = 0 \qquad (70)$$

This scheme is stable in L_2 if $\Delta t/\Delta x^2 \leqq C$ (C = arbitrary positive constant). If the same lumping as (a) is made, the stability condition is not improved essentially. However, if lumping is made only to t+Δt, then the next scheme is obtained;

$$\frac{6u(t+\Delta t\ x)-u(t,x+\Delta x)-4u(t\ x)-u(t,\quad \Delta x)}{6\,\Delta\, t} + \frac{u(t,x+\Delta x)-u(t,x-\Delta x)}{2\,\Delta\, x} = 0 \qquad (71)$$

This scheme is consistent with (69) if $\Delta t/\Delta x$ is kept constant. As is easily proved, this scheme is stable in uniform norm if $\Delta t/\Delta x \leqq 1/3$.

In these examples, stability is improved by lumping. Similar result is obtained in the wave equation if two-level scheme is employed. However, such a conclusion does not hold generally because general method of lumping is not known. Anyway, the next assertion must be correct : by lumping, A_n is modified; if the modified A_n satisfies the conditions in II and IV, then the general theory of this paper can be adopted.

V – 5. Some results for $0.0 \leqq \theta < 0.5$ (linear problems)

In order to deal with the case $0.0 \leqq \theta < 0.5$, which is not referred to yet, the method of expansion by eigenfunctions is effective if A_n is linear and independent of t and any element of S_n can be expressed by linear combination of

eigenfunctions of A_n. Such a situation is realized, for example, if A_n is self-adjoint in S_n.

Then, a simple stability condition is given by

$$\Delta t \leqq \min_i \frac{2(C + \mathrm{Re}\,\lambda_{ni})}{(1-2\theta)|\lambda_{ni}|^2} \quad \text{for } 0.0 \leqq \theta \leqq 0.5 \quad (1 \leqq i \leqq \dim(S_n)) \tag{72}$$

where λ_{ni} is the i'th eigenvalue of A_n and C is an arbitrary non-negative constant that satisfies $C + \mathrm{Re}\lambda_{ni} > 0$. It is also to be noted that $\mathrm{Re}\lambda_{ni} \geqq$ '. In the case that $\theta = 0.0$, $C = 0.0$ and $\mathrm{Re}\lambda_{ni} > 0$, this result is identical with that of Reference 22.

Similar methods such as method of Fourier series and Fourier transform are also available if spatial mesh size is uniform and the coefficients of A are constant (Reference 23).

Some results obtained by these methods are shown for typical linear evolution equations;

(a) heat equation : (a) in V − 1 $\quad \Delta t \leqq \dfrac{2}{(1-2\theta)\,\lambda_{n\,\max}}$ (73)

(b) wave equation : (b − 2) in V − 1 $\quad \Delta t \leqq \dfrac{C}{(1-2\theta)\,\lambda_{n\,\max}}$ (74)

(c) Schrodinger's equation : (d) in V − 1 $\quad \Delta t \leqq \dfrac{C}{(1-2\theta)\,\lambda_{n\,\max}^2}$ (75)

(d) first order wave equation : (c) in V − 1 (in this case, piecewise linear shape function with uniform spatial mesh size Δx is employed.) $\quad \Delta t \leqq \dfrac{C\Delta x^2}{1-2\theta}$ (76)

where $\lambda_{n\,\max}$ is the largest eigenvalue of B_n, which is assumed to be self adjoint and positive bounded below, and C is an arbitrary positive constant. If $B = -d^2/dx^2$ and piecewise linear shape function with uniform mesh size Δx is used, then $\lambda_{n\,\max} = 12/\Delta x^2$. In (74), (75) and (76), there is no definite threshold value for stability as in (73) because C is arbitrary, and Δt must be made sufficiently small until reasonable results are obtained in practical computation.

The present condition only assures that the stability is guaranteed if $n \to \infty$ and $\Delta t \to 0$ with the above-mentioned conditions preserved. In such cases, the order of C_3 in (36) or (37) is equal to that of C.

As is seen from these results, the stability condition changes greatly if the equation changes. This is an important point to be recognized especially in the calculation of mixed type phenomena, in which different types of equations must be treated together. Furthermore, the conditions required are often too severe to carry out practical calculations. In such cases, the scheme for $0.5 \leqq \theta \leqq 1.0$ should be used.

V – 6. Two-level scheme for wave equation

In this sub-section, some consideration is given to two-level scheme for wave equation because such scheme has been often discussed (References 3, 4, 5 and 7). The scheme treated here is given as follows;

$$\frac{u_n(t+\Delta t) - 2u_n(t)+u_n(t-\Delta t)}{\Delta t^2} + B_n[\theta_2\,\theta_1 u_n(t+\Delta t)+(1-\theta_1)u_n(t) + (1-\theta_2)\,\theta_1 u_n(t)+(1-\theta_1)u_n(t-\Delta t) = 0 \tag{77}$$

where θ_1 and θ_2 are constants ($0 \leqq \theta_1, \theta_2 \leqq 1$), B_n is the approximation operator of B given in V – 1, and the source term f_n is omitted here because it does not affect the stability.

By use of a similar method to that of section IV, the unconditional stability in the following norm is obtained if $0.5 \leqq \theta_1, \theta_2 \leqq 1.0$;

$$\sqrt{\left\| \frac{u_n(t+\Delta t)-u_n(t)}{\Delta t} \right\|_H^2 + \left| \theta_1 u_n(t+\Delta t)+(1-\theta_1)u_n(t) \right|_{H_B}^2} \tag{78}$$

Therefore, the stability in the total energy is obtained, where velocity is replaced by its finite difference analog.

If the method of expansion by eigenfunction is employed, a bit more information can be obtained;

$$\begin{aligned} &\text{unconditionally stable} && \text{if } 3\theta_3 + \theta_4 \geqq 1 \text{ and } \theta_3 \geqq \theta_4 \\ &\Delta t \leqq \frac{2}{\sqrt{1-3\theta_3-\theta_4}\,\lambda_{n\,\max}} && \text{if } 3\theta_3 + \theta_4 < 1 \text{ and } \theta_3 \geqq \theta_4 \end{aligned} \tag{79}$$

where $\theta_3 = \theta_1\theta_2$ and $\theta_4 = (1-\theta_1)(1-\theta_2)$, and $\lambda_{n\,\max}$ is the largest eigenvalue of B_n. This result includes some of the results given in References 3, 4, 5, and 7. The latter condition (79) is, however, weaker than the former (78) because this stability is essentially stability of u_n and not that of velocity.

V – 7. Some remarks on the present method for more general nonlinear problems

As was already referred to, the present method is valid to nonlinear problems if (CA) holds. However, nonlinear problems are very complex and it cannot be expected that A satisfies (CA) in general nonlinear problems. Therefore, there arises a question whether the present method is applicable to such problems in which the global solution may not exist for arbitrary initial values and source terms. To give definite answer to it is probably very difficult, but the following consideration seems to still be possible.

The condition (CA) is fairly general for linear evolution equations, for it is a kind of stability condition for the system. On the other hand, such nonlinear equations are often obtained by linearization of the originally nonlinear ones.

Therefore, it can be expected that fairly wide classes of nonlinear evolution equations satisfy (CA) at least locally. If it is true, the present procedure can be continued until (CA) is violated.

Next, the computational procedure becomes very complicated because the approximate equations are usually nonlinear if A is nonlinear. To avoid this difficulty, such techniques as linearization at each step (piecewise linear procedure), extrapolation and iteration can be employed. Newton-Raphson method and predictor-corrector method are such examples (Reference 25). As for their validity, much investigation will be necessary.

In nonlinear problems, sometimes discontinuous solutions arise even for sufficiently smooth initial values. Therefore, approximate methods should be able to treat such problems correctly. The experience of the finite difference methods shows that the concept of artificial viscosity is effective especially to such problems. That is, the instability due to nonlinear effect can be annealed by such damping terms (References 31 and 32).

In the numerical experiments given in the next section, a simple nonlinear shock wave problem is treated by use of the above-mentioned methods and their effectiveness is investigated.

VI. NUMERICAL EXPERIMENTS

In order to demonstrate the validity of the present method, some numerical experiments were performed for several evolution equations. Some of the typical results are shown in this section. Especially some weak solutions are treated numerically, though the theoretical justification is not given in this paper.

In the present experiments, only one dimensional (in space) problems are treated by use of the piecewise linear shape function with uniform mesh size Δx. The initial values are taken such that they coincide with the exact initial values at nodal points.

The equations treated here are heat equation, wave equation, first order wave equation, equation of coupled sound and heat flow, Schrodinger's equation and nonlinear hyperbolic equation. The last one is treated purely experimentally because (CA) is not generally satisfied.

The main points of experiments to be observed are dependence of the scheme on θ and A. It can be generally expected that the scheme for $\theta = 0.5$ gives the best results to smooth solutions and $\theta = 1.0$ gives most stable results. It is also to be recognized that the method of solving the equation remains the same even if A differs.

In the present calculations, almost all the problems are solved by single precision (32 bits on HITAC 5020E) arithmetic and linear simultaneous equations are solved by Gaussian elimination method for band matrices.

(a) heat equation : $$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0 \ , \ u(0,x) = \sin(\pi x) \quad (0 \leq x \leq 1) \qquad (80)$$

The exact solution is $\sin(\pi x)\exp(-\pi^2 t)$. In Fig. 3, the results of $u(t,0.5)$ for $\Delta x = 0.1$ and $\Delta t = 0.01$ are shown in the case $\theta = 0.0$, 0.5 and 1.0. As is seen from them, the approximate solution for $\theta = 0.0$ is unstable. On the other hand, the

approximate solutions for θ = 0.5 and 1.0 are stable and the results for θ = 0.5 is the best.

In Fig. 4, the results for θ = 0.0 are shown for Δx = 0.1 and several values of Δt. The theory given in V – 5. predicts that the scheme is stable if $\Delta t/\Delta x^2 \leqq 1/6$, and this is ascertained by this experiment. By the way, the present solution belongs to a special class; that is, the approximate solution is an eigenfunction of the approximation operator corresponding to the smallest eigenvalue and it coincides with the exact eigenfunction at nodes. Therefore, its convergence occurs even if the above-mentioned stability condition is not satisfied. However, this situation is violated fairly rapidly by roundoff errors, and the phenomena is well represented in this figure. If the calculation is performed by the single precision in the case of Δt = 0.01, the approximate solution becomes unstable in about 7 steps, while it becomes unstable in about 15 steps in the double precision. That is, the same principle as that of power method governs the present phenomena and higher modes become dominant after some steps as is shown in Fig. 5 in the case of Δt = 0.005. Such a situation is much severer if the initial value originally contains higher modes.

Table 1 shows the results for $\Delta t/\Delta x^2$ = 1/6 in. which the scheme is stable for all values of θ. It is seen from them that highly accurate result is obtained for θ = 1.0. Similar phenomena is observed in the case of the finite difference scheme (68) for $\Delta t/\Delta x^2$ = 1/6 (Reference 23), and the present result is approved because the finite element scheme coincides with the finite difference scheme (68) in this case. The high accuracy is due to the small local truncation error of this scheme, but such a situation is no longer expected if a source term exists. In general, the scheme for θ = 0.5 gives the best results.

(b) wave equation: $$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0 \ , \ u(0,x)=\sin(\pi x), \ \frac{\partial u}{\partial t} = 0 (0 \leqq x \leqq 1) \qquad (81)$$

The exact solution is $\sin(\pi x)\cos(\pi t)$. In this case, the scheme for (b-2) in V-1 is employed and the scheme for θ = 0.5 coincides with one-step β-scheme for $\beta = 1/4$ (Reference 3).

Figure 6 shows the results of u(t, 0.5) in the case of $\Delta t = \Delta x$ = 0.1. Again, the scheme is unstable for θ = 0.0 and best for θ = 0.5. For θ = 1.0, the result is highly contractive and much smaller value of Δt is needed to improve the result. On the other hand, the scheme for θ = 0.5 gives excellent result because of its conservation of total energy and of small local truncation error.

Figure 7 shows the effect of Δt for θ = 0.0, where Δx = 0.1 in all cases. The result can be improved and the scheme becomes stable gradually if the value of Δt decreases, but it does not seem that there exists a definite threshold value of Δt as in the case of heat equation.

(c) first order wave equation: $$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0 \ , \ u(0,x)=u_0(x) \ (-\infty < x < \infty) \qquad (82)$$

The exact solution is $u_0(x-t)$ and discontinuous solution arises if u_0 is discontinuous. In the following examples, the interval for numerical calculation is taken sufficiently broad so that it approximates the infinite interval.

Figure 8 shows some results of convergence of approximate solution for $\Delta t/\Delta x$ = 1.0. The initial value is so chosen that it approximates the step function with unit height. The procedure is unstable for θ = 0.0 as is predicted in V-5, and is stable for θ = 0.5 and 1.0. For θ = 0.5, overshoot and oscillation is observed in the

neighborhood of the wave front, and the height of the overshoot does not seem to vanish even if mesh sizes are decreased. For $\theta = 1.0$, such oscillation is not observed but the rise of the wave is not so sharp as that of $\theta = 0.5$. Figure 9 shows the effect of the value of θ to the same problem. As is seen from it, the oscillation vanishes if the value of θ is taken a little larger than 0.5.

On the other hand, the scheme for $\theta = 0.5$ gives best result for continuous solution as is shown in Fig. 10, where u_0 is given as follows:

$$u_0(x) = x(1-x) \quad \text{for } 0 \leq x \leq 1\ , \quad u_0(x) = 0 \quad \text{for } x<0 \text{ and } x>1$$

In general, the scheme for $\theta = 0.5$ should be used for smooth solutions, but sometimes greater values of θ should be used for not sufficiently smooth solutions.

(d) Schrodinger's equation: $$\frac{\partial u}{\partial t} + i\left(-\frac{\partial^2 u}{\partial x^2} + V(x)\right) u = 0 \tag{83}$$

Table 2 shows the results for a steady state solution, where $V(x) = 0.0$, and the initial value and the exact solution are $u(0,x) = \sin(\pi x)$ and $u(t,x) = \sin(\pi x) \exp(i\pi^2 t)$ respectively. As is seen from these results for $|u(t, 0.5)|^2$, the result for $\theta = 0.5$ is best, while the result for $\theta = 0.0$ is unstable and the result for $\theta = 1.0$ is highly contractive.

Figures 11 and 12 show the calculated profile ($|u(t,x)|^2$) of Gaussian wave packet by the schemes for $\theta = 0.5$ and 1.0. In this case, the scheme for $\theta = 0.0$ gives rapidly diverging solution. The initial value and the potential are given as follows:

$$u(0,x) = (2\pi y^2)^{-1/4} \exp\left(-\frac{(x-0.16)^2}{4y^2} + i\,50\pi x\right) \quad (0 < x < 0.64)$$

$$V(x) = 5000\pi^2 \qquad (0.32 \leq x \leq 0.384)$$

$$= 0 \qquad \text{(otherwise)}$$

where $y = 0.035$ and the values of u at both ends are 0. The mesh sizes are: $\Delta x = 0.004$, $\Delta t = \Delta x^2$. Figure 13 shows the change of the shape of the wave packet after bouncing back by potential and rigid wall. The tendency of the finite element solution for $\theta = 0.5$ is in good agreement with that of the finite difference solution by Goldberg et al. (Reference 30). On the other hand, the result for $\theta = 1.0$ is much more attractive again.

(e) equation fo coupled sound and heat flow: (e) in V-1

The initial value is given as follows:

$$u^1(0,x) = 1.0, \quad u^2(0,x) = 1/\sqrt{3}, \quad u^3(0,x) = 1/\sqrt{3} \quad \text{for } x < 0$$

$$u^1(0,x) = 0.0, \quad u^2(0,x) = 0.0\ , \quad u^3(0,x) = 0.0 \quad \text{for } x > 0$$

In this case, the speed of propagation of shock wave is 1.0 if $\sigma = 0.0$. This infinitesimal shock decays gradually due to the heat conduction as is shown in Fig. 14, where the calculation was done by the scheme for $\theta = 0.5$. Figure 15 shows the profile of the shock at t = 165, and the result is compared with the finite difference solution by Richtmyer and Morton (Reference 23). Their agreement is fairly good and it seems that the effect of coupling is well represented.

(f) nonlinear hyperbolic equation: $$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \epsilon \Delta x \frac{\partial^2 u}{\partial x^2} \quad (\epsilon \geq 0) \tag{84}$$

The last term in (84) is added as artificial viscosity and the equation is treated as a parabolic one if $\epsilon \neq 0.0$ (References 31 and 32). In this case, the stability of the type (32) is assured if u has compact support, because (49) can be proved. However, (34) is not generally expected.

In order to treat the nonlinearity, this equation is linearized in each step [t, t + Δt] by using U, the value of the approximate solution at t;

$$\frac{\partial u}{\partial t} + U \frac{\partial u}{\partial x} + u \frac{\partial U}{\partial x} - U \frac{\partial U}{\partial x} = \epsilon \Delta x \frac{\partial^2 u}{\partial x^2} \tag{85}$$

The problems treated here are propagation of shock wave and rarefaction wave where the initial values are taken as step functions with unit height.

Figure 16 shows the result of shock wave at t = 100 for several values of ϵ, where $\Delta t = 1.0$, $\Delta x = 1.0$, $\theta = 1.0$, and the shock speed is 0.5. It is seen that a fairly good result can be obtained for $\epsilon = 0.0$ in this case but better result can be obtained if the value of ϵ is suitably chosen. In this calculation, the shock wave calculated reaches its steady state after about 20 steps.

Figure 17 shows the rarefaction wave calculated by setting $\epsilon = 0.0$ and $\theta = 1.0$. In this case, the problem is not so serious as that of shock wave and a fairly good result is obtained without using the artificial viscosity.

VII. CONCLUDING REMARKS

Some consideration is made with respect to a simple finite element scheme for initial value problems, and a criterion for its convergence and stability is given. Its applicability is also demonstrated experimentally, and it is shown that fairly wide classes of evolution equations can be treated by a unified method. There are many things left to be done, but the functional analysis and the numerical experiments will be efficiently employed for such investigations.

REFERENCES

1. Zienkiewicz, O. C., Cheung, Y. K. The Finite Element Method in Structural and Continuum Mechanics, McGraw–Hill, 1967.

2. Oden, J. T. "A General Theory of Finite Elements: I. Topological Considerations; II. Applications", Int. J. Num. Meth. Engrg, I, 205–221 & 247–259, 1969.

3 Newmark, N. M. "A Method of Computation for Structural Dynamics", J. Eng. Mech. Div., Proc. ASCE, 85, 67–94, 1959.

4. Chan, S. P., Cox, H. L., Benfield, W. A. "Transient Analysis of Forced Vibrations and Complex Structural Mechanical Systems", J. Aeronaut. Soc., 66, 457–460, 1962.

5. Leech, J. W., Hsu, P. T., [illegible], E. W. "Stability of a Finite-Difference Method for Solving Matrix Equations", AIAA J., 3, 2172–2173, 1965.

6. Johnson, D. E. "A Proof of the Stability of the Houbolt Method", AIAA J., 4, 1450–1451, 1966.

7. Nickell, R. E. 'On the Stability of Approximation Operators in Problems of Structural Dynamics", Int. J. Solids Structures, 7, 301–319, 1971.

8. Visser, W. "A Finite Element-Method for the Determination of Non-Stationary Temperature Distribution and Thermal Deformations", Proc. Conf. Matrix. Meth. Struct. Mech., Wright-Patterson AFB, Ohio, 925–944, 1965.

9. Wilson, E. L., Nickell, R. E. "Application of the Finite Element Method to Heat Conduction Problems", Nucl. Engrg Des., 4, 276–286, 1966.

10. Douglas, J., Dupont, T. "Galerkin Methods for Parabolic Equations", SIAM J., Numer. Anal., 7, 575–626, 1970.

11. Descloux, J. "On the Numerical Integration of the Heat Equation", Num. Meth., 15, 371–381, 1970.

12. Zienkiewicz, O. C., Parekh, C. J. "Transient Field Problems; Two-dimensional Analysis by Isoparametric Finite Elements", Int. J. Num. Meth. Engrg, 2, 61–71, 1970.

13. Argyris, J. H., Scharpf, D. W. "Finite Elements in Time and Space". Nucl. Engrg. Des., 10, 456–464, 1969.

14. Yoshida, K. Functional Analysis, Springer, 1968.

15. Lions, J. L. Quelques Methodes de Resolution des problemes aux Limtes non Lineaires, Gauthier, 1969.

16. Komura, Y. "Nonlinear Semi-Groups in Hilbert Space", J. Math. Soc. Japan, 19, 493–507, 1967.

17. Kato, T. "Nonlinear Semigroups and Evolution Equations", J. Math. Soc. Japan, 19, 508–520, 1967.

18. Mikhlin, S. G. The Problem of the Minimum of a Quadratic Functional, Holden-Day, 1965.

19. Arantes Oliveira, E. R. "Theoretical Foundation of the Finite Element Method", Int. J. Solids Structures, 4, 929–952, 1968.

20. Caraso, A., Parter, S. V. "An Analysis of Boundary-Value Techniques' for Parabolic Problems", Math. Comp. 24, 315–340, 1970.

21. Kikuchi, F., Ando, Y. "A Finite Element Method for Friedrichs' Symmetric Positive Systems", a paper to be presented at 21st Japan Nat. Congr. Appl. Mech., Tokyo, 1971.

22. Varga, R. S. Matrix Iterative Analysis, Prentice-Hall, 1962.

23. Richtmyer, R. D., Morton, K. W. Difference Methods for Initial-Value Problems, Interscience, 1967.

24. Lax, P. D., Richtmyer, R. D. "Survey of the Stability of Linear Finite Difference Equation", CPAM, 9, 267–293, 1956.

25. Isaacson, E., Keller, H. B. Analysis of Numerical Methods, John Wiley & Sons, 1966.

26. Kreiss, H. O. "Uber Implizite Differenzmethoden fur partielle Differentialgleichungen", Num. Math., 5, 24–47, 1963.

27. Fujita, H. "Contribution to the Theory of Upper and Lower Bounds in Boundary Value Problems", J. Phys. Soc. Japan, 10, 1–8, 1955.

28. Mizohata, S. Theory of Partial Differential Equations (Japanese title: Hembibunhoteishiki-ron), Iwanami, 1965.

29. Oden, J. T., Kross, D. A. "Analysis of General Coupled Thermo-Elastic Problems by the Finite Element Method", Proc. 2nd Conf. Matrix Meth. Struct. Mech., Wright-Patterson AFB, Ohio, 1968.

30. Goldberg, A., Schey, H. M., Schwartz, J. L. "Computer Generated Motionpicture of One Dimensional Quantum Mechanical Transmission and Reflection", Am. J. Phys., 35, 177–186, 1967.

31. von Neumann, J., Richtmyer, R. D. "A Method for the Numerical Calculation of Hydrodynamic Shocks", J. Aopl. Phys., 21, 232–237, 1950.

32. Lax, P. D. "Weak Solution of Nonlinear Hyperbolic Equations and Their Numerical Computation", CPAM, 7, 159–193, 1954.

STEP θ	0.0	0.5	1.0	EXACT
1	9.834E−01	9.836E−01	9.837E−01	9.837E−01
5	9.198E−01	9.204E−01	9.210E−01	9.210E−01
10	8.460E−01	8.472E−01	8.483E−01	8.483E−01
50	4.334E−01	4.364E−01	4.394E−01	4.393E−01
100	1.878E−01	1.904E−01	1.930E−01	1.930E−01
150	8.138E−02	8.309E−02	8.481E−02	8.480E−02
200	3.527E−02	3.626E−02	3.726E−02	3.726E−02
250	1.528E−02	1.582E−02	1.637E−02	1.637E−02
300	6.623E−03	6.904E−03	7.192E−03	7.192E−03

$\Delta x = 1/10$

$\Delta t = 1/600$

Table 1: Comparison of u(0.5) for θ = 0.0, 0.5 and 1.0 —— heat equation ——

Δt	0.1			0.05		
θ	0.0	0.5	1.0	0.0	0.5	1.0
t = 0.1	1.99E+00	1.00E+00	5.02E−01	1.56E+00	1.00E+00	6.43E−01
0.5	9.71E+04	1.00E+00	3.20E−02	3.31E+19	1.00E+00	1.09E−01
1.0	3.42E+25	1.00E+00	1.03E−03		1.00E+00	1.20E−02
2.0	overflow	1.00E+00	1.05E−06	overflow	1.00E+00	1.44E−04
3.0		1.00E+00	1.08E−09		1.00E+00	1.72E−06
4.0		1.00E+00	1.11E−12		1.00E+00	2.07E−08
5.0		1.00E+00	1.13E−15		1.00E+00	2.48E−10
EXACT	1.0000					

$\Delta x = 0.1$

Table 2: Comparison of $|u(0.5)|^2$ for θ = 0.0, 0.5 and 1.0 —— Schrodinger's equation ——

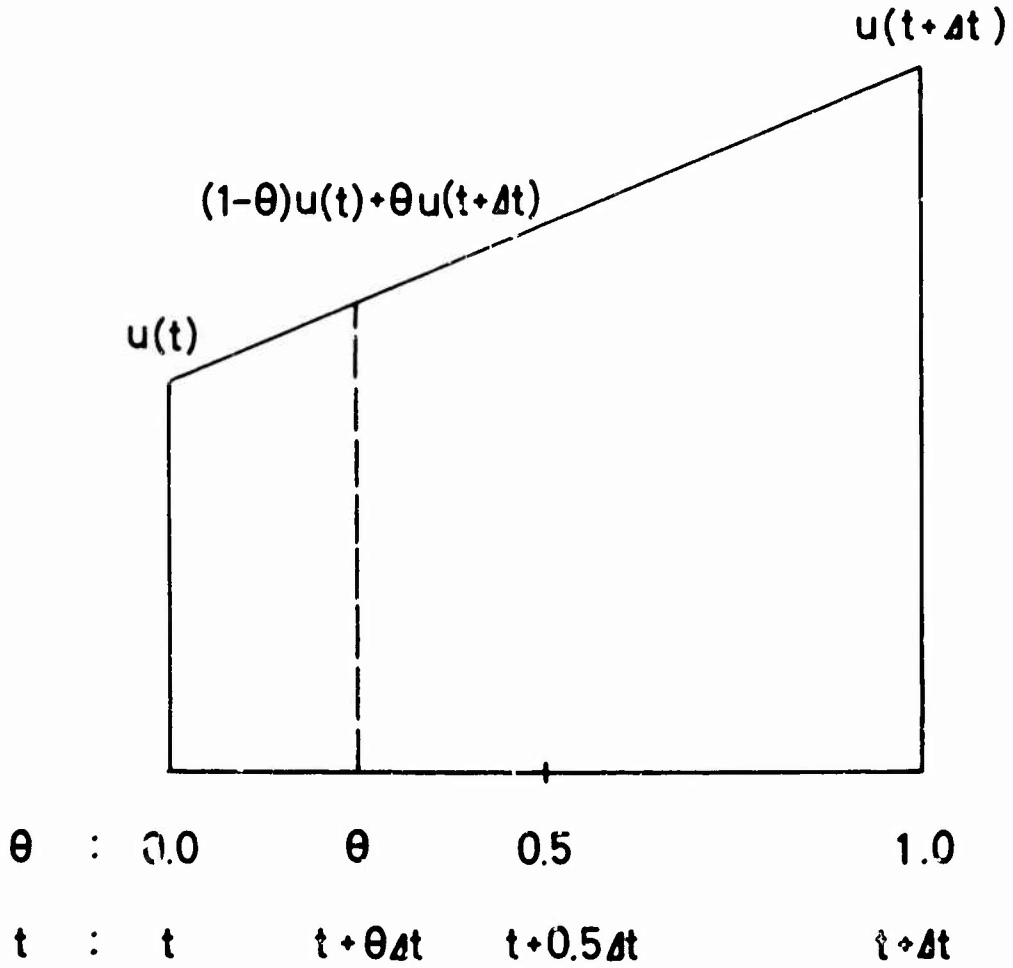


Figure 1. Parameter "θ"

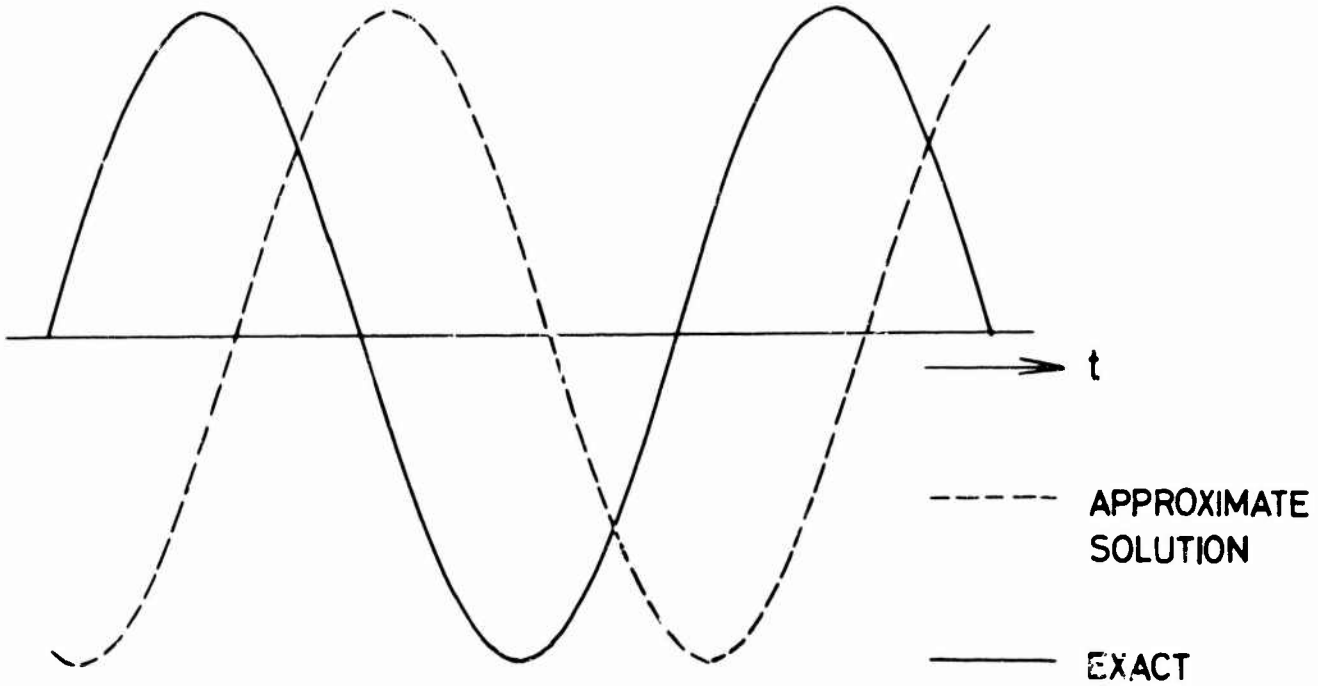


Figure 2. Behavior of an Approximate Solution When the Norm Is Conserved

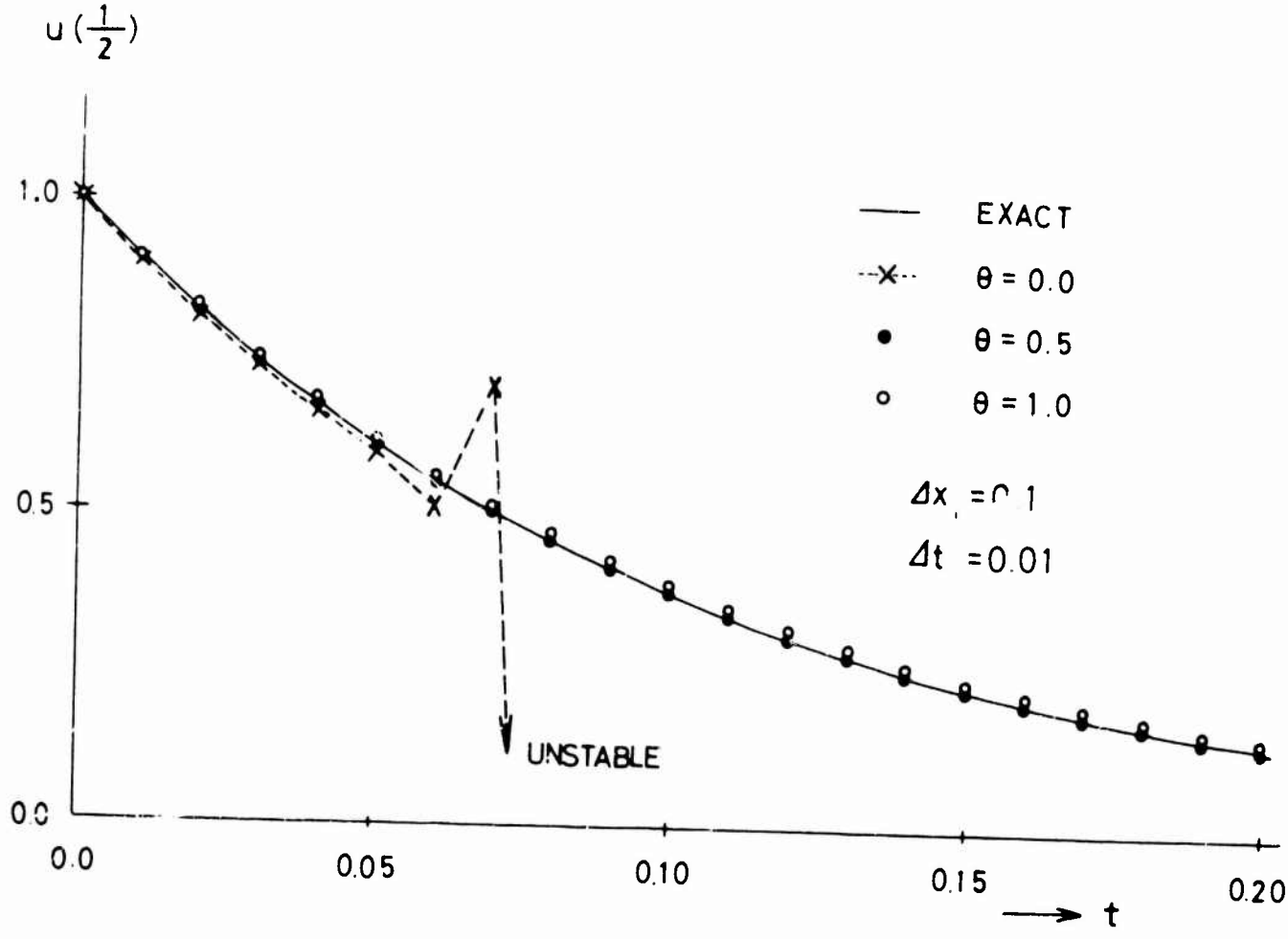


Figure 3. Approximate Solutions of a One-Dimensional Heat Flow Problem

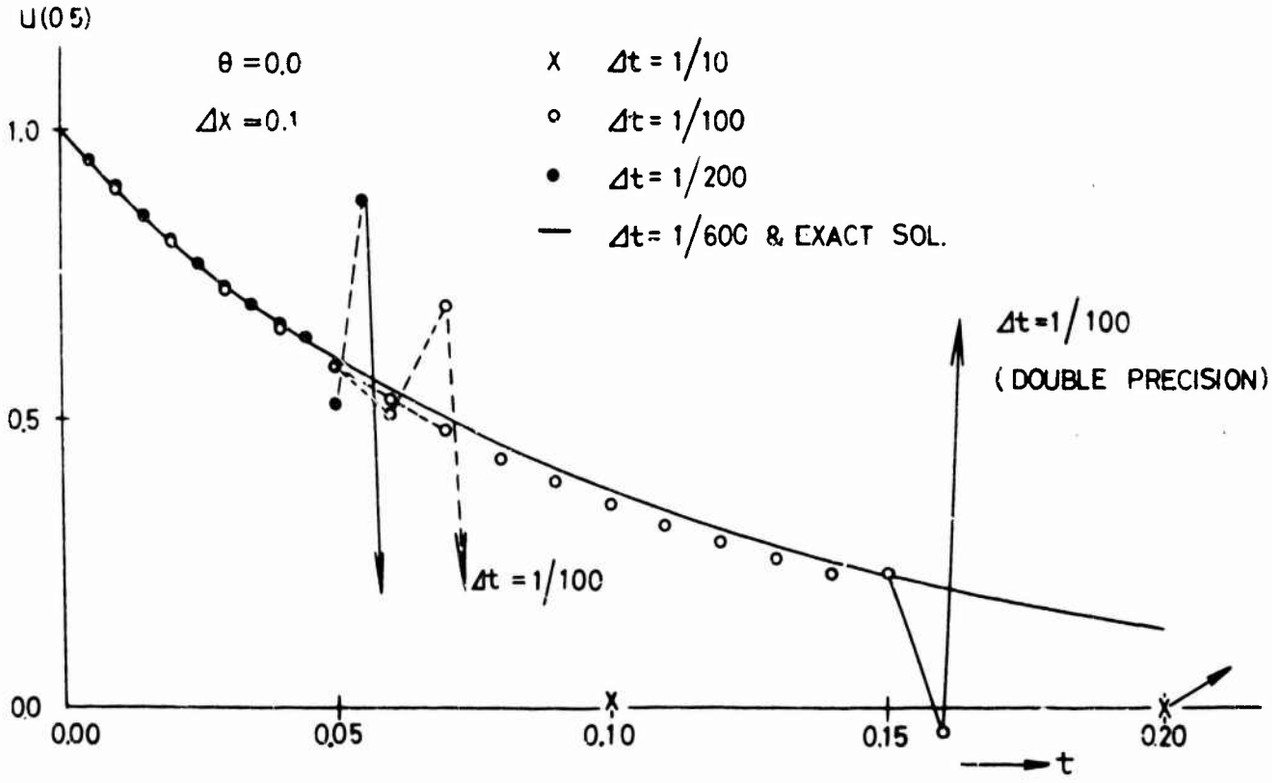
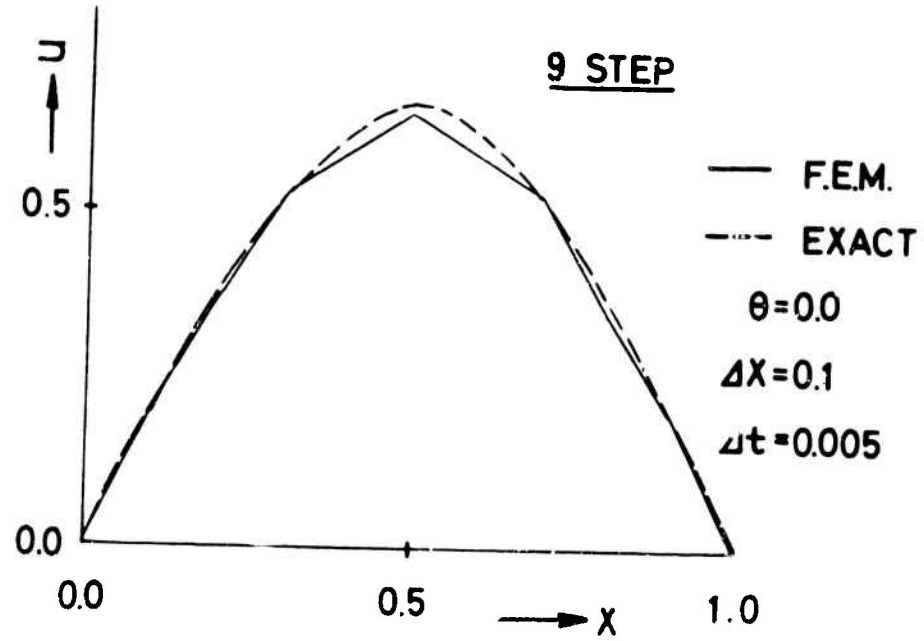


Figure 4. Approximate Solutions of a One-Dimensional Heat Flow Problem — θ = 0.0 —



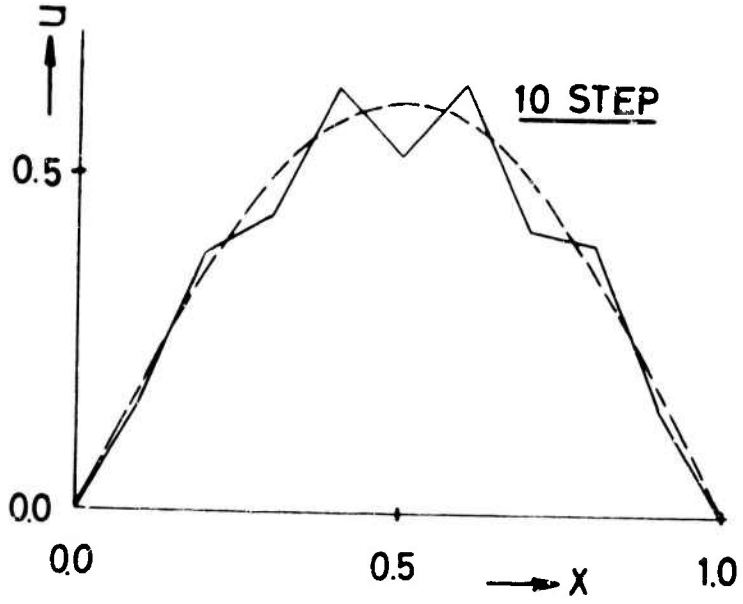


Figure 5. Instability Due to Mingling of Higher Modes — Heat Equation, θ = 0.0 —

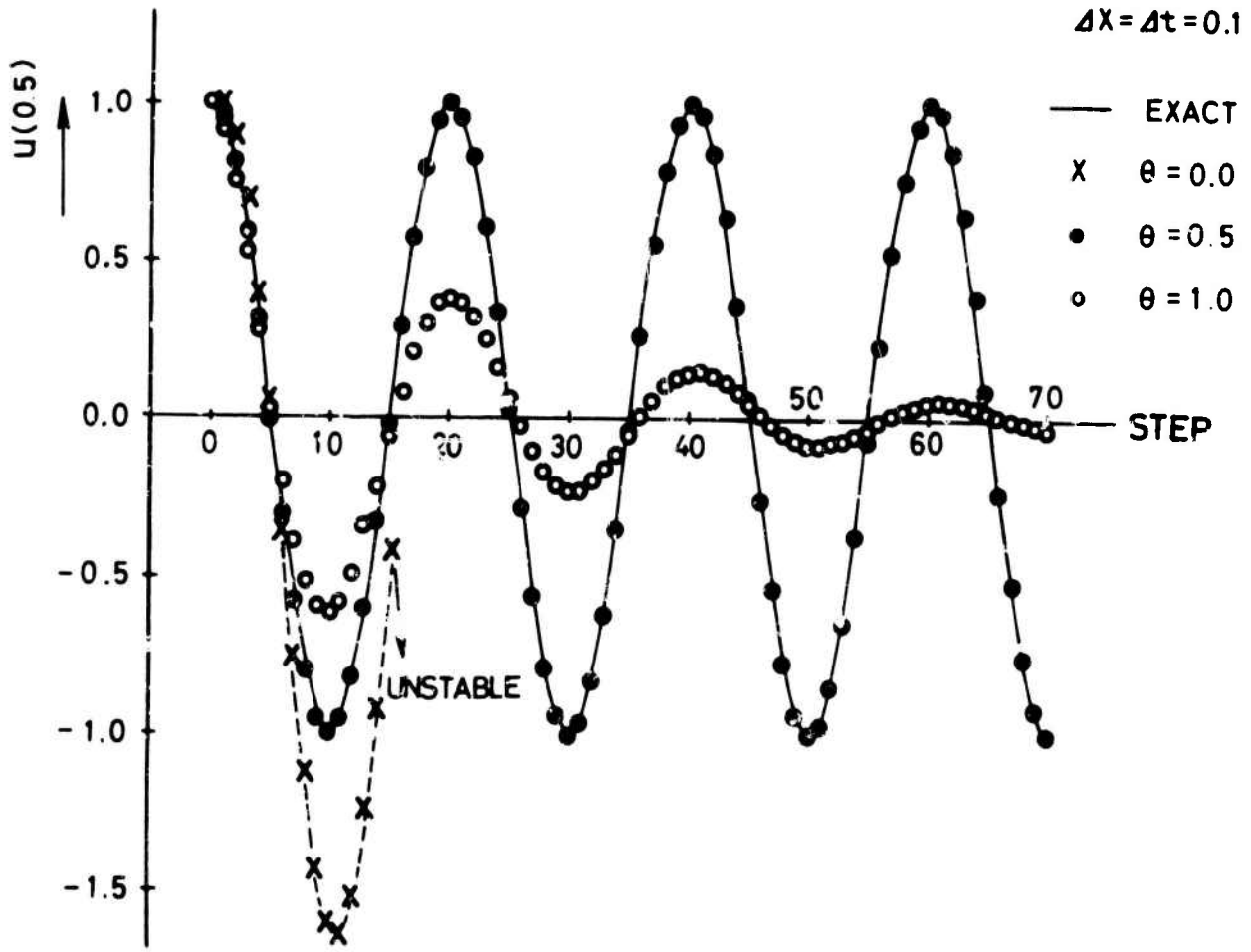


Figure 6. Approximate Solutions of a One-Dimensional Vibration Problem

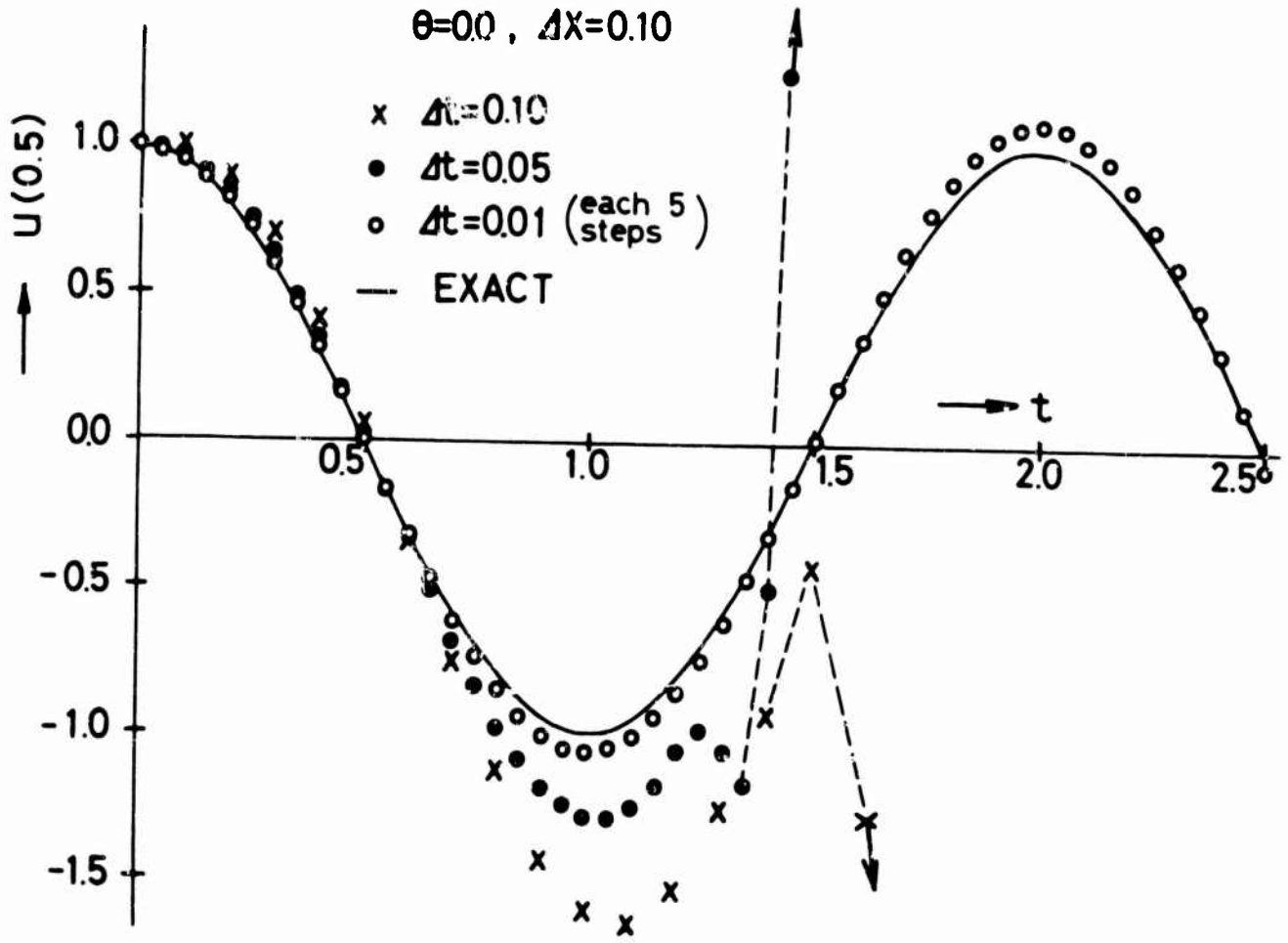


Figure 7. Approximate Solutions of a One-Dimensional Vibrational Problem —— θ = 0.0 ——

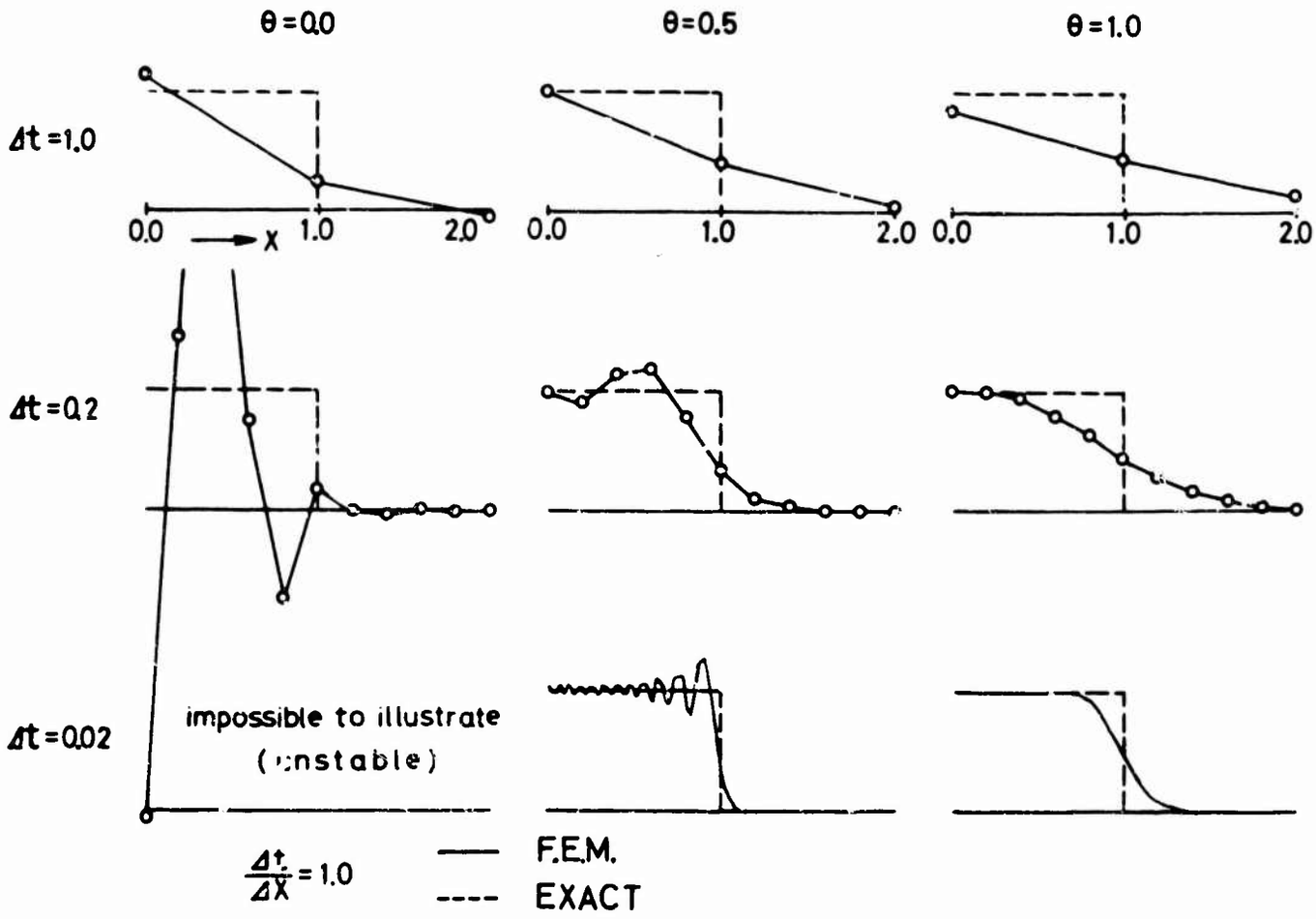


Figure 8. Propagation of a Discontinuous Wave — First Order Wave Equation —

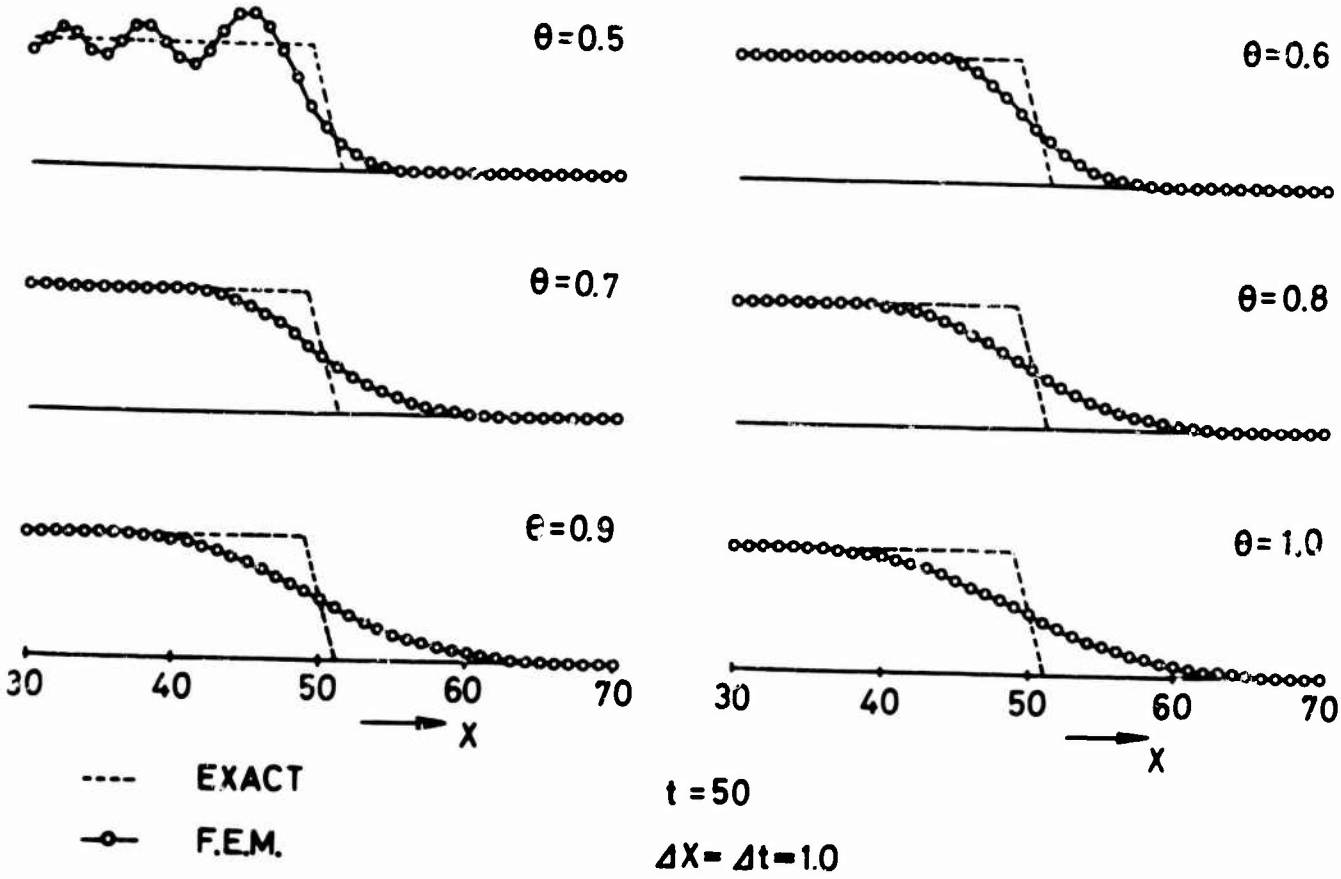


Figure 9. Effect of θ — First Order Wave Equation —

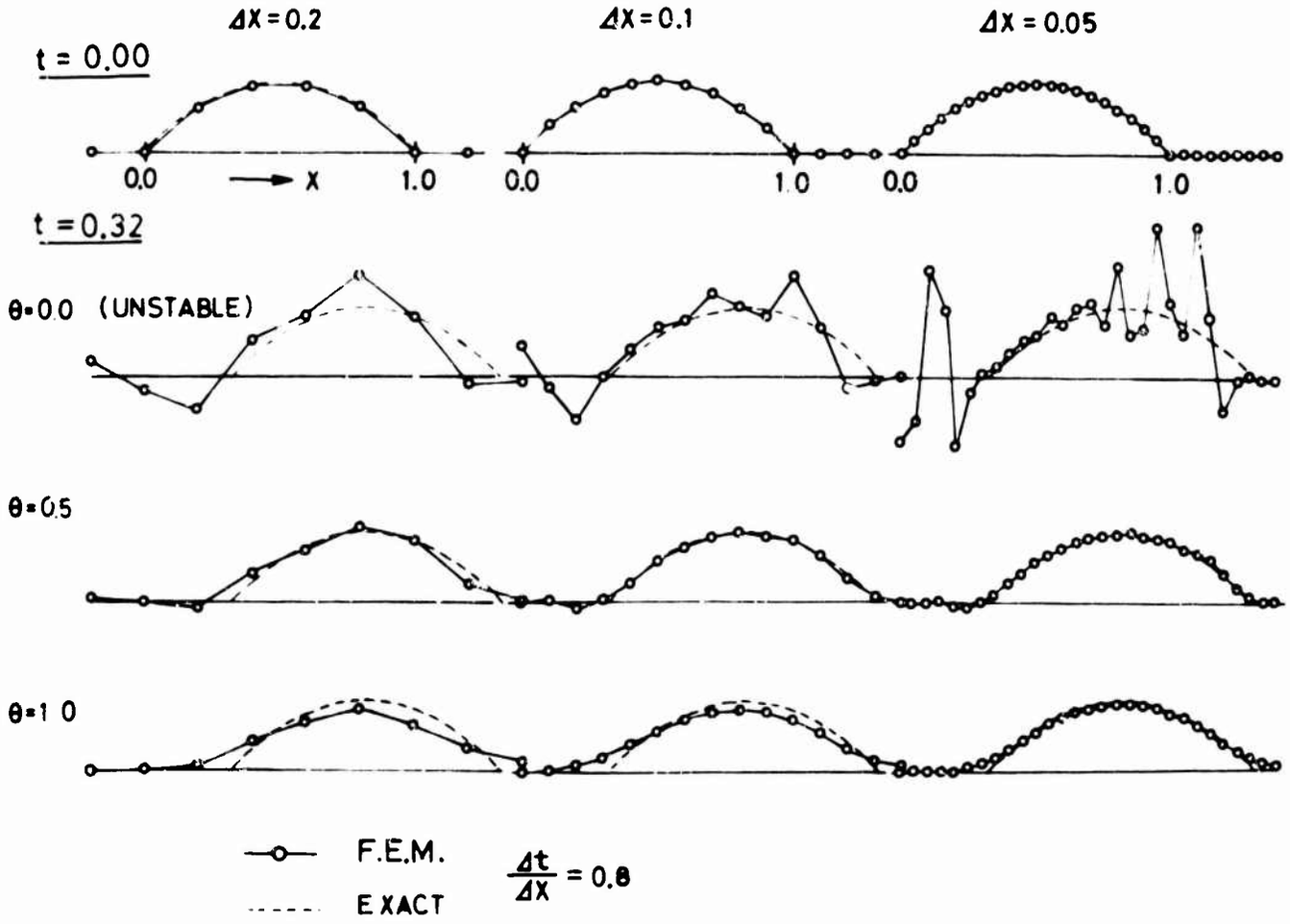


Figure 10. Propagation of a One-Dimensional Wave With Compact Support
— First Order Wave Equation —

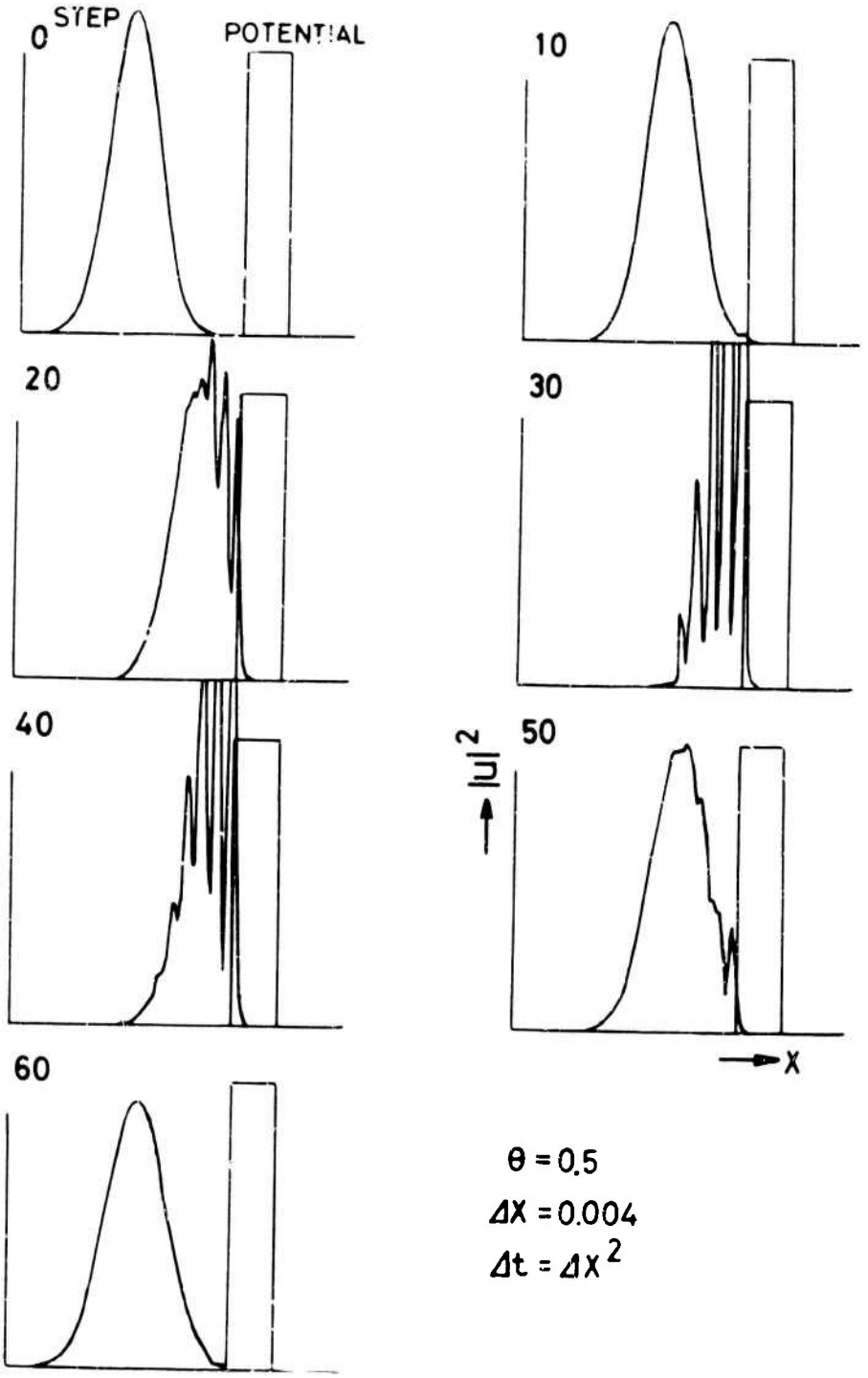


Figure 11. Gaussian Wave Packet Scattering From A Square Barrier — θ = 0.5 —

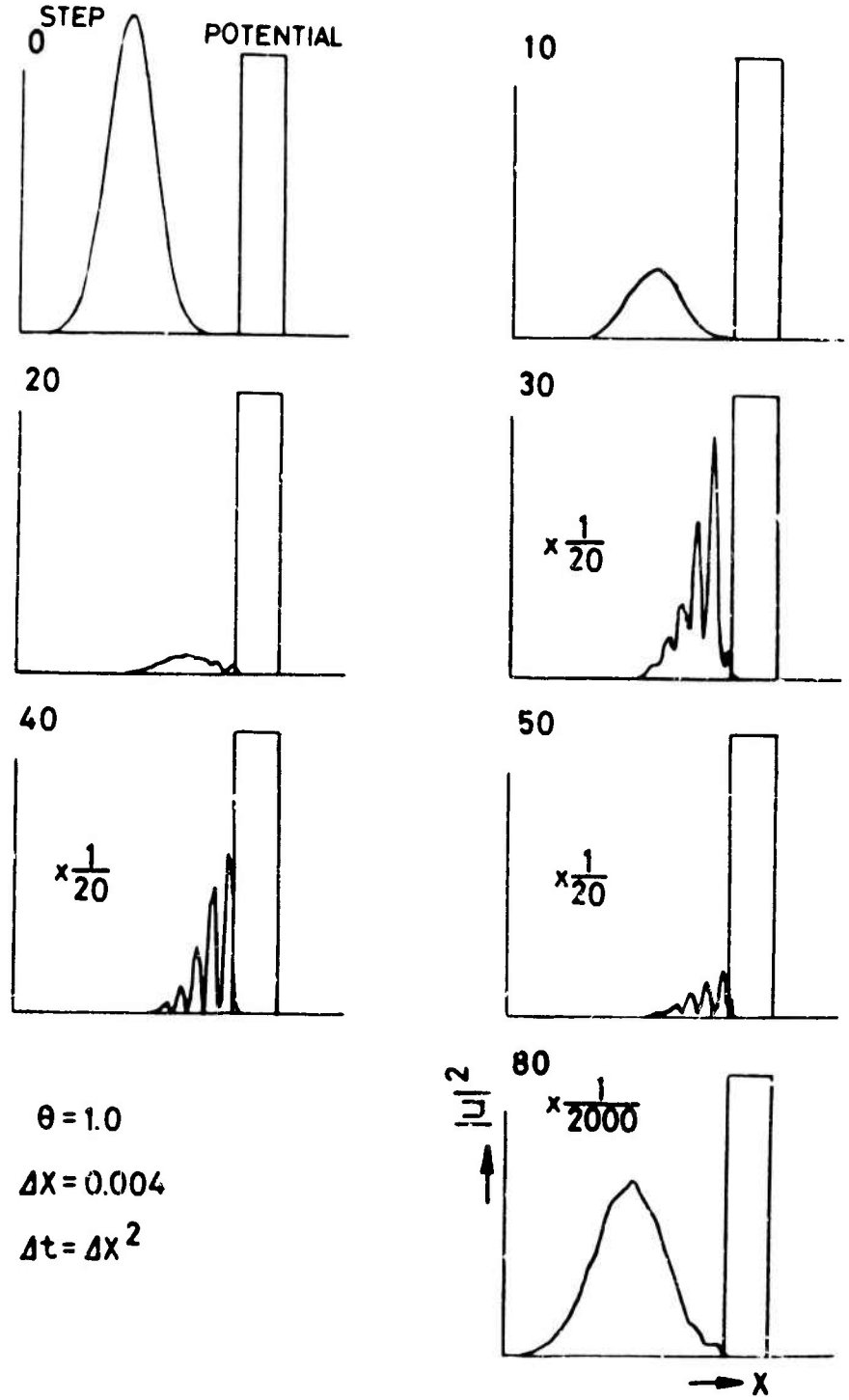


Figure 12. Gaussian Wave Packet Scattering From A Square Barrier — θ = 1.0 —

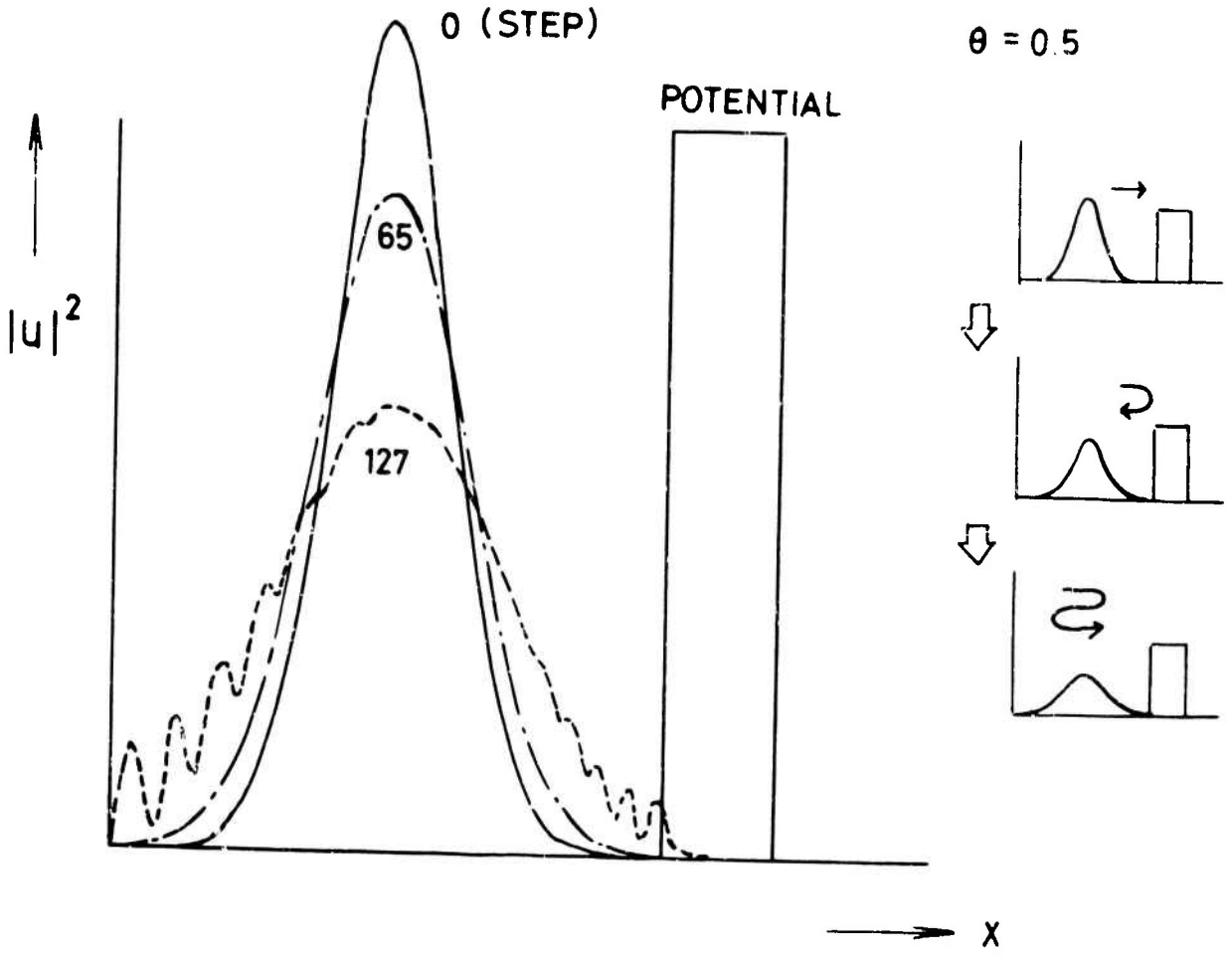


Figure 13. Gaussian Wave Packet Scattering from a Square Barrier and a Rigid Wall — θ = 0.5 —

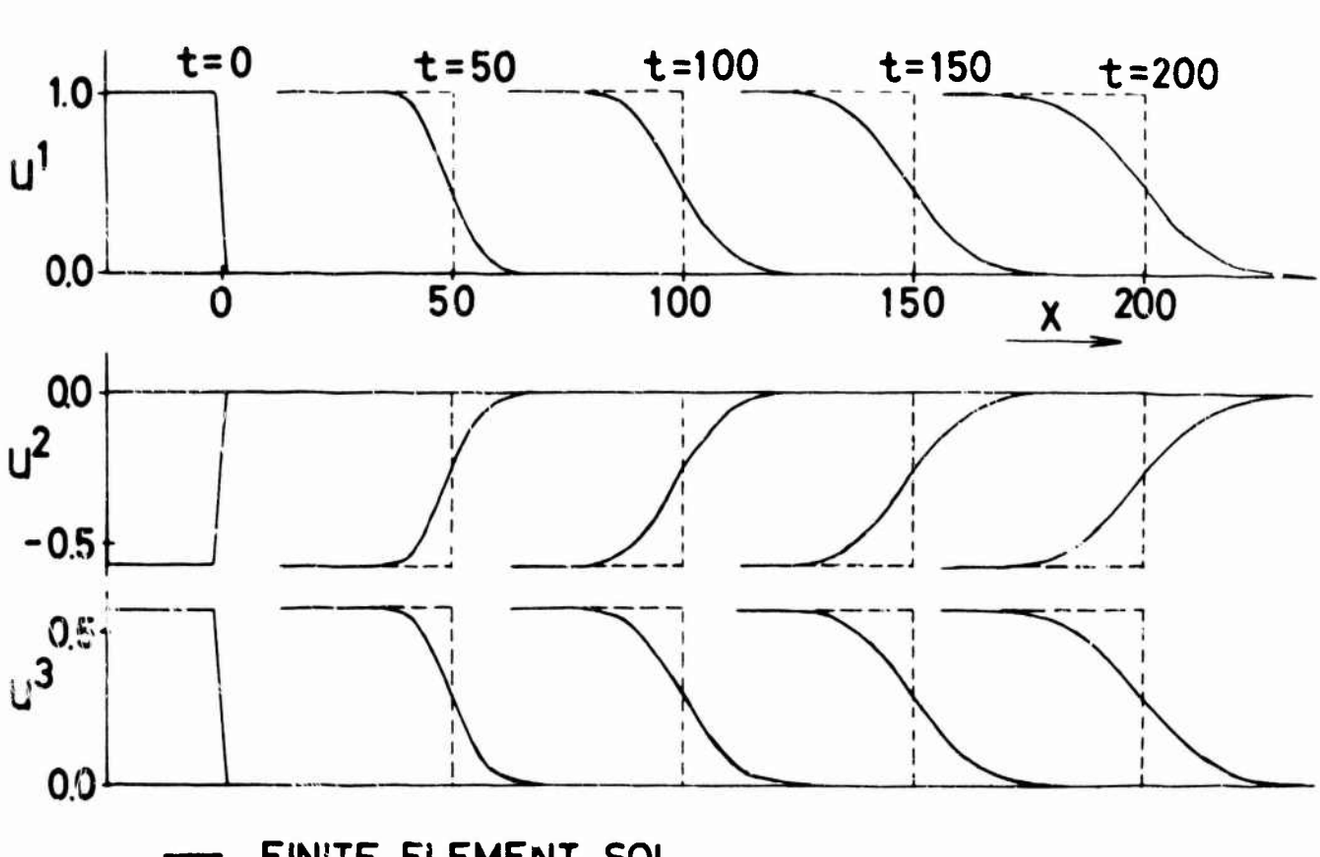


Figure 14. Calculated Profiles of Coupled Sound and Heat Flow — $\theta = 0.5$ —

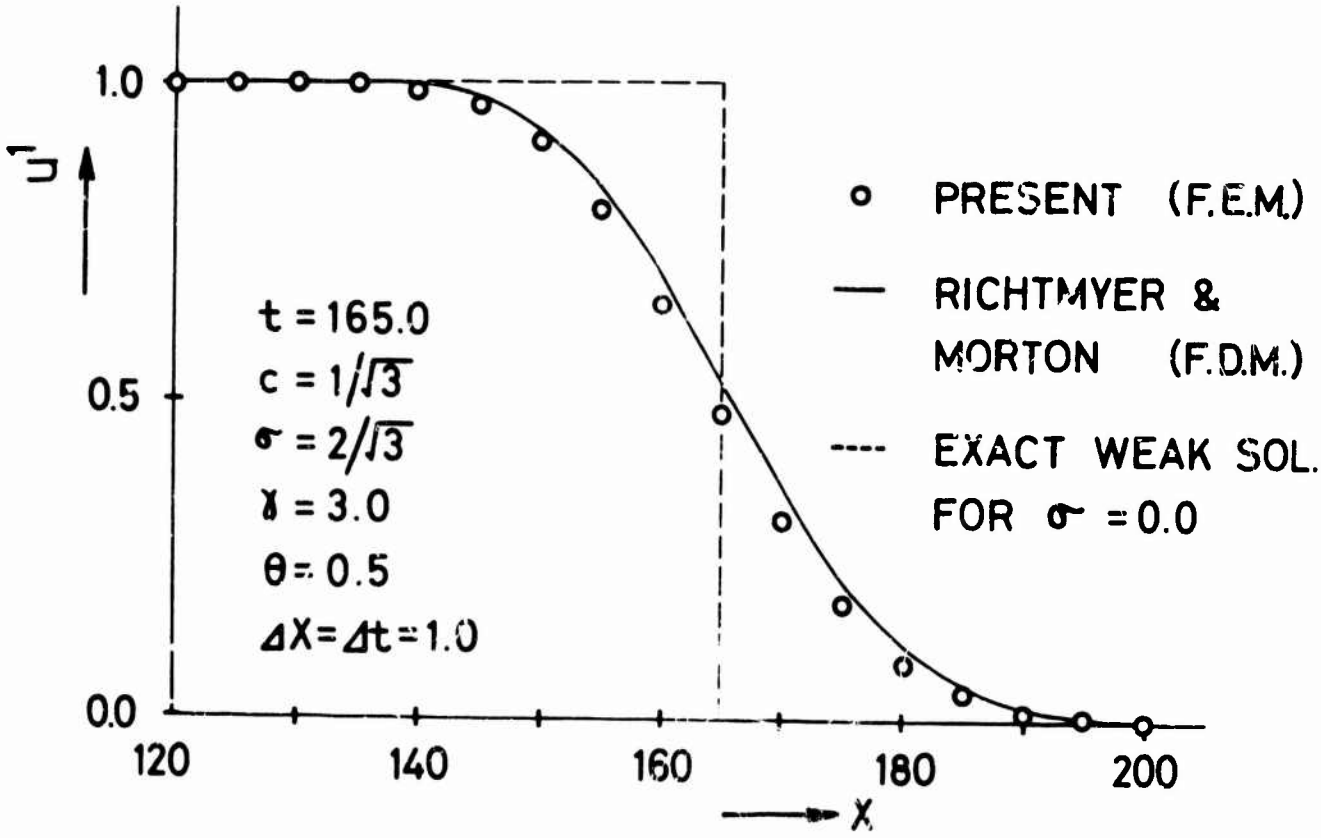


Figure 15. Calculated Profile of an Initially Sharp Sound Wave — θ = 0.5 —

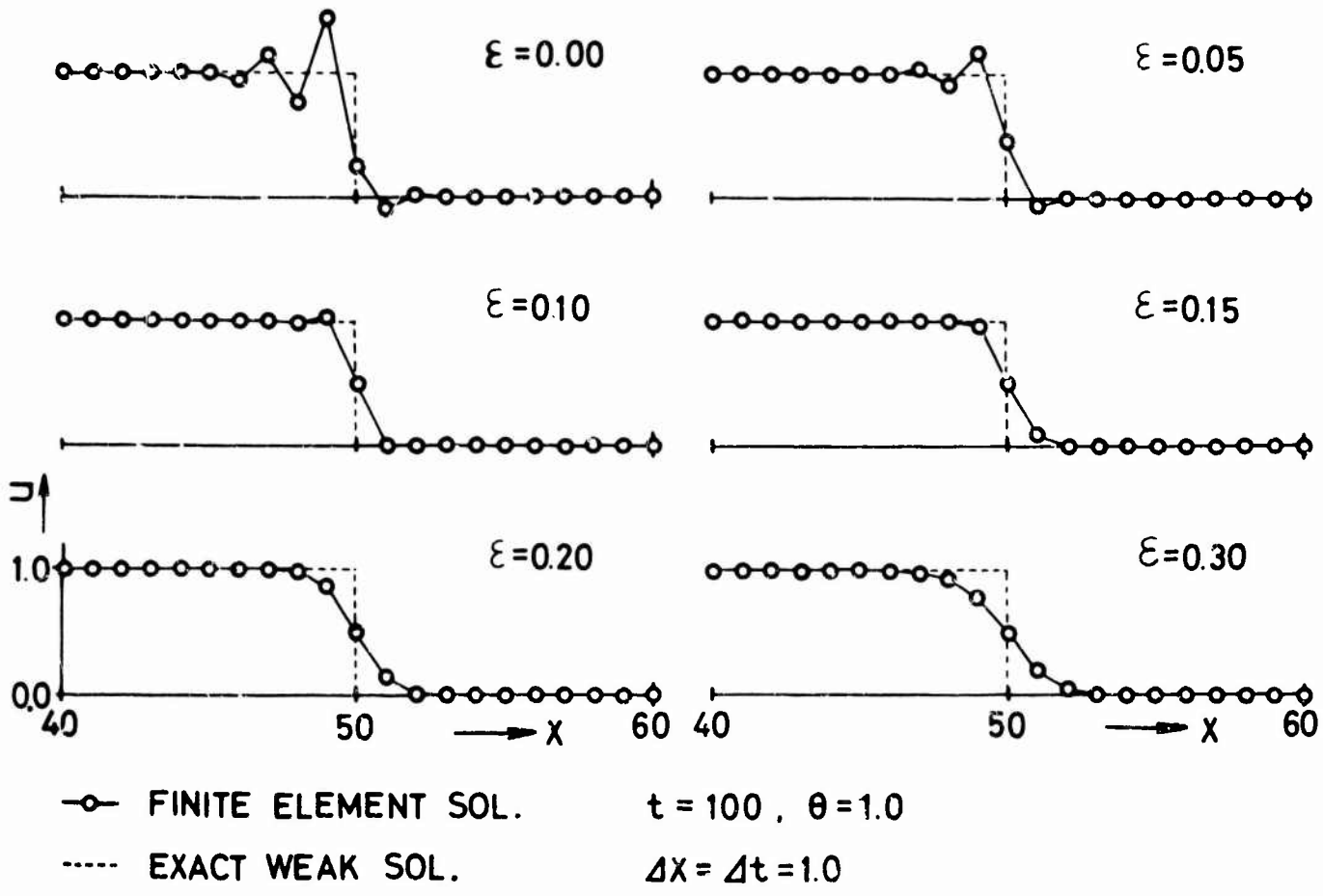


Figure 16. Calculated Profiles of a Nonlinear Shock Wave — θ = 1.0 —

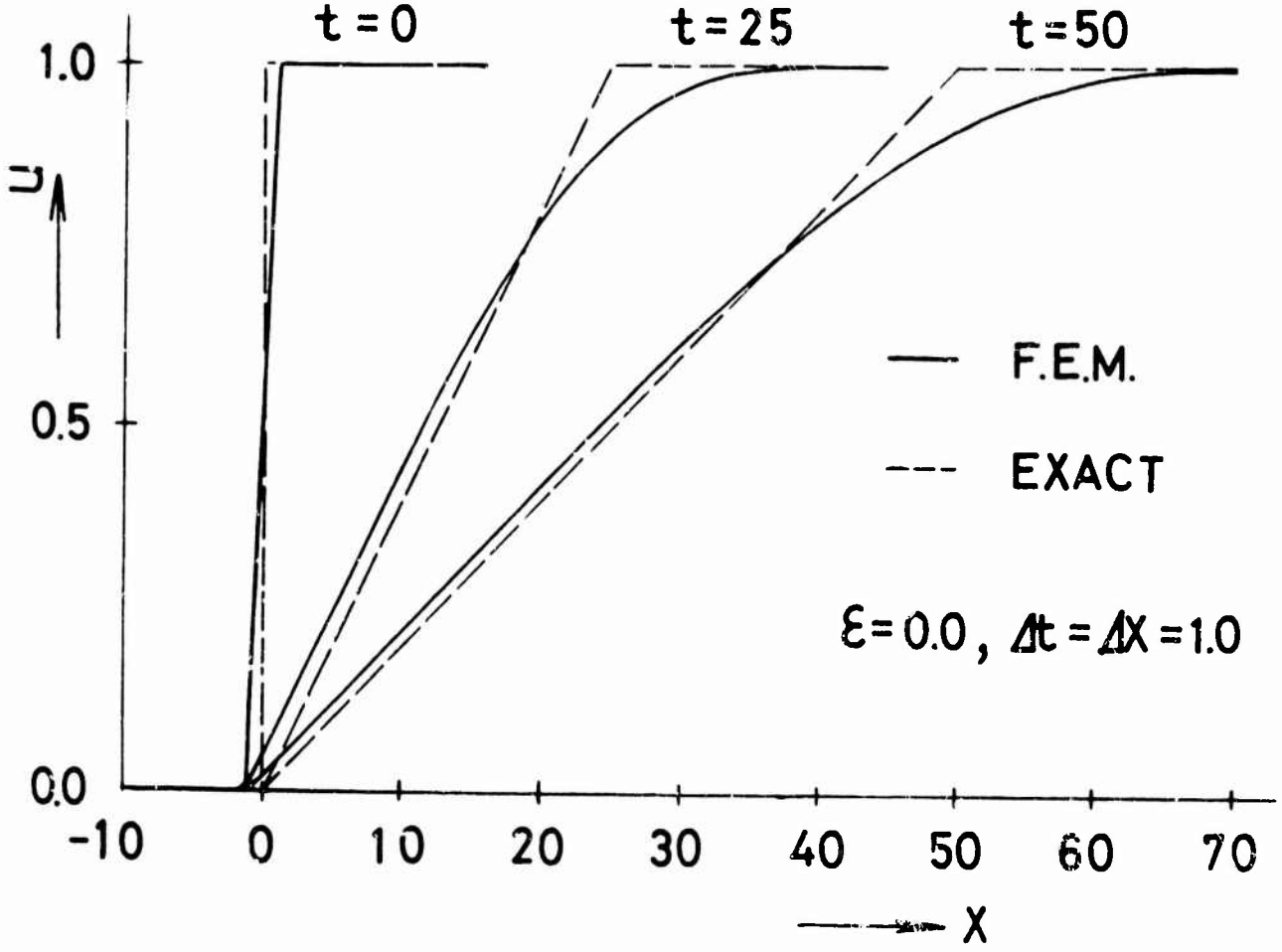


Figure 17. Calculated Profile of a Rarefaction Wave -- θ = 1.0 —

FINITE ELEMENT GALERKIN METHOD SOLUTIONS TO SELECTED ELLIPTIC AND PARABOLIC DIFFERENTIAL EQUATIONS

Mustafa M. Aral*, Paul G. Mayer**, and C. V. Smith, Jr.***
Georgia Institute of Technology
Atlanta, Georgia

In this paper a discussion of the use of variational analysis versus Galerkin formulation in developing integral finite element equations is presented. In view of this discussion finite element Galerkin type equations are developed for selected elliptic and parabolic equations. In developing these equations special emphasis is given to the inter element boundary conditions. Two examples are included in the area of fluid mechanics for illustrative purposes.

*Formerly, Graduate Research Assistant, School of Civil Engineering, Faculty member at the Mathematics Department, Middle East Technical University, Ankara, Turkey.
**Professor of Civil Engineering.
***Associate Professor of Aerospace Engineering.

SECTION I

INTRODUCTION

The finite element technique appears to have had its origin in the area of structural analysis, beginning, at least for most engineers in the U.S. and in the U.K. with the paper by Turner, Clough, Martin, and Topp[1][1]. For several years afterwards, the necessary stiffness or flexibility matrices were derived from basic concepts such as the principle of virtual work or Castiglianos's Theorem as applied to the individual finite elements. The so-called global or overall matrix was formed by using familiar ideas from the analysis of discrete structures such as trusses and frames.

A paper by Melosh[2] appears to have been one of the first to demonstrate that the finite element displacement method can be viewed as an application of the principle of stationary potential energy. After the finite element method was given this firm basis on energy principles, applications began to appear in areas other than solid mechanics and structural analysis [3], [4], [5], [6]. It was recognized that any problem whose solution could be associated with a stationary value of some functional could be solved by the finite element technique. Unfortunately, this seeming dependence on the existence of a variational principle led many investigators to formulate statements which were not, in reality, true variational principles [7], [8], [9]. Although these quasivariational principles and restricted variational principles give useful results, the fact remains that they do not possess a rigorous foundation in the classical calculus of variations. Therefore, it would appear that their use should be avoided whenever possible, especially when a sound alternative is available which gives exactly the same results.

There have been at least two earlier formulations of the finite element technique without recourse to a stationary principle. Oden and Kross [10] developed the equations of coupled thermoelasticity by means of what they described as energy balance equations for each element. Szabo and Lee [11] discussed the development of a plate bending element by Galerkin's method.

The procedure illustrated in this paper is the method of weighted residuals, primarily the Galerkin method, extended to systems in which the external and internal boundary continuity conditions are not identically satisfied by the trial functions. The procedure requires only a knowledge of the differential equations and boundary conditions which must be satisfied in a given domain and boundary. If a true, quasi, or restricted variational principle exists the Galerkin method can be made to give exactly the same result; and in the case of quasi or restricted principles, this result will have a more firm theoretical basis.

This paper presents formulations in two dimensional domains for the following differential equations with appropriate boundary conditions:

---

[1]Numbers in brackets refer to the list of references.

$$\frac{\partial}{\partial x}\left(K_x \frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial \phi}{\partial y}\right) = 0 \qquad (1)$$

$$\frac{\partial}{\partial x}\left(K_x \frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial \phi}{\partial y}\right) - K_t \frac{\partial \phi}{\partial t} = 0 \qquad (2)$$

$$\frac{\partial}{\partial x}\left(K_x \frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial \phi}{\partial y}\right) - u\frac{\partial \phi}{\partial x} - v\frac{\partial \phi}{\partial y} - \frac{\partial \phi}{\partial t} = 0 \qquad (3)$$

Equations (1) is an elliptic differential equation describing, for example, steady incompressible potential flow, or steady fluid seepage through a anisotropic, non-homogeneous medium, Zienkiewicz, Mayer, and Cheung[3]. This equation has also beer considered by Martin[5] through a true variational principle; it is presented again only to establish the correspondence between Galerkin's method and the variational principle.

Equation (2) is a parabolic differential equation which might represent unsteady heat conduction[4]. The finite element solution can be compared to exact solutions for simple geometries, thereby demonstrating the accuracy of the numerical integration scheme used to solve the differential equations in time.

Equation(3) is a parabolic equation which describes the unsteady diffusion of the concentration of a substance in a fluid flow of specified velocities(u,v).

SECTION II

TRUE, QUASI, AND RESTRICTED VARIATIONAL PRINCIPLES

This section presents a brief description of the meaning of the terms true, quasi, and restricted variational principles. The ideas are drawn very heavily from a most outstanding series of papers by Finlayson and Scriven [7], [8], [9], and anyone interested in a more thorough discussion is encouraged to read the indicated references.

It is appropriate to begin with the concept of a true variational principle as defined within the framework of the classical calculus of variations. First there must exist a functional which associates a definite real number with every function (or set of functions) belonging to the class of admissible comparison functions. Next it must be realized that if two admissible comparison functions are different, then their partial derivatives with respect to <u>all</u> independent variables will also be different. Finally, the variational principle states that if the functional is made stationary by the techniques of the calculus of variations, then the resulting Euler equations and natural boundary conditions will indeed describe the physical behavior of the system under consideration.

As an example of the three ideas associated with the true variational principle, consider the following functional:

$$J[\phi] = \int_0^L \frac{1}{2} K(x) \left[\frac{d\phi(x)}{dx} \right]^2 dx \qquad (4)$$

The class of admissible comparison functions might be specified as containing all functions defined on the closed interval $0 \leq x \leq L$ which are continuous have continuous first derivatives, and have specified values at $x = 0$ and $x = L$. If ϕ_1 and ϕ_2 are related by

$$\phi_1(x) = \phi_2(x) + \eta(x) \qquad (5)$$

where $\eta(x)$ is simply the difference between $\phi_1(x)$ and $\phi_2(x)$, then the derivatives with respect to the independent variable (x) will be related by

$$\frac{d\phi_1}{dx} = \frac{d\phi_2}{dx} + \frac{d\eta}{dx} \qquad (6)$$

This means that if $\eta(x)$ is not identically equal to zero, then $d\eta/dx$ cannot

be identically equal to zero. Finally, if this functional is made stationary, the resulting Euler equation will be

$$\frac{d}{dx}\left(K\frac{d\phi}{dx}\right) = 0 \tag{7}$$

which does describe, for example, one dimensional steady state heat transfer for a nonhomogeneous material.

If a physical system is supposedly described by a variational principle which lacks any of the three requirements of a true variational principle, then ipso facto, that principle can not be a true variational principle. In this case it is useful to have terminology which indicates that the so-called "principle" is not exactly what one might think.

As an example, suppose the problem is one dimensional steady state heat transfer for a nonhomogeneous material with temperature dependent conductivity. Further assume that the temperature has specified values at $x = 0$ and $x = L$. The differential equation is

$$\frac{d}{dx}\left[K(\phi,x)\frac{d\phi(x)}{dx}\right] = 0. \tag{8}$$

One could now define the quantity δJ as follows

$$\delta J \equiv \int_0^L \frac{d}{dx}\left[K(\phi,x)\frac{d\phi(x)}{dx}\right]\delta\phi(x)\,dx \tag{9}$$

and enunciate the "variational principle" that

> "Among all functions $\phi(x)$ which are defined on the closed interval $0 \leq x \leq L$, are continuous, have continuous second derivatives, and have specified values at $x = 0$ and $x = L$, the correct (exact) function ϕ will make $\delta J = 0$ for arbitrary $\delta\phi$."

This is clearly a correct statement, but not very profound and really not worthy of dignifying as a "variational principle". With this example, the point is that because k is a function of ϕ there is no functional $J(\phi)$ whose first variation has the form given in Equation (9). This type of formulation was called a quasi-variational principle by Finlayson and Scriven [7] because the first requirement of a true variational principle is obviously missing.

As another example, consider the problem of one dimensional transient diffusion described by the following differential equation, initial condition, and boundary conditions:

$$K \frac{\partial^2 \phi(x,t)}{\partial x^2} = \frac{\partial \phi(x,t)}{\partial t} \quad , \quad 0 < x < L \, , \, t > 0 \qquad (10)$$

$$\phi(x,0) = 0 \quad , \quad 0 \leq x \leq L \, , \quad t = 0 \qquad (11)$$

$$\phi(0,t) = \phi_0 (1 - e^{-\alpha t}) \, , \quad x = 0 \, , \quad t > 0 \qquad (12)$$

$$\phi(L,t) = \phi_L (1 - e^{-\alpha t}) \, , \quad x = L \, , \quad t > 0 \qquad (13)$$

The boundary conditions have been taken as shown in order to provide continuity with the initial condition at t = 0, x = 0, and t = 0, x = L. If a functional J is defined as follows

$$J[\phi, \psi] = \int_0^t \int_0^L \left[\frac{1}{2} K \left(\frac{\partial \phi}{\partial x} \right)^2 - \psi \phi \right] dx \, dt \qquad (14)$$

then one can state the "variational principle" that

> "Among all functions $\phi(x,t)$ which are defined in the domain $0 \leq x \leq L$ and $t \geq 0$, are continuous, have continuous first derivatives with respect to x, and satisfy Equations (11) through (13), the correct function ϕ will make $\delta J = 0$ for arbitrary $\delta\phi$ and fixed ψ subject to the condition that $\psi(x,t) = \partial\phi(x,t)/\partial t$ after the variation if performed."

In this example there is a functional (for a given value of t), and a superficial resemblance to a stationary condition leading to an Euler equation which does indeed describe the physical behavior of the system. However, the variational process is not carried out according to the classical calculus of variations, with the result that the functional is not stationary for the exact ϕ. While the function ϕ is allowed to vary, the derivative $\partial\phi/\partial t$ is held fixed. This rather unclassical maneuver is accomplished by introducing the "alias variable" [7] $\psi(x,t)$ as an independent function during the variation and not identifying ψ as equal to $\partial\phi/\partial t$ until after the variation. This type of variational principle has been called a restricted variational principle because of the restriction imposed upon the allowable variation.

Although there are demonstrated deficiencies in rigor associated with both quasi and restricted variational principles, these formulations do

give useful results, particularly in the area of developing approximate solutions for difficult problems for which no true variational principle exists. For this reason, any analyst would be justifiably inclined toward using these principles, especially if no other procedure were available. However, the fact is that another technique is available. As shown by Finlayson and Scriven [7] through the consideration of many examples, a straightforward application of the general method of weighted residuals (or, in particular, Galerkin's method) is completely equivalent to the application of a quasi or restricted principle. Furthermore, the method of weighted residuals provides the answer without all the internal inconsistencies which are intrinsic to quasi and restricted variational principles. Using the closing words of Reference [7];

> "When approximate solutions are in order the applied scientist and engineer are better advised to turn immediately to direct approximation methods for their problems, rather than search for or try to understand quasi-variational formulations and restricted variational principles."

SECTION III

FINITE ELEMENT GALERKIN METHOD

In the preceding section it was shown that the method of weighted residuals should be preferred over quasi or restricted variational principles. However, the conventional application of the method of weighted residuals involves trial functions which are continuous over the entire domain, resulting in a formulation not strictly applicable to the finite element technique in which discontinuities may exist. There have been many papers in the past which have successfullly considered the problem of developing approximate solutions with certain discontinuities in the trial functions [12], [13]. However, the attention has been given to problems for which a true variational principle exists; and the previous papers have been primarily directed toward demonstrating how to modify the variational principle so as to expand the class of admissible functions to include functions with the specified discontinuities.

Reference [14] shows that the method of weighed residuals can also be generalized to allow for discontinuities in the assumed solution functions. This then completely frees the finite element method from any dependence on a variational principle and allows the rigorous use of the finite element concepts for problems with no true variational principle.

For a simple example of Galerkin's method extended to allow for discontinuities, consider the following problem:

$$\text{for} \quad 0 < x < L \,, \quad \frac{d^2\phi}{dx^2} = 0 \tag{15}$$

$$\text{at} \quad x = 0 \,, \quad \phi = \phi_0 \tag{16}$$

$$\text{at} \quad x = L \,, \quad \frac{d\phi}{dx} = q_L \tag{17}$$

where ϕ_o and q_1 are specified constants. This set of equations could describe one dimensional steady state heat conduction based on the Fourier law of conduction for a homogeneous material with temperature independent conductivity. At x = 0, the temperature is specified; at x = L, the heat flux is specified.

If this problem is to be solved by the finite element Galerkin method (certainly not recommended in this case), then it is necessary to recognize that the final solution must be continuous and posses continuous first derivatives in the domain. These additional requirements are established either from a knowledge of the physics of the problem or from a knowledge of the character of solutions for the given differential equation. Therefore,

at the i^{th} internal node, with location denoted by x_i, it will be necessary to require

$$\phi^{i+} - \phi^{i-} = 0 \tag{18}$$

$$\left.\frac{d\phi}{dx}\right|^{i+} - \left.\frac{d\phi}{dx}\right|^{i-} = 0 \tag{19}$$

where

$$\phi^{i+} = \phi(x_i^+) \quad , \quad \phi^{i-} = \phi(x_i^-) \tag{20}$$

and

$$x_i^+ = x_i + \epsilon \quad , \quad x_i^- = x_i - \epsilon \quad , \quad 0 < \epsilon << 1 \tag{21}$$

After dividing the domain into N subdomains (or finite elements), the next step in this example solution is to assume ϕ in the form

$$\phi(x) = \sum_{r=1}^{N} {}_r\phi(x) \tag{22}$$

where

$${}_r\phi(x) = \begin{cases} \sum_i {}_ra_i \, {}_r\phi_i(x) \ , & \text{if } x \text{ is in the } r^{th} \text{ finite element} \\ 0 \ , & \text{if } x \text{ is not in the } r^{th} \text{ finite element} \end{cases} \tag{23}$$

In Equation (22), the summation over r represents a summation over the N finite elements which make up the domain. In Equation (23), the ${}_r\phi_i$ are completely specified functions of x, and the ${}_ra_i$ represents the as-yet unknown generalized coordinates which are to be determined by the solution technique.

It will be assumed in what follows that by proper choice of the functions ${}_r\phi_i$ and by proper identification of the ${}_ra_i$ it is possible to identically satisfy Equations (16) and (18). Then, according to Reference [14], the complete finite element Galerkin method equation will be

$$\sum_{r=1}^{N} \int_{rL} \left[\sum_i {}_ra_i \frac{d^2\,{}_r\phi_i}{dx^2} \right] \left[\sum_j \delta_r a_j \, {}_r\phi_j \right] dx$$

$$- \left\{ \left[\sum_i {}_Na_i \frac{d_n\phi_i}{dx}\bigg|_{x=L} \right] - q_L \right\} \left\{ \sum_j \delta_N a_j \, {}_N\phi_j(L) \right\} \qquad (24)$$

$$+ \sum_{n=1}^{N-1} \left[\sum_i {}_pa_i \frac{d_p\phi_i}{dx}\bigg|_{x_n^+} - \sum_i {}_ma_i \frac{d_m a_i}{dx}\bigg|_{x_n^-} \right] \left[\sum_j \delta_p a_j \, {}_p\phi_j(x_n^+) = \sum_j \delta_m a_j \, {}_m\phi_j(x_n^-) \right] = 0$$

The first term represents a summation over all the finite elements, with the r^{th} element having length ${}_rL$. The $\delta_r a_i$ are variations of those generalized coordinates which are unspecified. The third term represents a summation over all interior nodes; the left subscript p represents the "plus" side of the node, and the left subscript m represents the "minus" side of the node. It is also noted that

$$\sum_j \delta_p a_j \, {}_p\phi_j(x_n^+) = \sum_j \delta_m a_j \, {}_m\phi_j(x_n^-) \qquad (25)$$

since, by assumption, Equation (18) is satisfied at each interior node.

If desired, the solution could proceed directly from Equation (24) by requiring that the equation be satisfied for arbitrary values of the $\delta_r a_i$. However, for this particular problem, the equation can be put in a more simple form by integrating the first term by parts. The results will be

$$\sum_{r=1}^{N} \int_{rL} \left[\sum_i {}_ra_i \frac{d_r\phi_i}{dx} \right] \left[\sum_j \delta_r a_j \frac{d_r\phi_j}{dx} \right] dx - q_L \left[\sum_j \delta_N a_j \, {}_N\phi_j(L) \right] = 0 \qquad (26)$$

Equation (26) is immediately recognized as the true first variation of the familiar functional

$$V = \frac{1}{2}\int_L \left(\frac{d\phi}{dx}\right)^2 dx - q_L \phi(L) \qquad (27)$$

when ϕ is specified by Equation (22) and Equations (16) and (18) are satisfied. Since a true variational principle does exist for this problem, there is really no reason to introduce the Galerkin method formulation. The point is, however, that for problems with no true variational principle, the Galerkin procedure (or, more generally, the method of weighted residuals as presented in reference [14]) still remains rigorously valid.

The principle idea which has been illustrated with this example is the need for explicitly including interior boundary continuity equations in the Galerkin formulation. In other words, it is not enough to develop a solution which approximately satisfies only the differential equation in each finite element and the exterior boundary conditions. Also certain conditions must be satisfied across the boundaries between adjacent elements, and these conditions should be included in the approximation formula.

Suppose that the ${}_r\phi_i$ functions were chosen as linear functions of x, which is exactly the type of approximation utilized for the more complicated examples presented later in this paper. Then the differential equation is satisfied identically in each element; and if Equation (24) is written without the interior node terms, there is no way to establish the values of the ${}_ra_i$. However, the correct form of the Galerkin method equation, as given in Equation (24) or equivalently in Equation (26), does provide a set of equations which can be used to evaluate ${}_ra_i$.

SECTION IV

BASIC ANALYSIS

For each of the differential equations presented in Equations (1) through (3), the two dimensional domain is divided into triangular subdomains. For each subdomain the origin of the local axes which defines the local coordinates of the element nodes is placed at the centroid of the element and the principle axes are inclined in the direction of local anisotropy. Considering the triangular element shown in Figure 1, the variation of the dependent variable ϕ in the r^{th} element is approximated by a linear polynimial of the form,

$$ {}_r\phi = \sum_i {}_rN_i \, \phi_i \tag{28} $$

where

$$ {}_rN_i = {}_ra_i + {}_rb_i \, {}_rx + {}_rc_i \, {}_ry \tag{29} $$

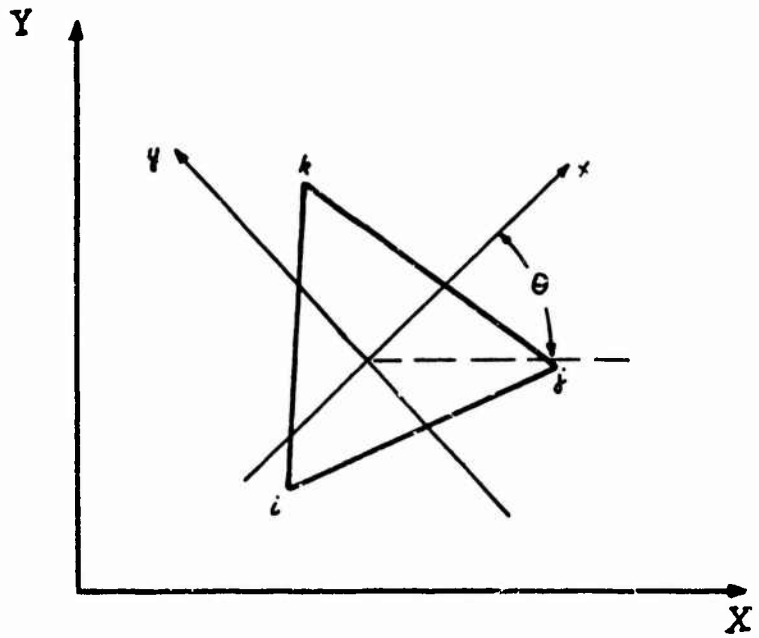


Figure 1. A Triangular Element

and

$$ {}_r a_i = ({}_r x_j \, {}_r y_k - {}_r x_k \, {}_r y_j) \frac{1}{2 \, {}_r A} \tag{30} $$

$$ {}_r b_i = ({}_r y_j - {}_r y_k) \frac{1}{2 \, {}_r A} \tag{31} $$

$$ {}_r c_i = ({}_r x_k - {}_r x_j) \frac{1}{2 \, {}_r A} \tag{32} $$

where ${}_r A$ is the area of the r^{th} triangular element (see reference [15] for details). Similarly a linear variation over each subdomain is assumed for the weighting functions, with a modification of the weighting function for Equation (3), Aral [16]. Then the method of weighted residuals is applied to the subdivided domain, paying particular attention to the continuity requirements which must be satisfied across the boundaries between the elements. The result is a set of algebraic equations in the case of Equation (1), or a set of ordinary differential equations in time, in the case of Equations (2) and (3), which are then solved for the system response.

First consider Equation (1). The complete problem is assumed as follows:

$$ \frac{\partial}{\partial x} \left(K_x \frac{\partial \phi}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_y \frac{\partial \phi}{\partial y} \right) = 0 , \quad \text{in the domain } A \tag{33} $$

$$ \phi = \bar{\phi}(s) , \quad \text{on the boundary denoted by } S_D \tag{34} $$

$$ K_n \frac{d\phi}{dn} = \bar{q}(s) , \quad \text{on the boundary denoted by } S_N \tag{35} $$

Also, across any internal boundaries, there must be continuity in values of ϕ and values of $K \, d\phi/dn$. In Equation (34), $\bar{\phi}(s)$ is a specified function of arc length parameter s on boundaries with a Dirichlet type boundary condition. In Equation (35), n denotes the outer normal direction and $\bar{q}(s)$ is a specified function on boundaries with a Neumann type boundary condition. Across interior boundaries, n denotes an arbitrarily directed normal direction; by definition this normal direction is an outward normal for the negative side of the inner boundary and an inward normal for the positive side of the boundary.

The choice of the ${}_r\phi$ as given by Equations (28) through (32) makes it possible to easily provide continuity in values of ϕ across internal boundaries. Furthermore, it will be assumed that on S_D the specified $\bar{\phi}(s)$ is such that Equation (34) can be identically satisfied. Therefore, the complete finite element Galerkin method equation will be

$$\sum_r \iint_{rA} \left\{\left[\frac{\partial}{\partial x}\left({}_rK_x \frac{\partial}{\partial x}\right)+\frac{\partial}{\partial y}\left({}_rK_y \frac{\partial}{\partial y}\right)\right]\left[\sum_\ell {}_rN_\ell \phi_\ell\right]\right\}\left\{\sum_m {}_rN_m \delta\phi_m\right\} d_rA$$

$$-\sum_b \int_{bS_N}\left[{}_bK_n \frac{d}{dn}\left(\sum_\ell {}_bN_\ell \phi_\ell\right)-\bar{q}\right]\left[\sum_m {}_bN_m \delta\phi_m\right] d_bs \qquad (36)$$

$$+\sum_{m+p} \int_{S_{int}}\left[{}_pK_n \frac{d}{dn}\left(\sum_\ell {}_pN_\ell \phi_\ell\right)-{}_mK_n \frac{d}{dn}\left(\sum_\ell {}_mN_\ell \phi_\ell\right)\right]\left[\sum_i {}_pN_i \delta\phi_i = \sum_i {}_mN_i \delta\phi_i\right] ds = 0$$

The first term represents a summation over all finite elements. The second term represents a summation over all those finite elements which are adjacent to the S_N boundary. The last term is a summation over all interior boundaries, with the left subscript p denoting the plus side and the left subscript m denoting the minus side of the boundary.

Green's theorem can be applied to the first term in Equation (36), resulting in the following completely equivalent form of the Galerkin equation:

$$\sum_r \iint_{rA} \left\{\left[{}_rK_x \sum_\ell \frac{\partial_r N_\ell}{\partial x}\phi_\ell\right]\left[\sum_m \frac{\partial_r N_m}{\partial x}\delta\phi_m\right]\right.$$

$$\left.+\left[{}_rK_y \sum_\ell \frac{\partial_r N_\ell}{\partial y}\phi_\ell\right]\left[\sum_m \frac{\partial_r N_m}{\partial y}\delta\phi_m\right]\right\} d_rA \qquad (37)$$

$$-\sum_b \int_{bS_N} \bar{q}\left[\sum_m {}_bN_m \delta\phi_m\right] d_bs = 0$$

When going from Equation (36) to Equation (37), it is necessary to recall that Green's theorem involves outward normal directions, while the normals on interior boundaries are arbitrarily inward or outward depending on whether the region is positive or negative with respect to that inner boundary. Also it has been necessary to note that

$$K_n \frac{d\phi}{dn} = K_x \frac{\partial \phi}{\partial x} n_x + K_y \frac{\partial \phi}{\partial y} n_y \qquad (38)$$

where n_x and n_y are the components of a unit vector in the direction specified by a vector $\underline{n}$.

Equation (37) is exactly what would result from a finite element evaluation of the true variational principle for this problem. This furnishes one more demonstration of the equivalence between the complete Galerkin equation, Equation (36), and a true variational principle.

Requiring that Equation (37) be satisfied for arbitrary $\delta\phi_m$ leads to the usual matrix formulation of the form

$$[S]\{\phi\} = \{F\} \qquad (39)$$

where (S) is the global matrix of coefficients (analogous to the stiffness matrix) which incorporates the properties of the materials in the domain and the geometry, $\{\phi\}$ is the vector of unknown ϕ's at the nodes, and $\{F\}$ is the global "load" vector. The details of the development of Equation (39) and possible solution techniques are conventional and will not be discussed in this paper.

The same type of formulation can be carried out for the parabolic partial differential equation described by Equation (2), a problem for which no true variational principle exists. A complete statement of the problem is assumed to be

$$\frac{\partial}{\partial x}\left(K_x \frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial \phi}{\partial y}\right) - K_* \frac{\partial \phi}{\partial t} = 0 \quad , \text{ in } A \, , \; t > 0 \qquad (40)$$

$$\phi = \bar{\phi}(s,t) \quad , \text{ on } S_0 \, , \; t > 0 \qquad (41)$$

$$K_n \frac{d\phi}{dn} = \bar{q}(s,t) \, , \text{ on } S_N \, , \; t > 0 \qquad (42)$$

$$\phi = \phi_0(x,y) \quad , \text{ in } A \quad , \; t = 0 \qquad (43)$$

with no great concern regarding continuity of boundary and initial data. The interior continuity requirements associated with Equations (33) through (35) are applicable again. In addition, ϕ should be a continuous function of time t.

Since the solution ϕ is a function of both space and time the approximate solution within the r^{th} element is now taken in the form

$$ {}_r\phi(x,y,t) = \sum_i {}_rN_i(x,y)\,\phi_i(t) \tag{44} $$

It will be assumed that Equations (41) and (43) and internal continuity of ϕ in both space and time are identically satisfied. Then the finite element Galerkin equation will be

$$ \int_0^t \Bigg\langle \sum_r \iint_{rA} \left\{ \left[\frac{\partial}{\partial x}\left({}_rK_x \frac{\partial}{\partial x}\right) + \frac{\partial}{\partial y}\left({}_rK_y \frac{\partial}{\partial y}\right) - {}_rK_t \frac{\partial}{\partial t} \right] \left[\sum_\ell {}_rN_\ell \phi_\ell \right] \right\} \left\{ \sum_m {}_rN_m \delta\phi_m \right\} d_rA $$

$$ - \sum_b \int_{bS_N} \left[{}_bK_n \frac{d}{dn}\left(\sum_\ell {}_bN_\ell \phi_\ell \right) - \bar{q} \right] \left[\sum_m {}_bN_m \delta\phi_m \right] d_bs $$

$$ + \sum_{m-p} \int_{S_{int}} \left[{}_pK_n \frac{d}{dn}\left(\sum_\ell {}_pN_\ell \phi_\ell \right) - {}_mK_n \frac{d}{dn}\left(\sum_\ell {}_mN_\ell \phi_\ell \right) \right] \tag{45} $$

$$ \left[\sum_i \rho N_i \delta\phi_i = \sum_i m N_i \delta\phi_i \right] ds \Bigg\rangle dt = 0 $$

After application of Green's theorem, the Galerkin equation will be

$$ \int_0^t \Bigg\langle \sum_r \iint_{rA} \Bigg\{ \left[{}_rK_x \sum_\ell \frac{\partial {}_rN_\ell}{\partial x} \phi_\ell \right] \left[\sum_m \frac{\partial {}_rN_m}{\partial x} \delta\phi_m \right] $$

$$ + \left[{}_rK_y \sum_\ell \frac{\partial {}_rN_\ell}{\partial y} \phi_\ell \right] \left[\sum_m \frac{\partial {}_rN_m}{\partial y} \delta\phi_m \right] \Bigg\} d_rA \tag{46} $$

$$ + \sum_r \iint_{rA} \left[{}_rK_t \sum_\ell {}_rN_\ell \frac{\partial \phi_\ell}{\partial t} \right] \left[\sum_m {}_rN_m \delta\phi_m \right] d_rA $$

$$ - \sum_b \int_{bS_N} \bar{q} \left[\sum_m {}_bN_m \delta\phi_m \right] d_bs \Bigg\rangle dt = 0 $$

Requiring that Equation (46) be satisfied for arbitrary $\delta\emptyset_m$ leads to a system of differential equations which can be written in the matrix form

$$[S]\{\phi\} + [P]\left\{\frac{\partial\phi}{\partial t}\right\} = \{F\} \tag{47}$$

To proceed with the transient part of the solution, it will be assumed that $\partial\emptyset/\partial t$ associated with each degree of freedom of discrete system varies linearly within a time increment (Δt), as first suggested by Clough and Wilson [17]. Thus, from a direct integration over the time interval, Δt for all nodal points, the following equation for $\emptyset$ at the end of a time interval can be obtained.

$$\{\phi\}_t = \{\phi\}_{t-\Delta t} + \left(\left\{\frac{\partial\phi}{\partial t}\right\}_{t-\Delta t} + \left\{\frac{\partial\phi}{\partial t}\right\}_t\right)\frac{\Delta t}{2} \tag{48}$$

Thus, if the initial values of $\emptyset$ are known, Equations (47) and (48) can be solved simultaneously to obtain the values of $\emptyset$ at the time $(t + \Delta t)$.

This simultaneous time and space solution process can be formulated neatly for computer applications, thus decreasing the algebra, computer storage and time. At time t substituting Equation (48) into Equation (47), one can write

$$\left([P]\frac{2}{\Delta t} + [S]\right)\{\phi\}_t = [P]\left(\left\{\frac{\partial\phi}{\partial t}\right\}_{t-\Delta t} + \frac{2}{\Delta t}\{\phi\}_{t-\Delta t}\right) + \{F\} \tag{49}$$

Again substituting Equation (47) into Equation (49), this time at time $(t - \Delta t)$, one can write

$$\left([P]\frac{2}{\Delta t} + [S]\right)\{\phi\}_t = \left([P]\frac{2}{\Delta t} - [S]\right)\{\phi\}_{t-\Delta t} + 2\{F\} \tag{50}$$

If one defines

$$[S]^* = [P]\frac{2}{\Delta t} + [S] \tag{51}$$

and

$$\{F\}^* = \{F\} + \frac{2}{\Delta t}[P]\{\phi\}_{t-\Delta t} \qquad (52)$$

then from Equation (53) one can solve for the unknowns $\{\phi\}^*$

$$[S]^*\{\phi\}^* = \{F\}^* \qquad (53)$$

Once the $\{\phi\}^*$'s are determined, the problem reduces to solving Equation (54) for the values of ϕ_t which are the nodal values of the function sought at time t.

$$\{\phi\}_t = 2\{\phi\}^* - \{\phi\}_{t-\Delta t} \qquad (54)$$

Incrementing by Δt and repeating the same process, continuous solutions can be obtained in time and space coordinates for unsteady problems.

The solution process for the convective diffusion problem, Equation (3), is slightly modified based on a true variational principle for a simplified problem. For the case of unsteady diffusion with uniform flow in a homogeneous medium, reference [16] shows that there exists a true variational principle with the functional

$$I = \iint_A \left[\frac{1}{2}K_x\left(\frac{\partial\phi}{\partial x}\right)^2 + \frac{1}{2}K_y\left(\frac{\partial\phi}{\partial y}\right)^2\right]\exp\left(-\frac{ux}{K_x} - \frac{vy}{K_y}\right)dA \qquad (55)$$

The physical interpretation for the magnitude of I is not clear, but the fact remains that $\delta I = 0$ does supply the governing equations for a system with specified ϕ on the boundary. It should be noted that in the term $\exp(-ux/K_x - vy/K_y)$, the x and y values are evaluated in a global or overall coordinate system.

For the case of transient diffusion with uniform flow in a homogeneous medium, the differential equation will be

$$K_x\frac{\partial^2\phi}{\partial x^2} + K_y\frac{\partial^2\phi}{\partial y^2} - u\frac{\partial\phi}{\partial x} - v\frac{\partial\phi}{\partial y} - \frac{\partial\phi}{\partial t} = 0 \quad , \text{in } A \, , \; t > 0 \qquad (56)$$

with boundary conditions given by Equations (41) through (43) and continuity requirements in space and time as specified for the previous example. Once again it will be assumed that Equations (41) through (43) and internal continuity of ϕ in both space and time are identically satisfied. Now the finite element Galerkin type equation will be taken as follows

$$\int_0^t \Big\langle \sum_r \iint_{rA} \Big\{ \Big(K_x \frac{\partial^2}{\partial x^2} + K_y \frac{\partial^2}{\partial y^2} - u \frac{\partial}{\partial x} - v \frac{\partial}{\partial y} - \frac{\partial}{\partial t} \Big) \Big(\sum_\ell {}_rN_\ell \, \phi_\ell \Big) \Big\}$$

$$\otimes \Big\{ \sum_m {}_rN_m \, \delta\phi_m \, \exp\Big[-\frac{u}{K_x} ({}_rx + {}_ra) - \frac{v}{K_y} ({}_ry + {}_rb) \Big] \Big\} d_rA$$

$$- \sum_b \int_{bS_N} \Big\{ K_n \frac{d}{dn} \Big(\sum_\ell {}_bN_\ell \, \phi_\ell \Big) - \bar{q} \Big\}$$

$$\otimes \Big\{ \sum_m {}_bN_m \, \delta\phi_m \, \exp\Big[-\frac{u}{K_x} ({}_bx + {}_ba) - \frac{v}{K_y} ({}_by + {}_bb) \Big] \Big\} d_bs \qquad (57)$$

$$+ \sum_{m \neq p} \int_{S_{int}} \Big\{ K_n \frac{d}{dn} \Big(\sum_\ell {}_pN_\ell \, \phi_\ell \Big) - K_n \frac{d}{dn} \Big(\sum_\ell {}_mN_\ell \, \phi_\ell \Big) \Big\}$$

$$\otimes \Big\{ \sum_i {}_pN_i \, \delta\phi_i \, \exp\Big[-\frac{u}{K_x} ({}_px + {}_pa) - \frac{v}{K_y} ({}_py + {}_pb) \Big] \Big\} ds \Big\rangle \, dt = 0$$

In Equation (57), the integral over the r^{th} finite element or the b^{th} exterior boundary or a given interior boundary is normally expressed in terms of the appropriate local coordinate systems. Therefore it is necessary to transform the global coordinates to local coordinates by the following equations which account for the translation:

$$x = {}_rx + {}_ra \quad , \qquad y = {}_ry + {}_rb \qquad (58)$$

Note that there need be no coordinate rotations for this case of homogeneous material.

After application of Green's theorem, the Galerkin-type equation will be

$$\int_0^t \Big\langle \sum_r \iint_A \Big[\Big(K_x \sum_\ell \frac{\partial_r N_\ell}{\partial x} \phi_\ell \Big) \Big(\sum_m \frac{\partial_r N_m}{\partial x} \delta\phi_m \Big) + \Big(K_y \sum_\ell \frac{\partial_r N_\ell}{\partial y} \phi_\ell \Big) \Big(\sum_m \frac{\partial_r N_m}{\partial y} \delta\phi_m \Big) \Big] \exp\Big[-\frac{u}{K_x} ({}_r x + {}_r a) - \frac{v}{K_y} ({}_r y + {}_r b) \Big] d_r A$$

$$+ \sum_r \iint_{rA} \Big(\sum_\ell {}_r N_\ell \frac{\partial \phi_\ell}{\partial t} \Big) \sum_m {}_r N_m \delta\phi_m \exp\Big[-\frac{u}{K_x} ({}_r x + {}_r a) - \frac{v}{K_y} ({}_r y + {}_r b) \Big] d_r A \qquad (58)$$

$$- \sum_b \bar{q} \Big\{ \sum_m {}_b N_m \delta\phi_m \exp\Big[-\frac{u}{K_x} ({}_b x + {}_b a) - \frac{v}{K_y} ({}_b y + {}_b b) \Big] \Big\} d_b s \Big\rangle dt$$

Requiring that Equation (58) be satisfied for arbitrary $\delta\phi_m$ leads to system of differential equations which can be written in the form given by Equation (47). If the flow is steady, then the [S] and [P] matrices are constants, and the solution proceeds as described earlier.

The numerical results presented in Example 2 of the next section are based on Equation (59) with the simplifying assumption that ${}_r a = 0$ and ${}_r b = 0$; that is the translation between local and global coordinates has been ignored. This means that the interior boundary continuity requirement for K_n $d\phi/dn$ is not satisfied. However, the results agree very well with the known finite difference solution.

SECTION V

NUMERICAL EXAMPLES

In this section, numerical examples to the differential equations treated in detail in earlier sections will be given. These examples are chosen in the field of fluid mechanics. The variety of possibilities for the application of the computer program written for the solution of these problems is obvious [16]. Also obviously, the appropriate employment of the mathematical form is by no means limited to the problem types specifically treated in this study. Indeed, any physical phenomena properly modelled by elliptic and parabolic differential equations, accompanied by appropriate boundary conditions can be investigated using the techniques developed in this study.

Example One: An engineering application of the finite element method developed in this paper can be the study of ground water seepage under a dam as demonstrated earlier in a paper by Zienkiewicz, Mayer and Cheung [3]. If the earth material under this dam is layered and anisotropic, it becomes very difficult for an engineer to estimate the quantity of seepage passing under this dam.

In this problem an arbitrarily conceptualized dam is located on a layered and anisotrpic media. The problem may be the determination of equipotential lines under a constant head. A schematic description of this problem is given in Figure 2.

Within each layer the mathematical model describing this problem is

$$K_x \frac{\partial^2 \phi}{\partial x^2} + K_y \frac{\partial^2 \phi}{\partial y^2} = 0 \qquad (60)$$

where K_x and K_y are permeabilities of the medium in the principle local directions, respectively. ϕ in this model represents the equipotential function. As the boundary conditions of the problem, ϕ is assumed to vary from ten on the upstream side of the dam to zero on the downstream side of the dam. Under the dam and at the impervious boundary the normal derivative of ϕ is equal to zero since there is no flux across these boundaries.

Finite element results for this problem can be seen on Figure 3. In this problem, the continuum analyzed was idealized by 297 elements and the computer execution time was 15 seconds.

Example Two: Another application of the finite element method can be the study of dispersion in a porous medium. This phenomenon can be modelled by the convective dispersion equation presented in the earlier section. The two-dimensional form of this equation in cartesian coordinates is

$$D_x \frac{\partial^2 \phi}{\partial x^2} + D_y \frac{\partial^2 \phi}{\partial y^2} - u \frac{\partial \phi}{\partial x} - v \frac{\partial \phi}{\partial y} = \frac{\partial \phi}{\partial t} \qquad (61)$$

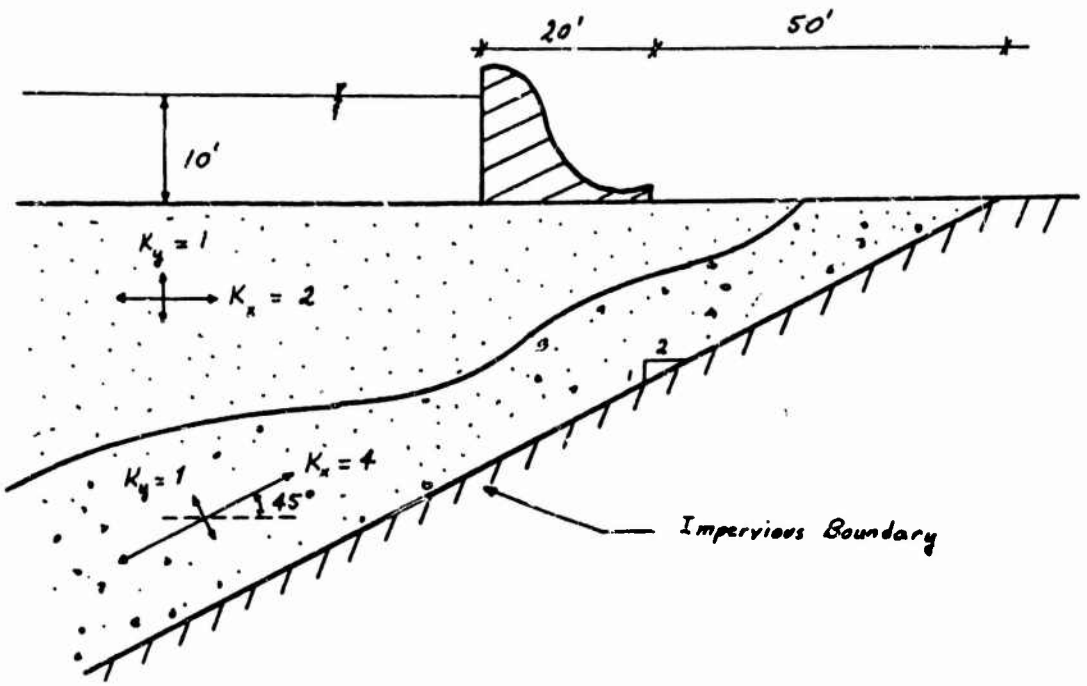


Figure 2. Concrete Dam on Non-Homogeneous and Anisotropic Media

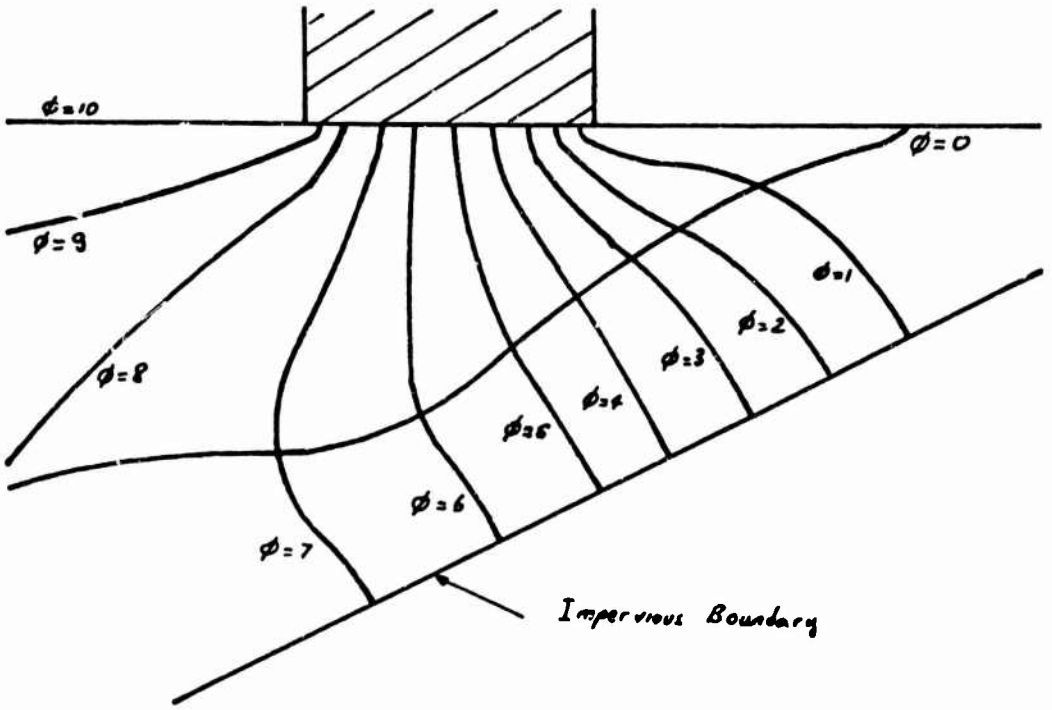


Figure 3. Equipotential Line Configuration Under a Dam on Non-Homogeneous and Anisotropic Media

where ϕ is the concentration of the dispersing mass, D_x and D_y are the dispersion coefficients and u and v are the seepage velocities in (x) and (y) directions respectively. Shamir and Harlemen [18] studied steady and unsteady problems of this type in detail and presented a numerical scheme for solving such problems. In this example, a typical unsteady problem studied in their report was solved by the finite element method.

The problem is to determine the one-dimensional dispersion of a tracer (concentration) introduced to the porous medium at a constant rate at x = 0 cm. There is a constant rate of seepage in the x-direction. The velocity distribution is assumed to be uniform in y-direction. Initially, the distribution of this tracer in the porous medium is zero. A schematic description of this problem is given in Figure 4.

The dimensionless form of the unsteady convective diffusion equation in one-dimensional studies is

$$\frac{\partial^2 \phi}{\partial \xi^2} - \lambda \frac{\partial \phi}{\partial \xi} = \frac{\partial \phi}{\partial \tau} \tag{62}$$

where

$$\phi = \frac{\phi}{\phi_o}, \quad \xi = \frac{x}{L}, \quad \lambda = \frac{uL}{D}, \quad \tau = \frac{Dt}{L^2} \tag{63}$$

where L is the length of the medium, t is the time coordinate and ϕ_o is the amount of concentration at x = 0 which is kept constant throughout the analysis. In applying the numerical solution the following data, which are taken from the above report, are used.

$$0 \leq x \leq 10 \text{ cm}$$

$$u = 0.1 \text{ cm/sec} \tag{64}$$

$$D = 0.01 \text{ cm}^2/\text{sec}$$

Thus the convective diffusion equation to be solved and the transformed boundary conditions take the form

$$\frac{\partial^2 \phi}{\partial \xi^2} - 100 \frac{\partial \phi}{\partial \xi} = \frac{\partial \phi}{\partial \tau} \tag{65}$$

$$\text{I.C.}: \quad \text{at} \quad \tau = 0 \qquad \phi = 0 \tag{66}$$

$$\text{B.C.1}: \quad \text{at} \quad \xi = 0 \qquad \phi = 1 \tag{67}$$

$$\text{B.C.2}: \quad \text{at} \quad \xi = 1 \qquad \frac{\partial \phi}{\partial \xi} = 0 \tag{68}$$

Numerical results and a comparison with the exact solution [18] are given in Figure 5.

In this example, 200 elements were used in idealizing the continuum given in Figure 4. The computer execution time was 0.16 seconds per time step. Although the finite element results are very accurate for this particular example, extreme care should be given to the coefficients appearing in the exponential terms in the solution of this type of problems. Very large numbers appearing in the exponent create truncation and roundoff errors in computer computations which may be significant and may result in unstable solutions.

This example was actually a laboratory model of flow of a certain concentration in a confined aquifer. A practical application of this example could be the study of the flow of some concentration in a confined aquifer between two rivers, of course, the appropriate constants of the problem may have to be changed for a specific application.

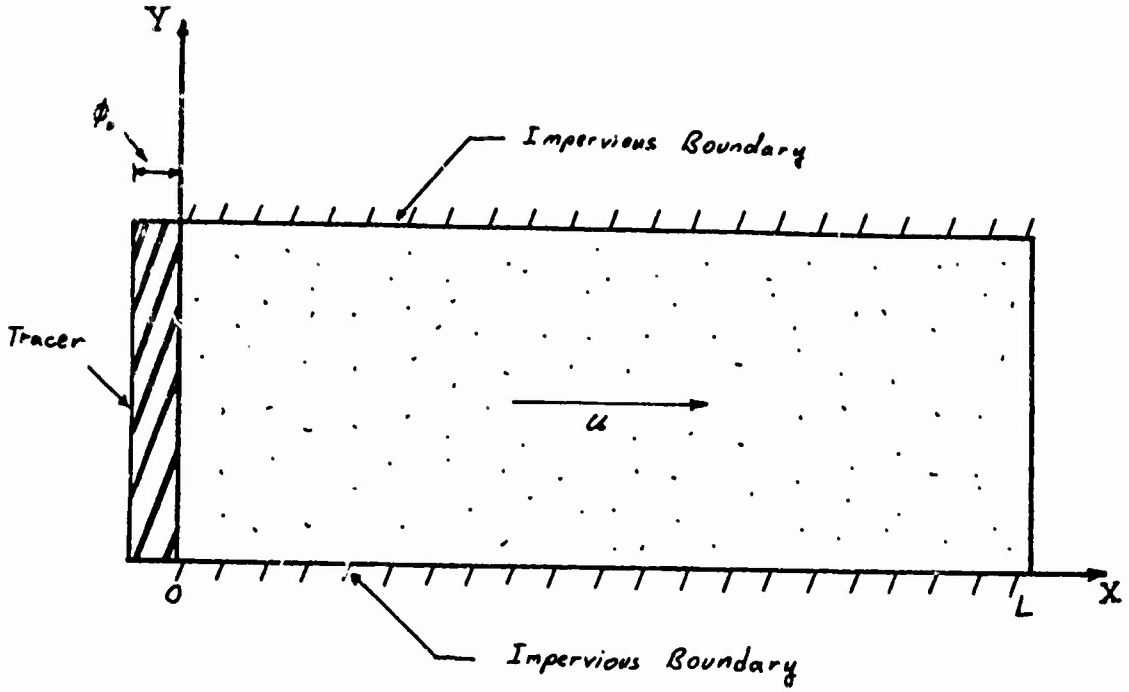


Figure 4. Longitudunal Dispersion Model

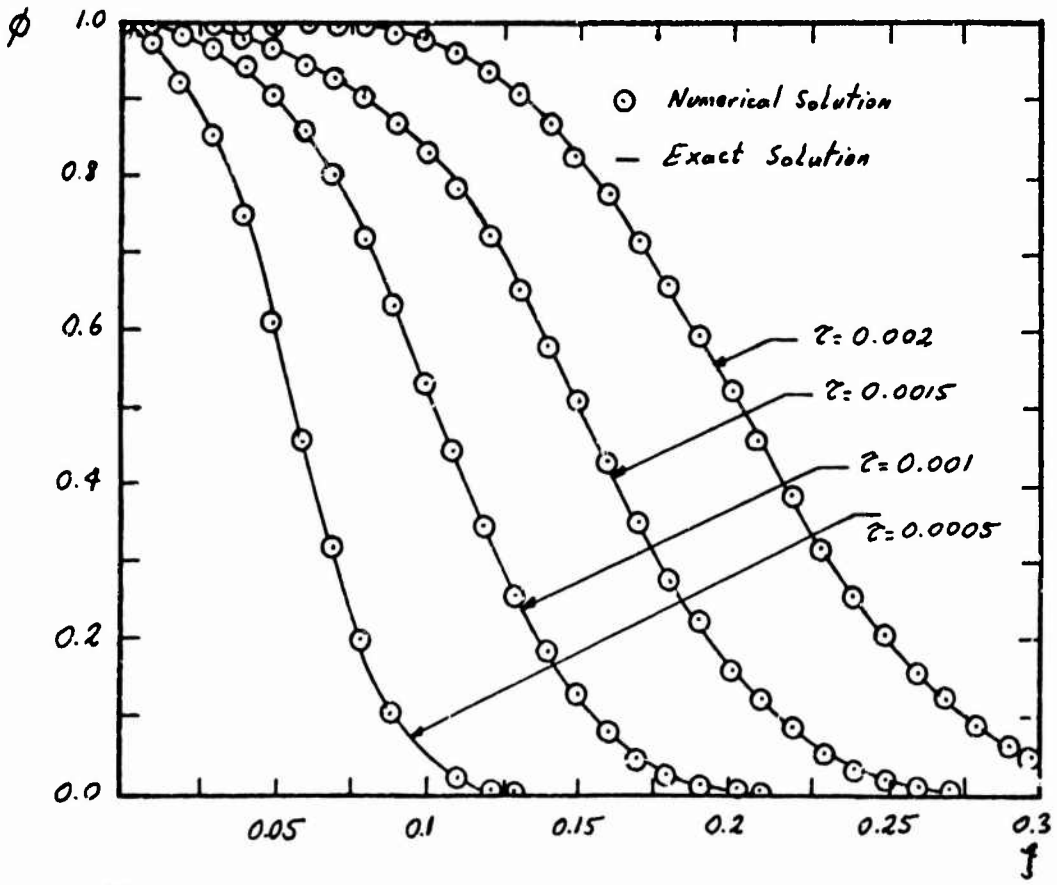


Figure 5. Convective Diffusion in Porous Media

SECTION VI

CONCLUDING REMARKS

In this paper a technique has been presented for the formulation of approximate solution equations for systems with no true variational principle. While the results are not necessarily any more accurate than earlier presented solutions, the formulations possess a rigor which is lacking in those solutions based on quasi or restricted variational principles. It is now clear that the finite element method can be used to provide solutions to many significant problems in fluid mechanics, heat transfer, and other nonstructural areas.

SECTION VII

REFERENCES

1. Turner, M. J., and Clough, R. W., Martin, H. C., and Topp, L. J., "Stiffness and Deflection Analysis of Complex Structures", Journal of Aerospace Sciences, v. 23, n. 9, pp. 805-823, September 1956.

2. Melosh, R. J., "Basis for Derivation of Matrices for the Direct Stiffness Method", AIAA J., v. 1, n. 7, pp. 1631-1637, July 1963.

3. Zienkiewicz, O. C., Mayer, P. G., and Cheung, Y. K., "Solution of Anisotropic Seepage Problems by Finite Elements", Proceedings, American Society of Civil Engineers, 92, EM1, pp. 111-120, 1966.

4. Wilson, E. L., and Nickell, R. E., "Application of the Finite Element Method to Heat Conduction Analysis", Nuclear Engineering and Design, North-Holland Publishing Company, Amsterdam, pp. 276-186, 1966.

5. Martin, H. C., "Finite Element Analysis of Fluid Flows", Proceedings of the Second Conference on Matrix Methods in Structural Mechanics, pp. 517-535, October 1968.

6. Visser, W., "A Finite-Element Method for the Determination of Non-Stationary Temperature Distributions and Thermal Deformations", Proceedings of Conference on Matrix Methods in Structural Mechanics, pp. 925-943, October 1965.

7. Finlayson, B. A., and Scriven, L. E., "On the Search for Variational Principles", Int. J. Heat Mass Transfer, v. 10, pp. 799-821.

8. Finlayson, B. A., and Scriven, L. E., "The Method of Weighted Residuals - A Review", Appl. Mech. Rev., v. 19, pp. 735-748, September 1966.

9. Finlayson, B. A., and Scriven, L. E., "The Method of Weighted Residuals and Its Relation to Certain Variational Principles for the Analysis of Transport Processes", Chem. Engng. Sci., v. 20, pp. 395-404, May 1965.

10. Oden, J. T., and Kross, D. A., "Analysis of General Coupled Thermo-elasticity Problems by the Finite Element Method", Proceedings of the Second Conference on Matrix Methods in Structural Mechanics, pp. 1091-1120, October 1968.

11. Szabo, B. A., and Lee, G. C., "Stiffness Matrix for Plates by Galerkin's Method", J. Engr. Mech. Div., ASCE, pp. 571-585, June 1969.

12. Prayer, W., "Variational Principles of Linear Elastostatics for Discontinuous Displacements, Strains, and Stresses", Recent Progress in Applied Mechanics, The Folke Odgvist Volume, ed. by Broberg, Hult, and Niordson, John Wiley, pp. 463-474, 1967.

13. Herrmann, L. R., "Finite-Element Bending Analysis for Plates", J. Engng. Mech. Div., ASCE, v. 93, pp. 13-26, October 1967.

14. Smith, C. V., Jr., "On the Method of Weighted Residuals, with Application to the Finite Element Technique", submitted for publication.

15. Zienkiewicz, O. C., and Cheung, Y. K., The Finite Element Method in Structural and Continuum Mechanics, McGraw-Hill, 1967.

16. Aral, Mustafa M., Application of Finite Element Analysis in Fluid Mechanics, Ph.D. thesis, Georgia Institute of Technology, Atlanta, Georgia, September 1971.

17. Wilson, E. L., and Clough, R. W., "Dynamic Response by Step-by-Step Matrix Analysis", Symposium on the Use of Computers in Civil Engineering, Lisbon, Portugal, 1962.

18. Shamir, U., and Harleman, D. R. F., "Numerical and Analytical Solutions of Dispersion Problems in Homogeneous and Layered Aquifers", M.I.T. Dept. of Civil Eng., Hydrodynamics Lab. Report 89, May 1966.

A BOUND THEOREM IN EIGENVALUES AND ITS PRACTICAL APPLICATIONS

Bruce M. Irons*
University of Wales
Swansea

Gabe Treharnet
Ove Arup and Partners
London

A familiar but undervalued theorem is presented, with applications. The theorem is that for positive-definite systems, the eigenvalues are bounded by the extreme element eigenvalues, and also by the extreme eigenvalues of infinitesimal elements. Familiar applications are to the relative stiffnesses of different formulations of the same element, to a roundoff criterion, and to the stability of elastoplastic solutions. Less familiar applications are to the noise dissipation and possible instability of first order marching solutions, and to the instability of creep solutions.

NOMENCLATURE

A, B, M, K are square symmetric matrices.

x, v, w are eigenvectors.

a_i, b_i are contributions to $\mathbf{x}^T\mathbf{A}\mathbf{x}$ and $\mathbf{x}^T\mathbf{B}\mathbf{x}$ from element i: $b_i > 0$.

λ, γ, Γ are eigenvalues.

ω is the over-relaxation factor.

D, $\mathbf{D}_p$ are modulus matrices relating stress and strain.

INTRODUCTION

The theorem discussed here gives bounds for eigenvalues in a form especially suited to finite element problems. It is already known, (Reference 9), it is indeed trivial, but it is not as familiar as it should be. Its use in large-scale computing often enables us to comment on the result by means of a very small calculation. Its use in teaching helps us to emphasize the most important aspects of eigenvalues, and gives students more physical insight. The purpose of this paper is to urge its inclusion in finite element courses.

*Senior Lecturer.
†Engineer.

RAYLEIGH'S PRINCIPLE

It is fruitful to define an eigenvalue as the stationary value of a Rayleigh quotient. If $\mathbf{Ax} = \lambda\mathbf{Bx}$, then the Rayleigh quotient for a slightly perturbated vector $x + dx$ is

$$\frac{(x+dx)^T\mathbf{A}(x+dx)}{(x+dx)^T\mathbf{B}(x+dx)} = \frac{\mathbf{x}^T\mathbf{Ax} + \mathbf{x}^T\mathbf{A}dx + dx^T\mathbf{Ax} + dx^T\mathbf{A}dx}{\mathbf{x}^T\mathbf{Bx} + \mathbf{x}^T\mathbf{B}dx + dx^T\mathbf{Bx} + dx^T\mathbf{B}dx}$$

$$= \frac{\lambda\,(\mathbf{x}^T\mathbf{Bx} + \mathbf{x}^T\mathbf{B}dx + dx^T\mathbf{Bx}) + dx^T\mathbf{A}dx}{(\mathbf{x}^T\mathbf{Bx} + \mathbf{x}^T\mathbf{B}dx + dx^T\mathbf{Bx}) + dx^T\mathbf{B}dx}$$

which equals λ to second order in dx. The quadratic forms $\mathbf{x}^T\mathbf{Ax}$ and $\mathbf{x}^T\mathbf{Bx}$ are energies, in general. Thus an eigenvector $\mathbf{x}$ is usually a set of displacements which give the ratio of two energies a stationary value, which is the corresponding eigenvalue λ. The Rayleigh quotient cannot lie outside the range λ_{min} to λ_{max}, because $\mathbf{x}^T\mathbf{Ax}/\mathbf{x}^T\mathbf{Bx}$ is a well-behaved function of the $\mathbf{x}$.

THE BOUND THEOREM FOR ELEMENT EIGENVALUES

Consider $\lambda = (a_1 + a_2 + \ldots + a_n) / (b_1 + b_2 + \ldots + b_n)$, where all the b_i are positive. If every a_i/b_i is greater than m and less than M, therefore, $mb_i < a_i < Mb_i$. Thus

$$\frac{m(b_1+b_2+\ldots+b_n)}{(b_1+b_2+\ldots+b_n)} < \frac{a_1+a_2+\ldots+a_n}{b_1+b_2+\ldots+b_n} = \lambda < \frac{M(b_1+b_2+\ldots+b_n)}{(b_1+b_2+\ldots+b_n)}$$

or $m < \lambda < M$. Here a_i denotes the contribution to $\mathbf{x}^T\mathbf{Ax}$ from element i and b_i denotes the contribution to $\mathbf{x}^T\mathbf{Bx}$ from element i. But a_i/b_i is itself bounded by the lowest and highest eigenvalues of element i, namely λ^i_{min} and λ^i_{max}. Thus we have the theorem:

$$\lambda^i_{min} \text{ (over all elements)} < \lambda_{min} \text{ (for structure)}$$

$$< \lambda_{max} \text{ (for structure)} < \lambda^i_{max} \text{ (over all elements)} \qquad (1)$$

For example, the highest frequency of an idealized structure is less than the highest frequency of its smallest element.

THE BOUND THEOREM FOR INFINITESIMAL ELEMENTS

We now consider a_i and b_i as the contributions to $\mathbf{x}^T\mathbf{Ax}$ and $\mathbf{x}^T\mathbf{Bx}$ of an infinitesimal element i. The extreme eigenvalues of all such infinitesimal elements within a finite element bound the extreme eigenvalues of that element. Further, the extreme eigenvalues of all the infinitesimal elements in a structure bound the extreme eigenvalues of that structure. This is the most useful form of the theorem.

APPLICATION TO A ROUNDOFF CRITERION

A useful criterion of roundoff sensitivity in a Gaussian reduction is the "Diagonal energy" (the sum of the N diagonal contributions to the strain energy) divided by the N^2 terms of the strain energy (Reference 1):

$$\frac{U_d}{U} = \frac{\Sigma_i K_{ii} x_i^2}{\Sigma_i \Sigma_j K_{ij} x_i x_j} \qquad (i, j = 1 \text{ to } N) \qquad (2)$$

The bandwidth also affects the situation, although to a smaller extent. Now if u_d/u for each element is bounded – as it is if the element matrices are positive definite – then U_d/U is also bounded regardless of how many elements are assembled. A long cylinder with radial motions only is a case in point, or a beam on an elastic foundation: so is the problem of least-squares surface fitting: so is the case, usually, in marching solutions. Unfortunately, most element stiffness matrices are semidefinite.

RELATIVE STIFFNESS OF TWO ELEMENT FORMULATIONS

Suppose we have two formulations of the same problem, e.g. one by potential energy, the other by complementary energy. The first gives assembled stiffness $\mathbf{K}_1$ and the second $\mathbf{K}_2$. If

$$\mathbf{x}^T\mathbf{K}_1\mathbf{x} < \mathbf{x}^T\mathbf{K}_2\mathbf{x} \qquad (3)$$

for all $\mathbf{x}$, we say that the first formulation is less stiff than the second (Reference 2). Extending our ideas of positive-definiteness, we write $\mathbf{K}_1 < \mathbf{K}_2$. This condition follows if $\mathbf{k}_1 < \mathbf{k}_2$ for all elements, i.e. if all the eigenvalues from $\mathbf{k}_1\mathbf{x} = \lambda\mathbf{k}_2\mathbf{x}$ are less than 1, for all elements. The situation is complicated by two facts:

1. For certain $\mathbf{x}$, for example the rigid body motions, the element forces $\mathbf{k}_1\mathbf{x}$ and $\mathbf{k}_2\mathbf{x}$ are zero. These are ignored, as $a_i = b_i = 0$.

2. Certain other $\mathbf{x}$ – for example, the constant strain conditions – give $\mathbf{k}_1\mathbf{x} = \mathbf{k}_2\mathbf{x}$. Indeed, if both formulations are to converge, the unit values of λ must appear.

ROUGH ESTIMATE OF THE CRITICAL STEP-LENGTH IN TEMPERATURE TRANSIENTS

According to Leech (Reference 3) the critical step-length is about 1/6 of the smallest period in dynamic problems. Small elements have high frequencies. Therefore the expense of a marching solution increases as the mesh is refined, not only because there are more variables, but also because the time step decreases. It is useful to be able to cost this effect in advance.

In a temperature transient, we have $\mathbf{C}\tau = -\mathbf{Q}\dot{\tau}$ where τ = temperatures, $\mathbf{C}$ = conductivities, and $\mathbf{Q}$ = heat capacities. The eigen-situation here is

$$\mathbf{C}\tau = -\lambda \mathbf{C}\tau \tag{4}$$

so that $$\tau(t) = e^{-\lambda t}\tau(0) \tag{5}$$

That is, all the temperatures decrease together exponentially, at the same rate. To estimate the biggest λ, we consider the smallest element, we guess the temperature pattern that is most rapidly evanescent as in Figure 1, and we compute λ as a Rayleigh quotient, evidently a very small calculation.

When trapezoidal integration is used, the solution should be stable. What uninformed workers take for instability is usually noise decaying very slowly -- although with nonlinearities this can cause instability. Many workers (e.g. C. Parekh, C. Taylor, B. Irons at Swansea, K. Fullard at C.E.G.B. Berkeley, Gloucestershire) have independentiy proposed re-start techniques to eliminate noise. For example, one steps from $\tau(0)$ to $\tau(t)$ normally, then re-starts at $\tau^*(t/2) = \frac{1}{2}[\tau(0)+\tau(t)]$ having eliminated most of the noise. Figure 2 shows that the spectral effectiveness of this, the simplest algorithm, is surprisingly good. Occasionally without re-starts a solution diverges slowly (in say 200 steps). Presumably roundoff generates more noise per step than is dissipated (Reference 8).

THE CONVERGENCE OF THE ITERATIVE PROCESS IN AN INCREMENT OF PLASTIC DEFORMATION

Following Gallagher and Marcal (References 4, 5) we take $\mathbf{K}$ as the elastic stiffness and $\mathbf{K}_p$ as the plastic incremental stiffness. During the iteration we compute the residual forces $\mathbf{R}$ which, added to the external loads, would exactly balance the internal loads due to the incorrect state of stress in the structure (Reference 11). We then add to the current deflections $\boldsymbol{\delta}$ of the structure a correction, giving:

$$\boldsymbol{\delta}^* = \boldsymbol{\delta} + \omega \mathbf{K}^{-1}\mathbf{R} \tag{6}$$

where ω is the over-relaxation factor. If we are fortunate, and $\mathbf{K}_p^{-1} = \mathbf{K}^{-1}$ exactly, then $\boldsymbol{\delta}^*$ is correct. In general, however, there will again be a nonzero residual:

$$\mathbf{R}^* = \mathbf{K}_p \text{ multiplied by error in } \boldsymbol{\delta}^*$$

$$= (\mathbf{I} - \omega \mathbf{K}_p \mathbf{K}^{-1})\,\mathbf{R}$$

Similarly, $$\mathbf{R}^{**\cdots*} = (\mathbf{I} - \omega \mathbf{K}_p \mathbf{K}^{-1})^n \mathbf{R} \tag{7}$$

If $\mathbf{R}$ iterates towards zero, the process converges. With this technique we can avoid creating and inverting $\mathbf{K}_p$. But we must consider the conditions for convergence if we are to use the method confidently. We consider first the eigenvectors $\mathbf{v}_i$ such that $\mathbf{K}_p\mathbf{v}_i = \lambda_i \mathbf{K}\mathbf{v}_i$, but it is more convenient to work with the force perturbation, $\mathbf{w}_i = \mathbf{K}\mathbf{v}_i$. For Equation (7) we now have

$$(\mathbf{I} - \omega \mathbf{K}_p \mathbf{K}^{-1})\mathbf{w}_i = (1 - \omega\lambda_i)\mathbf{w}_i \tag{8}$$

which is a scalar multiplier. If Equation (7) is to converge, then, the factor must lie between -1 and 1, giving

$$0 < \omega\lambda_i < 2 \tag{9}$$

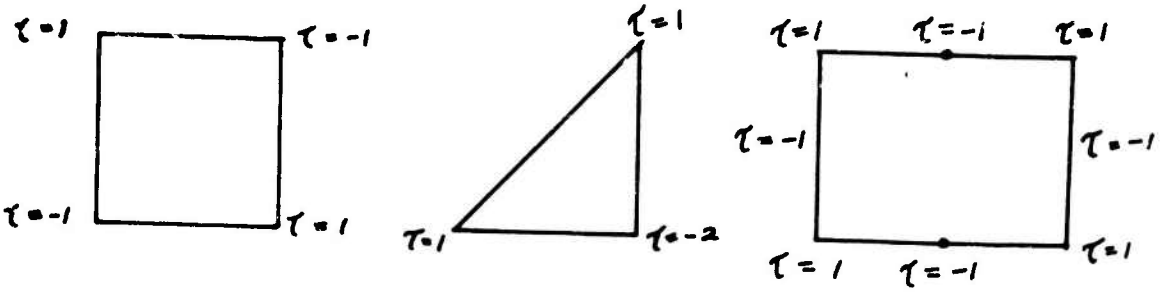


Figure 1. Inteiligent Guesswork in Approximating the Noisiest Eigenvector

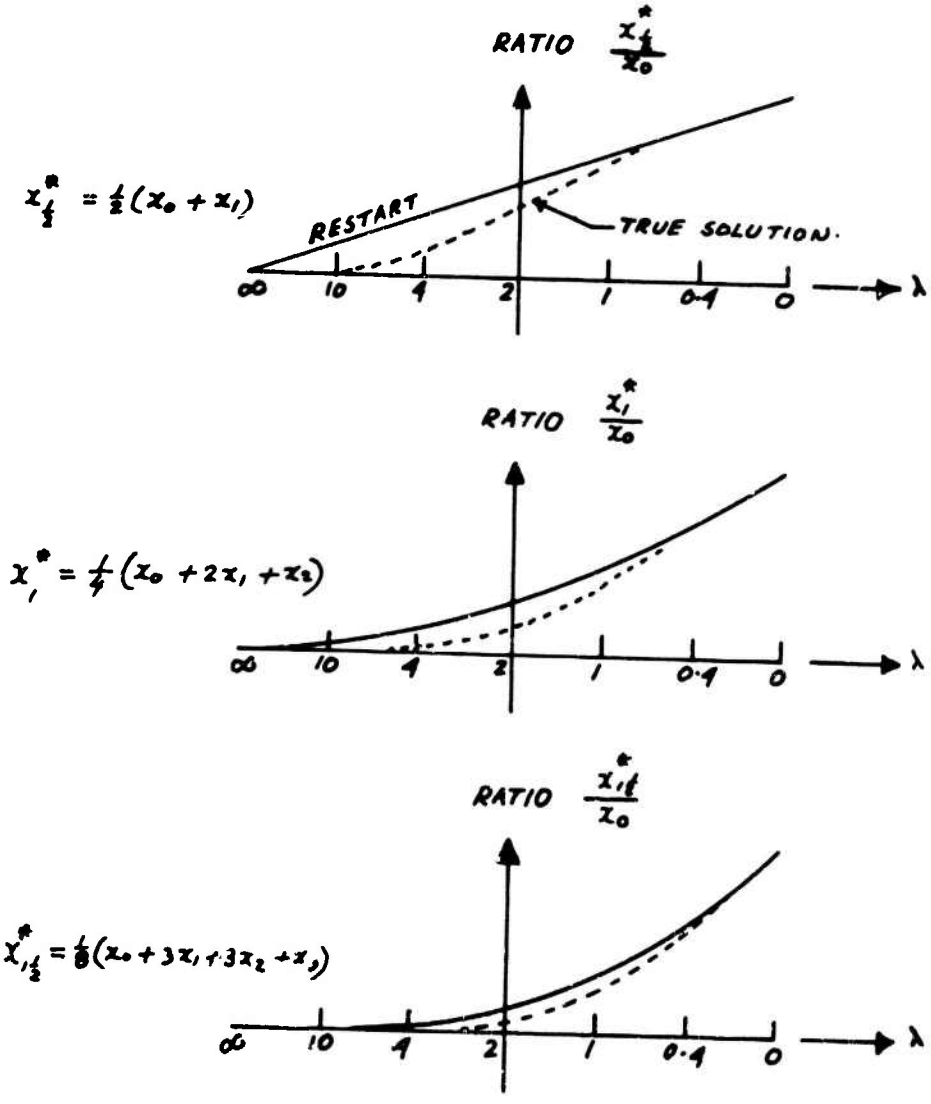


Figure 2. Performance of Different Noise-Rejecting Algorithms

As before, the λ_i are bounded by the extreme eigenvalues for individual elements. Presumably, $\lambda_i = 1$ for unyielded elements and $\lambda_i < 1$ for yielded elements. A more useful but wider bound depends only on the properties of an infinitesimal region of material. We choose, not too carefully, a stress increment which appears to give the smallest purely elastic strain energy or the largest purely plastic strain energy. We then take the ratio as an adequate bound on λ.

This is a nontrivial result (Reference 7). A numerical process may converge well in simple cases (e.g. Jacobi relaxation) but fail utterly in problems of any real size. But in the case of plasticity it appears that, if the process converges for a single constant-stress element, then it converges for a large structure. This conclusion holds generally. If an iterative or marching algorithm always succeeds with an infinitesimal element, then it should succeed with a large structure.

Professor G. Maier of Milan has remarked (private communication) that although this conclusion is universally true, the assumptions are unrealistic. Between iterations a small region can switch from elastic to plastic or vice versa, so that the w_i change radically; however the bounds on λ_i remain. As observed by Nayak (Reference 10) this behaviour is not rare, and indeed iteration with tangential stiffnesses is hazardous because the switching can become chaotic. However, we can prove that if Equation (9) is satisfied the perturbation

$$p = \Sigma a_i w_i$$

tends towards zero even if the a_i and w_i change from one iteration to the next.

We consider

$$\mathbf{p}^T\mathbf{K}^{-1}\mathbf{p} = \Sigma a_i^2 \mathbf{w}_i^T\mathbf{K}^{-1}\mathbf{w}_i$$

because the $\mathbf{w}_i$ are orthogonal. Each iteration effectively multiplies $\mathbf{w}_i^T\mathbf{K}^{-1}\mathbf{w}_i$ by the factor $F_i = (1 - \omega\lambda_i)^2$ giving a modified $\mathbf{p}^T\mathbf{K}^{-1}\mathbf{p}$:

$$\Sigma a_i^2 F_i \mathbf{w}_i^T\mathbf{K}^{-1}\mathbf{w}_i \leq F_{max}\mathbf{p}^T\mathbf{K}^{-1}\mathbf{p}$$

Therefore if $F_i \leq F_{max} < 1$ for all iterations, i.e. if Equation (9) is satisfied, $\mathbf{p}^T\mathbf{K}^{-1}\mathbf{p}$ tends towards zero and the iteration converges.

We note from Equation (9) that a given value of ω secures convergence for $0 < \lambda < 2/\omega$, and that for convergence $0 < \omega < 2/\lambda_{max}$. One would intuitively expect under-relaxation to be safer in all practical cases; Equation (9) confirms this.

Further, we can now choose a near-optimum value of ω. In the present case, λ_{min} is known from the stress-strain at the worst point in the structure and $\lambda_{max} = 1$ for the elastic regions. Thus the value of ω that minimizes the greatest absolute value of $(1-\omega\lambda)$, the convergence ratio, is given by

$$1 - \omega\lambda_{min} = -(1 - \omega\lambda_{max})$$

$$\text{or} \quad \omega = 2/(\lambda_{min} + \lambda_{max}) \tag{10}$$

NONSYMMETRIC MATRICES

If the problem generates a nonsymmetric matrix K_p, D. J. Naylor has suggested that $K = \frac{1}{2}(K_p + K_p^T)$ might be shown to give convergence. The theorem gives no support here, as the eigenvalues can be complex. It is not known how successful such approaches would be. But the loss of many useful theorems must discourage the most intrepid worker from voluntarily involving himself with nonsymmetric matrices.

STABILITY IN A MARCHING SOLUTION OF CREEP–TIME PROBLEMS

This is another case that depends directly on the stress-strain laws rather than on some postulated behaviour of the smallest finite element. The problem is technically important, because most workers prefer the simplest form of integration (in which the stress is assumed constant over each time interval) and the intervals must be short to avoid instability: longer intervals would usually give sufficient accuracy, even with this crude type of integration, except perhaps in the early time-steps (Reference 6).

The effective stress:

$$\bar{\sigma}^2 = \frac{3}{2}\sigma_{ij}\sigma'_{ij} \text{ where } \sigma'_{ij} = \sigma_{ij} - \frac{1}{3}\sigma_{kk}$$

$$= \tfrac{1}{2}[(\sigma_x - \sigma_y)^2 + (\sigma_y - \sigma_z)^2 + (\sigma_z - \sigma_x)^2 + 6\tau_{xy}^2 + \ldots] \qquad (11)$$

is associated with an effective strain in the neighborhood of some time t:

$$\bar{\epsilon} = \bar{\epsilon}\,(\bar{\sigma}, t)$$

$$\text{thus} \quad d\,\epsilon_x^c = \frac{d\bar{\epsilon}}{\bar{\sigma}}(\sigma_x - \tfrac{1}{2}\sigma_y - \tfrac{1}{2}\sigma_z) \text{ etc.} \qquad (12)$$

$$d\gamma_{xy}^c = \frac{d\bar{\epsilon}}{\bar{\sigma}}(3\tau_{xy}) \text{ etc.} \qquad (13)$$

as in the Prandtl–Reuss equations. Thus at the start of an interval δ we have the stresses $\boldsymbol{\sigma}$ and we have the elastic stiffness $\mathbf{K}$. We can also postulate a creep stiffness $(\delta t)^{-1}\mathbf{K}_c$, although this concept is not pursued for the moment. As before we could consider eigenvectors of displacements or of perturbating forces, but in this case it is more profitable to consider vectors of perturbating stresses. We thus consider an infinitesimal element of material, whose stresses are perturbated from the true solution, so that the stress and the strain perturbations relax exponentially together: the most critical condition is the case giving the fastest rate of decay.

This is essentially the same concept as used above, but we now modify it slightly to take account of the integration process. Consider the problem in one degree of freedom, $\dot{y} = -\lambda y$, solved using steps of δt, over which y is regarded as constant:

$$y_{i+1} = y_i - \lambda y_i \delta t$$

Leaving aside questions of accuracy, the condition for stability is

$$0 < \lambda \delta t < 2$$

Returning to the problem of creep flow, we consider a problem in plane strain ($\epsilon_z = 0$) giving both elastic and creep strains:

$$\text{elastic} \quad \epsilon_z = \sigma_z / E$$

$$\text{creep} \quad \epsilon_z = \sigma_z \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t}$$

But the total $\epsilon_z = 0$ at all times, so that

$$\frac{\sigma_z}{E} = -\sigma_z \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t}$$

$$\sigma_z = -\lambda \sigma_z \text{ where } \lambda = E \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t}$$

Substituting λ in Equation (14)

$$E \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t} \delta t < 2$$

Thus we derive the limiting step length as required:

$$\delta t < \frac{2}{E \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t}} \tag{15}$$

THE THREE–DIMENSIONAL CASE IN PLASTICITY AND IN CREEP

In the case of plasticity it is reasonably straightforward to express the result in a general three-dimensional form:

$$\mathbf{D}\,\boldsymbol{\epsilon} = \lambda\, \mathbf{D}_p \boldsymbol{\epsilon}$$

where $\mathbf{D}$ is the elastic modulus matrix and $\mathbf{D}_p$ is the plastic modulus matrix. This gives bounds for the λ of Equation (10).

The case of creep is more difficult. We postulate an interacting set of infinitesimal elements, with no external load involved in the perturbation of the stresses on these elements. Thus we have

$$\delta \boldsymbol{\epsilon}_e = \mathbf{D}^{-1} \delta \boldsymbol{\sigma}$$

$$(\delta \epsilon_c)_{kl} = \frac{\partial^2 \bar{\epsilon}}{\partial \bar{\sigma} \partial t} \frac{\partial \bar{\sigma}}{\partial \sigma_{ij}} \delta \sigma_{ij} \frac{\partial \bar{\sigma}}{\partial \sigma_{kl}}$$

$$= \mathbf{D}_c^{-1} \delta \sigma$$

The total strain rate is

$$\delta\dot{\epsilon} = \mathbf{D}_c^{-1} \delta\dot{\sigma} + \mathbf{D}^{-1} \delta\dot{\sigma} \tag{17}$$

If we assume that the given perturbation decays according to exp $(-\lambda t)$:

$$\delta\dot{\epsilon} = [\mathbf{D}_c^{-1} - \lambda \mathbf{D}^{-1}] \delta\dot{\sigma} \tag{18}$$

We can now construct an energy of perturbations:

$$\tfrac{1}{2} \int \delta\dot{\sigma}^T \delta\dot{\epsilon} \, d(\text{vol}) = \tfrac{1}{2} \int \delta\dot{\sigma}^T [\mathbf{D}_c^{-1} - \lambda \mathbf{D}^{-1}] \delta\dot{\sigma} \, d(\text{vol}) \tag{19}$$

whose derivative with respect to any relevant nodal variable gives an unbalanced force. This is zero, and defines the eigenvalues which control the convergence.

As before, these eigenvalues are bounded by the extreme eigenvalues of the infinitesimal elements:

$$\mathbf{D}_c^{-1} \delta\dot{\sigma} = \lambda \mathbf{D}^{-1} \delta\dot{\sigma} \tag{20}$$

$$\text{or} \quad \mathbf{D} \delta\dot{\epsilon} = \lambda \mathbf{D}_c \delta\dot{\epsilon} \tag{21}$$

and hence $\delta t < 2/\lambda_{max}$.

In the isotropic case the criterion was used automatically in the programs of Reference 6. Instability was successfully avoided, and the indications were that the judgments were somewhat pessimistic throughout, but not excessively so.

CONCLUSION

We have discussed the applications of an easily taught and easily remembered theorem. Some of its applications are exceptionally useful. Therefore the theorem should be taught in finite element courses.

REFERENCES

1. Irons, B. M., "Roundoff Criteria in Direct Stiffness Solutions," AIAA Journal, Vol. 6, pp 1308–1312, July 1968.

2. Khanna, J., and Hooley, R. F., "Comparison and Evaluation of Stiffness Matrices", AIAA Journal, Vol. 4, pp 2105–2111, December 1966.

3. Leech, J. W., Pao-Tan Hsu, and Mack, E. W., "Stability of a Finite-Difference Method for Solving Matrix Equations", AIAA Journal, Vol. 3, pp 2172–2173, November 1965.

4. Dupuis, G. A., Pfaffinger, D. D., and Marcal, P. V., "Effective Use of The Incremental Stiffness Matrices in Nonlinear Geometric Analysis", IUTAM Conference on High-Speed Computing of Elastic Structures, University of Liege, Belgium, August 1970.

5. Gallagher, R. H., and Marcal, P. V., (to be published).

6. Treharne, G., "Applications of The Finite Element Method to The Stress Analysis of Materials Subject to Creep", Ph.D. Dissertation, University of Wales, 1971.

7. Argyris, J. H., and Scharpf, D. W., "Methods of Elastoplastic Analysis", Symposium on Finite Element Techniques, University of Stuttgart, June 1969.

8. Kan, D. K. Y., "Solutions Techniques for Large Finite Element Problems Concerning Creep and Temperature Transients in Gas Turbine Blades", Ph.D. Dissertation, University of Wales, 1971, (to be submitted).

9. Tong, P., paper presented at University of Waterloo in conference on finite elements, March 1971.

10. Nayak, G. C., "Plasticity and Large Deformation Problems by the Finite Element Method", Ph.D. Dissertation, University of Wales (Swansea), 1971.

11. Zienkiewicz, O. C., "The Finite Element Method in Engineering Science", McGraw-Hill, 1971.

ON DERIVATION OF STIFFNESS MATRICES WITH C° ROTATION FIELDS FOR PLATES AND SHELLS

Senol Utku*
Duke University

The use of C° deflection (displacements and rotations) fields over triangular meshes in the middle plane of a plate for the total potential energy functional always requires the retention of the transverse shear strain energy in order to have the stiffness matrix of the supported structure non-singular when the degrees of freedom in the transverse direction are not suppressed. This situation is similar to the case of sandwich plates and shells where the transverse shear strain energy is retained because of its importance. Since the transverse shear strain energy is a function of transverse shear moduli, rather than in-plane shear modulus, it should be computed directly from the deflection fields, but not from the corresponding stress couple fields. This paper shows the method of computation of transverse shear strain energy directly from the deflection fields. Applying the method to C° deflection fields, a mathematical foundation is laid for a widely used plate and shell element by Martin, Melosh and Utku.

Introduction

After the celebrated work of Turner, Clough, Martin and Topp[1] in 1956 for solving two dimensional Elasticity problems with C° trial displacement fields and the Ritz procedure, it took about ten years to do a similar but not exactly the same thing for the plate bending problems. In 1965, Melosh reported[2] that C° displacement fields for plate bending problems were of very poor convergence, and suggested the use of hypothetical spar beams of mysteriously adjusted stiffness around the periphery of triangular elements to represent the transverse shear rigidity, and ignore the one resulting from the mathematical procedure. He mentions[2] successful and monotonically converging solutions of several plate bending problems by using the spar beam concept and the right

* Associate Professor of Civil Engineering

triangular elements. In 1966, Utku gave[3] a mathematical derivation of stiffness matrices for thin plates and shells using C° deflection (the term deflection in this paper is used to mean both displacements and rotations) fields over arbitrary triangular meshes with or without curvature. The shell curvatures are handled by approximating the middle surface section within a triangle by a parabolic surface, and the slow convergence is overcome by replacing the portion of the transverse shear stiffness matrix of an element, which is associated with the rotations, by quantities insuring equilibrium among the forces represented by the columns of the final transverse shear stiffness matrix (although there are infinitely many ways of generating such quantities, the one named the "equilibrium algorithm" is arrived at by analogy to the behavior of clamped beams). The transverse shear stiffness matrix obtained by the "equilibrium algorithm" is identical with that of Ref. (2). The monotonic and rapid convergence characteristics of this element with various meshes are demonstrated in Ref. (3) for both plates and shells. In 1967, Utku and Melosh reported[4] that when the triangle is obtuse, the transverse shear matrix of Refs. (2) and (3) becomes indefinite, and the curved element stiffness matrix of Ref. (3) becomes indefinite when the apex of the paraboloid approximating the shell middle surface is not on the normal at the centroid of the base triangle. To prevent the indefiniteness of the transverse shear stiffness matrix, they suggested[4] to modify the transverse force portion of the matrix arbitrarily by a process they named the "constant trace scheme". In Ref. (4), extensive results are given to demonstrate the monotonic and rapid convergence of the resulting matrix in both plates and shells. In 1967, Martin arrived[5] at the same transverse shear stiffness matrix of Refs. (2) and (3) by a physical model of cover plates and spar beams with a sound engineering reasoning. However, his matrix is also indefinite when the triangle is obtuse.

Today, several computer programs, such as ELAS[6], SAMIS[7], and DYNAL[8], use the flat shell element of Ref. (4) which uses C° trial deflection fields in conjunction with the "equilibrium algorithm" and the "constant trace scheme". Because of the latter modifications, of course, one is never sure of the true class of the trial fields yielding the final stiffness matrix. In spite of this weakness, this element 1) has been the first conforming element for arbitrary triangulation in both plates and shells, 2) is still the only available element which represents the behavior of non-Kirchhoffian plates and shells by using the transverse shear moduli (rather than in-plane shear modulus), and therefore 3) is the only element suitable for sandwich plates and shells.

This work is an attempt to furnish the mathematical foundations of this widely used plate and shell element by eliminating from its definition all physical analogies and engineering approximations. For this purpose, first a method is given for the computation of transverse shear strain energy directly from the deflection fields. Then, this method is applied to C° deflection fields over triangular meshes to explain the behavior of the Martin-Melosh-Utku element.

Finally the behavior of the new element is compared with that of the Martin-Melosh-Utku element.

Computation of Transverse Shear Strain Energy Directly from the Deflection Fields

Let x,y,z denote a right-handed Cartesian coordinate system located at the centroid of a triangular thin shell element of thickness t. Let 1,2,3 denote the vertex labels of the triangle. It is assumed that the element is sufficiently small to justify the assumption that vertices 1,2 and 3 lie in the xy plane. Let $\underline{i}$, $\underline{j}$, $\underline{k}$ denote the unit vectors of the coordinate system (see Fig. 1). Let

$$\underline{d} = u\,\underline{i} + v\,\underline{j} + w\,\underline{k} \tag{1}$$

denote the displacement vector. It is assumed that the transverse shear strains $[\gamma] = [\gamma_{xz}\ \gamma_{yz}]$ and the transverse shear stresses $[\tau] = [\tau_{xz}\tau_{yz}]$ are related with

$$\begin{Bmatrix}\tau_{xz}\\ \tau_{yz}\end{Bmatrix} = \begin{bmatrix}D'_{11} & D'_{12}\\ D'_{21} & D'_{22}\end{bmatrix}\begin{Bmatrix}\gamma_{xz}\\ \gamma_{yz}\end{Bmatrix}$$

which may be rewritten as

$$\{\tau\} = [D']\{\gamma\} \tag{2}$$

where [D'] is as displayed above. The transverse shear moduli matrix [D'] is positive definite and symmetric. The linear strain displacement relations are

$$\begin{Bmatrix}\gamma_{xz}\\ \gamma_{yz}\end{Bmatrix} = \begin{Bmatrix}u_{,z} + w_{,x}\\ v_{,z} + w_{,y}\end{Bmatrix} \tag{3}$$

where a comma in the subscript indicates partial differentiation with respect to the quantity following the comma. Since the thin shell or plate assumptions imply that $u_{,z}$ = const., $v_{,z}$ = const., and $w_{,z} = 0$, the rotations of middle surface normals about x and y axes, θ_x and θ_y, may be expressed as

$$\begin{Bmatrix}\theta_x\\ \theta_y\end{Bmatrix} = \begin{Bmatrix}-v_{,z}\\ u_{,z}\end{Bmatrix} \tag{4}$$

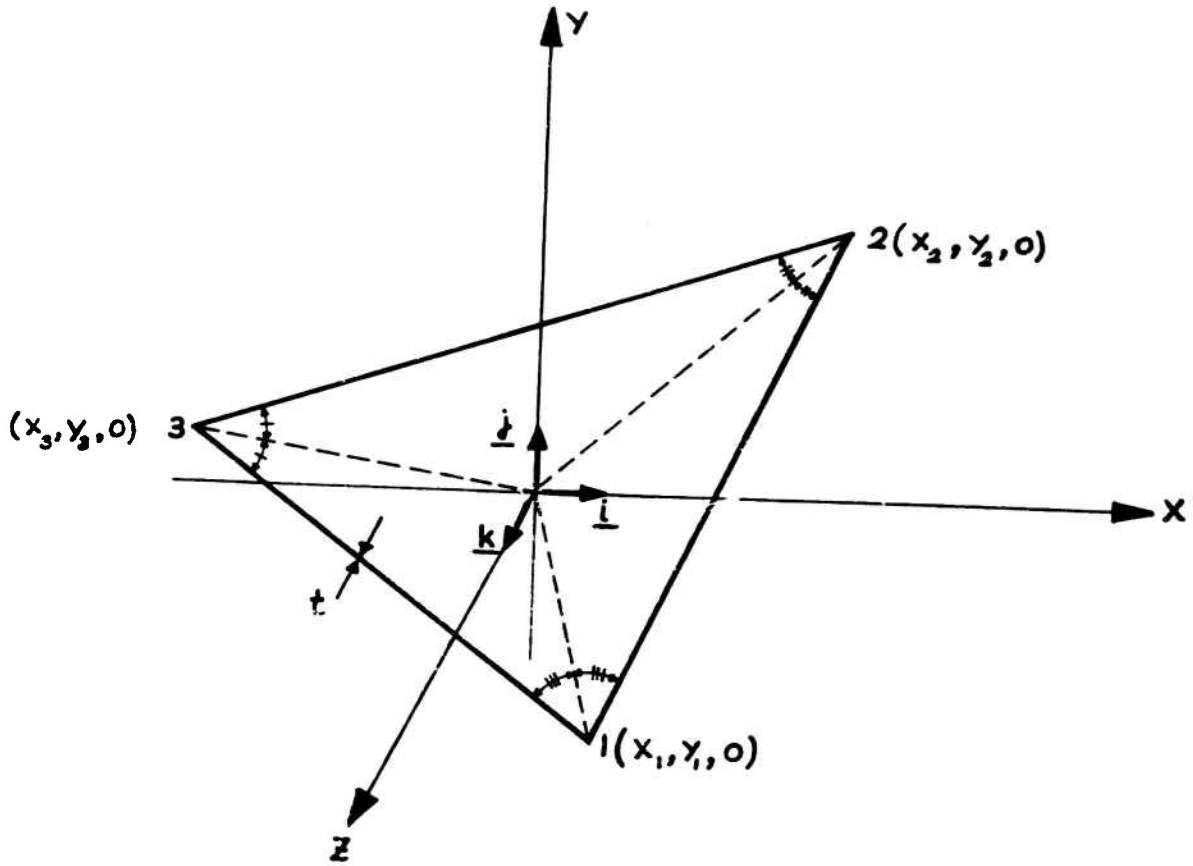


Figure 1. Definition Sketch

Substituting θ_x and θ_y from (4) into (3) one obtains

$$\begin{Bmatrix} \gamma_{xz} \\ \gamma_{yz} \end{Bmatrix} = \begin{Bmatrix} \theta_y + w_{,x} \\ -\theta_x + w_{,y} \end{Bmatrix} \tag{5}$$

If the behavior is truly Kirchhoffian $\gamma_{xz} = \gamma_{yz} = 0$. Denoting the transverse displacement with w' in the case of Kirchhoffian behavior, one may write from (5)

$$\begin{Bmatrix} \theta_x \\ \theta_y \end{Bmatrix} = \begin{Bmatrix} w'_{,y} \\ -w'_{,x} \end{Bmatrix} \tag{6}$$

Let the difference between non-Kirchhoffian and Kirchhoffian transverse displacements be denoted by w* such that

$$w^* = w - w' \tag{7}$$

Using (6) and (7), one may rewrite (5) as

$$\{\gamma\} = \begin{Bmatrix} \gamma_{xz} \\ \gamma_{yz} \end{Bmatrix} = \begin{Bmatrix} w^*_{,x} \\ w^*_{,y} \end{Bmatrix} \tag{8}$$

Since the strain energy density for transverse shear is $\frac{1}{2}[\gamma][D']\{\gamma\}$ using (8), the transverse shear strain energy of the triangle may be expressed as

$$U_S = \frac{1}{2} \int_A [w^*_{,x} \; w^*_{,y}] \begin{bmatrix} D'_{11} & D'_{12} \\ D'_{21} & D'_{22} \end{bmatrix} \begin{Bmatrix} w^*_{,x} \\ w^*_{,y} \end{Bmatrix} dA \tag{9}$$

where dA is the area element, and A is the area of the triangle. If one uses (9) for the transverse shear strain energy in the total potential energy functional, for monotonic convergence, it is necessary that w* is at least of C^0 over the triangular mesh. In other words, one should have w* (but not its derivatives) continuous across the interelement boundaries.

Triangular Element with C° Deflection Fields for Non-Kirchhoffian Plates or Shells

Let $x_i, y_i, z_i, u_i, v_i, w_i, \theta_{xi}, \theta_{yi}$, and θ_{zi}, i=1,2,3 denote the values of Cartesian coordinates and the deflection components at the vertices of the triangle as referred to the x,y,z coordinate system (see Fig. 1). If one assumes that, excluding w and θ_z the deflection components vary linearly in the triangle with x, and y, the values of these components at a point with coordinates x, and y in the triangle may be expressed by linear interpolation as

$$[u \quad v \quad \theta_x \quad \theta_y] = \frac{1}{2A} [x \quad y \quad 1] \begin{bmatrix} -y- \\ -x- \\ -r- \end{bmatrix} \begin{bmatrix} | & | & | & | \\ \bar{u} & \bar{v} & \bar{\theta}_x & \bar{\theta}_y \\ | & | & | & | \end{bmatrix} \tag{10}$$

where $[-y-] = [y_2-y_3 \quad y_3-y_1 \quad y_1-y_2]$, $[-x-] = [x_3-x_2 \quad x_1-x_3 \quad x_2-x_1]$,

$[-r-] = [x_2y_3-x_3y_2 \quad x_3y_1-x_1y_3 \quad x_1y_2-x_2y_1]$, $[-\bar{u}-] = [u_1 \quad u_2 \quad u_3]$,

$[-\bar{v}-] = [v_1 \quad v_2 \quad v_3]$, $[-\bar{\theta}_x-] = [\theta_{x1} \quad \theta_{x2} \quad \theta_{x3}]$, and

$[-\bar{\theta}_y-] = [\theta_{y1} \quad \theta_{y2} \quad \theta_{y3}]$. Now if one assumes that similar linear interpolation applies in all the triangles of the whole triangular mesh, the resulting deflection field is of C°, since the deflection components, but not their derivatives, are continuous across the interelement boundaries.

The total strain energy associated with the triangle, U, may be written as the sum of the membrane, bending, and transverse shear strain energies U_M, U_B, and U_S:

$$U = U_M + U_B + U_S \tag{11}$$

and U_M, U_B, and U_S may be expressed as

$$U_M = \frac{1}{2} [\, q \,][K_M]\{q\} \tag{12a}$$

$$U_B = \frac{1}{2} [\, q \,][K_B]\{q\} \tag{12b}$$

$$U_S = \frac{1}{2} [\, q \,][K_S]\{q\} \tag{12c}$$

where $[K_M]$, $[K_B]$, and $[K_S]$ are the membrane, the bending and the transverse shear stiffness matrices of the triangular element, and $[q] = [u_1\ u_2\ u_3 \mid v_1\ v_2\ v_3 \mid w_1\ w_2\ w_3 \mid \theta_{x1}\ \theta_{x2}\ \theta_{x3} \mid \theta_{y1}\ \theta_{y2}\ \theta_{y3}]$.

Let the tangential strains $[\varepsilon] = [\varepsilon_x\ \ \varepsilon_y\ \ \gamma_{xy}]$ and the tangential stresses $[\sigma] = [\sigma_x \sigma_y \tau_{xy}]$ be related with

$$\{\sigma\} = [D]\{\varepsilon\} \tag{13}$$

where the material matrix [D] is positive definite and symmetric. Using the linear interpolation rule of (10), and stress-strain relations (13), $[K_M]$ and $[K_B]$ may be computed[3] as

$$[K_M] = \frac{t}{4A}\begin{bmatrix} P & R & 0 & 0 & 0 \\ R^T & Q & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad [K_B] = \frac{t^3}{48A}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Q & -R^T \\ 0 & 0 & 0 & -R & P \end{bmatrix} \tag{14a,b}$$

where

$$[P] = [M]^T[D][M]\ , \quad [R] = [M]^T[D][N]\ , \quad [Q] = [N]^T[D][N] \tag{14c,d,e}$$

and

$$[M] = \begin{bmatrix} -y\rightarrow \\ -0\rightarrow \\ -x\rightarrow \end{bmatrix}\ , \quad [N] = \begin{bmatrix} -0\rightarrow \\ -x\rightarrow \\ -y\rightarrow \end{bmatrix}\ , \quad [-0\rightarrow] = [0\ 0\ 0], \quad [0] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{14f,g,h,i}$$

In (14a) the contributions of local curvatures, if there are any, are ignored.

In order to compute U_S in terms of the vertex deflections using (9), one needs w' or its first derivatives as implied by (7). A consistent w' distribution in the triangle with the linear rotation distributions may be expressed as

$$w' = [x^2\ \ y^2\ \ xy\ \ x\ \ y\ \ 1]\begin{Bmatrix} a_1 \\ a_2 \\ \vdots \\ a_6 \end{Bmatrix} \tag{15}$$

where the coefficients a_i, i=1,....5 may be expressed in terms of the vertex rotations as follows: Using (15) in (6) one may write

$$\theta_x = [2y \quad x \quad 1]\begin{Bmatrix} a_2 \\ a_3 \\ a_5 \end{Bmatrix} \tag{16a}$$

and

$$\theta_y = -[2x \quad y \quad 1]\begin{Bmatrix} a_1 \\ a_3 \\ a_4 \end{Bmatrix} \tag{16b}$$

and evaluating these at the vertices, after inversion one obtains:

$$\begin{Bmatrix} a_2 \\ a_3 \\ a_5 \end{Bmatrix} = \frac{1}{4A}\begin{bmatrix} \longrightarrow x \longrightarrow \\ (2) \longrightarrow y \longrightarrow \\ 4A/3 \longrightarrow 1 \longrightarrow \end{bmatrix}\begin{Bmatrix} \bar{\theta} \\ \theta_x \end{Bmatrix} \tag{17a}$$

and

$$\begin{Bmatrix} a_1 \\ a_3 \\ a_4 \end{Bmatrix} = -\frac{1}{4A}\begin{bmatrix} \longrightarrow y \longrightarrow \\ (2) \longrightarrow x \longrightarrow \\ 4A/3 \longrightarrow 1 \longrightarrow \end{bmatrix}\begin{Bmatrix} \bar{\theta} \\ \theta_y \end{Bmatrix} \tag{17b}$$

where $[\longrightarrow 1 \longrightarrow] = [111]$. It is observed that a_3 is computed in two different ways in (17a) and (17b). To assure single valuedness, the average of the two may be taken as a_3, so that

$$\begin{Bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{Bmatrix} = \frac{1}{4A}\left[\begin{array}{cc} \longrightarrow 0 \longrightarrow & \longrightarrow y \longrightarrow \\ \longrightarrow x \longrightarrow & \longrightarrow 0 \longrightarrow \\ \longrightarrow y \longrightarrow & \longrightarrow x \longrightarrow \\ \longrightarrow 0 \longrightarrow & 4A/3 \longrightarrow 1 \longrightarrow \\ 4A/3 \longrightarrow 1 \longrightarrow & \longrightarrow 0 \longrightarrow \end{array}\right]\left\{\begin{array}{c} \bar{\theta} \\ \theta_x \\ -\theta_y \end{array}\right\} \tag{18}$$

Calling the value of w' at the origin w'_c and substituting a_i, i=1...,5, from (18) into (15), one may write

$$w'-w'_o = \frac{1}{4A}[x^2 \quad y^2 \quad xy \quad x \quad y]\left[\begin{array}{cc} \longrightarrow 0 \longrightarrow & \longrightarrow y \longrightarrow \\ \longrightarrow x \longrightarrow & \longrightarrow 0 \longrightarrow \\ \longrightarrow y \longrightarrow & \longrightarrow x \longrightarrow \\ \longrightarrow 0 \longrightarrow & 4A/3 \longrightarrow 1 \longrightarrow \\ 4A/3 \longrightarrow 1 \longrightarrow & \longrightarrow 0 \longrightarrow \end{array}\right]\left\{\begin{array}{c} \bar{\theta} \\ \theta_x \\ -\theta_y \end{array}\right\} \tag{19}$$

By evaluating w' at the vertices, one obtains from (19)

$$\begin{Bmatrix} w'_1-w'_o \\ w'_2-w'_o \\ w'_3-w'_o \end{Bmatrix} = \frac{1}{4A}\begin{bmatrix} x_1^2 & y_1^2 & x_1y_1 & x_1 & y_1 \\ x_2^2 & y_2^2 & x_2y_2 & x_2 & y_2 \\ x_3^2 & y_3^2 & x_3y_3 & x_3 & y_3 \end{bmatrix}\left[\begin{array}{cc} \longrightarrow 0 \longrightarrow & \longrightarrow y \longrightarrow \\ \longrightarrow x \longrightarrow & \longrightarrow 0 \longrightarrow \\ \longrightarrow y \longrightarrow & \longrightarrow x \longrightarrow \\ \longrightarrow 0 \longrightarrow & 4A/3 \longrightarrow 1 \longrightarrow \\ 4A/3 \longrightarrow 1 \longrightarrow & \longrightarrow 0 \longrightarrow \end{array}\right]\left\{\begin{array}{c} \bar{\theta} \\ \theta_x \\ -\theta_y \end{array}\right\}$$

which may be reduced to

$$\begin{Bmatrix} w_1'-w_o' \\ w_2'-w_o' \\ w_3'-w_o' \end{Bmatrix} = \frac{1}{6}\begin{bmatrix} y_1 & & \\ & y_2 & \\ & & y_3 \end{bmatrix}\begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix}\begin{Bmatrix} \\ \bar{\theta}_x \\ \\ \downarrow \end{Bmatrix} - \frac{1}{6}\begin{bmatrix} x_1 & & \\ & x_2 & \\ & & x_3 \end{bmatrix}\begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix}\begin{Bmatrix} \\ \bar{\theta}_y \\ \\ \downarrow \end{Bmatrix} \qquad (20)$$

By using (7), this leads to

$$\begin{Bmatrix} \\ \bar{w}^* \\ \\ \downarrow \end{Bmatrix} = \begin{Bmatrix} \\ \bar{w} \\ \\ \downarrow \end{Bmatrix} - \frac{1}{6}\begin{bmatrix} y_1 & & \\ & y_2 & \\ & & y_3 \end{bmatrix}\begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix}\begin{Bmatrix} \\ \bar{\theta}_x \\ \\ \downarrow \end{Bmatrix} + \frac{1}{6}\begin{bmatrix} x_1 & & \\ & x_2 & \\ & & x_3 \end{bmatrix}\begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix}\begin{Bmatrix} \\ \bar{\theta}_y \\ \\ \downarrow \end{Bmatrix} - w_o\begin{Bmatrix} \\ 1 \\ \\ \downarrow \end{Bmatrix} \qquad (21)$$

where $[\!\!-\bar{w}^*\!\!\rightarrow] = [w_1^* \;\; w_2^* \;\; w_3^*]$ and $[\!\!-\bar{w}\!\!\rightarrow] = [w_1 \;\; w_2 \;\; w_3]$.

In order to be consistent with the interpolation rule used in obtaining the membrane and the bending stiffness matrices, one may try a C° field for w^*. This means that the difference between the Kirchhoffian and the non-Kirchhoffian transverse displacements is assumed varying linearly within a triangle. By using the linear interpolation rule of (10), one may write

$$w^* = \frac{1}{2A}[x \;\; y \;\; 1]\begin{bmatrix} -y- \\ -x- \\ -r- \end{bmatrix}\begin{Bmatrix} \\ \bar{w}^* \\ \\ \downarrow \end{Bmatrix} \qquad (22)$$

Substituting w^* from (22) into (8), and using (21) one obtains

$$\{\gamma\} = \frac{1}{2A}\begin{bmatrix} -y- \\ -x- \end{bmatrix}\left[\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \vdots -\frac{1}{6}\begin{bmatrix} y_1 & & \\ & y_2 & \\ & & y_3 \end{bmatrix}\begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix} \vdots \frac{1}{6}\begin{bmatrix} x_1 & & \\ & x_2 & \\ & & x_3 \end{bmatrix}\begin{bmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{bmatrix}\right]\begin{Bmatrix} \bar{w} \\ \cdots \\ \bar{\theta}_x \\ \bar{\theta}_y \end{Bmatrix} \qquad (23a)$$

which can be reduced into

$$\{\gamma\} = \frac{1}{2A}\left(\frac{1}{2}\begin{bmatrix} -y- \\ -x- \end{bmatrix}\left[\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \vdots \begin{matrix} -y_1 & -y_2 & -y_3 \end{matrix} \vdots \begin{matrix} x_1 & x_2 & x_3 \end{matrix}\right] + \frac{1}{2}\left[\begin{array}{c:c:c} -y- & -0- & \frac{2A}{3}-1- \\ -x- & -\frac{2A}{3}-1- & -0- \end{array}\right]\right)\begin{Bmatrix} \bar{w} \\ \cdots \\ \bar{\theta}_x \\ \bar{\theta}_y \end{Bmatrix} \qquad (23b)$$

Note that the transverse shear strains in (23b) are expressed as the arithmetic average of two terms. The approximation of the transverse shear strains by only the second term has been studied previously under the name of "average rotations algorithm"[4].

The use of $\{\gamma\}$ from (23b) in (9) gives the transverse shear stiffness matrix $[K_s]$ of (12c) as

$$[K_S] = \frac{t}{4A}\begin{bmatrix} 0 & & & & \\ 0 & 0 & & & \\ 0 & 0 & S_{11} & \text{Sym.} & \\ 0 & 0 & S_{21} & S_{22} & \\ 0 & 0 & S_{31} & S_{32} & S_{33} \end{bmatrix} \tag{24a}$$

where

$$[S_{11}] = \begin{bmatrix} | & | \\ y & x \\ \downarrow & \downarrow \end{bmatrix}[D']\begin{bmatrix} \text{—}y\text{→} \\ \text{—}x\text{→} \end{bmatrix}, \quad [S_{21}] = -\frac{1}{2}\begin{bmatrix} y_1 y_2 y_3 \end{bmatrix}[S_{11}], \tag{24b,c}$$

$$[S_{31}] = \frac{1}{2}\begin{bmatrix} x_1 x_2 x_3 \end{bmatrix}[S_{11}], \quad [S_{22}] = \frac{1}{4}\begin{bmatrix} y_1 y_2 y_3 \end{bmatrix}[S_{11}]\begin{bmatrix} y_1 y_2 y_3 \end{bmatrix}, \tag{24d,e}$$

and

$$[S_{32}] = -\frac{1}{4}\begin{bmatrix} x_1 x_2 x_3 \end{bmatrix}[S_{11}]\begin{bmatrix} y_1 y_2 y_3 \end{bmatrix}, \quad [S_{33}] = \frac{1}{4}\begin{bmatrix} x_1 x_2 x_3 \end{bmatrix}[S_{11}]\begin{bmatrix} x_1 x_2 x_3 \end{bmatrix} \tag{24f,g}$$

The total stiffness matrix of the element may now be written as the sum of $[K_M]$, $[K_B]$ and $[K_S]$:

$$[K] = \frac{t}{4A}\begin{bmatrix} P & & & & \\ R^T & Q & & & \\ 0 & 0 & S_{11} & \text{Sym.} & \\ 0 & 0 & S_{21} & \beta Q + S_{22} & \\ 0 & 0 & S_{31} & -\beta R + S_{32} & \beta P + S_{33} \end{bmatrix}, \quad \beta = \frac{t^2}{12} \tag{25}$$

which may be used in

$$U = \frac{1}{2}\{q\}[K]\{q\} \tag{26}$$

for the computation of the total strain energy of the triangular plate or shell element in terms of the vertex deflections.

Comparison of the New Element with the Martin-Melosh-Utku Element

The membrane and the bending portions of the new element stiffness matrix are identical with those of the Martin-Melosh-Utku element. Moreover, $[S_{11}]$ submatrix in both cases are the same. The difference exists in submatrices $[S_{21}]$, $[S_{31}]$, $[S_{22}]$, $[S_{32}]$, and $[S_{33}]$ of $[K_S]$. Using superscript e to distinguish quantities belonging to the Martin-Melosh-Utku element (e stands for the "equilibrium algorithm"), for non-obtuse triangles, one may observe that

$$[S^e_{11}] = [S_{11}] \tag{27a}$$

$$[S^e_{21}] = [S_{21}] + \frac{1}{6}\begin{Bmatrix} | \\ y \end{Bmatrix}[a_1\ b_1\ c_1] \ , \tag{27b}$$

$$[S^e_{31}] = [S_{31}] - \frac{1}{6}\begin{Bmatrix} | \\ x \end{Bmatrix}[a_1\ b_1\ c_1] \ , \tag{27c}$$

$$[S^e_{22}] = [S_{22}] + \frac{1}{6}\begin{Bmatrix} | \\ y \end{Bmatrix}[a_2\ b_2\ c_2] \ , \tag{27d}$$

$$[S^e_{32}] = [S_{32}] - \frac{1}{6}\begin{Bmatrix} | \\ x \end{Bmatrix}[a_2\ b_2\ c_2] \ , \tag{27e}$$

$$[S^e_{33}] = [S_{33}] - \frac{1}{6}\begin{Bmatrix} | \\ x \end{Bmatrix}[a_3\ b_3\ c_3] \ , \tag{27f}$$

where

$$\begin{aligned} a_i &= (S^e_{i1})_{31} - (S^e_{i1})_{21} \ , \\ b_i &= (S^e_{i1})_{21} - (S^e_{i1})_{32} \ , \qquad i = 1,2,3 \\ c_i &= (S^e_{i1})_{32} - (S^e_{i1})_{31} \ , \end{aligned} \tag{27g}$$

the second pair of subscripts indicating the row and the column number of the entry of the submatrix to be used.

The three scalars $(S^e_{i1})_{21}$, $(S^e_{i1})_{31}$ and $(S^e_{i1})_{32}$ of (27g) are the three off-diagonal elements in the lower part of submatrix $[S_{i1}]$, i=1,2,3. For i=1, it has been shown[4] that for non-obtuse triangles these three scalars are all non-positive, and for obtuse triangles one of them is positive (for obtuse triangles the Martin-Melosh-

Utku element stiffness matrix without the "constant trace scheme" modification is indefinite). For equilateral triangles, when i=1, these three scalars are all equal, implying that $a_1 = b_1 = c_1$, and therefore $[S^e_{21}]=[S_{21}]$ and $[S^e_{31}]=[S_{31}]$. In the case of equilateral triangle, in order to have also $[S^e_{22}] = [S_{22}]$, $[S^e_{32}] = [S_{32}]$ and $[S^e_{33}] = [S_{33}]$, the orientation of the triangle relative to the x,y,z coordinate system should also be special.

The comparisons of the results by the new element and by the Martin-Melosh-Utku element for several plate problems with various meshes under a single transverse concentrated load are given in Figures 2-8, where the transverse displacement under the concentrated load load is plotted against some measure of the mesh refinement (the larger the abscissa, the finer the mesh). The results with equilateral and bilateral triangular meshes indicate that the new element converges faster (Figs. 2 and 3). In the case of right bilateral triangles the two results are almost coincident (Fig. 4). For the case of obtuse triangles, the slightly larger results by the Martin-Melosh-Utku element with the "constant trace scheme" are of the same convergence rate with those of the new element (Figs. 5 and 6).

In Figs. 7 and 8, the better convergence characteristic of the Martin-Melosh-Utku element is obvious. The paradox between the behavior shown in Figs. 2-6, and Figs. 7-8 clearly indicates that the Martin-Melosh-Utku element is orientation dependent. It has a much higher convergence rate if the largest edge of a triangle is not quite coincident with the minimum bending direction. Presently, the mathematical explanation of this quite remarkable behavior is not complete for formal presentation.

Conclusions

From the study presented in this paper, the following conclusions may be drawn:

1. The transverse shear strain energy is a function of the quantity w*which represents the difference between the Kirchhoffian and the non-Kirchhoffian transverse displacements.

2. Conforming element representations for the non-Kirchhoffian behavior of thin plates and shells are possible by C° rotation, tangential displacement and w* fields, as shown in this work.

3. The stiffness matrices with C° deflection fields as derived in this work closely follow the behavior of the Martin-Melosh-Utku element in general.

4. The convergence characteristics of the Martin-Melosh-Utku element can be substantially improved by properly orienting the element.

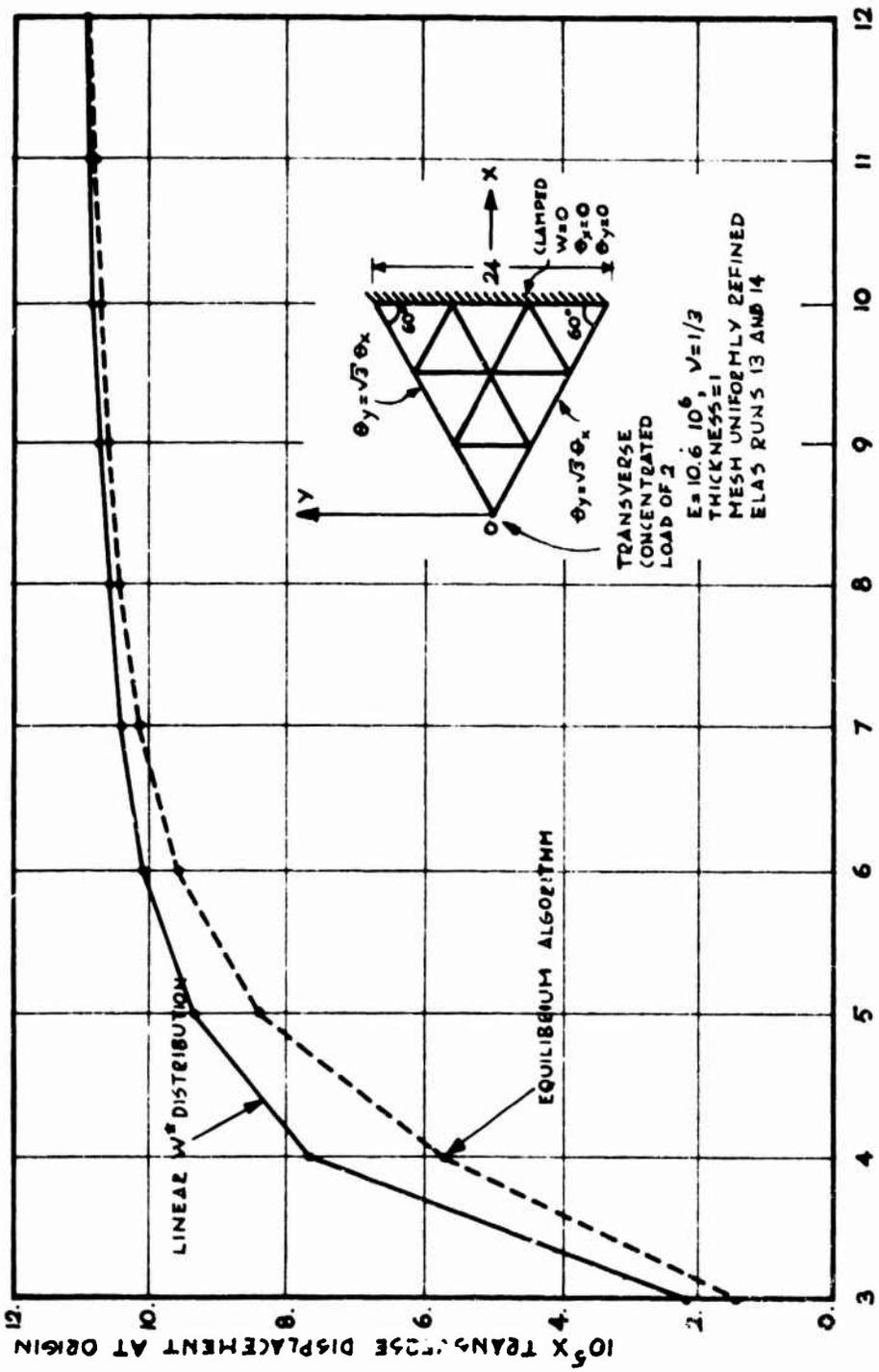


Figure 2. Convergence in Equilateral Triangular Mesh

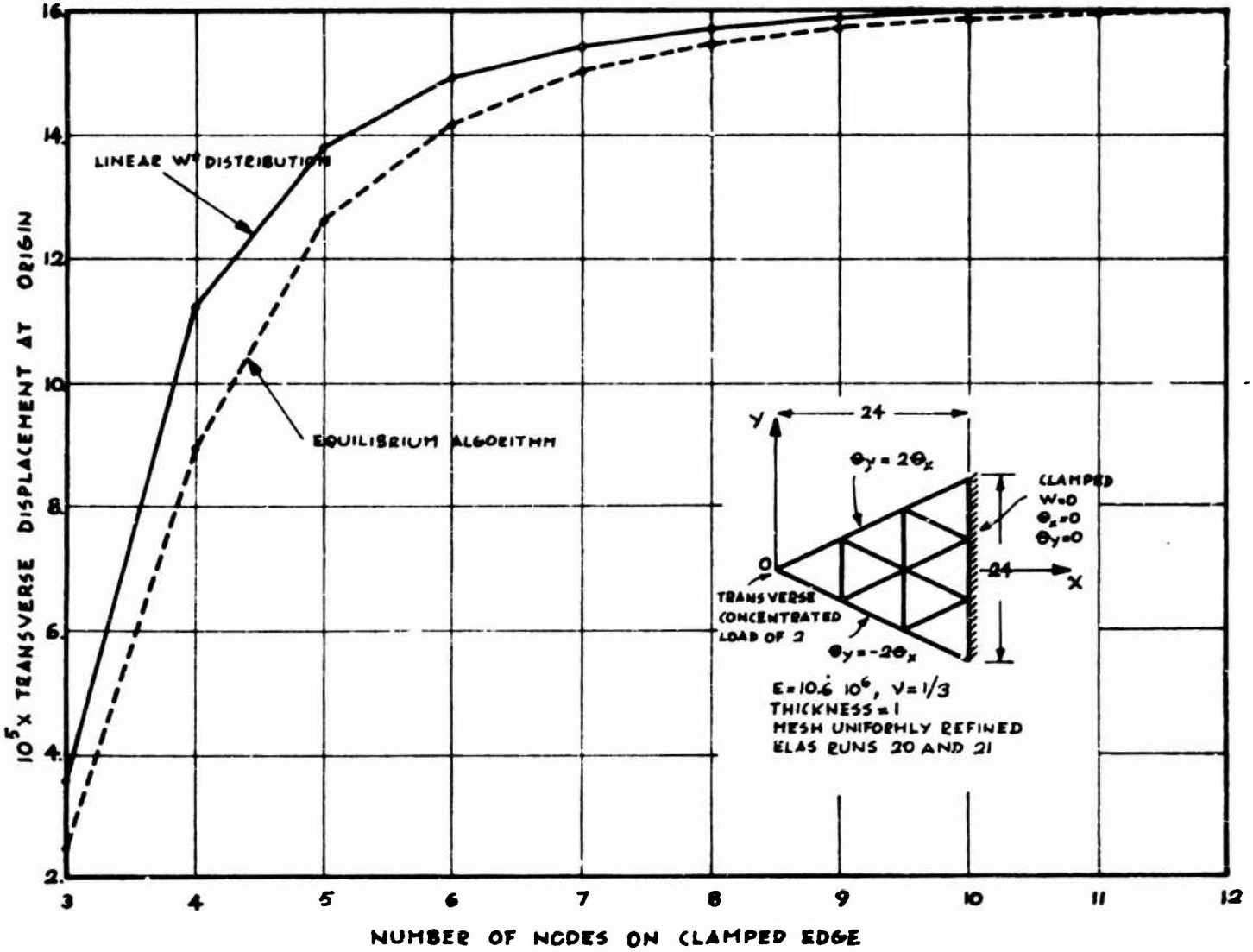


Figure 3. Convergence in Bilateral Triangular Mesh

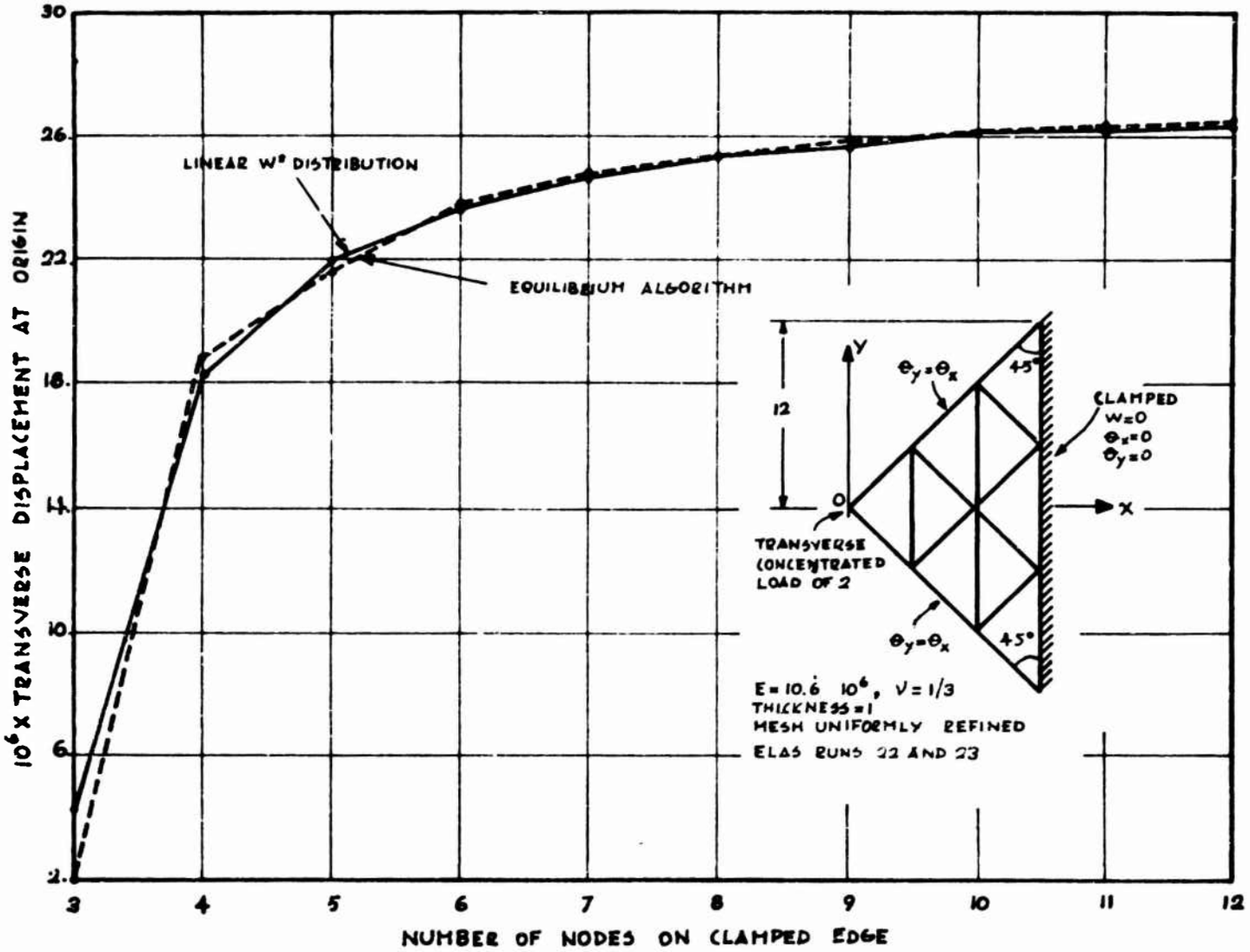


Figure 4. Convergence in Right Bilateral Triangular Mesh

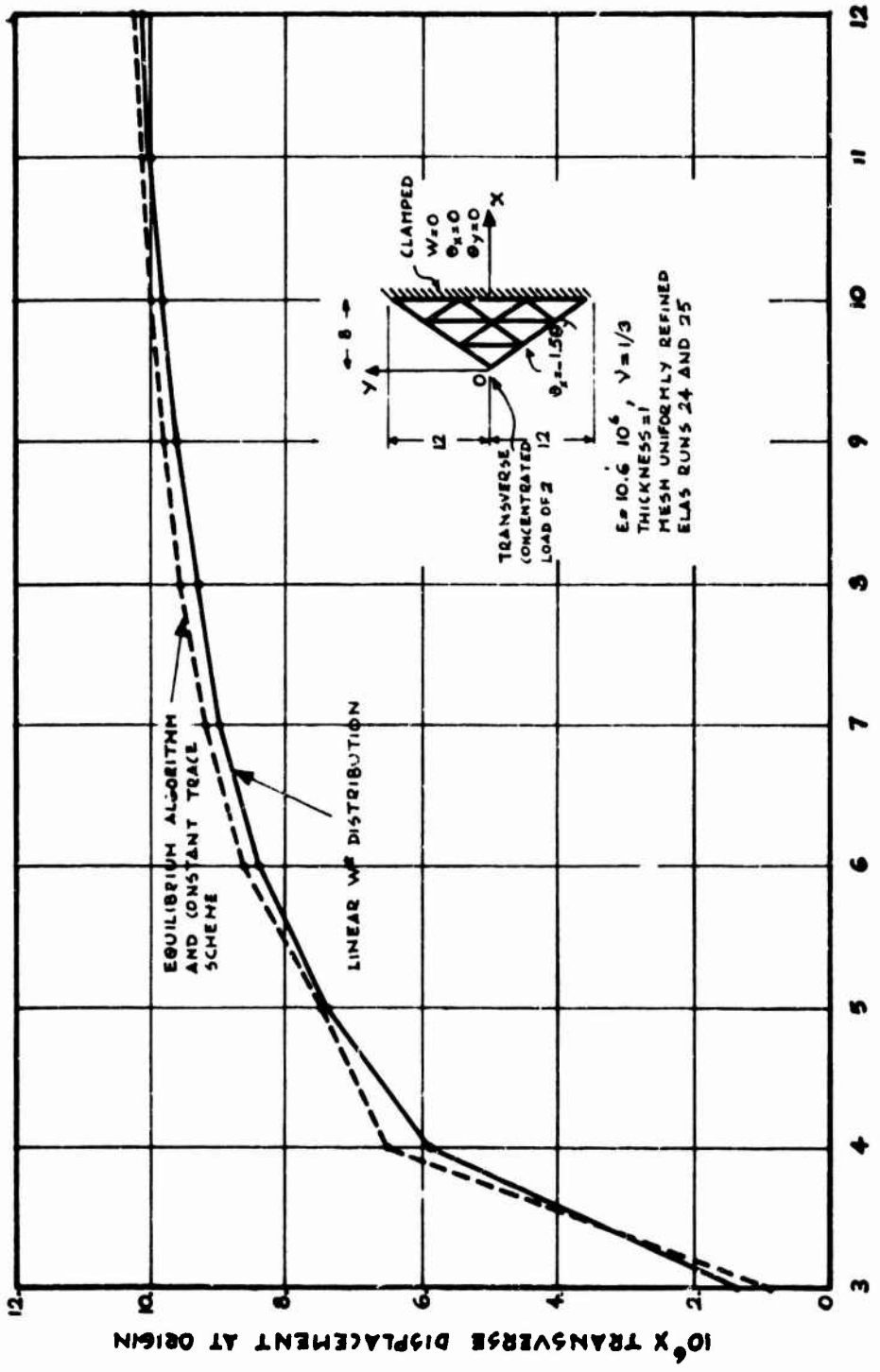


Figure 5. Convergence in Obtuse Bilateral Triangular Mesh

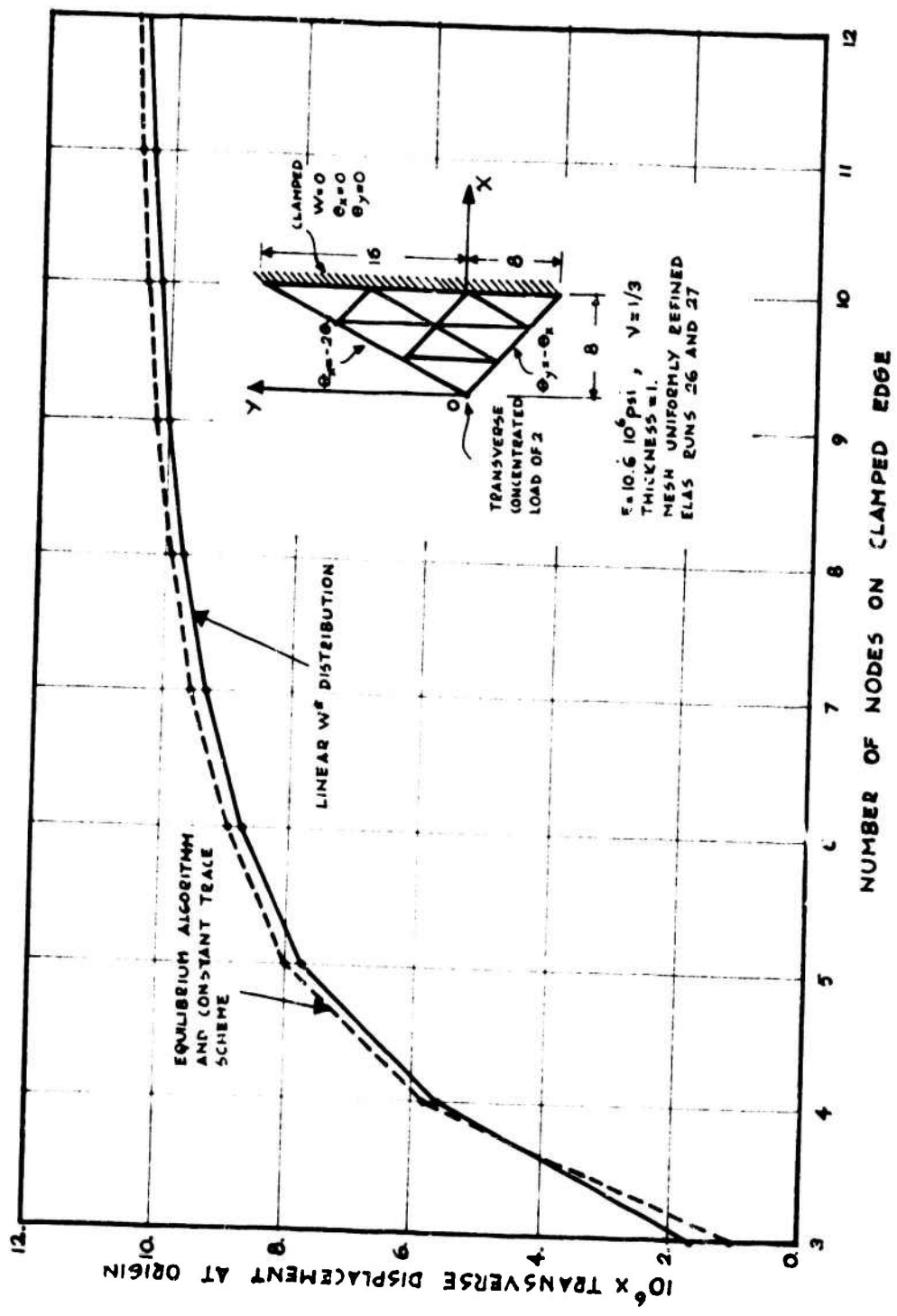


Figure 6. Convergence in Obtuse Triangular Mesh

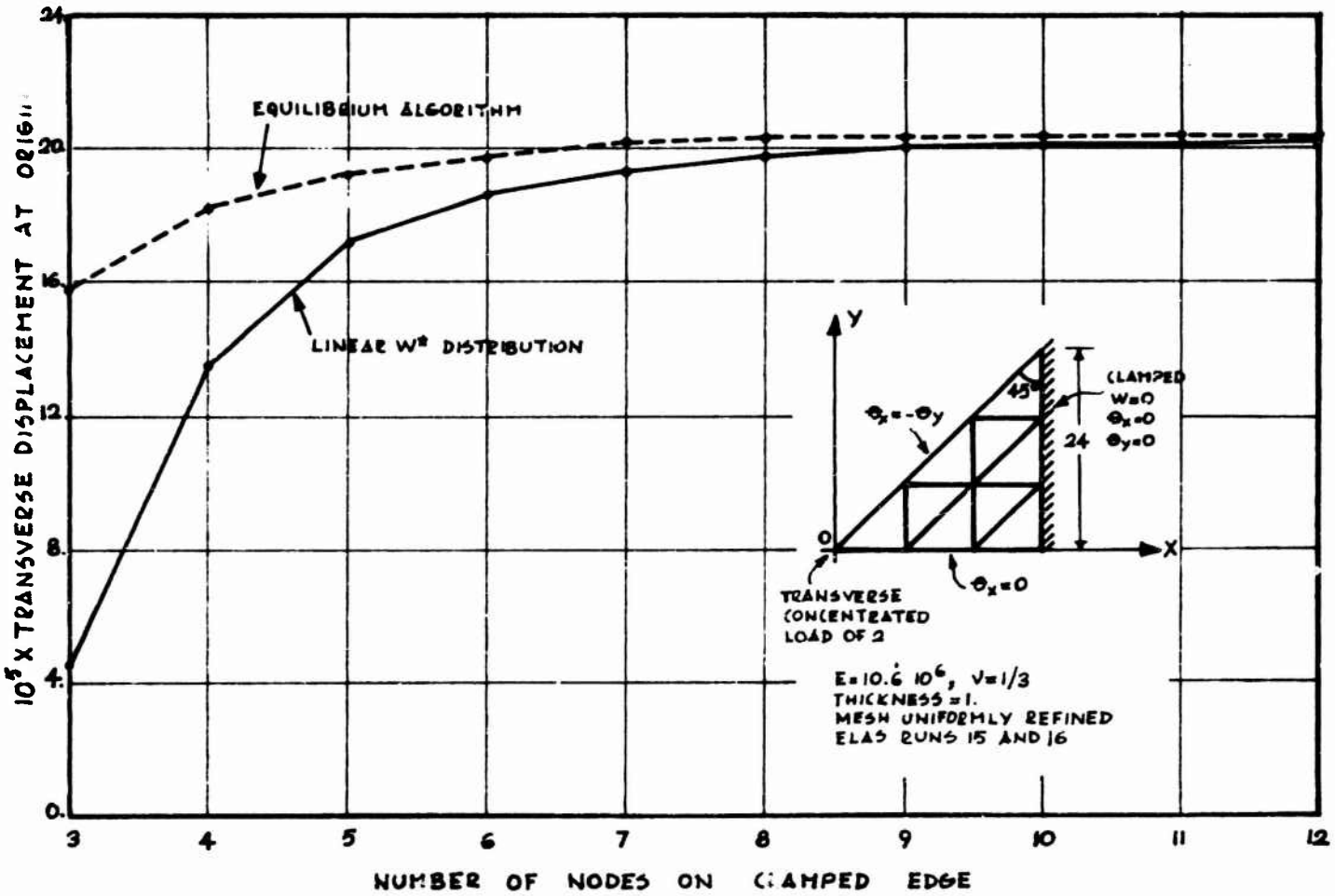


Figure 7. Convergence in Right Triangular Mesh with Proper Orientation

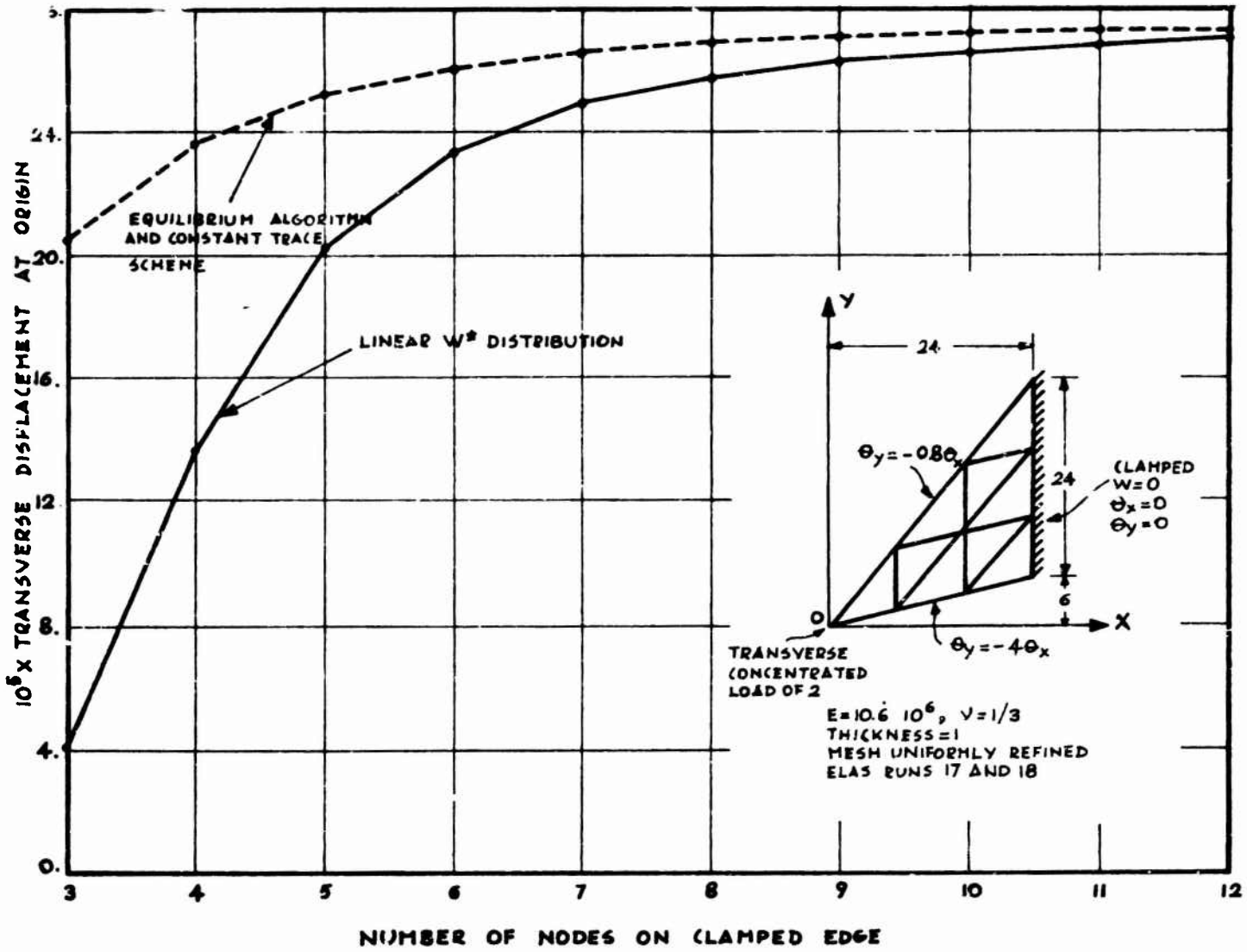


Figure 8. Convergence in Obtuse Triangular Mesh with Proper Orientation

Acknowledgement

The author wishes to acknowledge the computer time support from the Duke University for the convergence studies presented herein.

References

1. Turner, M.J., R.W. Clough, H.C. Martin, and L.J. Topp, "Stiffness and Deflection Analysis of Complex Structures", Journal of the Aeronautical Sciences, 23, No. 9, pp. 805-823, September 1956.

2. Melosh, R.J., "A Flat Triangular Shell Element Stiffness Matrix", AFFDL-TR-66, 1966, pp. 503-509, Air Force Flight Dynamics Lab., Wright-Patterson Air Force Base, Ohio.

3. Utku, Senol, "Stiffness Matricess for Thin Triangular Elements of Nonzero Gaussian Curvature", paper 66-530, AIAA 4th Aerospace Sciences Meeting, Los Angeles, Ca., June 27-29, 1966. See also AIAA Journal, Vol. 5, No. 9, pp. 1659-1667, September 1967; and also JPL TR 32-1153.

4. Utku, Senol and Robert J. Melosh, "Behavior of Triangular Shell Element Stiffness Matrices Associated with Polyhedral Deflection Distributions", paper 67-114, AIAA 5th Aerospace Sciences Meeting, New York, N.Y., January 23-26, 1967. See also AIAA Journal, Vol. 6, No. 2, pp. 374-376, February 1968; and also JPL TR 32-1217, January, 1968.

5. Martin, H.C., "Stiffness Matrix for a Triangular Sandwich Element in Bending", Technical Report 32-1158, Jet Propulsion Laboratory, Pasadena, California, October 1967.

6. Utku, Senol and F.A. Akyuz, "ELAS - A General Purpose Computer Program for the Equilibrium Problems of Linear Structures, User's Manual", TR 32-1240, Vol. I, Jet Propulsion Laboratory, Pasadena, California, February 1968.

7. Melosh, R.J. and H.N Christiansen, "Structural Analysis and Matrix Interpretive System (SAMIS) Program: Technical Report", TM 33-311, Jet Propulsion Laboratory, Pasadena, California, November 1966.

8. DYNAL Program of McDonnell-Douglas Automation Company, St. Louis, Missouri, 1969.

SESSION 3. FINITE ELEMENTS

Session Chairman

R. H. Mallett*

Bell Aerospace Corp.
Buffalo, New York

*Now with Westinghouse Electric Corp., Madison, Pa

COMPARISON OF TWO HIGH-PRECISION TRIANGULAR FINITE ELEMENTS FOR ARBITRARY DEEP SHELLS

by

G. Richard Cowper* and Garry M. Lindberg**
National Aeronautical Establishment
National Research Council of Canada

and

Mervyn D. Olson***
Department of Civil Engineering,
University of British Columbia.

Two approaches to the finite element analysis of arbitrary deep shells are presented and compared. One approach involves the derivation of a suitable transformation to link shallow shell elements together to form deep shells. The second approach involves the derivation of a curvilinear element based on the general tensorial formula for the strain energy of a thin shell. Both approaches use higher order interpolation functions and both result in triangular finite elements with 36 degrees of freedom, 12 at each node. Results confirm that both approaches give accurate predictions of stresses as well as displacements. Solutions rapidly converge as the number of elements is increased and engineering accuracy is always attained with just a few elements.

1.0 INTRODUCTION

The finite element analysis of shells has been studied extensively during the past decade. An extensive critical survey of developments up to 1969 has been given by Gallagher [1], while more limited surveys have been given by Zienkiewicz [2], Key [3], Hartung [4], and by the authors [5]. It is fair to say that no one element has won general acceptance, and also that many of the available elements impose limitations on the shape of the shell, for example, the limitation to shells of revolution.

During the past few years the authors have developed a number of triangular high-precision finite elements for plate bending [6,7], plane stress [8], shallow shells [8] and cylindrical shells [9]. These elements use higher order interpolation polynomials and result in smaller overall problem size, give accurate predictions of stresses, and satisfy sufficient conditions to guarantee rapid convergence to the exact solutions. This family is now enlarged to permit the analysis of deep shells of general shape.

---

*Senior Research Officer; **Associate Research Officer; ***Assistant Professor.

Two approaches are presented. The first is derivation of a curvilinear triangular element based on the general tensorial formula for the strain energy of a thin shell. The second is the derivation of a suitable transformation to link shallow shell elements together to form deep shells. Although the first approach was expected to yield better results, the second approach promised some advantages such as the avoidance of numerical integration and of the complexities of tensor notation. Moreover, available computer routines for shallow shells could be exploited in the second approach. Following a description of the two approaches, the derived elements are tested on certain problems and their performance is compared.

2.0 THE GENERAL SHELL ELEMENT (CURSHL)

The following derivation is very brief but complete details are available [10]. Let α,β be curvilinear coordinates on the shell surface, not necessarily orthogonal nor principal. The same displacement functions which have been used successfully in earlier high-precision elements are also adopted for the general shell element. The normal displacement w is taken as a restricted quintic polynomial in the coordinates α,β, while the tangential displacements u, v are taken as complete cubic polynomials in α,β. In keeping with the assumed displacement functions, the generalized displacements are the values of u, $\partial u/\partial\alpha$, $\partial u/\partial\beta$, v, $\partial v/\partial\alpha$, $\partial v/\partial\beta$, w, $\partial w/\partial\alpha$, $\partial w/\partial\beta$, $\partial^2 w/\partial\alpha^2$, $\partial^2 w/\partial\alpha\partial\beta$, $\partial^2 w/\partial\beta^2$ at the three vertices of the element, a total of 36 degrees of freedom per element. Centroidal displacements u_c, v_c, are used during the development of the stiffness matrix but are later eliminated by static condensation.

The choice of displacement functions and generalized displacements assures that u, v, w, and first derivatives of w are continuous between elements. Conformity, however, requires that the displacement vector and the rotations of the normal be continuous. If the shell is smooth, continuity of the displacement vector is equivalent to continuity of the components u, v, w. In general shell theories, the formulas for rotations involve the first derivatives of w plus products of u, v with the shell curvature. Therefore continuity of u, v, w, and first derivatives of w implies continuity of rotations, provided that the shell curvature is continuous. In Donnell-Vlasov theory and in shallow shell theory the rotations are approximated by the first derivatives of w. In these cases, continuity of rotations is equivalent to continuity of the first derivatives of w regardless of the continuity of shell curvatures.

The highest derivatives of displacement which appear in the expression for strain energy are the second derivatives of w and first derivatives of u, v. Since the interpolation function for w contains a complete quartic polynomial, it follows from Taylor's theorem that the finite element can represent any distribution of second derivatives of w with an error of order h^3, where h is a typical linear dimension of the element. Likewise, the first derivatives of u, v, are represented with an error of order h^3,

since the interpolation functions for u, v are complete cubic polynomials. Then, according to the theorem of minimum potential energy [11,6] the element should have a discretization error in the strain energy of order h^6.

Some authorities [1,12,13] advocate the use of equal order polynomials for the displacements u, v, w. This is not necessary for conformity nor is it efficient. Using quintic polynomials for all of u, v, w, would result in only marginal improvements in accuracy because the order of the error in the strain energy would still be limited by the accuracy of w. Moreover, this marginal improvement would be bought at the price of a 50% increase in the degrees of freedom, from 12 per node to 18 per node. On the other hand, the use of equal order polynomials for u, v, w, facilitates the handling of non-smooth junctions of shells.

Little consideration has been given to obtaining exact rigid-body modes, a point which has often been over-emphasized. Rather, the focus of attention has been the error in the strain energy. Making the error in strain energy acceptably small automatically ensures that rigid-body motions are adequately represented.

Lack of space precludes more than a sketch of the computational procedure. Let {W} be the vector of nodal displacements of an element and let {F(α,β)} be the vector of displacements and their derivatives at a general point α,β, within the element, thus

$$\{F\}^T = \{u,\ u_\alpha,\ u_\beta,\ v,\ v_\alpha,\ v_\beta,\ w,\ w_\alpha,\ w_\beta,\ w_{\alpha\alpha},\ w_{\alpha\beta},\ w_{\beta\beta}\} \quad (1)$$

where the subscripts on u, v, w, denote derivatives. The vector {F} can be related to {W}, thus

$$\{F\} = [S]\{W\} \quad (2)$$

where [S] is a matrix of interpolation polynomials. Let {e} be the vector of membrane and bending strains,

$$\{e\}^T = \{\varepsilon_{11},\ \varepsilon_{12},\ \varepsilon_{22},\ \kappa_{11},\ \kappa_{12},\ \kappa_{22}\} \quad (3)$$

and let {e} be related to {F} by

$$\{e\} = [B]\{F\} \quad (4)$$

The form of [B] depends on which shell theory is used. The computer program which has been developed allows the options of using either Koiter-Sanders theory [14,15], Donnell-Vlasov theory [16] or shallow shell theory [16,17]. In all three cases continuous Kirchoff constraints are applied.

The strain energy of the element can be written

$$U_e = \tfrac{1}{2}\int\int\{e\}^T[E]\{e\}\sqrt{a}\ d\alpha d\beta \quad (5)$$

where [E] is the matrix of bending and stretching rigidities and a is the determinant of the metric tensor. The matrix [E], which

is the same for each of the above three shell theories, is limited in the program to isotropic materials, but variable thickness can be accommodated. It follows from (2), (4), and (5) that

$$U_e = \tfrac{1}{2}\{W\}^T(\int\int[S]^T[B]^T[E][B][S]\sqrt{a}\,d\alpha d\beta)\{W\} \qquad (6)$$

and hence the stiffness matrix is given by

$$[K] = \int\int[S]^T[B]^T[E][B][S]\sqrt{a}\,d\alpha d\beta \qquad (7)$$

For the sake of generality the tensorial form of shell theory was used in setting up the matrices [B] and [E]. The quantities u, v, w, therefore are tensor components of displacement, in contrast to the physical components which are used in other high-precision elements. Computer output routines, which convert tensor components to physical components, have been developed.

Because of the complexity of formula (7) and because so many of the terms of matrices [B] and [E] may be variables, it was decided to abandon the closed-form integration which has been used in previous high-precision elements and to use numerical integration instead to evaluate (7). A newly-developed 13-point numerical integration formula, which has an error of order h^8, is used [10]. A numerical integration formula with error of order h^8 was selected to retain the energy-bounding property of conforming elements. For conforming finite elements of the displacement type the discretization error in strain energy is known to be negative. Since the discretization error of the element is of order h^6, the error of numerical integration should be small compared with the discretization error and the total error in strain energy should still be negative. Hence the calculated strain energy should still be a lower bound on the true strain energy. As will be seen later the attempt to retain the energy-bounding property has not been entirely successful.

The matrix [B] involves many geometric quantities dependent on the shape of the shell, such as metric and curvature tensors and Christoffel symbols. The values of these quantities at the pivotal points of the integration formula must be known in order to evaluate (7). It is assumed that the shell surface is defined by the equations

$$x = x(\alpha,\beta), \qquad y = y(\alpha,\beta), \qquad z = z(\alpha,\beta) \qquad (8)$$

where x, y, z are Cartesian coordinates of a general point on the shell. The required geometric quantities can be computed from x, y, z and their derivatives using standard formulae [18]. Derivatives up to third order are required in Koiter-Sander's theory while second order derivatives of x, y, z suffice in Donnell-Vlasov and shallow shell theory. This data is fed into the computer program from a user-supplied subroutine which must return the values of x, y, z and their derivatives at any arbitrarily given point α,β. If, as is generally the case, the shell surface is of simple form then the exact equations of the surface can be used in setting up the subroutine. On the other hand, the representation of the shell by a fitted polynomial surface is not precluded. In all applications

to date the exact equations of the surface have been used.

The computation of consistent load vectors is similar to the computation of the stiffness matrix. The computer program can accommodate thermal loads as well as arbitrarily distributed body forces on the shell surface. Data on the loads is fed into the program from a second user-supplied subroutine which must return the values of all components of load at an arbitrarily given point α,β.

3.0 THE TRANSFORMED SHALLOW SHELL ELEMENT (TSS)

The following derivations are necessarily brief, but full details are available in Ref. 19. This work is a natural extension of the simplified transformation proposed in Ref. 5. The main requirement for applying shallow shell elements to an arbitrary deep shell is a suitable transformation between the shallow shell degrees of freedom and those for the deep shell. The shallow shell element of Ref. 8 is used here because it proved to be superior both in ease of formulation and in accuracy to most other shallow shell elements. Further, it leads to the same deep shell degrees of freedom as for the general shell element (Sec. 2) and hence is directly comparable with it.

The relevant geometry is shown in Figure 1, where x, y, z are Cartesian global coordinates, ξ, η, ζ, are Cartesian local coordinates, and α,β are curvilinear shell coordinates lying in the mid-surface of the shell. At this point, the latter need not be orthogonal. The ξ, η coordinates are to be used as the base plane coordinates in the shallow shell formulation and the ξ, η axes are defined to go through the element corner nodes 1, 2, 3 as shown.

Again it is assumed that the shell surface is defined by equations (8) and that derivatives of x, y, z up to second order are available. This leads to a natural computation process in which the element corner nodes are located first by specifying their shell coordinates (α,β), and then their global coordinates (x, y, z) are calculated from equations (8). The element geometry may then be determined from these latter coordinates.

Some results from analytical geometry follow. The global coordinates of the ξ, η, ζ, system origin are

$$\begin{aligned} x_o &= (1-\rho)x_1 + \rho x_2 \\ y_o &= (1-\rho)y_1 + \rho y_2 \\ z_o &= (1-\rho)z_1 + \rho z_2 \end{aligned} \tag{9}$$

where the parameter ρ is

$$\rho = \frac{(x_3-x_1)(x_2-x_1) + (y_3-y_1)(y_2-y_1) + (z_3-z_1)(z_2-z_1)}{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2} \tag{10}$$

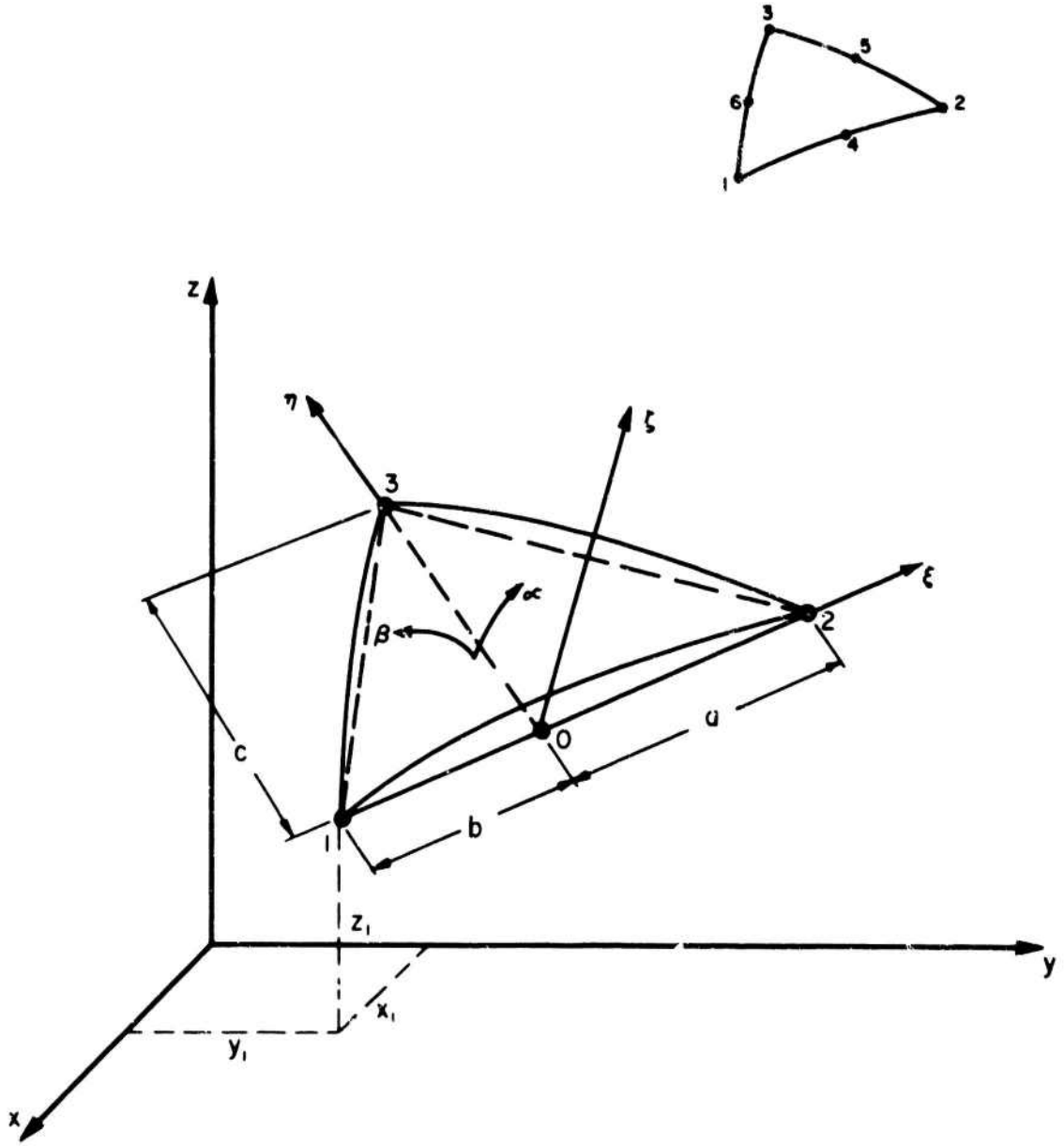


Figure 1. Geometry and Co-Ordinate Systems For Transformed Shallow Shell Formulation

The subscripts denote the element corner nodes 1, 2, 3 of Figure 1. The element dimensions a, b, c are then given by

$$a = (1-\rho)\sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}$$

$$b = \rho\sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2} \qquad (11)$$

$$c = \sqrt{(x_3-x_0)^2 + (y_3-y_0)^2 + (z_3-z_0)^2}$$

The following relations between the global and local Cartesian coordinates are obtained

$$\xi = [(x_2-x_1)(x-x_0) + (y_2-y_1)(y-y_0) + (z_2-z_1)(z-z_0)]/(a+b)$$

$$\eta = [(x_3-x_0)(x-x_0) + (y_3-y_0)(y-y_0) + (z_3-z_0)(z-z_0)]/c \qquad (12)$$

$$\zeta = b_1 (x-x_0) + b_2 (y-y_0) + b_3 (z-z_0)$$

where

$$b_1 = [(y_2-y_1)(z_3-z_0) - (z_2-z_1)(y_3-y_0)]/(a+b)c$$

$$b_2 = [(z_2-z_1)(x_3-x_0) - (x_2-x_1)(z_3-z_0)]/(a+b)c \qquad (13)$$

$$b_3 = [(x_2-x_1)(y_3-y_0) - (y_2-y_1)(x_3-x_0)]/(a+b)c$$

Combining equations (8) and (12) then yields the base plane coordinates as explicit functions of the shell coordinates, written symbolically as

$$\xi = \xi(\alpha,\beta)$$
$$\eta = \eta(\alpha,\beta) \qquad (14)$$

Now all the derivatives required in transforming the generalized displacements from local coordinates to shell coordinates may be derived from the above equations.

3.1 Transformation Matrices

The shallow shell formulation of Ref. 8 begins with generalized displacements written relative to the base plane coordinates, w, w_ξ, w_η, etc., and these are now transformed to deep shell ones. The displacement normal to the shell w is a scalar and hence, using equations (14), its derivatives transform simply as

$$(w, w_\alpha, w_\beta, w_{\alpha\alpha}, w_{\alpha\beta}, w_{\beta\beta})^T = [R_2](w, w_\xi, w_\eta, w_{\xi\xi}, w_{\xi\eta}, w_{\eta\eta})^T \qquad (15)$$

where $[R_2]$ is given in Table 1. The tangential displacements u, v are defined to be parallel to the coordinate axes and hence must be transformed as vectors. Again using results from analytical geometry and the element's shallowness leads to the approximation

$$(u, u_\xi, u_\eta, v, v_\xi, v_\eta)^T = [R_1](\tilde{u}, \tilde{u}_\alpha, \tilde{u}_\beta, \tilde{v}, \tilde{v}_\alpha, \tilde{v}_\beta)^T \quad (16)$$

where $\tilde{u}$ and $\tilde{v}$ are in the directions of α and β, and $[R_1]$ is given in Table 1. Combining all these results and dropping the "tilda" notation for simplicity, the complete transformation for one element is

$$\{W_1\} = [R]\{W_2\} \quad (17)$$

where

$$\{W_2\}^T = (u_1, u_{\alpha 1}, u_{\beta 1}, v_1, v_{\alpha 1}, v_{\beta 1}, w_1, w_{\alpha 1}, w_{\beta 1}, w_{\alpha\alpha 1}, w_{\alpha\beta 1}, w_{\beta\beta 1}, u_2, \ldots\ldots, u_3, \ldots\ldots, u_c, v_c) \quad (18)$$

is the generalized displacement vector in shell coordinates and [R] is the new "rotation" matrix given in Table 1. The 38 × 38 stiffness matrix relative to shell coordinates is then

$$[K] = [R]^T [T_1]^T [k] [T_1] [R] \quad (19)$$

where $[T_1]$, $[k]$, $\{W_1\}$, the consistent load vector and the stiffness matrix condensation procedure are the same as in Ref. 8.

It may be seen that the foregoing transformation breaks down if the Jacobian of equations (14) vanishes. This occurs for example at the pole of a shell of revolution with most coordinate systems. This difficulty is easily overcome by reverting to the global Cartesian coordinates of Ref. 8 at such a point, since any element in that region will also be shallow with respect to the latter coordinates. That is, the present rotation matrices $[R_1]$ and $[R_2^{-1}]$ are merely replaced by $[R_1]$ and $[R_2]$ given in Table 1 of Ref. 8 for a pole node.

3.2 Shell Curvatures

The shell element curvatures may be determined locally by specifying the shell elevations above the ξ,η base plane at the mid-side nodes 4, 5, 6 (Fig. 1) and using them in the quadratic function, equation (5) of Ref. 8. Note that the elevations are zero at the corner nodes 1, 2, 3. The locations of the mid-side nodes are determined first in the shell coordinates simply as averages of the corner node coordinates. The global coordinates of these nodes are then easily obtained from equations (8) and are substituted into equations (12) to yield the three elevations ζ_i and the corresponding base plane coordinates ξ_i, η_i for i = 4,5,6.

TABLE 1: TRANSFORMATION MATRIX [R]

$$[R] = \begin{bmatrix} [R_1]_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & [R_2^{-1}]_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & [R_1]_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & [R_2^{-1}]_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & [R_1]_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & [R_2^{-1}]_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & [R_3]_c \end{bmatrix}$$

where

$$[R_2] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \xi_\alpha & \eta_\alpha & 0 & 0 & 0 \\ \theta & \xi_\beta & \eta_\beta & 0 & 0 & 0 \\ 0 & \xi_{\alpha\alpha} & \eta_{\alpha\alpha} & \xi_\alpha^2 & 2\xi_\alpha\eta_\alpha & \eta_\alpha^2 \\ 0 & \xi_{\alpha\beta} & \eta_{\alpha\beta} & \xi_\alpha\xi_\beta & \xi_\alpha\eta_\beta+\xi_\beta\eta_\alpha & \eta_\alpha\eta_\beta \\ 0 & \xi_{\beta\beta} & \eta_{\beta\beta} & \xi_\beta^2 & 2\xi_\beta\eta_\beta & \eta_\beta^2 \end{bmatrix}$$

$$[R_3] = \begin{bmatrix} \xi_\alpha/a & \xi_\beta/b \\ \eta_\alpha/a & \eta_\beta/b \end{bmatrix} \qquad a = \sqrt{\xi_\alpha^2 + \eta_\alpha^2} \qquad b = \sqrt{\xi_\beta^2 + \eta_\beta^2}$$

$J = \xi_\alpha\eta_\beta - \xi_\beta\eta_\alpha$, $\xi_\alpha = \partial\xi/\partial\alpha$, etc.

Subscripts $[\ ]_i$ mean evaluated at node i or centroid c.

Table 1 (cont'd)

$$[R_1] = \begin{bmatrix}
\xi_\alpha/a & 0 & 0 \\
\eta_\alpha^2(\xi_{\alpha\alpha}\eta_\beta - \xi_{\alpha\beta}\eta_\alpha)/Ja^3 & & \\
+\xi_\alpha\eta_\alpha(\eta_\alpha\eta_{\alpha\beta} - \eta_\beta\eta_{\alpha\alpha})/Ja^3 & \xi_\alpha\eta_\beta/Ja & -\xi_\alpha\eta_\alpha/Ja \\
\eta_\alpha^2(\xi_\alpha\xi_{\alpha\beta} - \xi_\beta\xi_{\alpha\alpha})/Ja^3 & & \\
+\xi_\alpha\eta_\alpha(\xi_\beta\eta_{\alpha\alpha} - \xi_\alpha\eta_{\alpha\beta})/Ja^3 & -\xi_\alpha\xi_\beta/Ja & \xi_\alpha^2/Ja \\
\eta_\alpha/a & 0 & 0 \\
\xi_\alpha^2(\eta_\beta\eta_{\alpha\alpha} - \eta_\alpha\eta_{\alpha\beta})/Ja^3 & & \\
+\xi_\alpha\eta_\alpha(\xi_{\alpha\beta}\eta_\alpha - \xi_{\alpha\alpha}\eta_\beta)/Ja^3 & \eta_\alpha\eta_\beta/Ja & -\eta_\alpha^2/Ja \\
\xi_\alpha^2(\xi_\alpha\eta_{\alpha\beta} - \xi_\beta\eta_{\alpha\alpha})/Ja^3 & & \\
+\xi_\alpha\eta_\alpha(\xi_{\alpha\alpha}\xi_\beta - \xi_\alpha\xi_{\alpha\beta})/Ja^3 & -\xi_\beta\eta_\alpha/Ja & \xi_\alpha\eta_\alpha/Ja \\
\xi_\beta/b & 0 & 0 \\
\eta_\beta^2(\xi_{\alpha\beta}\eta_\beta - \xi_{\beta\beta}\eta_\alpha)/Jb^3 & & \\
+\xi_\beta\eta_\beta(\eta_\alpha\eta_{\beta\beta} - \eta_\beta\eta_{\alpha\beta})/Jb^3 & \xi_\beta\eta_\beta/Jb & -\xi_\beta\eta_\alpha/Jb \\
\eta_\beta^2(\xi_\alpha\xi_{\beta\beta} - \xi_\beta\xi_{\alpha\beta})/Jb^3 & & \\
+\xi_\beta\eta_\beta(\xi_\beta\eta_{\alpha\beta} - \xi_\alpha\eta_{\beta\beta})/Jb^3 & -\xi_\beta^2/Jb & \xi_\alpha\xi_\beta/Jb \\
\eta_\beta/b & 0 & 0 \\
\xi_\beta^2(\eta_\beta\eta_{\alpha\beta} - \eta_\alpha\eta_{\beta\beta})/Jb^3 & & \\
+\xi_\beta\eta_\beta(\xi_{\beta\beta}\eta_\alpha - \xi_{\alpha\beta}\eta_\beta)/Jb^3 & \eta_\beta^2/Jb & -\eta_\alpha\eta_\beta/Jb \\
\xi_\beta^2(\xi_\alpha\eta_{\beta\beta} - \xi_\beta\eta_{\alpha\beta})/Jb^3 & & \\
+\xi_\beta\eta_\beta(\xi_\beta\xi_{\alpha\beta} - \xi_\alpha\xi_{\beta\beta})/Jb^3 & -\xi_\beta\eta_\beta/Jb & \xi_\alpha\eta_\beta/Jb
\end{bmatrix}$$

Note that these points do not necessarily coincide with the mid-side points of the base plane triangle. These results, together with the zero elevations at the three corner nodes, are then sufficient to complete the quadratic fit to the shell shape and determine the 6 coefficients. Finally, the three shell curvatures required in the strain energy calculation are obtained simply by differentiation.

Alternately, the shell curvatures for each element may be specified externally. This is the approach used for the results presented herein, since all example applications have constant curvatures anyway.

4.0 RESULTS

Both elements have been extensively tested [10,19] and only selected results are given below to emphasize the major characteristics of the two approaches and illustrate their effectiveness.

4.1 Rigid Body Modes

Rigid body modes have been calculated for a typical spherical shell element. The angular coordinates of the element analyzed are given at the top of Table 2, Poisson's ratio was taken as $\nu = 0.3$, $Et/(1-\nu^2)$ was set equal to 1 and R/t was fixed at 50. Three different sizes of element were analyzed, with R = 1,10 and 100. The first eight eigenvalues obtained using TSS and CURSHL (K-S) (Koiter-Sanders shell theory) are given in Table 2. Note that the first 'non-rigid' body mode, the seventh eigenvalue, has the same predicted value for all sizes of element for the two approaches. It is always substantially larger than the last 'rigid' body mode.

For TSS, λ_7 is always some seven orders of magnitude larger than λ_6, indicating that the first six eigenvalues are indeed rigid body modes. This is not surprising, since zero strain modes are polynomials in Donnell-Vlasov (D-V) or in shallow shell theory and these are represented exactly in TSS. The negative eigenvalues are extremely small and are probably due to round-off error.

For CURSHL (K-S), a smaller separation between λ_7 and λ_6 occurs, and this separation varies with the element size. This is expected, since zero strain modes are not exactly represented by polynomials in Koiter-Sanders shell theory and since an approximate numerical integration scheme is used. However, CURSHL (K-S) still gives excellent numerical results.

4.2 Uniformly Pressurized Sphere

This is a membrane type problem, with R/t = 50 and $\nu = 0.3$ (Fig. 2). Because of symmetry, only a wedge of the shell extending between the equator and the pole had to be analyzed. A uniformly spaced gridwork was used and results were obtained for N = 2,4,8 and 16. Typical non-dimensionalized strain energy, displacements

Finite Element Co-ordinates, Radians

i	1	2	3
α_i	0	0.174533	0
β_i	1.57080	1.74533	1.74533

Eigenvalues In Ascending Magnitude CURSHL (K-S)

No.	R = 1 t = 0.02	R = 10 t = 0.2	R = 100 t = 2
1	$.861 \times 10^{-14}$	$.206 \times 10^{-15}$	$.208 \times 10^{-17}$
2	$.765 \times 10^{-13}$	$.839 \times 10^{-15}$	$.838 \times 10^{-17}$
3	$.348 \times 10^{-12}$	$.655 \times 10^{-14}$	$.659 \times 10^{-16}$
4	$.462 \times 10^{-11}$	$.693 \times 10^{-13}$	$.697 \times 10^{-15}$
5	$.282 \times 10^{-10}$	$.504 \times 10^{-11}$	$.690 \times 10^{-13}$
6	$.969 \times 10^{-10}$	$.148 \times 10^{-10}$	$.203 \times 10^{-12}$
7	$.488 \times 10^{-8}$	$.487 \times 10^{-8}$	$.487 \times 10^{-8}$
8	$.731 \times 10^{-8}$	$.724 \times 10^{-8}$	$.720 \times 10^{-8}$
$\frac{\lambda_6}{\lambda_7}$	1.98×10^{-2}	3.02×10^{-3}	4.15×10^{-5}

Eigenvalues In Ascending Magnitude - TSS

No.	R = 1 t = 0.02	R = 10 t = 0.2	R = 100 t = 2
1	$-.736 \times 10^{-18}$	$.580 \times 10^{-18}$	$-.386 \times 10^{-18}$
2	$-.768 \times 10^{-18}$	$.676 \times 10^{-18}$	$.409 \times 10^{-18}$
3	$-.324 \times 10^{-17}$	$.167 \times 10^{-17}$	$-.139 \times 10^{-17}$
4	$.226 \times 10^{-16}$	$.209 \times 10^{-16}$	$.300 \times 10^{-16}$
5	$-.735 \times 10^{-15}$	$-.166 \times 10^{-15}$	$-.165 \times 10^{-14}$
6	$.121 \times 10^{-14}$	$.267 \times 10^{-14}$	$.245 \times 10^{-14}$
7	$.493 \times 10^{-8}$	$.493 \times 10^{-8}$	$.493 \times 10^{-8}$
8	$.747 \times 10^{-8}$	$.747 \times 10^{-8}$	$.747 \times 10^{-8}$
$\frac{\lambda_6}{\lambda_7}$	2.45×10^{-7}	5.41×10^{-7}	4.98×10^{-7}

Table 2. Rigid Body Modes for a Spherical Shell Segment Using Exact Curvatures: $\nu = 0.3$, $Et/(1-\nu^2)=1$, $R/t=50$

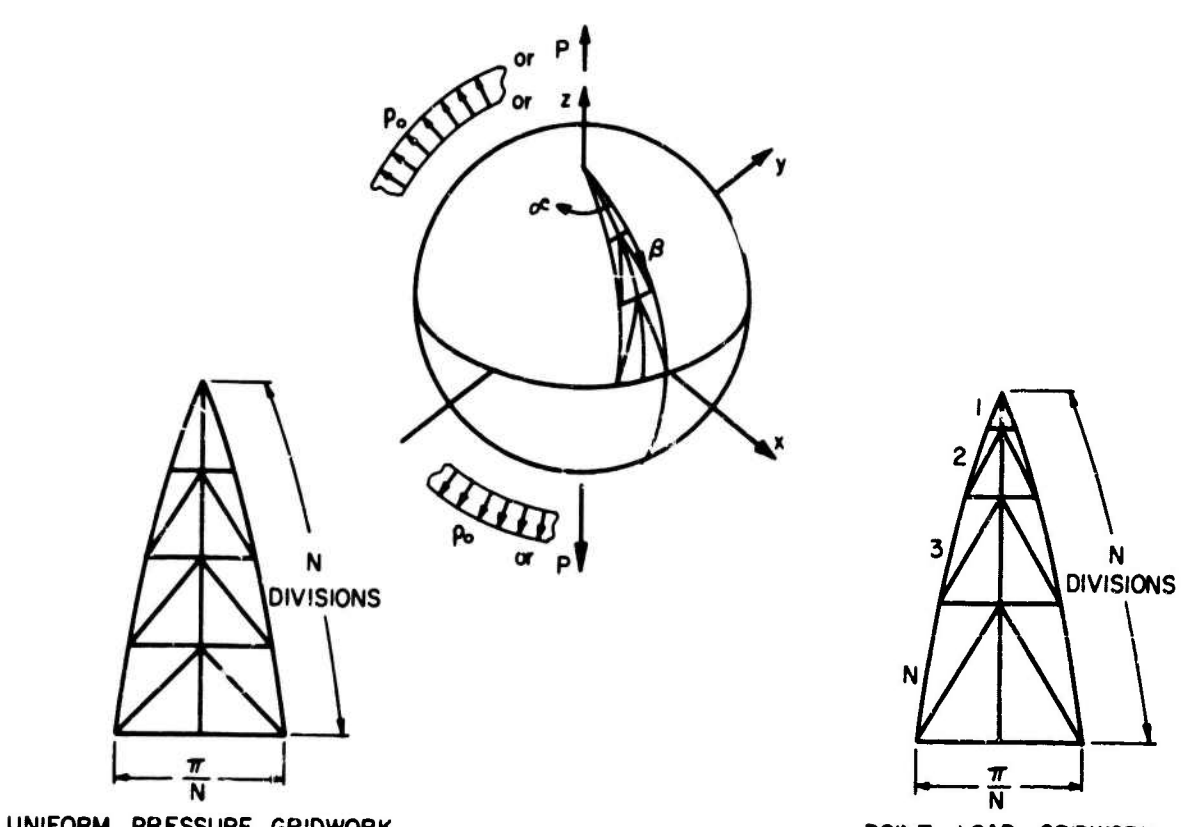


Figure 2. Spherical Shell Program
R/t = 50, ν = 0.3

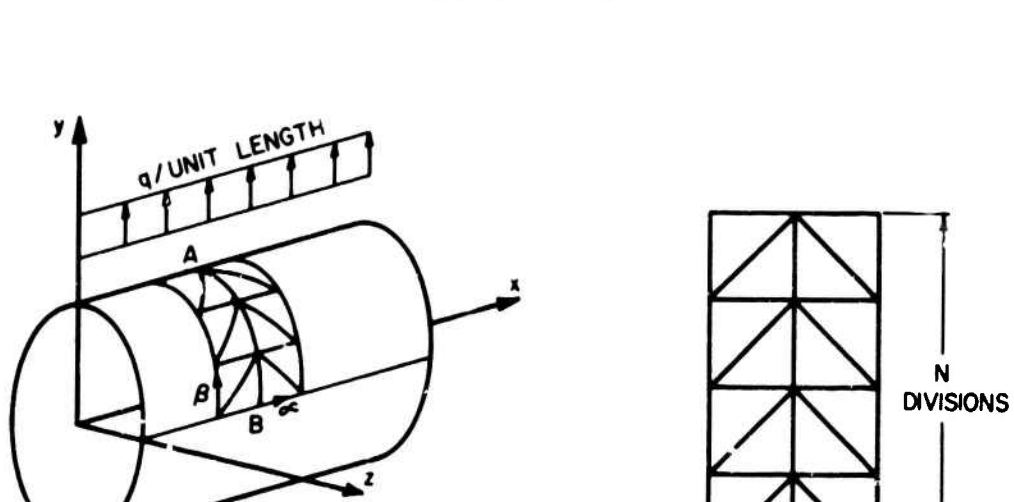


Figure 3. Infinitely Long Pinched Cylindrical Shell
R/t = 10, ν = 0.3

and stress resultants are given in Table 3.

The exact solution is a uniform radial outward displacement. Although the shape functions of both CURSHL and TSS can exactly represent this displacement, neither element quite achieves an exact solution. In the case of CURSHL, this is due to the approximate integration scheme used. The errors are extremely small however. TSS also gives excellent results and the strain energy convergence rate is slightly better than N^{-2}. Exact results are not obtained because TSS produces only an approximately conforming assemblage of elements.

4.3 Pinched Spherical Shell

The pinched sphere is a difficult problem in that there are regions of large bending stresses, regions where membrane action predominates and a region of high stress concentration. Regrettably, the problem is axi-symmetric but the authors have been unable to find a meaningful non-axi-symmetric problem for which a reasonable 'exact' solution exists. Again $R/t = 50$ and $\nu = 0.3$ were used and only a wedge of the shell was analyzed. A non-uniform gridwork spacing was used where the ratios of the sides of successive elements were taken as 1:2:3:4 .....N (Fig. 2). Two solutions were found using CURSHL, one with Koiter-Sanders (K-S) and one with Donnell-Vlasov (D-V) shell theory. These results together with the TSS results are given in non-dimensional form in Table 4 and Figures 4 to 9.

The displacement under the load (proportional to strain energy) is given in Table 4 and is plotted in Figure 4. All three results converge very rapidly, with CURSHL (K-S) converging to the 'asymptotic' solution given by Koiter [20]. Both TSS and CURSHL (D-V) are converging to a common result, some 0.5 percent below the Koiter solution, as a result of differences in shell theories. TSS slightly overshoots the solution for small values of N, not surprising since it is not fully conforming. The convergence rate of CURSHL (K-S) is close to N^{-4}. This is more rapid than that found for most point load problems [6,7,8] and probably occurs because of the non-uniform spacing of the gridworks. Table 4 also gives values for displacements and stresses at various points in the shell, and convergence of all these results is very rapid. Again, CURSHL (K-S) converges to different values than the other two results.

Displacement distributions near the load are given in Figure 5 for N = 5 and 10. These are compared to a shallow shell theory solution from Timoshenko [21], since it is extremely difficult to calculate distributions from the Koiter 'asymptotic' solution. Clearly, the Timoshenko solution is not precise enough, since it is approximately three percent lower than the Koiter solution under the load. The finite element solutions follow the shape of the Timoshenko solution closely but lie slightly above as expected. For N = 10, TSS and CURSHL (K-S) are indistinguishable, while for N = 5, minor differences are observed. CURSHL (D-V) has not been

	Deflection At Pole $\frac{Etw_p}{p_o R^2}$		Deflection at Equator $\frac{Etw_e}{p_o R^2}$		Stress Resultant At Pole $\frac{N_{\alpha p}}{p_o R} = \frac{N_{\beta p}}{p_o R}$	
N	CURSHL(K-S)	TSS	CURSHL(K-S)	TSS	CURSHL(K-S)	TSS
2	.350003	.3505	.349972	.3205	.50006	.4773
4	.350001	.3522	.350000	.3435	.50001	.4862
8	.350000	.3520	.350000	.3484	.50000	.4955
16		.3507		.3496		.4988
Exact	0.350000		0.350000		0.50000	

	Stress Resultants At Equator $N_{\alpha e}/p_o R$		$N_{\beta e}/p_o R$		Strain Energy $EtU/p_o^2 R^4$	
N	CURSHL(K-S)	TSS	CURSHL(K-S)	TSS	CURSHL(K-S)	TSS
2	.50006	.5291	.50001	.4884	2.65561544	2.3458
4	.50000	.4975	.50000	.5026	2.65561517	2.5724
8	.50000	.4989	.50000	.5011	2.65561511	2.6344
16		.4997		.5003		2.6503
Exact	0.50000		0.50000		2.65561509	

Table 3. Convergence of Displacements and Stresses of Uniform Pressure Loaded Spherical Shell: $R/t = 50$, $\nu = 0.3$

	Deflection Under Load Etw_p/P			Stress Resultant Under Load $RN_{\alpha p}/P = RN_{\beta p}/P$		
N	CURSHL(K-S)	CURSHL(D-V)	TSS	CURSHL(K-S)	CURSHL(D-V)	TSS
2	12.790	12.727	18.100	1.649	1.681	7.047
4	20.673	20.567	20.946	7.853	8.002	8.645
6	21.085	20.979	21.109	9.341	9.565	9.762
8	21.162	21.056	21.113	9.716	9.994	10.069
10	21.184	21.078	21.109	9.839	10.161	10.195
15	21.197	21.091	21.102	9.883	10.286	10.294
20	21.199	21.093	21.100	9.853	10.313	10.32
Koiter	21.2004					
Timoshenko	20.654					

	Deflection At Equator Etw_e/P			Stress Resultant At Equator $RN_{\beta p}/P$		
N	CURSHL(K-S)	CURSHL(D-V)	TSS	CURSHL(K-S)	CURSHL(D-V)	TSS
2	-.2994	-.2993	-.0867	.0611	.0615	.3071
4	-.2084	-.2083	-.2256	.1628	.1627	.1873
6	-.2069	-.2072	-.2161	.1602	.1602	.1686
8	-.2069	-.2068	-.2114	.1596	.1596	.1636
10	-.2069	-.2068	-.2094	.1594	.1593	.1617
15	-.2069	-.2068	-.2079	.1592	.1592	.1602
20	-.2069	-.2068	-.2069	.1591	.1591	.1597
Membrane	-0.2069			0.1592		

Table 4. Convergence of Displacements and Stresses of Point Loaded Spherical Shell; R/t = 50, ν=0.3

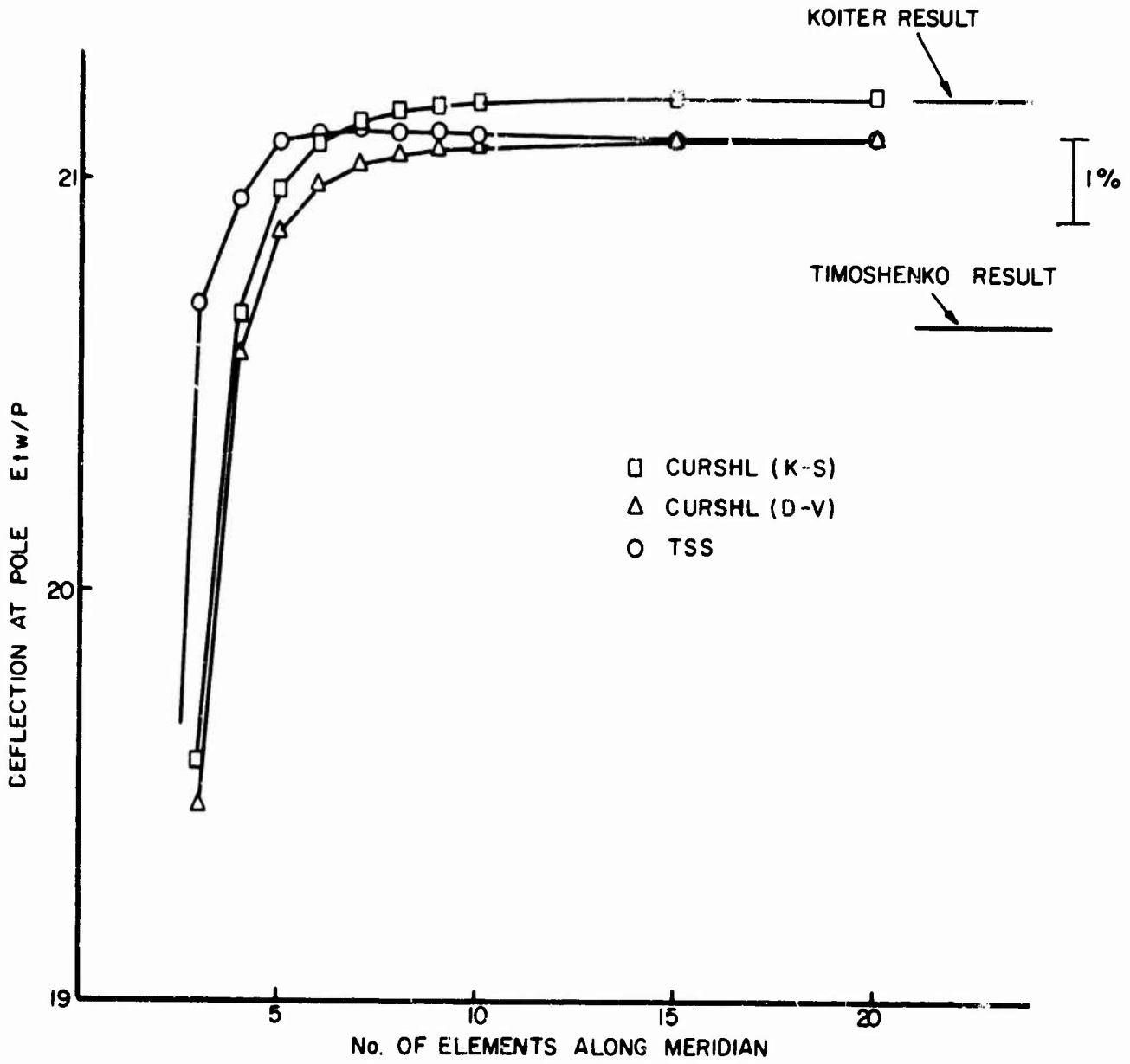


Figure 4. Displacement Under Point Load on Spherical Shell
R/t = 50, ν = 0.3

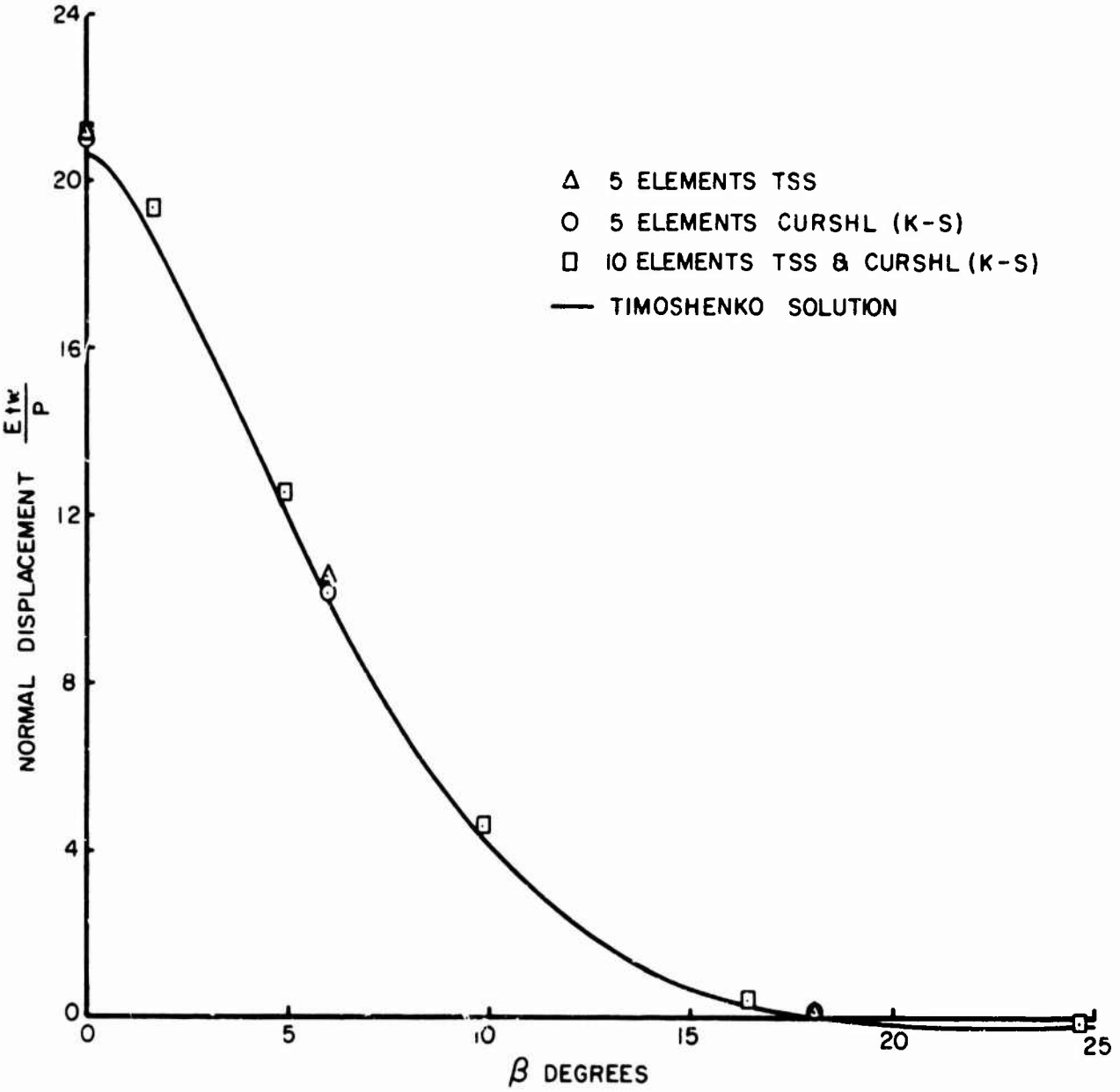


Figure 5. Displacement Near Pole of Point Loaded Sphere
R/t = 50, ν = 0.3

plotted since it is essentially identical with TSS.

The displacements are compared with the membrane solution well away from the load in Figure 6. For N = 10, both TSS and CURSHL (K-S) give 'exact' distributions, while for N = 5, there are slight differences. It appears that the membrane solution is valid for values of β greater than 40°. It is remarkable that the finite element results are so accurate well away from the load, since the displacements are over one order of magnitude smaller than the displacement under the load.

Bending stress and stress resultant distributions near the point load are given in Figures 7 and 8 for N = 5 and 10, and are compared with a Flügge 'asymptotic' solution [22]. These stresses are very accurately predicted by both methods and the only differences between TSS and CURSHL (K-S) occur near the point load.

Away from the point load, bending stresses die out and stress resultants may be compared with a membrane solution. This is done in Figure 9, where it may be seen that CURSHL (K-S) gives excellent results. TSS results lie close to the membrane solution for N_β but are still not quite converged for N_α. It must be remembered, however, that these stresses are much smaller than the critical stresses near the point load.

4.4 Infinitely Long Pinched Cylindrical Shell

This problem is a sensitive indicator of differences in shell theory. A Poisson's ratio of $\nu = 0.0$ was used and all x-direction degrees of freedom were constrained to be zero so that only a strip of the shell had to be analyzed (Fig. 3). An R/t ratio of 10 was chosen so that the strain energy of stretching would not be too small relative to the strain energy of bending. Uniform gridworks were used for a quarter of the cylinder and solutions have been obtained for N = 1,2,4,8 and 16.

The non-dimensional results obtained with CURSHL (K-S), CURSHL (D-V) and TSS are given in Table 5 together with exact results. All three results converge rapidly to the exact solutions, although the performance of CURSHL is significantly better than that of TSS. A plot of strain energy (directly proportional to displacement under the load) reveals that CURSHL is converging at a rate close to its predicted asymptotic value, N^{-6}, while TSS is converging at a rate of only N^{-2}. TSS converges to the Donnell-Vlasov theory results. This is not surprising since the shallow shell elements used in the transformation process are based on Donnell-Vlasov type equations for bending strains.

4.5 Timing Considerations

Computer timings and cost analyses are exceedingly difficult to present in a useful manner. Both elements have been programmed on IBM 360/67 computers, but different operating systems were used. Therefore, times given are not directly comparable. Double

	Deflection Under Line Load Etw_A/qR			Deflection 90° From Line Load Etw_B/qR		
	CURSHL(K-S)	CURSHL(D-V)	T.S.S.	CURSHL(K-S)	CURSHL(D-V)	T.S.S.
1	43.8757	33.1989	26.5908	-35.7181	-26.7111	-21.7275
2	86.9064	50.9793	47.1083	-79.3085	-44.1400	-40.7192
4	89.5872	51.9687	50.9115	-81.6603	-44.8811	-43.9443
8	89.6583	51.9949	51.7278	-81.7207	-44.8989	-44.6623
16	-	-	51.9286	-	-	-44.8400
Exact K-S	89.6597			-81.7219		
Exact D-V		51.9954			-44.8992	

	Stress Resultant Under Line Load N_A/q			Bending Moment Under Line Load M_A/qR		
	CURSHL(K-S)	CURSHL(D-V)	T.S.S.	CURSHL(K-S)	CURSHL(D-V)	T.S.S.
1	-7.675	-5.6685	4.2566	.23044	.21347	.09713
2	-3.338	-1.969	2.2108	.31616	.26077	.21582
4	-0.533	-0.0853	0.9945	.31839	.26181	.24997
8	-0.067	0.2616	0.5244	.31832	.26180	.25875
16	-	-	0.3751	-	-	.26103
Exact K-S	0.0			+0.31831		
Exact D-V		0.3183			+0.26180	

Table 5. Convergence of Displacements and Stresses of Infinitely Long Pinched Cylindrical Shell: $R/t = 10$, $\nu = 0.0$.

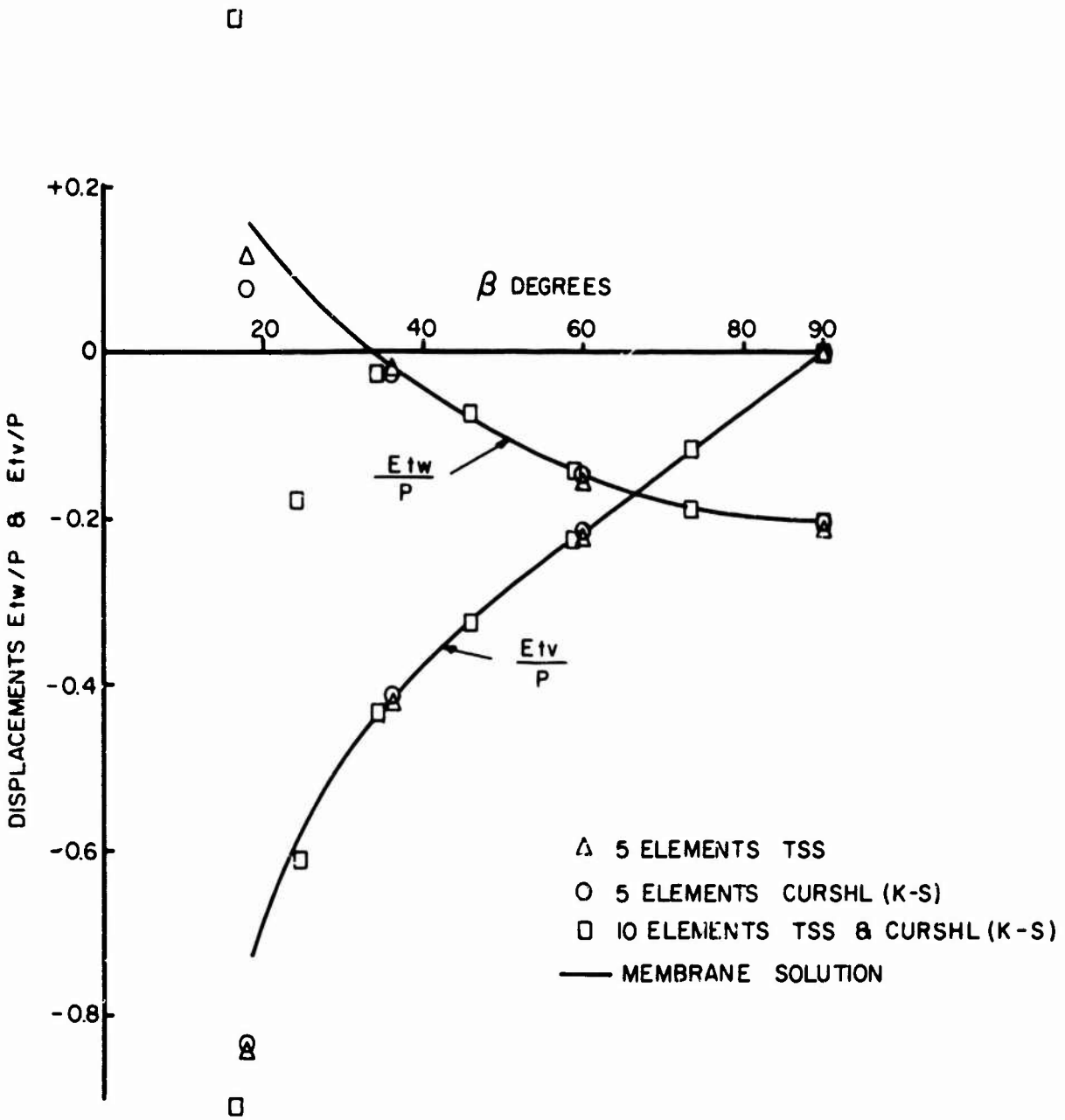


Figure 6. Displacements Remote From Pole of Point Loaded Sphere
R/t = 50, ν = 0.3

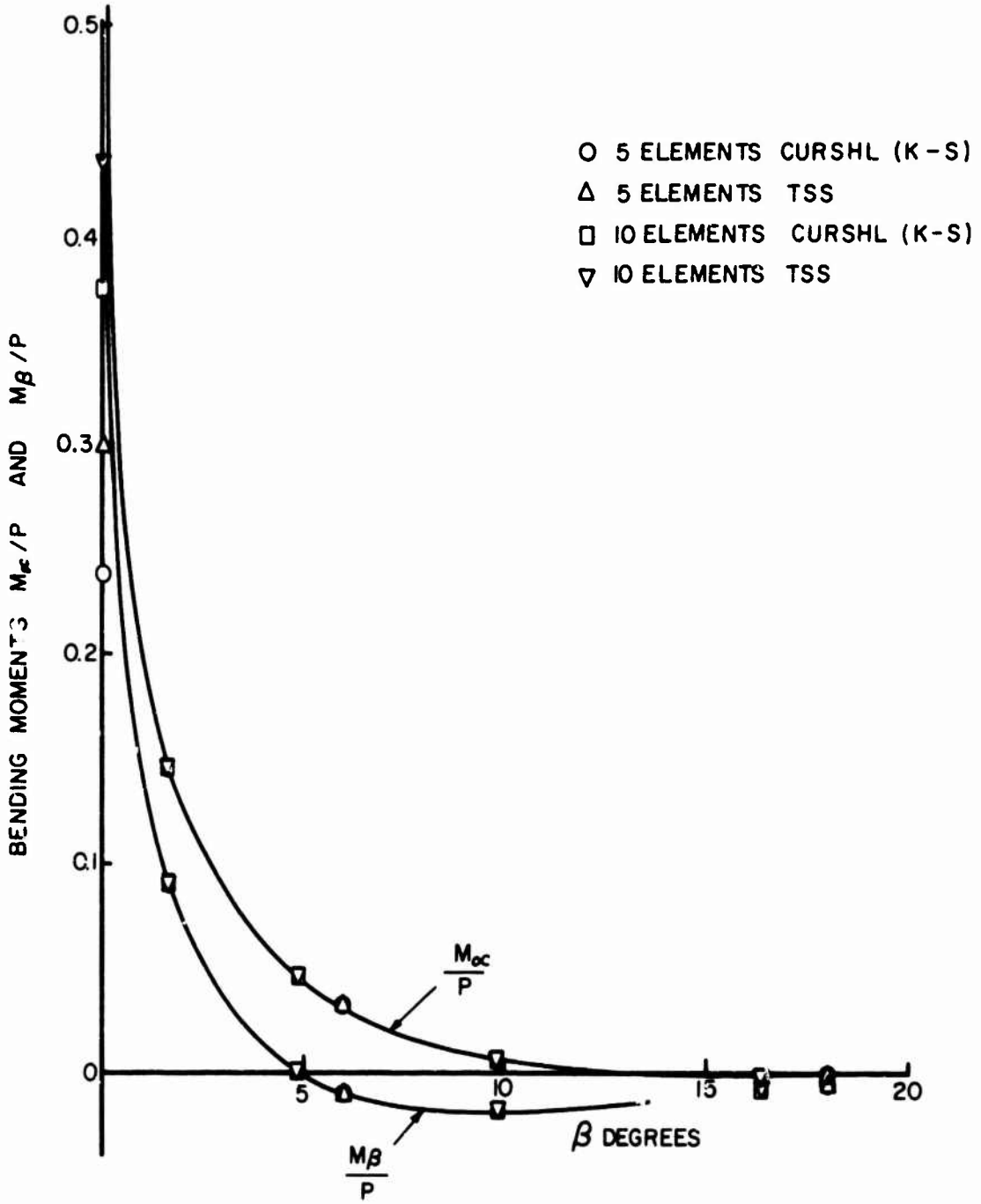


Figure 7. Moments Near Pole of Point Loaded Sphere
R/t = 50, ν = 0.3

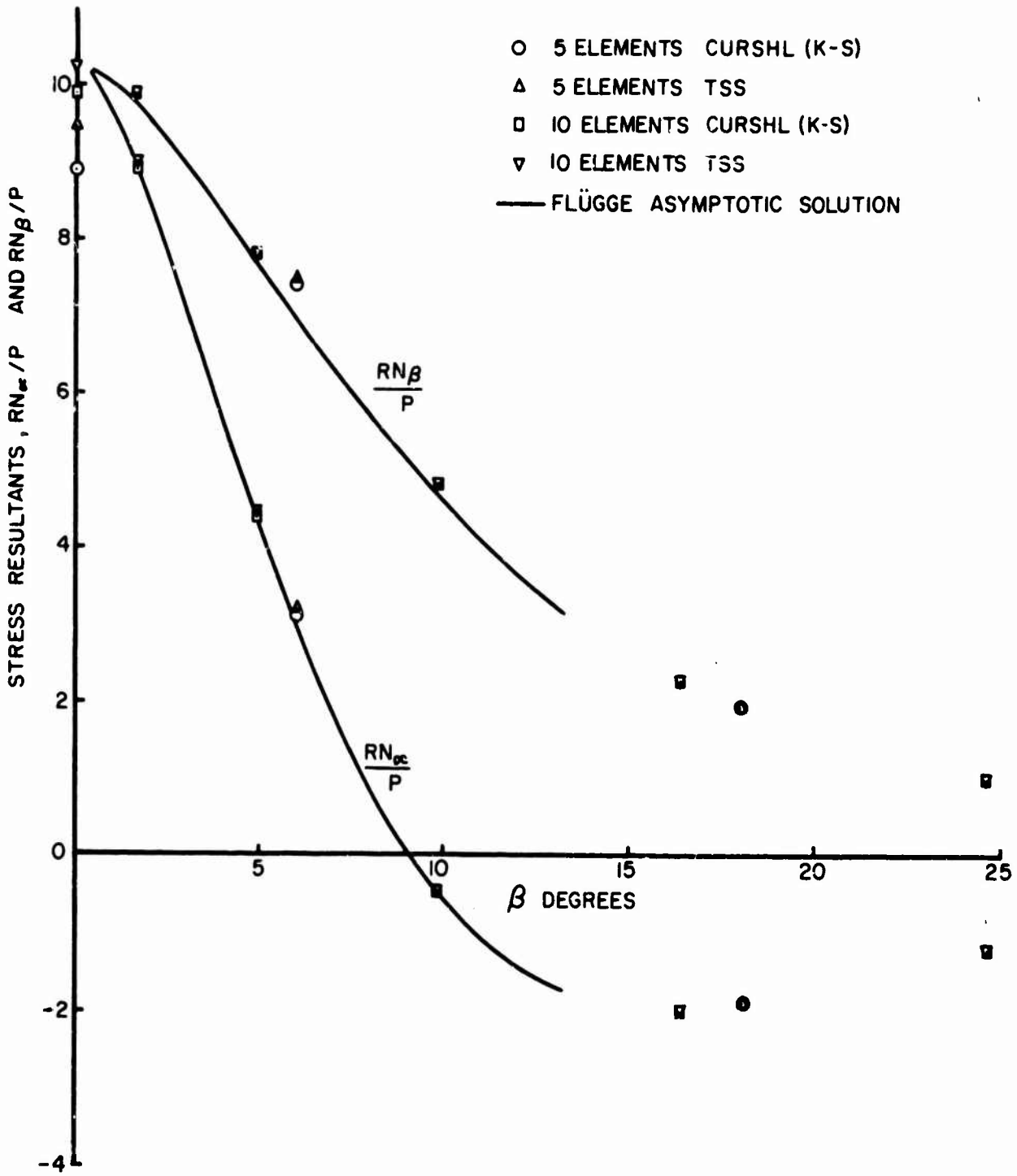


Figure 8. Stress Resultants Near Pole of Point Loaded Sphere
R/t = 50, ν = 0.3

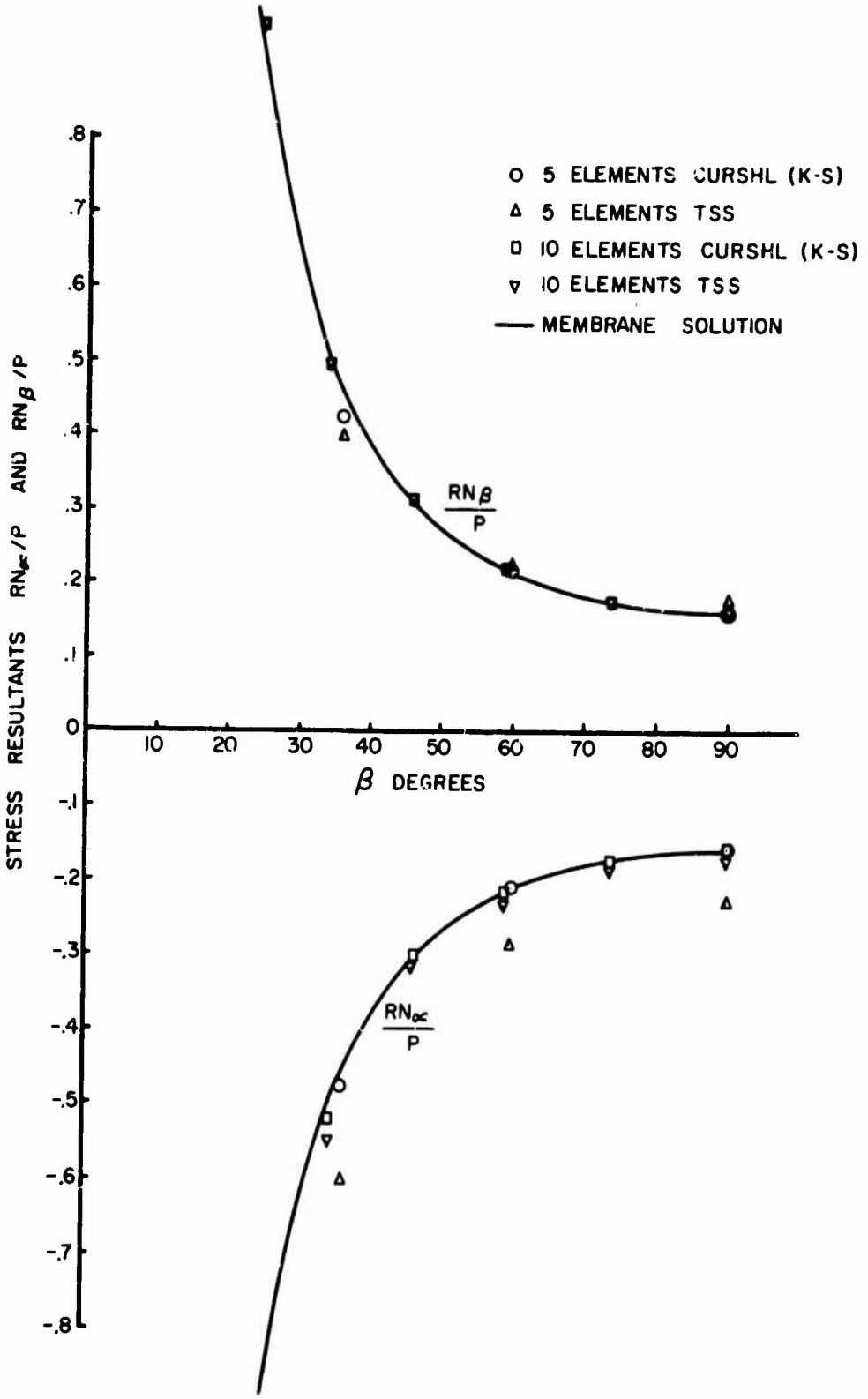


Figure 9. Stress Resultants Remote From Pole of Point Loaded Sphere: R/t = 50, ν = 0.3

precision arithmetic was used throughout.

CURSHL has been run on the IBM System/360 Time Sharing System using full H level Fortran. CURSHL (K-S) takes approximately 1.6 seconds of VMR time to generate stiffness matrix, while CURSHL (D-V) takes only 1.3 seconds. It must be remembered that an approximate integration scheme has been used; times would be increased if a more precise integration scheme were desired.

TSS has been run on the Michigan Terminal System (MTS) and only part of the program is in H level Fortran. TSS takes approximately 1.5 seconds of CPU time to generate one stiffness matrix. This includes an exact, closed form generation of the shallow shell stiffness matrix involved. Equation solving times are exactly the same since the same degrees of freedom are used in both approaches.

4.6 Comparison with other Elements

Space precludes extensive comparison of TSS, CURSHL and other elements. As well, writers seldom agree on the choice of test problems, making direct comparisons difficult. However, mention must be made of several other elements. Argyris and Scharpf [12] have developed SHEBA, a general triangular shell element that uses complete fifth order polynomials for u, v and w. This leads to an excessive, 18 degrees of freedom per node (see Section 2) and results using this element should be very similar to TSS or CURSHL at the expense of much greater computational effort. Key [3] has developed SLADE, a shell code based on a quadrilateral shell element which includes shear deformation. The element has nine degrees of freedom per node and is restricted to shells of revolution. SLADE gives good results but comparisons [10,19] indicate that the present approaches are superior. In addition, there are no shape limitations.

A third approach is that of Ahmad et al [23] who generate three-dimensional isoparametric finite elements into shell elements by restricting the variation of displacements through the thickness of each element. While valid for quite thick shells, their elements do not appear to compete with CURSHL or TSS for thin shells. Several other shell elements have been developed and are discussed at greater length in Ref. 10 and 19.

5.0 CONCLUSIONS

Two approaches to the finite element solution of arbitrary deep shells have been presented and compared. Both the arbitrary deep shell (CURSHL), and the transformed shallow shell (TSS) give excellent accuracy and both converge rapidly to the exact solutions as the number of elements used is increased. Good predictions of stress resultants and bending moments are obtained as well as of displacements. Engineering accuracy is always achieved with the use of a few elements.

CURSHL is a more flexible code with two general shell theories, Koiter-Sanders and Donnell-Vlasov, available as options, whereas TSS gives only Donnell-Vlasov theory solutions. While Donnell-Vlasov theory can be in error in special cases (as in Sec. 4.4), it is generally accepted that it is adequate for most practical shell problems where bending occurs in limited regions.

The predicted asymptotic rate of convergence for CURSHL of N^{-6} has been numerically confirmed and this rate is significantly greater than N^{-2}, the rate found for the non-conforming TSS approach. However, for a small number of elements, the accuracy of both methods is quite similar (though CURSHL generally gives better results). It has been demonstrated that the approximate numerical integration used in CURSHL is adequate and it is not necessary to explicitly include rigid body modes in an element. TSS is potentially the faster code, and it certainly is easier to program. With the verification of the validity and accuracy of these two approaches, perhaps the quest for 'a good, accurate, arbitrary triangular general shell finite element' may now be at an end.

REFERENCES

1. Gallagher, R., "Analysis of Plate and Shell Structures", Proc. of Symposium on Application of Finite Element Methods in Civil Engineering, Vanderbilt University, Nashville, Tennessee, Nov. 13-14, 1969.

2. Zienkiewicz, O.C., "The Finite Element Method: From Intuition to Generality", Applied Mechanics Reviews, Vol. 23, March 1970.

3. Key, S.W., "The Analysis of Thin Shells with a Doubly Curved Arbitrary Quadrilateral Finite Element", Paper presented at the Lockheed Conference on Computer Oriented Analysis of Shell Structures, Lockheed Palo Alto Research Lab., Palo Alto, Calif., Aug. 10-14, 1970.

4. Hartung, R.F., "An Assessment of Current Capability for Computer Analysis of Shell Structures", Air Force Flight Dynamics Lab., Air Force Systems Command, Wright Patterson Air Force Base, Dayton, Ohio, Technical Report AFFDL-TR-67-194, Preliminary Draft, Feb. 1970.

5. Lindberg, G.M., Olson, M.D., and Cowper, G.R., "New Developments in the Finite Element Analysis of Shells", National Research Council of Canada DME/NAE Quarterly Bulletin No. 1969(4).

6. Cowper, G.R., Kosko, E., Lindberg, G.M., and Olson, M.D., "A High Precision Triangular Plate-Bending Element", National Research Council of Canada, Aeronautical Report LR-514, Dec. 1968.

7. Cowper, G.R., Kosko, E., Lindberg, G.M., and Olson, M.D., "Static and Dynamic Applications of a High Precision Triangular Plate-Bending Element", AIAA Journal, Vol. 7, No. 10, pp. 1957-1965, 1969.

8. Cowper, G.R., Lindberg, G.M., and Olson, M.D., "A Shallow Shell Finite Element of Triangular Shape", Int. Journal of Solids and Structures, Vol. 6, pp. 1133-1156, 1970.

9. Lindberg, G.M., and Olson, M.D., "A High Precision Triangular Cylindrical Shell Finite Element", AIAA Journal, Vol. 9, No. 3, pp. 530-532, 1971.

10. Cowper, G.R., "CURSHL: A High-Precision Finite Element for Shells of Arbitrary Shape", National Research Council of Canada, Aeronautical Report, 1971 (in preparation).

11. McLay, R.W., "Completeness and Convergence Properties of Finite Element Displacement Functions: A General Treatment", AIAA Paper No. 67-143, AIAA 5th Aerospace Sciences Meeting, New York, Jan. 1967.

12. Argyris, J.H., and Scharpf, D.W., "The SHEBA Family of Shell Elements for the Matrix Displacement Method", Aeronautical Journal, Vol. 72, No. 694, pp. 873-883, 1968.

13. Greene, B.E., Jones, R.E., McLay, R.W., and Strome, D.R., "Dynamic Analysis of Shells Using Doubly-Curved Finite Elements", Proc. of the Second Conference on Matrix Methods in Structural Mechanics, AFFDL-TR-68-150, Wright-Patterson Air Force Base, October 1968.

14. Koiter, W.T., "A Consistent First Approximation in the General Theory of Elastic Shells", Proc. of IUTAM Symposium on the Theory of Thin Elastic Shells, Delft, 1959.

15. Budiansky, B., and Sanders, J.L., Jr., "On the 'Best' First-Order Linear Shell Theory", Progress in Applied Mechanics, (The Prager Anniversary Volume), The Macmillan Co., New York, 1963.

16. Novozhilov, V.V., "Thin Shell Theory", Sec. 17, P. Noordhoff Ltd., Gronigen, The Netherlands, 2nd Ed., 1964.

17. Green, A.E., and Zerna, W. "Theoretical Elasticity", Chapter XI, Oxford, Clarendon Press, 1954.

18. Ibid., Chapter I.

19. Olson, M.D., "Analysis of Arbitrary Shells Using Shallow Shell Finite Elements", Dept. of Civil Engineering Structural Research Series Report, University of British Columbia, 1971 (in preparation).

20. Koiter, W.T., "A Spherical Shell Under Point Loads at its Poles", Progress in Applied Mechanics (The Prager Anniversary Volume), Macmillan, New York, 1963.

21. Timoshenko, S.P., and Woinowsky-Krieger, S., "Theory of Plates and Shells", 2nd Ed., pp. 558-561, McGraw-Hill, New York, 1959.

22. Flügge, W., "Stresses in Shells", pp. 350-353, Springer-Verlag, Berlin/Gottingen/Heidelberg, 1960.

23. Ahmad, S., Irons, B.M. and Zienkiewicz, O.C., "Analysis of Thick and Thin Shell Structures by Curved Finite Elements", Int. J. for Num. Meth. in Eng., Vol. 2, No. 3, 1970, pp. 419-451.

IMPROVEMENTS OF FINITE ELEMENT SOLUTIONS FOR STRUCTURAL AND NON STRUCTURAL APPLICATIONS

G. SANDER* and P. BECKERS**

Laboratoire de Techniques Aeronautiques et Spatiales

University of Liege, Belgium

The paper presents a new formulation which allows the derivation of finite element properties in terms of stress functions. It is applicable to both displacement and equilibrium models and leads for certain classes of problems to a substantial reduction in the number of unknowns. The choice of the stress functions as unknowns is identified with an automatic selection procedure of redundancies of minimal diffusion. Therefore this method can be interpreted as a force method in which the numerical search for the redundancies is avoided. It is shown to apply particularly well in certain non structural problems. Independently the effect of the introduction of internal degrees of freedom (bubble functions) in finite elements is evaluated. The conclusion is that a few of such modes can improve significantly the finite element properties and bring them practically in coincidence with assumed stress models.

* Associate Professor

** Assistant

1. INTRODUCTION.

The finite element method provides a powerful tool for structural analysis. Many problems of linear and non linear elasticity have been successfully solved and applications to other domains such as fluid mecnanics and heat conduction have been proposed (Ref. 14). These analyses imply the solution of large systems of linear algebraic equations. If, in aerospace applications, the cost of these solutions is not often a limiting factor, especially when compared to the cost of preparing input data and presenting the output, the size of the proolems that can be handled is still limited by the presently available computers and programmes. This limitation is due not only to the core storage size but also to the danger of round-off error propagation. Therefore a need exists for improved finite elements as well as improved algorithms of solution to minimize the size of the systems of equations to handle.

Most of the applications of the finite element method have been conducted using either conforming or non conforming finite elements while the algorithm of solution uses displacements as unknowns. Various attempts have been made to use other formulations. The force method which was proposed since the early days of the finite element method has always been handicapped by the need of a numerical search for the redundancies. The equilibrium elements, although their assumptions seem more appealing to the structural analyst, have received only few practical applications due to the higher number of unknowns they involve (Ref. 1,3,5β). Therefore the possibilities offered by the dual analysis of a given problem, that is comparing displacement and equilibrium solutions, has remained unapplied.

The present paper proposes an alternative formulation of the finite element properties which uses a different set of variables : contour deformations can be used for displacement models and stress function modes for equilibrium models. In both cases the natural unknowns are stress function modes. Furthermore the choice of these unknowns can be interpreted as an automatic selection procedure of redundancies of minimal diffusion in the structure (Ref. 11). It follows that, when using these variables, the numerical search for the redundancies in the force method is avoided and the algorithm of solution becomes identical to that of the displacement method. The duality between the two sets of variables in both displacement and equilibrium models and the two solution methods becomes more complete.

A simple criterion is proposed to decide which formulation is the most economic. The problem of introducing the boundary conditions is examined for structural as well as for non structural applications.

Independently the effect of the introduction of internal degrees of freedom by bubble functions in the finite elements (Ref. 3,5) is evaluated. The conclusion is that a few of such modes can improve significantly the finite element properties and bring them practically in coincidence with assumed stress models.

2. FORMULATION FOR CONFORMING DISPLACEMENT MODELS

They are derived by application of the minimum total energy principle

$$\delta(U + P) = \delta\left(\frac{1}{2}\int_{vol} \varepsilon' H \varepsilon \, dvol - \int_{\Gamma_\sigma} \bar{t}' u \, d\Gamma_\sigma\right) = 0 \qquad (1)$$

where U is the strain energy

P is the potential energy (in which the body forces are not considered here)

ε is the column of the strain components

H is a symmetric matrix of elastic coefficients

t is the column of surface tractions on Γ_σ

u is the displacement vector

Γ_σ is the part of the boundary where surface tractions are prescribed.

The prime denotes transposition and the bar prescribed values.

The principle may be applied only if the displacement field satisfies a priori the compatibility conditions :

$\varepsilon_{ij} = D_j u_i$ or in matrix form, $\varepsilon = \partial u$ in the volume

$u = \bar{u}$ on the part Γ_u of the boundary where the displacements are prescribed.

The displacement field in a finite element is discretized in the form

$$u = M a \qquad (2)$$

where M is a matrix of assumed modes and a the column of their n_a (unknown) intensities. The column of the strain tensor components is written

$$\varepsilon = \partial M \, a \qquad (3)$$

and the strain energy becomes

$$U = \frac{1}{2}\int_{vol} a' \, \partial M' \, H \, \partial M \, a \, dvol = \frac{1}{2} a' \, I \, a \qquad (4)$$

When using the minimum total energy principle, the stresses are only defined by

$$\sigma = \frac{\partial U}{\partial \varepsilon} \qquad (5)$$

With the discretization adopted, the stress field is described by a set of parameters b

$$b = \frac{\partial U}{\partial a} = I\ a = \int_{vol} \partial M'\ \sigma\ dvol \qquad (6)$$

which are weighted averages of the stresses and yield therefore only a "weak" knowledge of the stress state in the element. From (6) the matrix I is clearly a generalized form of the stress-strain relations. A detailed description of the stresses is often derived from (3)

$$\sigma = H\ \partial M\ a \qquad (7)$$

This distribution is however only a particular example of the stress fields satisfying (6) and in interpreting the stress output, it is worth to remember that any other stress field satisfying (6) is equally consistent. This means that stress distributions can be superimposed to (7) provided they yield a zero value to the generalized stresses (6).

The compatibility conditions in the volume are automatically satisfied by the assumptions (2). To insure conformity along the interfaces, a number n_c of interface displacement modes of intensities q_c has to be defined in such a way that the displacements along the element boundaries are uniquely defined. As the displacement field has to be unique at a vertex common to different elements, q_c always contains the local values of the displacement components at the vertices. In addition other local values, derivatives or integrals of displacements can be used as boundary modes along the interfaces. The expression of the boundary modes in terms of the parameters a_i is denoted

$$q_c = C\ a \qquad (8)$$

where the matrix C is called the kinematic matrix of the element. It is not necessarily a square matrix and hence not necessarily invertible.

In the absence of body forces, the virtual work (V.W.) consist only in the work done by the prescribed surface tractions.

$$V.W. = \int_{\Gamma_\sigma} u'\bar{t}\ d\Gamma_\sigma = a'b = a' \left[\int_{\Gamma_\sigma} M'\bar{t}\ d\Gamma_\sigma \right] \qquad (9)$$

or alternatively

$$V.W. = q'_c\ g'_c = a'\ C'\ g_c \qquad (10)$$

where g_c is the set of generalized loads conjugate to the interface modes. Equation (9) yields another expression for the generalized stresses b in terms

of surface tractions, while comparison of (9) and (10) shows that

$$b = C' g_c \tag{11}$$

The transpose of the kinematic matrix is a static matrix relating the generalized stress parameters to the applied loads.

The problem of the derivation of the finite element can now be presented as that of finding the minimum of the functional :

$$(\tfrac{1}{2} a' I a - q'_c g_c)_{min} \tag{12}$$

subject to the constraints (8). Three cases have now to be distinguished :

a) $n_a = n_c$

The number of parameters of the displacement field equals the number of interface connection modes. The matrix C is square, non singular. This allows to solve for the constraints. Inverting (8) and substituting into (12) yields

$$(\tfrac{1}{2} q'_c C^{-1\prime} I \bar{C}^{-1} q_c - q'_c g_c)_{min}$$

which furnishes the familiar stiffness relation

$$g_c = K_c q_c = C^{-1\prime} I C^{-1} q_c \tag{13}$$

between the interface modes and the corresponding generalized forces. The latter can now be interpreted, using (9) and (11)

$$g_c = C^{-1\prime} \int_{\Gamma_\sigma} H' \bar{t} \, d\Gamma_\sigma \tag{14}$$

They are weighted averages of the prescribed surface tractions, giving again a "weak" knowledge of the stress state on Γ_σ

b) $n_a > n_c$

In this case the homogeneous system

$$C a = 0 \tag{15}$$

has $n_1 = n_a - n_c$ non trivial solutions. Let s_i denote such solutions and group them in a matrix S.

$$S = (s_1, s_2, s_3 \ldots s_{n_1})$$

The non zero solutions of (15) can be written

$$a = Sw$$

where w is any arbitrary vector. Such modes of the parameters a_i exist with zero displacement prescribed along the interfaces. For this reason they are sometimes called "bubble" modes of displacement. The constraints (8) can now be solved in the form

$$a = T\,q_c + S\,a_x \tag{16}$$

where Tq_c is a particular solution and Sa_x the complementary solution. In general the matrices T and S could be obtained numerically, by a Gauss Jordan elimination technique for instance. However this operation can always be avoided by one of the following methods.

Add to (8) a set of n_i internal values of the displacements q_i and form an enlarged connection matrix

$$q = \begin{vmatrix} q_c \\ \\ q_i \end{vmatrix} = \begin{vmatrix} C \\ \\ C_i \end{vmatrix} a = C^* a \tag{17}$$

C^* is now non-singular and the general procedure described above can be applied. The displacements q_i can finally be eliminated by the standard condensation process.

Another simpler possibility arises when the displacement field can be a priori split in two parts, the first controlling the interface modes and the second the "bubble" modes

$$u = M\,a = M_c\,a_c + M_b\,a_b \tag{18}$$

In this case $$q_c = C_c\,a_c \tag{19}$$

is invertible and the elimination of the modes a_b requires simply a condensation of the matrix I. The stiffness matrix turns out to be

$$K_c = C_c^{-1\prime}\,\left[\, I_{cc} - I_{cb}\, I_{bb}^{-1}\, I_{bc} \,\right] C_c^{-1} \tag{20}$$

if I_{cc}, I_{bb}, I_{bc} are the submatrices of I corresponding to a_c and a_b. This latter procedure represents a significant saving of numerical work compared to the first one.

c) $n_a < n_c$

The constraints (3) cannot be inverted and this case was ruled out of previous formulations (Ref. 1). It is however possible to derive a finite element in this case.

Considering the static relation (11) the number n_b of generalized stress parameters b is smaller than the number n_g of generalized interface loads g_c.
This means that $n_k = n_g - n_b$ non-zero solutions exist to the problem

$$C' g_c = 0 \tag{21}$$

These solutions correspond to combinations of loads applied along the interfaces which do not generate stresses in the element. By analogy with the terminology adapted for equilibrium models these modes might be called "spurious static modes".

To derive the finite element properties, let us split the constraints in two parts

$$\begin{aligned} q_c^* &= C^* a \\ q_c &= C_r a \end{aligned} \tag{22}$$

where q_c^* is a subset of q_c composed of n_a components arbitrarily selected while q_r forms the complementary subset of q_c^*. The matrix C^* is non-singular and (22) yields

$$a = C^{*-1} q_c^*$$

Substitution in the expression of the minimum principle yields

$$(\frac{1}{2} q_c^{*\prime} C^{*-1\prime} I C^{*-1} q_c^* - q_c^{*\prime} g_c^* - q'_r g_r) \text{ min} \tag{23}$$

with the remaining constraints

$$q_r = C_r C^{*-1} q_c^* \tag{24}$$

Introducing these remaining constraints in (23) affected by a column of lagrangian multiplier λ, it turns out that

$$\begin{aligned} K_c^* q_c^* &= g_c^* + C^{*-1\prime} C_r \lambda \quad \text{with} \quad K_c^* = C^{*-1\prime} I C^{*-1} \\ g_r &= \lambda \\ q_r &= C_r C^{*-1} q_c \end{aligned} \tag{25}$$

This procedure yields a stiffness matrix K_c^* which can be used if the satisfaction of the remaining constraints is achieved at the structural level. It appears often in practice (Ref. 2,3) that the remaining constraints can be eliminated in a group of finite elements. This allows the derivation of super-elements free of such constraints. This procedure justifies a posteriori and generalizes the derivation of certain previously developped elements like the conforming quadrilateral plate bending elements (Ref. 4,5) or its triangular version (Ref. 2,6). In fact a complete family of such elements exists up to any degree (Ref. 2).

The distinction between the three cases described above is based on the comparison between the numbers n_a and n_c. However it is possible that bubble modes and spurious static modes exist simultaneously in a model. Therefore it is more correct to base the distinction on the number of solutions of the homogeneous problems (15) and (21) and, if such solutions exist, apply the corresponding treatment. In general, a model will be free of remaining constraints if

$$n_a - n_i = n_c$$

There is another possible formulation for displacement models which although it appears less physical, can be useful in certain circumstances and allows a better understanding of the next section.

Consider the description of the strain field in (3). It does not depend upon the rigid body modes of the displacement field and therefore it can be described in terms of a reduced set a^* of parameters

$$\varepsilon = \overline{\partial M}\, a^*$$

The relation between the two sets of parameters can be written in general

$$a^* = D\, a \qquad (26)$$

but in some cases a^* is simply a subset of a. The strain energy is now written

$$U = \frac{1}{2} a^{*\prime} \left[\int_{vol} \overline{\partial \Pi}' \, H \, \overline{\partial H} \, dvol \right] a^{*} = \frac{1}{2} a^{*\prime} I^{*} a^{*} \qquad (27)$$

where I^{*} is now non-singular.
Using (26) the relation with (4) is

$$I = D' I^{*} D$$

showing that I^{*} is a non singular kernel of I.

Instead of expressing the continuity of the displacements between elements in terms of displacement modes, one could adopt a system of contour deformation modes providing the same continuity. Let us denote by d_c the set of these contour deformation modes. They can be expressed in terms of the parameters of the deformation by

$$d_c = Q a^{*} \qquad (28)$$

where Q is another kinematic matrix. The properties of the finite element are obtained by minimization of

$$\left(\frac{1}{2} a^{*\prime} I^{*} a^{*} - d'_c f_c \right) \min \qquad (29)$$

subject to the constraints (28) and where f_c denotes the generalized contour forces conjugate of the contour deformation modes d_c.
Substituting (28) in (29) and using the fact that I^{*} is non singular, we obtain

$$I^{*} a^{*} = a' f$$

$$d_c = Q a^{*} = Q I^{*-1} Q' f = F f \qquad (30)$$

where F is now a flexibility matrix.

Apparently the discussion of the properties of the kinematic matrix Q is unnecessary in this formulation where we obtain in any case a flexibility matrix. However the difficulty is only hidden as will appear in the next section, where the same problem arises in the derivation of equilibrium elements.

Finally the following relations between the two formulations are easily obtained

$$d_c = Q\, a^* = Q\, D\, a = a\, D\, C^{-1}\, q_c = C^*\, q_c$$

where C^* is a square matrix relating the contour deformation modes to the interface displacement modes.
By virtual work we also have

$$V.W. = g'_c\, q_c = f'_c\, d_c = f'_c\, C^*\, q_c$$

$$g_c = C^{*t}\, f_c$$

3. FORMULATION FOR STRESS-DIFFUSING EQUILIBRIUM MODELS

Stress diffusing equilibrium models are derived by application of the minimum complementary energy principle

$$\delta(\psi + Q) = \delta\left(\frac{1}{2}\int_{vol} \sigma'\, H^{-1}\sigma\, dvol - \int_{\Gamma_u} t'\bar{u}\, d\Gamma_u\right) = 0$$

where ψ and Q denotes respectively the complementary stress energy and the potential energy, while σ is a column of the stress tensor components. The application of this principle implies the satisfaction of the equilibrium equations

$$D_j\, \sigma_{ij} + X_i = 0 \qquad \text{in the volume}$$

$$t_i = \ell_j\, \sigma_{ji} = \bar{t}_i \qquad \text{on } \Gamma_\sigma$$

(ℓ_j are the direction cosines of the outward normal).
In the following the case without body forces only will be considered for simplicity.

The part of the stress field satisfying the homogeneous equilibrium equations can be discretized by two different procedures :

a) Each stress component is represented by a linear combination of assumed modes $P_i\ (x_m)$ of intensity b_i

$$\sigma = P\ b \tag{31}$$

These modes must be adjusted in such a way that the equilibrium equations in the volume are satisfied.

b) The other procedure is relatively new and reveals extremely fruitful for the development of finite elements as well as it provides a new method of solution of the global problem.

The stress tensor can be derived from a set of stress functions

$$\sigma_{ij} = \epsilon_{imr}\ \epsilon_{jns}\ \phi_{rs,\ mn} \qquad \text{(ij not summed)} \tag{32}$$

where, in the three dimensional case ϕ_{rs} is a symmetric tensor of stress functions, while ϵ_{ijk} is the alternator symbol. Note that for membrane problems ϕ_{rs} reduces to the Airy stress function and for plate bending problems to two stress functions (Ref. 7).

These stress functions can be approximated as a linear combination of modes $H_{i(rs)}\ (x_m)$ of intensity a_i

$$\phi_{rs} = \sum_k H_{k(rs)}\ (x_m)\ a_k$$

or, in matrix form

$$f = H\ a \tag{33}$$

The stress field derived according to (30) automatically satisfies the homogeneous equilibrium equations

$$\sigma = \partial^2 H\ a$$

The number of parameters a_i is always larger than that of the b_i by the number of integration constants of the system (30) which are the "rigid body modes" of the stress functions. In some cases b can be simply a subset of a, but in general the relation between the two sets is denoted

$$b = D\ a \tag{34}$$

The complementary strain energy ψ can be alternatively written

$$\psi = \frac{1}{2}\int_{vol} b'\ P'\ H^{-1}\ P\ b\ dvol = \frac{1}{2} b' J^*\ b$$

$$\psi = \frac{1}{2}\int_{vol} a'\ \partial^2 H'\ H^{-1}\ \partial^2 H\ a\ dvol = \frac{1}{2}\ a'\ J\ a \tag{35}$$

where J^* and J are two flexibility matrices.

The generalized strains can be defined by

$$\pi_b = \frac{\partial \psi}{\partial b} = J^* v = \int_{vol} P \varepsilon' \, dvol \tag{36}$$

$$\pi_a = \frac{\partial \psi}{\partial a} = J a = \int_{vol} \partial^2 M \varepsilon' \, dvol \tag{37}$$

They are weighted averages of strains and this corresponds only to a "weak" knowledge of the strain field. The parameters b_i are independent and each of them represents a non-zero contribution to the strain energy. The matrix J^* is therefore non-singular. As the parameters a_i contains the "rigid body modes" of the stress functions which do not contribute to the strain energy, the matrix J is singular. Using (33) the relation between the two is

$$J = D' J^* D \tag{38}$$

which shows that J^* is a non-singular kernel of J. In view of (36) and (37) both can be interpreted as generalized strain-stress relations.

Let us turn next to the equilibrium requirements along the boundary. They concern only the interfaces as we have ruled out the body forces. Using the first description of the stress field (29), the surface tractions can be expressed by

$$t = L \sigma = L P b \tag{39}$$

where L is a matrix of the direction cosines ℓ_j. The surface traction modes described by t have to be uniquely determined by a set of n_c forces denoted g_c. These can be local values or derivatives, integrals or any combinations of them provided that along each edge the definition is unique.
The expression of these generalized forces in terms of the stress fields parameters is written

$$g_c = C b \tag{40}$$

where C is called the static matrix of the element. It is in general a rectangular matrix, but even if it is square, the relation (40) cannot be inverted since the generalized forces are a priori not independent since they satisfy the global equilibrium equations. Note that in contrast to the formulation of the displacement elements, the stress tensor has no unique value at a vertex. Therefore in pure equilibrium models, there are no variables associated with a vertex and the forces g_c are essentially interface variables.

The virtual work consists only in the work done by the prescribed displacements and can be written alternatively

$$V.W. = \int_{\Gamma_u} u' \, t \, d\Gamma_u = \int_{\Gamma_u} u' \, L \, P \, b \, d\Gamma_u = \pi'_b \, b \tag{41}$$

$$V.W. = q'_c \, g_c = q'_c \, C \, b \tag{42}$$

where q_c are the generalized (average) displacements conjugate to g_c. Comparison of the two expressions yields

$$\pi_b = C' \, q_c \tag{43}$$

The transpose static matrix turns out to be a kinematic matrix relating the generalized (average) strains to the generalized interface displacements.

The problem of the derivation of the finite element reduces now to that of finding the minimum of the functional

$$\left(\frac{1}{2} b' \, J^* \, b - g'_c \, q_c \right) \min \tag{44}$$

subject to the constraints (40). It yields

$$J^* b = C' \, q_c \tag{45}$$

and as J^* is non singular

$$g_c = C \, b = C \, J^{*-1} \, C' \, q_c = K_c \, q_c \tag{46}$$

which is the familiar form of the stiffness matrix for an equilibrium element. This derivation of the stiffness matrix for equilibrium models in terms of generalized forces appears formally identical with the derivation of the flexibility matrix for conforming displacement elements in terms of contour deformations. In this former case it was apparently unnecessary to examine the properties of the kinematic matrix Q to obtain a flexibility matrix. Here also a stiffness matrix is obtained whatever the properties of the static matrix C are. The difficulty is however only hidden and the distinction between these cases is necessary if one wishes to use in practice the relations (46) or by analogy (30).

Consider first the static matrix defined in (40). The homogeneous system can have non trivial solutions representing internal stress modes which exist with zero prescribed surface tractions on the interfaces. These n_i modes are "bubble" stress modes of the element. It is however useless to define here internal generalized forces to determine these modes in a similar way to the internal displacements defined in (17). In fact the solution of the constraints (40) is not necessary and the bubble stress modes do not contribute to the complementary potential energy. Therefore the reduction of these modes, whatever they are, is automatically included in the expression (46).

Consider next the kinematic matrix defined in (43). The homogeneous system has always non-trivial solutions or, more physically, displacement modes of the interfaces which do not produce generalized strains. Among these are the rigid body modes of elements. If all these solutions in number n_k are grouped in a matrix Z, the solutions of the homogeneous system (43) can be expressed by

$$q_c = Z\, w \tag{47}$$

where w is any arbitrary vector of dimension n_k. The virtual work of such modes is zero

$$V.W. = g'_c\, q_c = g'_c\, Z\, w = 0 \tag{48}$$

As w is arbitrary, it implies n_k relations to be satisfied by the generalized forces

$$Z'\, g_c = 0 \tag{49}$$

If n_k equals the number n_r of rigid body modes these relations are simply the global equilibrium equations. If n_k is larger than n_r the stiffness matrix defined in (46) cannot be used unless the $n_k - n_r$ remaining constraints are satisfied. The displacement modes corresponding to these remaining constraints have been called spurious kinematic modes (Ref. 8).

Here again it appears often that such constraints can be eliminated in a group of elements, leading to a super-element free of such constraints. The procedure has been applied implicitly in the derivation of a certain number of equilibrium membrane elements (Ref. 2,5). The advantage of the present formulation is to show that the procedure is general and that the problem is strictly the same as in displacement elements. Finally it appears that the condition to fulfil to avoid remaining constraints (or spurious kinematic modes) is that

$$n_b - n_i = n_c - n'_r \tag{50}$$

The strict similarity between the derivation of displacement and equilibrium elements becomes even more transparent if the second formulation, in terms of stress functions, is used. In that case the requirements of surface traction continuity has to be translated into requirements of stress function continuity. They certainly involve the continuity of the stress function local values but in addition, they eventually involve the continuity of some of its derivatives. Those requirements are especially simple to formulate in the case of membrane and plate bending problems if the Southwell slab and-

logies are recalled (Ref. 7) and specific examples are given in the following. In general it requires to insure the single valuedness of m_c local value of the stress functions and its derivatives denoted f_c. These m_c stress function modes are equivalent to the n_c forces g_c. They can be expressed in terms of the intensities of the assumed stress function modes a_i by

$$f_c = T\, a \tag{51}$$

The matrix T is another static matrix of the element. Note that the m_c interface connection modes necessarily define the "rigid body" modes of the stress functions and that their number m_c is equal to n_c. In contrast with the first formulation, the generalized stress function modes f_c are not necessarily interface variables only, as the continuity often requires a single value of the stress function at a vertex.

The virtual work of the surface tractions can be alternatively written

$$\begin{aligned} V.W. &= \int_{\Gamma_u} u'\, t\, d\Gamma_u = \int_{\Gamma_u} u'\, \partial^2\, M\, a\, d\Gamma_u = \pi'_a\, a \\ V.W. &= f'_c\, d_c = a'\, T'\, d_c \end{aligned} \tag{52}$$

where d_c are generalized displacements conjugate of f_c. Comparison of the two expressions reveals that

$$\pi_a = T'\, d_c \tag{53}$$

which again indicates that the transpose static matrix is a kinematic matrix relating the strains π_a to the displacements d_c.

The properties of the finite element are now obtained by finding the minimum of

$$(\tfrac{1}{2}\, a'\, J\, a - f'_c\, d_c)\ \min \tag{54}$$

subject to the constraints (51). As J is singular, the procedure applied in the first formulation does not apply and the inversion of the relation (51) is necessary. Comparison of (54) and (12) shows that the situation is formally identical to that of deriving a displacement element if f_c

plays the role of q_c and d_c that of g_c. The solution of the constraints (51) requests again the distinction between three cases.

a) $n_a = n_c$

In this case it turns out that

$$a = T^{-1} f_c$$

and $$d_c = T^{-1}{}' J T^{-1} f_c = F_c f_c \tag{55}$$

where F_c is the flexibility matrix of the element in terms of the stress function modes.

b) $n_a > n_c$

The homogeneous system

$$T a = 0 \tag{56}$$

has n_i non-trivial solutions which are internal or "bubble" stress function modes. The procedure for their elimination can be formally identical to that described for the displacement models.

c) $n_a < n_c$

The homogeneous system

$$T' a = 0 \tag{57}$$

has n_k non-trivial solutions corresponding to spurious kinematic deformation modes. A flexibility matrix can still be derived by selecting arbitrarily a subset of independent values of stress function modes to invert (51). In this case the flexibility matrix cannot be used without the n_k remaining constraints. Again these constraints can often be eliminated in a group of elements.

Finally the two formulations for the equilibrium elements are related to each other by the following relations

$$g_c = C b = C J T^{-1} f_c = C^* f_c \tag{58}$$

Note that although the matrix C^* is square it is always singular as the forces g_c are not independent. By virtual work

$$V.W. = g'_c q_c = f'_c C^{*\prime} q_c = f'_c d_c \tag{59}$$

it appears

$$d_c = C^{*\prime} q_c \tag{60}$$

which completes the analogy.

4. DUALITY BETWEEN DISPLACEMENT AND EQUILIBRIUM MODELS

Before discussing the relative advantages of the two formulations, proposed above, their practical meaning in elasticity is illustrated in an example. For this purpose the formulation of the Southwell slab analogies is briefly recalled (Ref. 7).

A first analogy allows to show that the function used to describe the deflection $w(x,y)$ of a Kirchhoff plate can be used as the Airy stress function $\phi(x,y)$ of a membrane and that the continuity requirements for this function in order to derive a conforming displacement plate bending element are precisely the same as those required for the stress function in order to derive an equilibrium membrane element. Therefore the derivation of the flexibility matrix for an equilibrium element in the form required by the stress function formulation is identical to that of the stiffness matrix for a conforming plate bending element in the displacement formulation. The deflection has simply to be interpreted as the Airy stress function and the stress-strain relations substituted for the moment-curvature relations. At this time it may be better to replace the word displacement by deformation function in order to make the analogy more transparent. Note that as a by-product the first analogy has recently allowed to define new families of triangular and quadrangular equilibrium membrane elements and at the same time of conforming plate bending elements up to any degree, by forming super-elements composed of triangles (Ref. 2,3). In addition sub-equilibrium and hyper-equilibrium membrane elements are recognized as the analog respectively of non conforming and hyper-conforming plate bending elements.

A second analogy allows to show that the displacement functions u, v used in conforming membrane elements are precisely the stress functions required to derive equilibrium plate bending elements and that the continuity conditions of these two functions are also identical in the two problems. Therefore the families of conforming membrane elements of triangular and quadrangular shape recognized by many authors (Ref. 12) furnish immediately the flexibility matrices for equilibrium plate bending elements as required by the stress function method.

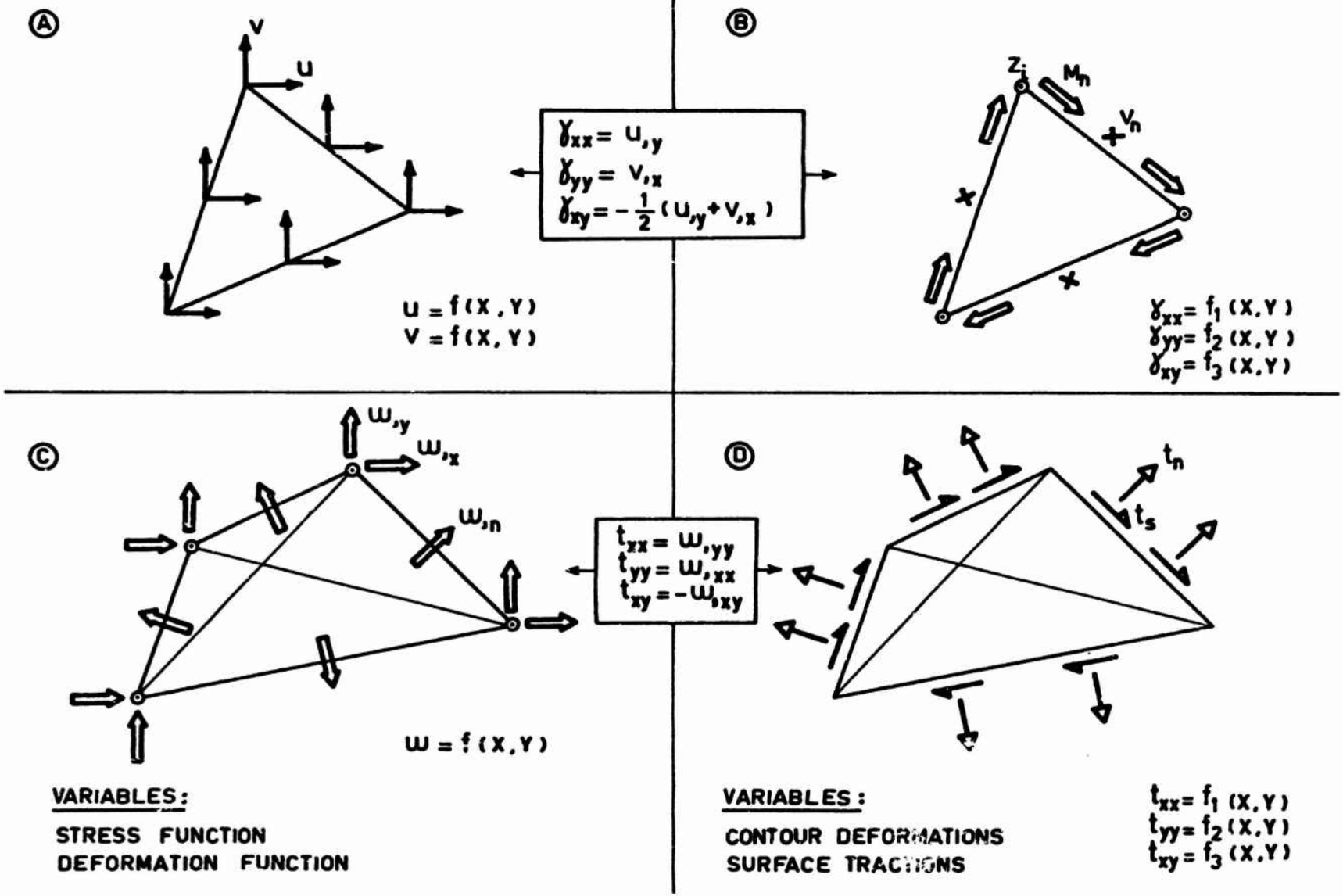
The two analogies are illustrated in figure 1. The elements (A) and (D) correspond to dual membrane elements, one conforming, the other equilibrium for which the choice of the unknowns leads to a stiffness matrix. (B) and (C) are dual plate bending elements for which the choice of the unknowns also yields a stiffness matrix. Another interpretation of the analogies is possible : (C) and (D) represent the same equilibrium membrane

element formulated either in terms of stress function and leading to a flexibility matrix, or in terms of surface tractions and yielding a stiffness matrix. But (C) and (D) also represent the two formulations of the conforming plate bending element either in terms of deformation functions (displacements) or in terms of contour deformations. The same double interpretation holds for (A) and (B).

The interpretation of these contour deformations follows immediately by application of the Southwell analogies : in the case (B) a corner load corresponds to the variation of the angle at a vertex, a bending moment to the extension of the edge, etc (Ref. 7).

Finally the following interpretation of the local values of the stress functions or deformation functions is important for what follows. In a group of equilibrium membrane elements, say of type (C), formulated in terms of stress functions, the Airy stress function is single-valued along the interfaces in the absence of loads. Therefore a local value of the stress function represents a self stressing state (Ref. 11). This is easy to realize if one imagines the stress state produced by one single local value different from zero. By analogy a local value of the deformation function (a local displacement) can be interpreted as a self straining state. By virtual work considerations the variables conjugate of the stress function modes are generalized cuts while those conjugate of the contour deformations are stress functions.

THE SOUTHWELL ANALOGIES



VARIABLES:

STRESS FUNCTION
DEFORMATION FUNCTION

VARIABLES:

CONTOUR DEFORMATIONS
SURFACE TRACTIONS

FIGURE 1

5. METHODS OF SOLUTION OF THE GLOBAL PROBLEM

Considering the strict analogy between the derivation of displacement and equilibrium models, the methods of solution of the global problem will be examined only for one category of models and the results translated for the other. As the equilibrium models are less well known, we shall concentrate on them in the following.

The complementary energy principle applied to an assemblage of equilibrium finite elements can be written, using the formulation (44) in terms of surface tractions

$$\left(\frac{1}{2} \sum_k b'_k J_k^* b_k - g' q \right) \min \qquad (k = \text{element index}) \qquad (61)$$

while using the stress function formulation (51)

$$\left(\frac{1}{2} \sum_k f'_k F_k f_k - d' f \right) \min \qquad (62)$$

Assembling the elements is achieved by stating that along each interface the sum of the generalized loads equals the applied external loading

$$g = \sum_k L_k g_k = \sum_k L_k C_k b_k \qquad (63)$$

where L_k are localizing or ooolean matrices and C_k the static matrices defined in (40).

In the absence of interface loading, the stress functions are continuous across the interfaces, what is written

$$f_k = \bar{L}_k f \qquad (64)$$

where $\bar{L}_k$ are other localizing matrices. Along the loaded interfaces the relation (55) allows to express that the dislocations or jumps of the stress functions equal the loads

$$g = \sum_k g_k = \sum_k L_k C_k^* f_k \qquad (65)$$

The solution of the minimization problem (61) or (62) is obtained by substitution of the constraints (63) or (64). In the first formulation we obtain the classical results

$$\sum_k J^*_k b_k = \sum_k C'_k L'_k q$$

$$g = \sum_k L_k g_k = \sum_k L_k C_k b_k = \left(\sum_k L_k C_k J^{*-1}_k C'_k L'_k \right) q = \sum_k L_k K_k L_k q = K_s q \qquad (66)$$

corresponding to the displacement method applied to equilibrium finite elements. The unknowns are effectively interface displacement modes. K_s is the global stiffness matrix obtained by a suitable localizing of the element stiffnesses.

In the second formulation, we obtain by the same operation, neglecting first the eventually prescribed loads

$$\left(\sum_k \bar{L}'_k F_k \bar{L}_k \right) f = F_s f = d \qquad (67)$$

where F_s is the structural flexibility matrix for the same assemblage of equilibrium elements. The unknowns are here interface stress function modes. As these have been interpreted as self-stressing states, this solution is in fact in the spirit of the force method, but in which the search for the redundancies is avoided. The second member d represents contour deformations or generalized cuts conjugate to the self stressing states.

The figure 2 translates the results obtained for equilibrium models into the corresponding ones for displacement models. With the help of the analogies described above, the figure is self explanatory.
In the two formulations denoted force method it is important to remember that the search for the self stressing states is avoided by the particular choice of the stress functions as variables.

The choice for one method or the other should be based on the number of variables involved in the particular application considered. The same computer program can obviously be used to assemble the four systems of equations and for their solution.

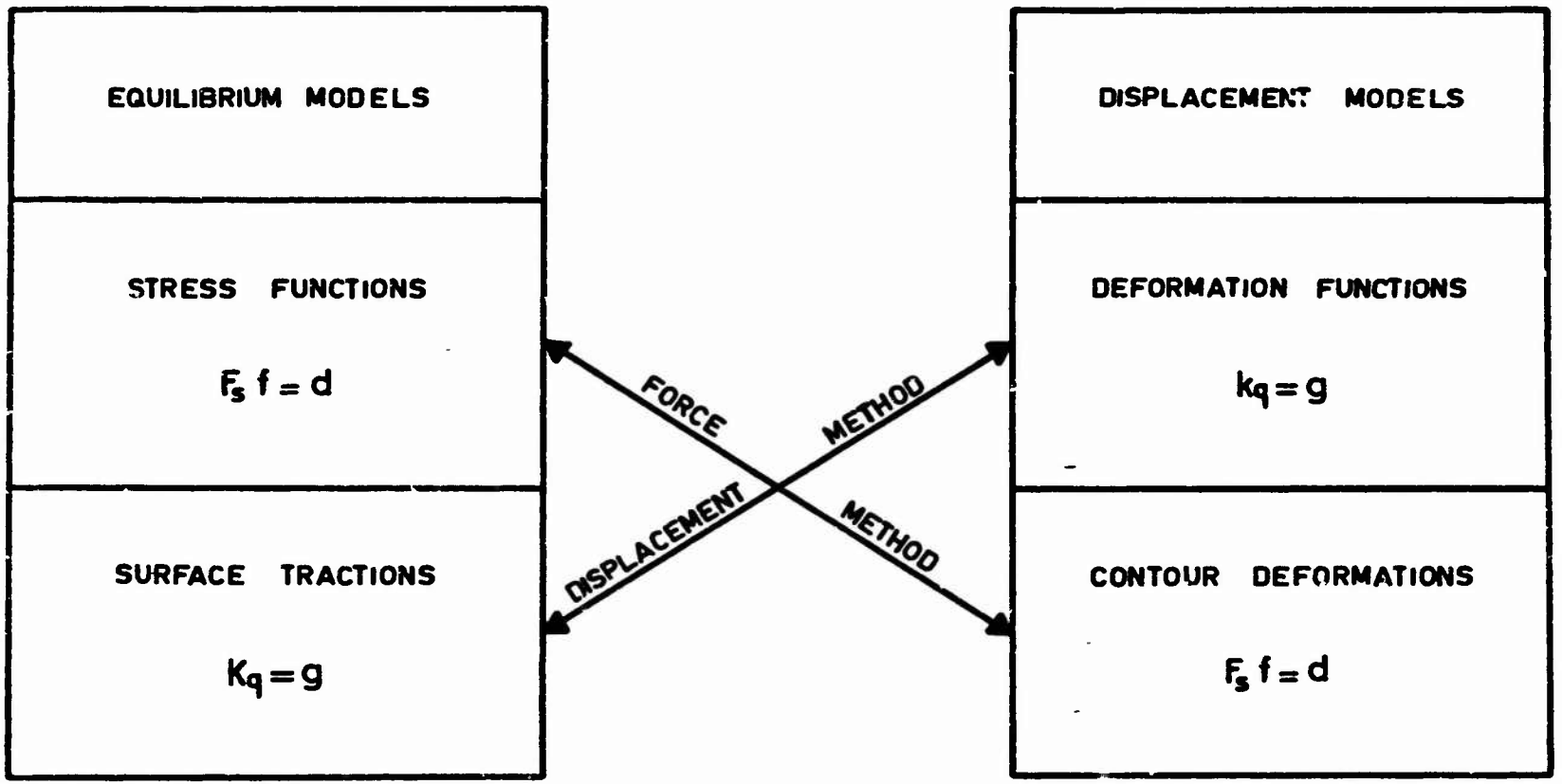

EQUILIBRIUM MODELS
STRESS FUNCTIONS
$F_s f = d$
SURFACE TRACTIONS
$K_q = g$
DISPLACEMENT MODELS
DEFORMATION FUNCTIONS
$k_q = g$
CONTOUR DEFORMATIONS
$F_s f = d$
FORCE METHOD
DISPLACEMENT METHOD

FIGURE 2

6. INTRODUCTION OF THE BOUNDARY CONDITIONS

The introduction of the boundary conditions has not been considered in the preceding formulation. These conditions can be prescribed forces or displacements; they can be zero or non-zero prescribed or the average over some region can be prescribed.

In the two forms of displacement methods, prescribed forces appear in the second member and prescribed displacements lead to a reduction of the number of unknowns. In case of prescribed average values of forces or displacement (like for instance a prescribed global shear force in a box beam) the corresponding linear constraints between the variables has to be written and added to the system of equations.

In the two new forms of force method, the boundary conditions have to be expressed in terms of stress functions or of contour deformations. When these conditions involve only the contour of the assemblage of finite elements the change of variables is relatively easy. However if loads are applied inside the structure, dislocations have to be introduced in the stress function field. This can be achieved in the following way.

In the expression (62) for the potential energy, we separate the terms pertaining to unloaded and loaded inter element boundaries and use for the latter their expression in terms of forces $\bar{g}$ and displacements $\bar{\bar{q}}$

$$d' f = d'_u f_u + \bar{g}' \bar{\bar{q}} \qquad \text{(subscript u stands for unloaded)} \tag{68}$$

Using then (57) and (63)

$$g = \sum_k L_k C^*_k f_k = \sum L_k C^*_k \bar{L}_k f = C^*_s f \tag{69}$$

where C^* is a global matrix translating the stress functional modes in terms of surface traction modes.

Substitution of (68) into (61) yields

$$\begin{vmatrix} F_s & C^*_s \\ C^*_s & 0 \end{vmatrix} \begin{vmatrix} f \\ -\bar{q} \end{vmatrix} = \begin{vmatrix} 0 \\ \bar{d} \end{vmatrix} \tag{70}$$

It appears therefore that the introduction of the boundary conditions corresponds to the addition of constraint equations in number proportional to the number of loaded interfaces. Hence it is obvious that if all the inter-

faces are loaded, the stress function formulation becomes unattractive.

The introduction of the boundary conditions in terms of stress functions is illustrated by two examples in figures 3 to 8 for a cantilever U beam composed of membrane elements and for the bending of a centrally loaded point supported square plate.

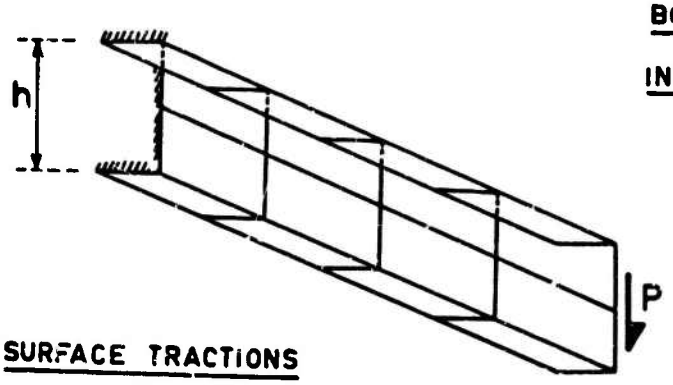
Taking first the cantilever U beam, figure 3 represents the boundary conditions as they are used in a classical equilibrium analysis formulated in terms of generalized force and solved by a displacement method. These boundary conditions for the normal and tangential surface tractions n, t are translated into boundary conditions expressed in terms of curvature of the stress function $F(x,y)$ as illustrated by figure 4. On the same figure the arbitrary fixation of the "rigid body" modes of the stress function is also indicated. Note that F_x stands for $\frac{\partial F}{\partial x}$ etc. From these boundary conditions it is easy to derive the set of constraints represented in figure 5. This operation is easily done by hand on the example but could also result from equation 69.

Turning next to the plate bending problem, the boundary conditions in terms of surface tractions are indicated in figure 6 and translated in terms of stress functions in figure 7 where the arbitrary fixation of the "rigid body" modes is also represented. One notes that the introduction of a concentrated load inside the plate requires to impose a dislocation of the stress function u,v, along a line arbitrarily selected from the load to an external boundary. The angle of opening due to the dislocation is proportional to the load. The deformed shape of the stress function is indicated by the dashed lines. In figure 8 the detailed expressions of the constraints corresponding to the concentrated load are represented. The upper part of the figure indicates the general conditions to satisfy by the stress functions on both sides of the dislocation. The lower part of the figure illustrates the application in the case where the dislocation is bordered by triangular elements with a linear stress function field. It appears that the final form of the constraints expressed point by point is extremely simple and can be introduced by adding a certain number of relations of the same type between nodal values. This observation allows the systematic treatment of the constraints by the computer program.

CANTILEVER U BEAM

BOUNDARY CONDITIONS

IN EQUILIBRIUM ANALYSIS



a) **SURFACE TRACTIONS**

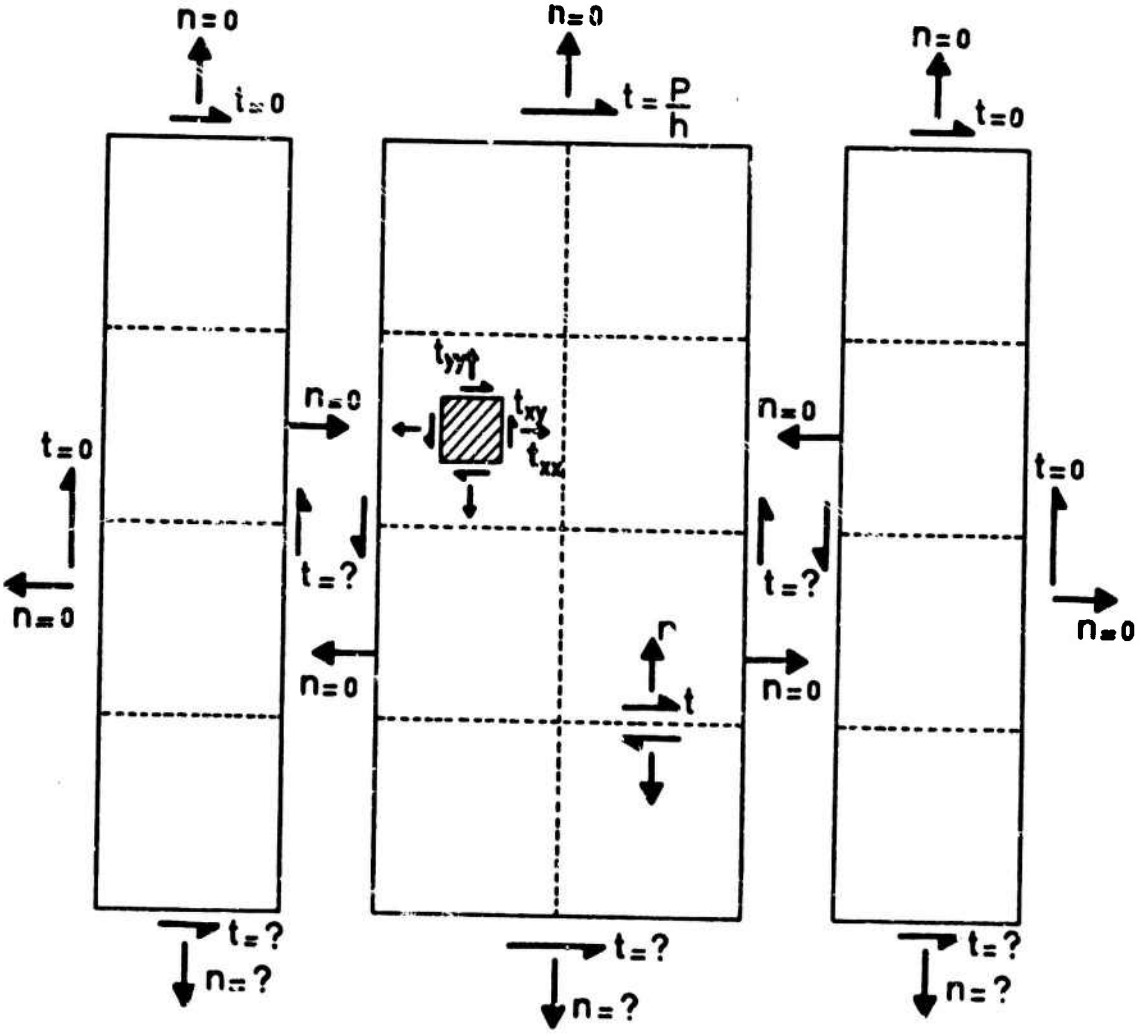
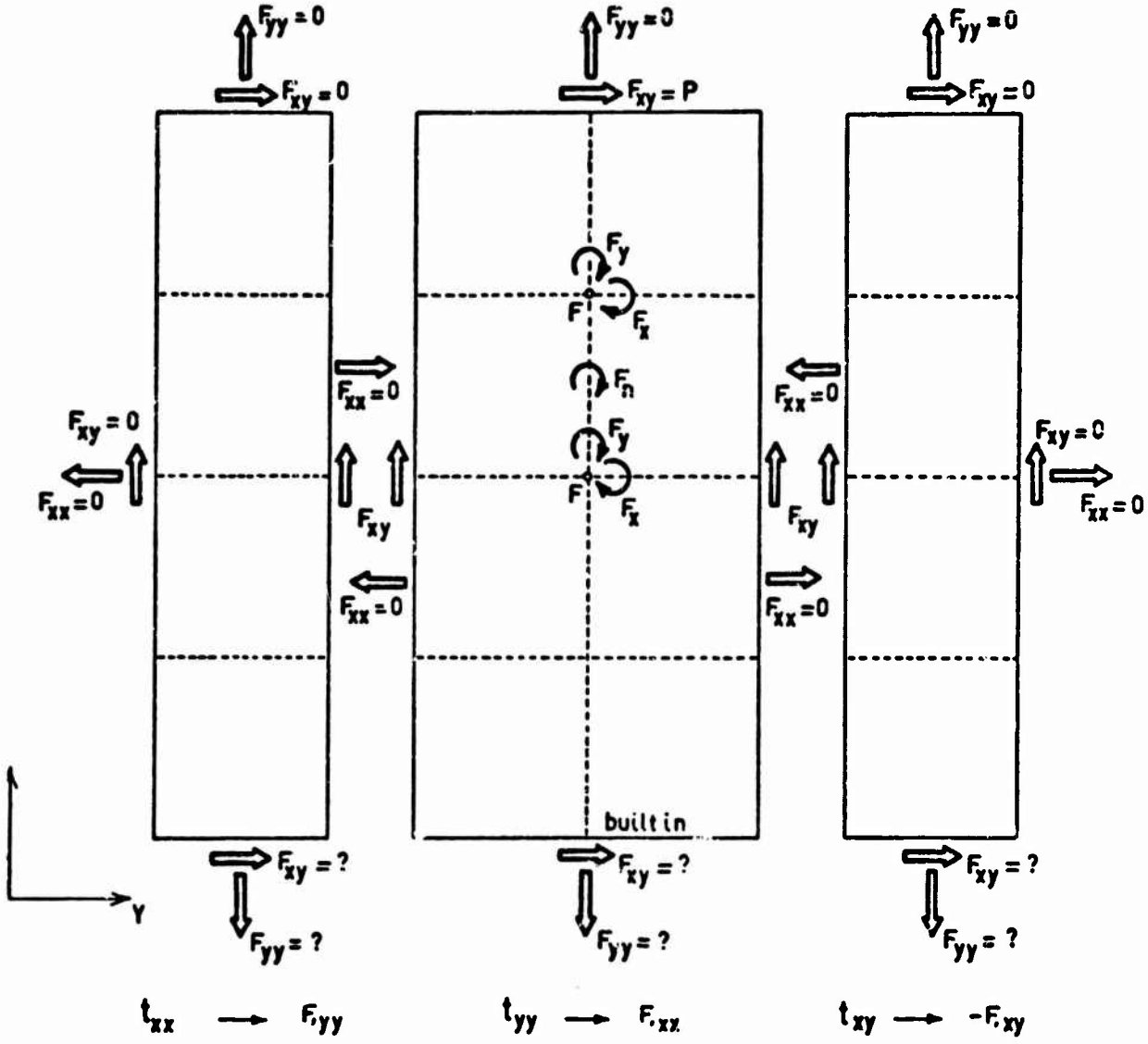


FIGURE 3

b) BOUNDARY CONDITIONS IN TERMS OF STRESS FUNCTION F(x,y)



FIXATION OF
THE RIGID BODY MODES

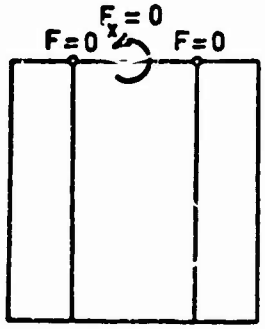


FIGURE 4

c) CONSTRAINTS IMPOSED TO THE STRESS FUNCTION F(x.y)

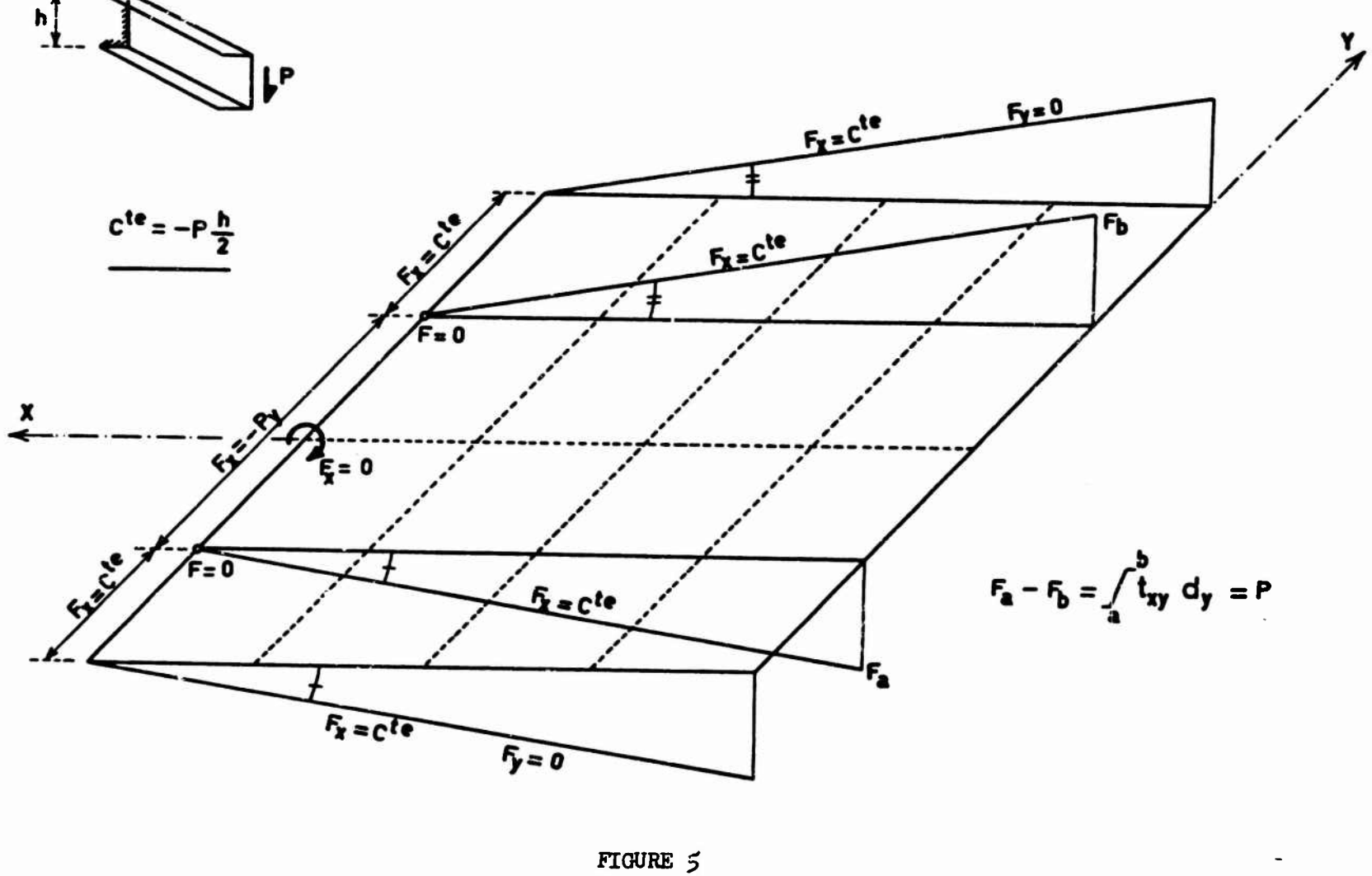
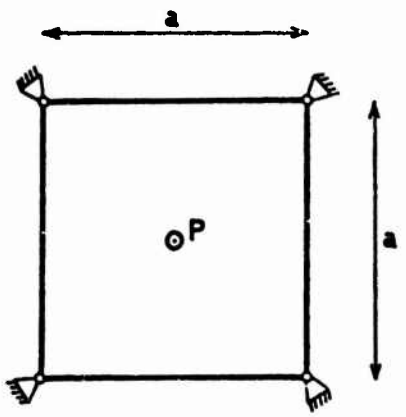


FIGURE 5

POINT SUPPORTED SQUARE PLATE

CENTRAL CONCENTRATED LOAD



$$\frac{\partial M_x}{\partial X} + \frac{\partial M_{xy}}{\partial Y} = V_x \qquad \frac{\partial V_x}{\partial X} + \frac{\partial V_y}{\partial Y} = -P = 0$$

$$\frac{\partial M_y}{\partial Y} + \frac{\partial M_{xy}}{\partial X} = V_y$$

$$K_n = V_n + \frac{\partial M_{sn}}{\partial s}$$

$$Z_i = M_{sn_{i+\varepsilon}} - M_{sn_{i-\varepsilon}}$$

a) BOUNDARY CONDITIONS IN TERMS OF SURFACE TRACTIONS

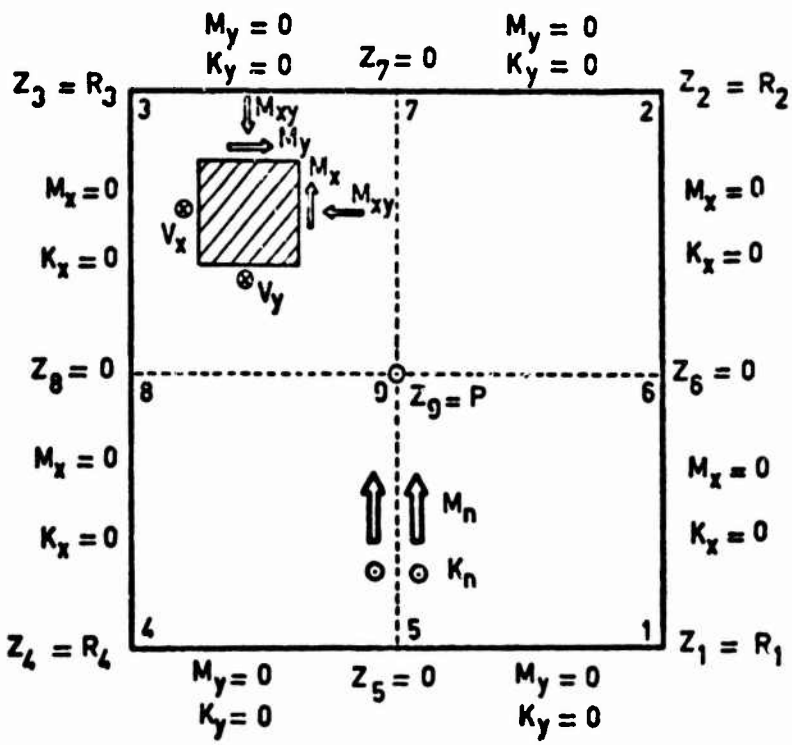


FIGURE 6

b) BOUNDARY CONDITIONS IN TERMS OF STRESS FUNCTIONS U AND V

$$M_x = \frac{\partial V}{\partial Y} = V_y \qquad K_x = -U_{yy}$$

$$M_y = \frac{\partial U}{\partial X} = U_x \qquad K_y = -V_{xx}$$

$$M_{xy} = -\frac{1}{2}(U_y + V_x) \qquad Z_{j_{xy}} = 2M_{xy_{j-\varepsilon}} = U_{y_{x=ij}} + V_{x_{y=jk}}$$

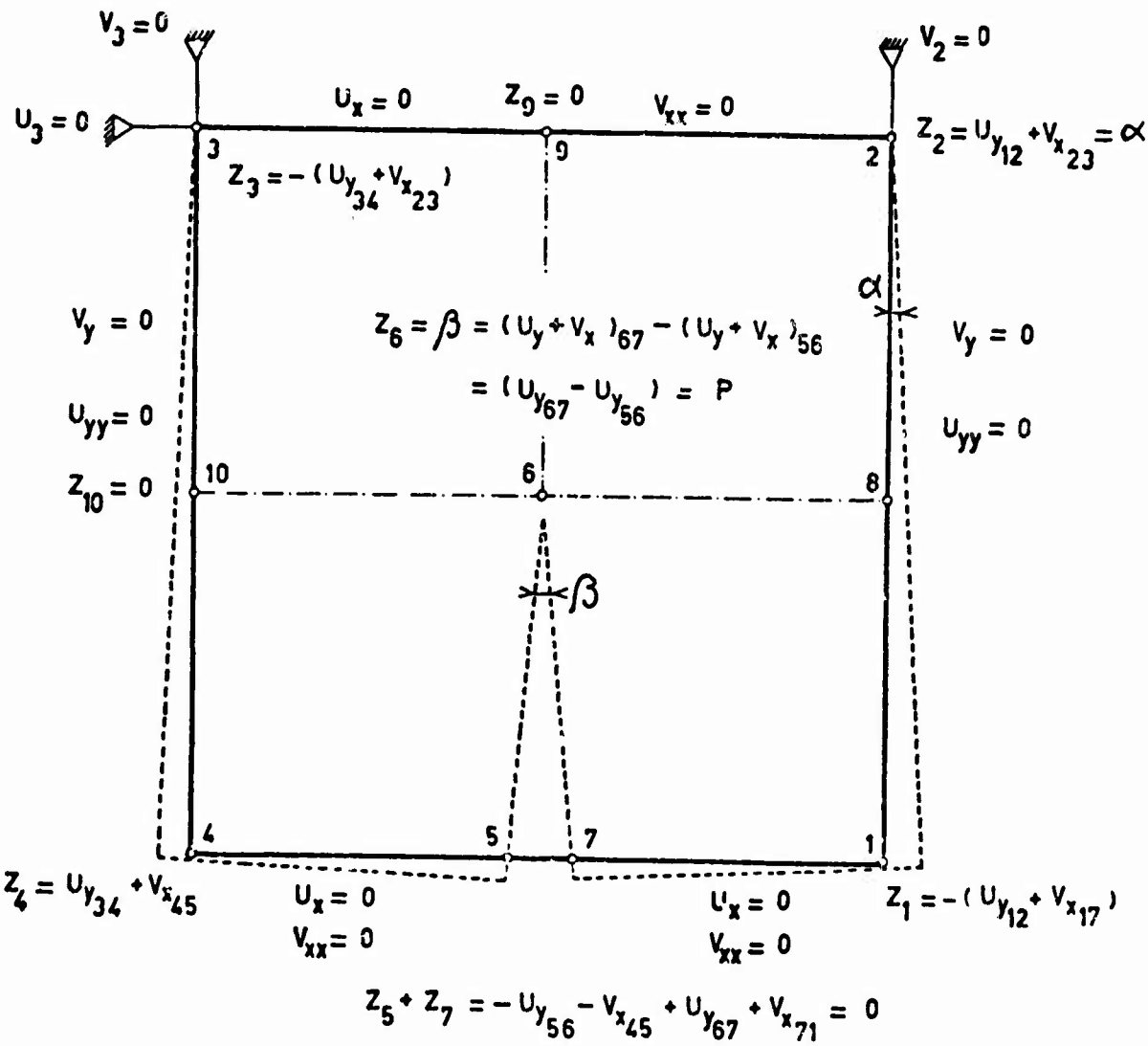
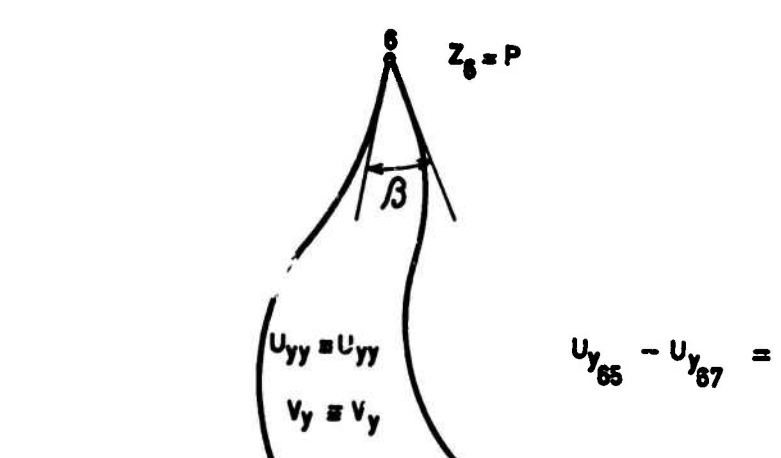
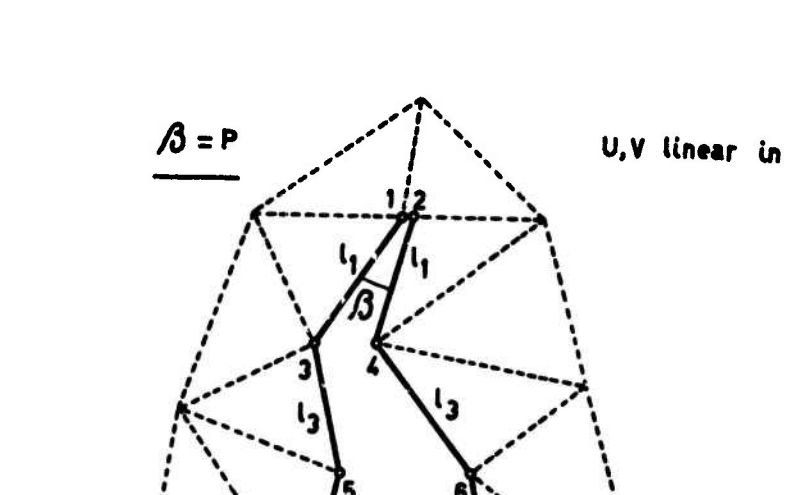


FIGURE 7





CONSTRAINTS

$$I \begin{cases} U_i = U_{i+1} + l_i\beta \\ V_i = V_{i+1} \end{cases} \qquad i = 1, 3, 5$$

$$II \begin{cases} \dfrac{\Delta V_{97}}{l_{97}} - \dfrac{\Delta U_{57}}{l_{57}} = \dfrac{\Delta V_{8,10}}{l_{8,10}} - \dfrac{\Delta U_{68}}{l_{68}} \quad \text{or} \quad Z_7 = Z_8 \\ \text{etc.... for (5,6) (3,4) (1,2)} \end{cases}$$

FIGURE 8

7. CRITERION FOR SELECTING THE METHOD OF SOLUTION

The number of equations to be solved is obviously different in the two formulations available for displacement and equilibrium models as the topology of the connections is dependent of the choice of the variables. It is also dependent of the boundary conditions which can either decrease the number of equations (as for clamped boundaries in displacement methods) or increase it (as in case of additional constraints). A simple criterion that can be used to select the most efficient method is based on the hyperstaticity index. It is defined by :

$$\Delta = \frac{n_p - n_f}{n_f} \tag{70}$$

where n_p is the total number of strain or stress parameters in the structure and n_f is the number of free, unprescribed degrees of freedom in the formulation examined. This index is the ratio between the numbers of unknowns involved in the two formulations of the same problem. Therefore if the index Δ is larger than 1., the formulation examined is more efficient than the other.

In the case of the displacement method, consider the global static matrix of the assembled structure (55), and suppose that all the lines corresponding to reactions against prescribed displacements have been suppressed. The number n_p of parameters b is larger than the number n_f of the free degrees of freedom (they are equal only if the structure is isostatic). This connection matrix C can be split in two parts corresponding to a particular solution and a set of self-stressing states

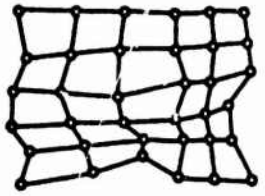
$$g = C\, b = C_1\, b_1 + C_2\, b_2$$

where C_1 is a square matrix and C_2 a rectangular matrix of dimension $n_f \times (n_p - n_f)$. It is clear that the displacement method is well adapted if $n_f < (n_p - n_f)$ that is if the number of redundancies is larger than the number of unknown displacements.

An application of this criterion is illustrated in figure 9 for a membrane stretching problem idealized by a 5 x 5 grid. Displacement and equilibrium models of increasing degree of sophistication are considered.

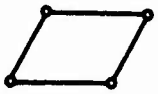
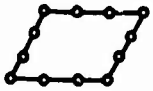
THE HYPERSTATICITY INDEX FOR A MEMBRANE STRETCHING PROBLEM

5×5 GRID



25 ELEMENTS
36 NODES
60 INTERFACES

(2 D.O.F PAR NODAL POINT)

	MODEL	d	n_b	n_p	FREE EDGES		CLAMPED EDGES	
					n_f	Δ	n_f	Δ
DISPLACEMENT		0	5	125	72	.810	32	2.91
		1	13	325	192	.720	112	1.90
		2	21	525	312	.700	192	1.79
EQUILIBRIUM		0	5	125	120	.07	80	.56
		1	13	325	240	.37	160	1.03
		2	21	525	360	.47	240	1.19

d = DEGREE OF THE STRESS OR STRAIN FIELD

n_b = NUMBER OF PARAMETERS OF THE STRAIN OR STRESS FIELD

n_p = TOTAL NUMBER OF STRAIN OR STRESS FIELD PARAMETERS

n_f = NUMBER OF UNPRESCRIBED UNKNOWNS

Δ = HYPERSTATICITY INDEX

FIGURE 9

The hyperstaticity index is established for the displacement method in two limit cases of boundary conditions : all the contour is either free or clamped. Had the index been established on the basis of the force method, the values of Δ would have been the inverse of the present one.

It appears that the advantage of the displacement method is the largest for the clamped plate idealized by the simplest displacement element (Δ=2.91) while the force method is by far the best choice in the case of a free contour and the simplest equilibrium element (Δ=.07). Note that this element is very similar to the well known shear panel. Contrary to an established opinion, a force method would be more economic even for displacement models in the case of free boundaries. It is recalled that in the present context the search for the redundancies is avoided in the force method by the special choice of the variables.

From this example it appears that the choice of the solution method is strongly influenced not only by the type of element used, the mesh size and connectivity of the structure, but also by the boundary conditions. In the cases where these are such that constraints are to be added to the system of equations, their number should be introduced in the evaluation of the hyperstaticity index.

8. NON-STRUCTURAL APPLICATIONS

The structural problems encountered in aerospace applications are characterized by the complexity of their geometry. Their physical subdivision is usually so fine that even the simplest elements yield already a good approximation of the displacements. The representation of the stress field is however relatively crude. In the boundary conditions one usually finds a large number of concentrated loads. These characteristics explain that for displacement models the displacement method has been exclusively used, while for bars and shear panels which are very similar to the simplest equilibrium elements, the classical force method has been preferred although it requires a numerical search for the redundancies. But usually the hyperstaticity index of the structures idealized by bars and shear panels is so small that some advantage remains even if the additional work of the Gauss-Jordan elimination is considered.

It has been shown in the preceding section that the new formulation of the force method proposed in this paper, as it avoids the search for the redundancies, can be more economic in a large class of problems. This is

especially true for analysis by equilibrium models which suffer in the displacement method of its relatively larger number of equations. The utilisation of these models has been very limited for this reason and this despite the fact that they yield often better stresses which are at the same time of easier interpretation. Comparison of the energy content at the element level as well as globally is also sometimes fruitful.

Nevertheless the practical use of the stress function formulation for equilibrium models still presents a difficulty which is peculiar to structural problems. That is the need for translating the boundary conditions from applied forces into local values and dislocations of the stress functions. This step is presently achieved by hand but the development of a simple and automatic algorithm is studied.

There is however a large class of problems, that might be called non-structural, to which the new force method can be readily applied. Among them plate bending and plane elasticity problems, fluid flows, heat conduction problems etc, can be mentionned. The boundary conditions met in such problems are easier to express in the variables required by the new force method while at the same time the domains studied are often less physically subdivided than in structural applications. The number of dislocations to introduce in the stress function field (or its equivalent) is therefore reduced.

The complete duality between finite element models and methods of solution is illustrated in the case of potential fluid flow and heat conduction problems in figure 10 which is self explanatory. More details can be found in Reference 13.
Note that passing from one set of variables to the other and from "displacement" to "equilibrium" models is much simpler in these applications, which involve scalar fields, than in elasticity, and only require formally the change of the constitutive equations. In particular the topology of the connection remains the same for displacement and equilibrium models in a given formulation.

FORMULATIONS FOR POTENTIAL FLUID FLOWS AND HEAT CONDUCTION PROBLEMS

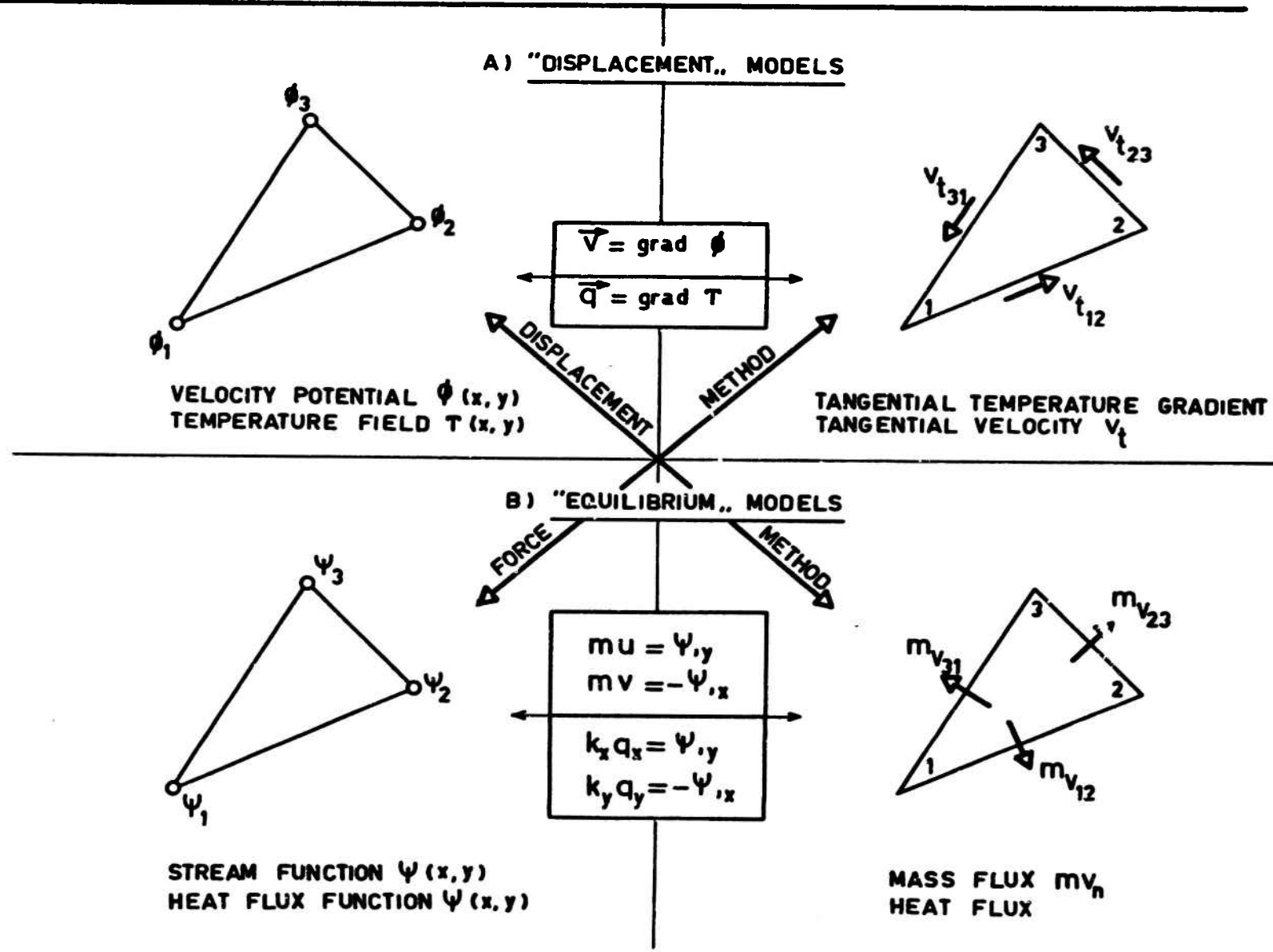


FIGURE 10

9. EFFECT OF INTERNAL DEGREES OF FREEDOM ON FINITE ELEMENT PROPERTIES.

In sections 2 and 3 the possibility of introducing internal displacement modes (bubble functions) has been mentionned. An attempt is made here to evaluate the effect of such modes on the finite element properties. Four membrane elements are used in this investigation :

- CTM denotes the classical Conforming Triangular Membrane element
- CPM denotes the Conforming Parallelogram Membrane element using a bilinear or biquadratic displacement field.
- CQM denotes the Conforming Quadrilateral Membrane super-element formed by assembling four triangles CTM and condensation of the internal displacements.
- HPM denotes the Hybrid Parallelogram Membrane element with assumed displacements along the boundaries and assumed stresses inside.

The displacement field in the three conforming elements is represented in the form

$$u = P_1 (x,y) + f_2 (x,y) \cdot P_2 (x,y)$$

where $P_1 (x,y)$ is the classical polynomial which contains a number of parameters just equal to the number of generalized displacements on the contour of the element (6 or 12 for the triangles, 8 or 16 for the quadrangles).

$f_2 (x,y) = o$ is the product of the equations of the interfaces, that is a function vanishing on the boundaries

$P_2 (x,y)$ is another polynomial representing bubble or internal displacement modes. The elements have been tested without bubble modes and then the degree of P_2 has been increased from zero to four. Note that zero corresponds already to one internal mode.

In the hybrid element the contour modes are linear or quadratic while the stress field inside the element derives from an Airy stress function represented by a complete polynomial of degree ranging from 2 to 6. The stress field is therefore of degree o to 6.

In all cases the internal degrees of freedom have been eliminated by condensation. The elements are either square or equilateral triangle of side length equal to 1., the Young's modulus is 10^4, the thickness is 1. and the Poisson's ratio is .3. All the eigenvalues of the stiffness matrices have been computed for each of these elements. The first three eigenvalues are zero and correspond to the rigid body modes. In addition in the hybrid element additional zero eigenvalues are found when the degree of the stress field is not sufficient. They correspond to spurious kinematic deformation modes (Ref. 3). The first non zero and the maximum eigenvalues are given in figure 11 together with the trace of the stiffness matrices.

It appears from these results that when the number of internal degrees of freedom grows, the conforming and hybrid elements converge, as expected by the theory, to the same values which correspond to the exact solution of a membrane with the constraints of linear or quadratic variation of the displacements prescribed along the edges. The hybrid elements yield eigenvalues which are very close of the exact solution even for the simplest stress field that can be used without kinematic modes. The convergence of the other elements is very fast so that it is seen that a few bubble modes allow to satisfy almost exactly the equilibrium inside the element.

In conforming elements, an internal field which is one degree higher than that of the interfaces brings the major part of the possible improvement. It can affect predominantly the minimum eigenvalue in which case it will influence the global response of the structure or the maximum eigenvalue in which case it reveals more important for a correct stress representation inside the element. Note that the simplest triangular element is unaffected by such modes. Comparison of the numerical work necessary to build a HPM or a CPM element with a few bubble modes shows little difference.

The conclusion is that the solution with hybrid models or with conforming models containing a few internal modes is practically coincident. The choice of one model or the other has therefore little influence on the results which are governed only by the degree of the displacements along the contour.

EFFECT OF INTERNAL DEGREES OF FREEDOM

			linear interface displacements			quadratic interface displacements		
Model	d	n_b	λ_{min}	λ_{max}	trace	λ_{min}	λ_{max}	trace
HPM	0		-	-	-	-	-	-
	1		75.973	285.71	745.35	-	-	-
	2		75.973	285.71	745.35	-	-	-
	3		78.638	285.71	750.68	31.058	395.87	2180.3
	4		78.638	285.71	750.68	31.401	396.66	2236.9
	5		78.855	285.71	751.12	31.401	396.86	2254.0
	6		78.856	285.71	751.12	31.408	398.74	2256.9
CQM	1	2	113.960	285.71	821.33	-	-	-
	2	10	80.953	285.71	755.31	38.978	412.62	2553.1
	3	26	78.997	285.71	751.40	31.554	401.04	2283.6
	4	50	78.925	285.71	751.26	31.417	399.21	2267.1
CPM	-	0	98.901	285.71	791.21	33.610	943.38	4114.3
	0	2	79.794	285.71	752.99	33.610	467.15	2604.5
	1	8	79.263	285.71	751.93	31.435	403.23	2313.7
	2	18	78.952	285.71	751.31	31.412	399.53	2273.0
	3	32	78.913	285.71	751.23	31.410	399.19	2265.4
	4	50	78.902	285.71	751.21	31.403	398.97	2263.3
CTM	1	0	95.940	343.62	593.40	-	-	-
	2	0	95.940	343.62	593.40	16.078	1022.9	2967.0
	3	2	95.940	343.62	593.40	15.777	857.65	2423.8
	4	6	95.9[illegible]	343.62	593.40	15.629	843.26	2362.2
	5	12	95.[illegible]	343.62	593.40	15.601	841.42	2354.1

d = degree of the polynomial $P_2(x,y)$ or, in HPM, degree of the stress field

n_b = number of condensed internal degrees of freedom

FIGURE 11

REFERENCES

1. SANDER G. and FRAEIJS de VEUBEKE B.

   "Upper and Lower Bounds to Structural Deformations by Dual Analysis in Finite Elements"

   USAF Tech. Report. - AFFDL-TR-66-199 - 1966

2. SANDER G.

   "Application of the Dual Analysis Principle"

   Proc. IUTAM Coll. on "High Speed Computing of Elastic Structures" Ed. FRAEIJS de VEUBEKE, Liège, 1970 (in Press)

3. FRAEIJS de VEUBEKE B., SANDER G. and BECKERS P.

   "Dual Analysis by Finite Elements - Linear and Nonlinear Applications"

   USAF Tech. Report - AFFDL - 72-93, 1973.

4. FRAEIJS de VEUBEKE B.

   "A Conforming Finite Element for Plate Bending"

   Int. J. Solids and Struct., Vol. 4, 1968, Pergamon

5. SANDER G.

   "Dual Analysis of a Multiwek Swept Back Wing Model"

   Aircraft Engineering, Febr. 1968

6. CLOUGH R. and TOCHER J.

   "Finite Element Stiffness Matrices for Analysis of Plates in Bending"

   Proc. USAF Conf. DAYTON, Report AFFDL-TR-66-80, 1966

7. FRAEIJS de VEUBEKE B. and ZIENKIEWICZ O.C.

   "Strain Energy Bounds in Finite Element Analysis by Slab Analogies"

   J. of Strain Analysis, Vol 2, n° 4, 1967

8. FRAEIJS de VEUBEKE B.

   "Displacements and Equilibrium Models in the Finite Element Method"

   Chap. 9 in "Stress Analysis" Ed. Zienkiewicz, John Wiley, 1965

9. SANDER G.

   "Application de la Méthode des Eléments Finis à la Flexion des Plaques"

   Coll. des Publications Fac. Sc. Appl. - Univ. Liège, n° 15, 1969

10. KOSKO E.

    "The Equivalence of Force and Displacement Methods in the Analysis of Elastic Structures"

    Proc. USAF Conf. DAYTON, Ohio, AFFDL-TR-66-80, 1966

11. FRAEIJS de VEUBEKE B.

    "Basis of a Well Conditionned Force Program for Equilibrium Models the Southwell Analogies"

    USAF Report AFFDL-TR-67-60, 1967

12. ZIENKIEWICZ, IRONS, ERGATOUDIS, [illegible] SCOT

"Isoparametric and Associated Element Families for Two and Three Dimensional Analysis"

Chap.13 in "Finite Element Method" Ed. Holand and Bell, TAPIR, 1969

13. BECKERS P.

"Les fonctions de tensions dans la méthode des éléments finis"

Dc. Thesis, Univ. Liège 1971

To be published in Coll. Publ. Fac. Sc. Appl., Liège

14. ZIENKIEWICZ, O.C.

"The finite element method in continuum mechanics"

2nd edition Mac Graw Hill, 1971.

BASIS FOR ELEMENT INTERCHANGEABILITY IN FINITE ELEMENT PROGRAMS

John Robinson and Gernot W. Haggenmacher
Lockheed-California Company, Burbank, California

The concept of characteristic matrices of any finite element for elastostatic analysis is presented. To demonstrate the concept the characteristic matrices of a simple beam element are derived in the Appendix. This approach permits complete interchangeability of different types of elements between independently developed analysis programs irrespective of the analysis method or basic element assumption. The analogy between element derivation procedures and the equality of principal matrices are developed. A wide spread adoption of this concept would have far reaching implications in, facilitating communications, exchanging of new developments, and speed-up practical applications in the field of finite elements.

INTRODUCTION

The element research and development over the past decade has been largely restricted to element stiffness matrices based on deformation assumptions. However, over the past few years element flexibility matrices based on stress assumptions have appeared with increasing frequency in the published literature. It is therefore appropriate to show the analogy between element derivation procedures and matrices and their interchangeability in the force, displacement and combined methods of analysis.

The data required to incorporate any element into an elastostatic analysis system, force, displacement or combined method, can be conveniently organized into four characteristic matrices. These matrices define the elastic behavior, the spatial assembly into the structure, and the required output information for each element. This concept was used for the force method. [1,2] Further characteristic matrices appear when extending the concept to vibration and non-linear analysis. The method of element development is independent of the method used to form the equations for the complete structure. Certain analogies which exist in the finite element methods, including methodology and elements, have been reported. [3,4,5,6,7,8] In the present paper an analogy is developed which covers the element derivation process for both stress and strain assumptions within the element. The formal presentation of this analogy demonstrates that it is possible to achieve complete interchangeability of the characteristic matrices of any type of element between the major methods of finite element analysis. This work complements the review of methodology relationships in Reference 2.

1. RELATIONS BETWEEN FINITE ELEMENT METHODS

Three basic finite element methods of analysis are usually distinguished, namely, force, displacement and combined (Lagrangian Multiplier) method. In the force method the nodal equilibrium equations are written in terms of the generalized element forces. These forces are solved first and then the nodal displacements are determined. In the displacement method the nodal equilibrium equations are written in terms of the generalized nodal displacements. These displacements are solved first and then

the element forces are determined. In the combined method the system of equations to be solved contains both the element forces and nodal displacements simultaneously as one set of unknown variables. The full solution is obtained in one pass. A full derivation of the relations between the three methods was given in Reference 2. The equations relevant to the present work will be summarized subsequently.

In the force method the nodal equilibrium equations, in all unconstrained degrees of freedom, are expressed in terms of the assembled generalized element for variables, $\{F\}$, that is,

$$[E]\{F\} = \{L\} \tag{1}$$

where

$\{L\}$ = vector of external applied generalized nodal forces expressed in the global system.

$[E]$ = matrix which transforms the element force variables into equivalent nodal forces in the global system and allocates them to the appropriate node.

The total element deformations, $\{d\}$, corresponding to the element force variables and initial deformations, are given by

$$\{d\} = \{d_e\} + \{d_o\} = [D]\{F\} + \{d_o\} \tag{2}$$

where

$[D]$ = matrix of assembled element natural flexibility matrices.

$\{d_o\}$ = initial, stress-free element deformations (thermal, etc.)

$\{d_e\}$ = element stress-related deformations.

In the displacement method of analysis the nodal equilibrium equations, in all unconstrained degrees of freedom, are expressed in terms of the nodal displacements, $\{r\}$, expressed in the global system, that is,

$$[K]\{r\} = \{L\} + [a]^T\{F_o\} \tag{3}$$

where

$[K]$ = structural stiffness matrix which transforms the nodal displacements into equivalent nodal forces in the global system and allocates them to the appropriate node.

$\{F_o\}$ = initial element force variables due to initial strains

The element forces are given by

$$\{F\} = [k]\{d_e\} \tag{4}$$

where,

$[k]$ = matrix of assembled element non-singular stiffness matrices.

The continuity conditions, relating the independent element deformations to nodal displacements are given by

$$[a]\{r\} = \{d\} = \{d_e\} + \{d_o\} \tag{5}$$

where,

$[a]$ = transformation matrix relating the nodal displacements and the total independent element deformations $\{d\}$.

In References 2 and 9 it was shown that

$$[a] = [E]^T \tag{6}$$

and

$$[k] = [D]^{-1} \tag{7}$$

The structural stiffness matrix is given by,

$$[K] = [a]^T[k][a] = [E][D]^{-1}[E]^T \tag{8}$$

Substituting Equation (6) into (5) gives,

$$\{d_e\} = [E]^T\{r\} - \{d_o\} \tag{9}$$

Using Equation (2) and (9)

$$[D]\{F\} + \{d_o\} - [E]^T\{r\} = \{0\} \tag{10}$$

Writing Equations (1) and (10) into a single system yields the Combined Method.[2/]

$$\left[\begin{array}{c|c} D & -E^T \\ \hline -E & 0 \end{array}\right]\left[\begin{array}{c} F \\ \hline r \end{array}\right] = -\left[\begin{array}{c} d_o \\ \hline L \end{array}\right] \tag{11}$$

or, in a slightly different form, using Equations (6) and (7),

$$\left[\begin{array}{c|c} I & -ka \\ \hline -a^T & 0 \end{array}\right]\left[\begin{array}{c} F \\ \hline r \end{array}\right] = -\left[\begin{array}{c} kd_o \\ \hline L \end{array}\right] \tag{12}$$

In the combined method the forces and nodal displacements are evaluated simultaneously in one pass.

2. CONCEPT OF CHARACTERISTIC ELEMENT MATRICES

Following this preliminary review of finite element methods, the paper will now concentrate on the common aspects of element derivation and formulation, using both strain and stress fields as the basic assumptions.

It will be shown that a set of characteristic matrices is required to incorporate any type of finite (structural) element into an analysis system based on either the force, displacement or combined method. This concept was discussed for the force method in References 1 and 2. The work of many authors has contributed to the formulation of this unified approach.

Characteristic matrices for elastostatic analysis are developed in this paper, namely;

1. The Element Natural Elastic Matrix (Stiffness and Flexibility).
2. The Initial Stress-Free Deformation Vector (Thermal).
3. The Element Assembly Matrix.
4. The Element Output Matrix.

Further characteristic matrices, to be considered in later work, are the Consistent Element Mass [16] and Inverse Mass [17,18] Matrices for vibration analysis, and the Geometric Stiffness Matrices [16] for various types of large order deflection analysis.

For any element, regardless of the analysis method, a force-deformation relationship is needed, which leads to the first two characteristic matrices, that is, the element natural elastic matrix and initial stress-free deformation vector. This relationship can be written in either of the following two forms,

$$\{F_m\} = [k_m]\{d_m\} - \{F_{mo}\} \tag{13}$$

or

$$\{d_m\} = [D_m]\{F_m\} + \{d_{mo}\} \tag{14}$$

where

$[k_m]$ = element natural stiffness matrix (non-singular)

$[D_m]$ = element natural flexibility matrix (non-singular)

$\{F_m\}$ = independent generalized elastic forces for element m.

$\{d_m\}$ = independent generalized deformations for element m

$\{F_{mo}\}, \{d_{mo}\}$ = initial values of force and deformations.

The matrices $[k]$ and $[D]$ of Equations (7) and (2) are assembled from the element natural stiffness and flexibility matrices for all elements, respectively.

It is important for this parallel study of strain and stress elements and their interchangeability, that the generalized element deformation and force variables, $\{d_m\}$ and $\{F_m\}$ respectively, do not include the zero strain and zero stress states, and are thus independent quantities. The relationships of Equations (13) and (14) are derived in a local reference system for the element. The choice of this system is important in two ways; not only can this simplify the derivation but it usually helps the understanding of the physics of the element. The variables and reference system should represent the behavior of the element in a realistic 'natural' way. This philosophy has long been expounded by Professor Argyris.

The use of independent variables is the essential requirement for the present issue of interchangeability. Equation (13) is the form to be used in the displacement method, and Equation (14) that in the force method. Either relation can however, be derived based on strain or on stress assumptions.

It should be realized that if the element deformations $\{d_m\}$ are chosen as independent variables in Equations (13) and (14), then $\{F_m\}$ are forces corresponding to these deformations (in character and "line of action"), and vice versa.

The force-deformation relations for any type of element will be derived using virtual principles and certain basic assumptions, namely:

1. The assumption of a deformation or strain field within the element, which must satisfy continuity, but the corresponding stresses do no necessarily satisfy equilibrium. This is the basis of strain, displacement, compatible and Hybrid II type elements.
2. The assumption of a force or stress field within the element, which must satisfy equilibrium, but the corresponding strains will not necessarily satisfy continuity. This is the basis of stress, force, equilibrium and Hybrid I type elements.

Regardless of which assumption is made, the resulting force-deformation relation can be used for any of the finite element methods.

The <u>third characteristic matrix</u> is the element assembly matrix

To assemble the structural stiffness matrix in the displacement method the independent strain variables $\{d_m\}$ for each element have to be transformed into equivalent nodal displacements $\{\Delta_m\}$ in the global system, including rigid body freedoms, using the transformation (assembly) matrix $[a_m]$, that is,

$$\{d_m\} = [a_m]\{\Delta_m\} \tag{15}$$

To assemble the equilibrium equations in the force method the independent force variables $\{F_m\}$ for each element have to be transformed into equivalent nodal forces $\{Q_m\}$ in the global system, including the dependent forces (static supports), using the transformation (assembly) matrix $[E_m]$, that is,

$$[E_m]\{F_m\} = \{Q_m\} \tag{16}$$

The fourth characteristic matrix is the element output matrix.

In all methods of analysis the output data include generalized stress and/or strain quantities within each discrete element, it is therefore appropriate and convenient to define an element output matrix. $[S_m]$. This will relate the independent variables to the required output information which is suitable for structural engineering design.

3. THE STRAIN ELEMENT

3.1 FORCE-DEFORMATION RELATION

A generalized total strain field, $\{\varepsilon_\varepsilon\}$, is expressed in terms of a set of independent deformation variables, $\{d_{m\varepsilon}\}$, by the equation

$$\{\varepsilon_\varepsilon\} = [T_{\varepsilon d}]\{d_{m\varepsilon}\} \tag{17}$$

where the j^{th} column of $[T_{\varepsilon d}]$ gives the strain distribution, which satisfies compatibility, for a unit value of $d_{m\varepsilon j}$. The set of deformation variables does not contain the zero strain state (rigid body displacements) as is the case in many displacement element derivations.

To obtain the strain field in terms of the element deformation variables it is usual to assume displacement functions expressed by the following equation:

$$\{u\} = [T_{uB}]\{B\} \tag{18}$$

where,

$\{u\}$ = generalized displacements in the local axes system.

$\{B\}$ = function coefficients, the linear terms (such as, b_1, b_2x, b_3y) correspond to the zero strain state.

Appropriate differentiation of these functions, see Table 1, results in the total strain field of Equation (17). The deformation variables, $\{d_{m\varepsilon}\}$, and the coefficients $\{B\}$ are directly related by the equation

$$\{d_{m\varepsilon}\} = [T_{dB}]\{B\} \tag{19}$$

Applying the principle of virtual deformations to the j^{th} deformation variable establishes the relation

$$\tilde{d}_{m\varepsilon j}\, F_{m\varepsilon j} = \int_V \{\tilde{\varepsilon}_\varepsilon\}_j^T \{\sigma\}\, dV \tag{20}$$

where,

$d_{m\varepsilon j}$ = the j^{th} independent element deformation variable.

$\{\varepsilon_e\}_j = f(d_{m\varepsilon j})$, the corresponding j^{th} total strain field.

$F_{m\varepsilon j}$ = element force corresponding to $d_{m\varepsilon j}$

$\{\sigma\}$ = function of $\{\varepsilon_e\} = \{\varepsilon_\varepsilon\} - \{\varepsilon_o\}$

V = Volume of element

$\sim$ indicates virtual quantities

Subscript ε indicates quantities derived from strain assumptions.

If there are $\propto$ independent deformation variables then there are $\propto$ linear equations in the form of Equation (20), therefore the following system of equations results,

$$\left[\tilde{d}_{m\varepsilon}\right]\{F_{m\varepsilon}\} = \int_V [\tilde{\varepsilon}_\varepsilon]^T\{\sigma\}\, dV \tag{21}$$

where $\left[\tilde{d}_{m\varepsilon}\right]$ is a diagonal matrix of virtual deformation variables.

Using Equation (17), the complete set of virtual strain systems is given by,

$$[\tilde{\varepsilon}_c] = [T_{\varepsilon d}]\left[\tilde{d}_{m\varepsilon}\right] \tag{22}$$

Substituting Equation (22) into (21) gives,

$$\left[\tilde{d}_{m\varepsilon}\right]\{F_{m\varepsilon}\} = \left[\tilde{d}_{m\varepsilon}\right]\int_V [T_{\varepsilon d}]^T\{\sigma\}\, dV \tag{23}$$

or,

$$\{F_{m\varepsilon}\} = \int_V [T_{\varepsilon d}]^T\{\sigma\}\, dV \tag{24}$$

Equation (24) defines the true element force $F_{m\varepsilon j}$ corresponding to the chosen deformation variable $d_{m\varepsilon j}$, such that the product $d_{m\varepsilon j}.F_{m\varepsilon j}$ is the work done by the j^{th} deformation.

The stresses $\{\sigma\}$ are a function of the stress-related strains $\{\varepsilon_e\}$ which are the difference between total and initial strains, that is,

$$\{\sigma\} = \{\sigma(\varepsilon_e)\} = \{\sigma(\varepsilon_t - \varepsilon_o)\} \tag{25}$$

Refer to Figure 1.

For linear stress-strain behavior, the stresses are,

$$\{\sigma\} = [T_{\sigma\varepsilon}]\{\varepsilon_e\} = [T_{\sigma\varepsilon}]\left(\{\varepsilon_t\} - \{\varepsilon_o\}\right) \tag{26}$$

where, $[T_{\sigma\varepsilon}]$ expresses the classical stress-strain relations.

Therefore, substituting Equations (17) and (26) into (24), yields

$$\{F_{m\varepsilon}\} = \left(\int_V [T_{\varepsilon d}]^T [T_{\sigma\varepsilon}][T_{\varepsilon d}]\, dV\right)\{d_{m\varepsilon}\} - \int_V [T_{\varepsilon d}]^T [T_{\sigma\varepsilon}]\{\varepsilon_o\}\, dV \tag{27}$$

or

$$\{F_{m\varepsilon}\} = [k_{m\varepsilon}]\{d_{m\varepsilon}\} - \{F_{m\varepsilon o}\} \tag{28}$$

where

$$\{F_{m\varepsilon o}\} = \int_V [T_{\varepsilon d}]^T [T_{\sigma\varepsilon}]\{\varepsilon_o\}\, dV \tag{29}$$

are the initial (fictitious) element forces which would produce the same element deformations as the stress-free initial strains $\{\varepsilon_o\}$ (thermal) and,

$$[k_{m\varepsilon}] = \int_V [T_{\varepsilon d}]^T [T_{\sigma\varepsilon}][T_{\varepsilon d}]\, dV \tag{30}$$

is the natural stiffness matrix based on strain assumptions.

This natural stiffness matrix, as well as the initial forces, are for a strain (displacement) element as well as a hybrid element based on an internal strain field.

To obtain the external initial element deformations $\{d_{m\varepsilon o}\}$ due to the initial strains $\{\varepsilon_o\}$, it is only necessary to visualize that the same deformations could be produced elastically, thus the corresponding forces are,

$$\{F_{m\varepsilon o}\} = [k_{m\varepsilon}]\{d_{m\varepsilon o}\} \tag{31}$$

Hence,

$$\{d_{m\varepsilon o}\} = [k_{m\varepsilon}]^{-1}\{F_{m\varepsilon o}\} \tag{32}$$

where $\{F_{m\varepsilon o}\}$ is given by Equation (29) and $[k_{m\varepsilon}]$ by Equation (30). In Equation (29) the function for $\{\varepsilon_o\}$ is not specified, and may actually be given numerically. $\{d_{m\varepsilon o}\}$ is the initial stress free deformation vector (second characteristic matrix).

Using Equation (32), Equation (28) can now be written as,

$$\{F_{m\varepsilon}\} = [k_{m\varepsilon}]\left(\{d_{m\varepsilon}\} - \{d_{m\varepsilon o}\}\right) = [k_{m\varepsilon}]\{d_{m\varepsilon\varepsilon}\} \tag{33}$$

where $\{d_{m\varepsilon\varepsilon}\}$ are the stress related deformations.

3.2 STRAIN ELEMENT ASSEMBLY MATRIX

The element nodal displacements in the local system, $\{\delta_m\}$, including rigid body freedoms, can be expressed in terms of the coefficients $\{B\}$ by substituting the proper nodal coordinates into Equation (18), that is,

$$\{\delta_m\} = [T_{\delta B}]\{B\} \tag{34}$$

Solving for $\{B\}$ gives,

$$\{B\} = [T_{\delta B}]^{-1}\{\delta_m\} \tag{35}$$

Substituting from Equation (35) into (19) results in,

$$\{d_{m\varepsilon}\} = [T_{d\delta\varepsilon}]\{\delta_m\} \tag{36}$$

where

$$[T_{d\delta\varepsilon}] = [T_{dB}][T_{\delta B}]^{-1} \tag{37}$$

= element displacement assembly matrix in the local system based on strain assumptions.

The relation between the corresponding force variables and nodal forces $\{q_m\}$ in the local element system, including dependent nodal forces (static supports), can be found by applying the principle of virtual displacements. Therefore,

$$[\tilde{\delta}_m]\{q_m\} = [\tilde{d}_{m\varepsilon}]^T\{F_{m\varepsilon}\} \tag{38}$$

From Equation (36)

$$[\tilde{d}_{m\varepsilon}] = [T_{d\delta\varepsilon}][\tilde{\delta}_m] \tag{39}$$

Hence, substituting from Equation (39) into (38)

$$\{q_m\} = [T_{qF\varepsilon}]\{F_{m\varepsilon}\} \tag{40}$$

where,

$$[T_{qF\varepsilon}] = [T_{d\delta\varepsilon}]^T \tag{41}$$

= element force assembly matrix in the local system based on strain assumptions.

When assembling the elements, the nodal forces and displacements for each element are transformed into the global system, denoted by $\{Q_m\}$ and $\{\Delta_m\}$ respectively, using a simple coordinate transformation matrix, $[T_c]$, that is,

$$\{Q_m\} = [T_c]\{q_m\} \tag{42}$$

$$\{q_m\} = [T_c]^T\{Q_m\} \tag{43}$$

and similarly,

$$\{\Delta_m\} = [T_c]\{\delta_m\} \tag{44}$$

$$\{\delta_m\} = [T_c]^T\{\Delta_m\} \tag{45}$$

Therefore, using the coordinate transformations of Equations (45) and (42) in Equations (36) and (40), respectively,

$$\{d_{m\varepsilon}\} = [a_{m\varepsilon}]\{\Delta_m\} \tag{46}$$

where,

$$[a_{m\varepsilon}] = [T_{d\delta\varepsilon}][T_c]^T \tag{47}$$

= element displacement assembly matrix in the global system.

and

$$\{Q_m\} = [E_{m\varepsilon}]\{F_{m\varepsilon}\} \tag{48}$$

where

$$[E_{m\varepsilon}] = [T_c][T_{qF\varepsilon}] = [a_{m\varepsilon}]^T \tag{49}$$

= element force assembly matrix in the global system.

3.3 STRAIN ELEMENT OUTPUT MATRIX

In a strain element the corresponding stress field is defined by Equation (26), that is,

$$\{\sigma_e\} = [T_{\sigma\varepsilon}]\{\varepsilon_e\} = [T_{\sigma\varepsilon}]\left(\{\varepsilon_e\} - \{\varepsilon_o\}\right) \tag{50}$$

Substituting from Equation (17) into (50) gives,

$$\{\sigma_e\} = [T_{\sigma\varepsilon}]\left([T_{\varepsilon d}]\{d_{m\varepsilon}\} - \{\varepsilon_o\}\right) \tag{51}$$

It appears however that the real stress distribution in the element cannot be dependent on the quite arbitrary function of $\{\varepsilon_o\}$, but rather more appropriately on the element forces given by Equation (33).

In the case of a simple axial element the stress can only be constant, irrespective of the complexity of the initial strain distribution $\{\varepsilon_o\}$. This is contrary to Equation (51). This equation, though mathematically derived from previous relationships, gives a result which is inconsistent in a physical sense. The stresses can only be those which are consistent with the element forces. The elastic deformations are given by,

$$\{d_{m\varepsilon e}\} = \{d_{m\varepsilon}\} - \{d_{m\varepsilon o}\} \tag{52}$$

where the deformation quantities are defined by Equations (32) and (46). Therefore, using the same transformation as in Equation (17) to give the elastic strains, that is,

$$\{\varepsilon_e\} = [T_{\varepsilon d}]\{d_{m\varepsilon e}\} \tag{53}$$

the stress distribution within the element is given by

$$\{\sigma_e\} = [T_{\sigma\varepsilon}]\{\varepsilon_e\} = [T_{\sigma\varepsilon}][T_{\varepsilon d}]\{d_{m\varepsilon e}\} \tag{54}$$

In the displacement method programs the stress output is evaluated using the computed element elastic deformations. Therefore, based on the specific output requirements and using Equation (54), the stress output is given by,

$$\{\sigma_{m\varepsilon}\} = [S_{m\varepsilon}]\{d_{m\varepsilon e}\} \tag{55}$$

where,

$\{\sigma_{m\varepsilon}\}$ = vector of required output data for a strain element

$[S_{m\varepsilon}]$ = element output matrix for a strain element.

From Equation (33),

$$\{F_{m\varepsilon}\} = [k_{m\varepsilon}]\left(\{d_{m\varepsilon}\} - \{d_{m\varepsilon o}\}\right)$$

or

$$\{F_{m\varepsilon}\} = [k_{m\varepsilon}]\{d_{m\varepsilon e}\} \tag{56}$$

Hence, from Equation (56)

$$\{d_{m\varepsilon e}\} = [k_{m\varepsilon}]^{-1}\{F_{m\varepsilon}\} \tag{57}$$

Substituting from Equation (57) into (55) gives the stress ouput in terms of force variables, that is,

$$\{\sigma_{m\varepsilon}\} = [S_{m\varepsilon}][k_{m\varepsilon}]^{-1}\{F_{m\varepsilon}\} \tag{58}$$

The element output matrix will now be denoted by $[S_{m\varepsilon}]$ with the appropriate subscripts of F or D to denote in which method it is to be used (force or displacement.

Therefore, for the displacement method, Equation (55),

$$[S_{m\varepsilon D}] = [S_{m\varepsilon}] \tag{59}$$

and the force method, Equation (58),

$$[S_{m\varepsilon F}] = [S_{m\varepsilon}][k_{m\varepsilon}]^{-1} \tag{60}$$

The relationship between the two forms of output matrices is given by comparison of Equation. (59) and (60), that is,

$$[S_{m\varepsilon F}] = [S_{m\varepsilon D}][k_{m\varepsilon}]^{-1} \tag{61}$$

In the case of the combined method the form of the output matrix can be either of the two forms depending on the choice made by the individual developing the program.

4. THE STRESS ELEMENT

4.1 FORCE-DEFORMATION RELATION

A generalized total stress field, $\{\sigma_\sigma\}$, is expressed in terms of a set of independent force variables, $\{F_{m\sigma}\}$, in the form

$$\{\sigma_\sigma\} = [T_{\sigma F}]\{F_{m\sigma}\} \tag{62}$$

where, the j^{th} column of $[T_{\sigma F}]$ gives the stress distribution, which must satisfy equilibrium, for a unit value of $F_{m\sigma j}$. The set of force variables does not contain the zero stress state. Such a generalized stress-field can be assumed directly, subject to the constraint that it satisfies stress-equilibrium conditions. Alternatively such a stress field can be obtained by assuming stress-functions $\{\Phi\}$, that is,

$$\{\Phi\} = [T_{\Phi B}]\{B\} \tag{63}$$

where

$\{B\}$ = function coefficients, the linear terms (such as, b_1, $b_2 x$, $b_3 y$) correspond to the zero stress state.

The stress field (which satisfies equilibrium) is then obtained by appropriate

differentiation of the stress functions as indicated in Table 1.

The independent force variables, $\{F_{m\sigma}\}$, and the coefficients, $\{B\}$, are then directly related by,

$$\{F_{m\sigma}\} = [T_{FB}]\{B\} \tag{64}$$

Applying the principle of virtual forces to the j^{th} force variable establishes the relation

$$\tilde{F}_{m\sigma_j}\, d_{m\sigma_j} = \int_V \{\tilde{\sigma}_\sigma\}_j^T \{\varepsilon\}\, dV \tag{65}$$

where

$F_{m\sigma_j}$ = the j^{th} independent element force variable

$\{\sigma_\sigma\}_j = f(F_{m\sigma_j})$ the corresponding j^{th}, stress field

$d_{m\sigma_j}$ = element deformation corresponding to $F_{m\sigma_j}$

$\{\varepsilon\} = \{\varepsilon_0\} + \{\varepsilon_e\}$ total generalized strains corresponding to $\{\sigma\}$

$\{\varepsilon_0\}$ = initial strains, free thermal expansions, lack of fit, etc. Stress-free initial strains.

$\{\varepsilon_e\} = \{\varepsilon(\sigma)\}$ stress related strain

V = Volume of element

$\sim$ indicates virtual quantities

Subscript σ indicates quantities relative to an element based on stress assumptions.

If there are β independent force variables then there are β linear equations in the form of Equation (65), therefore, the following system of equations results.

$$\left[\tilde{F}_{m\sigma}\right]\{d_{m\sigma}\} = \int_V [\tilde{\sigma}_\sigma]^T \{\varepsilon\}\, dV \tag{66}$$

where, $\left[\tilde{F}_{m\sigma}\right]$ is a diagonal matrix of virtual force variables. Using Equation (62), the complete set of virtual stress systems is given by,

$$[\tilde{\sigma}_\sigma] = [T_{\sigma F}]\left[\tilde{F}_{m\sigma}\right] \tag{67}$$

Substituting Equation (67) into (66) gives,

$$\left[\tilde{F}_{m\sigma}\right]\{d_{m\sigma}\} = \left[\tilde{F}_{m\sigma}\right]\int_V [T_{\sigma F}]^T \{\varepsilon\}\, dV \tag{68}$$

or,

$$\{d_{m\sigma}\} = \int_V [T_{\sigma F}]^T \{\varepsilon\}\, dV \tag{69}$$

Equation (69) defines the actual element deformation, $d_{m\sigma j}$, corresponding to a chosen force variable, $F_{m\sigma j}$. This deformation is that through which the force variable does complementary work but is caused by the total strain distribution, $\{\varepsilon\}$. As a non-trivial example, if the force variable is a pressure or stress (lb/in.^{2}) then the corresponding deformation is a volumetric displacement (in.^{3}).

The complete strains can be expressed in terms of the stresses and the initial strains

$$\{\varepsilon\} = \{\varepsilon_o\} + \{\varepsilon_e\} \tag{70}$$

For linear stress-strain behavior (see also Equation (26))

$$\{\varepsilon_e\} = [T_{\varepsilon\sigma}]\{\sigma_\sigma\} \tag{71}$$

where,

$$[T_{\varepsilon\sigma}] = [T_{\sigma\varepsilon}]^{-1} \tag{72}$$

Therefore, introducing Equations (70) and (71) into (69),

$$\{d_{m\sigma}\} = \int_V [T_{\sigma F}]^T \{\varepsilon_o\}\, dV + \left(\int_V [T_{\sigma F}]^T [T_{\varepsilon\sigma}][T_{\sigma F}]\, dV\right)\{F_{m\sigma}\} \tag{73}$$

or,

$$\{d_{m\sigma}\} = \{d_{m\sigma o}\} + \{d_{m\sigma e}\} \tag{74}$$

where,

$$\{d_{m\sigma o}\} = \int_V [T_{\sigma F}]^T \{\varepsilon_o\}\, dV \tag{75}$$

= corresponding initial deformations (thermal)

$$\{d_{m\sigma e}\} = [D_{m\sigma}]\{F_{m\sigma}\} \tag{76}$$

and

$$[D_{m\sigma}] = \int_V [T_{\sigma F}]^T [T_{\varepsilon\sigma}][T_{\sigma F}]\, dV \tag{77}$$

- natural flexibility matrix based on stress assumptions.

This natural flexibility matrix, as well as the initial deformations, are for a stress element as well as for a Hybrid I element (Reference 10, Equation 8) based on an internal stress field.

4.2 STRESS ELEMENT ASSEMBLY MATRIX

Corresponding to each independent force variable there is an equilibrated set of boundary force distributions which are directly obtained from the values at the boundaries of the assumed internal stress field. For the i^{th} boundary

$$\{N_{bi}\} = [T_{N\sigma i}]\{\sigma_\sigma\} \tag{78}$$

where,

$\{N_{bi}\}$ = generalized forces on the i^{th} element boundary.

Substituting from Equation (62) into (78) gives

$$\{N_{bi}\} = [T_{NFi}]\{F_{m\sigma}\} \tag{79}$$

where,

$$[T_{NFi}] = [T_{N\sigma i}][T_{\sigma F}] \tag{80}$$

= boundary force distributions for unit values of the independent force variables.

The boundary forces, $\{N_{bi}\}$, are used, in one form or another, to form the equilibrium equations for the assembled structure. In the usual formulation of nodal equilibrium it is necessary to establish a set of nodal forces which are equivalent to the boundary forces. In stress elements this can be achieved by integrating the boundary forces and distributing them to the nodes while still maintaining equilibrium. The element nodal forces which are equivalent to the i^{th} boundary forces, $\{q_{mi}\}$, are given by,

$$\{q_{mi}\} = [T_{qNi}]\{N_{bi}\} \tag{81}$$

and contain dependent forces (static supports). Substituting Equation (79) into (81) and summing for all boundaries results in,

$$\{q_m\} = \left(\sum_{i=1}^{b} [T_{qNi}][T_{NFi}]\right)\{F_{m\sigma}\} \tag{82}$$

or,

$$\{q_m\} = [T_{qF\sigma}]\{F_{m\sigma}\} \tag{83}$$

where,

$$[T_{qF\sigma}] = \sum_{i=1}^{b} [T_{qNi}][T_{NFi}] \tag{84}$$

$\{q_m\}$ = equivalent equilibrated nodal forces

b = number of boundaries.

A more systematic procedure is to use the principle of virtual displacements to find the equivalent nodal forces which are in equilibrium with the boundary forces. This requires that a set of kinematically possible boundary displacement functions be established. These are not consistent with the assumed stress field. This procedure is the basis of Pian's Hybrid I elements.[10,11] For the i^{th} boundary,

$$\{u_{bi}\} = [T_{u\delta i}]\{\delta_m\} \tag{85}$$

where,

$\{\delta_m\}$ = nodal displacements in the local element system including rigid body freedoms.

$\{u_{bi}\}$ = boundary displacement functions.

Application of the principle of virtual displacements, for virtual nodal displacements, results in,

$$\left[\tilde{\delta}_m\right]\{q_m\} = \sum_{i=1}^{b} \int_{S_i} [\tilde{u}_{bi}]^T \{N_{bi}\}\, dS_i \tag{86}$$

Substituting into Equation (86) from (79) and (85),

$$\left[\tilde{\delta}_m\right]\{q_m\} = \left[\tilde{\delta}_m\right]\left(\sum_{i=1}^{b} \int_{S_i} [T_{u\delta i}]^T [T_{NFi}]\, dS_i\right)\{F_{m\sigma}\} \tag{87}$$

where,

S_i = surface variable on the i^{th} boundary.

This gives,

$$\{q_m\} = [T_{qF\sigma}]\{F_{m\sigma}\} \tag{88}$$

where,

$$[T_{qF\sigma}] = \sum_{i=1}^{b} \int_{S_i} [T_{u\delta i}]^T [T_{NFi}]\, dS_i \tag{89}$$

The $[T_{qF\sigma}]$ matrix is referred to as the element force assembly matrix in the local system based on stress assumptions. This matrix transforms the element force variables into an equivalent set of nodal forces in the local element system including dependent forces.

The relation between the corresponding deformations and nodal displacements in the local element system can be found by applying the principle of virtual forces for virtual force variables. Therefore,

$$\left[\tilde{F}_{m\sigma}\right]\{d_{m\sigma}\} = \left[\tilde{F}_{m\sigma}\right]\left(\sum_{i=1}^{b} \int_{S_i} [T_{NFi}]^T [T_{u\delta i}]\, dS_i\right)\{\delta_m\} \tag{90}$$

Hence,

$$\{d_{m\sigma}\} = [T_{d\delta\sigma}]\{\delta_m\} \tag{91}$$

where,

$$[T_{d\delta\sigma}] = \sum_{i=1}^{b} \int_{S_i} [T_{NFi}]^T [T_{u\delta i}]\, dS_i \tag{92}$$

= element displacement assembly matrix in the local system based on stress assumptions.

Referring to Equations (89) and (92) it can be seen that,

$$[T_{d\delta\sigma}] = [T_{qF\sigma}]^T \tag{93}$$

In other words, the matrix which relates the element nodal displacements in the local system, including rigid body freedoms, to the element deformations corresponding to the force variables is the transpose of the element local assembly matri··

In some cases the two methods of obtaining equivalent nodal forces will be identical, as is the case for the rectangular membrane element given in Reference 10, with the stress field:

$$\sigma_x = a_1 + a_2 y$$
$$\sigma_y = a_3 + a_4 x$$
$$\tau_{xy} = a_5$$

Substituting from Equation (88) into (42) gives,

$$\{Q_m\} = [E_{m\sigma}]\{F_{m\sigma}\} \tag{94}$$

where,

$$[E_{m\sigma}] = [T_c][T_{qF\sigma}] \tag{95}$$

= element force assembly matrix in the global system based on stress assumptions.

Substituting Equation (45) into (91) gives,

$$\{d_{m\sigma}\} = [a_{m\sigma}]\{\Delta_m\} \tag{96}$$

where,

$$[a_{m\sigma}] = [T_{d\delta\sigma}][T_c]^T = [E_{m\sigma}]^T \tag{97}$$

= element displacement assembly matrix in the global system based on stress assumptions.

To esteblish equilibrium models, [5,6,24] that is, force contimuity or bcundary force

equilibrium between adjacent element boundaries, the boundary force distributions should be equilibrated directly. These must be resolved into a suitable reference system which will also result in an element assembly matrix. In this case the concept of a node must be considered in a more general sense.

4.3 STRESS ELEMENT OUTPUT MATRIX

In a stress element the stress field is given by Equation (62), that is,

$$\{\sigma_\sigma\} = [T_{\sigma F}]\{F_{m\sigma}\} \tag{98}$$

This equation gives the stress distributions within the stress element in terms of the local element coordinates and generalized force variables. In the force method programs the stress output is evaluated directly from the computed element force variables, $\{F_{m\sigma}\}$, by means of Equation (62), using $[T_{\sigma F}]$ at specific locations. The stress output is given by,

$$\{\sigma_{m\sigma}\} = [S_{m\sigma}]\{F_{m\sigma}\} \tag{99}$$

where

$\{\sigma_{m\sigma}\}$ = vector of required output data for a stress element

$[S_{m\sigma}] = [T_{\sigma F}]$ at specified locations within the element.

To use this element in the displacement method, obtain from Equations (74) and (76),

$$\{F_{m\sigma}\} = [D_{m\sigma}]^{-1}\left(\{d_{m\sigma}\} - \{d_{m\sigma o}\}\right) = [D_m]^{-1}\{d_{m\sigma e}\} \tag{100}$$

Hence, substituting from Equation (100) into (99), the stress output in terms of deformations is given by,

$$\{\sigma_{m\sigma}\} = [S_{m\sigma}][D_{m\sigma}]^{-1}\left(\{d_{m\sigma}\} - \{d_{m\sigma o}\}\right) \tag{101}$$

or

$$\{\sigma_{m\sigma}\} = [S_{m\sigma}][D_{m\sigma}]^{-1}\{d_{m\sigma e}\} \tag{102}$$

The element output matrix will now be denoted by $[S_{m\sigma}]$ with subscripts F or D to denote in which method it is to be used (force or displacement). Therefore for the force method

$$[S_{m\sigma F}] = [S_{m\sigma}] \tag{103}$$

and for the displacement method

$$[S_{m\sigma D}] = [S_{m\sigma}][D_{m\sigma}]^{-1} \tag{104}$$

The relationship between the two forms of output matrices is,

$$[S_{m\sigma D}] = [S_{m\sigma F}][D_{m\sigma}]^{-1} \tag{105}$$

In the case of the combined method either form can be used as desired.

5. SUMMARY AND PRACTICAL SIGNIFICANCE OF CHARACTERISTIC ELEMENT MATRICES

The concept of characteristic element matrices has considerable practical significance. If every computer program for structural analysis was written on the basis of characteristic matrices for any element, complete interchangeability of elements between the various systems could be obtained. This would be irrespective of the analysis method (force, displacement, combined) or the basic element assumption (stress, strain). An analysis system of this nature has been developed at the Lockheed-California Company. Consider now a program based on the displacement method for static analysis which was written to accept four characteristic matrices for an element (elastic, initial deformation, assembly, output). Now let anyone who develops a stress element, write the subroutine for this element to return the characteristic matrices $[D_{m\sigma}]$, $\{d_{m\sigma_0}\}$, $[E_{m\sigma}]$ and $[S_{m\sigma F}]$. This subroutine could then be used directly to introduce the new stress element into the displacement method system by changing a few cards and using a subroutine of the form shown in Figure 2(a). Accompanying the element subroutine would be a report of the element derivation and a check example. The converse situation of developing a strain element and having a program based on the force method would require an element subroutine of the form shown in Figure 2(b). The complete interchangeability of element matrices is shown in Figure 3.

A great many publications have appeared on elements for use in finite element programs. Various types of elements have been developed and published in many international journals and conference proceedings. However, the use of these elements in practice is limited mainly because analysis programs have not been written to readily and rapidly add new elements. In addition, the various programs using even the same method have different ways of introducing elements. The concept of element characteristic matrices presents a unified scheme for element introduction and interchangeability in finite element programs, which would facilitate communications and lead to a wider and more rapid use of new element developments. The concept is demonstrated in the Appendix.

6. ELEMENT SINGULAR STIFFNESS MATRIX

Most researchers and practitioners using the displacement method adopt singular stiffness matrices for an element. It is therefore appropriate to show how this matrix is formed simply using its natural stiffness matrix and its assembly matrix.

The energy within an element has the quadratic form

$$U_m = \tfrac{1}{2} \lfloor \Delta_m \rfloor [k_{mg}] \{\Delta_m\} = \tfrac{1}{2} \lfloor d_m \rfloor [k_m] \{d_m\} \tag{106}$$

where,

$[k_{mg}]$ = element singular stiffness matrix in the global system.

In the case of a strain element,

$$\{d_m\} = \{d_{m\varepsilon}\} = [a_{m\varepsilon}] \{\Delta_m\} \tag{107}$$

and

$$[k_m] = [k_{m\varepsilon}] \tag{108}$$

Substituting from Equations (107) and (108) into (106) gives the element singular stiffness matrix,

$$[k_{mg}] = [k_{m\varepsilon g}] = [a_{m\varepsilon}]^T [k_{m\varepsilon}] [a_{m\varepsilon}] \tag{109}$$

In the case of a stress element,

$$\{d_m\} = \{d_{m\sigma}\} = [E_{m\sigma}]^T \{\Delta_m\} \tag{110}$$

and

$$[k_m] = [D_{m\sigma}]^{-1} \tag{111}$$

Substituting from Equations (110) and (111) into (106) gives,

$$[k_{mg}] = [k_{m\sigma g}] = [E_{m\sigma}] [D_{m\sigma}]^{-1} [E_{m\sigma}]^T \tag{112}$$

Stress Elements	Strain Elements
Assumed Airy type stress functions	**Assumed displacement functions**
1. Axial elements (local axis x)	1. Axial elements (local axis x)
$\{\phi\} = \phi_1$	$\{u\}$ = u, when u is the displacement in the x-direction
2. Membrane elements (local axes x, y)	2. Membrane elements (local axes x, y)
$\{\phi\} = \phi_1$	$\{u\} = \{u\ v\}$, where u and v are the in-plane displacements in the x and y directions respectively
3. Plate bending elements (local axes x, y, z)	3. Plate bending elements (local axes x, y, z)
$\{\phi\} = \{\phi_1 \phi_2\}$	$\{u\}$ = w, where w is the displacement in the z-direction (normal to element plane)
4. Solid elements (local axes x, y, z)	4. Solid elements (local axes x, y, z)
$\{\phi\} = \{\phi_1 \phi_2 \phi_3\}$	$\{v\} = \{u\ v\ w\}$, where u, v and w are the displacements in the x, y and z directions respectively
Generalised stress field	**Generalised strain field**
1. Axial elements	1. Axial elements
$\sigma_x = \frac{\partial \phi_1}{\partial z}$	$\epsilon_x = \frac{\partial u}{\partial x}$
2. Membrane elements	2. Membrane elements
$\sigma_z = \frac{\partial^2 \phi_1}{\partial y^2}$	$\epsilon_x = \frac{\partial u}{\partial z}$
$\sigma_y = \frac{\partial^2 \phi_1}{\partial x^2}$	$\epsilon_y = \frac{\partial v}{\partial y}$
$\tau_{xy} = -\frac{\partial^2 \phi_1}{\partial x \partial y}$	$\gamma_{xy} = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}$
3. Plate bending elements	3. Plate bending elements
$M_x = \frac{\partial \phi_2}{\partial y}$	$\frac{1}{R_x} = \frac{\partial^2 w}{\partial x^2}$
$M_y = \frac{\partial \phi_1}{\partial z}$	$\frac{1}{R_y} = \frac{\partial^2 w}{\partial y^2}$
$M_{xy} = \frac{1}{2}\left(\frac{\partial \phi_1}{\partial y} + \frac{\partial \phi_2}{\partial x}\right)$	$\frac{1}{R_{xy}} = \frac{\partial^2 w}{\partial x \partial y}$
4. Solid elements	4. Solid elements
$\sigma_x = \frac{\partial^2 \phi_3}{\partial y^2} + \frac{\partial^2 \phi_2}{\partial z^2}$	$\epsilon_z = \frac{\partial w}{\partial z}$
$\sigma_y = \frac{\partial^2 \phi_3}{\partial x^2} + \frac{\partial^2 \phi_1}{\partial z^2}$	$\epsilon_y = \frac{\partial v}{\partial y}$
$\sigma_z = \frac{\partial^2 \phi_2}{\partial x^2} + \frac{\partial^2 \phi_1}{\partial y^2}$	$\epsilon_z = \frac{\partial w}{\partial z}$
$\tau_{xy} = -\frac{\partial^2 \phi_3}{\partial x \partial y}$	$\gamma_{xy} = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}$
$\tau_{yz} = -\frac{\partial^2 \phi_1}{\partial y \partial z}$	$\gamma_{yz} = \frac{\partial w}{\partial y} + \frac{\partial v}{\partial z}$
$\tau_{zx} = -\frac{\partial^2 \phi_2}{\partial z \partial x}$	$\gamma_{zx} = \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}$

TABLE 1 FIELD-FUNCTION RELATIONS

	STRESS ELEMENT	STRAIN ELEMENT
BASIC RELATIONS	$d_{m\sigma} = D_{m\sigma} F_{m\sigma} + d_{m\sigma o}$ $Q_m = E_{m\sigma} F_{m\sigma}$ $\sigma_{m\sigma} = S_{m\sigma F} F_{m\sigma}$	$F_{m\varepsilon} = k_{m\varepsilon} d_{m\varepsilon} - F_{m\varepsilon o}$ $d_{m\varepsilon} = a_{m\varepsilon} \Delta_m$ $\sigma_{m\varepsilon} = S_{m\varepsilon D}(d_{m\varepsilon} - d_{m\varepsilon o})$
INTERCHANGEABLE RELATIONS	$F_{m\sigma} = k_{m\sigma} d_{m\sigma} - F_{m\sigma o}$ $d_{m\sigma} = a_{m\sigma} \Delta_m$ $\sigma_{m\sigma} = S_{m\sigma D}(d_{m\sigma} - d_{m\sigma o})$	$d_{m\varepsilon} = D_{m\varepsilon} F_{m\varepsilon} + d_{m\varepsilon o}$ $Q_m = E_{m\varepsilon} F_{m\varepsilon}$ $\sigma_{m\varepsilon} = S_{m\varepsilon F} F_{m\varepsilon}$
IDENTITIES	$k_{m\sigma} = D_{m\sigma}^{-1}$ $F_{m\sigma o} = D_{m\sigma}^{-1} d_{m\sigma o}$ $a_{m\sigma} = E_{m\sigma}^{T}$ $S_{m\sigma D} = S_{m\sigma F} D_{m\sigma}^{-1}$	$D_{m\varepsilon} = k_{m\varepsilon}^{-1}$ $d_{m\varepsilon o} = k_{m\varepsilon}^{-1} F_{m\varepsilon o}$ $E_{m\varepsilon} = a_{m\varepsilon}^{T}$ $S_{m\varepsilon F} = S_{m\varepsilon D} k_{m\varepsilon}^{-1}$

TABLE 2 ELEMENT RELATIONSHIPS AND IDENTITIES

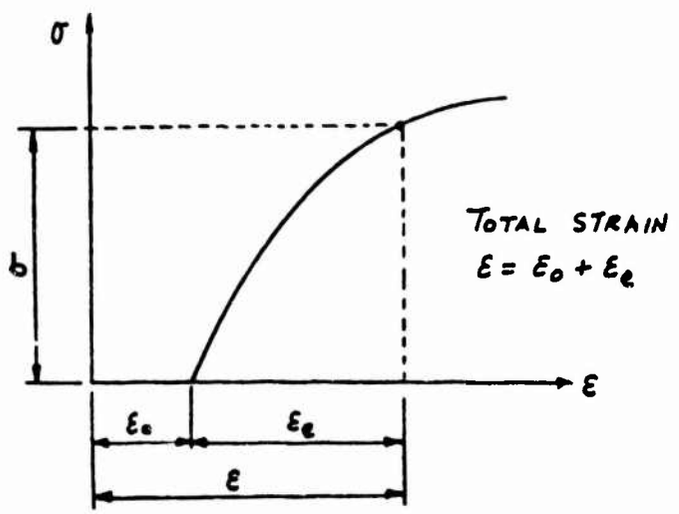


FIGURE 1 TOTAL STRESS-STRAIN RELATION

```
      SUBROUTINE DISEL ( k_mσ , F_mσo , a_mσ , S_mσD )
COMMENT ELEMENTS FOR DISPLACEMENT METHOD PROGRAM USING STRESS ELEMENTS
C     CALL THE FOLLOWING SUBROUTINE FOR A STRESS ELEMENT
      CALL STRESEL (D_mσ , d_mσo , E_mσ , S_mσF )
C     CALCULATE THE IDENTIES
      k_mσ = D_mσ^-1
      F_mσo = D_mσ^-1 d_mσo
      a_mσ = E_mσ^T
      S_mσD = S_mσF D_mσ^-1
      RETURN
      END
```

(a)

```
      SUBROUTINE FORSEL ( D_mε , d_mεo , E_mε , S_mεF )
C     ELEMENTS FOR FORCE METHOD PROGRAM USING STRAIN ELEMENTS
C     CALL THE FOLLOWING SUBROUTINE FOR A STRAIN ELEMENT
      CALL STRAEL ( k_mε , F_mεo , a_mε , S_mεD )
C     CALCULATE THE IDENTITIES
      D_mε = k_mε^-1
      d_mεo = k_mε^-1 F_mεo
      E_mε = a_mε^T
      S_mεF = S_mεD k_mε^-1
      RETURN
      END
```

(b)

FIGURE 2 TYPICAL SUBROUTINES FOR THE EXCHANGE OF ELEMENTS

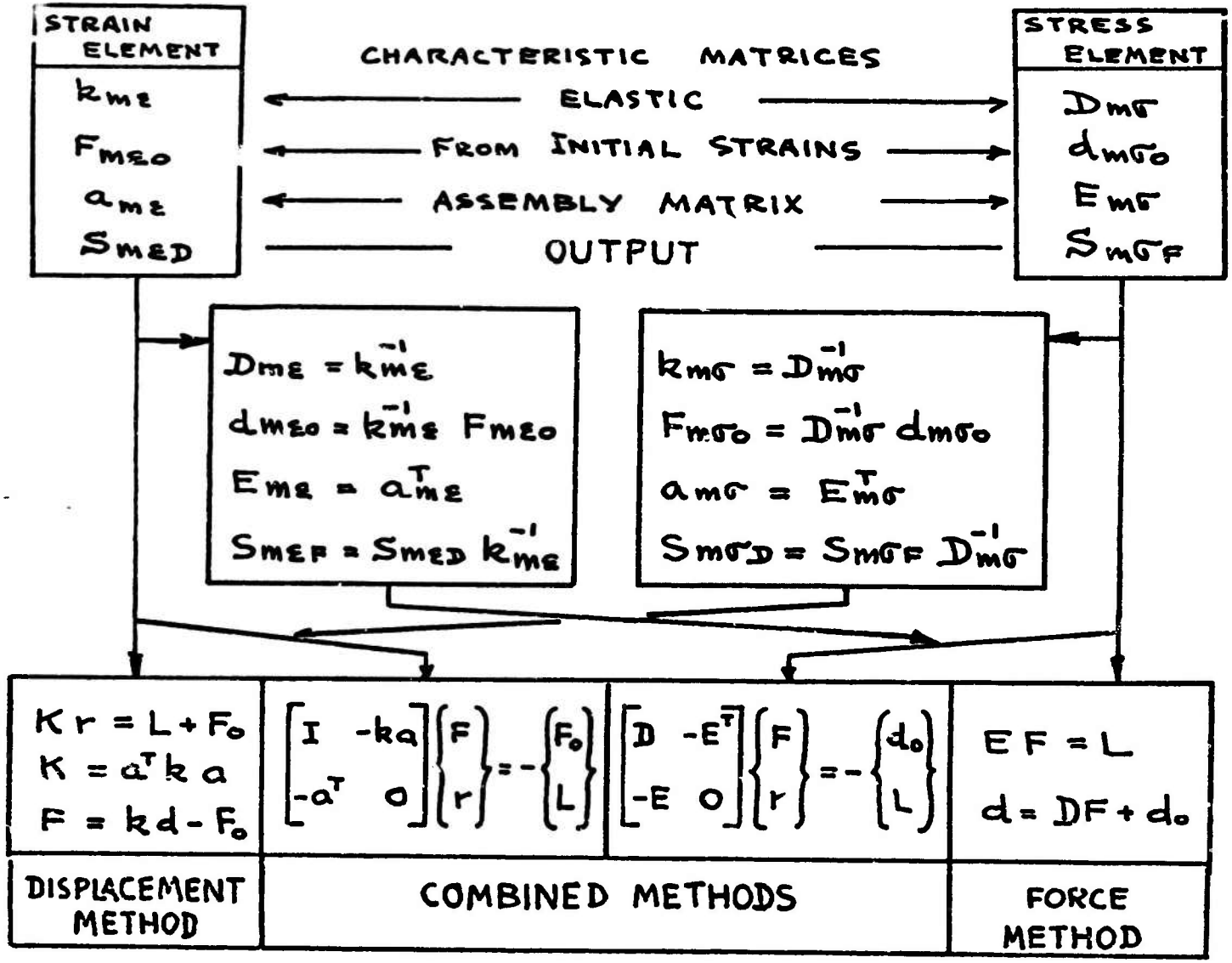


Figure 3. Interchange of Elements

REFERENCES

1. Robinson, J., Williams, R. and Haggenmacher, G. W., "Elastic and Equilibrium Matrices of General Semi-Monocoque Membrane Elements," The Aeronautical Journal, June 1970.

2. Robinson, J., and Haggenmacher, G. W., "Some New Developments in Matrix Force Analysis," Recent Advances in Matrix Methods of Structural Analysis and Design. Gallagher, R. H., Yamada, Y and Oden, J. T. (editors), University of Alabama Press, University of Alabama, 1971.

3. Argyris, J. H., and Kelsey, S., "Energy Theorems and Structural Analysis," Butterworths, London, 1960.

4. Kosko, E., "The Equivalence of Force and Displacement Methods in the Matrix Analysis of Elastic Structures," First Conference Matrix Meth. Struct. Mech., Wright-Patterson Air Force Base, Ohio, Ocotber 1965.

5. Fraeijs de Veubeke, B. M., "Bending and Stretching of Plates-Special Models for Upper and Lower Bounds," First Conf. Matrix Meth. Struct. Mech., Wright-Patterson Air Force Base, Ohio, October 1965.

6. Fraeijs de Veubeke, B. M., Basis of a Well Conditioned Force Program for Equilibrium Models via the Southwell Slab Analogies, Tech. Report AFFDL-TR-67-10, 1967.

7. Elias, Z. M., "Duality in Finite Element Methods," Proc. ASCE, Journal of the Engineering Mech. Div., Vol. 94, No. EM4, August 1968.

8. Pian, T. H. H., and Tong, P., "Basis of Finite Element Methods for Solid Continua," Int. J. Num. Meth. Engng., Vol. 1, No. 1, 1969.

9. Robinson, J. and Haggenmacher, G. W., "Optimization of Redundancy Selection in the Finite Element Force Method," AIAA Journal, Vol. 8, No. 8, August 1970.

10. Pian, T. H. H., "Derivation of Element Stiffness Matrices by Assumed Stress Distributions," AIAA Journal, Vol. 2, pp 1333-1336, July 1964.

11. Pian, T. H. H., "Element Stiffness Matrices for Boundary Compatibility and for Prescribed Boundary Stress," Proc. of First Conf. on Matrix Meth. Struct. Mech. Wright-Patterson Air Force Base, Ohio, 1965.

12. Pian, T. H. H., and Tong, P., "Rationalization in Deriving Element Stiffness Matrix by Assumed Stress Approach," Second Conf. on Matrix Meth. Struct. Mech. Wright-Patterson Air Force Base, Ohio, October 1968.

13. Allwood, R. T., and Cornes, G. M. M., "A Polygonal Finite Element for Plate Bending Problems Using the Assumed Stress Approach," Int. J. Num. Meth. Engng., Vol. 1, 135-149 (1969)

14. Fraeijs de Veubeke, B. M., and Sander, G., "An Equilibrium Model for Plate Bending," Int. J. Num. Meth. Engng., Vol. 4, No. 4, 1968.

15. Gallagher, R. H., Rattinger, I., and Archer, J. S., "A Correlation Study of Methods of Matrix Structural Analysis," AGARDograph 69, The Macmillan Company, New York, 1964.

16. Przemieniecki, J. S., "Theory of Matrix Structural Analysis," McGraw-Hill Book Company, New York, 1968.

17. Robinson , "A Derivation Procedure for the Dynamic Flexibility Matrix of a Triangular Bending Element," Aeronaut. J. 74, 337 (1970).

18. Robinson, J., and Petyt, M., "Dynamic Analysis of Structures Using the Rank Force Method," Int. J. Num. Meth. Engng., Vol. 3, No. 1, Jan-Mar. 1971, pp 103-117.

19. Robinson, J., "Structural Matrix Analysis for the Engineer," John Wiley and Sons, Inc., New York, 1966.

APPENDIX

The concept of characteristic matrices will be demonstrated using a two point beam, bending in one plane, based on both strain and stress assumptions. The required functions are expressed in a non-dimensional local coordinate system with the origin at mid-span. (See Figure A1). This choice of coordinate system, although not essential for the basic principle, is advantageous in the development.

A.1 STRAIN ELEMENT

The principal steps of Section 3 will now be followed.

Displacement function, Equation (18),

$$\{u\} = [T_{uB}]\{B\}$$

$$v = \begin{bmatrix} 1 & \xi & \xi^2 & \xi^3 \end{bmatrix} \{ b_1 \ldots b_4 \} \qquad \text{A.1.1}$$

Displacement derivatives,

$$\frac{dv}{dx} = \frac{1}{a} \begin{bmatrix} 1 & 2\xi & 3\xi^2 \end{bmatrix} \{ b_2\ b_3\ b_4 \} \qquad \text{A.1.2}$$

$$\frac{d^2v}{dx^2} = \frac{1}{a^2} \begin{bmatrix} 2 & 6\xi \end{bmatrix} \{ b_3\ b_4 \} \qquad \text{A.1.3}$$

Strain function, Equation (17),

$$\{\varepsilon_\varepsilon\} = [T_{\varepsilon d}]\{d_{m\varepsilon}\}$$

$$\varepsilon_x = -y\frac{d^2v}{dx^2} = -\frac{y}{a^2}\begin{bmatrix} 2 & 6\xi \end{bmatrix}\{b_3 \;\; b_4\}$$

$$\varepsilon_x = \frac{-y}{a^2}\begin{bmatrix} 1 & \xi \end{bmatrix}\{d_1 \;\; d_2\} \qquad \text{A.1.4}$$

where d_1 and d_2 are the chosen independent deformation variables. Equation (19) becomes,

$$\{d_{m\sigma}\} = [T_{dB}]\{B\}$$

$$\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \qquad \text{A.1.5}$$

The natural strain modes of Equation (A.1.4) are shown in Figure A6.

Stress-strain behaviour, Equation (26)

$$\{\sigma\} = [T_{\sigma\varepsilon}]\{\varepsilon_e\}$$

$$[T_{\sigma\varepsilon}] = E \quad \text{(Modulus of Elasticity)} \qquad \text{A.1.6}$$

Force-deformation relation, Equation (28),

$$\{F_{m\varepsilon}\} = [k_{m\varepsilon}]\{d_{m\varepsilon}\} - \{F_{m\varepsilon o}\}$$

Initial element forces, Equation (29),

$$\{F_{m\varepsilon o}\} = \int_V [T_{\varepsilon d}]^T [T_{\sigma\varepsilon}]\{\varepsilon_o\} dV$$

$$\{F_{m\varepsilon o}\} = \int_V \left(-\frac{y}{a^2}\right)\begin{bmatrix}1\\ \xi\end{bmatrix} E\,\varepsilon_o(\xi,y)\, dV \qquad \text{A.1.7}$$

Often $\{\varepsilon_o\}$ is assumed in the same form as $\{\varepsilon\}$, however, Equation A.1.7 can be integrated numerically or, in the case of thermal strains, as part of the thermal program.

Natural stiffness matrix, Equation (30),

$$[k_{m\varepsilon}] = \int_V [T_{\varepsilon d}]^T [T_{\sigma\varepsilon}][T_{\varepsilon d}] dV$$

$$[k_{m\varepsilon}] = \int_V \left(-\frac{y}{a^2}\right)\begin{bmatrix}1\\ \xi\end{bmatrix} E \left(-\frac{y}{a^2}\right)[1 \;\; \xi]\, dA\, a\, d\xi$$

$$= \int_{-1}^{1} \frac{E}{a^3}\begin{bmatrix}1 & \xi\\ \xi & \xi^2\end{bmatrix}\left(\int_A y^2 dA\right) d\xi$$

$$[k_{m\varepsilon}] = \frac{2}{3}\left(\frac{EI}{a^3}\right)\begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix} \qquad \text{A.1.8}$$

The simple form of $[k_{m\varepsilon}]$, that is the variables are uncoupled, should be noted. This results from the choice of natural coordinate axes and independent variables.

Element nodal displacements in the local system, see Figure A1,

$$\delta_1 = (v)_{\xi=-1} \quad , \quad \delta_2 = \left(\frac{dv}{dx}\right)_{\xi=-1}$$

$$\delta_3 = (v)_{\xi=1} \quad , \quad \delta_4 = \left(\frac{dv}{dx}\right)_{\xi=1} \qquad \text{A.1.9}$$

Using Equations (A.1.1) and (A.1.2), Equation (34)

$$\{\delta_m\} = [T_{\delta B}]\{B\}$$

becomes

$$\begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 0 & \frac{1}{a} & -\frac{2}{a} & \frac{3}{a} \\ 1 & 1 & 1 & 1 \\ 0 & \frac{1}{a} & \frac{2}{a} & \frac{3}{a} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \qquad \text{A.1.10}$$

Hence, Equation (35),

$$\{B\} = [T_{\delta B}]^{-1}\{\delta_m\}$$

$$[T_{\delta B}]^{-1} = \frac{1}{4}\begin{bmatrix} 2 & a & 2 & -a \\ -3 & -a & 3 & -a \\ 0 & -a & 0 & a \\ 1 & a & -1 & a \end{bmatrix} \qquad \text{A.1.11}$$

Element displacement assembly matrix in the local system, Equations (36) and (37),

$$\{d_{m\varepsilon}\} = [T_{d\delta\varepsilon}]\{\delta_m\}$$

$$[T_{d\delta\varepsilon}] = [T_{dB}][T_{\delta B}]^{-1}$$

Therefore, using Equations (A.1.5) and (A.1.11),

$$[T_{d\delta\varepsilon}] = \frac{1}{4}\begin{bmatrix} 0 & -2a & 0 & 2a \\ 6 & 6a & -6 & 6a \end{bmatrix} \qquad \text{A.1.12}$$

Coordinate transformation, Equation (45),

$$\{\delta_m\} = [T_c]^T\{\Delta_m\}$$

Referring to Figures A1, A2 and A3,

$$[T_c] = \begin{bmatrix} (\vec{m}_1\cdot\vec{n}_2) & 0 & 0 & 0 \\ (\vec{m}_2\cdot\vec{n}_2) & 0 & 0 & 0 \\ 0 & (\vec{m}_3\cdot\vec{n}_3) & 0 & 0 \\ 0 & 0 & (\vec{m}_1\cdot\vec{n}_2) & 0 \\ 0 & 0 & (\vec{m}_2\cdot\vec{n}_2) & 0 \\ 0 & 0 & 0 & (\vec{m}_3\cdot\vec{n}_3) \end{bmatrix} \qquad \text{A.1.13}$$

where, $\vec{m}_1, \vec{m}_2, \vec{m}_3$ are unit vectors in the global system

and $\vec{n}_1, \vec{n}_2, \vec{n}_3$ are unit vectors in the local system.

$$\vec{n}_1 = \frac{\vec{V2} - \vec{V1}}{abs(\vec{V2}-\vec{V1})} = \frac{\vec{V12}}{abs(\vec{V12})}$$

$$\vec{n}_3 = \frac{\vec{V12} \times \vec{V13}}{abs(\vec{V12} \times \vec{V13})}$$

$$\vec{n}_2 = \vec{n}_3 \times \vec{n}_1$$

A.1.14

Element displacement assembly matrix in the global system, Equations (46) and (47),

$$\{ d_{m\varepsilon} \} = [a_{m\varepsilon}] \{ \Delta_m \}$$

$$[a_{m\varepsilon}] = [T_{ds\varepsilon}] [T_c]^T$$

Using Equations (A.1.12) and (A.1.13)

$$[a_{m\varepsilon}] = \frac{1}{4} \begin{bmatrix} 0 & 0 & -2a(\vec{m}_3 . \vec{n}_3) & 0 & 0 & 2a(\vec{m}_3 . \vec{n}_3) \\ 6(\vec{m}_1 . \vec{n}_2) & 6a(\vec{m}_2 . \vec{n}_2) & 6a(\vec{m}_3 . \vec{n}_3) & -6(\vec{m}_1 . \vec{n}_2) & -6(\vec{m}_2 . \vec{n}_2) & 6a(\vec{m}_3 . \vec{n}_3) \end{bmatrix}$$

A.1.15

To obtain the output matrix determine the stress distributions within the element, Equation (54),

$$\{\sigma_e\} = [T_{\sigma\varepsilon}]\{\varepsilon_e\} = [T_{\sigma\varepsilon}][T_{\varepsilon d}]\{d_{m\varepsilon e}\}$$

$$\sigma_x = -y\left(\frac{E}{a^2}\right)\begin{bmatrix} 1 & \cdots & 3 \end{bmatrix}\{d_{m\varepsilon e}\} \qquad \text{A.1.16}$$

NOTE: $\{d_{m\varepsilon e}\} = \{d_{m\varepsilon}\} - \{d_{m\varepsilon o}\}$, Equation (52)

$\{d_{m\varepsilon}\} = \{d_1 \quad d_2\}$

$\{d_{m\varepsilon o}\} = [k_{m\varepsilon}]^{-1}\{F_{m\varepsilon o}\}$, Equation (32)

$\{F_{m\varepsilon o}\}$, Equation (29)

Stresses could also be directly calculated from the element nodal forces given by Equation (40).

Element output matrix for the displacement method, Equations (55) and (59),

$$\{\sigma_{m\varepsilon}\} = [S_{m\varepsilon}]\{d_{m\varepsilon e}\}$$

$$[S_{m\varepsilon D}] = [S_{m\varepsilon}]$$

$$\{\sigma_{m\varepsilon}\} = \{ (\sigma_x)_{\substack{x=-1 \\ y=c}} \quad (\sigma_x)_{\substack{x=1 \\ y=c}} \} \qquad \text{A.1.17}$$

From Equation (A.1.16)

$$[S_{m\varepsilon D}] = -c\left(\frac{E}{a^2}\right)\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \qquad \text{A.1.18}$$

Summary of the strain element characteristic matrices for the displacement method,

$[k_{m\varepsilon}]$	Equation A.1.8,	Natural Stiffness Matrix
$\{F_{m\varepsilon o}\}$	Equation A.1.7,	Initial Force Vector
$[a_{m\varepsilon}]$	Equation A.1.15,	Assembly Matrix
$[S_{m\varepsilon D}]$	Equation A.1.18,	Output Matrix

The strain element characteristic matrices for the force method are obtained using the transformations given in Figure 3, that is,

$$[D_{m\varepsilon}] = [k_{m\varepsilon}]^{-1} \quad \text{, Natural Flexibility Matrix}$$

$$\{d_{m\varepsilon o}\} = [k_{m\varepsilon}]^{-1}\{F_{m\varepsilon o}\} \text{, Initial Deformation Vector}$$

$$[E_{m\varepsilon}] = [a_{m\varepsilon}]^{T} \quad \text{, Assembly Matrix}$$

$$[S_{m\varepsilon F}] = [S_{m\varepsilon D}][k_{m\varepsilon}]^{-1} \text{, Output Matrix}$$

A.2 STRESS ELEMENT

The principal steps of Section 4 will now be followed.

Stress field, Equation (62),

$$\{\sigma_\sigma\} = [\,T_{\sigma F}\,]\{F_{m\sigma}\}$$

$$\{M\} = M = \begin{bmatrix} 1 & \xi \end{bmatrix}\{F_1 \;\; F_2\} \qquad \text{A.2.1}$$

$$\{\sigma_\sigma\} = \sigma_x = -\frac{y}{I}M = -\frac{y}{I}\begin{bmatrix} 1 & \xi \end{bmatrix}\{F_1 \;\; F_2\} \qquad \text{A.2.2}$$

Alternatively, assuming a stress function, Equation (63),

$$\{\Phi\} = [\,T_{\Phi B}\,]\{B\}$$

$$\Phi = \begin{bmatrix} 1 & \xi & y & \xi^2 & \xi y \end{bmatrix}\{b_1 \ldots b_5\} \qquad \text{A.2.3}$$

$$M = \frac{\partial \Phi}{\partial y} = \begin{bmatrix} 1 & \xi \end{bmatrix}\{b_3 \;\; b_5\} \qquad \text{A.2.4}$$

$$\sigma_x = -\frac{y}{I}\begin{bmatrix} 1 & \xi \end{bmatrix}\{b_3 \;\; b_5\} \qquad \text{A.2.5}$$

Equation (64),

$$\{F_{m\sigma}\} = [\,T_{FB}\,]\{B\}$$

becomes,

$$\begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_5 \end{bmatrix} \qquad \text{A.2.6}$$

The natural stress modes of Equation A.2.2 are shown in Figure A7.

Stress-strain behaviour, Equation (71),

$$\{\varepsilon_e\} = [T_{\varepsilon\sigma}]\{\sigma_\sigma\}$$

$$[T_{\varepsilon\sigma}] = \frac{1}{E} \qquad \text{A.2.7}$$

Force-deformation relation, Equations (74) and (76),

$$\{d_{m\sigma}\} = [D_{m\sigma}]\{F_{m\sigma}\} + \{d_{m\sigma o}\}$$

Initial element deformations, Equation (75),

$$\{d_{m\sigma o}\} = \int_V [T_{\sigma F}]^T \{\varepsilon_o\}\, dV$$

$$\{d_{m\sigma o}\} = \int_V \left(-\frac{y}{I}\right)\begin{bmatrix}1\\ \xi\end{bmatrix} \varepsilon_o(\xi, y)\, dV \qquad \text{A.2.8}$$

Natural flexibility matrix, Equation (77),

$$[D_{m\sigma}] = \int_V [T_{\sigma F}]^T [T_{\varepsilon\sigma}][T_{\sigma F}]\, dV$$

$$[D_{m\sigma}] = \int_V \left(-\frac{y}{I}\right)\begin{bmatrix}1\\ \xi\end{bmatrix} \frac{1}{E}\left(-\frac{y}{I}\right)\begin{bmatrix}1 & \xi\end{bmatrix} dA\, a\, d\xi$$

$$= \int_{-1}^{1} \frac{a}{EI^2}\begin{bmatrix}1 & \xi\\ \xi & \xi^2\end{bmatrix}\left(\int_A y^2\, dA\right) d\xi$$

$$[D_{m\sigma}] = \frac{2}{3}\left(\frac{a}{EI}\right)\begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix} \qquad \text{A.2.9}$$

Again, the simple form of $[D_{m\sigma}]$, resulting from the choice of coordinates and independent variables, should be noted.

Forces on the ith element boundary, Equation (79),

$$\{N_{bi}\} = [T_{NFi}]\{F_{mo}\}$$

From Equation (A.2.1),

$$M = -\frac{I}{y}\sigma_x = \begin{bmatrix} 1 & \xi \end{bmatrix}\{F_1 \; F_2\} \qquad \text{A.2.10}$$

The shear is given by,

$$Q = -\frac{1}{a}\frac{dM}{d\xi} = \frac{I}{ay}\frac{d\sigma_x}{d\xi} = -\frac{1}{a}\begin{bmatrix} 0 & 1 \end{bmatrix}\{F_1 \; F_2\} \qquad \text{A.2.11}$$

See Figures A4 and A5.

For the first boundary (node i = 1),

$$\{N_{b1}\} = \begin{bmatrix} N_{M1} \\ N_{Q1} \end{bmatrix} = \begin{bmatrix} -(1)_{\xi=-1} \\ -(Q)_{\xi=-1} \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & \frac{1}{a} \end{bmatrix}\begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \qquad \text{A.2.12}$$

For the second boundary (node i = 2),

$$\{N_{b2}\} = \begin{bmatrix} N_{M2} \\ N_{Q2} \end{bmatrix} = \begin{bmatrix} (M)_{\xi=1} \\ (Q)_{\xi=1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & -\frac{1}{a} \end{bmatrix}\begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \qquad \text{A.2.13}$$

The generalized forces on the boundaries have a positive vector sign convention.

The element nodal forces, in the local system, which are equivalent to the ith boundary forces, Equation (81),

$$\{ q_{mi} \} = [T_{qNi}] \{ N_{bi} \}$$

See Figures A1 and A5.

For the first boundary (node i = 1),

$$\{ q_{m1} \} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} N_{M1} \\ N_{Q1} \end{bmatrix}$$

A.2.14

For the second boundary (node i = 2),

$$\{ q_{m2} \} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} N_{M2} \\ N_{Q2} \end{bmatrix}$$

A.2.15

Element force assembly matrix in the local system, Equations (83) and (84),

$$\{ q_m \} = [T_{qF\sigma}] \{ F_{m\sigma} \}$$

$$[T_{qF\sigma}] = \sum_{i=1}^{b} [T_{qNi}] [T_{NFi}]$$

Substituting from Equations (A.2.12) and (A.2.13) into (A.2.14) and (A.2.15) respectively, gives,

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 0 & \frac{1}{a} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & -\frac{1}{a} \end{bmatrix} \right) \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$$

that is,

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{a} \\ -1 & 1 \\ 0 & -\frac{1}{a} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$$

A.2.16

The boundary displacement functions for the hybrid approach are given for the ith boundary by Equation (85),

$$\{ u_{bi} \} = [T_{u\delta i}] \{ \delta_m \}$$

For the first boundary (node i = 1),

$$\{u_{b1}\} = \begin{bmatrix} \left(\frac{dv}{dx}\right)_1 \\ (v)_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} \qquad \text{A.2.17}$$

For the second boundary (node i = 2),

$$\{u_{b2}\} = \begin{bmatrix} \left(\frac{dv}{dx}\right)_2 \\ (v)_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} \qquad \text{A.2.18}$$

Element force assembly matrix in the local system, hybrid approach, Equations (88) and (89),

$$\{q_m\} = [T_{qF\sigma}]\{F_{m\sigma}\}$$

$$[T_{qF\sigma}] = \sum_{i=1}^{b} \int_{S_i} [T_{u\delta i}]^T [T_{NFi}]\, dS_i$$

Substituting from Equations (A.2.12), A.2.13), (A.2.17) and (A.2.18),

$$\begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{a} \\ -1 & 1 \\ 0 & -\frac{1}{a} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \qquad \text{A.2.19}$$

In this example it can be seen that the element force assembly matrix in the local system is the same using either the stress integration approach or the hyorid approach. This is because for this element the set of nodal forces is unique.

Element force assembly matrix in the global system, Equations (94) and (95),

$$\{Q_m\} = [E_{m\sigma}]\{F_{m\sigma}\}$$

$$[E_{m\sigma}] = [T_c][T_{qF\sigma}]$$

Using Equations (A.1.13) and (A.2.19),

$$[E_{m\sigma}] = \begin{bmatrix} 0 & \frac{1}{a}(\vec{m}_1 \cdot \vec{n}_2) \\ 0 & \frac{1}{a}(\vec{m}_2 \cdot \vec{n}_2) \\ -(\vec{m}_3 \cdot \vec{n}_3) & (\vec{m}_3 \cdot \vec{n}_3) \\ 0 & -\frac{1}{a}(\vec{m}_1 \cdot \vec{n}_2) \\ 0 & -\frac{1}{a}(\vec{m}_2 \cdot \vec{n}_2) \\ (\vec{m}_3 \cdot \vec{n}_3) & (\vec{m}_3 \cdot \vec{n}_3) \end{bmatrix}$$

A.2.20

To obtain the output matrix use the stress field given by Equation (98),

$$\{\sigma_\sigma\} = [T_{\sigma F}]\{F_{m\sigma}\}$$

that is,

$$\sigma_x = -\frac{y}{I} \begin{bmatrix} 1 & \xi \end{bmatrix} \{F_1 \quad F_2\} \qquad \text{A.2.21}$$

Element output matrix for the force method, Equations (99) and (103),

$$\{\sigma_{m\sigma}\} = [S_{m\sigma F}]\{F_{m\sigma}\}$$

$$\{\sigma_{m\sigma}\} = \{ (\sigma_x)_{\substack{\xi=-1 \\ y=c}} \quad (\sigma_x)_{\substack{\xi=1 \\ y=c}} \} \qquad \text{A.2.22}$$

From Equation (A.2.22),

$$[S_{m\sigma F}] = -\frac{c}{I} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \qquad \text{A.2.23}$$

Summary of the stress element characteristic matrices for the force method,

$[D_{m\sigma}]$ Equation A.2.9, Natural Flexibility Matrix

$\{d_{m\sigma_o}\}$ Equation A.2.8, Initial Deformation Vector

$[E_{m\sigma}]$ Equation A.2.20, Assembly Matrix

$[S_{m\sigma F}]$ Equation A.2.24, Output Matrix

The stress element characteristic matrices for the displacement method are obtained using the transformations given in Figure 3, that is,

$$[k_{m\sigma}] = [D_{m\sigma}]^{-1}$$, Natural Stiffness Matrix

$$\{F_{m\sigma_o}\} = [D_{m\sigma}]^{-1}\{d_{m\sigma_o}\}$$, Initial Force Vector

$$[a_{m\sigma}] = [E_{m\sigma}]^{T}$$, Assembly Matrix

$$[S_{m\sigma D}] = [S_{m\sigma F}][D_{m\sigma}]^{-1}$$, Output Matrix

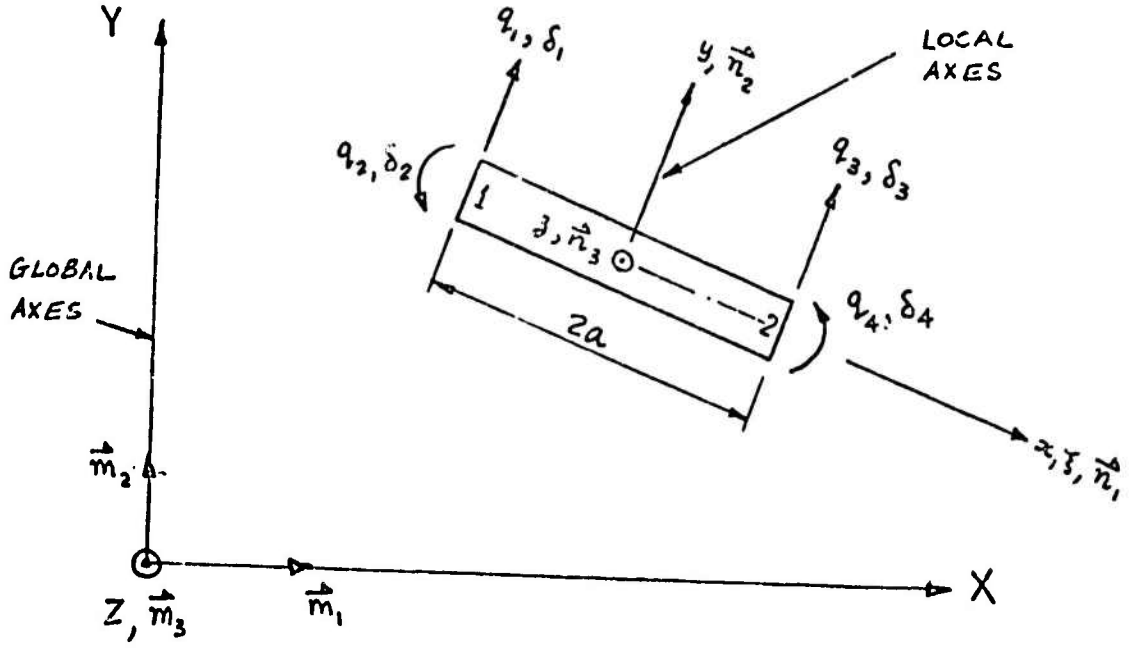


FIGURE A1 BEAM POSITIVE NODAL FORCES AND DISPLACEMENTS IN THE LOCAL SYSTEM.

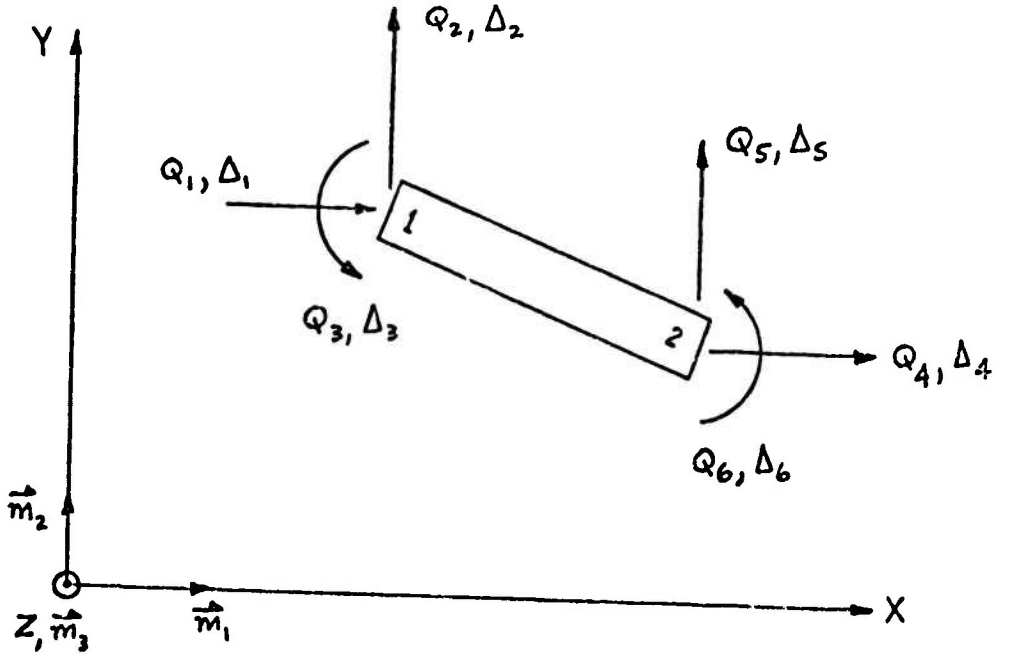


FIGURE A2 BEAM POSITIVE NODAL FORCES AND DISPLACEMENTS IN THE GLOBAL SYSTEM

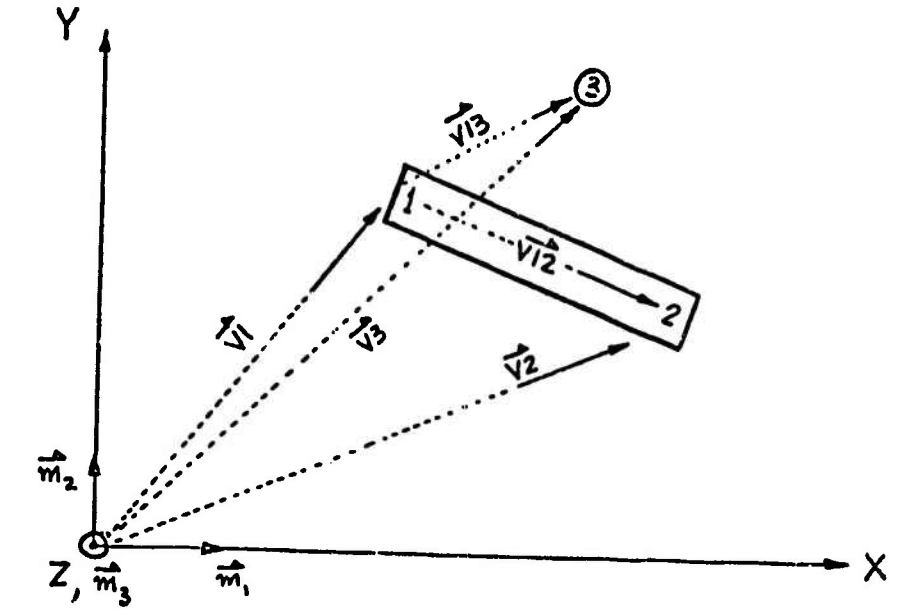


FIGURE A3 ELEMENT VECTORS

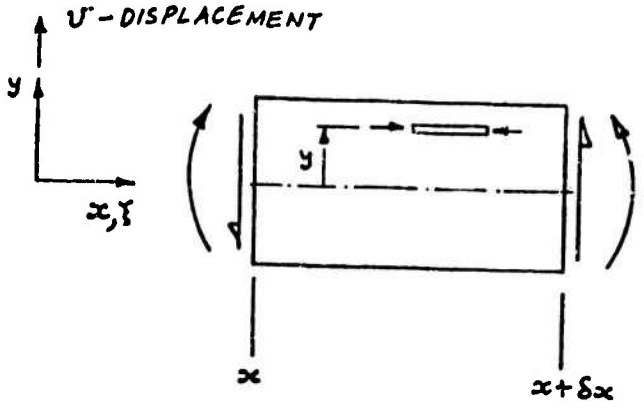


FIGURE A4 POSITIVE INCREMENTAL FORCES AND DISPLACEMENTS

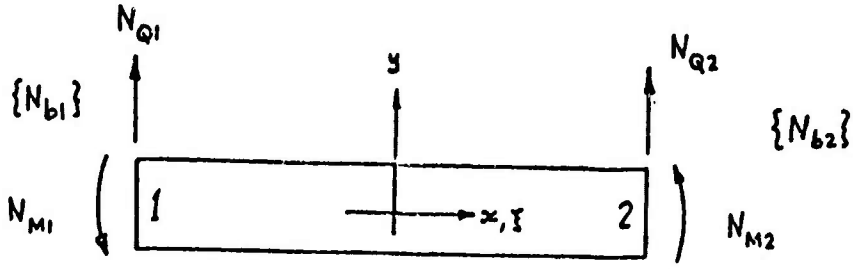
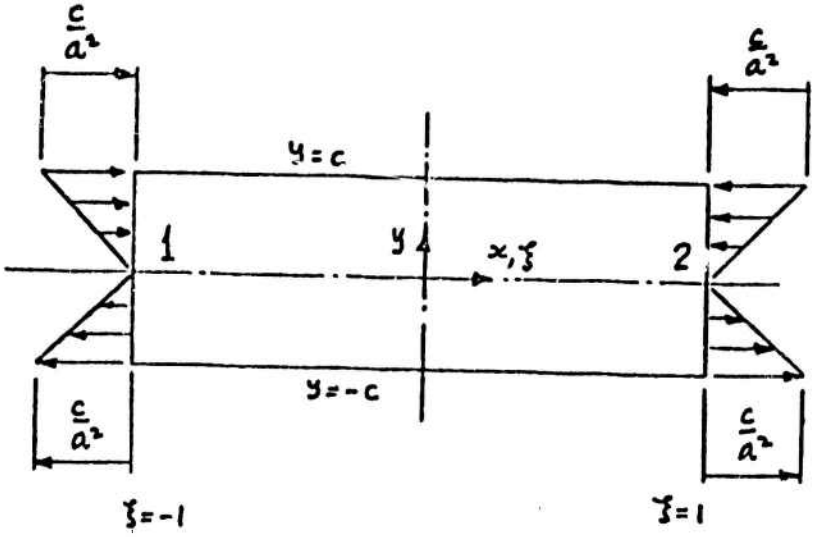
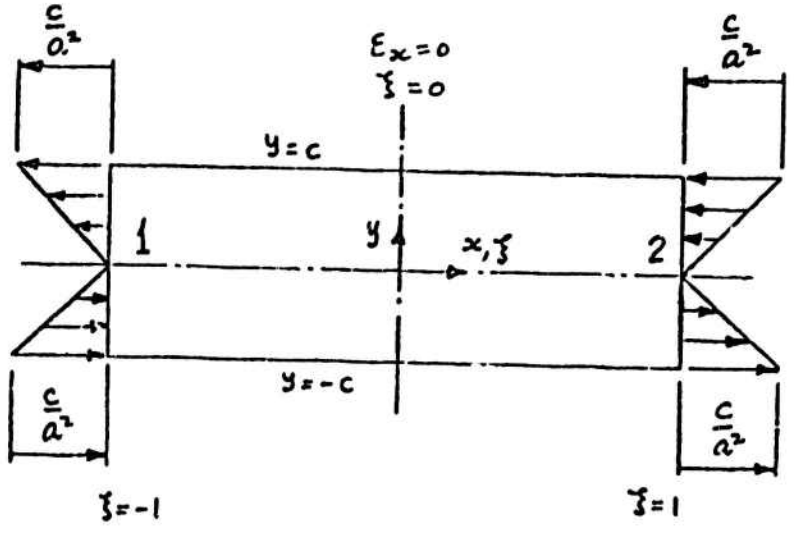


FIGURE A5 POSITIVE BOUNDARY FORCE DISTRIBUTIONS



$$\varepsilon_{xc} = -\frac{y}{a^2}$$

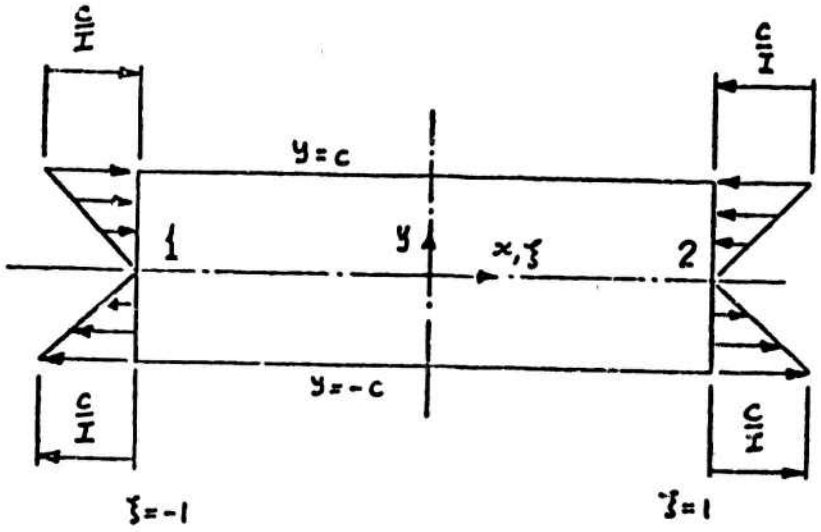
(a) FIRST INDEPENDENT STRAIN MODE, CONSTANT CURVATURE, $d_1 = 1$, $d_2 = 0$.



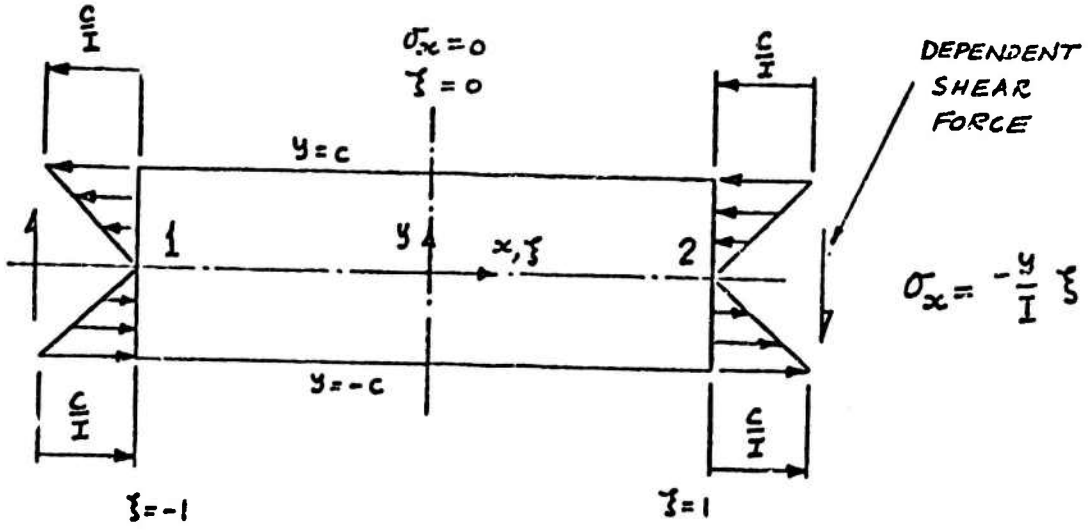
$$\varepsilon_x = -\frac{y}{a^2}\,\xi$$

(b) SECOND INDEPENDENT STRAIN MODE, LINEAR CURVATURE, $d_1 = 0$, $d_2 = 1$.

FIGURE A6 NATURAL STRAIN MODES



(a) FIRST INDEPENDENT STRESS MODE, CONSTANT MOMENT, $F_1 = 1$, $F_2 = 0$.



(B) SECOND INDEPENDENT STRESS MODE, LINEAR MOMENT, $F_1 = 0$, $F_2 = 1$.

FIGURE A7 NATURAL STRESS MODES

REFINED MIXED METHOD FINITE ELEMENTS FOR SHELLS OF REVOLUTION

Phillip L. Gould*
Subir K. Sen**

Washington University, St. Louis, Mo.

A rotational shell finite element with mixed stress and displacement variables is derived for the case of arbitrary static loading. Accurate representation of the shell geometry together with high-order approximations for the dependent variables are combined to formulate a comparatively precise element containing only the basic number of degrees of freedom. Representative examples are used to study the convergence of the element and to provide a basis of comparison with a similar displacement-type element.

* Associate Professor of Civil and Environmental Engineering
** Graduate Student, Civil and Environmental Engineering Department

INTRODUCTION

Finite elements for which the explicit nodal variables are a mixed set of stress and displacement terms have generally been derived from some form of the Hellinger-Reissner variational principle (1,2). For plate and shell problems, the mixed formulation appears to offer some significant advantages compared to the more conventional displacement approach (3). In the computation of moments, which are important design parameters, the advantage is quite obvious. In the mixed method the stress terms representing the moment resultants are determined directly as nodal variables, whereas in the displacement formulation they are computed by differentiating the approximate displacement functions and applying the constitutive relations. The mixed method therefore not only yields moments which are continuous over the inter-element boundary but also avoids the error usually introduced by differentiating the approximating functions. On the other hand, mixed-type elements usually have more nodal variables than displacement-type elements, but the difference is not large especially if transverse shear effects are included in both formulations.

The evolution of mixed-type finite elements for plates and shells has followed a somewhat similar pattern to that of displacement-type elements beginning with triangular plate bending elements with linear and constant moment and linear displacement functions (3,4) followed by the extension to linear moment and parabolic displacement fields (5). A generalization of the plate bending element to include vibration and instability problems has been noted (6). Just as in the displacement method, the first mixed formulation for shells utilized combined plane stress-plate bending flat elements to approximate the curved surface (7). Mixed-type triangular curved elements based on shallow shell theory using polynomial approximations of various degrees (8,9,10,11) and on non-shallow shell theory (8) have been proposed. Also, non-linear shallow shell elements have been described (8,11). However, in contrast to the pattern of development for the displacement-type elements the mixed-type axisymmetric shell element has received only limited attention to date (12).

It is apparent that axisymmetric problems can be treated using the general shell elements, but the inherent computational advantages provided by harmonic decoupling and a strongly banded set of algebraic equations make the use of specialized rotational shell elements attractive for a large class of problems. Accordingly, the objectives of this study are to derive and test a mixed-type curved element for a thin elastic shell of revolution under arbitrary static loading. Following an approach implemented for a displacement-type element (13,14), accurate geometric data will be utilized to the fullest extent and high-order polynomial approximations will be employed when necessary to achieve a comparatively precise solution with relatively few elements.

GEOMETRY

The geometry of the shell element is shown in Fig. 1. The equation of the meridian curve, $R = R(Z)$ is used to express the principal meridional radius of curvature R_φ as a function of the vertical coordinate Z

$$R_\varphi = - \frac{[1 + (R')^2]^{3/2}}{R''} \tag{1}$$

in which φ is the meridional curvilinear coordinate. In Eq. 1 the symbol $( )' = d( )/dZ$. The arc length for element i between nodes i and i+1, is

$$L_i = \int_{Z_i}^{Z_{i+1}} [1 + (R')^2]^{1/2} \, dZ \tag{2}$$

from which the nondimensional arc length variable

$$s = S/L_i \quad (o \leq s \leq 1) \tag{3}$$

is defined.

In the following derivation, it is necessary to evaluate numerous integrals which contain various combinations of the geometrical parameters R_φ and R. This is conveniently accomplished by representing each of these terms by a fourth-order Lagrangian interpolation polynomial in the variable s and then evaluating the subsequent integrals in closed form (13,15).

For closed shells of revolution with $\varphi = 0$ at the pole Eqs. 1 and 2 cannot be applied at the pole since $R' \rightarrow \infty$. A coordinate transformation for the cap element enables R_φ and L_i to be evaluated (15).

FINITE ELEMENT DERIVATION

Definition of Variables

For shells of revolution, it is convenient to take all variables in Fourier series form, to obtain the solution for each harmonic and then to combine the contributions of each harmonic by superposition. Accordingly, the displacements and rotations are taken as

$$\{D\} = \{D_f \mid D_c\} = \{U \; V \; W \mid B_\varphi \; B_\theta\} = \sum_{j=o}^{\infty} [\Theta_1] \{D^{(j)}(s)\} \tag{4a}$$

in which

$$[\Theta_1] = [\cos j\theta \; \sin j\theta \; \cos j\theta \; \cos j\theta \; \sin j\theta] \tag{4b}$$

and

$$\{D^{(j)}(s)\} = \{D_f^{(j)} \mid D_c^{(j)}\} = \{u^{(j)} \; v^{(j)} \; w^{(j)} \mid \beta_\varphi^{(j)} \; \beta_\theta^{(j)}\} \tag{4c}$$

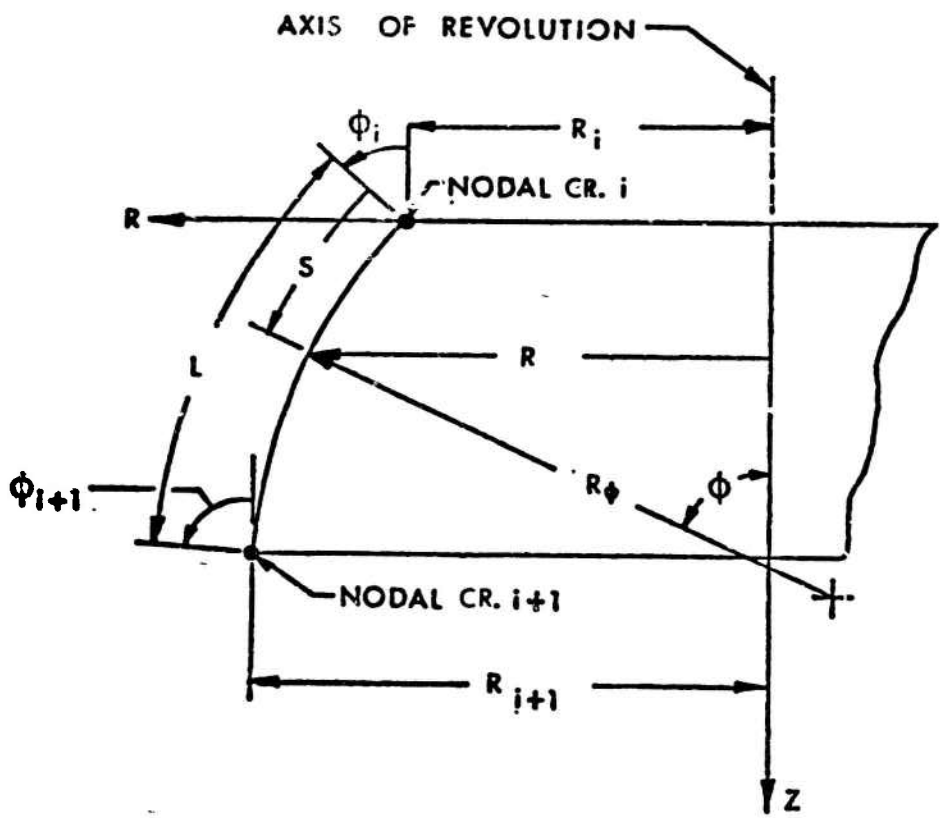


FIGURE 1 FINITE ELEMENT GEOMETRY

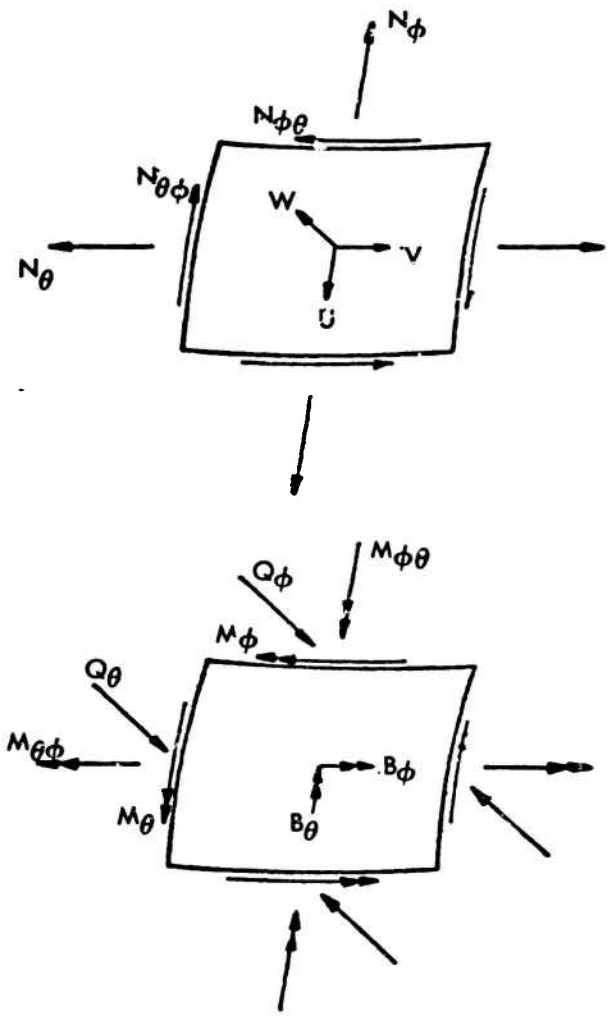


FIGURE 2 SIGN CONVENTION

In Eq. 4, U, V and W = the meridional, circumferential and normal displacements; β_φ and β_θ = the meridional and circumferential rotations, positive as shown in Fig. 2; and θ = the circumferential coordinate.

The strains and changes in curvature are taken as

$$\{H\} = \{H_f \mid H_c\} = \{E_\varphi \; E_{\varphi\theta} \; \Gamma_\varphi \; E_\theta \; E_{\theta\varphi} \; \Gamma_\theta \mid K_\varphi \; K_{\varphi\theta} \; K_\theta \; K_{\theta\varphi}\}$$

$$= \sum_{j=o}^{\infty} [\Theta_2] \{H^{(j)}(s)\} \tag{5a}$$

in which

$$[\Theta_2] = [\cos j\theta \; \sin j\theta \; \cos j\theta \; \cos j\theta \; \sin j\theta \; \sin j\theta \; \cos j\theta \; \sin j\theta \; \cos j\theta \; \sin j\theta] \tag{5b}$$

and

$$\{H^{(j)}(s)\} = \{H_{f\varphi}^{(j)} \mid H_{f\theta}^{(j)} \mid H_{c\varphi}^{(j)} \mid H_{c\theta}^{(j)}\}$$

$$= \{\epsilon_\varphi^{(j)} \epsilon_{\varphi\theta}^{(j)} \gamma_\varphi^{(j)} \mid \epsilon_\theta^{(j)} \epsilon_{\theta\varphi}^{(j)} \gamma_\theta^{(j)} \mid \kappa_\varphi^{(j)} \kappa_{\varphi\theta}^{(j)} \mid \kappa_\theta^{(j)} \kappa_{\theta\varphi}^{(j)}\} \tag{5c}$$

In Eq. 5, E_φ, E_θ and $E_{\theta\varphi}$ = the meridional, circumferential and in-plane shearing strains; Γ_φ and Γ_θ = the meridional and circumferential transverse shearing strains and K_φ, K_θ and $K_{\theta\varphi}$ = the meridional, circumferential and twisting changes in curvature. It is assumed that $E_{\varphi\theta} = E_{\theta\varphi}$ and $K_{\varphi\theta} = K_{\theta\varphi}$.

The stress resultants and moment resultants are taken as

$$\{T\} = \{T_f \mid T_c\} = \{N_\varphi \; N_{\varphi\theta} \; Q_\varphi \; N_\theta \; N_{\theta\varphi} \; Q_\theta \mid M_\varphi \; M_{\varphi\theta} \; M_\theta \; M_{\theta\varphi}\}$$

$$= \sum_{j=o}^{\infty} [\Theta_2] \{T^{(j)}(s)\} \tag{6a}$$

in which

$$\{T^{(j)}(s)\} = \{T_{f\varphi}^{(j)} \mid T_{f\theta}^{(j)} \mid T_{c\varphi}^{(j)} \mid T_{c\theta}^{(j)}\}$$

$$= \{n_\varphi^{(j)} n_{\varphi\theta}^{(j)} q_\varphi^{(j)} \mid n_\theta^{(j)} n_{\theta\varphi}^{(j)} q_\theta^{(j)} \mid m_\varphi^{(j)} m_{\varphi\theta}^{(j)} \mid m_\theta^{(j)} m_{\theta\varphi}^{(j)}\} \tag{6b}$$

In Eq. 6, N_φ, N_θ and $N_{\theta\varphi}$ = the meridional, circumferential and in-plane shearing stress resultants; Q_φ and Q_θ = the meridional and circumferential transverse shearing stress resultants; and M_φ, M_θ and $M_{\theta\varphi}$ = the meridional, circumferential and twisting moment resultants, positive as shown in Fig. 2. It is assumed that $N_{\varphi\theta} = N_{\theta\varphi}$ and $M_{\varphi\theta} = M_{\theta\varphi}$.

The surface loading is taken as

$$\{F\} = \{F_\varphi \; F_\theta \; F_n\} = \sum_{j=o}^{\infty} [\Theta_3] \{F^{(j)}_{(s)}\} \tag{7a}$$

in which

$$[\Theta_3] = [\cos j\theta \; \sin j\theta \; \cos j\theta] \tag{7b}$$

and

$$\{F^{(j)}_{(s)}\} = \{f^{(j)}_\varphi \; f^{(j)}_\theta \; f^{(j)}_n\} \tag{7c}$$

In Eq. 7, F_φ, F_θ and F_n = the distributed surface loading corresponding to the positive directions of U, V and W, respectively, as shown in Fig. 2. Also admissible are circumferentially distributed line loads applied at nodal circles, $R = R_i$, in the form

$$\{F_i\} = \{F_{\varphi i} F_{\theta i} F_{ni}\} = \sum_{j=o}^{\infty} [\Theta_3] \{F_i^{(j)}\} \tag{8a}$$

in which

$$\{F_i^{(j)}\} = \{f^{(j)}_{\varphi i} \; f^{(j)}_{\theta i} \; f^{(j)}_{ni}\} \tag{8b}$$

No applied couples are considered.

With all dependent variables and loading terms defined in separated form, the remaining derivation will proceed for a typical harmonic $j \geq 1$. For harmonic $j = 0$, the treatment is somewhat simplified since all terms dependent on $\sin j\theta$ are dropped.

Variational Formulation

Reissner's general variational theorem specialized for thin shell theory (16) serves as a basis for the mixed method finite element formulation envisaged in the present study. The theorem states that with the translational and rotational strain components satisfying the appropriate strain-displacement relationships, the governing equations of thin shell theory can be derived as the Euler equations of the variational problem. Prato (9) obtained a contracted form of the general variational principle by identically satisfying the stress resultant-strain relationships

$$\{T_f\} = f_1(H_f) \tag{9}$$

the moment equilibrium equations

$$\{Q_\varphi \; Q_\theta\} = f_2(T_c) \tag{10}$$

and the boundary conditions

$$\{T_c\}_b = \{\bar{T}_c\}_b \quad \text{on stress boundary } \sigma \tag{11a}$$

and

$$\{D_f\}_b = \{\bar{D}_f\}_b \quad \text{on displacement boundary } \delta \tag{11b}$$

In Eq. 11, the subscript b indicates the boundary value and the bar indicates a prescribed value of the variable.

For a typical harmonic $j \geq 1$ of an isotropic thin elastic shell of revolution subdivided into n discrete elements, the contracted functional takes the form

$$I^{(j)}(s) = \pi \sum_{i=1}^{n} \left[\left\{ \int_0^1 [W_M^{(j)} - W_B^{(j)} + U^{(j)} + q_\varphi^{(j)}(\gamma_\varphi^{(j)} - \beta_\varphi^{(j)}) + q_\theta^{(j)}(\gamma_\theta^{(j)} - \beta_\theta^{(j)})] R L_i ds \right.\right.$$

$$\left.\left. + \{F_i^{(j)}\}^T \{D_f^{(j)}\} R_i \right\} - \{\bar{T}_{f\varphi}^{(j)}\}_b^T \{D_f^{(j)}\}_b \bar{R}_\sigma + \{T_{c\varphi}^{(j)}\}_b^T \{\bar{D}_c^{(j)}\}_b \bar{R}_\delta \right] \tag{12a}$$

in which

$$W_M^{(j)} = \frac{Eh}{2(1-\mu^2)} [\epsilon_\varphi^{(j)2} + \epsilon_\theta^{(j)2} + 2\mu \epsilon_\varphi^{(j)} \epsilon_\theta^{(j)} + 2(1-\mu) \epsilon_{\theta\varphi}^{(j)2}] \tag{12b}$$

= Membrane strain energy density.

$$W_B^{(j)} = \frac{6}{Eh^3} [m_\varphi^{(j)2} + m_\theta^{(j)2} - 2\mu m_\varphi^{(j)} m_\theta^{(j)} + 2(1+\mu) m_{\theta\varphi}^{(j)2}]$$

$$+ \frac{1+\mu}{\lambda Eh} [q_\varphi^{(j)2} + q_\theta^{(j)2}] \tag{12c}$$

= Complementary strain energy density associated with bending and transverse shear deformation.

$$U^{(j)} = \{F^{(j)}(s)\}^T \{D_f^{(j)}\} \tag{12d}$$

= Potential energy density of the applied distributed loading

and

$\{\bar{T}_{f\varphi}^{(j)}\}_b$, $\{\bar{D}_c^{(j)}\}_b$ = Prescribed values of the force stress resultants and rotations at the boundaries where the horizontal radius $R(Z) = \bar{R}_\sigma$ and $\bar{R}_\delta$, respectively.

Also in Eqs. 12(b) and 12(c), h = the shell thickness, E = Young's modulus, μ = Poisson's ratio and λ = the shearing stress shape factor, commonly taken as 5/6.

The strain and rotation terms contained in Eq. 12 can be expressed in terms of the displacements using the strain-displacement relationships. For harmonic j of a rotational shell, these relationships take the form (17)

$$\{H^{(j)}(s)\} = [A^{(j)}(s)]\,\{D^{(j)}(s)\} \tag{13a}$$

in which

$$[A^{(j)}(s)] = \begin{bmatrix}
\frac{1}{L_i}\frac{d(\,)}{ds} & 0 & \frac{1}{R_\varphi} & 0 & 0 \\
-\frac{i}{2R} & \frac{1}{2L_i}\frac{d(\,)}{ds} - \frac{\cos\varphi}{2R} & 0 & 0 & 0 \\
-\frac{1}{R_\varphi} & 0 & \frac{1}{L_i}\frac{d(\,)}{ds} & 1 & 0 \\
\frac{\cos\varphi}{R} & \frac{i}{R} & \frac{\sin\varphi}{R} & 0 & 0 \\
-\frac{i}{2R} & \frac{1}{2L_i}\frac{d(\,)}{ds} - \frac{\cos\varphi}{2R} & 0 & 0 & 0 \\
0 & -\frac{\sin\varphi}{R} & -\frac{i}{R} & 0 & 1 \\
0 & 0 & 0 & \frac{1}{L_i}\frac{d(\,)}{ds} & 0 \\
0 & 0 & 0 & -\frac{i}{2R} & \frac{1}{2L_i}\frac{d(\,)}{ds} - \frac{\cos\varphi}{2R} \\
0 & 0 & 0 & \frac{\cos\varphi}{R} & \frac{i}{R} \\
0 & 0 & 0 & -\frac{i}{2R} & \frac{1}{2L_i}\frac{d(\,)}{ds} - \frac{\cos\varphi}{2R}
\end{bmatrix} \tag{13b}$$

As previously noted, the moment equilibrium equations, which take the form

$$q_\varphi^{(j)}(s) = \left[\frac{\cos\varphi}{R}\left(m_\varphi^{(j)} - m_\theta^{(j)}\right) + \frac{1}{L_i}\frac{dm_\varphi^{(j)}}{ds} + \frac{i}{R}\,m_{\theta\varphi}^{(j)}\right] \tag{14a}$$

$$q_\theta^{(j)}(s) = \left[\frac{2\cos\varphi}{R}\,m_{\theta\varphi}^{(j)} + \frac{1}{L_i}\,\frac{dm_{\theta\varphi}^{(j)}}{ds} - \frac{i}{R}\,m_\theta^{(j)}\right] \tag{14b}$$

for a rotational shell, are used to eliminate the transverse shear resultants from Eq. 12(a). Thus, the contracted form of Reissner's functional is expressible in terms of six dependent variables, three displacements ($u^{(j)}$, $v^{(j)}$, $w^{(j)}$), and three moment resultants ($m_\varphi^{(j)}$, $m_\theta^{(j)}$, $m_{\theta\varphi}^{(j)}$).

The solution for the dependent variables which will make the functional, Eq. 12(a), stationary will be assumed in the form of polynomial comparison functions. The functional contains only the first-order derivatives of the dependent variables; as such, it is sufficient that the functions have continuous first derivatives over the element domain and be continuous over the interelement boundaries (18,19). The natural boundary conditions provided by the last two terms of Eq. 12(a) require that the force stress resultants $\{\bar{T}_{f\varphi}^{(j)}\}_b$ and the rotations $\{\bar{D}_c^{(j)}\}_b$ be continuous across the interelement boundaries. This requirement, together with the continuity restrictions imposed on the dependent variables, cancel out these terms at all interior boundaries. Therefore, only the expressions along the exterior boundaries need be retained and any non-zero boundary values of the force stress resultants and/or the rotations may be specified through these terms. If the specified boundary quantities correspond to the explicit variables of the functional, the known value is directly used.

Approximations for Dependent Variables and Loading

The dependent variables are assumed in the following form over the element domain:

$$\{Y^{(j)}(s)\} = \{u^{(j)}\ v^{(j)}\ w^{(j)}\ m_\varphi^{(j)}\ m_\theta^{(j)}\ m_{\theta\varphi}^{(j)}\}$$

$$= \{y_1(s)\ y_2(s)\ y_3(s)\ y_4(s)\ y_5(s)\ y_6(s)\} \qquad (15a)$$

in which

$$y_\ell(s) = (1-s)y_\ell(o) + s\ y_\ell(1) + s(1-s)\sum_{m=1}^{\bar{\ell}} y_{\ell m}\ s^{m-1} \quad (\ell=1,6) \qquad (15b)$$

In Eq. 15, $y_\ell(s)$ = the ℓth element of vector $\{Y^{(j)}(s)\}$; the coefficients $y_\ell(o)$ and $y_\ell(1)$ = the nodal values of $y_\ell(s)$; $y_{\ell 1} \cdots y_{\ell\bar{\ell}}$ = coefficients of the higher-order terms; and $\bar{\ell}$ = indicator of the degree, $\bar{\ell}+1$, of each polynomial. It should be noted that the form of the polynomials specified in Eq. 15 allows a different order of approximation to be selected for each dependent variable, and that with the higher-order terms vanishing at the nodal circles, $s = 0$ and $s = 1$, the continuity of the variables across the nodes is maintained.

The load potential $U^{(j)}$ given in Eq. 12(d) is computed in a consistent manner for any arbitrary distribution of the surface loading (13,15). In the subsequent numerical analysis, the loading is assumed to be of the following form:

$$\{F^{(j)}(s)\} = \{f_1\ f_2\ f_3\} \qquad (16a)$$

in which

$$f_\ell(s) = (1-s)\ f_\ell(o) + s\ f_\ell(1) \qquad (\ell = 1,3) \qquad (16b)$$

and $f_\ell(o)$ and $f_\ell(1)$ represent the nodal values of the distributed loading.

Approximate Solution of System

The polynomials representing $u^{(j)}$, $v^{(j)}$ and $w^{(j)}$ given in Eq. 15(b) are substituted into Eq. 13 to get the strain-displacement relationships in terms of the comparison functions. Next, Eqs. 13 and 14 together with Eqs. 15(b) and 16(b) are substituted into Eq. 12. Then, after interchanging the order of variation and summation, the stationary condition for $I^{(j)}$, $\delta I^{(j)} = 0$, reduces to the set of algebraic equations

$$\sum_{i=1}^{n} \int_o^1 \frac{\partial I^{(j)}}{\partial y_\ell(o)} RL_i ds = 0 \qquad (\ell = 1,6) \tag{17}$$

$$\sum_{i=1}^{n} \int_o^1 \frac{\partial I^{(j)}}{\partial y_\ell(1)} RL_i ds = 0 \qquad (\ell = 1,6) \tag{18}$$

and

$$\sum_{i=1}^{n} \int_o^1 \frac{\partial I}{\partial y_{\ell m}} RL_i ds = 0 \qquad \begin{matrix}(\ell = 1,6)\\ (m = 1,\bar{\imath})\end{matrix} \tag{19}$$

Equations 17-19 may be combined into

$$\sum_{i=1}^{n} \left([M_i^{(j)}] \{\hat{Y}_i^{(j)}\} = \{\hat{F}_i^{(j)}\} + \{\tilde{F}_i^{(j)}\} \right) \tag{20a}$$

in which

$$\{\hat{Y}_i^{(j)}\} = \{y_1(o)\ y_2(o) \ldots y_6(o)\ y_1(1)\ y_2(1) \ldots y_6(1) | y_{11} \ldots y_{1\bar{1}} \ldots\ldots y_{6\bar{6}}\}$$

$$= \{\hat{Y}_{i1}^{(j)} | \hat{Y}_{i2}^{(j)}\} \tag{20b}$$

= the vector of nodal variables and coefficients of higher-order terms

$$\{\hat{F}_i^{(j)}\} = \{\hat{f}_1(o)\ \hat{f}_2(o)\ \hat{f}_3(o)\ o\ o\ o\ \hat{f}_1(1)\ \hat{f}_2(1)\ \hat{f}_3(1)\ o\ o\ o\ |$$

$$\hat{f}_{11} \ldots \hat{f}_{1\bar{1}}\ \hat{f}_{21} \ldots \hat{f}_{2\bar{2}}\ \hat{f}_{31} \ldots \hat{f}_{3\bar{3}}\ 0\ 0\ 0\}$$

$$= \{\hat{F}_{i1}^{(j)} | \hat{F}_{i2}^{(j)}\} \tag{20c}$$

= the external distributed load vector corresponding to $\{\hat{Y}_i^{(j)}\}$

$$\{\tilde{F}_i^{(j)}\} = \{F_i^{(j)} \circ\circ\circ F_{i+1}^{(j)} \circ\circ\circ \mid 0\} = \{\tilde{F}_{i1}^{(j)} \mid 0\} \tag{20d}$$

= the external line load vector corresponding to $\{\hat{Y}^{(j)}\}$

and

$[M_i^{(j)}]$ = the coefficient matrix which will be termed the <u>mixed</u> matrix.

In Eqs. 20(c,d), the large 0 represents a null vector of suitable dimension.

It is convenient to apply static condensation to Eq. 20(a) to reduce the number of unknowns (20). Considering any element i and partitioning the mixed matrix to conform with Eq. 20(b),

$$[M_i^{(j)}] = \left[\begin{array}{c|c} M_{i11}^{(j)} & M_{i12}^{(j)} \\ \hline M_{i21}^{(j)} & M_{i22}^{(j)} \end{array}\right] \tag{21a}$$

and solving for $\hat{Y}_{i2}^{(j)}$ in terms of $Y_{i1}^{(j)}$,

$$\{\hat{Y}_{i2}^{(j)}\} = [M_{i22}^{(j)}]^{-1} \{\hat{F}_{i2}^{(j)} - M_{i21}^{(j)} \hat{Y}_{i1}^{(j)}\} \tag{21b}$$

Eq. 20(a) reduces to

$$\sum_{i=1}^{n} \left([\hat{M}_i^{(j)}] \{\hat{Y}_{i1}^{(j)}\} = \{\hat{F}_{ei}^{(j)}\} \right) \tag{22a}$$

in which

$$[\hat{M}_i^{(j)}] = [M_{i11}^{(j)} - M_{i12}^{(j)} M_{i22}^{(j)-1} M_{i21}^{(j)}] \tag{22b}$$

= the element mixed matrix

and

$$\{\hat{F}_{ei}^{(j)}\} = \{\tilde{F}_{i1}^{(j)} + \hat{F}_{i1}^{(j)} - M_{i12}^{(j)} M_{i22}^{(j)-1} \hat{F}_{i2}^{(j)}\} \tag{22c}$$

= the vector giving the contribution of the loads acting on element i to the equivalent nodal forces at nodes i and i+1.

If the meridian of the shell is continuous, the variables $y_\ell(1)$ for element i are equal to the variables $y_\ell(o)$ for element i+1 and Eq. 22(a) may be directly assembled into a symmetric banded global set of equations

$$[\hat{M}^{(j)}]\ \{\hat{Y}_1^{(j)}\} = \{\hat{F}_e^{(j)}\} \tag{23}$$

If meridian discontinuities occur, appropriate coordinate transformations are required (13). The homogeneous boundary conditions for the displacements and the moment resultants are applied by eliminating the specified variables along with the corresponding rows and columns of the global mixed matrix $[\hat{M}^{(j)}]$ and rows of the load vector $\{\hat{F}_e^{(j)}\}$.

The preceding solution format is identical to the familiar direct stiffness method; however, the resulting system of equations differs in one respect. The global mixed matrix is not positive definite and the popular Choleski square root method cannot be used. Gaussian elimination was used for the numerical work in this study.

Once the global vector $\{\hat{Y}_1^{(j)}\}$ is determined, the local vectors $\{\hat{Y}_{i1}^{(j)}\}$ can be easily computed and then the higher-order coefficients $\{Y_{i2}^{(j)}\}$ found from Eq. 21(b). Internodal values of the displacements and moment resultants are calculated from Eq. 15(b). After a single differentiation of the comparison functions, the transverse shear stress resultants are determined from Eqs. 14(a) and 14(b) and the in-plane stress resultants from Eqs. 13 and the constitutive law.

Precision of Approximate Solution

The polynomial comparison functions for the dependent variables, Eqs. 15, may be truncated at any desirable value of ℓ. The algebraic simplicity of low-order expansions is very attractive and previous studies using mixed-type triangular shell elements have shown that solutions employing linear displacement functions (7,9) together with constant (7) and linear (9) moment functions both converge with a sufficient number of elements. When higher-order approximations were used with the triangular mixed-type elements, improved convergence was obtained at the expense of a more complicated derivation and an increased number of degrees of freedom (10,11). In the present rotational shell formulation, however, the high-order terms are incorporated with little added complexity and no increase in the size of the element mixed matrix.

Since the moment and force resultants have a derivative relationship to the displacements, these quantities usually exhibit poorer convergence than the displacements. The use of comparatively fewer high-order elements in place of a larger number of low-order elements will improve the convergence of these important design variables and also reduce the required computer time and storage. In the following example problems, two measures of convergence will be studied: i) increasing the number of elements for a constant order of approximation; and, ii) increasing the order of approximation for a constant number of elements.

Another item of interest which will be examined subsequently is the comparative merits of the mixed-type rotational shell element as compared to a displacement-type element. An available high-precision rotational shell element (14,15) is quite similar in many respects to the present element and provides a basis to evaluate the relative performance of the two types of elements for a selected group of problems.

EXAMPLES

Computer Program

A Fortran IV Computer program was written to analyze the example problems. The average running time on an IBM 360-50 computer (fast core) was 4-8 secs. per element per harmonic for quartic comparison functions with a reduction of approximately 1 sec. per element when quadratic expansions were used. The problems were also solved using the displacement formulation outlined in Ref. 15 and the time required was about the same as in the case of the mixed method for the same order of expansions of the variables.

Edge-Loaded Cylindrical Shell

Figs. 3 and 4 show results obtained from the analysis of an edge-loaded cylindrical shell, a familiar test problem in the finite element literature (22,14). The convergence of the meridional moment and transverse shear resultants are examined in Fig. 3 as the degree of the comparison function is increased, and in Fig. 4 as the number of element divisions is increased. Although the displacements, not shown, converge with the case 1 element breakdown using quadratic polynomials, Fig. 3 indicates that the meridional moment and transverse shear require a cubic approximation to give accurate results. The discontinuity in the computed shear force resultant at the first node point away from the free edge (Z = 0.25 in) is shown in Fig. 3. This discontinuity is quite large for quadratic approximation but is practically eliminated with quartic expansion. With an increased number of elements near the loaded edge, Fig. 4 shows that quadratic approximations are adequate to achieve convergence on the meridional moment and transverse shear, as well as the normal displacement.

Results obtained from a refined displacement-type element (14) using the case 1 element subdivision are given in Table 1 and it is seen that with cubic and quartic approximations, the mixed method solutions exhibit improved convergence over corresponding displacement method solutions.

The present problem has also been studied with respect to different order of approximations for displacements and moments and indications are that uniform expansions for all of the variables provide superior convergence.

Hydrostatically-Loaded Cylindrical Shell

A hydrostatically-loaded cylindrical shell is examined in Fig. 5 where 14 evenly spaced elements with increasing order of approximation are employed in the analysis. In Table 2, the results are compared with the displacement method solution (15) for various order of polynomial expansions. It may be observed that even the linear functions give good results for the meridional moment using the mixed-type element; however, the transverse shear is somewhat unsatisfactory since this quantity is constant over each element domain for a cylindrical shell if the meridional moment is linear, as indicated by Eq. 14(a). The quadratic approximation gives very good convergence for the moment and the shear. For the range of expansions considered in Table 2, the displacement method solutions, in general, exhibited less satisfactory convergence than the mixed formulation and the nodal discontinuities of the moment and transverse shear force were more pronounced.

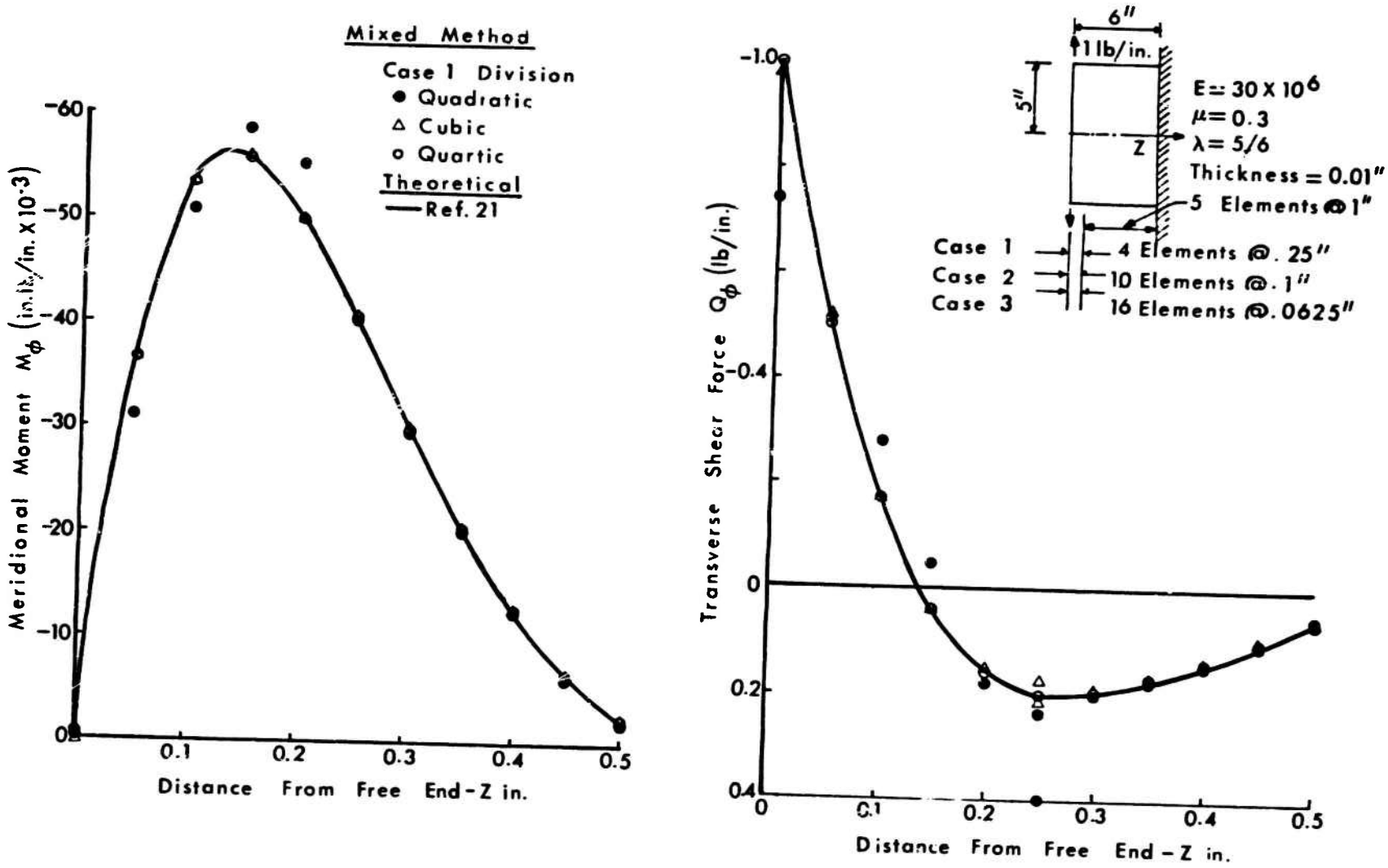


FIGURE 3. EDGE-LOADED CYLINDRICAL SHELL

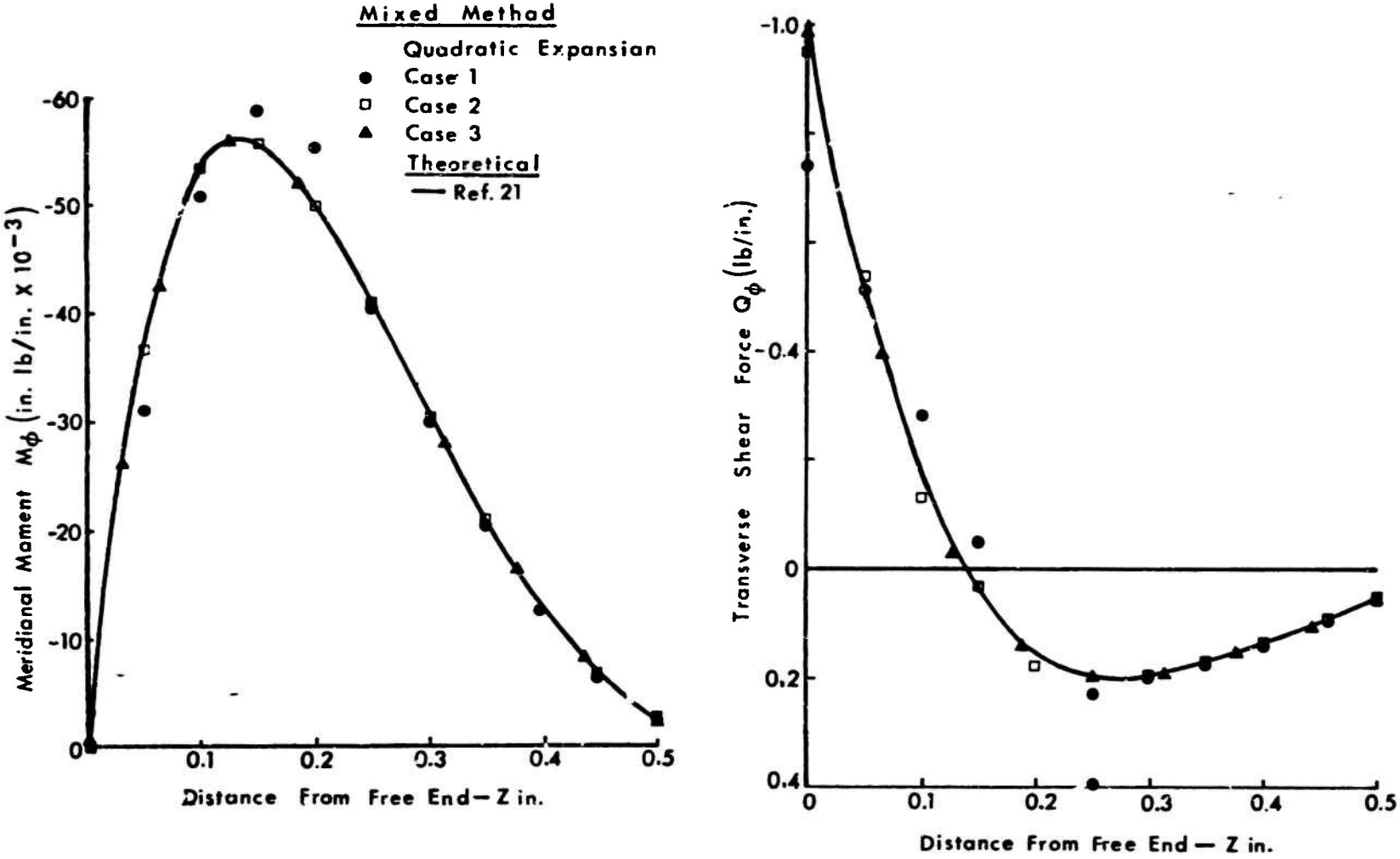


FIGURE 4. EDGE-LOADED CYLINDRICAL SHELL

TABLE 1. Edge Loaded Cylindrical Shell
(Case 1 Element Division)

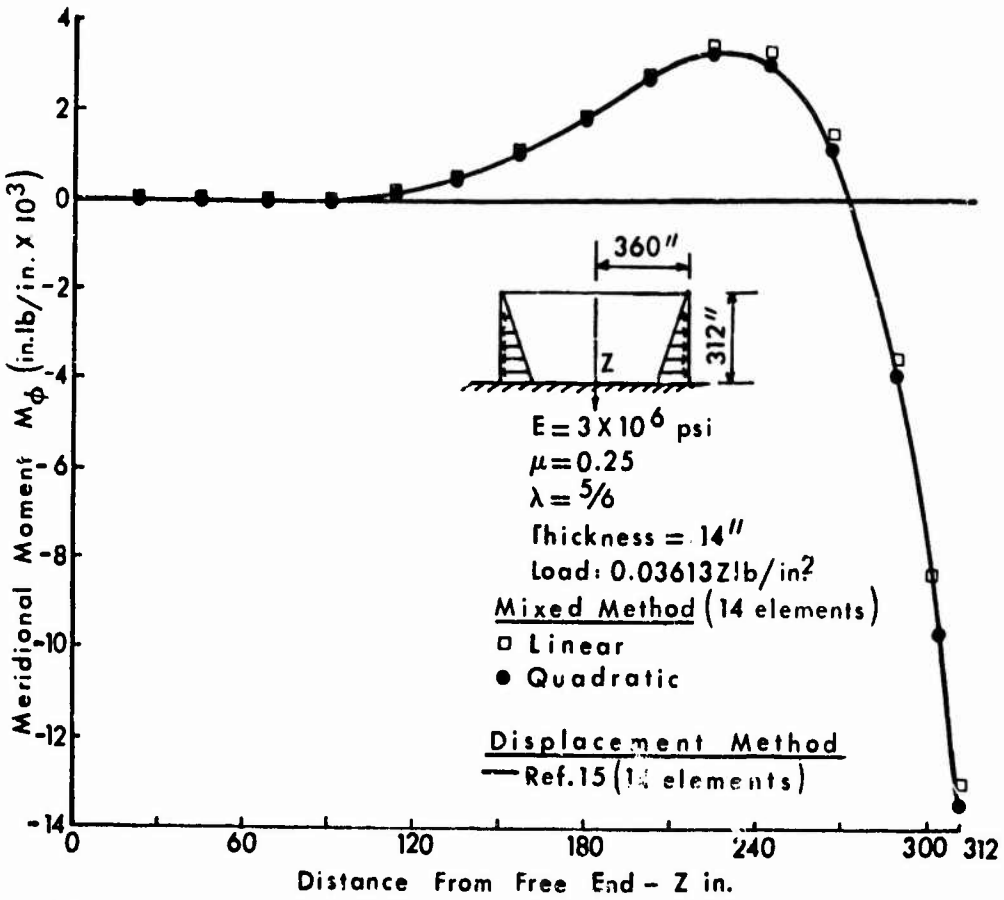
Dist. Z (in)	Meridional Moment M_φ (in.lb/in x 10^{-3})					Transverse Shear Force Q_φ (lb/in)				
	Theoret. Results (Ref.21)	Displ. Finite Elem.[a] Order of Expansion		Mixed Finite Elem. Order of Expansion		Theoret. Results (Ref.21)	Displ. Finite Elem.[a] Order of Expansion		Mixed Finite Elem. Order of Expansion	
		Cubic	Quartic	Cubic	Quartic		Cubic	Quartic	Cubic	Quartic
0.0	0.00	-28.50	- 5.82	0.00	0.00	-1.00	-2.24	-1.38	-0.97	-1.00
0.1	-53.23	-42.21	-51.84	-53.63	-53.22	-0.17	0.18	-0.25	-0.18	-0.17
0.2	-50.28	-53.14	-51.91	-50.06	-50.27	0.16	-0.28	0.32	0.17	0.16
0.3	-30.64	-29.98	-30.43	-30.55	-30.64	0.20	0.14	0.21	0.20	0.20
0.4	-13.02	-14.25	-13.15	-13.12	-13.01	0.14	0.19	0.14	0.15	0.14
0.5	- 2.59	- 1.09	- 2.44	- 2.61	- 2.61	0.07	0.14	0.07	0.07	0.07
0.6	1.67	1.05	1.59	1.67	1.66	0.02	0.03	0.03	0.02	0.02
0.7	2.40	2.58	2.49	2.37	2.39	-0.00	-0.12	-0.01	-0.00	-0.00
0.8	1.74	1.69	1.73	1.72	1.74	-0.01	-0.01	-0.01	-0.01	-0.01
0.9	0.88	0.93	0.92	0.87	0.88	-0.01	-0.01	-0.01	-0.01	-0.01
1.0	0.28	0.17	0.29	0.27	0.29	-0.00	-0.00	-0.00	-0.00	-0.00

a Results as per Ref. 15.

TABLE 2. Cylindrical Shell Under Hydrostatic Loading
(14 Uniformly Spaced Elements)

Dist. Z (in)	Meridional Moment M_φ (in.lb/in x 10^{3})					Transverse Shear Force Q_φ (lb/in x 10^{2})				
	Results as per Ref.15[b]	Displ. Finite Elem.[c] Order of Expansion		Mixed Finite Elem. Order of Expansion		Results as per Ref.15[b]	Displ. Finite Elem.[c] Order of Expansion		Mixed Finite Elem. Order of Expansion	
		Quadratic	Cubic	Quadratic	Cubic		Quadratic	Cubic	Quadratic	Cubic
0.0	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00
89.1	- 0.04	- 0.05	- 0.05	- 0.04	- 0.04	0.04	0.07	0.04	0.04	0.04
133.7	0.46	0.45	0.46	0.46	0.46	0.20	0.34	0.30	0.20	0.20
178.2	1.80	1.81	1.80	1.80	1.80	0.39	0.65	0.39	0.40	0.39
200.5	2.68	2.71	2.69	2.68	2.68	0.38	0.63	0.39	0.39	0.38
222.8	3.34	3.40	3.35	3.34	3.34	0.17	0.27	0.18	0.20	0.17
245.1	3.17	3.27	3.19	3.17	3.17	-0.39	-0.68	-0.37	-0.35	-0.39
267.4	1.20	1.36	1.22	1.20	1.20	-1.47	-2.51	-1.45	-1.42	-1.47
289.7	- 3.87	- 3.68	- 3.84	- 3.87	- 3.87	-3.19	-5.41	-3.16	-3.14	-3.19
312.0	-13.47	-12.43	-13.43	-13.47	-13.47	-5.49	-8.43	-5.67	-5.46	-5.50

b With sixth order expansions; c Results as per Ref. 15.



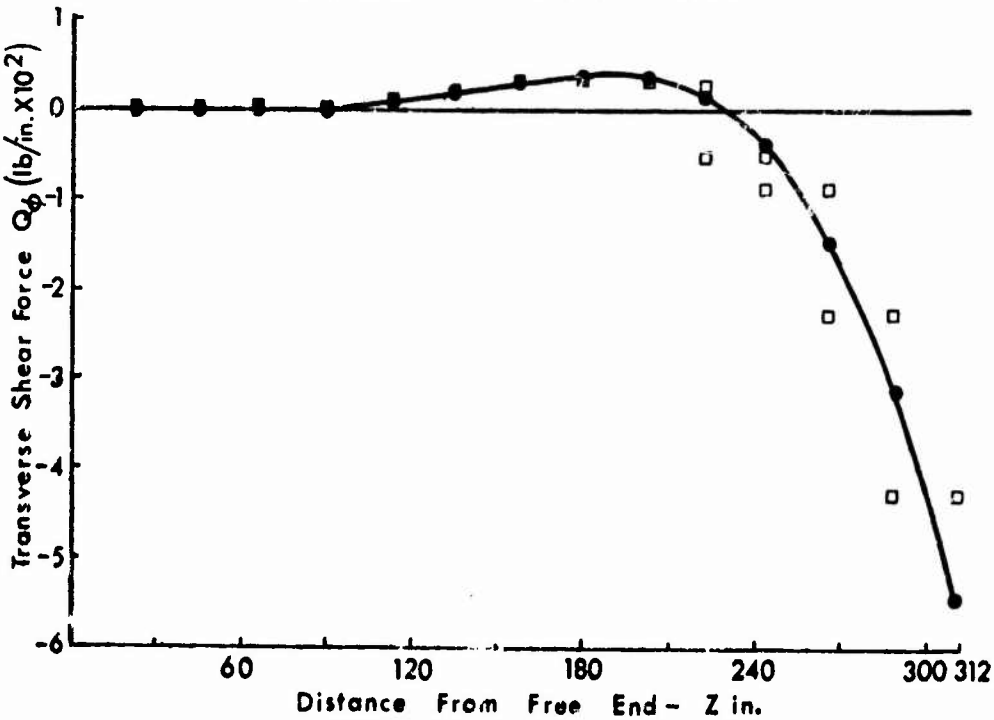


FIGURE 5 CYLINDRICAL SHELL UNDER HYDROSTATIC LOADING

This shell is moderately thick and the effect of shearing deformation is noticeable as the transverse shear and the meridional moment at the fixed end are approximately 3-4% lower than the corresponding values when the effect of shearing deformation is neglected (21).

Hyperboloidal Shell

A doubly curved shell of hyperboloidal shape subjected to an asymmetric loading is treated in Figs. 6 and 7. The shell is assumed to be fixed at the base. A quasistatic wind loading which is represented by the first 10 harmonics of the Fourier cosine expansion is considered (23). The meridional force and moment resultants obtained from the present formulation are compared with the results calculated by a displacement method solution with 200 conical-type elements (24). In Fig. 6, where 10 equally spaced elements are used (case 1 division), quartic expansions are required to find accurate moments near the base. For quadratic approximations of the variables, a 10 element idealization failed to predict the steep moment gradient in the vicinity of the base and a subsequent study with 20 equally spaced elements revealed that the computed moment at the base was 20% less than the actual value. However, when the number of elements near the base is increased (case 2 division), accurate results are obtained with quadratic approximations except for small inaccuracies in the meridional force near the free edge.

Results from a 10 element (case 1 division) displacement method solution (15) using quartic expansions are given in Table 3. Comparing these values with those obtained from Ref. 24, the meridional moments calculated using the displacement method are significantly less than the actual values near the base, while the moments found from the mixed solution are quite accurate. Both solutions yield comparatively accurate values for the meridional force.

SUMMARY AND CONCLUSIONS

A precise, mixed-type finite element analysis suitable for static analysis of shells of revolution has been presented. The formulation is based on a contracted form of Reissner's variational principle and includes the effect of transverse shear deformations. The shell element is in the form of a frustum of the meridional curve and fully incorporates the actual geometric data. The dependent variables in the formulation, the displacements and the moment resultants, have been represented by polynomial comparison functions which are continuous across the interelement boundaries and have continuous derivatives over the individual element domain. The continuity of the functions has been enforced by choosing the coefficients of the first-order terms to correspond to the nodal values of the variables and by forcing the higher-order terms to vanish at the nodal circles. The stationary condition of the functional yields a set of algebraic equations and static condensation permits the global mixed matrix to be expressed in terms of the nodal values of the variables alone, regardless of the order of the comparison functions.

The results of a number of test problems indicate that convergence can be achieved both by reducing the element size and by increasing the order of the comparison functions. However, from a computational standpoint, the use of high-order functions with fewer elements appears to be more efficient.

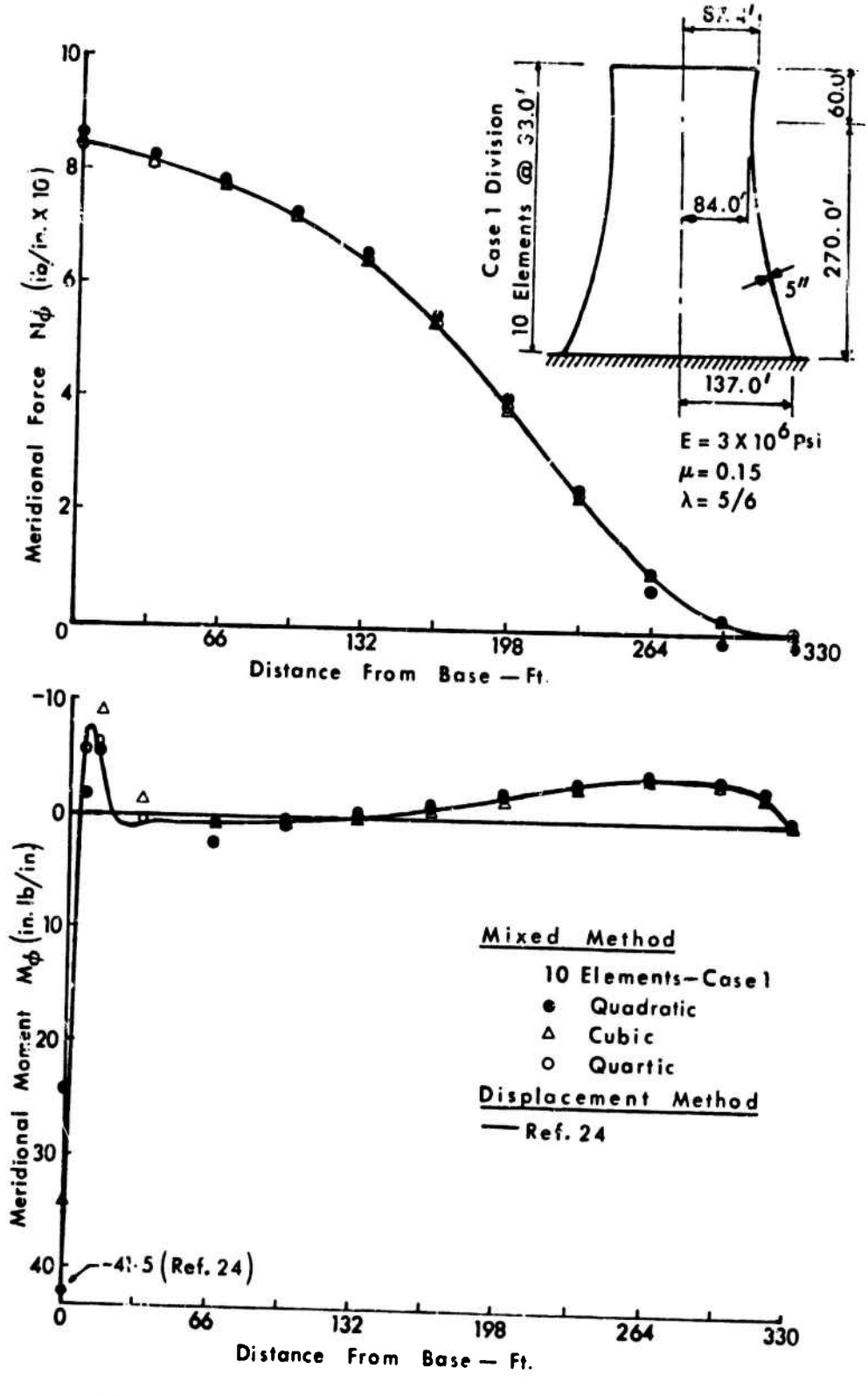


FIGURE 6 HYPERBOLOIDAL SHELL UNDER WIND LOAD

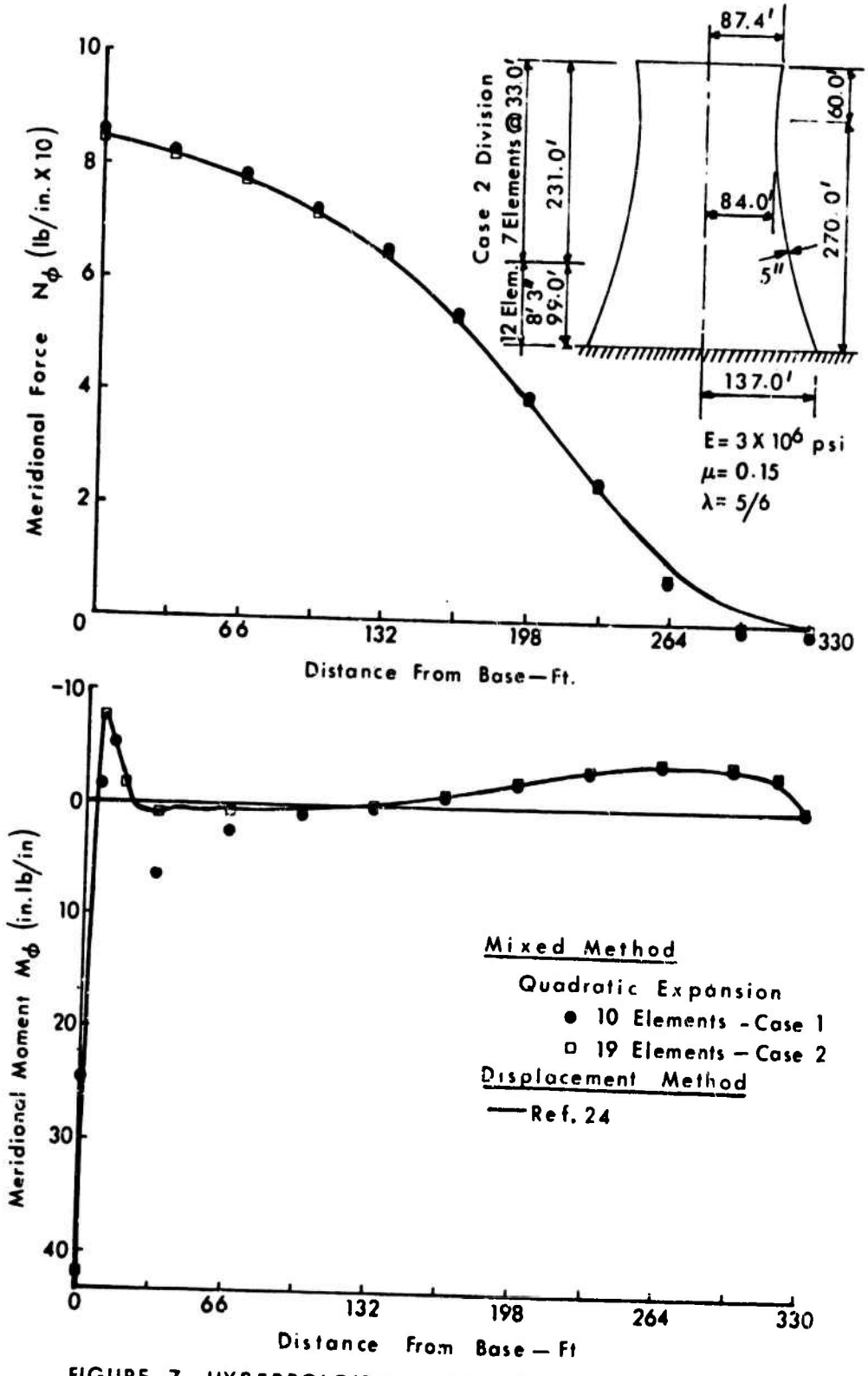


FIGURE 7 HYPERBOLOIDAL SHELL UNDER WIND LOAD

TABLE 3. Hyperboloidal Shell Under Wind Loading
(Quartic Expansion, 10 Elements)

Dist. from base (ft)	Meridional Moment M_φ (in. lb/in.)			Meridional Force N_φ (lb/in.)		
	Ref. 24[d]	Displ. Finite Elem.	Mixed Finite Elem.	Ref. 24[d]	Displ. Finite Elem.	Mixed Finite Elem.
0.0	41.50	18.02	41.98	84.30	84.22	84.35
6.6	- 6.20	3.22	- 5.78	83.60	84.02	83.91
13.2	- 5.40	- 4.52	- 6.22	83.30	83.08	83.15
19.8	- 0.30	- 5.40	1.46	82.60	82.51	82.58
26.4	0.90	0.36	1.26	82.00	82.25	82.15
33.0	0.60	5.39	0.30	81.00	81.25	81.20
66.0	0.60	0.46	0.64	77.30	77.78	77.68
99.0	0.40	0.43	0.43	72.60	72.45	72.32
132.0	0.00	- 0.08	- 0.08	64.40	64.73	64.51
165.0	- 0.90	- 0.86	- 0.88	53.50	53.76	53.42
198.0	- 2.00	- 1.92	- 1.98	39.10	39.62	39.08
231.0	- 3.50	- 3.17	- 3.28	24.10	24.29	23.49
264.0	- 4.40	- 4.10	- 4.11	11.10	11.18	10.28
297.0	- 4.40	- 3.62	- 3.93	2.80	2.92	2.34
330.0	0.00	0.05	0.00	0.00	0.00	0.02

d Scaled values

Also, the results from the mixed method analysis were compared with those obtained from a displacement method solution using a similar type comparison function. It was observed that the mixed method solution exhibited improved accuracy over the displacement method solution for the same number of elements and the same order of expansion.

ACKNOWLEDGMENT

The cooperation of the Washington University Computation Center, the financial support of the National Science Foundation under grant GK-19779, and the diligent computational assistance provided by Mr. Ed Dolata are gratefully acknowledged.

REFERENCES

1. Hellinger, E., "Die allgemeine Ansatz der Mechanik der Kontinua,"Art. 30, Encyklopädie Mathematischen Wissenschaften, mit Einschlus ihrer Anwendungen, Vol. 4, Part 4, Mechanik, 1914, pp. 601-694.

2. Reissner, E., "On a Variational Theorem in Elasticity," Journal of Mathematics and Physics, Vol. 29, 1950, pp. 90-95.

3. Herrmann, L. R., "A Bending Analysis for Plates," Proc. Conference on Matrix Methods in Structural Mechanics, AFFDL-TR-66-80, Wright Patterson Air Force Base, Ohio, 1965, pp. 577-604.

4. Herrmann, L. R., "Finite Element Bending Analysis for Plates," Journal of the Engineering Mechanics Division, ASCE, Vol. 93, No. EM5, Oct. 1967, pp. 13-26.

5. Visser, W., "A Refined Mixed-Type Plate Bending Element," AIAA Journal, Vol. 7, No. 9, Sept. 1969, pp. 1801-1803.

6. Cook, R. D., "Eigenvalue Problems with a'Mixed' Plate Element," AIAA Journal, Vol. 7, No. 5, May 1969, pp. 982-983.

7. Herrmann, L. R. and Campbell, D. M., "A Finite-Element Analysis for Thin Shells," AIAA Journal, Vol. 6, No. 10, Oct. 1968, pp. 1842-1847.

8. Prato, C. A., "A Mixed Finite Element Method for Thin Shell Analysis," Research Report R68-33, Dept. of Civil Engineering, Massachusetts Institute of Technology, 1968.

9. Prato, C. A., "Shell Finite Element Method Via Reissner's Principle," Int. J. Solids and Structures, Vol. 5, 1969, pp. 1119-1133.

10. Visser, W., "The Application of a Curved Mixed-Type Shell Element," IUTAM Symposium on High Speed Computing of Elastic Structures, Liege, Belgium, Aug. 1970.

11. Connor, J. and Will, D., "A Mixed Finite Element Shallow Shell Formulation," Proc. of the U.S.-Japan Seminar on Matrix Methods in Structural Analysis and Design, Tokyo, Sept. 1969, pp. 105-137.

12. Elias, Z. M., "Mixed Finite Element Method for Axisymmetric Shells," to be published in the International Journal for Numerical Methods in Engineering.

13. Brombolich, L. J. and Gould, P. L., "Finite Element Analysis of Shells of Revolution by Minimization of the Potential Energy Functional," Proc. of the Symposium on Application of Finite Element Methods in Civil Engineering, Vanderbilt University, Nashville, Tenn., 1969, pp. 279-307.

14. Brombolich, L. J. and Gould, P. L., "A High-Precision Curved Shell Finite Element," Proc. 12th Structures, Structural Dynamics and Materials Conf., AIAA and ASME, Anaheim, Calif., April, 1971.

15. Brombolich, L. J. and Gould, P. L., "High-Precision Finite Element Analysis of Shells of Revolution," Research Report No. 16, Structural Division, Dept. of Civil & Environmental Engineering, Washington University, St. Louis, Mo. 1970.

16. Reissner, E., "Variational Considerations for Elastic Beams and Shells," Journal of the Engineering Mechanics Division, ASCE, Vol. 88, No. EM1, Part I, Feb. 1962, pp. 23-57.

17. Washizu, K., "Variational Methods in Elasticity and Plasticity," Pergamon Press, 1968.

18. Strickland, J. A., "Geometrically Nonlinear Static and Dynamic Analysis of Shells of Revolution," IUTAM Symposium on High Speed Computing of Elastic Structures, Liege, Belgium, Aug. 1970.

19. Courant, R. and Hilbert, D., "Methods of Mathematical Physice," Vol. 1, Interscience, New York, 1953.

20. Przemieniecki, J. S., "Theory of Matrix Structural Analysis," McGraw-Hill Book Co., New York, 1968.

21. Timoshenko, S. and Woinowski-Krieger, S., Theory of Plates and Shells, 2nd Ed., McGraw-Hill Book Co., Inc., New York, 1959.

22. Grafton, P. E. and Strome, D. R. "Analysis of Axisymmetrical Shells by the Direct Stiffness Method," AIAA Journal, Vol. 1, No. 10, Oct. 1963, pp. 2342-2347.

23. Lee, S. L. and Gould, P. L., "Hyperbolic Cooling Towers under Wind Load," Journal of the Structural Division, ASCE, Vol. 93, No. ST5, Oct. 1967.

24. Hill, D. W. and Coffin, G. K., "Stresses and Deflections in Cooling Tower Shells Due to Wind Loading," Bulletin of the I.A.S.S., No. 35.

OPTIMIZATION OF FINITE ELEMENT SOLUTIONS

E.R. de Arantes e Oliveira (+)
Technical University of Lisbon

Abstract - The minimization of the discretization error corresponding to a given total number of nodes is considered for each one of the steps which may be considered in the application of the f.e.m. In what concerns the choice of the type of the element, the rule must be respected of the accuracy of the stresses being consistent with the accuracy of the displacements or tractions. In what concerns the topology of the mesh, it is shown that the best topology depends on the orientation of the mesh with respect to the iso-energetic lines. The problem of the best allocation of the nodes is also considered.

1 - INTRODUCTION

Convergence criteria have been studied in previous papers for sequences of approximate solutions generated by successive decompositions of a given domain into finite elements with decreasing size.

The present paper considers the problem of the maximization of the speed of convergence, or, more precisely, of which decisions must be made in order that the discretization error corresponding to a given total number of nodes be minimized.

Important decisions can be made in each one of the steps which may be considered in the application of the finite elements technique:

- Choice of the type of the element;
- Choice of the topology of the mesh;
- Allocation of the nodes.

(+) Professor of Civil Engineering.

The aim of the paper is to discuss some points in connexion with each one of the preceeding steps. Some simple rules resulting from such discussion will be tested by numerical examples.

2 - SUMMARY OF PRECEEDING RESULTS

Some concepts and results will be reminded in this Section which have already been presented elsewhere.[1]

The finite element technique can be used under four different modalities, according to the way in which the allowed fields are defined within each element and the inter-connexion between elements is ensured (see Table I).

TABLE I

definition of the allowed fields through / interconnexion by	displacements	stresses or strains
compatibility conditions	1st modality (classical)	2nd modality (hybrid models)
equilibrium conditions	3rd modality (equilibrium models)	

The first and second modalities are both connected with the so-called potential energy method. Both will thus be referred to, in the sequel, as potential energy approach. The third modality, connected with the complementary energy method, will be the complementary energy approach.

As it is well known, completeness with respect to a set $\mathcal{F}$, which contains the exact solution is a necessary condition for convergence in any modality. This means that, given a sequence of fields generated by successive decompositions of the domain into subdomains with decreasing

size, and an arbitrary field $f \in \mathcal{F}_0$, it must be possible to form a sequence $\{f_m\}$ of fields tending to f, by taking one field from each family.

Each field in the sequence is the so-called allowed field corresponding to f and to a given subdivision of the domain into finite elements. In the potential energy approach, the allowed field is the one which presents the same nodal displacements (generalized displacements) as f. In the complementary energy approach, it is the one which presents the same traction resultants and moments (generalized forces).

The notation used in [3] will be followed in the present paper.

3 - ORDER OF APROXIMATION

Consider a given subdivision of the domain Ω into subdomains Ω^e.

Let $\mathcal{F}$ denote the set of all the structural fields defined on Ω. Discontinuities of the stresses, strains and displacements are allowed on the subdomain boundaries but not within the subdomains.

As linearity is assumed, the distance between two fields belonging to $\mathcal{F}$ may be given by

$$d(f_1, f_2) = \sqrt{\sum_e U^e_{f_1 - f_2}} = \sqrt{\tfrac{1}{2} \sum_e \int_{\Omega^e} \underline{s}_1^T \underline{e}_2 \, d\Omega} = \sqrt{\tfrac{1}{2} \sum_e \int_{\Omega^e} \underline{s}_2^T \underline{e}_1 \, d\Omega} \qquad (1)$$

$\underline{s}_1$ and $\underline{s}_2$, $\underline{e}_1$ and $\underline{e}_2$, being the stress and strain vectors associated to f_1 and f_2.

In the [potential / complementary] energy approach, to each generating [compatible / equilibrated] field f, a unique generalized [compatible / equilibrated] fields f' corresponds which we call the allowed field corresponding to f. The set of all ordered pairs f-f' defines thus a linear bounded operator T, the domain and range of which are respectively the sets $\mathcal{F}$

and $\mathcal{F}'$ of the generating and generalized |compatible / equilibrated| fields. $\mathcal{S}$ and $\mathcal{S}'$ are subsets of $\mathcal{F}$. f' is the T-image of f and we write $f'=T(f)$.

The exact solution s'_a minimizes on $\mathcal{S}$ the total |potential / complementary| energy(*)

$$F=\sum_e\left(U^e-\int_{\Omega^e}\bar{f}^T u\,d\Omega-\int_{\Sigma^e\cap\Sigma_2}\bar{p}^T u\,d\Sigma\right) \quad \Big| \quad F=\sum_e\left(U^{*e}-\int_{\Sigma^e\cap\Sigma_2}p^T\bar{u}\,d\Sigma\right) \tag{2}$$

It is known that, in the linear case,

$$F(f)-F(s)=d^2(f,s) \tag{3}$$

where s is the exact solution and f is any generating |compatible / equilibrated| field.

An analogous expression

$$F(f')-F(s'_a)=d^2(f',s'_a) \tag{4}$$

is valid for the approximate solution s'_a and any generalized |compatible / equilibrated| field.

Let now s' be the T-image of s and assume that a generating |compatible / equilibrated| field s_a exists such that the approximate solution s'_a is the T-image of s_a (Fig. 3.1)

(*) The equation at |left / right| concerns the |potential / complementary| energy approach.

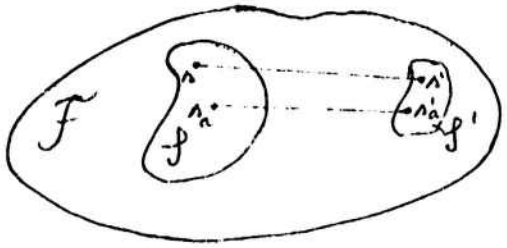


Fig. 3.1

By virtue of the total $\left|\begin{matrix}\text{potential}\\ \text{complementary}\end{matrix}\right|$ energy theorem,

$$F(\Lambda) \lessgtr F(\Lambda_a) \tag{5}$$

$$F(\Lambda'_a) \lessgtr F(\Lambda') \tag{6}$$

Let now

$$\delta F(\Lambda) = F(\Lambda) - F(\Lambda') \tag{7}$$

$$\delta F(\Lambda_a) = F(\Lambda_a) - F(\Lambda'_a) \tag{8}$$

Introducing in (5), there results

$$F'(\Lambda) + \delta F(\Lambda) - \delta F(\Lambda_a) \lessgtr F(\Lambda'_a) \tag{9}$$

and combining (6) and (8), we obtain

$$F'(\Lambda) + \delta F(\Lambda) - \delta F(\Lambda_a) \lessgtr F(\Lambda'_a) \lessgtr F(\Lambda') \tag{10}$$

Therefore

$$|F(\Lambda') - F(\Lambda'_a)| \lessgtr |\delta F(\Lambda)| + |\delta F(\Lambda_a)| \tag{11}$$

and, by virtue of (3),

$$d(\Lambda', \Lambda'_a) \lessgtr \sqrt{|\delta F(\Lambda)| + |\delta F(\Lambda_a)|} \tag{12}$$

Using now the triangular inequality, there results

$$d(\Lambda,\Lambda') \le d(\Lambda,\Lambda^*) + d(\Lambda',\Lambda^*) \le d(\Lambda,\Lambda^0) + \sqrt{|\delta\Gamma(\Lambda)| + |\delta\Gamma(\Lambda')|} \qquad (13)$$

Now,

$$d(\Lambda,\Lambda^*) = \sqrt{\tfrac{1}{2}\sum_e \int_{\Omega^e} \delta\underline{\varepsilon}^{*T} \underline{\underline{H}}\, \delta\underline{\varepsilon}^{*} d\Omega} = O(\delta\underline{\varepsilon}^{*}) = O(\delta\underline{\Lambda}^{*}) \qquad (14)$$

where $\delta\underline{\varepsilon}^{*}$ denotes the strain variation vector from Λ to Λ'

On the other hand, differentiating (2), we obtain

$$\delta F = \sum_e \int_{\Omega^e} \underline{\varepsilon}^{*T} \underline{\underline{H}}\, \delta\underline{\varepsilon}^{*} d\Omega - \int_{\Omega^e} \underline{f}^{T} \delta\underline{u} \, d\Omega - \int_{\Sigma-\Sigma_u} \underline{\bar{p}}^{T} \delta\underline{u} \, d\Sigma \;\Big|\; \delta F = \sum_e \left(\int_{\Omega^e} \underline{\varepsilon}^{*T} \underline{\underline{H}}\, \delta\underline{\varepsilon}^{*} d\Omega - \int_{\Sigma-\Sigma_u} \delta\underline{p}^{T} \underline{u} \, d\Sigma \right) \qquad (15)$$

and thus

$$\delta F = O(\delta\underline{\varepsilon}) + O(\delta\underline{u}) \quad \Big| \quad \delta F = O(\delta\underline{\Lambda}) + O(\delta\underline{p}) \qquad (16)$$

Introducing (14) in (13), there results

$$d(\Lambda,\Lambda') = O(\delta\underline{\varepsilon}^{*}) + \sqrt{O(\delta\underline{\varepsilon}^{*}) + O(\delta\underline{\varepsilon}^{*}) + O(\delta\underline{u}^{*})} \;\Big|\; d(\Lambda,\Lambda') = O(\delta\underline{\Lambda}^{*}) + \sqrt{O(\delta\underline{\Lambda}^{*}) + O(\delta\underline{\Lambda}^{*}) + O(\delta p^{*})} \qquad (17)$$

and, as $\left| \begin{matrix} O(\delta\underline{\varepsilon}^{*}) \\ O(\delta\underline{\Lambda}^{*}) \end{matrix} \right|$ is neglectable in comparison with $\left| \begin{matrix} \sqrt{O(\delta\underline{\varepsilon}^{*})} \\ \sqrt{O(\delta\underline{\Lambda}^{*})} \end{matrix} \right|$,

$$\left[d(\Lambda,\Lambda') \right]^2 = O(\delta\underline{\varepsilon}^{*}) + O(\delta\underline{\varepsilon}^{*}) + O(\delta\underline{\varepsilon}^{*}) + O(\delta\underline{u}^{*}) \;\Big|\; \left[d(\Lambda,\Lambda') \right]^2 = O(\delta\underline{\Lambda}^{*}) + O(\delta\underline{\Lambda}^{*}) + O(\delta\underline{\Lambda}^{*}) + O(\delta\underline{p}^{*}) \qquad (18)$$

4 - APPLICATION TO FINITE ELEMENTS

Let us consider now the expressions for the displacements and strains valid for a general element e :

$$\underline{\varepsilon} = \underline{\lambda}^{e}\, \underline{\varepsilon}^{*} \qquad (19)$$

$$\underline{u} = \underline{\varphi}^{e}\, \underline{u}^{*} \qquad (20)$$

According to what was written in [3], matrices $\underline{\lambda}^{e}$ and $\underline{\varphi}^{e}$ are independent of each other.

The allowed field corresponding to a given field f is the

one which presents the same nodal displacements as f . The corresponding generalized strain vector, $\underline{\varepsilon}^e$, is connected with the nodal displacement vector $\underline{u}^e$ by the generalized strain-displacement equation

$$\underline{\varepsilon}^e = \underline{D}^e \underline{u}^e \tag{21}$$

It will be proved in this Section, first, that δu is of the order of $(l^e)^{n+1}$ if the expression for each displacement component,

$$u_i = \underline{q}_i^e \underline{u}^e \tag{22}$$

includes an arbitrary complete polynomial with degree n , and if the $(n+1)^{th}$ derivatives of such displacement component corresponding to f are all bounded within Ω^e .

l^e is a characteristic dimension of element e . On the other hand, $\underline{q}_i^e$ denotes the i^{th} row of $\underline{q}^e$.

Indeed, if the $(n+1)^{th}$ derivatives are all bounded, u_i may be expressed within Ω^e by the following Taylor's expansion

$$u_i = u_i(0) + u_{i,j}(0)(x_j - x_j^0) + \cdots + \frac{1}{n!} u_{i,j\cdots k}^{(n)}(0)(x_j - x_j^0)\cdots(x_k - x_k^0) + \frac{1}{(n+1)!} u_{i,j\cdots l}^{(n+1)}(Q_i)(x_j - x_j^0)\cdots(x_l - x_l^0) \tag{23}$$

where 0 and Q_i are points within Ω^e. Q_i depends on the co-ordinates of the points where the value of the function is to be determined.

Let us consider now the tangential displacement field to u_i in 0 ,

$$u_{t_i} = u_i(0) + u_{i,j}(0)(x_j - x_j^0) + \cdots + \frac{1}{n!} u_{i,j\cdots k}^{(n)}(x_j - x_j^0)\cdots(x_k - x_k^0) \tag{24}$$

which is allowed because all the polynomia of the n^{th} degree were assumed to be allowed.

Subtracting (24) from (23), we obtain

$$u_i - u_{t_i} = \frac{1}{(n+1)!} u_{i,j\cdots l}^{(n+1)}(Q_i)(x_j - x_j^0)\cdots(x_l - x_l^0) = O((l^e)^{n+1}) \tag{25}$$

It could be similarly concluded that the difference of the k derivatives is of the order of $(h^e)^{m+1-k}$

$$u_i^{(k)} - u_{ti}^{(k)} = O[(h^e)^{m+1-k}] \tag{26}$$

Consider now the displacement field u_i which takes the same nodal displacements as u_{ti}.

For sake of dimensional homogeneity, φ_{ij}^{k} must be of the form

$$\varphi_{ij}^{k} = (h^e)^{k} \psi_{ij}^{k}\left(\frac{x_1}{h^e}, \frac{x_2}{h^e}, \ldots\right) \tag{27}$$

if the generalized displacement u_i^k is a k^{th} derivative of u_i.

As u_i^l and u_{ti} are both allowed fields, the difference can be expressed by

$$u_i^l - u_{ti} = \underline{\varphi}_i^1 (\underline{u}^2 - \underline{u}_t^2) + \varphi_{ij}^2 (u_j^2 - u_{tj}^2) \tag{28}$$

Therefore, by virtue of (26) and (27),

$$u_i^l - u_{ti} = (h^e)^1 \psi_{ij}^{1}\left(\frac{x_1}{h^e}, \frac{x_2}{h^e}, \ldots\right) O[(h^e)^{m+1-1}] = O[(h^e)^{m+1}] \tag{29}$$

Subtracting (29) from (25) we obtain finally

$$\delta u_i = u_i - u_i^l = O[(h^e)^{m+1}] \tag{30}$$

as required.

An analogous theorems can be proved for the strains, i.e., that $\underline{\delta\varepsilon}$ is of order $(h^e)^{m+1}$ if the expression for each strain component

$$\varepsilon_x = \underline{\alpha}_i^1 \pm^2 \tag{31}$$

includes an arbitrary complete polynomial with degree m, and if

the $(m+1)^{th}$ derivatives of the strains corresponding to f are all bounded within Ω^e.

Consider indeed Taylor's expansion for the strain component ε_λ within Ω^e

$$\varepsilon_\lambda = \varepsilon_\lambda(0) + \varepsilon_{\lambda,i}(0)(x_i - x_i^0) + \cdots + \frac{1}{m!}\varepsilon^{(m)}_{\lambda,ij\cdots k}(0)(x_i - x_i^0)\cdots(x_k - x_k^0) + \frac{1}{(m+1)!}\varepsilon^{(m+1)}_{\lambda,ij\cdots k}(\theta_\lambda)(x_i - x_i^0)\cdots(x_k - x_k^0)(x_l - x_l^0) \tag{32}$$

Consider the tangential field at 0,

$$\varepsilon_{t\lambda} = \varepsilon_\lambda(0) + \varepsilon_{\lambda,i}(0)(x_i - x_i^0) + \cdots + \frac{1}{m!}\varepsilon^{(m)}_{\lambda,ij\cdots k}(0)(x_i - x_i^0)\cdots(x_k - x_k^0) \tag{33}$$

Subtracting (33) from (32), we obtain

$$\varepsilon_\lambda - \varepsilon_{t\lambda} = \frac{1}{(m+1)!}\varepsilon^{(m+1)}_{\lambda,ij\cdots k}(\theta_\lambda)(x_i - x_i^0)\cdots(x_k - x_k^0)(x_l - x_l^0) \tag{34}$$

Let f_t denote the structural field whose strains coincide with $\varepsilon_{t\lambda}$ within Ω^e. Considering (34),

$$\|f - f_t\| = O[\ell^{m+1}] \tag{35}$$

ℓ being the largest value of ℓ^e on the set of subdomains Ω^e.

Consider now the allowed field corresponding to f

$$f' = T(f) \tag{36}$$

As f_t is already an allowed field,

$$f_t = T(f_t) \tag{37}$$

and, therefore, as T is a linear operator,

$$f' - f_t = T(f - f_t) \tag{38}$$

Now, if $\|T\|$ is the norm of T,

$$\|f' - f_t\| < \|T\| \cdot \|f - f_t\| \tag{39}$$

and, as $\|T\|$ is bounded,

$$\|f - f_\lambda\| = O[\lambda^{m+1}] \tag{40}$$

Therefore,

$$\|f - f^h\| \leq \|f - f_\lambda\| + \|f_\lambda - f^h\| = O[\lambda^{m+1}] \tag{41}$$

This means that

$$\delta_E = O[\lambda^{m+1}] \tag{42}$$

and the theorem is thus proved.

Analogous theorems could be proved for the complementary energy method. It could be concluded thus that δ_λ is of the order of $(\lambda)^{n+1}$, n being the degree of the arbitrary complete polynomial of highest degree included in the polynomial expressions of the stress components, and δ_p is of the order of $(\lambda)^{m+1}$, m being the degree of the arbitrary complete polynomial of highest degree included in the polynomial expression of the tractions. These last expressions concern of course the element boundaries. Boundedness of the $(n+1)^{th}$ and $(m+1)^{th}$ derivatives of the components are assumed.

5 - RULES CONCERNING THE CHOICE OF THE TYPE OF ELEMENT

Introducing (29) and (42) into (1), there results

$$d(\Delta, \Delta_\lambda) = O(\lambda^{n+1}) + O(\lambda^{m+1}) \tag{43}$$

Thus, if n and m are both larger than unity, completeness is ensured because $d(\Delta, \Delta_\lambda)$ tends to zero with λ.

An important point is that n must be equal to m.

Indeed, if m is larger than n, $d(\Delta, \Delta_\lambda)$ is of the order of $(\ell^e)^n$. This means that the accuracy of the displacements (or tractions) is unnecessarily high. The same can be said about the accuracy of

stresses if n is larger than m.

Now, in the first modality, the strains are related to the displacements by the strain-displacement equations

$$\underline{\varepsilon} = \underline{D}\,\underline{u} \tag{44}$$

in which $\underline{D}$ is a differential operator. The stresses can be determined in terms of the strains by means of the Hooke's law, and thus

$$\underline{\sigma} = \underline{H}\,\underline{D}\,\underline{u} \tag{45}$$

In that case, if the derivatives involved in $\underline{D}$ are of order up to p, it is easy to demonstrate that

$$n = m + p \tag{46}$$

Now, as p cannot be less than one, n can never become equal to m, as long as the first modality is used. This means that the first modality is not logical: the accuracy of the displacements and the accuracy of the stresses are not consistent with each other.

Take for instance the case of plane elasticity, for which $p = 1$, and assume that the allowed displacements are linear, which means that the allowed stresses are constant. n is thus equal to 1 and m equal to 0. Consistency would be achieved only if the displacements and stresses had both the same degree, which is only possible in the second modality, i.e., if hybrid elements are used.

In the case of plates, without transverse shear deformability, $p = 2$. Assuming the allowed transverse displacements to be of the third degree, the rotations are of the second degree, and as the rotations are included among the generalized displacements, $n = 2$. On the other hand, the moments are linear, and, thus, $m = 1$. The stresses and the rotations should however be both of the same degree. Hybrid elements ought to be used again. Remark however that using stresses of degree higher than the degree of the rotations will be useless. The same will happen in plane elasticity if stresses are used with degree higher than the degree of the displacements.

The same conclusions were reached by Pin Tong and Pian [5].

The third modality, in which tractions are related to stresses through the traction-stress equation

$$\underline{p} = \underline{N}\,\underline{\sigma} \tag{47}$$

is consistent under the point of view of n being equal to m.

Indeed as $\underline{N}$ is not a differential operator, the accuracy in stresses is the same as the accuracy in tractions, which implies $n = m$.

This explains why, for a given total number of unknowns, more accurate stresses are obtained with the second and third modalities than with the first.

Although no general discussion about the relative merits of simple and complex elements will be attempted in the present paper, it must be observed that the convergence theorems do not ensure that more accurate results will always be obtained if n and m become both larger and larger.

First, some of the derivatives of the stresses of the exact solution with order up to $n+1$ may be unbounded within one of the subdomains Ω^s.

If this happens, i.e., if some of the derivatives of the stresses of the exact solution happens to be unbounded within any of the subdomains Ω^s, n must be smaller than the order of such derivatives.

Such is the only condition for n if compatibility or equilibrium are not violated, respectively in the first modality and in the third.

Assume now that compatibility is violated across the elements boundaries (non-conforming elements in the first modality) or within the elements (second modality). Then, the derivatives of order up to $m+1$ of the field λ_a must be bounded also.

This means that, if some of the derivatives of the stresses corresponding to the field λ_a happens to be unbounded within any of the subdomains Ω^s, n must be smaller than the order of such derivatives.

We remark that it is possible to predict the boundedness of the derivatives of the stresses corresponding to λ_a by examining the

behaviour of the derivatives of the stresses corresponding to the approximate solution [3,4].

6 - RULES CONCERNING THE MESH

One of the greatest advantages of the finite element method is the almost complete freedom in the location of the nodes.

Such freedom is used mainly near singularities and on the boundary. Inside the domain, the elements are usually disposed according to regular patterns which are particularly unavoidable if the mesh is automatically generated. Some rules will be established in the present Section concerning the kind of pattern which should be used.

The two patterns represented in fig. 6.1a and b are to be compared with each other.

The first pattern, characterized by the fact that all the nodes are in the same conditions, has proved better than the second (union-jack mesh) under the point of view of convergence [2,3,4,6]. Supposing however that convergence is achieved, will it be possible to prove that the first is absolutely better than the second?

It will be concluded in the present Section that such statement cannot be made. The first pattern can indeed be better or worse than the second. It all depends on its orientation with respect to the lines of constant strain energy density (isoenergetic lines).

The proof is very simple.

Consider first the case of a function of two variables $f(x_1, x_2)$ which to be approximated on a two-dimensional domain Ω by a piecewise defined function f_1 corresponding to a certain subdivision of Ω into subdomains.

Consider now on such domain the lines f = constant and assume the domain subdivided according to the first pattern (fig. 6.2a).

By rotating segments CD and EF , we obtain the second pattern (fig. 6.2b), which is worse than the first, because squares NPRS and QRTU should be preferably subdivided in the direction perpendicular to the lines f=constant, along which f presents its steepest variation.

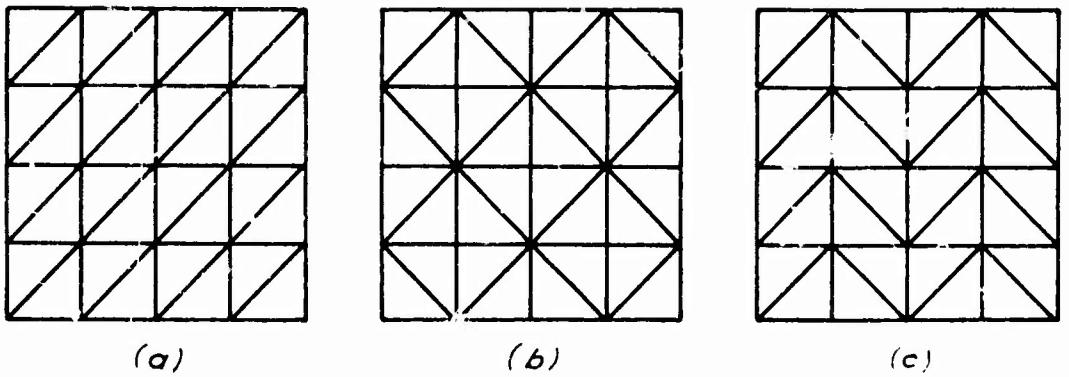


Fig. 6.1

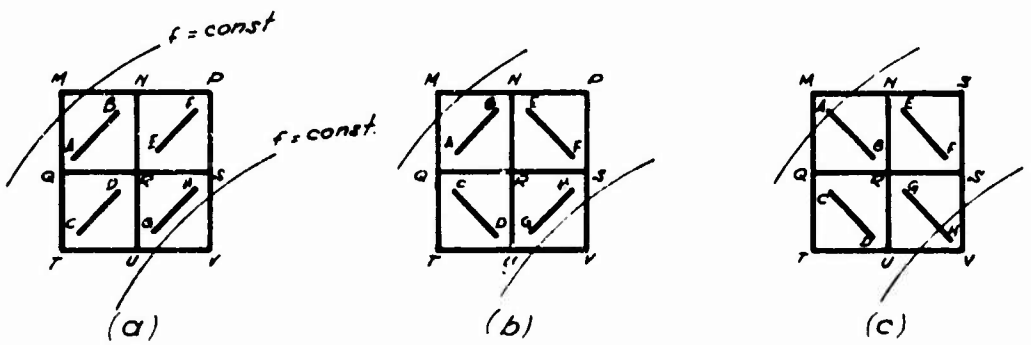


Fig. 6.2

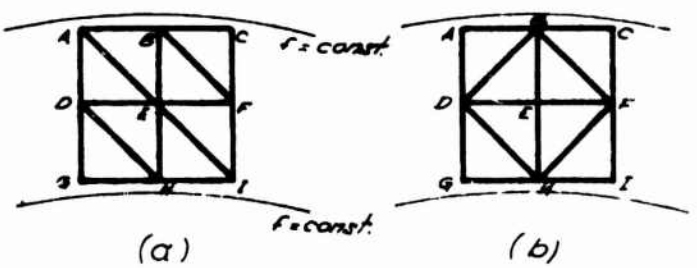


Fig. 6.3

Assuming for instance that the function f can take only constant values within each triangle, the difference between cases a and b is that, in the first case, f can take different values on triangles TQR and TRU, and, in the second case, can take different values on triangles TQU and QRU . However, as the function f takes the same values on both sides of diagonal QU , it is much more important that f takes different values on triangles TQR and TRU than on triangles TQU and QRU .

The approximation becomes still worse if segments AB and GH are also rotated (fig. 6.2c). The pattern of Fig. 6.2c is however nothing else than the first one rotated 90 degrees.

We conclude thus than the first pattern is better than the second if the diagonals are parallel to the lines f=constant, and worse if the diagonals are perpendicular to such lines.

If the second pattern is rotated 90 degrees, we obtain still the same pattern, which can be considered thus as isotropic.

It is important to remark that rotating both patterns 45 degrees, we obtain pratically the same approximation.

Indeed, referring to fig. 6. 3a and b, it cannot matter very much that diagonals AE and EI become diagonals DB and BF.

In practice, the lines f=constant will generally not be straight, and thus the mesh cannot result from the intersection of three families of parallel straight lines.

The preceeding discussion inspires however a very simple rule.

Such rule consists in drawing a set of lines f=constant and subdividing the space between them as in fig. 6.4a. We obtain thus a mesh similar to the one in fig. 6.1a.

It is clear that worse results would be obtained if diagonals DE, EF and FG were replaced by diagonals AH, BI and CJ as in fig. 6.4b, which corresponds to fig. 6.1a.

The mesh represented in fig. 6.4c corresponds to the pattern of fig. 6.1c. It leads pratically to the same results as the mesh of fig. 6.4a, since the subdivision of the triangles in a direction perpendicular to the lines f=constant has pratically no influence. However, the number of lines is considerably smaller. For each is

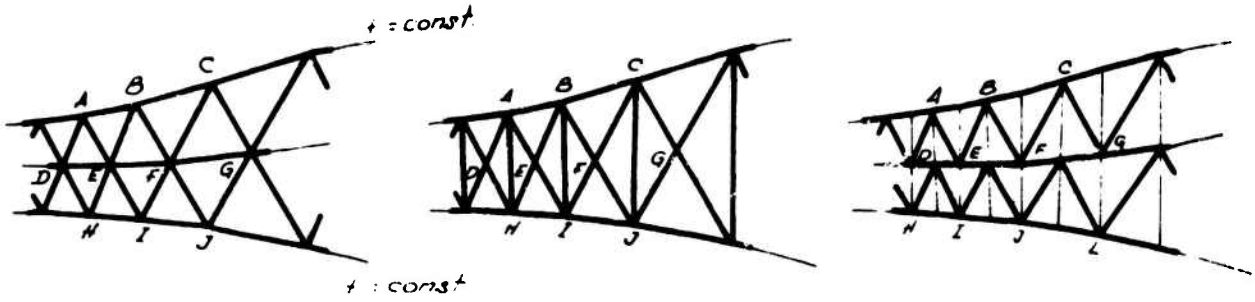


Fig. 6.4

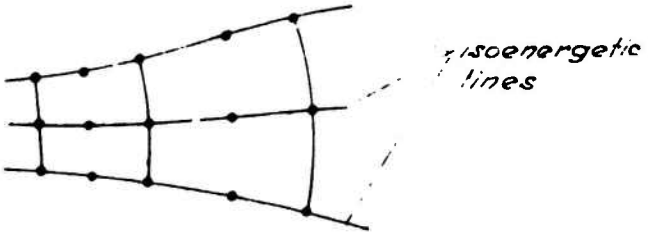


Fig. 6.5

thus to be rejected.

The conclusions which have been draw concern the approximation of one single function, and seem difficult to generalize to the case of structural fields, in which the distance depends on the approximation of the whole set of stresses, i.e., on the simultaneous approximation of several functions.

The generalization will be quite simple however if it is remarked that the approximation of the exact solution by the approximate solution is really what matters.

Assume indeed linearity, both geometrical and physical.

The case of the approximate solution being compatible will first be considered. As the distance between a general compatible solution and the exact solution is equal to the difference of the corresponding total potential energies, the difference between the exact and approximate solutions is also equal to the difference between the total potential energies of the approximate solution and the exact solution.

Assuming now that the prescribed incompatibilities(*) vanish, there results that the total potential energies of both the exact and approximate solutions equal their respective strain energies taken with reversed signe.

Thus, provide compatibility is not violated and the prescribed incompatibilities vanish, our interest is to minimize the difference between the strain energies of the exact and approximate solution.

This means that the strain energy density of the approximate solution must approximate the strain energy density of the exact solution in the best possible way. By other words, all that has been said becomes applicable it function f represents the strain energy density of the exact solution, and f_1 the strain energy density of the approximate solution.

The lines f =constant become the lines of constant strain energy density which will be called isoenergetic in the sequel.

(*) The prescribed incompatibilities are the displacements prescribed on the boundary, the prescribed initial strains (or initial stresses) and the prescribed displacement discontinuities.[347]

In case the approximate solution is equilibrated, the distance to the exact solution becomes the difference between the corresponding total complementary energies which, if the prescribed incompatibilities vanish, equal the complementary energies, both in case of the approximate and of the exact solution.

But, as linearity is assumed, the isoenergetic lines are also lines of constant complementary energy density and nothing changes.

The same conclusions would be reached for the less usual case of vanishing external forces and non-vanishing incompatibilities.

If the approximate solution is not compatible nor equilibrated no conclusions can be drawn. However, provided convergence is achieved, the approximate solution becomes more and more compatible (in the potential energy approach) or more and more equilibrated (in the complementary energy approach) as the size of the elements decreases, so that, after a certain degree of subdivision of the domain, we can reason as if the approximate solution was really compatible (in the potential energy approach) or equilibrated (in the complementary energy approach).

To close this Section, let us remark that the use of isoparametric[11] elements seems very convenient in the present context. The space between the isoenergetic lines will be indeed ideally covered by isoparametric elements as it is suggested in fig. 6.5.

7 - ALLOCATION OF THE NODES

Knowing that the nodes must be disposed along the isoenergetic lines is not enough.

In order that the topology of the mesh can be completely defined, decisions are indeed needed about how many isoenergetic lines are to be located at each of such lines.

And then, after the topology of the mesh has been defined, the isoenergetic lines must actually be selected and the nodes actually be distributed on them.

In the present paper, we shall satisfy ourselves with indicating how, in the linear case, the best allocation of the nodes can be determined starting from a given topology.

Assuming the approximate solution to be compatible, the total potential energy of the approximate solution will always be larger than the total potential energy of the exact one.

As the distance between both is equal to the difference of the corresponding total potential energies, it follows that the best approximate solution is the one which provides the minimum total potential energy.

The optimum allocation of the nodes can thus be determined by minimizing the total potential energy with respect to the nodal coordinates. A program based in this principle is known to have been prepared by Marçal and McNeice [9] at Brown University.

As the approximate solution results itself from the minimization of the total potential energy with respect to the nodal displacements, it results that the minimization of the total potential energy with respect both to the nodal displacements and nodal coordinates will provide the best possible solution corresponding to a given type of elements and a given topology.

Analogous considerations can be made if the approximate solution is equilibrated and not compatible. The total complementary energy must then be minimized instead of the total potential energy.

Assuming that the prescribed incompatibilities vanish, the best compatible approximate solution is the one which maximizes the strain energy, and the best equilibrated approximate solution is the one for which the strain energy is a minimum. This results from what was written above about the relations between the total potential and complementary energies and the strain energy, in case of vanishing prescribed incompatibilities.

If the prescribed external forces vanish, it is the best equilibrated solution which maximizes and the best compatible solution which minimizes the strain energy.

The rule established in Section 6 could thus be tested in a very simple example: the problem of the hollow cylinder[2] submitted to a uniform pressure on the outer surface (with $\nu = 0$) was solved using triangular finite elements with a linear displacement field.

Fig. 7.1a and b show the two meshes which have been considered The first one, which respects the rule of Section 6, led to a value of the strain energy 1% larger than the second.

8 - CONCLUSIONS

Two main rules have been established in the paper.

The first concerns the type of the element and requires the accuracy in stresses to be consistent with the accuracy in displacements (first and second modality) or with the accuracy in tractions (third modality).

Such rule which, in the present paper, results from a generalization of the convergence analysis presented in previous papers by the author, was indicated in 1969 by Fin Tong and Pian[5] for the case of hybrid elements, using a different reasonment. It implies that the first modality is inferior to the second and third.

The second rules concerns the topology of the mesh. It results from this rule that the mesh characterized by the same number of lines intersecting at each node is not absolutely better than the union-jack mesh. Il all depends on its orientation with respect to the isoenergetic lines.

It is fair to remark that the requirement of disposing the elements along the isoenergetic lines is not always easy to follow.

First, the isoenergetic lines are not known a priori. We can guess of course how they are before any computation is made. It seems however that it is easier to guess stress trajectories than isoenergetics. On the other hand, the properties of isoenergetics are not known. Theorems about isoenergetics should thus be investigated.

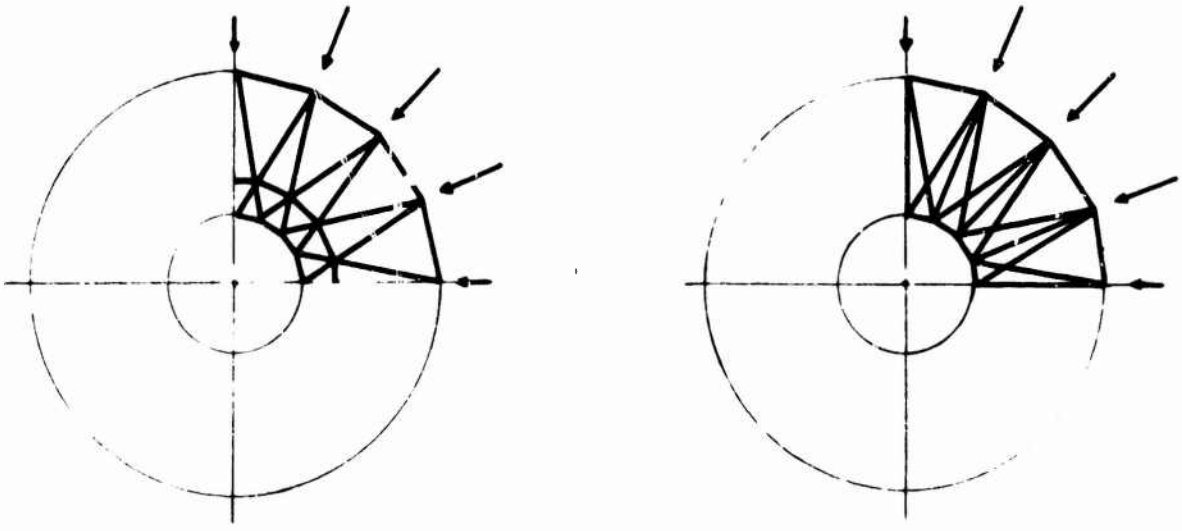
We can of course always resort to a preliminary coarse analysis which may give information about the variation of the strain energy density.

However, even if the isoenergetic lines are known, it is not always easy to dispose the elements along them. Such is for instance what happens in the classical problem of the infinitely large plate with a circular hole submitted to a uniform tension[9], the isoenergetic of which are represented in fig. 8.1.

In those cases, the rule can still be used within certain critical regions of the domain, mainly where stress concentrations appear.

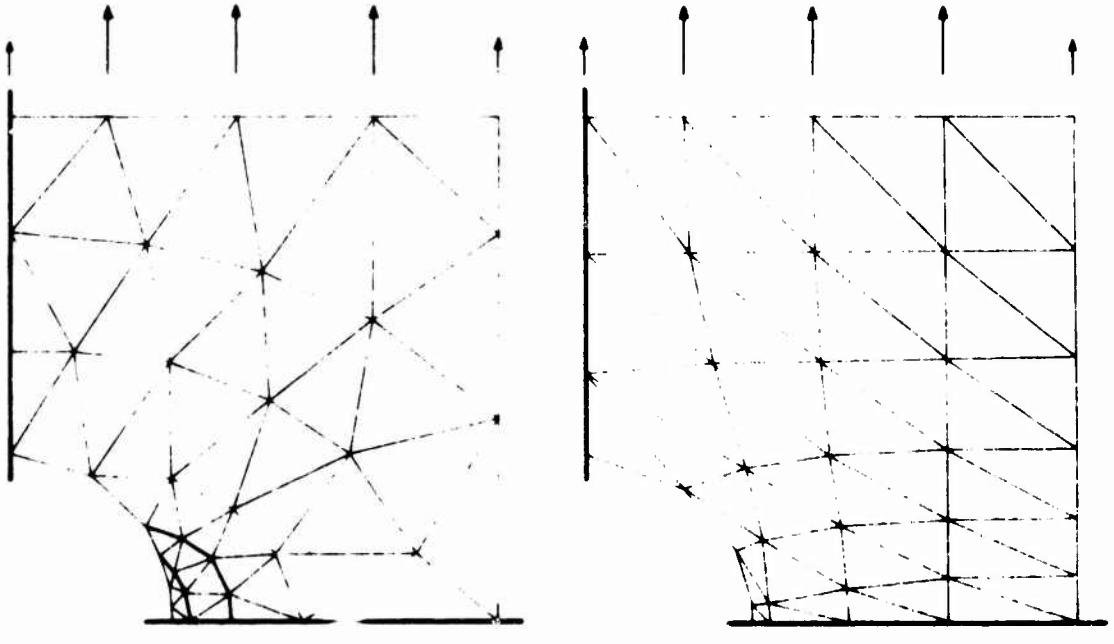
Fig. 8.2 shows how the rule has been used in the case of the plate with the finite width. In fig. 8.2a two isoenergetics have been used for disposing the elements near the point where the stress concentration factor is larger, while in fig. 8.2b the mesh has been arranged in a more traditional way. The first mesh led to a value of the strain energy 3.1% larger than the second. This shows that benefits have been drawn even if the rule was not fully adopted.

Reference was made in Section 8 to the possibility of determining the best coordinates of the nodes by minimizing the total potential energy with respect both to the nodal displacements and nodal coordinates. It is our intention to use such method in the near future to find and test new simple rules which may guide structural analysts in choosing their meshes in the best possible way.



(a) (b)

Fig. 7.1



(a) (b)

Fig. 8.2

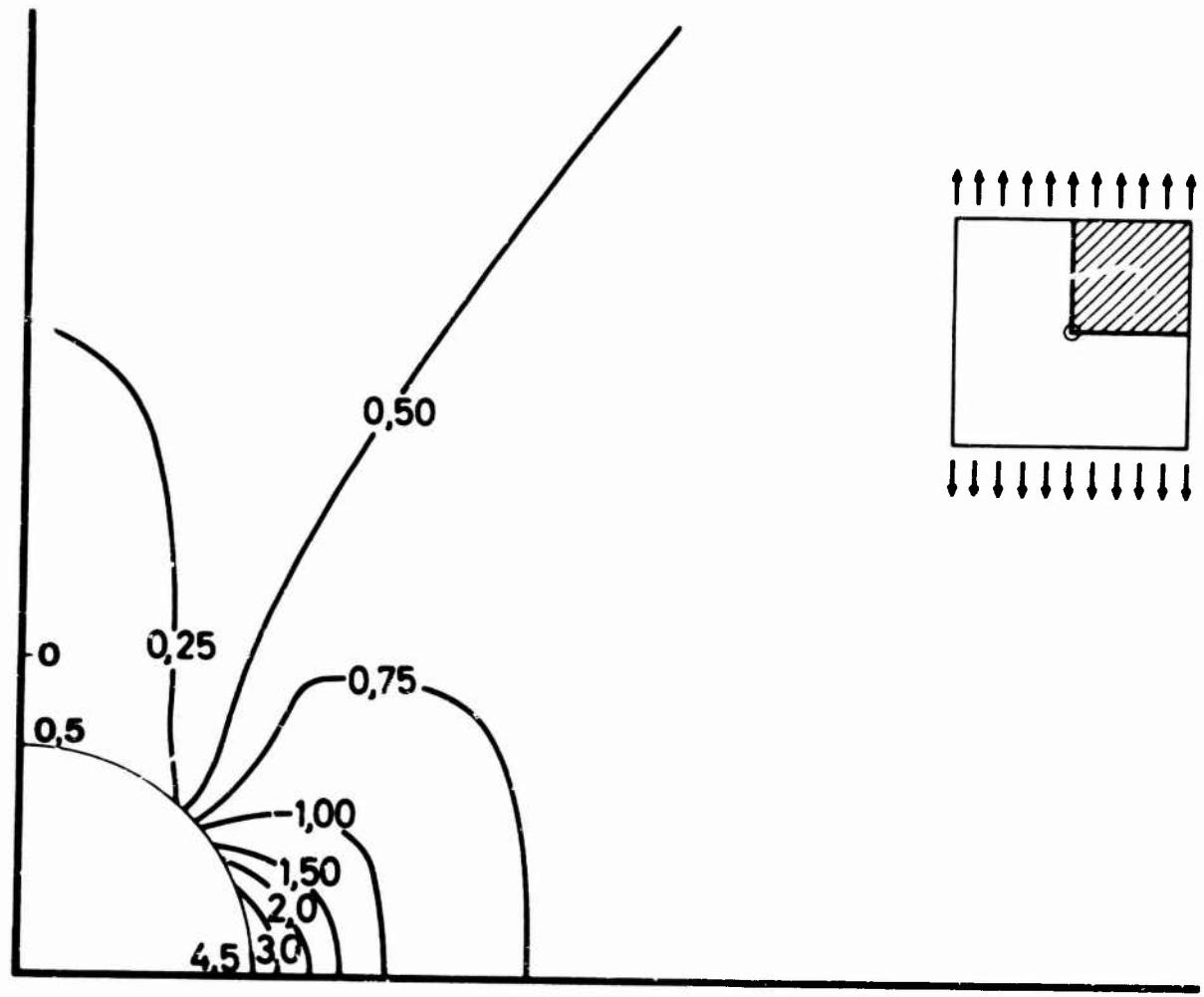


Fig. 8.1

REFERENCES

1 - OLIVEIRA, E.R. ARANTES - "Theoretical Foundations of the Finite Element Method". Int. J. Solids Structures, Vol. 4, No. 10, October, 1968, pp. 929-952.

2 - OLIVEIRA, E.R. ARANTES - "Completeness and Convergence in the Finite Element Method", Proc. of the 2nd Conference on Matrix Methods in Structural Mechanics, Wright-Patterson AFB, Ohio, 1968.

3 - OLIVEIRA, E.R. ARANTES - "The Convergence Theorems and their Role in the Theory of Structures", IUTAM Symposium on High Speed Computing of Elastic Structures, Liège, 1970.

4 - OLIVEIRA, E.R. ARANTES - "Lectures for the NATO Advanced Institute on Finite Element Methods in Continuum Mechanics", Lisbon, September, 1971.

5 - PIN TONG; PIAN, T.H. - "A Variational Principle and the Convergence of a Finite Element Method Based on Assumed Stress Distribution", Int. J. Solids Structures, Vol. 5, May, 1969, pp. 463-472.

6 - IRONS, B.M.; ZIENKIEWICZ, O.C.; OLIVEIRA, E.R. ARANTES - - Comment on the paper: Theoretical Foundation of the Finite Element Method". Int. J. Solids Structures, Vol.6, pp. 695-697.

7 - ERGATOUDIS, J.; IRONS, B.M. and ZIENKIEWICZ, O.C. - "Curved Isoparametric, Quadrilateral Elements for Finite Element Analysis". Int. J. Solid Struct. 4 31-42, 1968.

8 - TIMOSHENKO, S.; GOODIER, J.N. - "Theory of Elasticity", McGraw-Hill, New York, 1951.

9 - McNEICE, G.M., MARÇAL, P.V. - "Optimization of Finite Element Grids Based on Minimum Potential Energy"- - Report, Division of Engineering, Brown University.

EXTENDED INTERPOLATION IN FINITE ELEMENT ANALYSIS*

L. S. D. Morley**
Royal Aircraft Establishment, Farnborough, UK

June 1971

It is a widely held view that mid-side connection properties impose a disproportionately severe penalty upon the computational efficiency of finite element calculations. Here, an exploration is made of a technique which eliminates these undesirable connections by an extension of spacewise interpolation across the boundaries of triangular finite elements and yet retains much of the original accuracy. The technique is applied with success to the conforming plate bending finite element with cubically varying normal displacement along the triangle sides and also to the plane stress element with quadratically varying displacements. An important development is a constant bending moment finite element which has the simplest possible connection properties involving only the normal displacement at the vertices of the triangle.

** Senior Principal Scientific Officer, Structures Department.

1. INTRODUCTION

The finite element method is a piecewise epplicetion of the clessicel Rayleigh-Ritz technique whereby all the essential connection properties between elements are determined by the governing variational principle, e.g. minimum potential or complementery energy. Spacewise interpolation techniques are employed within each element in order to set up coordinate, or shepe, functions as a preliminary to the ectual calculation of the numericel velues associated with the connection properties. The question does not seem to have been explored, nowever, as to whether there may be advanteges in extending these spacewise interpolation techniques across and beyond the element boundaries.

Let us begin by looking at the conforming plate bending triangular finite element with normal displacements w which vary cubically along the element sides; early formulations are given by Clough and Tocher[1] and by Bazeley *et al*[2] although a more recent formulation[3] is used for the numerical work in the present paper. The mid-side connection property, which is required to ensure continuity of the normal derivative $\partial w/\partial n$ of the displacement across the element interface, is eliminated in these early formulations as the arithmetic mean of the appropriate directional derivatives taken at the nodes at each end of the side. This simple interpolation, used with this element, leads to an unduly overstiff structure and, accordingly, finds little favour in practical applications. On the other hand, while retention of the mid-side connection leads to very satisfactory numerical results, the accompanying inflation of the computer storage requirements coupled with an increase in execution time imposes a disproportionately severe penalty upon the computational efficiency. The trivial case which is shown in Figs.1a and 1b provides an example where the storage requirement, when it is exemplified by the global stiffness matrix in banded form without sub-structuring, turns out to be doubled. Although the size of the global stiffness matrix provides an inadequate measure for the efficiency of modern computing techniques it is, nevertheless, this kind of penalty which encourages the development of higher degree elements, see, e.g. Bell[4], with excess nodal continuities but where the mid-side connection quantities are easily eliminated at element level and without significant loss of accuracy.

In re-examining the conforming plate bending triangular finite element with w displacements which vary cubically along the element sides it was noticed that it is a relatively simple matter to arrange an extension of the finite element spacewise interpolation process so as to estimate the mid-side normal derivative $\partial w/\partial n$ more accurately than hitherto. This is achieved in explicit terms of the twelve nodal connection quantities of the two triangular elements which share the common side; it produces, for example, the situation shown in Fig.1c where the size of the global stiffness matrix is only twenty per cent above that which is required for the simple interpoletion of Fig.1b. Moreover, the numerical results which are calculated for the classical comparative problems of the bending of a square plate now show a negligible difference from those obtained with retention of the mid-side connection.

With this encouregement, the idea of an extended interpoletion is pursued to develop a trianguler plete bending element which has the simplest possible connection properties involving only the displacement w at the nodes. This development is centred around the constant bending moment element which hes attracted many different derivations of whet is really e strict equilibrium element, e.g.[5,6,7,8,9]. Recent derivations[8,9] ere by

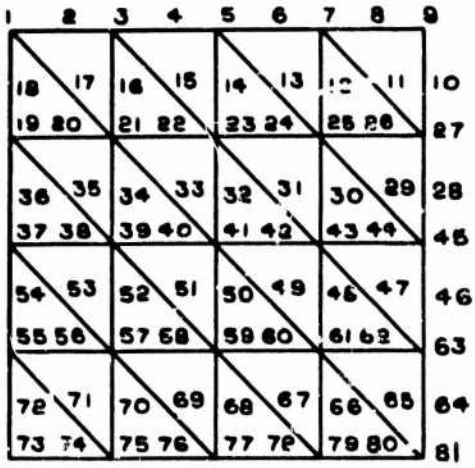


Fig.1a Mid-side connection points retained

Size of global stiffness matrix:
Rows = 131
Semi-bandwidth = 35
(131 × 35 = 4585)

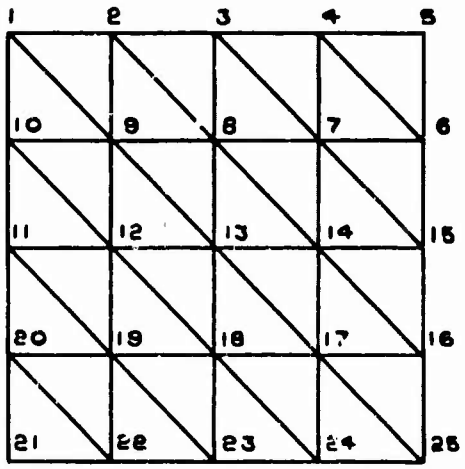


Fig.1b Mid-side connection points eliminated by simple interpolation

Size of global stiffness matrix:
Rows = 75
Semi-bandwidth = 30
(75 × 30 = 2250)

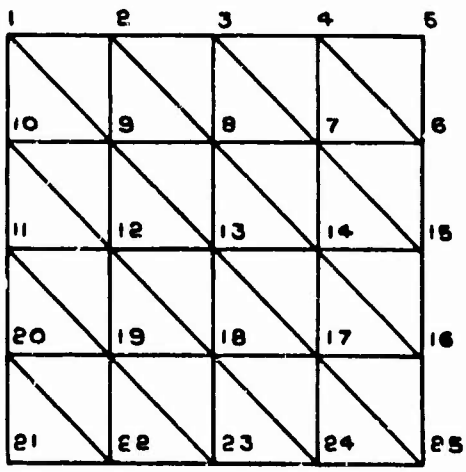


Fig.1c Mid-side connection points eliminated by extended interpolation

Size of global stiffness matrix:
Rows = 75
Semi-bandwidth = 36
(75 × 36 = 2700)

Fig.1 Element numbering for conforming plate bending triangular elements with cubically varying edge displacement w

way of a straightforward application of the theorem of minimum potential energy to a non-conforming quadratically varying w displacement with inter-element connections of w at each node and of $\partial w/\partial n$ at each mid-side as is shown in Fig.2a. Extended interpolation is now applied to express the mid-side $\partial w/\partial n$ in terms of the displacement w at nearby nodes. Although the interpolation destroys the strict equilibrium nature of this element our numerical examples show that it provides, more often than not, agreeable improvements in the accuracy of the interesting physical quantities as is compared with the situation where the mid-side connection points are retained. One explanation for this phenomenon is that the excess flexibility of the strict equilibrium element is suitably moderated by the kinematic constraint which is imposed by the extended interpolation.

The concluding application deals with the elimination of the mid-side connection properties which are normally required for the quadratically varying displacements u and v in the well known triangular plane stress finite element[10,11]. Comparative numerical results are given here for the square plate under uniform tension which contains a central circular hole. This problem is known[12,13] to be particularly sensitive because of the severe gradients which occur in the circumferential stress resultant near the hole boundary and, indeed, the comparative finite element results for linearly varying u,v displacements are found to be virtually worthless for the mesh under investigation.

The Appendix provides a listing of a Fortran subroutine which deals rudimentarily with the topological exercise of finding, for each triangular element, those nearby node numbers which are prerequisite to the application of the extended interpolation processes.

It is emphasized that this investigation is by nature exploratory. While extended interpolation is seen to reduce the size of the overall computational problem when it is measured by the global stiffness matrix, it does involve additional computational effort at element level and, in this respect, it bears resemblance with the sub-structuring technique. Notwithstanding, the extended interpolation is likely to enjoy particularly beneficial application to studies which concern buckling, vibration and optimum design. The stratagem in the elimination of the mid-side connection quantities hinges here upon an exact recovery of all global regular polynomial solutions as are permitted by the basic shape functions of the element itself. Finally, the application of extended interpolation is most ideally suited to an element mesh of equilateral triangles although it is by no means necessary, as is evidenced by the numerical examples treated herein, to impose such a limitation on a mesh which is already in accord with sound finite element practice.

Acknowledgement is gladly given to B. C. Merrifield and to D. R. Blackeby for unstinting labours in preparing computer programs, respectively for the plate bending and for the plane stress problems, also to Carol Hanson who assisted in the preparation of the subroutine which is listed in the Appendix.

2. PLATE BENDING ELEMENT WITH CUBICALLY VARYING w-DISPLACEMENT ALONG THE SIDES

The triangular finite element of arbitrary shape is shown in Fig.3 with nodes numbered 1 3 5. A point within the element is located in the usual way by non-dimensional area coordinates L_1, L_3 and L_5 where

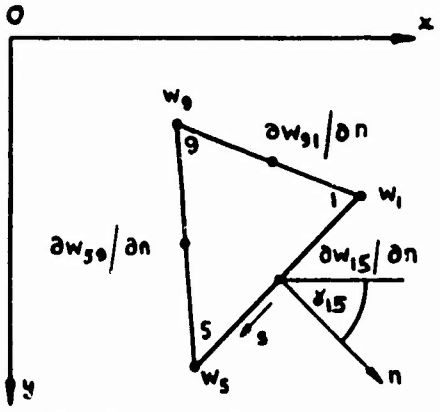


Fig. 2a Equilibrium element

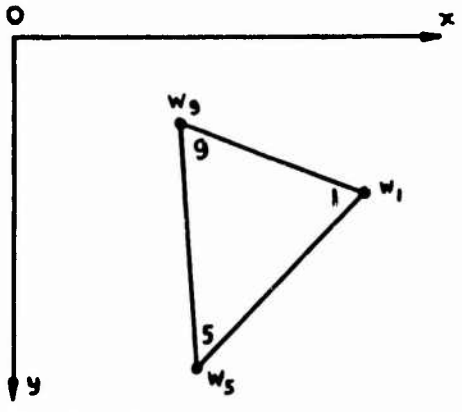


Fig. 2b With extended interpolation

Fig. 2 Connection properties of plate bending element with quadratically varying w displacement

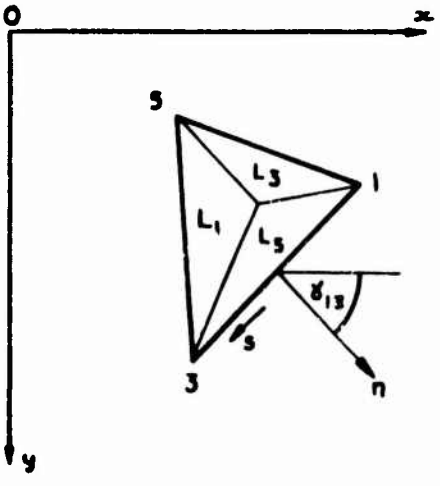


Fig. 3a Notation

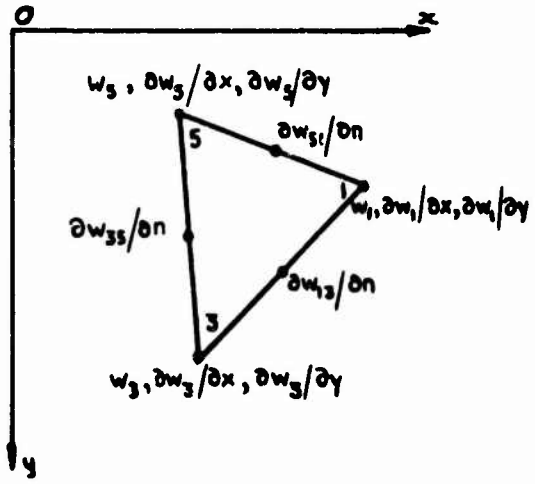


Fig. 3b Usual connection properties

Fig. 3 Plate bending element with cubically varying w displacement along the sides

$$L_1 = (a_1 + b_1 x + c_1 y)/2A,$$

$$a_1 = x_3 y_5 - x_5 y_3, \qquad b_1 = y_3 - y_5, \qquad c_1 = x_5 - x_3, \tag{2-1}$$

with the remaining expressions following by permutation of the suffices. The area of the element is denoted by A where

$$2A = b_3 c_5 - b_5 c_3 \ . \tag{2-2}$$

The theory of conforming triangular plate bending elements with cubically varying w-displacement along the sides is well documented[1,2,3]; the connection properties are illustrated in Fig.3b where the mid-side connection points are present in recognition of the correct quadratic variation of $\partial w/\partial n$ along the sides. From the very practical viewpoint of efficient computation, however, it is desirable to eliminate these mid-side connections because, as discussed in the Introduction, they are responsible for a disproportionately large increase in both the bandwidth and in the total number of rows in the global stiffness matrix. Their elimination by the simple expediency of linearizing[1,2] the normal derivative $\partial w/\partial n$ along the sides of the triangular elements is found to be unsatisfactory because the constraints which are thereby imposed lead to an unacceptably overstiff solution. Clearly, a more precise interpolation for this mid-side value of $\partial w/\partial n$ is worth investigating where, at least, a partial aim must be the exact recovery of all solutions where the displaced shape $w(x,y)$ is described in terms of the general surface cubic.

In what follows, the value of each mid-side $\partial w/\partial n$ is to be interpolated from the twelve nodal connection quantities of the two triangles which share each common side. The basic topology of this extended interpolation is denoted by nodal numbers 1 2 3 4 5 6 and is illustrated in Fig.4a where the focal element is picked out by node numbers 1 3 5. The separate components of this topology are shown in Fig.4b, while the Fig.4c provides an example of a permitted variation where the nodes 2 and 4 coalesce. The listing of a Fortran computer subroutine is given in the Appendix which, when it is provided with a table of focal element node numbers like

1	0	2	0	9	0
2	0	8	0	9	0
9	0	8	0	13	0

etc.

that are appropriate to the situation shown in Fig.1c, derives the table of node numbers

1	0	2	8	9	10
2	3	8	13	9	1
9	2	8	14	13	12

etc.

which define the basic topology of our extended interpolation. The 0 in this latter table signifies that the side 1 2 coincides with the plate boundary.

Consider the component four-point topology which is denoted by node numbers 1 2 3 5 as is illustrated in Fig.4b. No loss in generality is entailed by imposing a temporary translation and rotation of the coordinate axes xOy so that the side 1 3 now lies on the Oy axis with the Ox axis passing through the mid-side point as is shown in Fig.5a. Because of this convenient reorientation of the coordinate axes, the subsidiary formulae in equations (2-1) simplify to

$$\left.\begin{aligned} a_1 = a_3 = -x_5y_3, \quad a_5 = 0; \quad b_1 = y_3 - y_5 \quad b_3 = y_3 + y_5 \quad b_5 = -2y_3; \\ c_1 = -c_3 = x_5, \quad c_5 = 0 \end{aligned}\right\} \tag{2-3}$$

We start by fitting the ten term general surface cubic to the ten connection points which remain in triangle 1 3 5 of Fig.5b when the mid-side connection $\partial^2 w_{13}/\partial x^2$ is omitted. Thus

$$w(x,y) = \sum_{j=1,3,5} (N_j w_j + N_{jx}\partial w_j/\partial x + N_{jy}\partial w_j/\partial y) + N_{13}\partial w_{13}/\partial x \tag{2-4}$$

where, on making use of equations (2-3), the shape functions $N_1, N_{1x}, \ldots$ are given by

$$\left.\begin{aligned} N_1(x,y) &= L_1 - b_1(N_{1x} + N_{3x} + N_{5x})/2A - c_1(N_{1y} + N_{3y} + N_{5y})/2A - b_1N_{13}/2A, \\ N_{1x}(x,y) &= -c_1L_1L_5(L_3 - L_1), \\ N_{1y}(x,y) &= L_1(2b_1L_3L_5 + b_3L_1L_5 - b_5L_1L_3), \\ N_3(x,y) &= L_3 - b_3(N_{1x} + N_{3x} + N_{5x})/2A + c_1(N_{1y} + N_{3y} + N_{5y})/2A - b_3N_{13}/2A, \\ N_{3x}(x,y) &= c_1L_3L_5(L_3 - L_1), \\ N_{3y}(x,y) &= -L_3(2b_3L_1L_5 + b_1L_3L_5 - b_5L_1L_3), \\ N_5(x,y) &= L_5 - b_5(N_{1x} + N_{3x} + N_{5x})/2A - b_5N_{13}/2A, \\ N_{5x}(x,y) &= -c_1L_5^2(L_1 + L_3), \\ N_{5y}(x,y) &= L_5^2(b_1L_3 - b_3L_1), \\ N_{13}(x,y) &= 4c_1L_1L_3L_5, \end{aligned}\right\} \tag{2-5}$$

where A denotes the area of the triangle 1 3 5 and, from equations (2-2) and (2-3),

$$2A = -b_5c_3 = 2a_1 = 2a_3 = -2x_5y_3 \ . \tag{2-6}$$

At the middle of side 1 3 we note that the second differentials of these shape functions with respect to x are given by

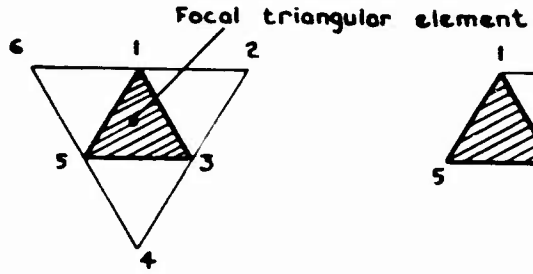


Fig. 4a Six point basic topology

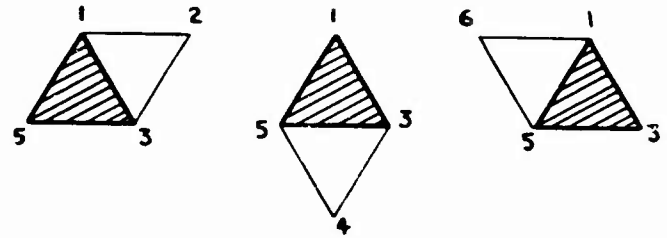


Fig. 4b Component four point topologies

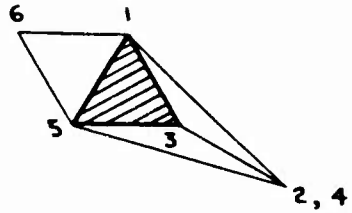


Fig. 4c Example of permitted variation in basic topology

Fig. 4 Topology of extended interpolation for element with cubically varying w displacement along the side

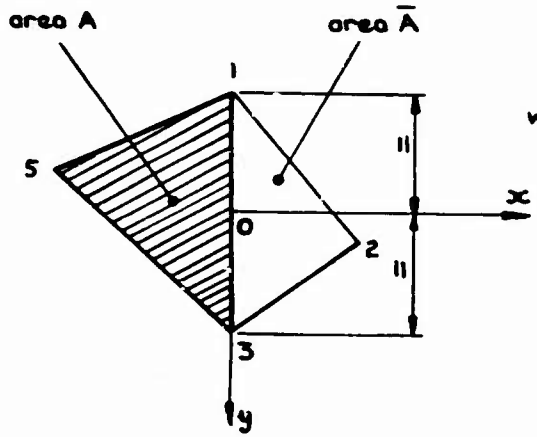


Fig. 5a Temporary change of coordinate axes

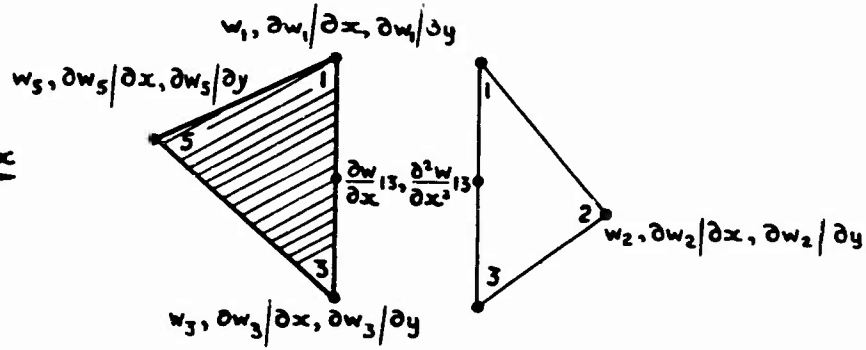


Fig. 5b Connection properties for extended interpolation procedure

Fig. 5 Coordinate axes and connection properties for extended interpolation procedure

$$\left.\begin{aligned}
\partial^2 N_1/\partial x^2 &= -3y_3(y_3 - y_5)/A^2, & \partial^2 N_3/\partial x^2 &= -3y_3(y_3 + y_5)/A^2,\\
\partial^2 N_1/\partial x^2 &= -y_5/A, & \partial^2 N_{3x}/\partial x^2 &= y_5/A,\\
\partial^2 N_{1y}/\partial x^2 &= -y_3(3y_3 + y_5)(y_3 - y_5)/2A^2, & \partial^2 N_{3y}/\partial x^2 &= y_3(3y_3 - y_5)(y_3 + y_5)/2A^2,\\
\partial^2 N_5/\partial x^2 &= 6y_3^2/A^2, & & \\
\partial^2 N_{5x}/\partial x^2 &= 2y_3/A, & \partial^2 N_{13}/\partial x^2 &= 4y_3/A.\\
\partial^2 N_{5y}/\partial x^2 &= -2y_3^2 y_5/A^2, & &
\end{aligned}\right\} \quad (2\text{-}7)$$

In an entirely similar way for the triangle 1 2 3 of Fig.5b the cubic $w(x,y)$ is defined by

$$\begin{aligned}
w(x,y) = {} & \bar{N}_1 w_1 + \bar{N}_{1x}\partial w_1/\partial x + \bar{N}_{1y}\partial w_1/\partial y\\
& + N_2 w_2 + N_{2x}\partial w_2/\partial x + N_{2y}\partial w_2/\partial y\\
& + \bar{N}_3 w_3 + \bar{N}_{3x}\partial w_3/\partial x + \bar{N}_{3y}\partial w_3/\partial y + N_{31}\partial w_{31}/\partial x
\end{aligned} \quad (2\text{-}8)$$

where the formulae for the shape functions are like those given in equation (2-5).

A satisfactorily high level of continuity is achieved across the interface 1 3 of the two triangles shown in Fig.5b by ensuring that, at the mid-side,

$$\partial^2 w_{13}/\partial x^2 = \partial^2 w_{31}/\partial x^2 \quad (2\text{-}9)$$

in addition to the continuities which equations (2-4) and (2-8) already provide across this interface. This equation (2-9) serves to define the interpolation formula for $\partial w_{13}/\partial x$ because, on substitution from equations (2-4) and (2-8), we obtain

$$I_{13}\partial w_{13}/\partial x = I_{13}\partial w_{31}/\partial x = \sum_{j=1,2,3,5} (I_j w_j + I_{jx}\partial w_j/\partial x + I_{jy}\partial w_j/\partial y) \quad (2\text{-}10)$$

where the constants $I_{13}, I_1, I_{1x}, \ldots$ are explicitly

$$\left.\begin{aligned}
I_{13} &= 4(1/A + 1/\bar{A}),\\
I_1 &= 3\{(y_3 - y_5)/A^2 + (y_2 - y_3)/\bar{A}^2\}, & I_{1x} &= (y_5/A + y_2/\bar{A})/y_3,\\
I_{1y} &= (3y_3 + y_5)(y_3 - y_5)/2A^2 + (y_2 + 3y_3)(y_2 - y_3)/2\bar{A}^2,\\
I_2 &= 6y_3/\bar{A}^2, & I_{2x} &= -2/\bar{A}, & I_{2y} &= -2y_2 y_3/\bar{A}^2,\\
I_3 &= 3\{(y_3 + y_5)/A^2 - (y_2 + y_3)/\bar{A}^2\}, & I_{3x} &= -I_{1x},\\
I_{3y} &= -(3y_3 - y_5)(y_3 + y_5)/2A^2 - (y_2 - 3y_3)(y_2 + y_3)/2\bar{A}^2,\\
I_5 &= -6y_3/A^2, & I_{5x} &= -2/A, & I_{5y} &= 2y_3 y_5/A^2,
\end{aligned}\right\} \quad (2\text{-}11)$$

where $\bar{A}$ is the area of the triangle 1 2 3

$$2\bar{A} = 2x_2y_3 \,. \qquad (2\text{-}12)$$

It can be confirmed that equation (2-10) provides the exact value for $\partial w_{13}/\partial x$ whenever $w(x,y)$ is a general surface cubic and, moreover, when the temporary reorientation of the coordinate axes is reversed, see equation (2-3), it is a simple matter then to derive the corresponding estimate for the normal derivative $\partial w_{13}/\partial n$.

The interpolation formulae for the remaining mid-side normal derivatives $\partial w_{35}/\partial n$ and $\partial w_{51}/\partial n$ of the focal triangle 1 3 5 may be written down in like manner. While shortage of space does not allow details to be given here, a matrix interpolation scheme can be arranged to eliminate all traces of $\partial w_{13}/\partial n$, $\partial w_{35}/\partial n$ and $\partial w_{51}/\partial n$ from the element stiffness matrix. Such a scheme has been set up and applied by Morley and Merrifield[3] to the conforming plate bending triangular element with cubically varying displacement w. Their element is essentially that described by Bazeley *et al*[2] and employs rational functions to supplement the ten term cubic description of the displacement $w(x,y)$ within the finite element; their bending moments are, however, calculated on a novel basis which makes use of a homogeneous equilibrium field within the element. Numerical results of calculations for the classical problems of the square plate are listed in Tables 1 and 2 where only a quarter of the plate is considered with the finite element mesh as shown in Fig.1; the plate has side length L and Poisson's ratio $\nu = 0.3$. It is to be remarked that there is little difference between the results which are secured by retaining the mid-side connection property and those obtained here with the aid of extended interpolation. What is more, the size of the global stiffness matrix now more nearly resembles that which is required for the non-conforming element of Bazeley *et al*[2] whose comparative results for the displacements are also quoted in Table 1. The results from extended interpolation are, incidentally, in good agreement with the exact values.

3. PLATE BENDING ELEMENT WITH QUADRATICALLY VARYING w-DISPLACEMENT

It is shown elsewhere[8] that if a straightforward application of the familiar theorem of minimum potential energy is made with a plate bending triangular element where the displacement $w(x,y)$ varies quadratically and where the (non-conforming) connection properties are as shown in Fig.2a, then this is exactly equivalent to a correct application of the theorem of minimum complementary energy. Furthermore, since this constant bending moment equilibrium element is used in conjunction with a complete set of coordinate/shape functions there is also an assuredness of convergence to the exact solution during progression to successively finer meshes. Encouraged by the success of our previous application of the idea of extended interpolation, it is tempting to enquire whether there is benefit to be gained here by eliminating the mid-side connection points of Fig.2a so as to provide the simplest possible connection properties for a plate bending element, i.e. those which concern merely the value of the displacement w at the three nodes of the triangle as is illustrated in Fig.2b. In the examination which is given below it is convenient, especially in view of the algebraic simplicity, to provide a complete derivation of the element pseudo-stiffness matrix. It should be noted, however, that in the equilibrium model the connection quantities,

Table 1

CENTRAL DEFLECTION OF A SQUARE PLATE USING CONFORMING PLATE BENDING TRIANGULAR ELEMENTS WITH CUBICALLY VARYING EDGE DISPLACEMENT w

	With extended inter-polation	With mid-side connection points	Non-conforming element[2]	Exact	Multiplier
UDL simply supported	0.004065	0.004064	0.00405	0.00406	q_0L^4/D
UDL clamped	0.001257	0.001258	0.00134	0.00126	q_0L^4/D
conc. load simply supported	0.01151	0.01152	0.01165	0.0116	PL^2/D
conc. load clamped	0.005488	0.005494	0.00572	0.00560	PL^2/D
global stiffness matrix					
rows	75	131	75		
semi-bandwidth	36	35	30		

Table 2

BENDING MOMENTS IN A SQUARE PLATE USING CONFORMING PLATE BENDING TRIANGULAR ELEMENTS WITH CUBICALLY VARYING EDGE DISPLACEMENT w

	Centre of side	Centre of plate		Corner reaction	Multiplier
	M_n	M_x	M_y	$\lvert 2M_{xy} \rvert$	
UDL simply supported		0.0488 0.0493* (0.0479)	0.0500 0.0502* (0.0479)	0.0688 0.0686* (0.065)	q_0L^2
UDL clamped	-0.0481 -0.0476* (-0.0513)	0.0240 0.0240* (0.0231)	0.0245 0.0243* (0.0231)		q_0L^2
conc. load simply supported				0.126 0.126* (0.122)	P
conc. load clamped	-0.1199 -0.1172* (-0.1257)				P

The first value, in each case, is obtained with the aid of extended interpolation. *The asterisked value is obtained with the retention of the mid-side connection property. (The value in parentheses is exact).

The numerical results in Tables 1 and 2 are obtained by considering a quarter of the square plate where the finite element mesh is as shown in Fig.1.

like w_1 and $\partial w_{15}/\partial n$, are really Lagrangian multipliers which serve to enforce traction continuity across the element interfaces and, consequently, their estimation by other than the strict variational process generally undermines this traction continuity as well as the bounded property of the variational principle.

The triangular finite element of arbitrary shape is shown in Fig.2 where it is now expedient to number the nodes 1 5 9. A point within the element is located by the non-dimensional area coordinates L_1, L_5 and L_9 where, cf. equation (2-1),

$$\left.\begin{aligned} L_1 &= (a_1 + b_1 x + c_1 y)/2A, \\ a_1 &= x_5 y_9 - x_9 y_5, \quad b_1 = y_5 - y_9, \quad c_1 = x_9 - x_5, \end{aligned}\right\} \tag{3-1}$$

with the remaining expressions following by permutation of the suffices. The area A of the triangular element is here

$$2A = b_5 c_9 - b_9 c_5 \ . \tag{3-2}$$

The coordinate s is taken clockwise around the boundary of the element and the length of the side joining nodes 1 and 5 is denoted by s_{15} where

$$s_{15}^2 = b_9^2 + c_9^2 \ . \tag{3-3}$$

The outward pointing normal n from this side subtends the angle γ_{15} with the Ox coordinate axis and it is easily verified that

$$\sin \gamma_{15} = -c_9/s_{15}, \quad \cos \gamma_{15} = -b_9/s_{15} \ . \tag{3-4}$$

The artifice described by Bazeley *et al*[2] is followed so that the actual displacement of the element is described by

$$w = w' + w^R \tag{3-5}$$

where w^R is the rigid body displacement

$$w^R(x,y) = w_1 L_1 + w_5 L_5 + w_9 L_9 \tag{3-6}$$

which contributes nothing to the curvatures and $w'(x,y)$ is the relative displacement of the element when it is regarded as supported at the nodes, i.e.

$$w_1' = w_5' = w_9' = 0 \ . \tag{3-7}$$

The normal slope of the displaced element surface at the mid-side 1 5 is denoted by $\partial w_{15}/\partial n$, see Fig.2a, with

$$\partial w_{15}/\partial n = (\partial w_{15}/\partial x) \cos \gamma_{15} + (\partial w_{15}/\partial y) \sin \gamma_{15} \tag{3-8}$$

where the derivatives $\partial w_{15}/\partial x$ and $\partial w_{15}/\partial y$ also refer to this mid-side point. It follows from equation (3-5), typically for this side, that

$$\partial w_{15}/\partial n = \partial w_{15}'/\partial n + \partial w_{15}^R/\partial n \ . \tag{3-9}$$

The quadratically varying relative deflection w' is now written

$$w'(x,y) = N_{15}\partial w'_{15}/\partial n + N_{59}\partial w'_{59}/\partial n + N_{91}\partial w'_{91}/\partial n \qquad (3\text{-}10)$$

with shape function

$$N_{15}(x,y) = -2AL_9(1 - L_9)/s_{15} \qquad (3\text{-}11)$$

and $N_{59}(x,y)$, $N_{91}(x,y)$ following by permutation of the suffices. The shape function $N_{15}(x,y)$ is zero at all three node points, as is required for support, while the first derivatives

$$\partial N_{15}/\partial x = b_9(2L_9 - 1)/s_{15}, \quad \partial N_{15}/\partial y = c_9(2L_9 - 1)/s_{15} \qquad (3\text{-}12)$$

are zero at the mid-sides 5 9 and 9 1; along the side 1 5, however, they are constant with

$$\partial N_{15}/\partial n = (\partial N_{15}/\partial x)\cos\gamma_{15} + (\partial N_{15}/\partial y)\sin\gamma_{15} = 1 \qquad (3\text{-}13)$$

which follows on using equations (3-3) and (3-4).

The variation of the virtual work contribution by the element is given by

$$\delta U^e = \{\delta q'\}^{eT}[k']^e\{q'\}^e + \{\delta q'\}^{eT}\{F^{*'}\}^e + \{\delta q^R\}^{eT}\{F^{*R}\}^e \qquad (3\text{-}14)$$

where the 3 by 1 column matrices of generalised displacements are

$$\{q'\}^e = (\partial w'_{15}/\partial n \;\; \partial w'_{59}/\partial n \;\; \partial w'_{91}/\partial n)^T, \quad \{q^R\}^e = (w_1 \;\; w_5 \;\; w_9)^T \; . \qquad (3\text{-}15)$$

The stiffness matrix is most easily derived by way of a preliminary application of Green's theorem to the customary surface integral for the strain energy so that, for the constant thickness element,

$$\{\delta q'\}^{eT}[k']^e\{q'\}^e = -\int M_n \partial\delta w'/\partial n \, ds = -\{\delta q'\}^{eT}(s_{15}M_{n_{15}} \;\; s_{59}M_{n_{59}} \;\; s_{91}M_{n_{91}})^T \; . \qquad \ldots\ (3\text{-}16)$$

This equation enjoys such a simple form because the Kirchhoff force is zero along each side of the element and because the relative displacement $w'(x,y)$ satisfies $\nabla^4 w' = 0$ over the whole surface of the element as well as $w' = 0$ at the nodes. Typically, $M_{n_{15}}$ is the normal bending moment which is acting on the side 1 5 and is calculated by substituting equations (3-10) into the usual moment-displacement relations together with

$$M_{n_{15}} = M_x \cos^2\gamma_{15} + M_y \sin^2\gamma_{15} - 2M_{xy}\sin\gamma_{15}\cos\gamma_{15}. \qquad (3\text{-}17)$$

It follows for the particular case of the isotropic plate that the first three constituents of the 3 by 3 relative stiffness matrix $[k']^e$ are given by

$$\left.\begin{aligned} k'_{11} &= Ds^2_{15}/A, \\ k'_{12} &= k'_{21} = D\{b_1^2(b_9^2+\nu c_9^2) + c_1^2(c_9^2+\nu b_9^2) + 2(1-\nu)b_1b_9c_1c_9\}/As_{59}s_{15}, \\ k'_{13} &= k'_{31} = D\{b_5^2(b_9^2+\nu c_9^2) + c_5^2(c_9^2+\nu b_9^2) + 2(1-\nu)b_5b_9c_5c_9\}/As_{91}s_{15}, \end{aligned}\right\} \qquad (3\text{-}18)$$

where substitutions are made from equations (3-4) and the flexural rigidity is defined in the usual way by $D = Eh^3/12(1 - \nu^2)$ with E as the Young's modulus, h the plate thickness and ν the Poisson's ratio. The element load matrices $\{F^{*'}\}^e$ and $\{F^{*R}\}^e$ are concerned only with the contributions which arise from applied bending moments at the plate boundary; the virtual work from any applied normal force is more easily incorporated at global level. As an illustrative example we consider the situation where only one side, e.g. 1 5, of the element coincides with the plate boundary where the mid-side normal bending moment is prescribed as

$$M_{n_{15}} = M^*_{n_{15}} \quad . \tag{3-19}$$

The element load matrices are then easily found to be

$$\left.\begin{aligned} \{F^{*'}\}^e &= s_{15} M^*_{n_{15}} (1 \quad 0 \quad 0)^T, \\ \{F^{*R}\}^e &= -(M^*_{n_{15}}/2A)(b_1 b_9 + c_1 c_9 \quad b_5 b_9 + c_5 c_9 \quad s_{15})^T. \end{aligned}\right\} \tag{3-20}$$

The normal derivatives of the relative displacement, i.e. the $\partial w'/\partial n$, at the mid-side points of the triangular element are now to be estimated in terms of the actual deflection w at the nodes of nearby elements. The basic topology for this extended interpolation is slightly more complicated than that which was required in section 2; it is described by twelve node numbers 1 2 3 4 5 6 7 8 9 10 11 12 as is shown in Fig.6 where the focal element is picked out by the node numbers 1 5 9. Variations on the topology are permitted provided that at least four triangular elements share each node which does not lie on the plate boundary; the topology is arbitrary for nodes actually on the boundary. The extended interpolation is to be arranged in such a way that exact values are derived for the mid-side normal derivatives $\partial w/\partial n$ in all cases where the true global displaced shape $w(x,y)$ is describable in terms of the general surface quadratic; moreover, the value of $\partial w/\partial n$ is required to be uniquely determined in any progression to an adjacent focal element. The Fortran computer subroutine, which is listed in the Appendix, serves also to define the basic topology for this new extended interpolation; when it is provided with a table of focal element node numbers like

```
1   0   0   0   2   0   0   0   9   0   0   0
2   0   0   0   8   0   0   0   9   0   0   0
9   0   0   0   8   0   0   0  13   0   0   0
etc.
```

which are again appropriate to the situation shown in Fig.1c, it calculates the required table of node numbers

```
1   0   0   0   2   3   8  13   9  12  10   0
2   0   3   7   8  14  13  12   9  10   1   0
9   1   2   3   8   7  14  17  13  18  12  10
etc..
```

Let us begin by considering the component six-point topology 1 3 5 7 9 11 which completely surrounds the focal element 1 5 9 as is shown in Fig.6b. Equation (3-10) immediately provides the relative deflections w'_3, w'_7 and w'_{11} where

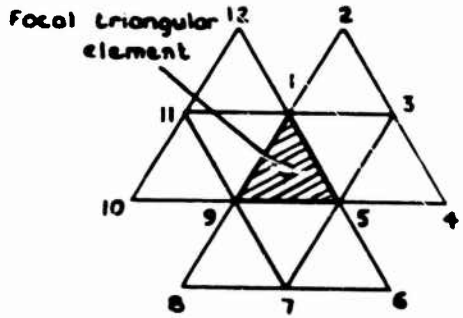


Fig. 6a Twelve point basic topology

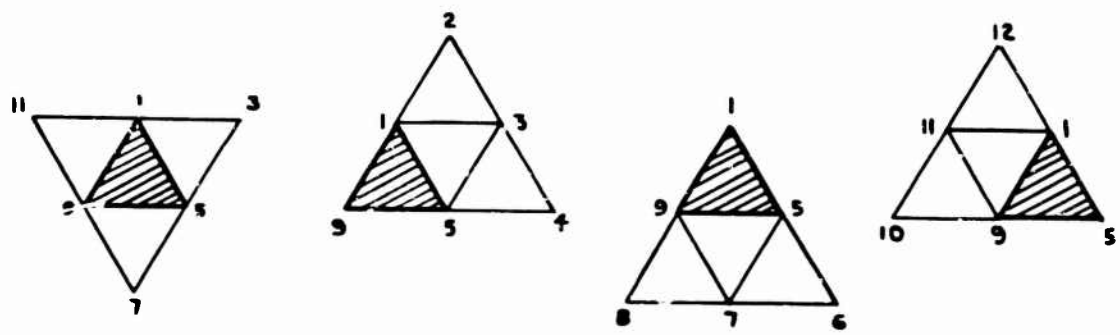


Fig. 6b Component six point topologies

Fig. 6 Topology of extended interpolation for element with quadratically varying displacements

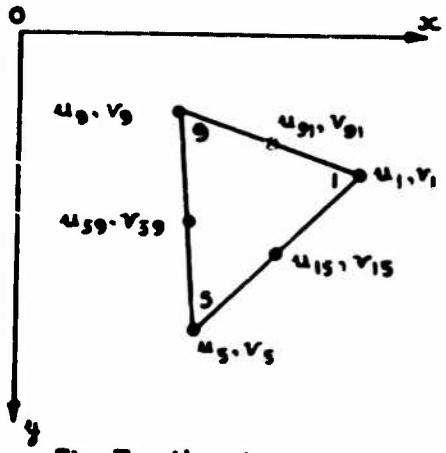


Fig. 7a Usual connection properties

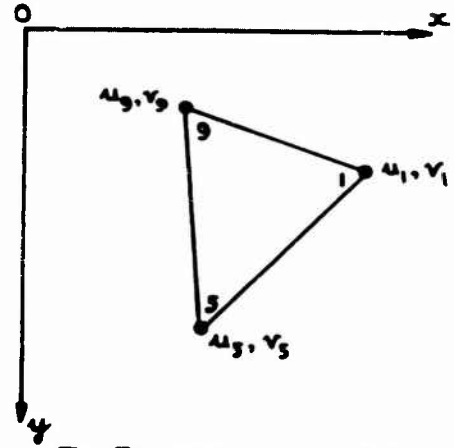


Fig. 7b With extended interpolation

Fig. 7 Connection properties of plane stress element with quadratically varying displacements

$$\begin{Bmatrix} w'_3 \\ w'_7 \\ w'_{11} \end{Bmatrix} = \begin{bmatrix} N_{15}(x_3,y_3) & N_{59}(x_3,y_3) & N_{91}(x_3,y_3) \\ N_{15}(x_7,y_7) & N_{59}(x_7,y_7) & N_{91}(x_7,y_7) \\ N_{15}(x_{11},y_{11}) & N_{59}(x_{11},y_{11}) & N_{91}(x_{11},y_{11}) \end{bmatrix} \begin{Bmatrix} \partial w'_{15}/\partial n \\ \partial w'_{59}/\partial n \\ \partial w'_{91}/\partial n \end{Bmatrix} . \quad (3\text{-}21)$$

The square matrix in this equation is required to be non-singular and it is in this context that the above mentioned topological restriction is important. There are also exceptional arrangements of mesh geometry which must be avoided, like a colinearity of the four nodes numbered 1 3 7 9 in the first of the topological descriptions of Fig.6b. Fortunately, however, the simple requirement that no finite element contains an obtuse angle is sufficient, although by no means necessary, to provide an essentially non-singular matrix. Such restriction on the triangular element shape is of minor practical importance, it is rather one more exhortation towards sound finite element practice. A simple inversion of the 3 by 3 matrix of equation (3-21) now gives

$$\begin{Bmatrix} \partial w'_{15}/\partial n \\ \partial w'_{59}/\partial n \\ \partial w'_{91}/\partial n \end{Bmatrix} = \begin{bmatrix} N_{15}(x_3,y_3) & N_{59}(x_3,y_3) & N_{91}(x_3,y_3) \\ N_{15}(x_7,y_7) & N_{59}(x_7,y_7) & N_{91}(x_7,y_7) \\ N_{15}(x_{11},y_{11}) & N_{59}(x_{11},y_{11}) & N_{91}(x_{11},y_{11}) \end{bmatrix}^{-1} \begin{Bmatrix} w'_3 \\ w'_7 \\ w'_{11} \end{Bmatrix} . \quad (3\text{-}22)$$

While this equation (3-22) provides a reasonable estimate for the value of $\partial w'_{15}/\partial n$, for example, it is noted from the component six-point topology 1 2 3 4 5 9 of Fig.6b that we may derive also

$$\begin{Bmatrix} \partial w'_{15}/\partial n \\ \partial w'_{59}/\partial n \\ \partial w'_{91}/\partial n \end{Bmatrix} = \begin{bmatrix} N_{15}(x_2,y_2) & N_{59}(x_2,y_2) & N_{91}(x_2,y_2) \\ N_{15}(x_3,y_3) & N_{59}(x_3,y_3) & N_{91}(x_3,y_3) \\ N_{15}(x_4,y_4) & N_{59}(x_4,y_4) & N_{91}(x_4,y_4) \end{bmatrix}^{-1} \begin{Bmatrix} w'_2 \\ w'_3 \\ w'_4 \end{Bmatrix} \quad (3\text{-}23)$$

which gives an equally reasonable, albeit generally different, estimate for $\partial w_{15}/\partial n$. Several courses of action are now available but our preference is to take the arithmetic mean of the two estimates with the assurance that the value is then always uniquely estimated, irrespectively of whether the focal triangle is 1 5 9 or 1 3 5. The equations (3-5) and (3-6) show that the relative displacements w'_3, w'_7, w'_{11} of equation (3-22), for example, may subsequently be transformed into actual displacements w by the matrix transformation

$$\begin{Bmatrix} w'_3 \\ w'_7 \\ w'_{11} \end{Bmatrix} = \begin{bmatrix} -L_1(x_3,y_3) & 1 & -L_5(x_3,y_3) & 0 & -L_9(x_3,y_3) & 0 \\ -L_1(x_7,y_7) & 0 & -L_5(x_7,y_7) & 1 & -L_9(x_7,y_7) & 0 \\ -L_1(x_{11},y_{11}) & 0 & -L_5(x_{11},y_{11}) & 0 & -L_9(x_{11},y_{11}) & 1 \end{bmatrix} \begin{Bmatrix} w_1 \\ w_3 \\ w_5 \\ w_7 \\ w_9 \\ w_{11} \end{Bmatrix} . \quad (3\text{-}24)$$

It remains to deal with cases such as where the side 1 5 is coincident with the plate boundary so that the node number 3, see Fig.6, does not even exist. The simplest case occurs when the mid-side kinematic boundary condition is prescribed

$$\partial w_{15}/\partial n = \partial w^{*}_{15}/\partial n \qquad (3\text{-}25)$$

because, from equations (3-5), (3-6) and (3-9) the normal derivative $\partial w'_{15}/\partial n$ of the corresponding relative displacement is then given directly in terms of nodal displacements by

$$\partial w'_{15}/\partial n = -\{(b_1 w_1 + b_5 w_5 + b_9 w_9)\cos\gamma_{15} + (c_1 w_1 + c_5 w_5 + c_9 w_9)\sin\gamma_{15}\}/2A + \partial w^{*}_{15}/\partial n \ . \qquad (3\text{-}26)$$

One way of dealing with the mid-side traction boundary conditions, like the normal bending moment which is prescribed by equation (3-19), is to substitute from the matrix equations (3-16) and (3-20) to obtain, in this instance,

$$\{k'_{11}\partial w'_{15}/\partial n + k'_{12}\partial w'_{59}/\partial n + k'_{13}\partial w'_{91}/\partial n + F^{*'}_1\}\,\delta\partial w'_{15}/\partial n = 0 \ . \qquad (3\text{-}27)$$

Equation (3-21) can then be replaced with

$$\begin{Bmatrix} 0 \\ w'_7 \\ w'_{11} \end{Bmatrix} = \begin{bmatrix} k'_{11} & k'_{12} & k'_{13} \\ N_{15}(x_7,y_7) & N_{59}(x_7,y_7) & N_{91}(x_7,y_7) \\ N_{15}(x_{11},y_{11}) & N_{59}(x_{11},y_{11}) & N_{91}(x_{11},y_{11}) \end{bmatrix} \begin{Bmatrix} \partial w'_{15}/\partial n \\ \partial w'_{59}/\partial n \\ \partial w'_{91}/\partial n \end{Bmatrix} + \begin{Bmatrix} F^{*'}_1 \\ 0 \\ 0 \end{Bmatrix} \qquad \ldots\ (3\text{-}28)$$

where the square matrix is inverted to provide a relationship similar to that of equation (3-22) but with the addition of a column vector of constants.

In what follows, it is convenie to suppose that all the above is summarised into a matrix interpolation scheme for the focal triangular element 1 5 9 in such a way that

$$\{q'\}^e = [I]\{q_I\}^e + \{i\} \qquad (3\text{-}29)$$

where the 3 by 1 column vector of mid-side connection quantities $\{q'\}^e$ is as given by equation (3-15), $[I]$ is a 3 by 12 rectangular matrix of constants, the column vector of nodal connection quantities $\{q_I\}^e$ is defined by

$$\{q_I\}^e = (w_1\ w_2\ w_3\ \ldots\ w_{12})^T, \qquad (3\text{-}30)$$

and $\{i\}$ is a 3 by 1 column vector of constants. Thus, when equation (3-29) is substituted into equation (3-14) the variation δU^e of the virtual work contribution by the element becomes

$$\delta U^e = \{\delta q_I\}^{eT}[I]^T[k']^e[I]\{q_I\}^e + \{\delta q_I\}^{eT}[I]^T[k']^e\{i\} + \{\delta q_I\}^{eT}[I]^T\{F^{*'}\}^e + \{\delta q^R\}^{eT}\{F^{*R}\}^e \qquad (3\text{-}31)$$

which is now in terms only of actual nodal displacements.

Although, as already noted, the interpolation of the mid-side value of the normal derivative $\partial w/\partial n$ destroys the strict equilibrium nature of this plate bending element, the Tables 3 and 4 show that it frequently provides agreeable improvements in the accuracy of the numerical results

Table 3

CENTRAL DEFLECTION OF A SQUARE PLATE USING TRIANGULAR ELEMENTS WITH QUADRATICALLY VARYING w-DISPLACEMENT

	With extended interpolation	With mid-side connection points	Non-conforming element[1]	Exact	Multiplier
UDL simply supported	0.00411 {0.00414}	0.00432 {0.00412}	0.00405	0.00406	q_0L^4/D
UDL clamped	0.00154 {0.00134}	0.00170 {0.00158}	0.00134	0.00126	q_0L^4/D
conc. load simply supported	0.01315 {0.01207}	0.01351 {0.01219}	0.01165	0.0116	PL^2/D
conc. load clamped	0.00727 {0.00611}	0.00776 {0.00628}	0.00572	0.00560	PL^2/D
global stiffness matrix					
rows	25 {81}	81 {289}	75		
semi-bandwidth	20 {36}	21 {37}	30		

{The values in curly brackets refer to a finer elemental mesh where the side lengths are half those shown in Fig.1.}

Table 4

BENDING MOMENTS IN A SQUARE PLATE USING TRIANGULAR ELEMENTS WITH QUADRATICALLY VARYING w-DISPLACEMENT

	Centre of side	Centre of plate		Corner reaction	Multiplier
	M_n	M_x	M_y	$\lvert 2M_{xy} \rvert$	
UDL simply supported		0.0465 0.0471* (0.0479)	0.0436 0.0447* (0.0479)	0.057 0.065* (0.065)	q_0L^2
UDL clamped	-0.0475 -0.0443* (-0.0513)	0.0209 0.0206* (0.0231)	0.0240 0.0235* (0.0231)		q_0L^2
conc. load clamped	-0.1164 -0.1065* (-0.1257)				P

The first value, in each case, is obtained with the aid of extended interpolation. *The asterisked value is obtained with the retention of mid-side connection property. (The value in parentheses is exact.)

Except where stated, the numerical results in Tables 3 and 4 are obtained by considering a quarter of the square plate where the finite element mesh is as shown in Fig.1.

as compared with the equilibrium situation where the mid-side connection points are retained throughout the analysis. The numerical results in these Tables again refer to the classical problems of the square plate where only a quarter portion is considered, firstly with the finite element mesh as is shown in Fig.1 and secondly with a finer mesh where the side lengths are halved; again, the plate has side length L and Poisson's ratio $\nu = 0.3$. Attention is drawn to the fact that the size of the banded global stiffness matrix is now only some thirty per cent of that which is required for retention of the mid-side connection points.

4. PLANE STRESS ELEMENT WITH QUADRATICALLY VARYING DISPLACEMENTS

For our final application of the idea of extended interpolation we turn to the problem of plane stress where the displacement components u and v are assumed to vary quadratically within the triangular finite element. The usual connection properties for this element are illustrated in Fig.7a where the nodes are numbered 1 5 9. Our aim is to eliminate the mid-side connection points with the aid of extended interpolation so as to provide the preferable situation which is shown in Fig.7b where the connection properties coincide with those of the simplest of all plane stress elements in which the displacement components u and v are linearly varying; it is, however, the intention to retain much of the superior accuracy which is associated with the more complicated element. Again, it is convenient to provide a complete rederivation of the element stiffness matrix.

The quadratically varying displacement components u and v may be expressed quite generally in the form

$$\left.\begin{aligned} u(x,y) &= N_1 u_1 + N_5 u_5 + N_9 u_9 + u'(x,y), \\ v(x,y) &= N_1 v_1 + N_5 v_5 + N_9 v_9 + v'(x,y), \end{aligned}\right\} \tag{4-1}$$

where u' and v' are relative displacements defined by

$$\left.\begin{aligned} u'(x,y) &= N_{15} u_{15} + N_{59} u_{59} + N_{91} u_{91}, \\ v'(x,y) &= N_{15} v_{15} + N_{59} v_{59} + N_{91} v_{91}, \end{aligned}\right\} \tag{4-2}$$

and N_1, N_{15} are shape functions given by

$$N_1(x,y) = L_1(2L_1 - 1), \quad N_{15}(x,y) = 4L_1 L_5 , \tag{4-3}$$

with the remaining shape functions following by permutation of the suffices. From equations (4-2) and (4-3) it is noted that the relative displacements enjoy the following properties

$$\left.\begin{aligned} &u'_1 = u'_5 = u'_9 = v'_1 = v'_5 = v'_9 = 0 , \\ &u'_{15} = u_{15}, \quad u'_{59} = u_{59}, \quad u'_{91} = u_{91} . \end{aligned}\right\} \tag{4-4}$$

The generalised strains are defined by

$$\{\varepsilon\} = \begin{Bmatrix} \partial u/\partial x \\ \partial v/\partial y \\ \partial u/\partial y + \partial v/\partial x \end{Bmatrix} = [B \quad B'] \begin{Bmatrix} q \\ q' \end{Bmatrix}^e \tag{4-5}$$

where

$$\{q\}^e = (u_1\ u_5\ u_9\ v_1\ v_5\ v_9)^T, \quad \{q'\}^e = (u_{15}\ u_{59}\ u_{91}\ v_{15}\ v_{59}\ v_{91})^T, \tag{4-6}$$

and the 3 by 6 matrices $[B]$ and $[B']$ are derived by taking partial derivatives of equations (4-1) and (4-2) so that

$$[B] = \begin{bmatrix} \partial N_1/\partial x & \partial N_5/\partial x & \partial N_9/\partial x & 0 & 0 & 0 \\ 0 & 0 & 0 & \partial N_1/\partial y & \partial N_5/\partial y & \partial N_9/\partial y \\ \partial N_1/\partial y & \partial N_5/\partial y & \partial N_9/\partial y & \partial N_1/\partial x & \partial N_5/\partial x & \partial N_9/\partial x \end{bmatrix} \tag{4-7}$$

and

$$[B'] = \begin{bmatrix} \partial N_{15}/\partial x & \partial N_{59}/\partial x & \partial N_{91}/\partial x & 0 & 0 & 0 \\ 0 & 0 & 0 & \partial N_{15}/\partial y & \partial N_{59}/\partial y & \partial N_{91}/\partial y \\ \partial N_{15}/\partial y & \partial N_{59}/\partial y & \partial N_{91}/\partial y & \partial N_{15}/\partial x & \partial N_{59}/\partial x & \partial N_{91}/\partial x \end{bmatrix} \tag{4-8}$$

with, typically,

$$\left.\begin{aligned} \partial N_1/\partial x &= b_1(4L_1 - 1)/2A, & \partial N_1/\partial y &= c_1(4L_1 - 1)/2A, \\ \partial N_{15}/\partial x &= 4(b_1L_5 + b_5L_1)/2A, & \partial N_{15}/\partial y &= 4(c_1L_5 + c_5L_1)/2A, \end{aligned}\right\} \tag{4-9}$$

The variation δU^e of the virtual work contribution by the element is given by

$$\delta U^e = \begin{Bmatrix} \delta q \\ \delta q' \end{Bmatrix}^T \begin{bmatrix} k & k' \\ k'^T & k'' \end{bmatrix}^e \begin{Bmatrix} q \\ q' \end{Bmatrix}^e + \begin{Bmatrix} \delta q \\ \delta q' \end{Bmatrix}^T \begin{Bmatrix} F^* \\ F^{*\prime} \end{Bmatrix}^e \tag{4-10}$$

where the 12 by 12 element stiffness matrix is derived from

$$\begin{bmatrix} k & k' \\ k'^T & k'' \end{bmatrix}^e = \iint_A \begin{bmatrix} [B]^T[D][B] & [B]^T[D][B'] \\ \text{sym} & [B']^T[D][B'] \end{bmatrix} dA \tag{4-11}$$

The 3 by 3 elasticity matrix $[D]$, taking as example an isotropic plate of constant thickness h, is given by

$$[D] = \frac{Eh}{1 - \nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & (1 - \nu)/2 \end{bmatrix} \tag{4-12}$$

and, in evaluating the area integrals of equation (4-11), it is helpful to note the standard forms

$$\iint_A L_1 dA = 2\iint_A L_1^2 dA = 4\iint_A L_1 L_5 dA = A/3. \qquad (4\text{-}13)$$

The 3 by 1 element load matrices $\{F^*\}^e$ and $\{F^{*'}\}^e$ are calculated in the usual manner. By way of illustration, if it is assumed that only the side 1 5 coincides with the plate boundary where the following tractions are prescribed

$$\sigma_{n_1} = \sigma^*_{n_1}, \quad \tau_{ns_1} = \tau^*_{ns_1}, \quad \sigma_{n_5} = \sigma^*_{n_5}, \quad \tau_{ns_5} = \tau^*_{ns_5}, \qquad (4\text{-}14)$$

then the load matrix $\{F^{*'}\}^e$, which is of special interest in the present context of extended interpolation is, in the absence of surface forces,

$$\{F^{*'}\}^e = -(s_{15}/3)\begin{Bmatrix} (\sigma^*_{n_1} + \sigma^*_{n_5}) \cos\gamma_{15} - (\tau^*_{ns_1} + \tau^*_{ns_5}) \sin\gamma_{15} \\ 0 \\ 0 \\ (\sigma^*_{n_1} + \sigma^*_{n_5}) \sin\gamma_{15} + (\tau^*_{ns_1} + \tau^*_{ns_5}) \cos\gamma_{15} \\ 0 \\ 0 \end{Bmatrix} . \qquad (4\text{-}15)$$

The mid-side displacement values like u_{15} and v_{15}, which constitute the column vector $\{q'\}^e$ of equation (4-6), are now to be estimated in terms of the displacements at the nodes of nearby elements. The general details of the mesh topology for the purpose of this plane stress extended interpolation are the same as are described in section 3, see Fig.6. The interpolation is again to be arranged in such a way that exact values are recovered for the mid-side displacements whenever the true displaced state $u(x,y)$, $v(x,y)$ is describable in terms of the general surface quadratic. It is of especial importance, also, that these mid-side displacements are uniquely determined so as to preserve the bounded property which belongs to correct applications of the theorem of minimum potential energy.

The six point component topology which is denoted by node numbers 1 3 5 7 9 11 is shown in Fig.6b to encircle the focal element 1 5 9 and the equations (4-2) for the relative displacements provide

$$\begin{Bmatrix} u'_3 \\ u'_7 \\ u'_{11} \end{Bmatrix} = \begin{bmatrix} N_{15}(x_3,y_3) & N_{59}(x_3,y_3) & N_{91}(x_3,y_3) \\ N_{15}(x_7,y_7) & N_{59}(x_7,y_7) & N_{91}(x_7,y_7) \\ N_{15}(x_{11},y_{11}) & N_{59}(x_{11},y_{11}) & N_{91}(x_{11},y_{11}) \end{bmatrix} \begin{Bmatrix} u_{15} \\ u_{59} \\ u_{91} \end{Bmatrix} . \qquad (4\text{-}16)$$

The sufficient conditions for the square matrix in this equation to be non-singular are identical with those already described for equation (3-21). The 3 by 3 matrix of equation (4-16) may then be inverted to give

$$\begin{Bmatrix} u_{15} \\ u_{59} \\ u_{91} \end{Bmatrix} = \begin{bmatrix} N_{15}(x_3,y_3) & N_{59}(x_3,y_3) & N_{91}(x_3,y_3) \\ N_{15}(x_7,y_7) & N_{59}(x_7,y_7) & N_{91}(x_7,y_7) \\ N_{15}(x_{11},y_{11}) & N_{59}(x_{11},y_{11}) & N_{91}(x_{11},y_{11}) \end{bmatrix}^{-1} \begin{Bmatrix} u'_3 \\ u'_7 \\ u'_{11} \end{Bmatrix} \tag{4-17}$$

While equation (4-17) provides a reasonable estimate for the value of the mid-side u_{15}, for example, the six point topology 1 2 3 4 5 9 of Fig.3b provides

$$\begin{Bmatrix} u_{15} \\ u_{59} \\ u_{91} \end{Bmatrix} = \begin{bmatrix} N_{15}(x_2,y_2) & N_{59}(x_2,y_2) & N_{91}(x_2,y_2) \\ N_{15}(x_3,y_3) & N_{59}(x_3,y_3) & N_{91}(x_3,y_3) \\ N_{15}(x_4,y_4) & N_{59}(x_4,y_4) & N_{91}(x_4,y_4) \end{bmatrix}^{-1} \begin{Bmatrix} u'_2 \\ u'_3 \\ u'_4 \end{Bmatrix} \tag{4-18}$$

which is an equally reasonable estimate. As in section 2, there are several alternatives which are now available but, again, our preference is to take the arithmetic mean of these two estimates for u_{15} noting that this ensures a uniquely determined value irrespectively of whether the focal triangle is 1 5 9 or 1 3 5. Next, the equations (4-1) show that relative displacements like u'_3, u'_7, u'_{11} of equation (4-17) are readily expressed in terms of actual displacements by means of the matrix transformation

$$\begin{Bmatrix} u'_3 \\ u'_7 \\ u'_{11} \end{Bmatrix} = \begin{bmatrix} -N_1(x_3,y_3) & 1 & -N_5(x_3,y_3) & 0 & -N_9(x_3,y_3) & 0 \\ -N_1(x_7,y_7) & 0 & -N_5(x_7,y_7) & 1 & -N_9(x_7,y_7) & 0 \\ -N_1(x_{11},y_{11}) & 0 & -N_5(x_{11},y_{11}) & 0 & -N_9(x_{11},y_{11}) & 1 \end{bmatrix} \begin{Bmatrix} u_1 \\ u_3 \\ u_5 \\ u_7 \\ u_9 \\ u_{11} \end{Bmatrix} . \tag{4-19}$$

The mid-side displacement components v are estimated in a similar way to that just described.

It remains to deal with the situations which arise when a side like 1 5 coincides with the plate boundary so that the node number 3 in the topology of Fig.6 does not exist. The simplest case occurs when kinematic boundary conditions are prescribed along this side for then

$$-u_{15} \sin \gamma_{15} + v_{15} \cos \gamma_{15} = u^*_{s_{15}} \tag{4-20}$$

and/or

$$u_{15} \cos \gamma_{15} + v_{15} \sin \gamma_{15} = u^*_{n_{15}} \tag{4-21}$$

where $u^*_{s_{15}}$ and $u^*_{n_{15}}$ are respectively the prescribed displacement components tangential and normal to the plate boundary at the mid-side point. When tractions are prescribed, it is necessary to select from equation (4-11) for the 12 by 12 element stiffness matrix the rows numbered 7 and/or 10 which, together with the rows numbered 1 and/or 4 of the element load matrix listed in equation (4-15), provide the sole global contributions to the variations δu_{15} and/or δv_{15}. Thus

$$k'_{11}u_1 + k'_{12}u_5 + k'_{13}u_9 + k'_{14}v_1 + k'_{15}v_5 + k'_{16}v_9$$
$$+ k''_{11}u_{15} + k''_{12}u_{59} + k''_{13}u_{91} + k''_{14}v_{15} + k''_{15}v_{59} + k''_{16}v_{91} = -F^{*'}_1$$
.... (4-22)

and/or

$$k'_{14}u_1 + k'_{24}u_5 + k'_{34}u_9 + k'_{44}v_1 + k'_{54}v_5 + k'_{64}v_9$$
$$+ k''_{41}u_{15} + k''_{42}u_{59} + k''_{43}u_{91} + k''_{44}v_{15} + k''_{45}v_{59} + k''_{46}v_{91} = -F^{*'}_4 .$$
.... (4-23)

The requisite equations from (4-20) to (4-23) then provide enough conditions to allow both u_{15} and v_{15} to be expressed in terms of the prescribed quantities and/or the displacement components at nodes 1 5 7 9 11 together with the displacement components at mid-sides 5 9 and 9 1. Attention is drawn to the fact, however, that care needs to be exercised in dealing with these boundary situations in plane stress in order to avoid the occurrence of singular conditions. For complete reassurance it is recommended that no attempt is made to seek extended interpolation contributions from the first two of the component six point topologies in Fig.6b so that, in the present typical case, unmoderated estimates for u_{59}, u_{91} and v_{59}, v_{91} are accepted from the two remaining component topologies. The overall mesh topology can, in consequence, be no longer arbitrary for nodes actually on the boundary, indeed the recommendation requires that <u>each boundary node is at the conjunction of at least three triangular finite elements</u>.

In that which follows, it is convenient to assume that the above is summarised by a matrix interpolation scheme for the focal triangular element 1 5 9 in such a way that

$$\{q'\}^e = [I]\{q_I\}^e + \{i\} \qquad (4\text{-}24)$$

where the column vector $\{q'\}^e$ is as described by equation (4-6), $[I]$ is a 6 by 24 matrix of constants with

$$\{q_I\}^e = (u_1\ u_2\ u_3 \cdots u_{12}\ v_1\ v_2\ v_3 \cdots v_{12})^T \qquad (4\text{-}25)$$

and $\{i\}$ is a 6 by 1 column vector of constants. It is worth noting that when no nodes are missing from the twelve point basic topology of Fig.6, such as in the case of focal triangles which do not touch the plate boundary, when the 6 by 24 matrix $[I]$ may be more conveniently written as

$$[I] = \begin{bmatrix} I' & 0 \\ 0 & I \end{bmatrix} \qquad (4\text{-}26)$$

where the rectangular matrix $[I']$ is a 3 by 12 matrix of constants. By expanding the matrices of equation (4-10) we obtain the following form for the variation δU^e of the strain energy

$$\delta U^e = \{\delta q\}^{eT}[k]^e\{q\}^e + \{\delta q'\}^{eT}[k']^{eT}\{q\} + \{\delta q\}^{eT}[k']^e\{q'\} + \{\delta q'\}^{eT}[k'']^e\{q'\} + \{\delta q\}^{eT}\{F^*\}^e + \{\delta q'\}^{eT}\{F^{*\prime}\}^e \quad (4\text{-}27)$$

where it is now a simple matter to substitute for $\{q'\}^e$ from equation (4-24) and for $\{\delta q'\}^{eT}$ from

$$\{\delta q'\}^{eT} = \{\delta q_I\}^{eT}[I]^T \quad (4\text{-}28)$$

in order to derive the preferred form for the variation.

The numerical example concerns a square plate of side length L which is under uniform tension of σ_n^* per unit run applied parallel to the Ox axis and is weakened by a central circular hole of diameter L/2. The Poisson's ratio is taken as $\nu = 0.3$. A relatively coarse finite element mesh, see Fig.8, is employed to generate the numerical values as listed in Table 5 which contains also comparisons from other methods of solution. This problem provides a particularly severe test of any numerical technique because of the severity of the stress gradients which occur in the vicinity of the hole and also across the minimum section where there is a substantial compressive stress at the external boundary. Indeed, a comparison between the results from the continuous function solution for the infinitely long strip due to Howland[12] with that for the square plate due to Hengst[13] is illuminating in this context. Table 5 shows that whereas our finite element results for quadratically varying u,v displacements with extended interpolation provide an approximation to the more accurate finite element solution where the mid-side connection quantities are retained, it is clear that the finite element results for linearly varying u,v displacements are virtually worthless for this relatively coarse mesh. The size of the banded global stiffness matrix is seen to be less than 25 per cent of that which is required when the mid-side connections are retained; the optimal semi-bandwidths which are quoted in Table 5 are calculated by a method developed by Morley and Merrifield[14].

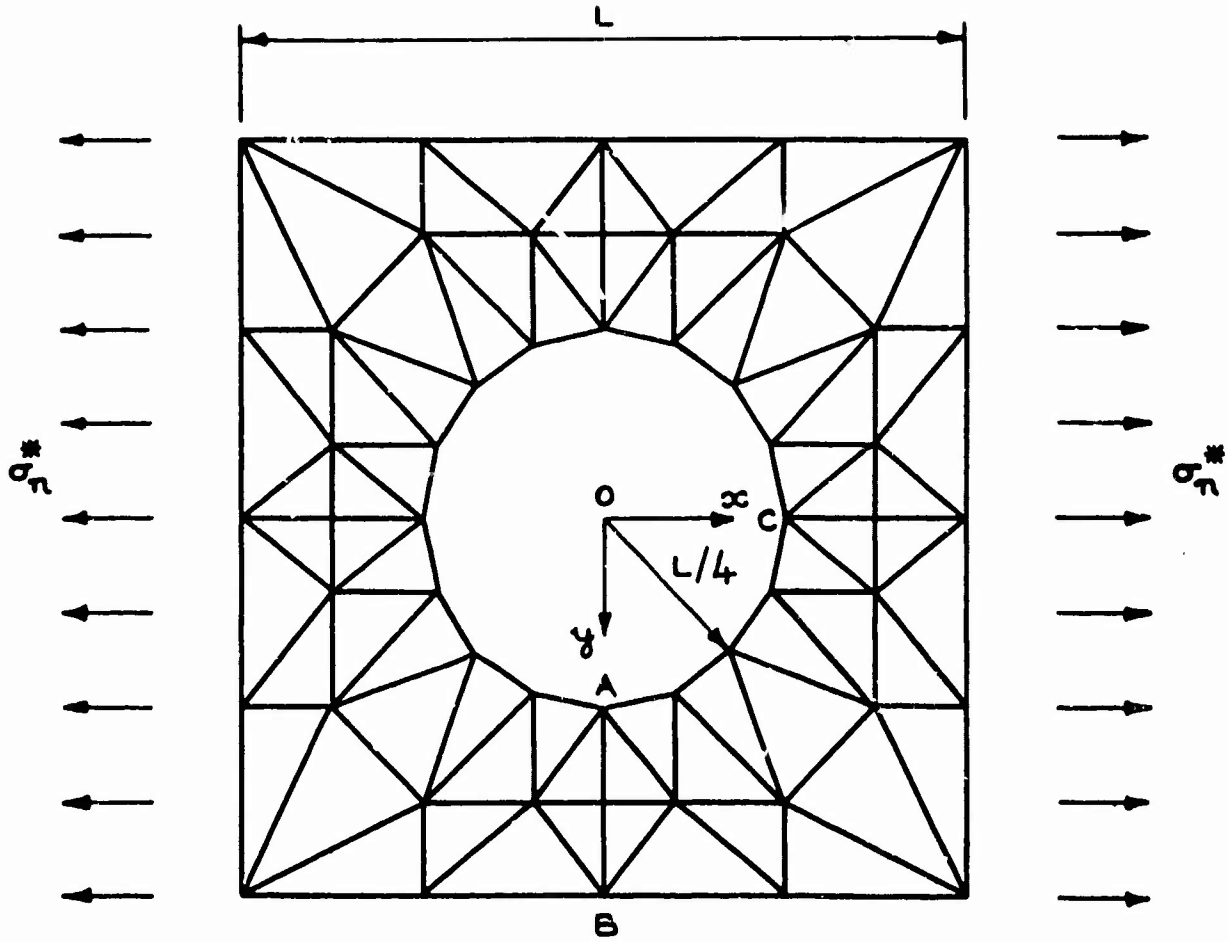


Fig. 8 Square plate under uniform tension with central circular hole

Table 5

STRESSES AND DISPLACEMENTS IN A SQUARE PLATE, UNDER UNIFORM TENSION, WITH CENTRAL CIRCULAR HOLE

	FINITE ELEMENT SOLUTIONS			CONTINUOUS FUNCTION SOLUTIONS		
	Quadratic u,v with extended interpolation	Quadratic u,v with mid-side connection points	Linear u,v	Infinite strip (Howland[12])	Square plate (Hengst[13])	Multiplier
σ_{s_A}	5.36	6.23	3.94	4.32	6.33	σ_n^*
σ_{s_B}	-0.42	-0.70	0.64	0.73		σ_n^*
σ_{s_C}	-2.81	-3.70	-1.41	-1.58	-4.11	σ_n^*
v_A	-0.884	-1.095	-0.674			$\sigma_n^* L/E$
u_C	1.525	1.734	1.233			$\sigma_n^* L/E$
global stiffness matrix						
rows	104	352	104			
optimal semi-bandwidth	40	48	16			

The displacements are appropriate to the condition that the centre of the hole does not translate or rotate.

Appendix A

FORTRAN SUBROUTINE FOR THE EXTENDED INTERPOLATION TOPOLOGY

A listing is now presented of a Fortran computer subroutine which, when it is given the basic information to define the focal-triangular elements, goes on to provide the topological description in the integer double array ELNO(NEL,JTYPE) which enables the various processes of extended interpolation to be readily performed. The notation is here

NEL = number of triangular finite elements in the plate
JTYPE = 6 or 12 depending upon the type of extended interpolation.

Thus, in the case of section 2 where a six-point extended interpolation is required for the example which is illustrated in Fig.1c, the subroutine requires node numbers in the integer double array ELNO(32,6) which provide the description of the focal elements, e.g.

1	0	2	0	9	0
2	0	8	0	9	0
9	0	8	0	13	0

etc.

The calling statement is then

CALL FILL ELNO 6 OR 12 (32,ELNO,6)

for the subroutine to return the following topological description in ELNO(32,6)

1	0	2	8	9	10
2	3	8	13	9	1
9	2	8	14	13	12

etc.

The 0 in this last table signifies that the side 1 2 coincides with the plate boundary. The procedure is similar for the twelve point topologies which are considered in sections 3 and 4.

```
      SUBROUTINE FILL ELNO 6 OR 12(NEL,ELNO,JTYPE)
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
C   NEL = NUMBER OF ELEMENTS.
C  ELNO = INTEGER DOUBLE ARRAY OF NODE NUMBERS TAKEN IN ORDERED
C         SEQUENCE. THE SUBROUTINE MUST BE SUPPLIED WITH AN ELNO WHICH
C         PRESCRIBES THE NODE NUMBERS FOR ONE FOCAL ELEMENT TO EACH ROW
C         OF JTYPE NUMBERS, EACH OF THESE FOCAL NODE NUMBERS IS FOLLOWED
C         BY ZEROS (ONE IF JTYPE=6, THREE IF JTYPE=12).
C JTYPE = PRESCRIBED AS 6 OR 12 TO FILL ELNO 6 OR 12 RESPECTIVELY.
      INTEGER ELNO(NEL,JTYPE),ELROW,ELCOL
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
      INTEGER JOIN(25,4)
C     NOTE: IF THE NUMBER OF ELEMENTS WITH A COMMON NODE EXCEEDS 25
C     THE FIRST DIMENSION OF JOIN MUST BE REDIMENSIONED ACCORDINGLY.
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
      INTEGER IDENT(500)
      DO 1 I=1,500
C     NOTE: IF THE LARGEST NODE NUMBER EXCEEDS 500 THEN THE LAST TWO
C     STATEMENTS MUST BE REDIMENSIONED ACCORDINGLY.
C * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
    1 IDENT(I)=1
      JK=4
      IF(JTYPE.EQ.6)JK=2
      DO 5 ELROW=1,NEL
      DO 5 ELCOL=1,1+2*JK,JK
      NODE=ELNO(ELROW,ELCOL)
      IF(IDENT(NODE).EQ.0)GO TO 5
      IDENT(NODE)=0
      NJOINS=0
      DO 2 I=ELROW,NEL
      DO 2 J=1,1+2*JK,JK
      IF(ELNO(I,J).NE.NODE)GO TO 2
      NJOINS=NJOINS+1
      JOIN(NJOINS,1)=I
      JOIN(NJOINS,2)=J
      J4=J+JK
      IF(J4.GT.JTYPE)J4=J4-JTYPE
      JOIN(NJOINS,3)=ELNO(I,J4)
      J8=J+2*JK
      IF(J8.GT.JTYPE)J8=J8-JTYPE
      JOIN(NJOINS,4)=ELNO(I,J8)
    2 CONTINUE
      DO 5 K=1,NJOINS
      I=JOIN(K,1)
      N=JOIN(K,2)
      N8=N+2*JK
      IF(N8.GT.JTYPE)N8=N8-JTYPE
      IN4=JOIN(K,3)
      IN8=JOIN(K,4)
      DO 4 KA=1,NJOINS
      IF(K.EQ.KA)GO TO 4
      J3=JOIN(KA,3)
      J4=JOIN(KA,4)
      IF(JTYPE.EQ.6)GO TO 3
      IF(IN4.EQ.J4)ELNO(I,N+2)=J3
      IF(IN8.EQ.J3)ELNO(I,N8+2)=J4
      GO TO 4
    3 IF(IN4.EQ.J4)ELNO(I,N+1)=J3
      IF(IN8.EQ.J3)ELNO(I,N8+1)=J4
    4 CONTINUE
      IF(JTYPE.EQ.6)GO TO 5
      N10=N+10
      IF(N10.GT.JTYPE)N10=N10-JTYPE
      IN2=ELNO(I,N+2)
      IN10=ELNO(I,N10)
      DO 5 KA=1,NJOINS
      IF(K.EQ.KA)GO TO 5
      J3=JOIN(KA,3)
      J4=JOIN(KA,4)
      IF(IN2.EQ.J4)ELNO(I,N+1)=J3
      IF(IN10.EQ.J3)ELNO(I,N10+1)=J4
    5 CONTINUE
      RETURN
      END
```

REFERENCES

No.	Author(s)	Title, etc.
1	R. W. Clough J.L. Tocher	Finite element stiffness matrices for analysis of plates in bending. Proc. Conf. Matrix Methods in Struct. Mech. AFFDL-TR-66-80, 515 (1966)
2	G. P. Bazeley Y. K. Cheung B. M. Irons O. C. Zienkiewicz	Triangular elements in plate bending - conforming and non-conforming solutions. Proc. Conf. Matrix Methods in Struct. Mech., AFFDL-TR-66-80, 547 (1966)
3	L. S. D. Morley B. C. Merrifield	On the conforming cubic triangular element for plate bending. In preparation.
4	K. Bell	Triangular plate bending elements. In: Finite Element Methods in Stress Analysis, edited by Holand and Bell. Tapir, 213 (1969)
5	L. R. Herrmann	Finite-element bending analysis for plates. Jour. Engng. Mech. Div. Am. Soc. Civ. Engrs., 93 EM5, 13 (1967)
6	K. Hellan	Analysis of elastic plates in flexure by a simplified finite element method. Acta Polytechnica Scandinavica. Civ. Engng. Bldg. Constructn. Series, No.46 (1967)
7	D. J. Allman	Triangular finite elements for plate bending with constant and linearly varying bending moments. IUTAM Symp. High-speed Computing Elastic Structures (Liège) (1970)
8	L. S. D. Morley	The constant-moment plate bending element. Jour. Strain Analysis, 6, 20 (1971)
9	T. Fujino	Analyses of hydrodynamic and plate structure problems by finite element methods. Japan - US Semin. Matrix Meth. Struct. Analysis Design, (Tokyo) (1969)
10	J. H. Argyris	Triangular elements with linearly varying strain for the matrix displacement method. Jour. Roy. Aero. Soc., 69, 711 (1965)
11	B. Fraeijs de Veubeke	Displacement and equilibrium methods in the finite element method. In: Stress Analysis, edited by O. C. Zienkiewicz and G. S. Hollister, Wiley, 145 (1965)
12	R. C. J. Howland	On the stresses in the neighbourhood of a circular hole in a strip under tension. Phil. Trans. Roy. Soc. London, A, 229, 49 (1930)
13	H. Hengst	Beitrag Zur Beurteilung des Spannungszustandes einer gelochten Scheibe. Ztschr.f.angew Math.und Mech., 18, 44 (1938)
14	L. S. D. Morley B. C. Merrifield	Bandwidth reduction of label sequences as encountered in finite element calculations. Internal Report of the Royal Aircraft Establishment, Farnborough (1970)

SESSION 4. DYNAMICS AND OPTIMIZATION

Session Chairman

R. J. Melosh*

Philco-Ford Corp.
Palo Alto, California

*now with Virginia Polytechnic Institute and State University, Blacksburg, Va.

The Transient Dynamic Analysis of Thin Shells by the Finite Element Method*

Samuel W. Key
Zelma E. Beisinger
Sandia Laboratories
Albuquerque, New Mexico 87115

A simple central difference time integration scheme along with a diagonal mass matrix is used to examine the transient dynamic response of linear elastic thin shells. The numerical stability of the explicit time integration scheme is examined and two inequalities are provided, the strongest of which defines a critical time step. Only calculations using a time step less than the critical time step are meaningful. Above the critical time step, the calculations diverge. The method is called conditionally stable. One inequality is based on membrane behavior and the other on bending behavior. Both of them are in terms of minimum mesh dimensions, shell densities and membrane and bending moduli. The diagonal mass matrix is generated from a consistent mass matrix in a rational manner. Three calculations are included to show the results of this work.

*This work was supported by the United States Atomic Energy Commission.

The Transient Dynamic Analysis of Thin Shells by the Finite Element Method

1. INTRODUCTION

In many respects, the transient dynamic linear analysis of thin shells is well in hand. There is a need, however, for continued improvement in computational speed while maintaining the accuracy of the results, particularly with two dimensional meshes being introduced and larger and larger problems being considered. Two methods are available to obtain improved computational times. A diagonal mass matrix can be used in place of the consistent or non-diagonal matrix prescribed by the finite element method, [1,2], and explicit time integration schemes in place of the implicit schemes can be used.

In the finite element literature, there is no clear cut preference for any one form of mass matrix nor any one method of integration. Although, the unconditionally stable implicit time integration schemes have won many adherents.

Wilson and Clough[3] adopted an implicit linear acceleration scheme, one of the Newmark β methods. It is a conditionally stable scheme. Clough and Felippa[4] in examining plate vibration problems conclude that lumped mass matrices are as accurate or more accurate than the non-diagonal mass matrices prescribed by the finite element method. In their lumped mass calculations, only the translational inertias are retained with the rotations eliminated by static condensation. Clough[5] in a survey of their work reiterates their use of the linear acceleration integration scheme. The paper contains a clear statement of their preference of the lumped mass matrix in place of the consistent mass matrix. Clough and Wilson[6] refer to a new linear acceleration scheme to integrate in time. It is the previous scheme coupled with a predictor-corrector scheme for the accelerations. The result is an unconditionally stable method.

Klein and Sylvester[7] adopt the unconditionally stable implicit time integration scheme of Chan, Cox and Benfield[8] along with the non-diagonal mass matrix of the finite element method.

Stricklin, et al.,[9] selected the unconditionally stable implicit time integration scheme of Houbolt[10], using a non-diagonal mass matrix for transient calculations. For the calculation of mode shapes and frequencies, they use a diagonal mass matrix, retaining the rotational degrees of freedom in the deflections normal to the shell,[11]. Their procedure

for generating the diagonal mass matrix is the same as the one adopted below but they provide no rationale for the choice of the scale factor introduced.

Fu[12] advocates the use of the de Vogelaere[13] method for integrating in time. The method is again an unconditionally stable implicit scheme.

Olson and Lindberg[14] in examining the vibration behavior of curved plates use a consistent mass matrix. Part of their results are obtained neglecting inplane inertia. Whether these degrees of freedom are eliminated by static condensation or simply set to zero to give a bending only response is unclear. The study is impressive.

Greene, Jones and Strome[15] in examining the vibrational modes of cylindrical panels consider both consistent and lumped mass matrices. In both cases, various freedoms are retained with the others being eliminated by static condensation. The result is a preference for the diagonal mass matrix. In none of their lumped mass calculations were rotational degrees of freedom retained.

One of the more informative pieces of work in the finite element literature on numerical time integration is a paper by Nickel[16], the result of an earlier entanglement with an unconditionally unstable implicit scheme.

Goudreau[17] has an extensive treatment of the behavior of various mass matrices in membrane and bending behavior along with an examination of several methods of integration in time. A diagonal mass matrix and a conditionally stable implicit time integration scheme are preferred. It is remarked

that the higher frequencies are inaccurate in the discrete equations and should be suppressed in any event. The remark is based on frequency error versus wave length plots for exact time integration. Key and Krieg[18] have re-examined this question in the light of discrete time integration and show that the higher modes may have quite accurate frequencies and should not necessarily be suppressed.

In explicit time integration schemes, the mass matrix M times the vector of accelerations $\ddot{q}$ occurs, $M\ddot{q}$. In order to find the accelerations which are used to advance the velocities and the displacements, this set of equations must be solved. If M is a diagonal mass matrix, then the solution is trivial. In the finite difference literature, the majority of applications treat only diagonal mass matrices. Implicit integration schemes involve the inverse of a weighted sum of the mass and stiffness matrices, $(b_1M + b_2K)^{-1}$, to find the accelerations. The coefficients b_1 and b_2 depend on the particular implicit scheme being used. If only the non-diagonal mass matrix prescribed by the finite element method is considered, no particular computational advantage is evidenced by either scheme because M^{-1} in the explicit schemes and $(b_1M + b_2K)^{-1}$ in the implicit schemes represent the same amount of computational effort. Thus, the use of a diagonal mass matrix must precede any claim of computational efficiency for the explicit time integration schemes. In the work that follows, a rational approach to generating a diagonal mass matrix from the non-diagonal matrix of the finite element method is discussed.

The major stumbling block to explicit time integration schemes is their conditional stability. Computationally, they are very fast but must use a time step below a certain critical value. If a value above the critical time step is used, then harmonic solutions are represented as exponentially growing solutions and the results are erroneous. Below the critical time step harmonic motion is represented as harmonic motion. The critical time step is invariably related to the shortest transit time between any two nodal points that exist in the finite element mesh. This corresponds to the highest frequency that the system will represent. For impulse problems and suddenly applied temperature and pressure, these frequencies are needed in the solution. Even when implicit schemes are used for these problems, they are used with time steps related to this same criteria just to keep these frequencies present and accurately represent them. They invariably give the same answers as explicit schemes for these problems but at a much greater computational expense. It should be noted that when only the lower modes of response are significant, then unconditionally stable implicit schemes can track them with a large time step while a conditionally stable explicit scheme must remain with what now looks to be a very small time step in relation to the response. However, if a modal solution is used, the explicit schemes again become competitive.

By examining stability, very good estimates of the critical time step are possible. Once this has been done, the explicit

schemes become very reliable and for the problems involving the high frequency response of the shell, they are equally as accurate as the implicit schemes and much faster computationally per time step. The work below considers a simple centered second order difference approximation to the accelerations and provides the critical time step expressions needed to make the computations.

2. STIFFNESS MATRIX

The finite element used for this analysis is a doubly curved arbitrary quadrilateral designed for shells where the reference surface is a portion of an axisymmetric surface. It is based on a minimum potential energy principle and a discrete Kirchhoff hypothesis. There are nine degrees of freedom at each mesh point; the circumferential, meridional and normal displacements along with their first derivatives in the surface variables are carried at each mesh point. The element is pictured in Figure 1. The details of its derivation are well documented in References [19,20,21].

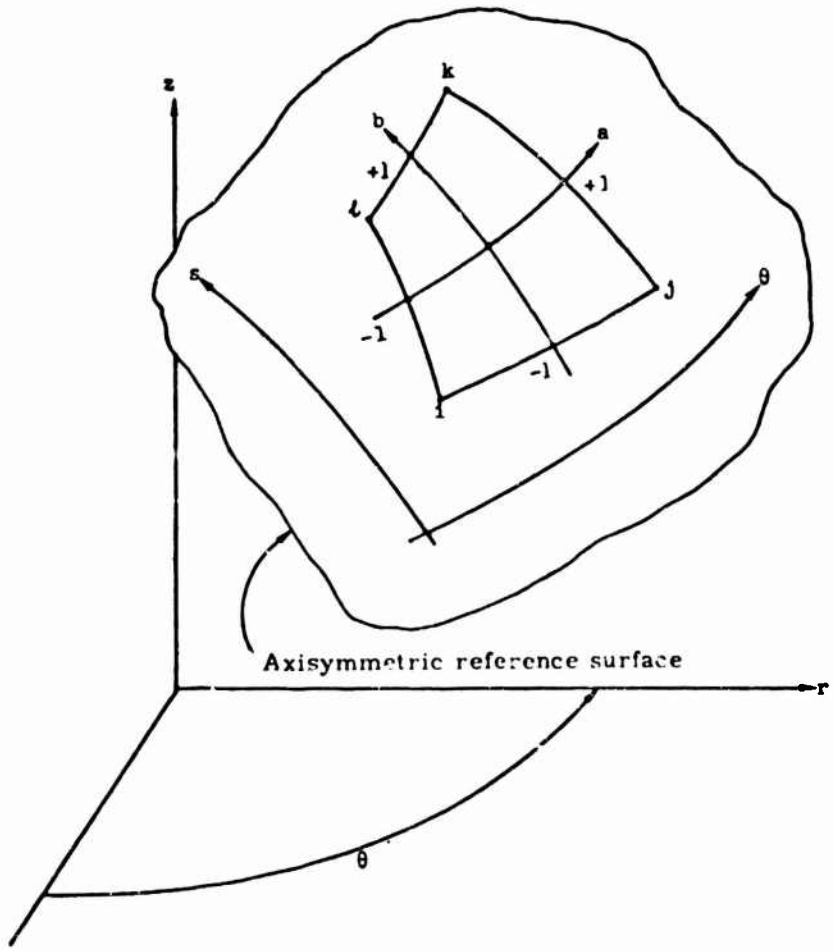


Figure 1. Element Geometry

3. MASS MATRIX

A diagonal or "lumped" mass matrix is used. The approach taken in obtaining it is based on results obtained in simple one dimensional membrane and bending problems which exist as special cases of the more general shell element introduced above.

By taking a cylindrical shell with a very large radius to thickness ratio and zero Poisson's ratio and striking it end on, the same results are obtained as if one were examining the one dimensional wave equation shown in Equation (2).

$$c_m^2 u_{xx} - u_{tt} = 0 \tag{2}$$

Here, u is the axial or meridional displacement of the cylinder, and x and t denote differentiation in space and time respectively. The sound speed c_m is given by the square root of the modulus E divided by the density ρ, $c_m = (E/\rho)^{\frac{1}{2}}$. The thickness of the shell cancels out. For the homogeneous problem considered with homogeneous boundary conditions, Hamilton's principle can be written as shown in Equation (3).

$$\delta \int_{t_1}^{t_2} \int_{x_1}^{x_2} \tfrac{1}{2}\left[(u_t)^2 - (c_m u_x)^2\right] dx dt = 0 \tag{3}$$

In this special case, the original polynomial assumptions contained in the finite element become a cubic in the space variable x. Using the Hermite interpolation form of the cubic polynomial and using an element extending from x_i to x_j, the displacement assumptions can be written as

$$u = u_i h_1(\eta) + u_j h_2(\eta) + u_{x_i} h_3(\eta)\left(\frac{x_j - x_i}{2}\right) + u_{x_j} h_4(\eta)\left(\frac{x_j - x_i}{2}\right) \quad (4)$$

where
$$h_1(\eta) = \tfrac{1}{4}(\eta^3 - 3\eta + 2)$$
$$h_2(\eta) = -\tfrac{1}{4}(\eta^3 - 3\eta - 2)$$
$$h_3(\eta) = \tfrac{1}{4}(\eta^3 - \eta^2 - \eta + 1)$$
$$h_4(\eta) = \tfrac{1}{4}(\eta^3 + \eta^2 - \eta - 1) \quad (5)$$

The local coordinate η runs from -1 to +1. The functions h_1, h_2, h_3, and h_4 are shown in Figure 2.

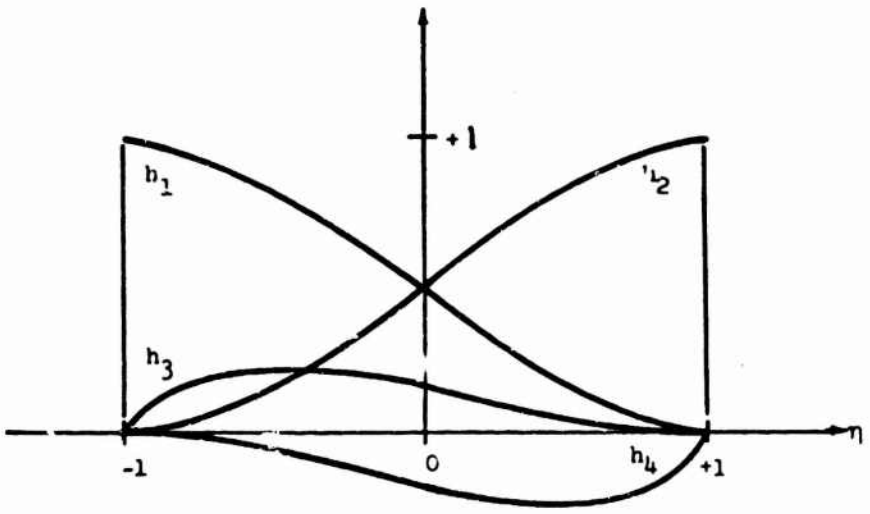


Figure 2. Interpolation functions h_1, h_2, h_3, h_4.

For convenience, the length of the element is written as ℓ in what follows.

The resulting element stiffness is given by Equation (6).

$$k_m = \begin{array}{c} \begin{matrix} u_i & u_{x_i} & u_j & u_{x_j} \end{matrix} \\ \left[\begin{array}{cccc} \frac{18}{15\ell} & \frac{1}{10} & -\frac{18}{15\ell} & \frac{1}{10} \\ \frac{1}{10} & \frac{4\ell}{30} & -\frac{1}{10} & -\frac{\ell}{30} \\ -\frac{18}{15\ell} & -\frac{1}{10} & \frac{18}{15\ell} & -\frac{1}{10} \\ \frac{1}{10} & -\frac{\ell}{30} & -\frac{1}{10} & \frac{4\ell}{30} \end{array}\right] \end{array} \qquad (6)$$

The resulting element mass matrix is given by Equation (7).

$$m = \begin{array}{c} \begin{matrix} u_i & u_{x_i} & u_j & u_{x_j} \end{matrix} \\ \left[\begin{array}{cccc} \frac{13\ell}{35} & \frac{11\ell^2}{210} & \frac{9\ell}{70} & \frac{-13\ell^2}{420} \\ \frac{11\ell^2}{210} & \frac{\ell^3}{105} & \frac{13\ell^2}{420} & \frac{-\ell^3}{140} \\ \frac{9\ell}{70} & \frac{13\ell^2}{420} & \frac{13\ell}{35} & \frac{-11\ell^2}{210} \\ \frac{-13\ell^2}{420} & \frac{-\ell^3}{140} & \frac{-11\ell^2}{210} & \frac{\ell^3}{105} \end{array}\right] \end{array} \qquad (7)$$

These are well-known results, and are easily obtained by standard methods in the finite element literature [22,23,24].

Combining the elements into a complete statement of the problem gives Hamilton's principle in the form of Equation (8).

$$\delta \int_{t_1}^{t_2} \tfrac{1}{2}\left[\dot{q}^T M \dot{q} - c_m^2\, q^T K_m q\right] dt = 0 \qquad (8)$$

Here, q is a vector of nodal displacements and displacement gradients and $\dot{q}$ is a vector of nodal velocities and velocity gradients. Lagrange's equations provide the equations of motion(9)

$$M\ddot{q} + c_m^2 K_m q = 0 \qquad (9)$$

The same cylinder subjected to a purely radial load will behave as if one were examining the lateral motion of a uniform beam as shown in Equation (10).

$$c_b^2\, w_{xxxx} + w_{tt} = 0 \qquad (10)$$

Here w is the radial or normal displacement of the cylindrical shell; x and t are again differentiations in space and time, respectively. The constant c_b is given by the square root of the modulus E times the thickness h squared, all divided by 12 times the density ρ, $c_b = (Eh^2/12\rho)^{\frac{1}{2}}$.

For the homogeneous problem considered with homogeneous boundary conditions, Hamilton's principle can be written as shown in Equation (11).

$$\delta \int_{t_1}^{t_2} \int_{x_1}^{x_2} \tfrac{1}{2}\left[(w_t)^2 - (c_b w_{xx})^2\right] dx dt = 0 \tag{11}$$

Note that rotatory kinetic energy is omitted; it is given by the Expression (12).

$$\int_{t_1}^{t_2} \int_{x_1}^{x_2} \tfrac{1}{2}\left[\frac{h^3}{12} (W_{xt})^2\right] dx dt \tag{12}$$

For bending, the original polynomial assumptions contained in the finite element become a cubic in the space variable x.

Just as before, this is given by Equation (13) for each element.

$$w = w_i h_1(\eta) + w_j h_2(\eta) + w_{x_i} h_3(\eta)\left(\frac{x_j - x_i}{2}\right) + w_{x_j} h_4(\eta)\left(\frac{x_j - x_i}{2}\right) \tag{13}$$

The resulting element stiffness matrix is given by Equation (14).

$$k_b = \begin{array}{c} \begin{matrix} w_i & w_{x_i} & w_j & w_{x_j} \end{matrix} \\ \left[\begin{array}{cccc} \frac{12}{\ell^3} & +\frac{6}{\ell^2} & -\frac{12}{\ell^3} & +\frac{6}{\ell^2} \\ +\frac{6}{\ell^2} & \frac{4}{\ell} & -\frac{6}{\ell^2} & \frac{2}{\ell} \\ -\frac{12}{\ell^3} & -\frac{6}{\ell^2} & \frac{12}{\ell^3} & -\frac{6}{\ell^2} \\ +\frac{6}{\ell^2} & \frac{2}{\ell} & -\frac{6}{\ell^2} & \frac{4}{\ell} \end{array}\right] \end{array} \tag{14}$$

This is a well-known result and easily obtained by standard methods in the finite element literature [22,23,24]. Without rotatory energy terms, the element mass matrix is the same as in the membrane case given by Equation (7).

Combining the elements into a complete statement of the problem gives Hamilton's principle in the form of Equation (15)

$$\delta \int_{t_1}^{t_2} \tfrac{1}{2} \left[\dot{q}^T M \dot{q} - c_b^2 \, q^T K_b q \right] dt = 0 \tag{15}$$

The resulting equations of motion are given by the Expression (16).

$$M\ddot{q} + c_b^2 K_b q = 0 \tag{16}$$

The mass matrices indicated in Equations (9) and (16) are the ones prescribed by the finite element method. They are symmetric and banded and have the same non-zero populations as the stiffness matrices K_m and K_b. Explicit time integration schemes developed in the finite difference literature depend on their ease of application on diagonal or lumped mass matrices, so that M^{-1} is trivial. In order to take advantage of these schemes, it is desirable to replace M with an equivalent diagonal mass matrix M_d. There is of course a considerable amount of literature in the finite difference method on how to create diagonal mass matrices directly and a small amount of work in the finite element literature as well, some of which was covered above.

Consider first only the nodal displacements and their respective mass terms as shown in Equation (17).

$$m = \begin{bmatrix} \frac{13\ell}{35} & \cdot & \frac{9\ell}{70} & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \frac{9\ell}{70} & \cdot & \frac{13\ell}{35} & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \tag{17}$$

If the off-diagonal terms are added to the diagonal terms, then the result is shown in Equation (18).

$$m_{11} + m_{13} = \left(\frac{13}{35} + \frac{9}{70}\right)\ell = \frac{\ell}{2}$$

$$m_{31} + m_{33} = \left(\frac{9}{70} + \frac{13}{35}\right)\ell = \frac{\ell}{2} \tag{18}$$

This is not unexpected if rigid body translational kinetic energy is going to be correct for either form of mass matrix. Following this lead, the gradient terms are added to obtain the diagonal inertias shown in Equation (19).

$$m_{22} + m_{24} = \left(\frac{1}{105} - \frac{1}{140}\right)\ell^3 = \frac{\ell^3}{420}$$

$$m_{24} + m_{44} = \left(\frac{1}{105} - \frac{1}{140}\right)\ell^3 = \frac{\ell^3}{420} \qquad (19)$$

As it will turn out, the important aspect of this exercise is that these terms should scale as ℓ^3 rather than ℓ. A diagonal mass matrix is constructed as shown in Equation (20)

$$m_d = \begin{bmatrix} \frac{\ell}{2} & & & \\ & \frac{\alpha\ell^3}{420} & & \\ & & \frac{\ell}{2} & \\ & & & \frac{\alpha\ell^3}{420} \end{bmatrix} \qquad (20)$$

Here, the gradient masses have been multiplied by a parameter α. The bending problem has rigid body rotations as well as the rigid body translations. However, if α is used to give the correct rotational kinetic energy for a single element rotating about its center of gravity, then a negative α will result giving negative gradient inertias. The membrane problem has no other rigid body motions but translation.

It remains to select α. The maximum frequencies of the membrane and bending problems serve as a guide for this selection.

A study of the maximum eigenvalue as a function of mesh length and gradient inertia parameter is contained in Tables I and II for the membrane and bending problems, respectively. A one-hundred element, uniform grid is used with free-free boundary conditions to approximate the maximum eigenvalue in an infinite mesh. The maximum eigenvalue is obtained by the power method. The nondiagonal mass matrix is used as a standard in each case.

A gradient inertia parameter α equal to 0.75 in the diagonal mass matrix is seen to provide essentially the same maximum eigenvalues as the nondiagonal mass matrix for the membrane problem.*

A gradient inertia parameter α equal to 0.45 is seen to provide essentially the same maximum eigenvalues as the nondiagonal mass matrix.* It seems reasonable to expect that further work with these ideas will lead to better criteria for choosing specific values of the gradient inertia parameter α. Certainly more important is how the entire frequency spectrum is affected by α, rather than just the maximum frequency.

To obtain a mass matrix suitable for the original shell element, these ideas are carried over just as they stand. The equivalent entries in the consistent mass matrix are computed and added to form the diagonal terms of a diagonal mass matrix. The same gradient inertia scaling is used; one value for membrane and another value for bending. No rotatory inertia is included. The selection of the scaling values comes from the frequency considerations just covered.

* Tables I and II have been corrected from the original. They do not now support these choices of the gradient inertia scaling parameters. However, all of the conclusions of the paper remain the same. Only the details of the calculations will change with new values of these parameters.

TABLE I

Membrane Eigenvalues

Maximum Eigenvalues of $M^{-1}K_m$ and $M_d^{-1}K_m$ as a Function of Element Length and Gradient Inertia Scaling

		Element Length, ℓ										
		0.25	0.50	0.75	1.00	1.25	1.50	2.00	3.00	4.00	5.00	6.00
$M^{-1}K_m$		504	168	56	42	27	18	10	5	2.6	1.7	1.2
$M_d^{-1}K_m$	$\alpha = 0.50$	2177	547	242	137	87	61	34	15	8.6	5.5	3.8
	$\alpha = 0.75$	1461	365	161	91	58	40	23	10	5.8	3.7	2.6
	$\alpha = 1.00$	1094	273	123	69	44	30	17	8	4.3	2.8	1.9

TABLE II

Bending Eigenvalues

Maximum Eigenvalues of $M^{-1}K_b$ and $M_d^{-1}K_b$ as a Function of Element Length and Gradient Inertia Scaling

		Element Length, ℓ										
		0.25	0.50	0.75	1.00	1.25	1.50	2.00	3.00	4.00	5.00	6.00
$M^{-1}K_b$		$.64\times10^6$	40,200	7,950	2510	1030	497	157	31	10	4.0	1.9
$M_d^{-1}K_b$	$\alpha = .25$	2.78×10^6	160,000	31,900	10,090	5130	1990	631	124	39	16.1	7.8
	$\alpha = .45$	1.44×10^6	89,760	17,700	5610	2300	1110	351	69	22	9.0	4.3
	$\alpha = 1.0$	0.65×10^6	40,400	7,990	2530	1030	499	158	31	10	4.0	1.9

The same procedure is used in Reference [11] to form a diagonal mass matrix from the non-diagonal mass matrix prescribed by the finite element method. However, rotatory inertia is retained in Reference [11] in contrast to the approach taken here. The details on the choice of the scaling parameters are omitted.

4. TIME INTEGRATION

Due to its computational speed and entirely satisfactory performance when operating with a time step below the critical one, a central difference time integration is used. Without reference to either the membrane or bending behavior, the equations of motion are separated into the two first order equations shown in Equation (21).

$$\begin{aligned} \dot{p} + c^2 M^{-1} K q &= 0 \\ \dot{q} - p &= 0 \end{aligned} \tag{21}$$

Here, p is a velocity vector. Using a time step of Δt and a subscript of n to denote a point in time, these equations are differenced to obtain the Equations (22)

$$\begin{aligned} p_{n+\frac{1}{2}} &= p_{n-\frac{1}{2}} - \Delta t c^2 M^{-1} K q_n \\ q_{n+1} &= q_n + \Delta t p_{n+\frac{1}{2}} \end{aligned} \tag{22}$$

Thus, given the velocities p at $n-\frac{1}{2}$ and the displacement q at n, the new velocities at $n + \frac{1}{2}$ and the new displacements at n+1 are computed with one pass. In theory, this is entirely equivalent to using a second central difference expression for the accelerations $\ddot{q}$ in either Equations (9) or (16) and using the old displacements at n and n-1 to get the new displacements at n+1. In practice, however, (22) is more accurate due to the finite word lengths of computers, [25].

It is also much more informative to know a velocity and a displacement at a given point in time rather than two successive displacements, albeit only a short calculation to get the velocities.

This integration scheme is only conditionally stable. To examine stability, it is more convenient to use the all-displacement version as shown in Equation (23).

$$q_{n+1} - 2q_n + q_{n-1} + \Delta t^2 c^2 M^{-1} K q_n = 0 \tag{23}$$

A solution in the form of Equation (24) is sought.

$$q_n = v e^{\beta n \Delta t} \tag{24}$$

Here β is undetermined and v is an arbitrary displacement shape. Substituting Equation (24) into Equation (23) provides the results in Equation (25).

$$(e^{\beta \Delta t} - 2 + e^{-\beta \Delta t} + \Delta t^2 c^2 M^{-1} K) v = 0 \tag{25}$$

This is an eigenvalue problem and for each eigenvalue there is a separate β. Denoting any given eigenvalue of $M^{-1}K$ by λ, the appropriate polynomial for β is given in Equation (26).

$$e^{\beta \Delta t} - 2 + e^{-\beta \Delta t} + \Delta t^2 c^2 \lambda = 0 \tag{26}$$

Since the motion in Equation (21) is bounded for all time, $|e^{\beta \Delta t}|$ must be less than 1. This leads to the requirement that $\Delta t^2 c^2 \lambda$ must be less than or equal to 4.

Since the maximum eigenvalue of $M^{-1}K$ provides the strictest condition, the limits on Δt become those in Equation (27).

$$\Delta t^2 \leq \frac{4}{c^2 \lambda_{max}} \tag{27}$$

Previous work in this area[26] shows that for the membrane problem, λ is proportional to $1/\ell^2$, the mesh length in a uniform mesh, and for the bending problem λ is proportional to $1/\ell^4$. Thus, empirical constants related to the differencing scheme or displacement assumptions and independent of element length and the constant c will give the expressions in Equation (28).

$$\Delta t \leq a_1 \frac{\ell}{c_m}$$

$$\Delta t \leq a_2 \frac{\ell^2}{c_b} \tag{28}$$

The first inequality is for the membrane problem, the second is the bending problem. If a combined membrane and bending response calculation is being made, then the strictest constraint will limit Δt. Table I shows that λ_{max} does vary as $1/\ell^2$ and a value of a_1 equal to 0.209 results for the membrane problem. In Table II, λ_{max} varies as $1/\ell^4$ and a value of a_2 equal to 0.0267 results for the bending problem.

The choices in the gradient inertia parameters based on a matching of the maximum eigenvalues give the nondiagonal and diagonal mass matrix the same stability constraints on time step size for this integration method. To apply these results to the general shell element, several adaptations must be made. For layered shells, an area density rather than space density is more appropriate. Equation (29) defines the area density r.

$$r(\theta,s) \equiv \int_{h^-}^{h^+} \rho(\theta,s,\zeta)d\zeta \tag{29}$$

The larger of the meridional or circumferential membrane moduli in the shell theory is used in place of the modulus E. Equation (30) defines this elastic constant.

$$C(\theta,s) \equiv \max\left(\int_{h^-}^{h^+} \frac{E_{\theta\theta}}{1-\nu_{s\theta}\nu_{\theta s}}\,d\zeta\ ,\ \int_{h^-}^{h^+} \frac{E_{ss}}{1-\nu_{s\theta}\nu_{\theta s}}\,d\zeta\right) \tag{30}$$

The larger of the meridional or circumferential bending moduli in the shell theory is used in place of $\frac{Eh^2}{12}$. Equation (31) defines this elastic constant

$$D(\theta,s) \equiv \max\left(\int_{h^-}^{h^+} \frac{\zeta^2 E_{\theta\theta}}{1-\nu_{s\theta}\nu_{\theta s}}\,d\zeta\ ,\ \int_{h^-}^{h^+} \frac{\zeta^2 E_{ss}}{1-\nu_{s\theta}\nu_{\theta s}}\,d\zeta\right) \tag{31}$$

Since meshing in general is irregular, in each element the minimum of the side lengths and diagonals is used in place of ℓ. Equation (32) defines this minimum distance δ.

$$\delta = \min\ (\overline{ij}, \overline{jk}, \overline{k\ell}, \overline{\ell i}, \overline{ik}, \overline{j\ell}) \tag{32}$$

Thus, in each element a critical time step is computed, one based on membrane response and the other on bending response. The minimum obtained for the entire mesh must govern the integration. Equation (32) restates the time step constraints for the general shell element.

$$\Delta t \leq \min\left(0.209\ \frac{\delta}{\sqrt{C/r}}\ ,\ 0.0267\ \frac{\delta^2}{\sqrt{D/r}}\right) \tag{32}$$

To account for boundary conditions, nonuniform meshing, nonhomogeneous materials, membrane and bending coupling and to improve accuracy, a value of Δt equal to 0.8 of critical is customarily used. For example, critical time step estimates have been worked out for the axial bending in cylindrical shells by Sobel[27], a case where the circumferential curvature makes the axial bending stiffer. It is reflected by a slight decrease in the critical time step for bending.

At a mesh length to thickness ratio of 2.25, both criteria give the same critical time step. The critical time step for smaller length to thickness ratios is controlled by the bending and for larger ratios by the membrane behavior.

5. NUMERICAL EXAMPLES

Three examples of the results of this work are enclosed. The first example is a very simple problem for which an exact solution is known. It is very difficult numerically since the solution is a propagating discontinuity. The second example is a bending problem where the stresses generated at a clamped boundary are examined. The last example is a cone problem without an exact solution but for which numerous other computations are available with which to make a comparison.

By putting a unit step in pressure on the end of a cylindrical shell of very large radius to thickness ration and zero Poisson's ratio a unit step in stress should propagate down the cylinder just as in a half space. Figure 3 shows the numerical results of this calculation along with the exact solution, the propagating discontinuity in stress.

Figures (3a) and (3b) show the temporal behavior of the stress pulse as it passes a fixed point along the cylinder. Figures (3c) and (3d) show the shape of the pulse at a fixed point in time. The time step used was controlled by bending and is one half of the critical membrane time step of the lumped mass system. The lumped mass results are much to be preferred over those from the consistent mass calculations. The overshoots in stress are less for the lumped mass results. Even in this ideal circumstance, the consistent mass calculations required twice the time to compute than the lumped mass results required. The early arrival time of a significant amount

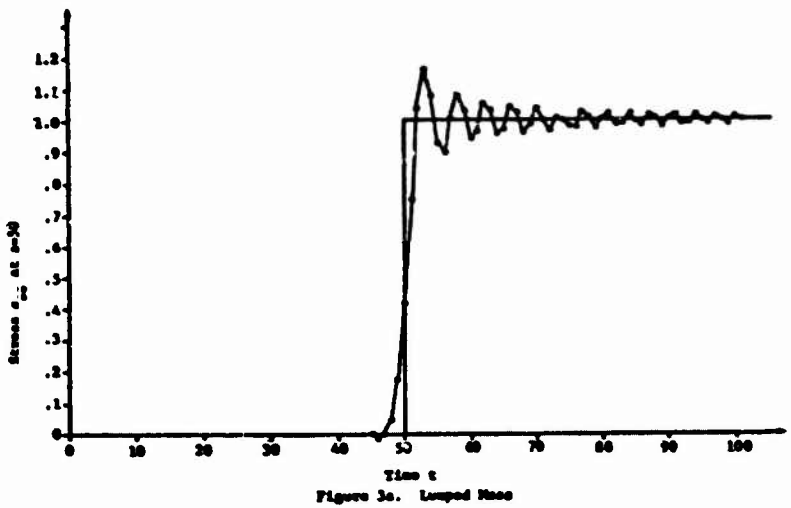


Figure 3a. Lumped Mass

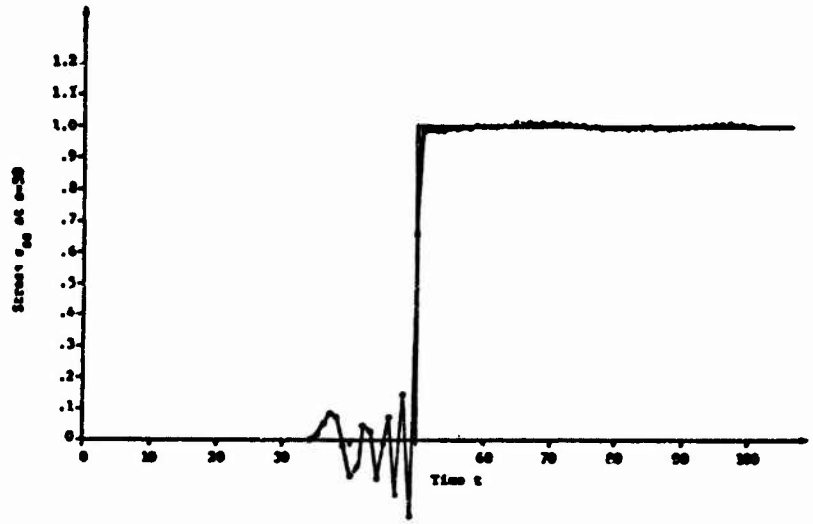


Figure 3b. Consistent Mass

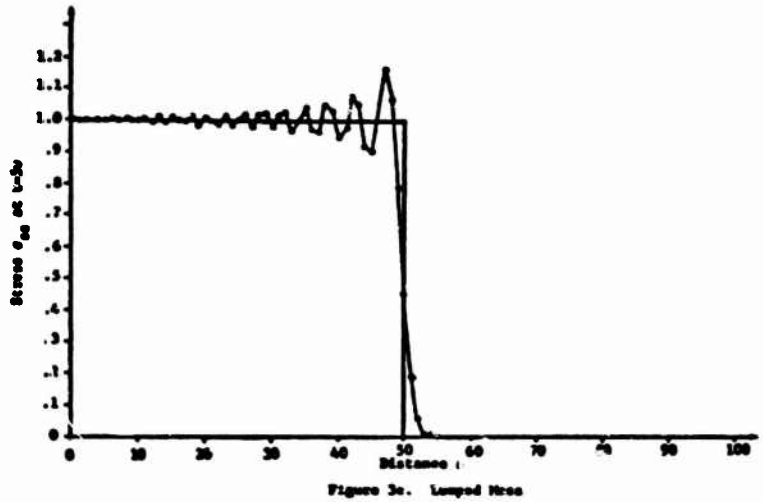


Figure 3c. Lumped Mass

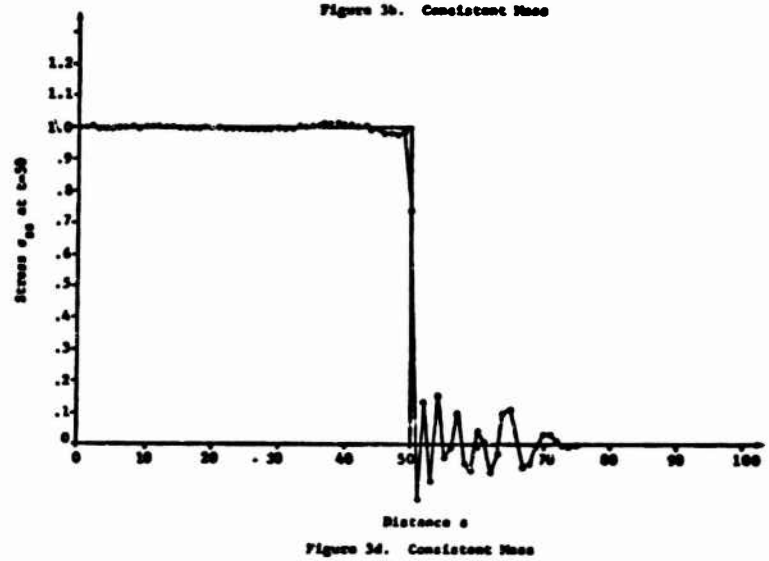


Figure 3d. Consistent Mass

Figure 3

The membrane response of a cylindrical shell struck on the end by a Heaviside pressure pulse. Modulus E=1, Poisson's ratio ν=0, density ρ=1, thickness h=1, radius r=10^{10}, length L=100, mesh spacing Δs=1, time step Δt=1/11.

of information in the consistent mass calculations cannot be tolerated in more complicated problems. One of the most valuable checks of the correctness of an involved calculation is the arrival times of stress pulses at far boundaries and interfaces. It is frequently necessary to establish delay times between the point of impact and some other location on the structure and calculation where stresses arrive early cannot be reliably used.

The smearing of the discontinuity and oscillations are typical of virtually all of the common numerical integration schemes. The overshoot and oscillations in the numerical solutions just behind the discontinuity can be reduced, but only at the expense of further smearing of the front with artificial viscosity or implicit damping accompanying some integration schemes[9] Shock matching or method of characteristics computer programs are able to improve on this solution. Both these techniques are used in one dimensional wave propagation calculations, but are inappropriate for the general shell problem of interest here.

A clamped cylindrical shell with a suddenly applied pressure responds early in time predominately in bending. Figure 4 shows the bending stresses at the clamped support and at two places removed from the support. A comparison is made with a bending solution developed by Forrestal, Sliter and Sagartz[27]. As can be seen, the agreement is quite good.

The last example is a problem from a series of check problems being assembled by the Lockheed Palo Alto Research Laboratories.

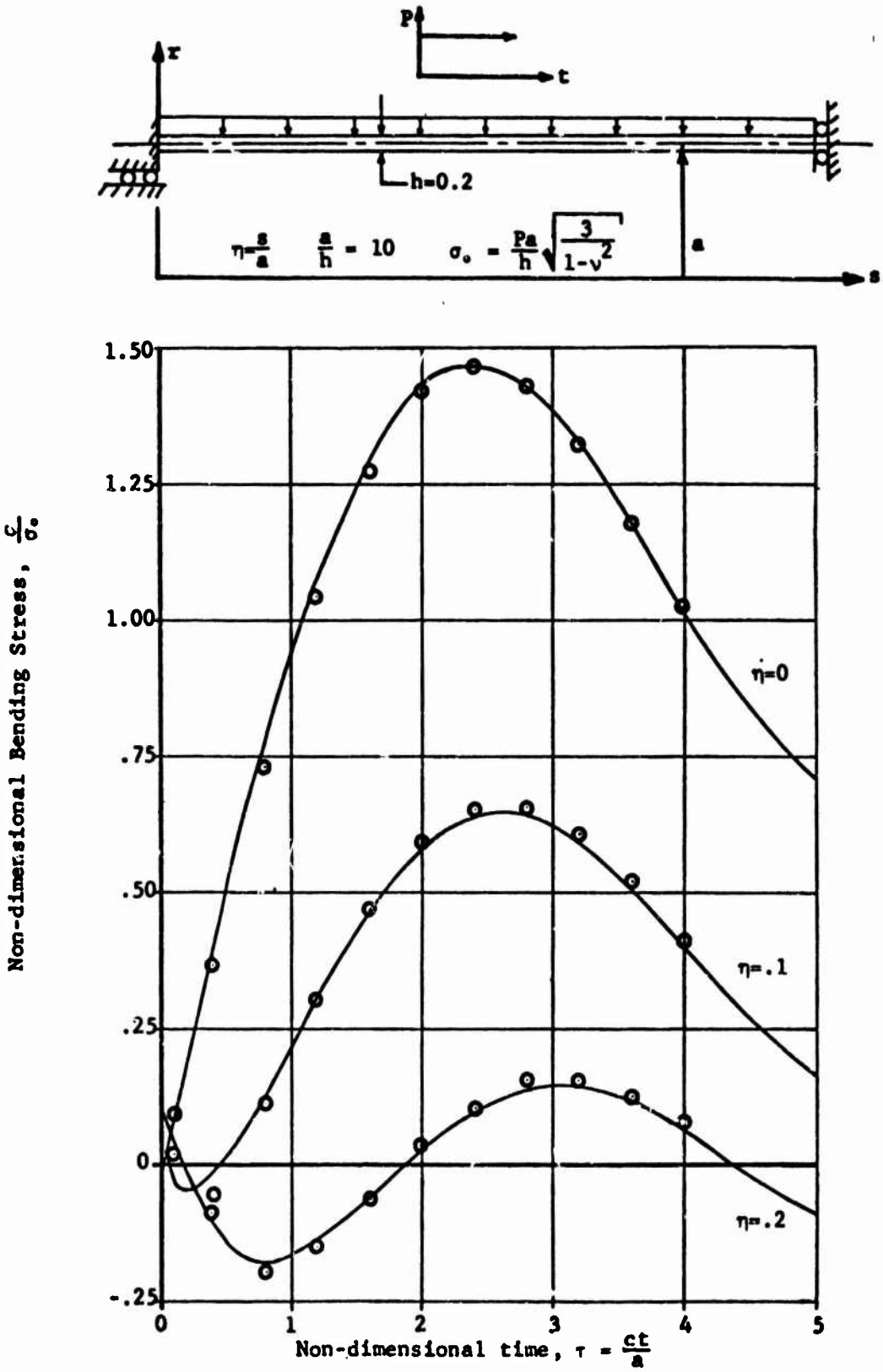


Figure 4. Bending response at a clamped support of a cylindrical shell.

It is a clamped-clamped conical frustum loaded with a half cosine initial radial velocity distribution. Figure 5 is a detailed statement of the problem. Figure 6 shows the mesh used and the point under the load for which the normal displacement is plotted. Figure 7 shows the normal deflection under the peak load as a function of time. The results from the present work and the other three computations are virtually identical. In view of the differences in these programs, both in space and time integrations, these results must be considered as correct and as a verification of all four efforts.

SLADE and the STAR program, a finite difference program at Lockheed, both use explicit time integration schemes, while DYNASØR and SABØR3/DRASTICII both use implicit time integration schemes. The critical time step for this problem is controlled by meridional bending at the small end of the cone. Expression (32) gives a critical time step of 1.03 μsec and a time step of 1 μsec was used in this analysis. A calculation made with a 1.06 μsec time step diverged.

While there does not presently exist enough information for timing comparisons with other schemes, it is clear that explicit methods are much faster per step than implicit methods. They are, however, constrained to function with very small time steps. It is interesting to note that whenever an explicit time integration method is used on a problem where high frequencies dominate, they are invariably

Material Properties

Homogeneous, isotropic, elastic

Young's Modulus = 3.52×10^6 psi

Poisson's Ratio = 0.286

Mass Density = 1.88×10^{-4} slugs/in^3

Boundary Conditions:

Clamped at Both Ends

Loading

Half Cosine Initial Velocity

Uniform along the Shell

V_o = 444. in/sec.

Meshing

18 x 15, Δt = 1 μsec

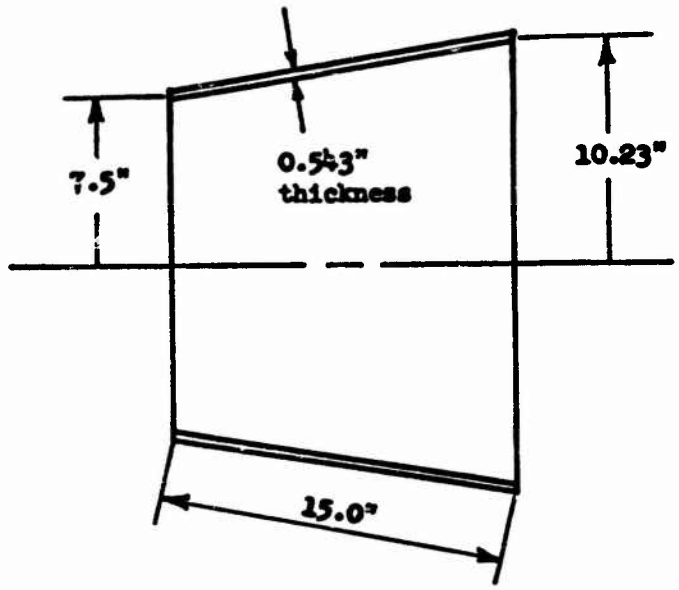


Figure 5 Dynamic Response of a Conical Shell

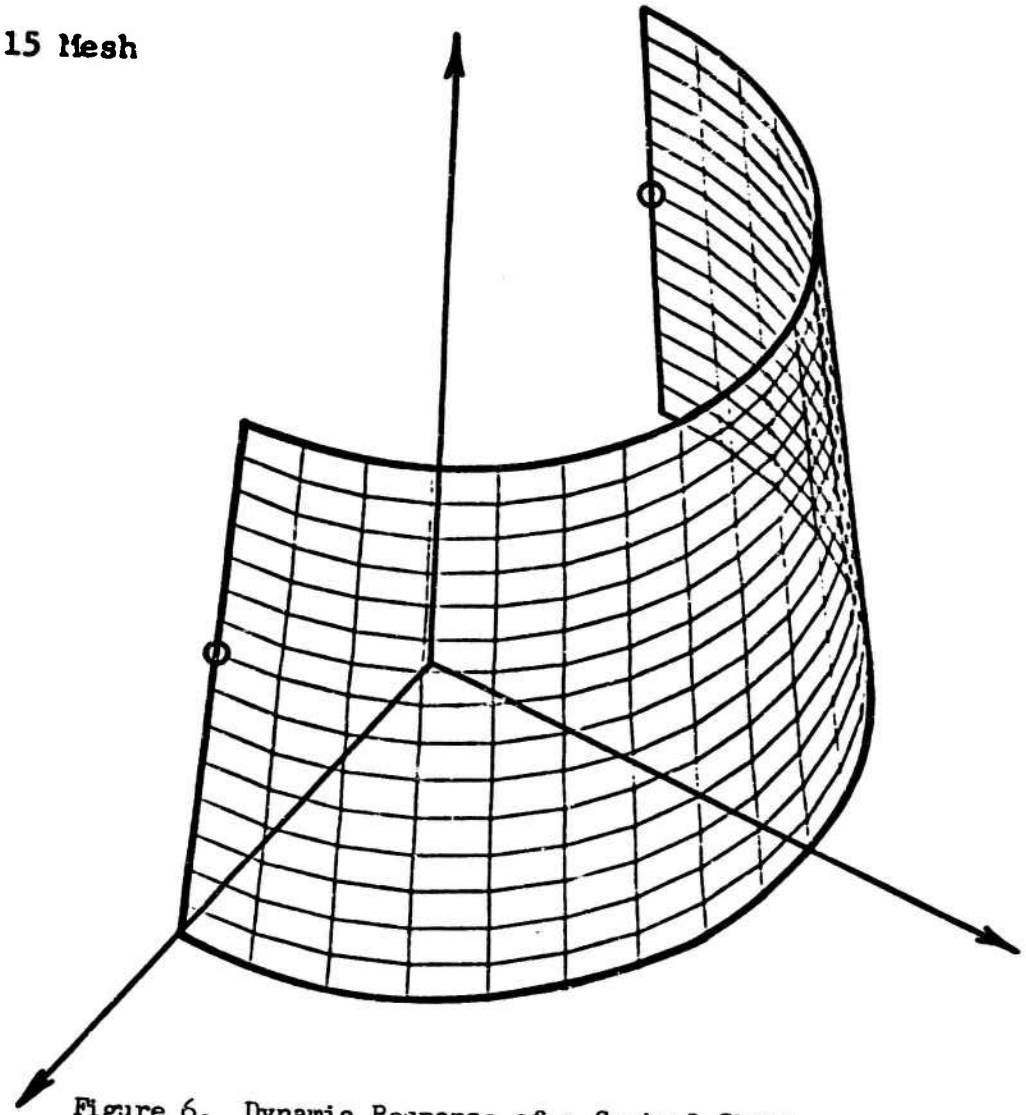


Figure 6. Dynamic Response of a Conical Shell

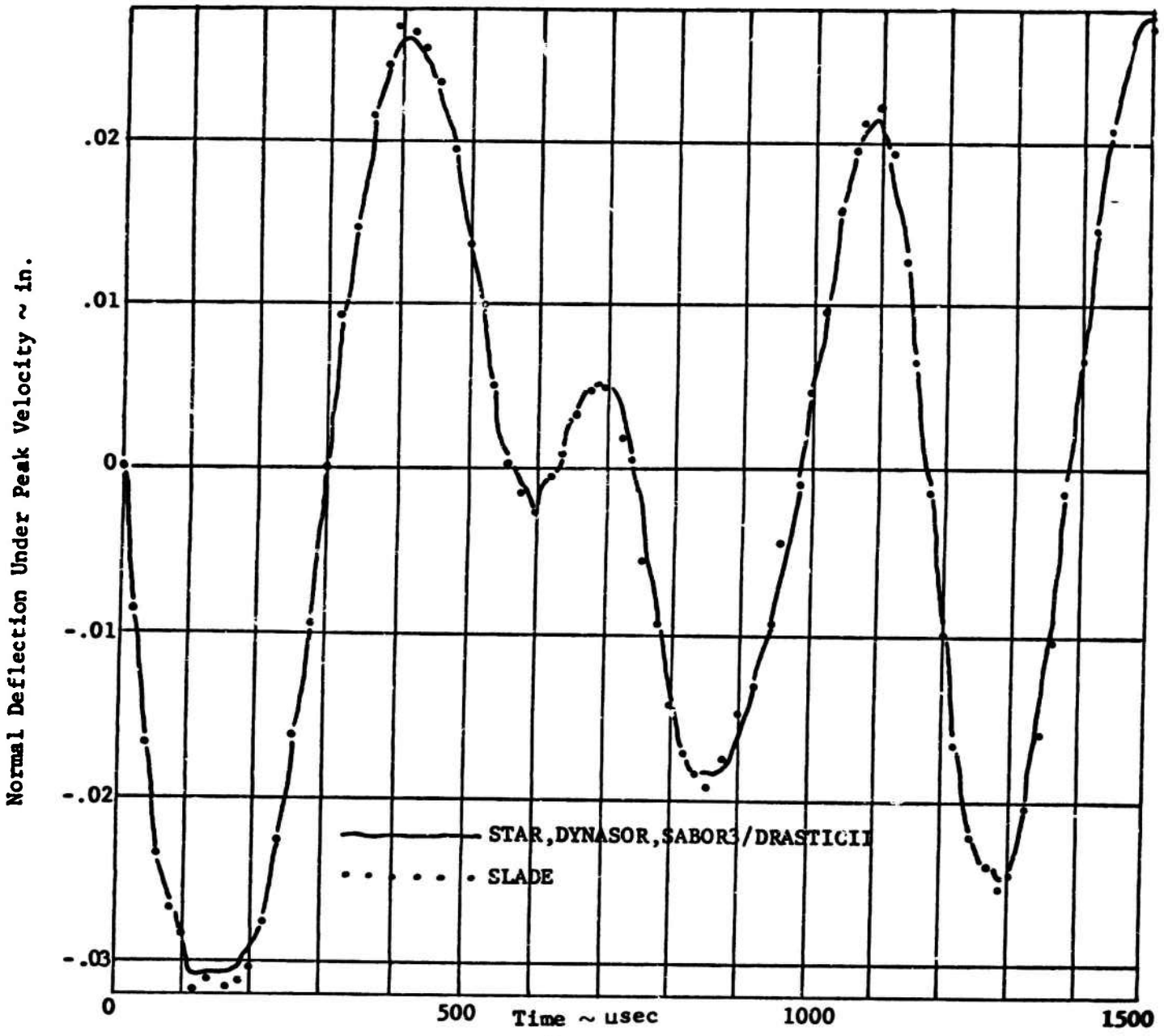

Normal Deflection Under Peak Velocity ~ in.
.02
.01
0
-.01
-.02
-.03
0
500
Time ~ usec
1000
1500
STAR,DYNASOR,SABOR3/DRASTICII
SLADE

Figure 7

used with time steps that are equal to or less than the critical time step of the explicit central difference time integrator. In this case, accuracies are the same. For the present results on a CDC6600 computer, 1890 equations of motion are integrated per central processor second. With nine equations at each mesh point, that is equivalent to 210 mesh points per central processor second or 756,000 mesh points per central processor hour. The integration subroutine is largely the product of the stiffness matrix times the displacement vector. This subroutine has also been coded in the CDC machine language. When the calculations are carried out with this version of the subroutine, 5700 equations of motion are integrated per central processor second, or 630 mesh points per central processor second.

6. CONCLUSIONS

The present work has demonstrated a very viable transient dynamic response approach. It is based on the simplest of mass matrices and time integration schemes. The frequently heard complaint of instability is easily resolved by a study of stability and in practice the time step constraint is contained internally in the computer program and used to check the users requested time step and reduce it as necessary. It is a one-time operation that remains valid throughout the integration. The diagonal mass matrix is indistinguishable from its nondiagonal parent in the results and is much easier to handle from a computational and storage standpoint.

While this work was undertaken to provide a mass matrix for the shell element in SLADE, the approach is applicable to a wide class of shell elements. The bending behavior of virtually all shell elements when taken as flat and in one dimension is that of a beam element with cubic displacement assumptions and the membrane behavior will be that of a bar element with linear, quadratic or cubic displacement assumptions. Whenever this is the case, the present results apply.

Acknowledgement

For many helpful discussion, the authors are indebted to Raymond D. Krieg, Sandia Laboratories, whose knowledge and experience in these matters overshadows all others we have encountered. Special mention must also be made of Walter A. Von Riesemann, Sandia Laboratories, whose generosity with his extensive collection of finite element literature has proven extremely helpful.

REFERENCES

1. J. S. Archer, "Consistent Mass Matrix for Distributed Mass Systems," Journal of the Structural Division, Proceed. ASCE, August 1963.

2. J. S. Archer, "Consistent Matrix Formulations for Structural Analysis Using Finite-Element Techniques," AIAA Journal, Vol. 3, October 1965.

3. E. L. Wilson and R. W. Clough, "Dynamic Response by Step-by-Step Matrix Analysis," Symposium on the Use of Computers in Civil Engineering, Lisbon, Portugal, October 1962.

4. R. W. Clough and C. A. Felippa, "A Refined Quadrilateral Element for Analysis of Plate Bending," Proceedings of the 2nd Conference on Matrix Methods in Structural Engineering, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, October 1968.

5. R. W. Clough, "Analysis of Structural Vibrations and Dynamic Response," Japan-U. S. Seminar on Matrix Methods of Structural Analysis and Design, Tokyo, Japan, August 1969.

6. R. W. Clough and E. L. Wilson, "Dynamic Finite Element Analysis of Arbitrary Thin Shells," Conference on Computer Oriented Analysis of Shell Structures, Lockheed Palo Alto Research Laboratory, Palo Alto, California, August 1970.

7. S. Klein and R. J. Sylvester, "The Linear Elastic Dynamic Analysis of Shells of Revolution by the Matrix Displacement Method," Proceedings of the 2nd Conference on Matrix Methods in Structural Engineering, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, October 1968.

8. S. P. Chan, H. L. Cox and W. A. Benfield, "Transient Analysis of Forced Vibrations of Complex Structural-Mechanical Systems," Journal of the Royal Aeronautical Society, Vol. 66, July 1962.

9. J. A. Stricklin, J. E. Martinez, J. R. Tillerson, J. H. Hong and W. E. Haisler, "Nonlinear Dynamic Analysis of Shells of Revolution by Matrix Displacement Method," AIAA/ASME 11th Structures, Structural Dynamics, and Materials Conference, Denver, Colorado, April 1970.

10. J. C. Houbolt, "A Recurrence Matrix Solution for the Dynamic Response of Elastic Aircraft," Journal of the Aeronautical Sciences, Vol. 17, September 1950.

11. L. B. McWhorter, and W. E. Haisler, "FAMSOR-A Finite Element Program for the Frequencies and Mode Shapes of Shells of Revolution," SC-CR-715125, Sandia Laboratories Contractor Report, Albuquerque, New Mexico, October 1970.

12. C. C. Fu, "A Method for the Numerical Integration of the Equations of Motion Arising From a Finite Element Analysis," JAM, Vol. 37, September 1970.

13. R. de Vogelaere, "A Method for the Numerical Integration of Differential Equations of Second-Order Without Explicit First Derivatives," Journal of Research, National Bureau of Standards, Vol. 54, No. 3, 1955.

14. M. D. Olson and G. M. Lindberg, "Vibration Analysis of Cantilevered Curved Plates Using a New Cylindrical Shell Finite Element," Proceedings of the 2nd Conference on Matrix Methods in Structural Engineering, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, October 1968.

15. B. E. Greene, R. E. Jones, D. R. Strome and R. W. McLay, "Dynamic Analysis of Shells Using Doubly-Curved Finite Elements," Proceedings of the 2nd Conference on Matrix Methods in Structural Engineering, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, October 1968.

16. R. E. Nickell, "On the Stability of Approximation Operators in Problems of Structural Dynamics," International Journal of Solids and Structures, Vol. 7, pp. 301-319, 1971

17. G. L. Goudreau, "Evaluation of Discrete Methods for the Linear Dynamic Response of Elastic and Viscoelastic Solids," Report No. 69-15, Department of Civil Engineering, University of California, Berkeley, California, June 1970.

18. S. W. Key and R. D. Krieg, "Comparison of Finite Element and Finite Difference Methods," ONR Symposium on Numerical and Computer Methods in Structural Mechanics, Urbana, Illinois, September, 1971.

19. S. W. Key, "The Analysis of Thin Shells With a Doubly Curved Arbitrary Quadrilateral Finite Element," Conference on Computer Oriented Analysis of Shell Structures, Lockheed Palo Alto Research Laboratory, Palo Alto, California, August 1970.

20. S. W. Key and Z. E. Beisinger, "The Analysis of Thin Shells by the Finite Element Method," IUTAM Symposium on High Speed Computing of Elastic Structures, Liege, Belgium, August 1970.

21. S. W. Key and Z. E. Beisinger, "SLADE: A Computer Program for the Static Analysis of Thin Shells," SC-RR-69-369, Sandia Laboratories, Albuquerque, N. M., November 1970.

22. H. C. Martin, "Introduction to Matrix Methods of Structural Analysis," McGraw Hill Book Co., New York, 1966.

23. O. C. Zienkiewicz, "The Finite Element Method in Structural and Continuum Mechanics," McGraw Hill Book Co., New York, 1967.

24. J. S. Przemieniecki, "Theory of Matrix Structural Analysis," McGraw Hill Book Co., New York, 1968.

25. R. D. Krieg and H. C. Monteith, "A Large Deflection Transient Analysis of Arbitrary Shells Using Finite Differences," Conference on Computer Oriented Analysis of Shell Structures, Lockheed Palo Alto Research Laboratory, Palo Alto, California, August 1970.

26. R. D. Krieg and S. W. Key, "UNIVALVE, A Computer Code for Analyzing Large Deflection Elastic-Plastic Response of Beams and Rings," SC-RR-66-2692, Sandia Laboratories, Albuquerque, New Mexico, January 1968.

27. L. H. Sobel, "Analytical Study of Asymmetric Wave Propagation in Shells," Progress Report No. 13, NASA Contract NAS1-9111, Lockheed Palo Alto Research Laboratory, Palo Alto, California, June 1970.

28. M. J. Forrestal, G. E. Sliter, and M. J. Sagartz, "Stresses Emanating from the Supports of a Cylindrical Shell Produced by a Lateral Pressure Pulse," JAM, to appear.

EFFECTS OF RISE TIME AND DAMPING ON FINITE ELEMENT ANALYSIS OF RESPONSE OF STRUCTURES

J. J. FARRELL* AND P. K. DAI**
TRW Systems Group

This paper discusses the results of some sensitivity studies on the effects of rise time and viscous damping on the output of a linear finite element code. The results are interpreted by means of modal analysis of the finite element model. Approximate methods are presented for the determination of the highest modal frequency of a uniform grid and spurious oscillations at the highest modal frequency are correlated to the shock spectrum of the forcing function. The effect of damping on shock propagation is described in modal terms. Finally, the effects of modal participation of the forcing function upon finite element models of continua are mentioned.

---

* Member of Technical Staff, TRW Systems Group
** Head, Protective Construction Section, TRW Systems Group

INTRODUCTION

There has been considerable interest in recent years, in the determination of dynamic response of half buried and/or buried structures to air blast and ground shock loading induced by the detonation of nuclear weapons or high explosives. These structures are generally of the thick wall construction as they are "hardened" against the blast and shock environments. The important task of determining the response of such structures is not only in the evaluation of structural integrity under the severe blast and shock environments but also in the understanding of transfer of shock environment from the surrounding medium into the structures as various mechanical, electrical and electronic components are mounted (either hard mounted or shock isolated) to the interior of the structures. Consequently, the response analysis must include both the gross response of the structure and the stress wave propagation in the material. In other words, stress in the thickness direction cannot be ignored.

Various numerical methods (e.g., spring-mass model and finite difference methods) have been used in the past in the study of this problem. The finite element approach has been used, in recent years, in the response analysis of this structure-medium interaction problem. In using the finite element method for analyzing such a problem, two questions were frequently raised:

- What are the criteria for selecting the grid size?
- How should damping be considered?

This paper will attempt to address itself to both of these questions. References 1) and 2) which refer to spring-mass models, provided the perspective from which the finite element discretization effects were assessed. Reference 3) describes the effects of proportional damping without particular reference to shock propagation.

CONSIDERATION OF RISE TIME IN THE SELECTION OF GRID SIZE

Typically, the blast and shock loading considered in the response analysis, as shown in Figure 1, has a short rise time, usually on the order of a fraction of a millisecond to a few milliseconds. When a shock load of this type is applied to a finite element model of a soil-structure complex as shown in Figure 2, spurious oscillations always occur. Such oscillations are spurious because they do not represent the physical motion of the continuum. Generally, they are caused by the dispersion and the filtering effects of the discrete elements. Such oscillations may be minized selecting the mesh size intelligently. Oscillations in one part of a mesh propagate into other regions as time passes. Consequently, the usual static procedure of placing more elements in regions of stress concentrations is not always the best method of selecting a mesh. This means that the grid of finite elements must be nearly uniform. Also, shock fronts in uniform grids tend to be "smeared out" across several elements. Such a filtering effect of the force pulse seems to indicate that if the rise time is short compared with the transit time across an element, similar response will be obtained when impulses of these pulses are nearly equal.

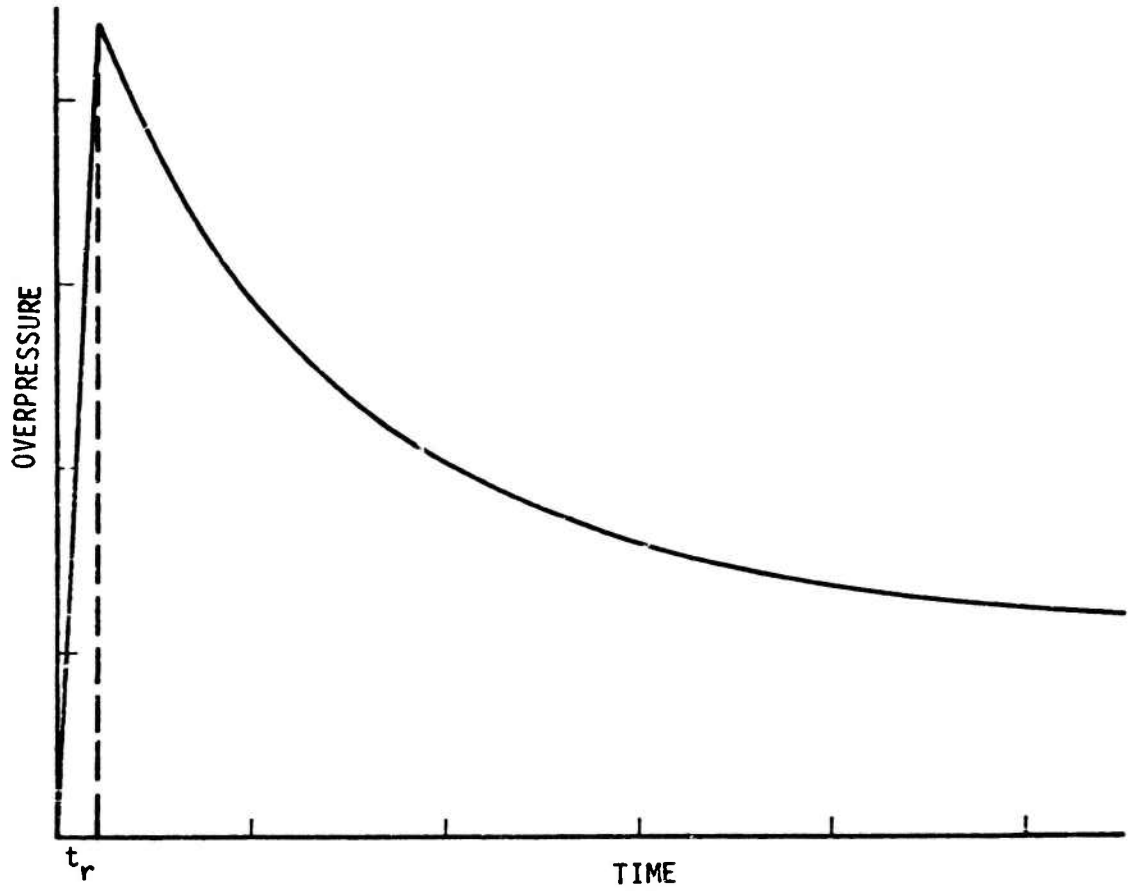


FIGURE 1. TYPICAL TIME HISTORY OF BLAST WAVE

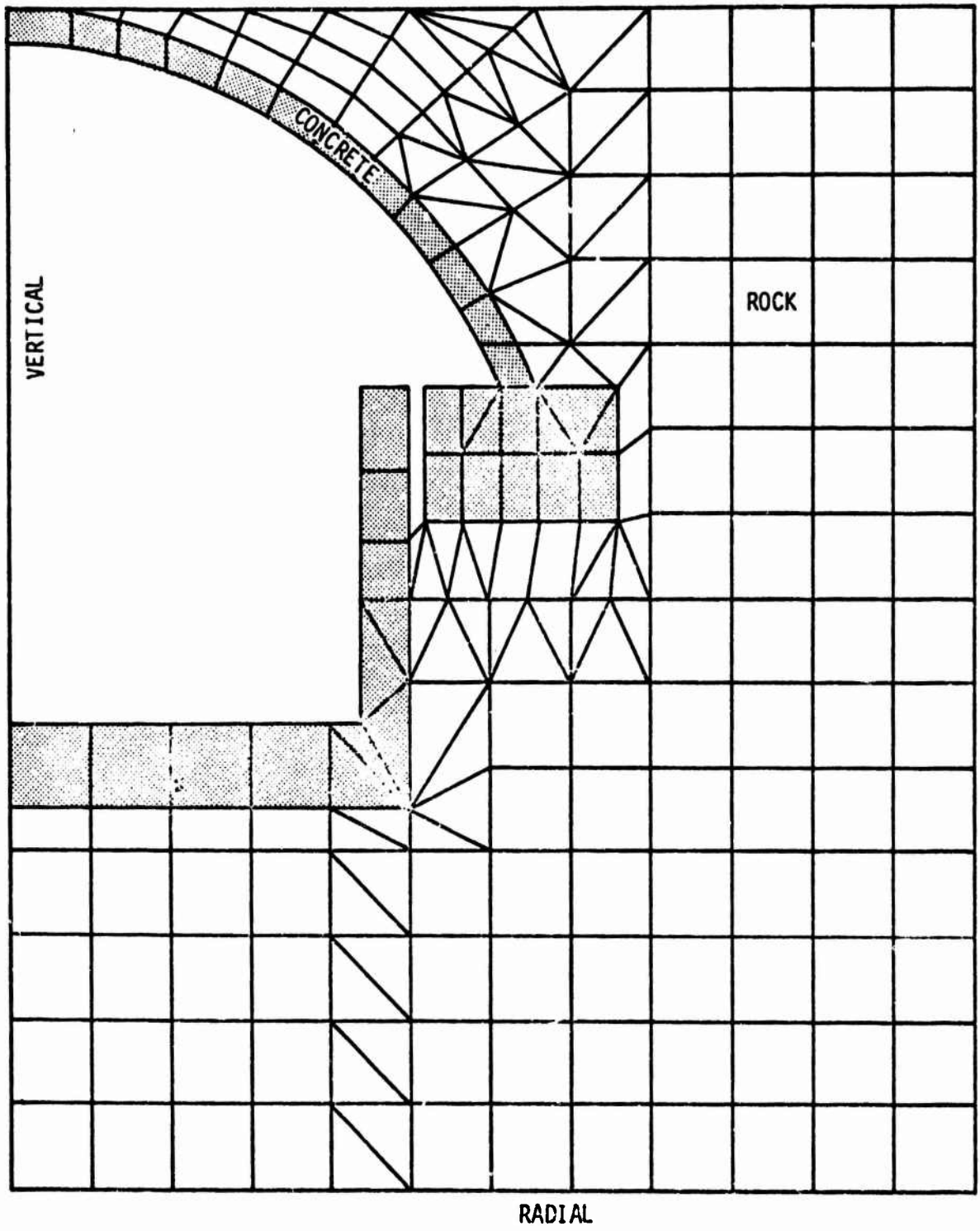


FIGURE 2. TYPICAL FINITE MODEL OF UNDERGROUND STRUCTURE

The frequencies associated with spurious oscillations are the highest modal frequencies of a grid of finite elements. For convenience of discussion, one-dimensional problems are used as examples. For a uniform one-dimensional grid, the mode shape corresponding to the highest frequency is approximately:

$$u = \{ 1 \ -1 \ \ 1 \ -1 \ \ldots\ldots \ 1 \ \ -1 \}$$

The Rayleigh quotient for the one-dimensional system can be computed to be:

$$Q(u) = \frac{\tilde{u}Ku}{\tilde{u}Mu} = \frac{4k}{m}\left(1 - \frac{1}{4n-2}\right)$$

where

k = stiffness of one element, $\frac{1-\nu}{(1+\nu)(1-2\nu)} \frac{EA}{L}$

m = mass at each node, ρAL

n = number of masses in the chains

ρ = density

ν = Poisson's ratio

E = Young's modulus

L = length of an element

A = cross-sectional area

As n becomes large, u becomes a good approximation to the highest mode shape and Q(u) approaches the value ω^2 of the highest mode. Consequently,

$$\omega^2 \simeq 4\ k/m$$

or

$$\omega = 2\ c_d/L$$

where c_d is the velocity of the dilatational wave (p-wave)

$$c_d = \sqrt{\frac{E}{\rho} \frac{(1-\nu)}{(1+\nu)(1-2\nu)}}$$

The transit time of the P-wave across an element is given by

$$t_e = L/c_d$$

and the period for the highest mode becomes:

$$\tau = \pi t_e$$

A series of calculations were made to consider a step pressure pulse with a rise time (t_r) varying from 0 to 2π. It was observed that relative minima for spurious oscillations occur approximately at the minima for the shock spectra of the highest frequency mode of a grid of finite elements. A strong correlation between the shock spectrum of a pressure pulse and the extent of the spurious oscillation can then be established. This enables one to "tune" the grid size to the relative minima of the shock spectra of the applied forces. For example, Figure 4 shows shock spectra for the calculation on a one-dimensional finite element grid as shown in Figure 3. It can be seen that if element sizes (t_r/τ) are selected in such a way that relative minima of shock spectra can be obtained the spurious oscillations can be minimized.

For good numerical accuracy, the time step for numerical integration can be taken to be less than or equal to 1/20 of the shortest element transit time. Figures 5, 6, and 7 show the time histories of stress which were computed. These figures illustrate two trends: First overshoot due to discretization is a function of depth; second, depending upon rise time overshoot may be greater at depth than at the surface. One explanation of these effects is that deeper elements participate more strongly in slightly lower frequency modes and hence, may have either higher or lower overshoot depending upon the ratio $\frac{t_r}{\tau}$ where τ is the most strongly represented mode. Another possibility which these figures suggest is that other rise conditions, e.g., versed sine or cycloid could be considered when waveforms other than step functions are to be modelled.

Figure 8 shows velocity waveforms. In general, velocity overshoots are less than stress overshoot.

DAMPING CONSIDERATION

Many of the today's finite element structural and continuum dynamics analyses assume the proportional damping, that is

$$C = \alpha M + \beta K$$

in

$$M\ddot{x} + Cx + Kx = F$$

Using the modal approach, one reduces the equation to

$$\ddot{q}_n + 2\gamma_n \dot{q}_n + \omega_n^2 q_n = \frac{Q_n}{m_n}$$

where

$$\gamma_n = 0.5 \left(\alpha + \beta \omega_n^2 \right)$$

This indicates that if $\beta \neq 0$, there is a possibility that low frequency modes might be underdamped while high frequency modes were being overdamped.

P-WAVE VELOCITY
10,000 FT/SEC

10 FT ELEMENT
NARROW BAND

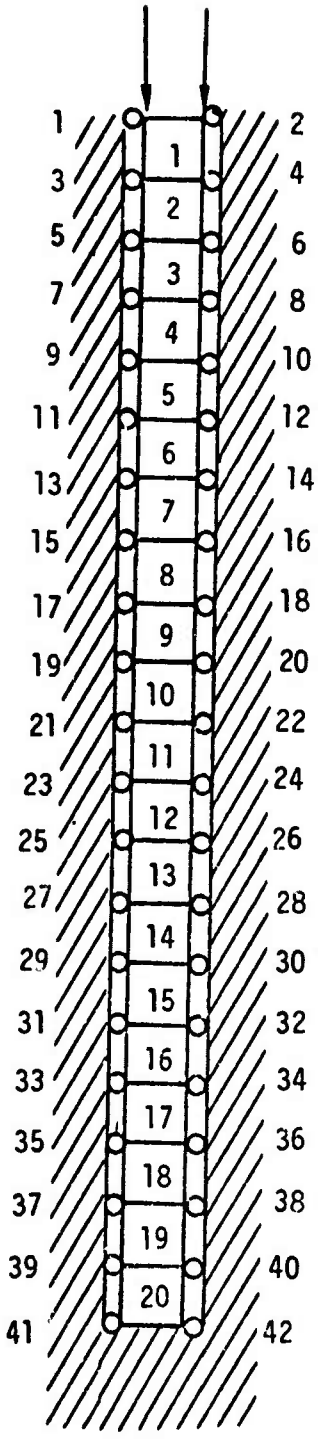


FIGURE 3. GEOMETRY OF GRID FOR RISE-TIME SENSITIVITY STUDY

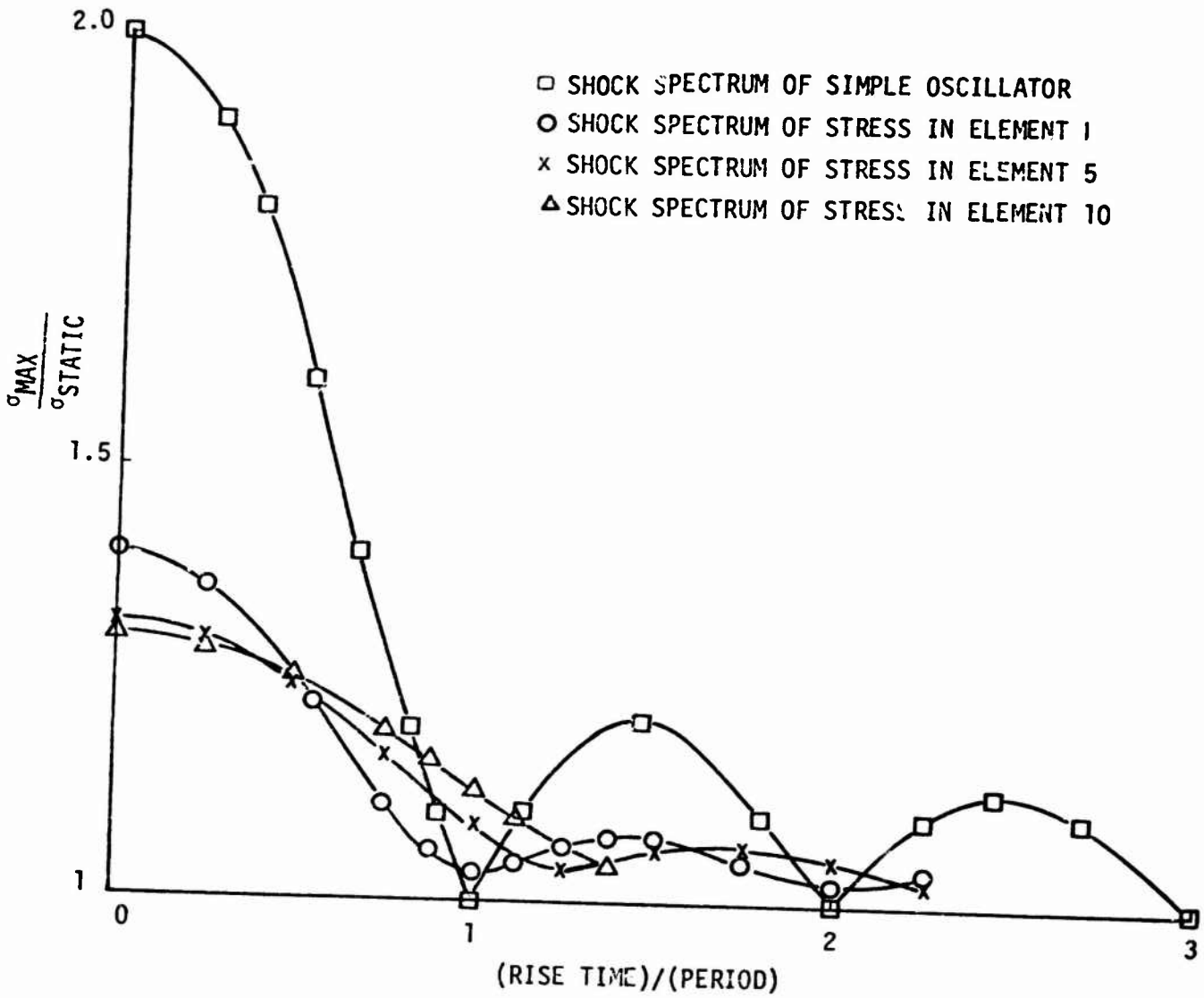


FIGURE 4. SHOCK SPECTRA FOR RISE-TIME SENSITIVITY STUDY

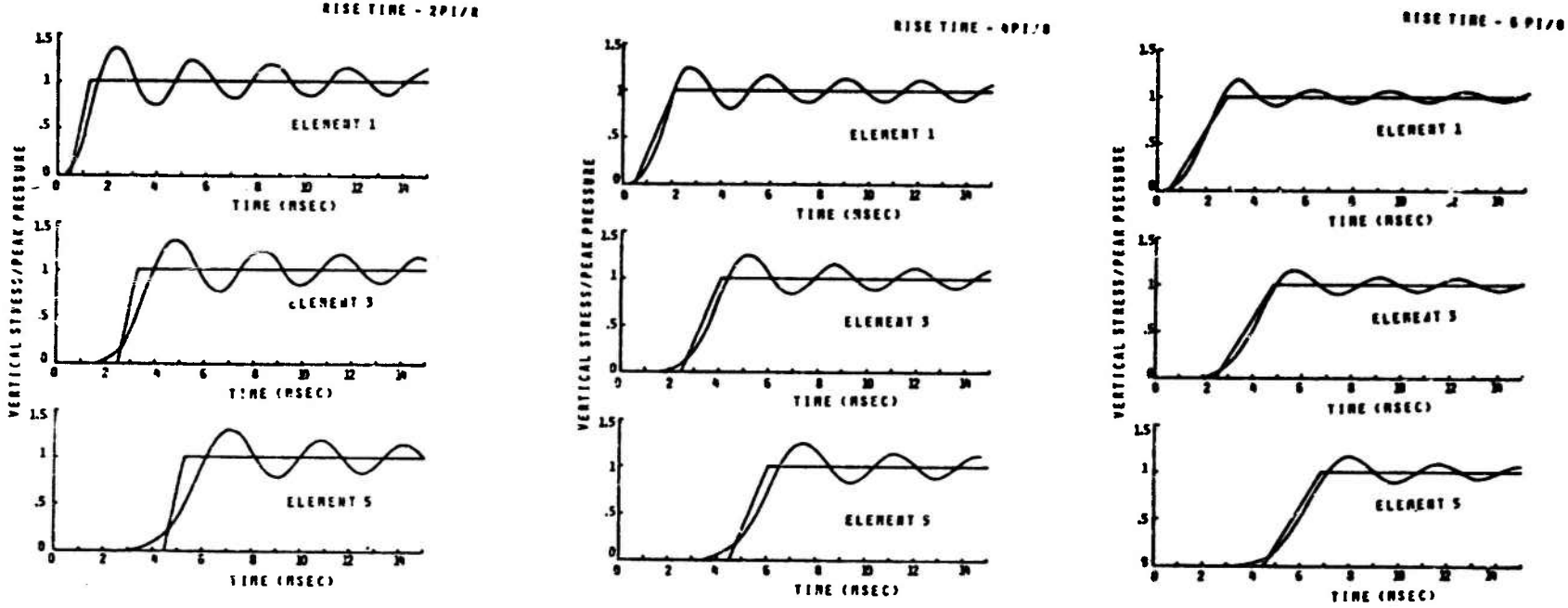


FIGURE 5. STRESS TIME HISTORIES RISE TIMES: $2\pi/8$, $4\pi/8$, $6\pi/8$

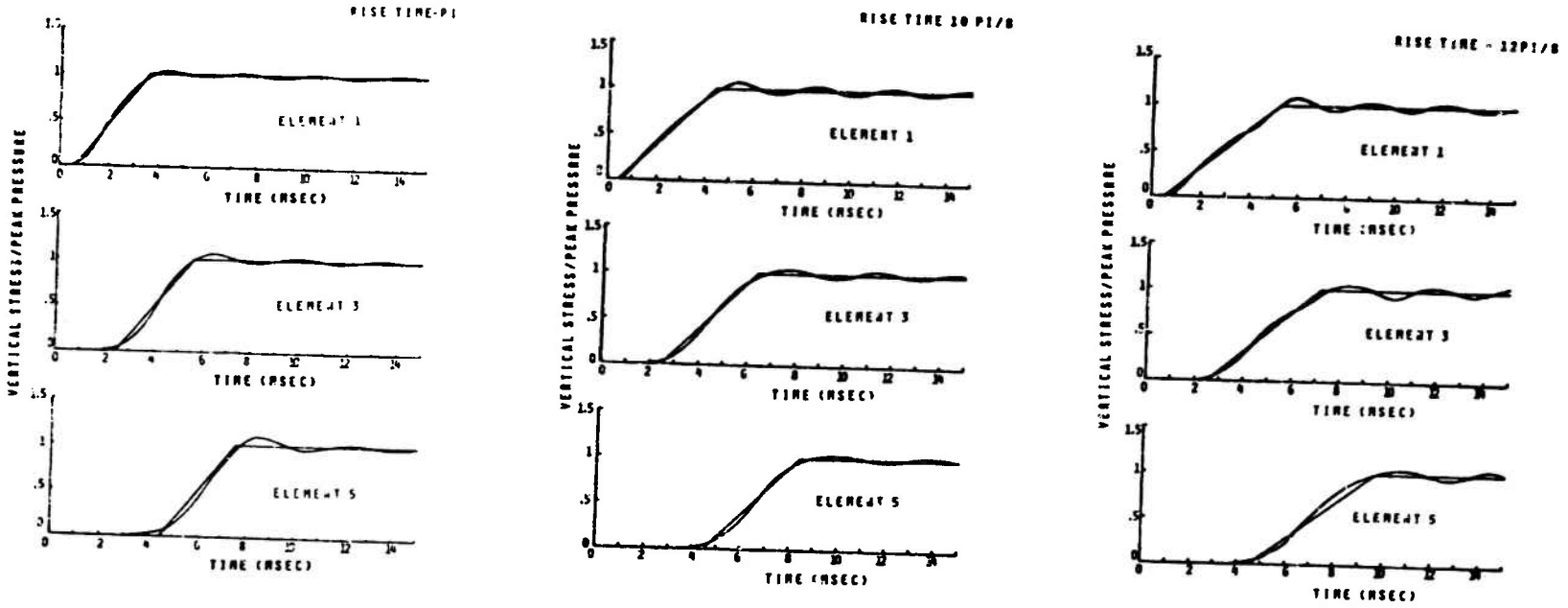


FIGURE 6. STRESS TIME HISTORIES RISE TIMES: π, $10\pi/8$, $12\pi/8$

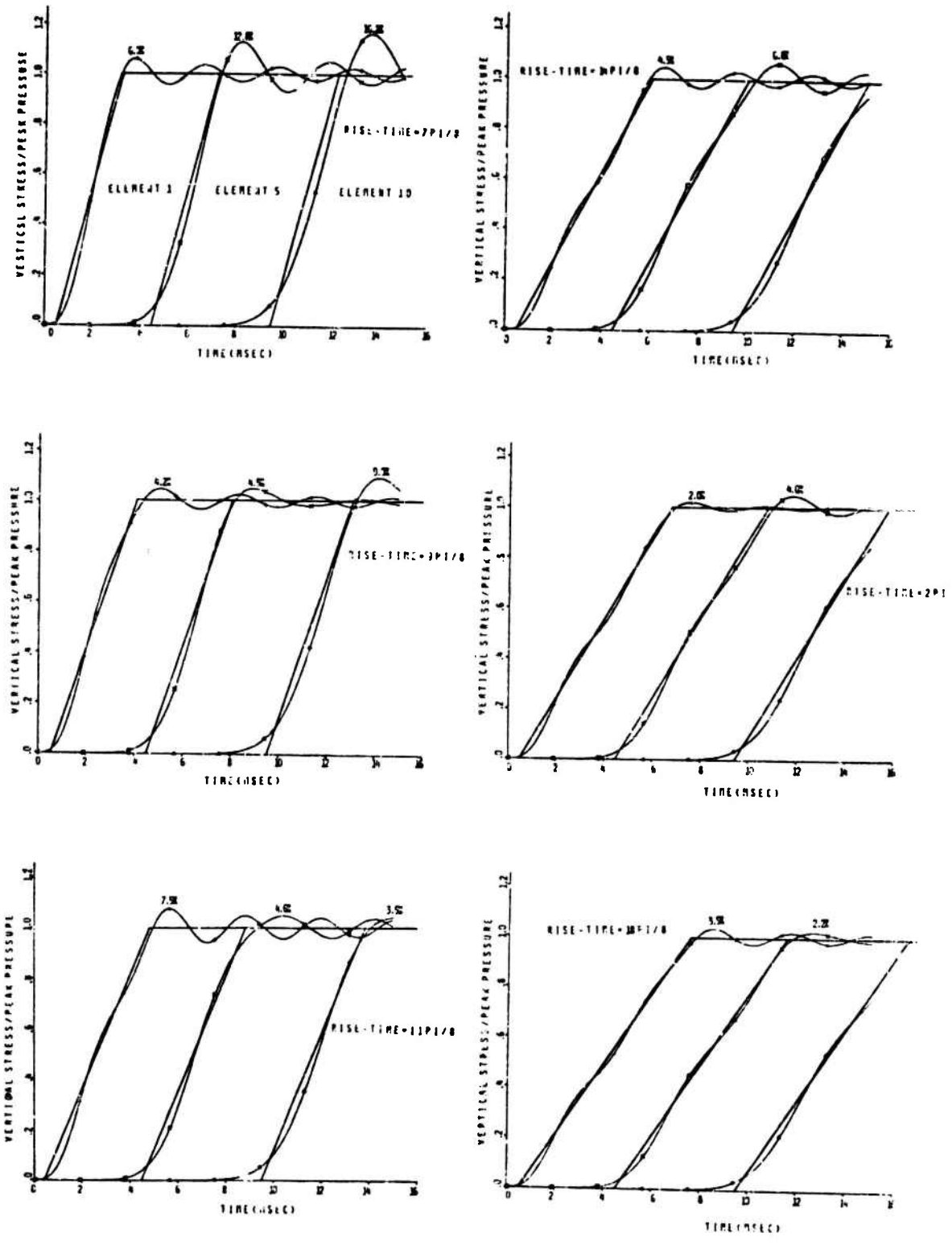


FIGURE 7. STRESS TIME HISTORIES
RISE TIMES: $7\pi/8$, $9\pi/8$, $11\pi/8$, $14\pi/8$, 2π, $18\pi/8$

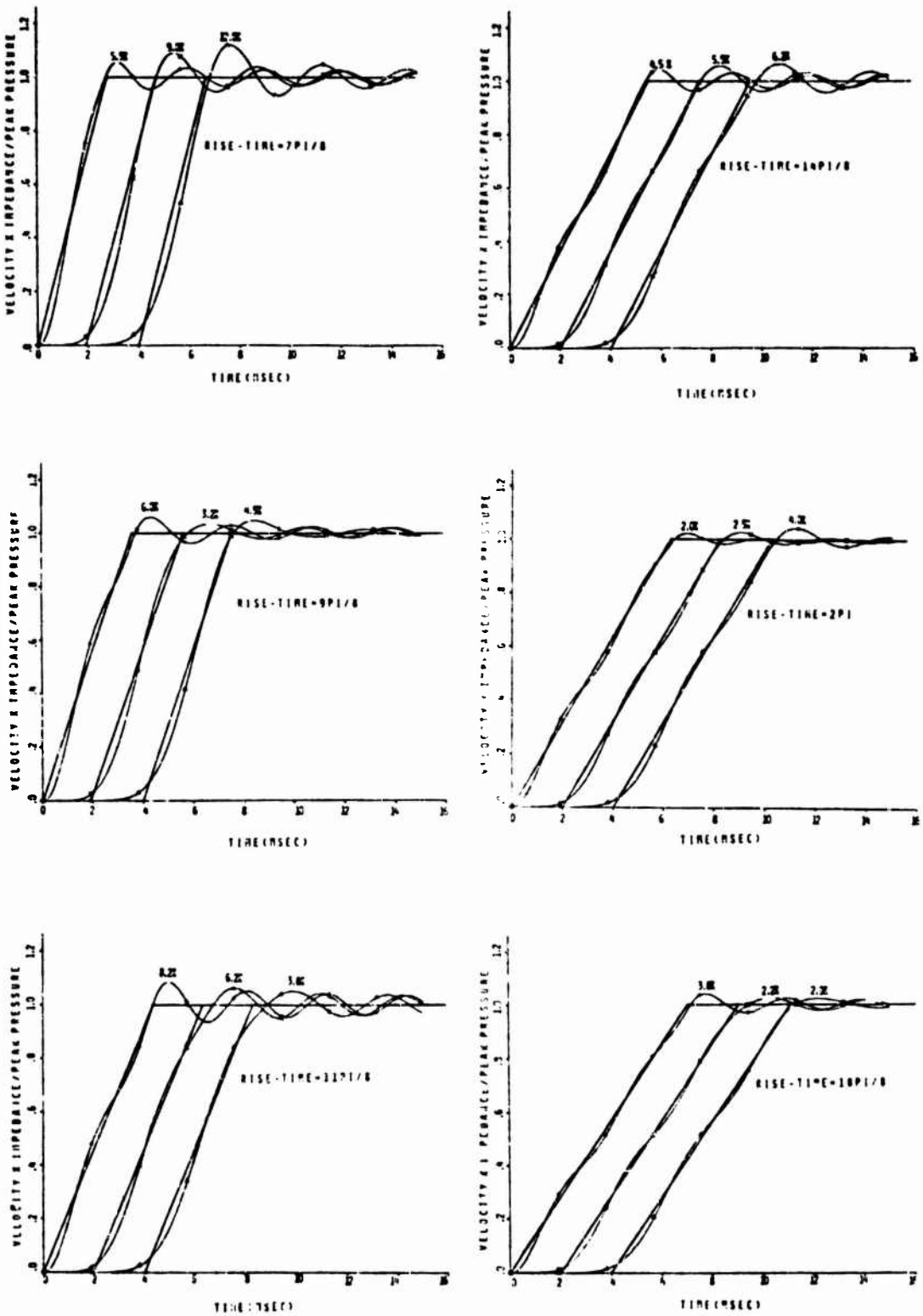


FIGURE 8. VELOCITY TIME HISTORIES
RISE TIMES: 7π/8, 9π/8, 11π/8, 14π/8, 2π, 18π/8

Critical damping occurs for the n-th mode when

$$\omega_n = \gamma_n$$

three cases were examined. If $\beta = 0$, $\alpha = 2\omega_n$ yields the critical damping. This indicates that for $\beta = 0$, modes with frequencies less than $\alpha/2$ will be overdamped and those with frequencies greater than $\alpha/2$ will be underdamped. If $\alpha = 0$, $\beta = 2/\omega_n$ yields the critical damping, and modes with $\omega_n < (2/\beta)$ are underdamped while higher modes, $\omega_n > (2/\beta)$ are overdamped. If α and β are both nonzero, regions are defined for α and β where the system is either overdamped or underdamped. It is obvious that of the damping coefficients α and β, only β has the effect of damping high frequency components. This means that if physical effects actually cause a high frequency cut off then β can be used to simulate such effects. For most structural analyses, one finds it appropriate to consider $C = \beta K$ if β can be chosen properly. A series of finite element calculations were made to study the sensitivity of finite element solution to the selection of coefficient β. Figure 9 presents some of the results. Three observations were made:

- Wave propagation phenomena can not be properly simulated by a grid of finite elements whose highest mode is critically damped.
- Structural damping is not effective in reducing peak spurious oscillations due to finite element discretization of a continuum.
- Calculation of high frequency structural response using the finite element method requires an accurate choice of the damping coefficient, β.

Although above discussions apply primarily to linear elastic one-dimensional P-wave propagation, similar observations have been made on the spurious shear oscillations, that is, such oscillations can also be minimized by using rise times proportional to the transit time of a shear wave across an element. Moreover, the elastic P-wave rise-time criteria have been successfully applied to axisymmetric and plane strain shock propagation in nonlinear media.

Ideally, one would prefer to choose a very fine mesh of elements so that all pertinent wave phenomena (dilitation, shear, Rayleigh, etc.) could be modelled with little spurious oscillation. Unfortunately, many blast and shock problems do not lend themselves to very fine grids. Sometimes one may use a priori knowledge of physics or test results to reduce the number of finite elements. For example, explosions in underground cavities are known to cause shock which has longer rise times at greater ranges. Hence, a grid having larger elements at greater range is appropriate. On the other hand, the modelling of the high frequency response

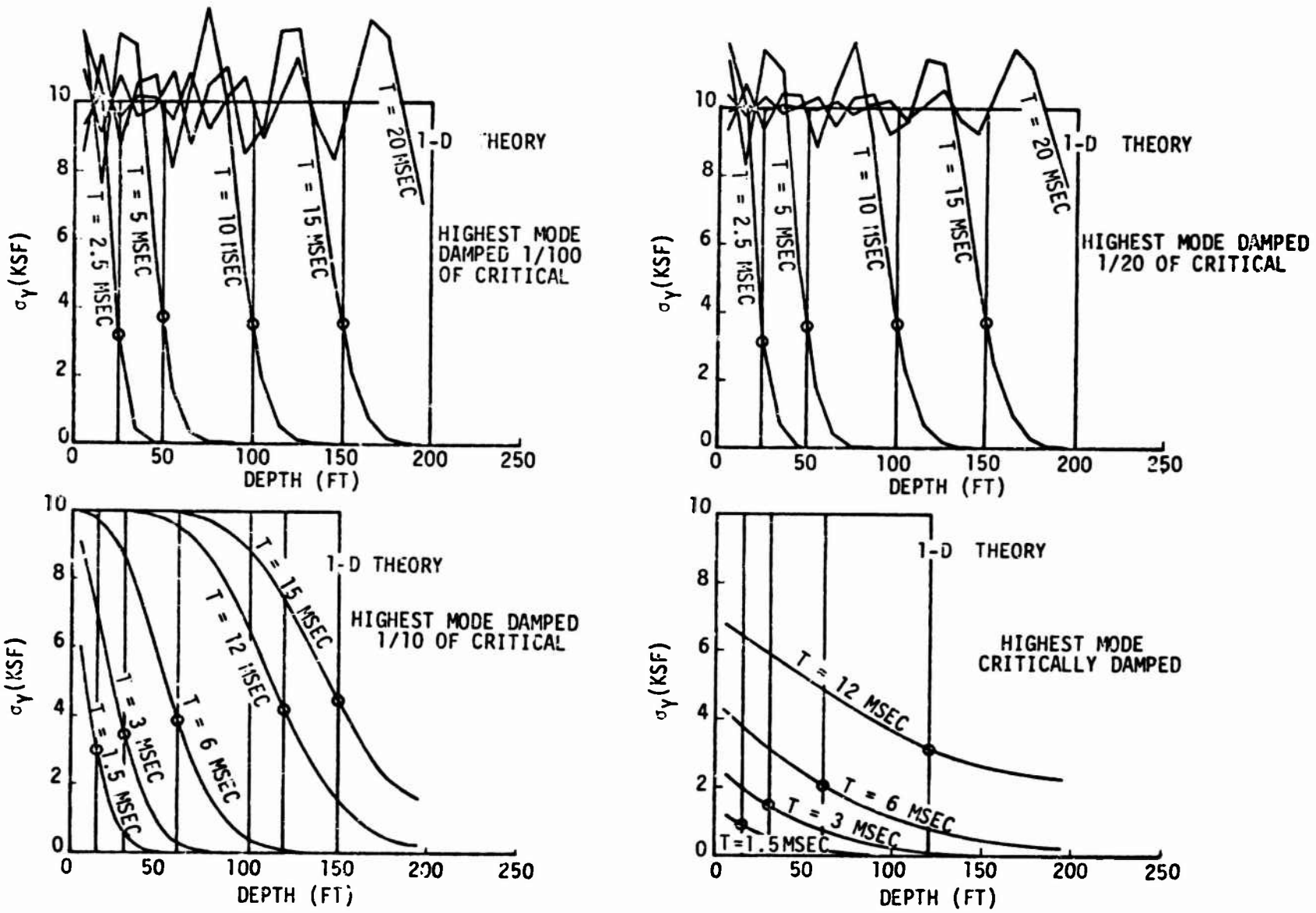


FIGURE 9. STRESS SPATIAL DISTRIBUTIONS FOR VISCOUS EFFECTS

of a stress gage in a cavity requires a different approach. Figure 10 shows these two problems which are modelled by the same grid under different loading conditions. In one case (the contained explosion) the grid is appropriate. In the other case (the high frequency stress gage) the grid is inappropriate. The reasons for the differences are threefold.

- The physics of underground explosions, as verified by test, is such that the high frequency content of the groundshock is attenuated with range (by both linear and nonlinear effects).
- The goals of the two analyses could be different i.e., low frequency displacements for the underground explosion vs high frequency changes in stress.
- The loading conditions are such that the grid adequately models the participation of the near cavity elements for the underground explosion while the coarse exterior grid limits the high frequency participation of the cavity under an external load.

CONCLUSION

If one considers the following effects, one has a better chance of achieving ones computational goals economically by finite element analyses:

- Finite element grids cannot propagate shocks whose frequency content is higher than the highest modal frequency of the grid.
- Spurious oscillation associated with the highest modal frequencies of a finite element mesh may be minimized by tuning the grid to the forcing function by shock spectral considerations.*
- The modal participations of the forcing functions and critical elements should be weighed carefully against computational goals when grids are made nonuniform for economic reasons.
- Structural damping becomes highly important for high frequency shocks; sharp shocks cannot propagate in and overdamped mode.

---

* Vector plots showing the spatial velocity field or stress fields are most easily interpreted if spurious oscillations are suppressed. Moreover spurious oscillations on codes with nonlinear material finite elements can cause irreversible nonlinear effects.

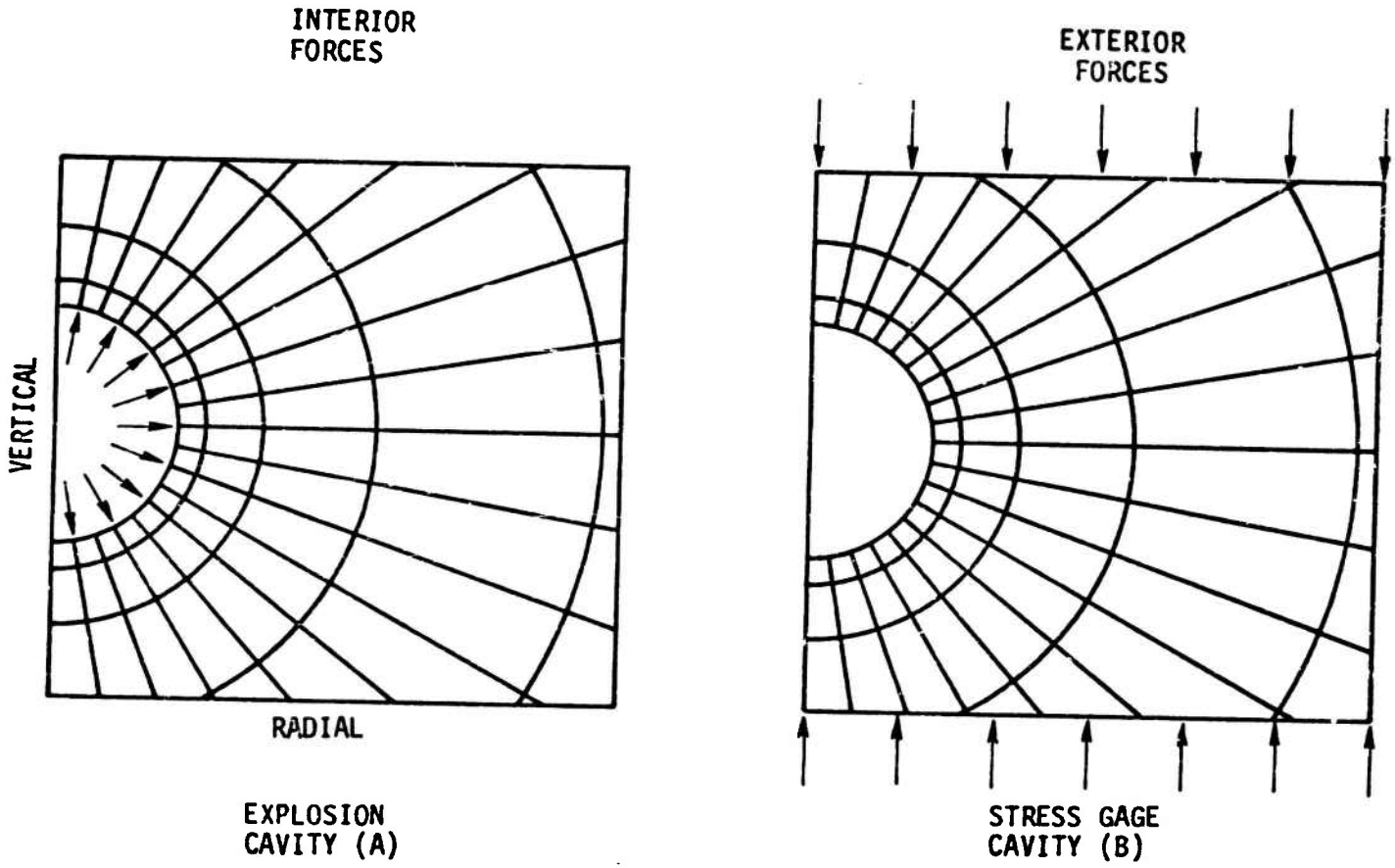


FIGURE 10. EXAMPLES OF MODAL PARTICIPATIONS OF LOAD

REFERENCES

1. Ang, Alfredo H. S., "Numerical Approach for Wave Motions in Nonlinear Solid Media, Proceedings of Conference on Matrix Methods in Structural Mechanics," Wright-Patterson AFB, Ohio, October 1965.

2. Smith, Rodney H. and Newmark, N. M., "Numerical Integration for One-Dimensional Stress Waves," University of Illinois Civil Engineering Studies, Structural Research Series No. 162, Urbana, Illinois, August 1958.

3. Wilson, Edward L., "A Computer Program for the Dynamic Stress Analysis of Underground Structures," Report No. 68-1 Structural Engineering Laboratory, College of Engineering, University of California, Berkeley, California, January 1968.

ACKNOWLEDGEMENTS

This study was performed with the support of USAF, SAMSO contract No. F04701-71-C-0065. The authors wish to thank Messrs. B. Menold and T. Wang of TRW's Software and Information Systems Division for their computational efforts.

AN APPLICATION OF FINITE ELEMENT METHODS TO PANEL FLUTTER OPTIMIZATION

Terry A. Weisshaar
University of Maryland*

This paper presents the results of an aeroelastic optimization study. The weight of a panel in high Mach number supersonic flow is minimized subject to the requirement that the critical aerodynamic parameter for flutter be held within prescribed limits. Panel equilibrium is described with finite element techniques. The study shows that finite element techniques, when used together with standard optimization methods, yield accurate results to this difficult problem. The numerical results of the investigation are compared to other converged numerical results to illustrate accuracy. The optimization mechanism itself is studied and discussed to provide qualitative results which may be applied to other aeroelastic optimization problems. The results of the paper show that, with an effective finite element model, weight savings and mass distributions found with finite element methods are comparable to those found by more complicated methods.

Nomenclature

A_{ij}, a_{ij} = System and element aerodynamic matrix elements, respectively (nondimensional)

a = Plate chordwise dimension

$D(x)$ = Plate chordwise bending stiffness

g_i = Constraint boundary derivatives $(\partial\lambda_{cr}/\partial t_i)$

grad( ) = Gradient of a function with respect to design variables

J = Objective function

K_{ij}, k_{ij} = System and tapered element stiffness matrix elements, respectively (nondimensional)

$m(x)$ = Mass per unit area

MR = Mass ratio (equation 13)

M_{ij}, m_{ij} = System and tapered element mass matrix elements, respectively (nondimensional)

n = Number of equal-length tapered elements used in panel model

q_o = Dynamic pressure (equation 5)

q_i = Nondimensional panel nodal displacements

S_i = Elements of step direction vector

* Assistant Professor, Aerospace Engineering

t_i = Design variable, nondimensional nodal thickness

$T(x)$ = Dimensional thickness

w_i = Elements of the gradient of the unconstrained objective function

x = Nondimensional chordwise coordinate ($x = x/a$)

ω = Panel oscillation frequency

Y_i = Tapered element nondimensional nodal displacement (Fig. 2)

δ_i = Ratio of face-sheet mass to total mass, reference panel

ϵ = Optimization step size

λ_o = Aerodynamic parameter (equation 5)

τ = Time

ω_r = Reference frequency, $\pi^2(D_o/m_o a^4)^{\frac{1}{2}}$ (equation 4)

$\{\ \}$ = Column matrix

$[\ ]$ = Row matrix

$[\ ]$ = Square matrix

1.0 Introduction

The aeroelastic optimization of structures involves the combination of two well-developed, sophisticated scientific disciplines; aeroelasticity and optimization. This paper will study a panel flutter optimization problem and discuss the results together with some of the difficulties encountered. The term "panel flutter optimization", as used in this paper, refers to the search for a least-weight design of a panel in high Mach number supersonic flow. The search for this optimum design is constrained by the requirement that a critical aerodynamic parameter for flutter be held within specified design limits.

Panel flutter optimization studies are of recent origin. Ashley and McIntosh (Ref. 1) presented a differential equation approach to panel flutter optimization. In their analysis, the equilibrium differential equations and the associated eigenvalues were treated as constraints. Variational calculus methods provided the equations necessary for panel optimality. However, they did not present a solution to their equations. The differential equation, with variational calculus, approach used in Ref. 1 has both advantages and difficulties. The primary advantage of this approach is that a great deal of useful information about the mathematical properties of the optimum solution may be found. If a solution can be found, it often provides useful insight into other, more sophisticated problems. The main difficulties of this approach arise from the face that only relatively simple structures may be described with differential equations. Also, although the governing equations are easily found, their solution is usually difficult. Two-point, nonlinear boundary value problems are encountered and often must be solved by numerical techniques. The convergence of these techniques are not always assured.

Turner (Ref. 2) approached the panel flutter optimization in a different manner. His panel model consisted of a series of finite elements, each with a different constant thickness. He then used the resulting matrix equations and variational methods to iteratively solve for a least-weight design. The numerical calculations included only several design variables and the weight savings were insignificant. More importantly, however, Turner studied the governing equations and deduced the presence of a least-weight design whose mass distribution is symmetric about the panel midchord. While this approach sacrifices some solution accuracy, it does eliminate some of the difficulties encountered with the differential equation approach. However, Turner's approach to the optimization portion of the problem is also subject to convergence difficulties.

Another recent study (Ref. 3) presented a converged numerical solution to a panel flutter optimization problem. This study used the differential equation approach of Ref. 1. There are several characteristics of this solution which determined the choice of a finite element model for the present study. First of all, the boundary conditions of the problem show that the panel thickness, in the absence of a minimum thickness requirement, must be zero at both the leading and trailing edges. Also, the resulting least-weight thickness distribution changes drastically from one panel chordwise position to another. A panel which is composed of uniform thickness elements is incapable of satisfying the zero-thickness boundary conditions. Because of this, an element whose thickness varies linearly from one end to the other was chosen to model the system. A preliminary study in Ref. 3 showed good results were possible with this "tapered element" approach. The study was, however, not comprehensive. The present study seeks to incorporate all the best features of the previous work in panel flutter optimization to demonstrate a method whose attributes are ease of application and accuracy of results.

2.0 The Panel Flutter Problem

Figure 1 shows a panel with supersonic airflow on one side and "dead" air on the other. The panel rests on simple supports and is of sufficiently large extent in the spanwise direction that its behavior may be considered that of a one-dimensional or semi-infinite panel. If the Mach number is greater than 1.6, then one-dimensional, linearized, quasi-steady, supersonic theory will adequately describe the airloads generated by panel oscillation (Ref. 4). The elastic and inertial behavior of the panel is described by one-dimensional plate equations. Thus, the plate equations are identical to beam equations with $D(x)$ substituted for the bending stiffness and $m(x)$ substituted for mass per unit length. This study will consider a panel composed of two variable thickness face-sheets with a uniform sheet of nonstructural material sandwiched between them.

Because of the sandwich construction, the plate bending stiffness is proportional to the face-sheet thickness. If D_o represents a reference plate stiffness, then

$$\frac{D(x)}{D_o} = \frac{T(x)}{T_o} = t(x) \qquad (1)$$

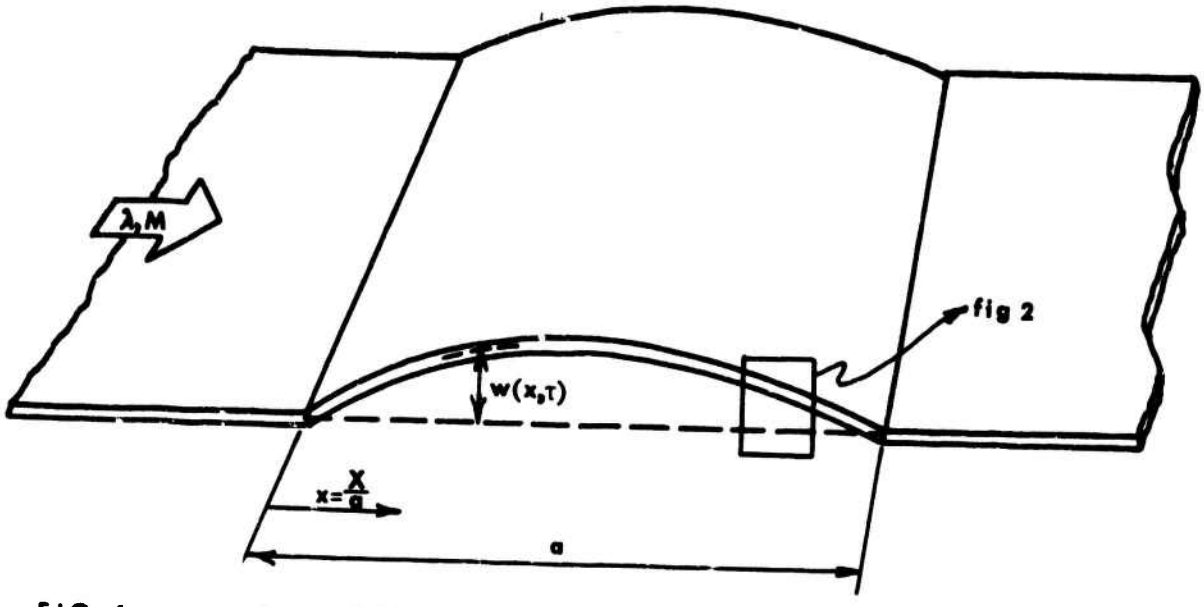


FIG 1 ONE-DIMENSIONAL PANEL FLUTTER MODEL

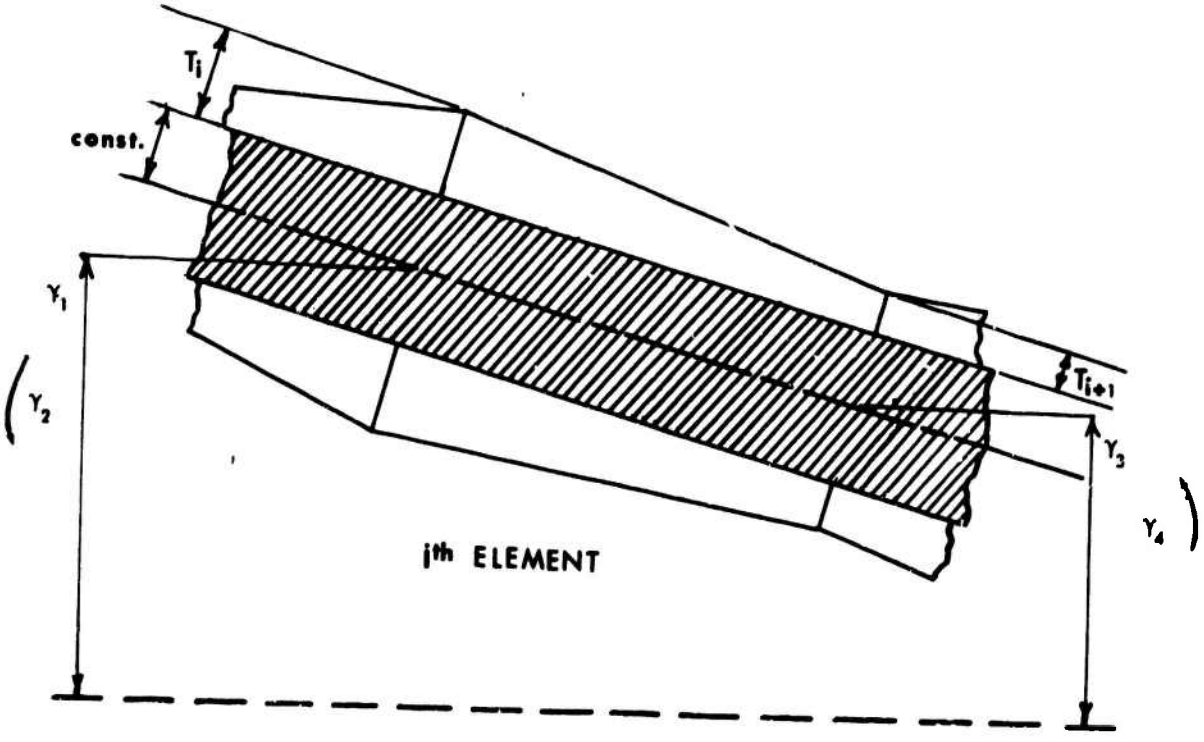


FIG 2 FINITE ELEMENT MODEL

The variable $t(x)$ is a nondimensional face-sheet thickness. Similarly, the mass per unit area may be nondimensionalized.

$$\frac{m(x)}{m_o} = \delta_1 t(x) + (1 - \delta_1) \tag{2}$$

The constant δ_1 is the ratio of the total weight of the face-sheets of a uniform-thickness, reference panel to the total weight of the reference panel.

Consider a panel centerline displacement of the form

$$W(x,\tau) = \xi(x)e^{i\alpha\tau} \tag{3}$$

where α is, in general, a complex number $\alpha = \omega + i\beta$. The differential equation of equilibrium for the panel can be written in nondimensional form as (Ref. 4, pp. 418-422),

$$(t\xi'')'' + \lambda_o \xi' + \left(\frac{i\alpha}{\omega_r} g_\alpha\right)\xi(x)\pi^4 - \frac{\alpha^2}{(\omega_r)^2}\pi^4(\delta_1 t + (1 - \delta_1))\xi(x) = 0 \tag{4}$$

where ω_r is a reference frequency. The symbol $(\,)'$ denotes differentiation with respect to x. The parameter λ_o is an aerodynamic parameter given by

$$\lambda_o = 2q_o a^3/D_o(M^2 - 1)^{\frac{1}{2}} \tag{5}$$

while g_α, the aerodynamic damping parameter, is given by

$$g_\alpha = (M^2 - 2)\rho U/m_o\omega_r(M^2 - 1)^{3/2} \tag{6}$$

where ρ and U are the air density and speed, respectively.

If $t(x)$ is equal to unity, equation (4) is the equilibrium equation of a reference panel, with constant thickness face-sheets, oscillating in supersonic flow. For a given value of g_α, the value of the frequency, α, is a function of λ_o. The stability problem is simplified considerably if g_α is equal to zero. Since the case where g_α equals zero represents a realistic problem (Ref. 4, p. 422), this assumption will be used throughout this paper. For a certain range of values, with zero aerodynamic damping,

$$0 \leq \lambda_o < \lambda_{cr}$$

all values of α are real and distinct. However, for the value $\lambda_o = \lambda_{cr}$ the first two frequencies merge or become equal. For $\lambda_o > \lambda_{cr}$, the lowest two frequencies, α_1 and α_2, are complex conjugates of one another. Because of equation (3), the motion, for values of λ_o above λ_{cr}, will be divergent with time. Thus λ_{cr} corresponds to a situation where the panel motion is neutrally stable. Neutral stability, for a uniform thickness or reference panel, is found to occur (Ref. 4, p. 422) when

$$\lambda_{cr} \cong 3.52\pi^4 \tag{7}$$

The design problem involves the search for a least-weight panel design whose critical value λ_{cr} is the same as that of the heavier reference panel.

Olson (Ref. 5) solved the uniform thickness, one-dimensional, panel flutter problem with a finite element description of the panel and the airloads. His results showed that λ_{cr} could be determined accurately with only a few elements. The present study uses Olson's consistently derived airload matrix.

2.1 The Finite Element Model

For oscillatory motion of the type given in equation (3), the nondimensional panel equilibrium matrix equation, equivalent to equation (4), is (Ref. 5)

$$\left[-\left(\frac{\alpha}{\omega_r}\right)^2 \pi^4 \left[M_{ij}\right] + \left[K_{ij}\right] + \lambda_o \left[A_{ij}\right] \right] \left\{ q_i \right\} = \left\{ 0 \right\} \tag{8}$$

The variables q_i are nondimensional displacements and rotations at the panel node points. The matrices $[M_{ij}]$, $[K_{ij}]$, and $[A_{ij}]$ are nondimensional mass, stiffness and aerodynamic matrices, respectively. A typical element for the panel is shown in Fig. 2. This element has a face-sheet thickness which varies linearly from one node point to another. The panel model consists of a series of n of these tapered elements, connected to each other at their node points. The modal deflection $\xi(x)$ of an element is assumed to be a third order polynomial, just as was done by Turner in Ref. 2. The element stiffness and mass matrices were rederived in nondimensional form by considering the strain energy and kinetic energy of the tapered element. In one form or another, these tapered element matrices are well-known. The nondimensional element matrices for the tapered element shown in Fig. 2 are given by

$$\left[m_{ij}\right] = \frac{z f_1}{420} \begin{bmatrix} 156 & 22z & 54 & -13z \\ & 4z^2 & 13z & -3z^2 \\ & & 156 & -22z \\ \text{(Symmetric)} & & & 4z^2 \end{bmatrix}$$

$$+ \frac{\delta_1 z (t_{p+1} - t_p)}{840} \begin{bmatrix} 72 & 14z & 54 & -12z \\ & 3z^2 & 14z & -3z^2 \\ & & 240 & -30z \\ \text{(Symmetric)} & & & 5z^2 \end{bmatrix} \tag{9a}$$

$$p = 1, 2, \cdots n$$

where $f_1 = 1 + \delta_1 (t_p - i)$ and $z = i/n$.

$$\left[k_{ij}\right] = \frac{2t_p}{z^3}\begin{bmatrix} 6 & 3z & -6 & 3z \\ & 2z^2 & -3z & z^2 \\ & & 6 & -3z \\ \text{(Symmetric)} & & & 2z^2 \end{bmatrix} + \frac{t_{p+i} - t_p}{z^3}\begin{bmatrix} 6 & 2z & -6 & 4z \\ & z^2 & -2z & z^2 \\ & & 6 & -4z \\ \text{(Symmetric)} & & & 3z^2 \end{bmatrix} \tag{9b}$$

The element airload matrix is, from Ref. 5,

$$\left[a_{ij}\right] = \begin{bmatrix} \frac{1}{2} & \frac{z}{10} & \frac{1}{2} & -\frac{z}{10} \\ & 0 & \frac{z}{10} & -\frac{z^2}{60} \\ & & \frac{1}{2} & \frac{z}{10} \\ \text{(anti-symmetric)} & & & 0 \end{bmatrix} \tag{9c}$$

Element equilibrium is, in terms of these matrices, given by

$$\left[-\left(\frac{\alpha}{\omega_r}\right)^2 \pi^4 \left[m_{ij}\right] + \left[k_{ij}\right] + \lambda_o \left[a_{ij}\right]\right]\left\{\gamma_j\right\} = \left\{0\right\} \tag{10}$$

By using a compatibility matrix and the simple support boundary conditions, these matrices may be used to form the system matrices $[M_{ij}]$, $[K_{ij}]$, and $[A_{ij}]$ in equation (8). The nondimensional thickness parameters at each node point, t_i, are design variables. The value of λ_{cr} is determined in the manner described in Section 2.0. For a given value of δ_1, the value of λ_{cr} is a function of the design variables, t_i.

$$\lambda_{cr} = \lambda(\delta_1, t_i) \tag{11}$$

2.2 The Objective Function

The object of this study is to reduce panel weight by varying only the face-sheet thickness. A measure of the panel face-sheet weight is given by

$$J = \int_0^1 t(x)dx \tag{12}$$

The integral J represents the ratio of the face-sheet weight of a variable thickness panel to the face-sheet weight of the reference panel. The ratio of

the total weight of a variable thickness panel to that of the reference panel is called the mass ratio, MR.

$$MR = \delta_1 \int_0^1 t(x)dx + (1 - \delta_1) \tag{13}$$

Since the nonstructural mass is not disturbed, the only way to reduce the mass ratio is to reduce J. With a tapered element approach, the objective function is given by

$$J = \frac{1}{n}\left((t_1 + t_{n+1})/2 + \sum_{i=2}^{n} t_i\right) \tag{14}$$

For the reference panel, J is equal to unity. The objective of this study will be to find a minimum value of J, as defined by equation (14), subject to the flutter parameter constraint.

2.3 The Flutter Parameter Constraint

The search for a minimum to the objective function defined in equation (14) is constrained by the requirement that the flutter parameter be held within design limits. Theoretically, the design requirements are that the least weight panel have a value λ_{cr} equal to that of a similar, uniform thickness, reference panel. But, because of linear approximations used in the optimization procedure, the flutter constraint will, in practice, be an inequality constraint.

$$\begin{aligned} &\lambda_1 < \lambda_{cr} < \lambda_2 \\ &\lambda_1 = (0.995)\lambda_{ref} \\ &\lambda_2 = (1.005)\lambda_{ref} \end{aligned} \tag{15}$$

The value λ_{ref} is the flutter parameter for the reference panel. Thus, a 0.5% variance on either side of the reference value is permissible. If, at any time, the value of λ_{cr} for the variable thickness panel falls outside these limits, the design must be modified to bring it back within design specifications of equation (15). λ_{ref} will be determined by finite element analysis.

The derivatives of the theoretical design constraint, λ_{cr} = constant, with respect to changes in design variables, are necessary to the optimization procedure. These derivatives are calculated by perturbing, one at a time, each value of t_i by an amount Δt_i. The value of λ_{cr} due to this perturbation is then used to calculate the derivative.

$$\frac{\partial \lambda_{cr}}{\partial t_i} \simeq \frac{\Delta \lambda_{cr}}{\Delta t_i} \tag{16}$$

2.4 The Optimization Problem

Two possible approaches to the optimization portion of the problem are available. The first approach is to convert the constrained minimization problem to an unconstrained minimization problem. This method is called elimination (Ref. 6). The constraint that λ_{cr} be fixed at a certain value during optimization may be expressed mathematically as

$$d\lambda_{cr} = \sum_{i=1}^{n+1} \frac{\partial \lambda_{cr}}{\partial t_i} dt_i = 0 \tag{17}$$

In terms of finite, small, design variable changes Δt_i, a truncated Taylor series gives

$$\Delta\lambda_{cr} = \sum_{i=1}^{n+1} \frac{\partial \lambda_{cr}}{\partial t_i} \Delta t_i = \sum_{i=1}^{n+1} g_i \Delta t_i = 0 \tag{18}$$

Equation (18) may be used to eliminate one of the design variable changes.

Since the objective function is given by equation (14), changes in J are given by

$$\Delta J = (\Delta t_i + \Delta t_{n+1})/2n + \sum_{i=2}^{n} \Delta t_i/n \tag{19a}$$

For simplification let us define new coefficients such that

$$\Delta J = \sum_{i=1}^{n+1} w_i \Delta t_i \tag{19b}$$

These coefficients w_i are elements of the gradient of J. Equation (18) may be used to eliminate a design variable change, call it Δt_k, from equation (19b). Equation (19b) then becomes

$$\Delta J = \sum_{i=1}^{n+1} G_i \, t_i = \sum G_i J \{\Delta t_i\} \qquad i \neq k \tag{20a}$$

where

$$G_i = (w_i g_k - w_k g_i)/g_k \tag{20b}$$

The matrix elements G_i are also those of a gradient of J, but with the constraints included.

A combination of design variable changes must be determined which reduces the value of the objective function. That is, a vector of design variable changes Δt_i must be formed such that ΔJ, given in equation (20a) is negative. If one thinks of searching through a multi-dimensional design space, then a "direction vector" is found to guide the search from one design to another, lesser weight design. The step direction vector given by $\{S\}$ can be formed in many different ways, each with its own advantages and disadvantages. In terms of $\{S\}$, the design variables changes are given by

$$\{\Delta t_i\} = \epsilon \{S_i\} \tag{21}$$

where ϵ is a constant which determines the step size. Note that in the unconstrained problem, $\{S\}$ is an n-dimensional vector. The change Δt_k can then be calculated from equation (18).

Rubin (Ref. 7) presented an automated method for solving constrained natural frequency problems. He chose his step direction vector as

$$\left\{S_i\right\} = \left\{\frac{i}{|G_i|_{max}} t_i G_i\right\} \qquad i \neq k \tag{22}$$

This method was used with good results in Ref. 3 for a panel flutter optimization problem with a large minimum thickness constraint.

Many different optimization techniques are discussed in a recent book by Fox (Ref. 8). One method which was found effective in this study was the Fletcher-Reeves or conjugate gradient method (Ref. 8, pp. 87-89). This method eliminates some difficulties which are encountered with the steepest descent method. The Fletcher-Reeves method chooses $\{S\}$ as follows.

$$\left\{S_i\right\}_{new} = -\left\{G_i\right\}_{new} + \beta \left\{S_i\right\}_{old} \tag{23}$$

The constant β is found from the relation

$$\beta = \sum_{i=1}^{n+1} (G_i)^2_{new} \Big/ \sum_{i=1}^{n+1} (G_i)^2_{old} \qquad i \neq k \tag{24}$$

In addition to the elimination method, a second method of approaching the optimization problem exists. This approach does not directly eliminate the constraints. Instead, the constraints are incorporated into the direction finding problem directly. With this approach, the vector $\{S\}$ is an $n + 1$ dimensional vector. One such method of finding a step direction is the method of feasible directions (Ref. 9). This technique has been used with excellent results on other flutter optimization problems (Ref. 10). An easily readable explanation of this technique is also given in Ref. 8. The choice of a step direction is shown to be an optimization problem in itself.

A simplification of the feasible directions technique reduces it to a gradient projection method. A complete discussion of the mathematical aspects of this technique are given by Fox (Ref. 8, pp. 179-196). The step direction is determined by the relation

$$\{S_i\} = -\text{grad}\ (J) - c(\text{grad}\ (\lambda_{cr})) \tag{25}$$

where the elements of grad J are w_i and the elements of grad λ_{cr} are the $n + 1$ partial derivatives in equation (16). With equation (25), all components of grad J parallel to the normal to the constraint boundary λ_{cr} = constant are subtracted from that vector. The constant, c, in equation (25) is then determined by the relation

$$\lfloor S_i \rfloor \text{grad}\ (\lambda_{cr}) = 0 \tag{26}$$

It should be noted that if grad λ_{cr} is a linear function of grad J, then no progress can be made because $\{S\} = 0$.. The relation

$$\text{grad}\ (J) + c(\text{grad}\ (\lambda_{cr})) = 0 \tag{27}$$

is related to the Kuhn-Tucker condition for a local minimum. Geometrically, equation (27) expresses the fact that the gradient of J is expressible as a linear function of the gradient of the constraint, which is normal to the boundary λ_{cr} = constant. Thus, there is no possible way to reduce the objective function without violating the constraint (Ref. 8, pp. 167-171). This Kuhn-Tucker condition is helpful in determining how close one is to an optimum design.

3.0 Computational Approach

The previous discussion and the number of references cited shows that there is a wealth of information on all aspects of the problem. The aeroelastic analysis techniques with finite element methods, the finite element analysis itself and the optimization logic are all readily available. The combination of all these techniques into a computational design algorithm is the next step in the analysis.

The techniques necessary to analyze the panel flutter problem, to form the direction vector and to calculate the step size ϵ were programmed for the computer. The program has these basic features

(a) λ_{cr} is determined from the analysis of a uniform thickness panel with a number of elements, n, equal to that of the panel being optimized.

(b) The variables $g_i = \partial\lambda_{cr}/\partial t_i$ are calculated for the design.

(c) One of the three methods outlined in the previous section is used to find a step direction.

(d) A step size ϵ is calculated such that the objective function is reduced by a given percentage.

(a) The design variables are modified by an amount Δt_i.

(f) λ_{cr} for the new design is calculated.

(g) If λ_{cr} falls outside the design constraint limits, it is modified by a gradient method. If the desired change in λ_{cr} is

$$\Delta\lambda_{cr} = \sum_{i=1}^{n+1} g_i \, t_i \tag{28a}$$

and

$$\Delta t_i = \mu g_i \tag{28b}$$

then

$$\mu = \Delta\lambda_{cr} / \sum_{i=1}^{n+1} (g_i)^2 \tag{28c}$$

(h) If the design constraint is satisfied then the operation begins with step (b) and a new set of design variables.

Steps (f) and (g) are necessary because the functional relationship between and the design variables Δt_i is nonlinear. The design modifications Δt_i represent linear steps along a nonlinear constraint boundary. If these steps are too large, the design variables will result in a value λ_{cr} which grossly violates the design specifications. Therefore, the decrease in the objective function cannot be too rapid.

3.1 Results

The initial study of a uniform thickness panel and other symmetric panels showed that the elements g_i are symmetric about the midchord, $x = (1/2)$. That is

$$g_i(x) = g_i(1 - x) \tag{29}$$

Since the uniform thickness panel is symmetric and is the initial design, the optimization procedures outlined can never yield a nonsymmetric design. This fact agrees with Turner's comments in Ref. 2. This symmetry property can be used to reduce the amount of computation for an n element panel. If n is an even number, only $1 + n/2$ values of g_i are necessary.

The value of g_i at the leading edge is, for the reference panel, much smaller than other values g_i at different panel locations. This relative insensitivity of λ_{cr} to design variable changes at the leading and trailing edges can be a source of error. If the design variable is not perturbed sufficiently, the value of g_i may be subject to numerical roundoff errors. An examination of the optimization methods outlined shows qualitatively that the design variables associated with the smallest values g_i will

decrease in magnitude while those associated with the largest values g_i will increase in magnitude.

The majority of the investigation was done with four and six element models. In all cases the end thicknesses tended to decrease from one design step to another. The Fletcher-Reeves method and the feasible direction method decreased the end thicknesses much faster than the gradient method suggested by Rubin. Because of the possibility that inaccuracy in determining the values of g_i at the ends might lead to inaccuracy in determining the optimum design, the author wanted to eliminate these variables from the design process as quickly as possible. For this reason, Rubin's gradient method was not used except in the early stages of the investigation. Also, because the value of the design variables at the ends approached zero, as predicted by the analytic analysis, a minimum thickness constraint is necessary to guard against meaningless results. This constraint is expressed mathematically as

$$t_i \geq t_{min} \tag{30}$$

in this investigation, t_{min} was taken as 0.10.

Using a six-element model and identical starting conditions, the Fletcher-Reeves method and the feasible directions method were compared. Figure 3 shows the results of the first few, completely automated, design cycles. The values of mass ratio and λ_{cr} are plotted versus design cycle. Since δ_1 is equal to unity, the objective function and the mass ratio are identical. When λ_{cr} falls outside the allowable range, the design is modified. This results in a slight weight increase. From Fig. 3, it is seen that both methods are of equivalent accuracy. A 1.5% improvement or decrease in the objective function is requested for each design cycle.

The formulation and logic used in the feasible directions technique has a slight advantage over the Fletcher-Reeves method in that the gradient of λ_{cr} not only provides a measure of how the design process in progressing, but it provides qualitative information on the next design step. This information is useful when interaction with the computer is desired. The designer can intervene with the design process at any step and modify the design logic.

After the objective function had been reduced by about 10%, the author used the feasible direction method to achieve better designs. The amount of objective function improvement requested at this point was of the order of 0.25 to 0.50%. Larger requests for improvement were found to be inefficient because of unacceptable λ_{cr} fluctuations.

Typically, on all models analyzed, the design variables fluctuated as the optimum value of J was approached. This could occur if the value of the objective function is not sensitive to the constraint near the optimum while the design variables themselves are rather sensitive. This sensitivity was more pronounced with larger numbers of elements.

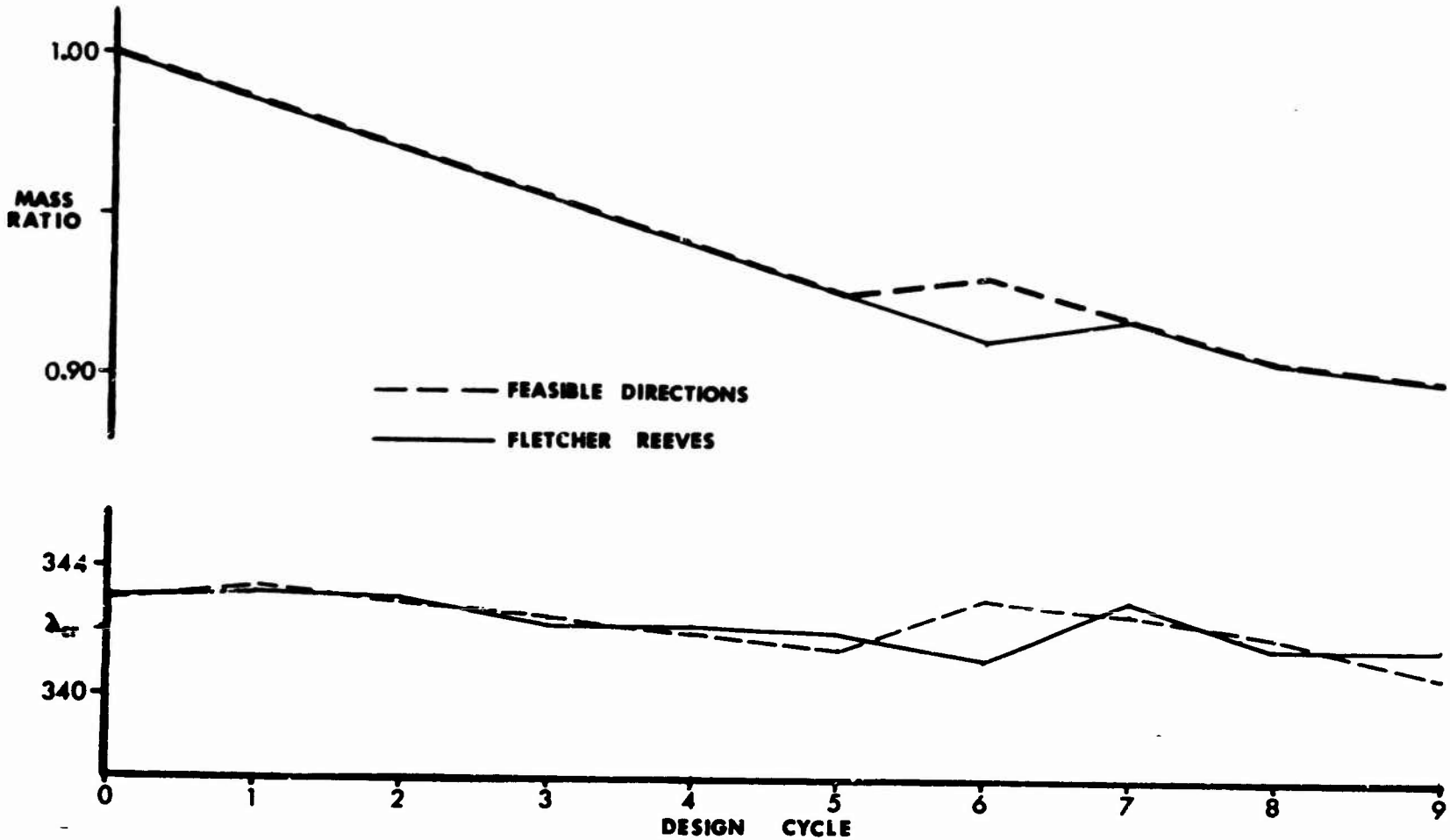


FIG 3 PROGRESSIVE COMPARISON OF TWO SIX ELEMENT OPTIMIZATION METHODS ($\delta_1 = 1.0$)

Figure 4 shows how a typical design progresses from a uniform panel to a lesser weight panel. Figure 5 shows the final designs found with 4, 6 and 10 elements. Some further improvement in these designs is possible but, with the techniques used, was thought to be uneconomical. Table 1 below shows, for a six-element model, the initial values of g_i and the final values of g_i compared to the gradient of J.

TABLE 1

A Comparison of g_i Values Before and After Panel Optimization
6-element Model, $\delta_1 = 1$, $t_{min} = 0.10$

n x grad J ($n \times w_i$)	g_i (Reference panel)	g_i (Final Design)	Design Variable
0.5	5.62	29.8	1
1.0	56.9	69.8	2
1.0	85.0	62.8	3
1.0	63.1	65.9	4
1.0	85.0	62.4	5
1.0	56.9	69.8	6
0.5	5.62	29.8	7

A measure of the accuracy of this method is provided by the comparison in Fig. 6. The converged solution for a panel obtained from Ref. 3 is compared to the best finite element solution obtained for a panel with identical parameters. The mass ratio for the converged solution is MR = 0.885 while the mass ratio for the finite element solution is slightly greater than 0.910. This is a difference of 2.5%.

Some difficulty was experienced when working with large numbers of elements. More difficulty was experienced with the calculation of the g_i elements when using ten elements than with four elements. This can be caused by a number of numerical difficulties. The important aspect of this problem is that it does exist and probably prevents the generation of more accurate results. This g_i calculation is an area where improvement possibly can be made. Rubin (Ref. 7) used analytical expressions for frequency gradients in his work. The use of a similar technique for the panel flutter problem to calculate g_i is hampered by the fact that flutter occurs when two frequencies merge. This factor considerably complicates the mathematical aspects of the problem.

4.0 Conclusions

The results of this study and the comparison with a converged numerical result shows that finite element techniques provide accurate models for the

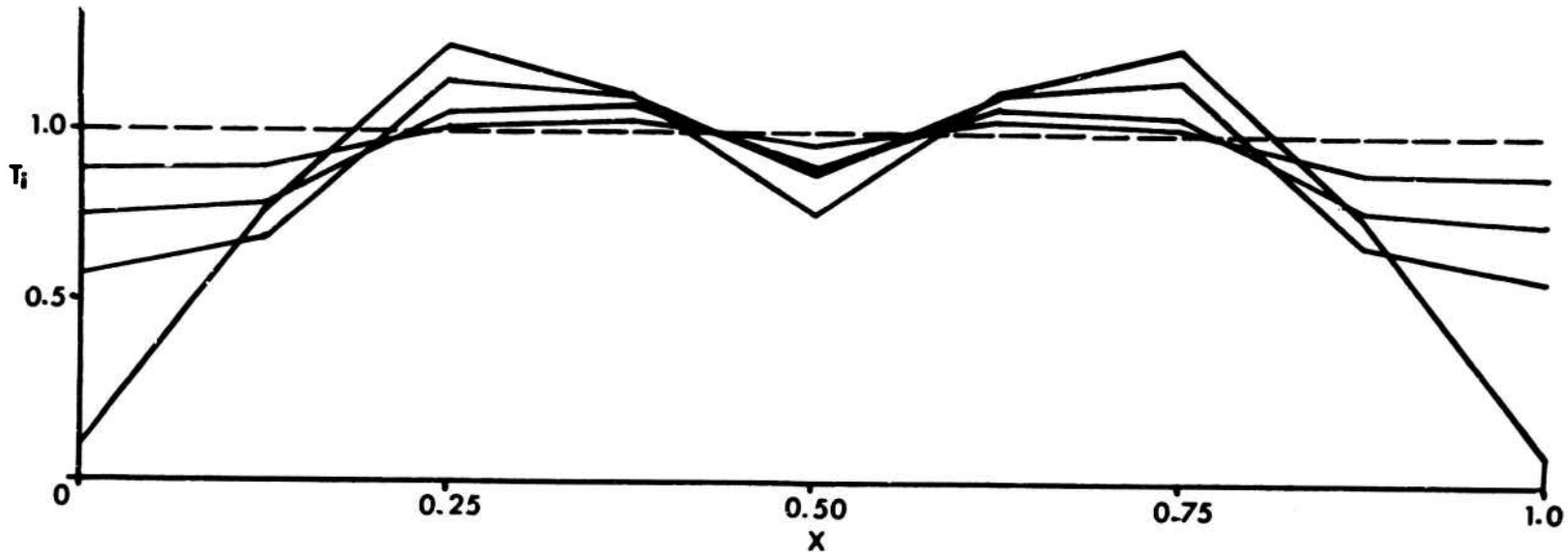


FIG 4 SEVERAL SUCCESSIVE EIGHT ELEMENT DESIGNS ($\delta_1 = 1.0$)

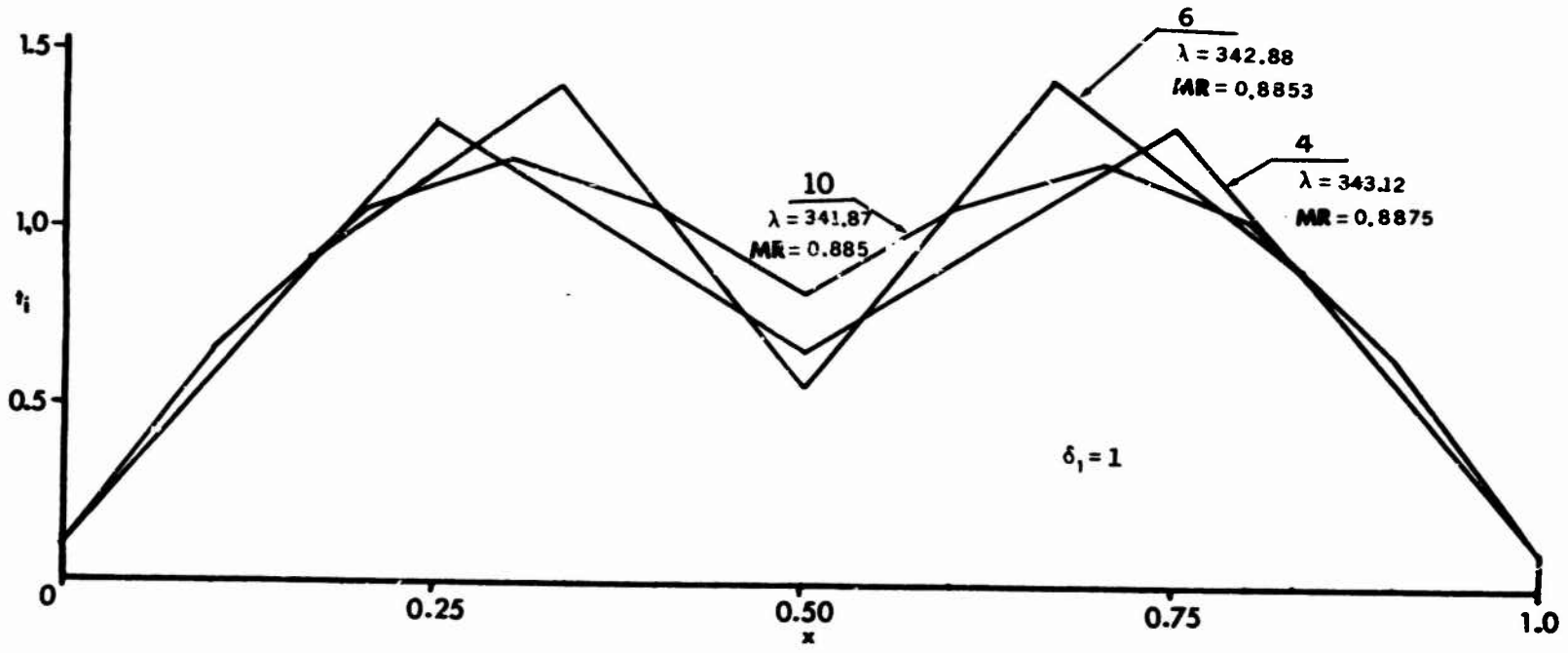


FIG 5 COMPARISON OF FINAL DESIGNS WITH FOUR, SIX AND TEN ELEMENT MODELS

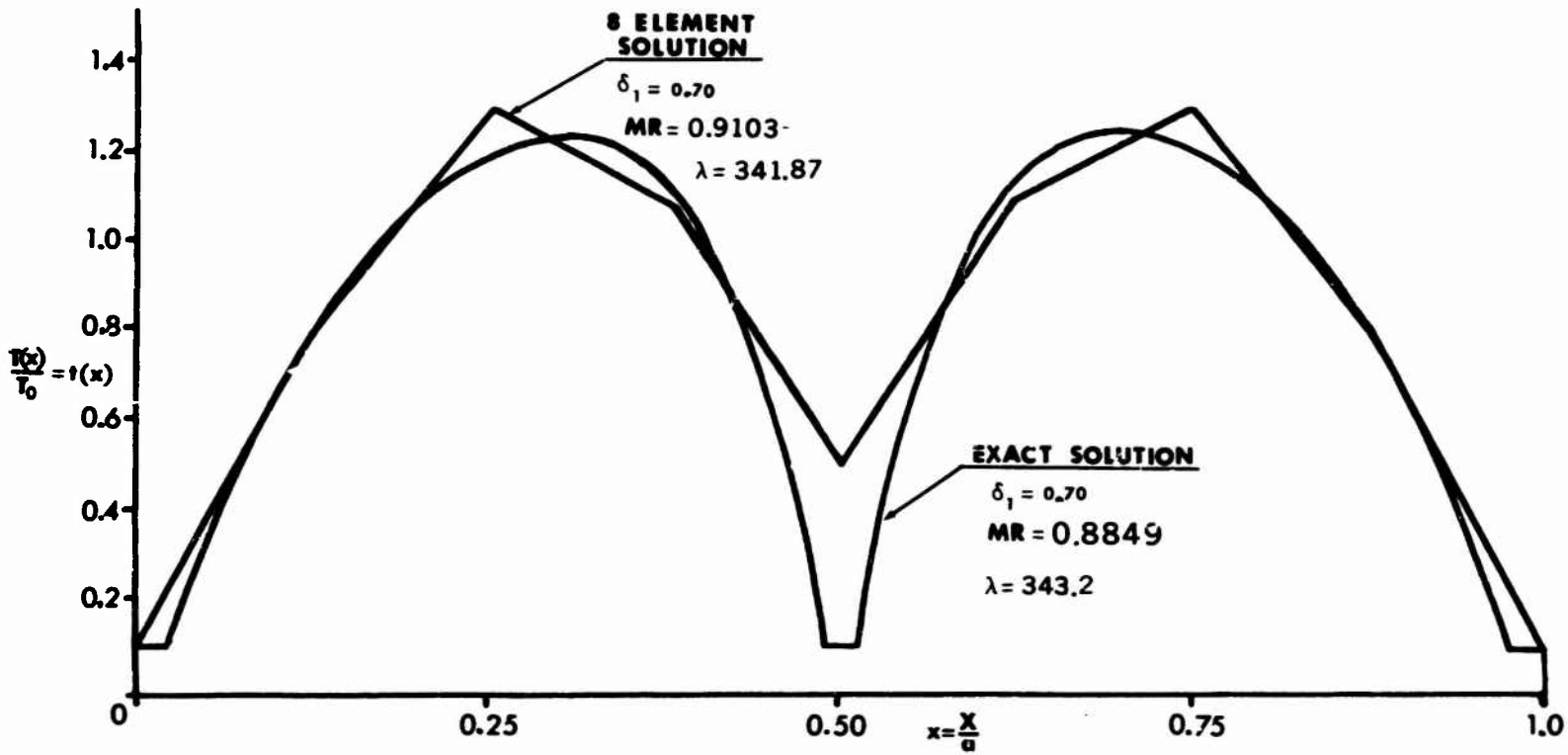


FIG 6 COMPARISON BETWEEN EXACT THICKNESS DISTRIBUTION AND EIGHT FINITE ELEMENT OPTIMIZATION

flutter optimization problem. Weight savings up to 12% have been shown. The ultimate accuracy of the method is dependent on the ability of the finite element model to adequately describe the actual optimal design. A recent flutter optimization study by Craig (Ref. 11) used a series of uniform thickness elements to model the panel. This approach is similar to Turner's except that Craig used a gradient projection technique in the optimization portion of the problem. Craig shows weight savings of around 4%. The present study uses an optimization technique similar to Craig's but, the use of the tapered element results in weight savings of more than twice that shown by Craig for a comparable problem. Thus, the main difference appears to be in the choice of the finite element model.

The great advantage of the finite element approach over the differential equation approach is that very little insight is necessary to generate good solutions. Convergence difficulties do not occur with finite element panel flutter optimization until near the optimal design. Some theoretical aspects of the problem are sacrificed with the finite element approach, but this is far overshadowed by the fact that a reasonable solution can be found. In addition, the model of the panel may be easily made more sophisticated and realistic by the addition of such effects as shear stiffness or the inclusion of different boundary conditions. This can be done without significant alterations in the computer program logic.

Additional improvements in optimization logic and the calculation of constraint gradients can undoubtedly be made. This should improve the solution and reduce the number of design cycles required. More importantly, this study shows that, given a suitable element, finite element techniques can be competitive with other, more complicated optimal structural design techniques.

Acknowledgements

The author wishes to acknowledge and thank the following group of people: Professor Holt Ashley, Stanford University, whose advice and encouragement suggested the investigation; Dr. Roy Craig, Jr., University of Texas, Austin, whose programming advice during the early stages of the study proved valuable; The Minta Martin Committee and Computer Science Facility at the University of Maryland for financial support; and Mr. Michael Hirtle, who helped the author with the computation and drew the figures.

References

1. H. Ashley and S. C. McIntosh, Jr., "Application of Aeroelastic Constraints in Structural Optimization," Proceedings of the 12th International Congress of Applied Mechanics, Springer, Berlin, 1969.

2. M. J. Turner, "Optimization of Structures to Satisfy Flutter Requirements," Volume of Technical Papers on Structural Dynamics, AIAA Structural Dynamics and Aeroelasticity Specialist Conference and ASME/AIAA 10th Structures, Structural Dynamics, and Materials Conference, AIAA, New Orleans, La., April 1969, pp. 1-8.

3. T. A. Weisshaar, An Application of Control Theory Methods to Optimization of Structures Having Dynamic or Aeroelastic Constraints, SUDAAR No. 412, Dept. of Aeronautics and Astronautics, Stanford University, October 1970.

4. R. L. Bisplinghoff and H. Ashley, Principles of Aeroelasticity, John Wiley & Sons, Inc., New York, 1962.

5. M. D. Olson, "Finite Elements Applied to Panel Flutter," AIAA Journal, Vol. 5, 1967, pp. 2267-2270.

6. A. E. Bryson and Y.-C. Ho, Applied Optimal Control, Blaisdell, Waltham, Mass., 1969.

7. C. P. Rubin, "Dynamic Optimization of Complex Structures," Volume of Technical Papers on Structural Dynamics, AIAA Structural Dynamics and Aeroelasticity Specialist Conference and ASME/AIAA 10th Structures, Structural Dynamics, and Materials Conference, AIAA, New Orleans, La., April 1969, pp. 9-14.

8. R. L. Fox, Optimization Methods for Engineering Design, Addison-Wesley, Reading, Mass., 1971.

9. G. Zoutendijk, Methods of Feasible Directions, Elsevier, Amsterdam, 1960.

10. S. C. McIntosh and L. Gwin, Unpublished Notes, Department of Aeronautics and Astronautics, April 1971.

11. R. R. Craig, Jr., "Optimization of a Supersonic Panel Subject to a Flutter Constraint - A Finite Element Solution," AIAA Paper No. 71-330, AIAA/ASME 12th Structures, Structural Dynamics and Materials Conference, Anaheim, California, April 19-21, 1971.

THE USE OF OPTIMALITY CRITERIA IN AUTOMATED STRUCTURAL DESIGN

Ronald A. Gellatly*
Bell Aerospace Company
Division of Textron, Inc.

&

Laszlo Berke**, Warren Gibson***
Air Force Flight Dynamics Laboratory

Research in the field of structural optimization in the past decade has led to the development of a number of methods for the automated design of structures of least weight subjected to a multiplicity of loads. These methods which have been principally based upon the use of various algorithmic forms of numerical search have achieved a considerable measure of success for smaller scale problems. Such methods do suffer from the major disadvantage that computational costs increase rapidly with problem size, tending to impose an economic limit on the complexity of the structures which can be optimized. In order to circumvent these economic limitations, the entire problem of structural optimization has been reviewed. A novel approach to the least weight design of indeterminate structures under multiple loading conditions with strength, displacement and fabricational constraints has been developed which overcomes many of the shortcomings of direct numerical search methods. Details of the new approach, which has dramatically reduced the number of cycles compared with mathematical programming formulations are presented along with examples of applications to the weight optimization of representative structures. The incorporation of other types of constraint conditions is also discussed.

I. INTRODUCTION

The last decade has been a period of triumph and tragedy for the technology of structural optimization. At the start of the 60's the initial concepts of performing automatic optimization of complex structures were being defined and translated into working programs. At the heart of this development were the finite element methods of structural analysis and the rapidly expanding electronic computer capabilities. Finite element methods of analysis had been demonstrated to be reliable, accurate, relatively rapid, general in application to all classes of structures and above all, largely automatic in use. Only the minimal quantity of basic data was needed to specify the problem and the analysis methods could generate without further assistance from the engineer all manner of information concerning stresses, displacements and other response phenomena. Although, in some respects, still in their infancy at that time, finite element methods clearly offered major advantages over other analysis procedures while possessing great development potential. Electronic computers were also in a similar state.

---

* Chief Engineer, Structural Systems Department
** Tech Manager, Structural Synthesis Group
*** Aerospace Engineer

Their basic potentialities had been eminently proven by the socalled first generation mechanics and the second generation computers were being widely used. Third and later generation machines were being promised by the manufacturers which would reduce computation costs and calculation times by orders of magnitudes while providing infinitely large rapid access storages. In this aura of general optimism, it was confidently assumed that the significant computational costs associated with analysis of early finite element problems involving less than a hundred degrees of freedom would be reduced, possibly to inconsequential levels for the majority of analysis problems. Unfortunately, the ambition of the engineer/analyst also increased with passage of time and he was no longer satisfied with the numerical results generated by the small numbers of degrees of freedom. Larger and more complex, problems were undertaken involving the handling of correspondingly larger matrix arrays.

At the same time some, but not all, of the anticipated advances in computer technology have taken place. Computers are large and fast but the net effect to this date has been that the requirements of analytical complexity have consistently exceeded the advances in computational capabilities made available by the successive generations of computers. The analyst has been forced to solve increasingly complex problems with the result that analysis times and costs have tended to increase rather than decrease. It is true that more detailed information has been generated for the more complex models but this has always been at the behest of the design engineer. Although the net result of this has been the opposite of what was generally foreseen, the view in the early 60's was that analysis cost could eventually become negligible.

The development of operations research methods arose through the requirements for determining optimal processes in a wide variety of industrial applications. Through application of various forms of mathematical programming techniques the determination of optimal values of complex functions of several variables was accomplished. From a prime application in process control, the use of mathematical programming quickly spread to a wide variety of other disciplinary activities.

Through use of the one or more of the many standard and nonstandard search algorithms, it is possible to determine the maximum or minimum value of some specified merit function subject to certain constraints, in an iterative manner. These methods of directed search are, in general, far superior to random trial and error procedures and guarantee some form of convergence, albeit to local minima. After each iteration, some form of evaluation of the merit function and constraints is necessary to determine the direction of the next stage of iteration.

In light of the apparent situation with regard to anticipated analysis costs in the late 50's, it was a logical and brilliant step to couple mathematical programming and finite element analysis together to generate an automated method for structural optimization (Reference 1). Mathematical programming methods of optimization required repeated analysis - and finite element methods provided a rapid automatic low cost procedure for such analyses.

Initially, this work met with a great degree of success, particularly with small scale problems. At last it was possible to come to grips with the real problems of design-stiffness as well as strength constraint conditions and fabricational considerations for structures subjected to a multiplicity of loading cases.

Flushed with the successful demonstration of the feasibility of these coupled procedures for constrained weight minimization, development proceeded apace. It appeared that their extension to larger order problems involving up to 100 or more degrees of freedom and/or variables could readily be accomplished. The field of

endeavor expanded, large scale problems with practical operational capabilities were developed (References 2 and 3), new formulations (Reference 4) of the structural problem investigated, new algorithms were used (Reference 5), more sophisticated types of problems were tackled (References 6 and 7), and the field seemed limitless. It was recognized with larger and more complex problems that the computational costs would increase. Larger analyses were being performed and some increase in the number of iterations to convergence was only to be expected. It was tacitly assumed by many that the anticipated improvements in computers and advances in analysis methods would keep, the costs of the analyses to economic levels. In addition, the actual increase in the number of iterative analyses would not be excessive.

With further development, a less satisfactory situation began to emerge. As indicated previously, the requirements for more sophisticated analyses had actually tended to increase rather than decrease computational costs. In addition, the number of iterations required and the number of analyses per iterative stage was found to increase more rapidly with problem size than had been first assumed, principally due to the explicit or implicity need to determine the derivatives of the constraint functions with respect to the design variables. Some improvements in this situation appeared with the reformulation of the basic constrained problem as an unconstrained one through the use of penalty functions of various types (References 8 and 9). A greater generality of merit and constraint functions was possible than with the earliest constrained formulations and the use of non-derivative search methods eliminated the need for so many function evaluations - but the gain was more apparent than real. Again, as in the case of the original constrained formulation development, a considerable measure of success was achieved initially. Certainly for the small test problems, the procedure was more efficient and the greater generality of merit function definition was an advantage. On the other hand, the actual improvements were relatively small, and the same difficulty of rapid rise in number of constraint evaluations with problem size was still present*. Although the explicit need for derivatives was eliminated, the implicit requirement had the same effect in the end.

In the 1960's, the mathematical programming approach almost completely dominated the structural optimization scene. The potential payoff in structural optimization aroused considerable general interest resulting in a considerable level of effort being expended by many researchers who approached the problem from a number of slightly different points of view with varying degrees of success. By the use of new search algorithms and new penalty functions or by various schemes for linearization of the problem attempts were made to overcome the critical rise in numbers of analyses with problem size. References 10 and 11 provide effective summaries of the many methods used with indications of their capabilities. Reference 10 documents the variations in approaches to optimization through the use of nonlinear programming methods whereas the papers of Reference 11 tend to highlight some of the barriers encountered in the practical use of nonlinear programming approaches to structural optimization.

In the most general terms, it was beginning to appear that there were remarkably low limits on the size of structure which could be handled economically by optimization programs. Even the most optimistic estimate was only about 150 variables or less. Thus, at the end of the 60's, a general need for structural design and optimization capabilities had been recognized but there appeared little immediate prospect for the development of more efficient nonlinear programming algorithms to overcome the economic barriers to widespread operational usage on real structures. The time was now ripe for a new approach to the problem, if structural optimization was to survive as an economic design potential rather than an interesting research toy.

---

* In Reference 10, it has been estimated that the number of function evaluations may go as high as $7N^3$, where the N is the number of variables.

II. NEW APPROACH TO OPTIMIZATION

A. Fully Stressed Design Methods

Since the avenue of nonlinear programming for structural optimization did not appear to offer the hope of large scale design applications, it was appropriate to reconsider some of the basic concepts which are used extensively in structural design.

Numerical search methods determine optimal systems in a purely empirical manner. That is, a set of rules is established which will guarantee a continuous and monotonic decrease in a prescribed merit function, without regard to the nature of that function. No preconditions concerning the optimum are specified apart from the criteria that is impossible or uneconomic to determine a further design which will be an improvement on the present design. Both the strength and weakness of mathematical programming reside in this formulation. The strength is the generality which this independence of problem type imparts; the weakness is that no use is made of any characteristics of the problem which would permit a more efficient solution.

For the structural problem, perhaps the generality and rigor are unnecessary. What is really required for practical design is a simple method (i.e. limited number of function evaluations) which may depend to some extent upon the specialized structural nature of the problem and which will operate with some measure of optimality. That is, a procedure which will rapidly produce designs better than a standard engineering approach but without rigorous guarantees of convergence on the optimum.

In structural design, the classic example of this type of approach is the concept of simultaneous failure. In this, it is assumed, intuitively, that the best design is one for which every possible mode of failure will occur at the same time. Thus, if there are an equal number of possible modes of failure (constraint conditions) and design variables, the result will be a set of simultaneous equations (possibly nonlinear) whose solution will yield the best design. If the numbers of constraints and variables are unequal, then a certain degree of judgement and even trial and error may be necessary. This corresponds to the determination of the vertices of constraint surfaces in the nonlinear programming with the basic difference that the desired vertex is found, not by a direct search, but by direct solution of the equations generated by the conditions assumed to obtain at the optimum. This method has its obvious limitations, but its simpler form of fully stressed design is widely known and used (Reference 12). This time-honored procedure is based upon the intuitively satisfying, although analytically unjustified assumption that in an optimal structure each member will be fully stressed under at least one of the loading conditions.

The basis for the selection of this type of criterion is obvious. Engineering reason clearly indicates that the best structure must be one in which each member is working to its utmost limit. For the case of a statically determinate structure under a single or multiple loading system, it has indeed been demonstrated that the fully stressed design has least weight (Reference 13). It is also fairly certain that this fact was demonstrated well after the concept was used in engineering design. On this basis, the extension of the fully stressed concept to multiple loaded indeterminate structures is a relatively simple step (Reference 14). Unfortunately, this also has proved to be an unjustified assumption (Reference 1). The paradox which exists about fully stressed design is that it may or may not be of minimum weight. True, there is no explicit reference to weight as a merit condition in the usual stress ratio method of determining a fully stressed design, but this alone does not necessarily negate the validity of the assumption. In many cases, a fully stressed design can be determined whose optimality with respect to weight can be verified by application of the empirical criteria used in the numerical search methods (Reference 15). In some cases, a

fully stressed design can equally be shown to be heavier than the least weight design (Reference 16) and in still other cases no fully stressed design can be found. That no apparent rules exist to differentiate between these cases, except through a post hoc examination of the fully stressed design has been demonstrated analytically (Reference 17).

In general, operational experience with fully stressed design methods does indicate in the vast majority of problems, not selected for their pathological behavior, that the resultant design - for stress limits only - is indeed either the optimum or close to it. It is also relatively easy to select problems for which the fully stressed design is remote from the optimum, but this is usually achieved by selecting unrealistic and disparate allowable stresses for members which provide parallel load paths in a redundant structure.

Using a simple stress-ratio redesign method, convergence on the fully stressed design will occur in one step for a statically determinate structure and will require a finite number of iterations for an indeterminate structure. The number of iterations required for convergence is another paradoxical problem being apparently controlled by characteristics which are presently poorly understood. One factor which is relatively clear, is that the number of iterations is not, in general, linked closely to the size and number of variables in the problem. It is this last fact which makes the use of fully stressed design so attractive. Each redesign step is simple and in many cases, the convergence is extremely rapid. The fact that a true least weight structure is not generated becomes of lesser importance when compared with the ease with which the improved design is reached (Reference 12).

Clearly, the basis for this type of optimization which is only applicable to strength considerations is the a priori specification of a set of conditions to be satisfied by the optimal design. That these criteria may not be rigorously applicable at the optimum is of little consequence. An a priori approach is in direct contract to a search algorithm in mathematical programming where the determination of an optimum is predicated on a strictly post hoc basis, e.g. when further improvement in merit is impossible, only then is the search terminated.

From consideration of the known characteristics of the fully stressed design procedure, it might be concluded that the rapid convergence and the relative independence of the number of iturations from problem size was largely attributable to the a priori specification of optimality criteria. The stress ratio method attempts to satisfy these arbitrary criteria and hence converges directly on the prescribed point in the design space without need for exploration of convexity or other use of the constreints.

With this recognition, a new avenue of approach to structural optimization becomes possible, through the use of optimality criteria. In use, this requires the definition of additional criteria for each response phenomena coupled with some procedure which will permit the simultaneous consideration of multiple criteria.

B. Displacement Limited Designs

In developing structural optimization programs, a heirarchy of priorities for the inclusion of different types of constraint conditions has become established. The first consideration has been strength followed closely by fabricational limitations (minimum material sizes) and stiffness constraints (displacement limits). Subsequently, buckling, both overall and local, frequency response and static and dynamic aeroelasticity conditions are considered.

In fully stressed design, the basis for the stress-ratio redesign method is the assumption that at the optimal point, an inequality stress constraint for an element becomes satisfied as an equality condition. The same basic logic may be used for a displacement constraint condition.

The first consideration of this approached was presented by Barnett and Hermann (Reference 18) for the optimal design of determinate trusses in which a value was specified for the displacement of one node point. Berke (Reference 19) re-evaluated this work and proposed its application to an indeterminate structure with a generalized constraint. From this starting point, the optimality criteria have been further generalized for indeterminate structures under multiple loading conditions with a multiplicity of displacement constraints (Reference 20). In deriving these criteria, provision is made for the simultaneous consideration of other redesign algorithms associated with additional criteria appropriate to the other constraint conditions.

The starting point for the development of this approach to minimum weight design is the consideration of a structure under the action of a single loading system.

The structure is represented by an assemblage of discrete elements for which, initially, one stress is sufficient to describe the response behavior of each element. For more complex elements involving multiple stress components, the same overall derivation will apply, but modified slightly to reflect the additional stress components. This will be demonstrated later. Single stress elements may be axial force members or shear panels. For an axial member, the appropriate design variable is selected to be its cross-sectional area, whereas for a plate or panel thickness is the chosen variable.

Application of external loading results in internal forces S defined as the produce of stress and cross-sectional area for axial members and stress and thickness for plates. The internal forces in each member S_i^P arise as the result of the application of an external loading system P which may consist of a single or many load components. Associated with the loading system will be displacements δ_i at all nodes of structure. The value of a genera ed displacement, which as a special case is a single nodal displacement, can be determined through use of a generalized virtual load method.

A virtual loading system Q, corresponding to the generalized displacement system in node point and direction, is applied to the structure. The virtual work Δ arising between the virtual system Q and the generalized displacement δ^* is given by

$$\Delta = \sum_{i=1} \frac{S_i^P \; S_i^Q \; L_i}{A_i \; E_i} \tag{1}$$

where S_i^P, S_i^Q are the forces due to the actual and virtual systems respectively

A_i is the (variable) cross-sectional dimension

L_i is the geometric dimension of the element and

E_i is the appropriate elastic modulus

L_i , the geometric dimension is defined by

$$L_i = V_i/A_i \tag{2}$$

where V_i is the volume of the element.

For plate elements L_i is the superficial area and for shear panels the appropriate value for E_i is the shear modulus. For the special case of a single displacement, the corresponding virtual loading system Q consists of a single unit load and hence Δ as calculated by Equation (1) is the value of displacement.

It is assumed now that the variables in the structure may be divided into two groups – active members and passive. Active members are those whose cross-sectional dimension A_i may be varied to achieve an optimized displacement limited design, whereas passive members will remain unchanged. The basis for the allocation of the individual member into the two grouping will be discussed later, and has no influence on the derivation.

With this division Equation (1) can be rewritten as

$$\Delta = \sum_{i=1}^{m} \frac{S_i^Q \, S_i^P \, L_i}{A_i \, E_i} + \sum_{i=m+1} \frac{S_i^P \, S_i^Q \, L_i}{A_i \, E_i} \tag{3}$$

where the first group contains m passive members.

Since in subsequent manipulations, the passive members are unaltered, the first term of Equation (3) is replaced by Δ_o

i.e $$\Delta = \Delta_o + \sum_{i=m+1} \frac{S_i^P \, S_i^Q \, L_i}{A_i \, E_i} \tag{4}$$

The weight of the structure W is given by

$$W = \sum_{i=1} V_i \, \rho_i = \sum_{i=1} A_i \, L_i \, \rho_i \tag{5}$$

where ρ_i is material density.

Introducing the active and passive member grouping Equation (5) becomes

$$W = W_o + \sum_{i=m+1} A_i \, L_i \, \rho_i \tag{6}$$

where W_o is the weight of the passive members.

A constraint on the allowable values of the virtual work of the load system Q can be prescribed as Δ^*, given by

$$\Delta^* = \Sigma \, Q_i \cdot S_i^* \tag{7}$$

where the summation is taken over all terms of the generalized system.

The minimum weight structure for which the generalized displacement will have the specified value Δ^* can now be determined by finding the stationary value of W subject to the equality constraint.

$$\Delta = \Delta^* \tag{8}$$

For optimization problems involving equality constraints, a Lagrange multiplier approach is ideally suited. Using a Lagrange multiplier λ, the problem can be expressed as the minimization of the function

$$F = (W_o + \sum_{i=m+1} A_i L_i \rho_i) + \lambda \left(\sum_{i=m+1} \frac{S_i^P S_i^Q L_i}{A_i E_i} + \Delta_o - \Delta^* \right) \tag{9}$$

For a minimum

$$\frac{\partial F}{\partial A_j} = 0 = L_j \rho_j - \lambda \frac{S_j^P S_j^Q L_j}{A_j^2 E_j} + \lambda \sum_{i=1} \left(\frac{\partial S_i^P}{\partial A_j} S_i^Q + \frac{\partial S_i^Q}{\partial A_j} S_i^P \right) \frac{L_i}{A_i E_i} \tag{10}$$

or

$$0 = L_j \rho_j - \lambda \frac{S_j^P S_j^Q L_j}{A_j^2 E_j} + \lambda \sum_{i=1} C_{ij} \tag{11}$$

It is to be noted that the summation involved in the last term of Equation (11) must be taken over all members active and passive. This term represents the redistribution of the internal forces associated with change in member size. For a statically determinate structure, all Cij are zero. For redundant structures, the terms $\partial S_i / \partial A_j$ will generally be small and will tend to diminish in size rapidly moving away from the j^{th} member.

From Equation (11)

$$A_j = \sqrt{\lambda} \sqrt{\frac{S_j^P S_j^Q L_j}{E_j (L_j \rho_j + \lambda \sum_{i=1} C_{ij})}} \tag{12}$$

Substituting Equation (12) into Equation (14) and noting Equation (8)

$$\Delta^* = \Delta_o + \frac{1}{\sqrt{2}} \sum_{j=m+1} \sqrt{\frac{S_j^P S_j^Q L_j}{E_j} (L_j \rho_j + \sum_{i=1} C_{ij})} \tag{13}$$

Rewriting Equation (13)

$$\sqrt{\lambda} = \left(\frac{1}{\Delta^* - \Delta_o} \right) \left(\sum_{j=m+1} L_j \sqrt{\frac{S_j^P S_j^Q \rho_j}{E_j} (1 + \frac{\lambda}{\rho_j L_j} \sum_{k=1} C_{jk})} \right) \tag{14}$$

Finally substituting Equation (14) into Equation (12)

$$A_i = \left(\frac{1}{\Delta^* - \Delta_o} \right) \left(\sum_{j=m+1} L_j \sqrt{\frac{S_j^P S_j^Q \rho_j}{E_j} (1 + \frac{\lambda}{\rho_j L_j} \sum_{k=1} C_{jk})} \right) \sqrt{\frac{S_i^P S_i^Q}{E_i \rho_i (1 + \frac{\lambda}{\rho_i L_i} \sum_{k=1} C_{ij})}} \tag{15}$$

Equation (15) now represents the criteria which the variables A_i must satisfy at the optimum for the desired displacement constrained minimum weight design.

Use of these relationships would involve the iterative determination of the constant λ and terms C_{ij} which are the local gradients of the stresses with respect to the variables. Their determination by finite differences or other means would be tantamount to a return to the computational situation which exists for nonlinear programming approaches. Since they are generally small, the error introduced by setting all C_{ij} to zero will be negligible. Equation (15) can now be rewritten as

$$A_j = \left(\frac{1}{\Delta^* - \Delta_o}\right) \sqrt{\frac{S_j^P S_j^Q}{E_j \rho_j}} \sum_{i=m+1} \sqrt{\frac{S_i^P S_i^Q \rho_i}{E_i}} \qquad (16)$$

In finite element analyses, the more usual form of output from a computer is stress rather than element force. Introducing elemental stress Equation (16) becomes

$$A_j = \left(\frac{1}{\Delta^* - \Delta_o}\right) A_j \sqrt{\frac{\sigma_j^P \sigma_j^Q}{E_j \rho_j}} \sum_{i=m+1} (A_i L_i \rho_i) \sqrt{\frac{\sigma_i^P \sigma_i^Q}{E_i \rho_i}} \qquad (17)$$

For a statically determinate structure, there is no redistribution of internal forces with variation in member size and all C_{ij} are identically zero. Hence, Equation (17) is exact, and will generate the active member sizes directly for a minimum weight structure with its critical displacement equal to the specified value Δ^* and ignoring any restraints on element sizes or stresses. If any members have been selected as passive, their fixed values will be reflected in the computation of the term Δ_o. This procedure is then analogous to a constrained minimization with one main constraint (the displacement limit) and a number of side constraints.

It has been long known that for a statically determinate structure subject to strength (stress) limitations only, the minimum weight design has all members either fully stressed or at their minimum allowable size. If such a member satisfies the inequality displacement limitation, no redesign is possible or required. If the displacement constraint is violated, certain (active) members must be increased in size. With reference to Equation (3), if the product $(S_i^P \cdot S_i^Q)$ is negative, then an increase in the member size (in the denominator) will decrease the negative component of Δ arising from this member and hence further increase the total value or Δ. On the other hand, reduction of the area is not possible since it will already have its minimum size to satisfy the stress or fabricational requirements. Thus, in effect, no variation in such a member size can be permitted, and this provides one criterion for the selection of a passive member in a statically determinate structure i.e. $(S_i^P S_i^Q)$ is negative. As a logical corollary to this definition, it also follows that any member for which an external condition would be governing, i.e. requiring a larger size than that demanded by the displacement constraint condition, must also be treated as a passive member.

For statically indeterminate structures, as indicated previously the relationship Equation (17) is only approximate due to the neglect of the C_{ij} terms.* As such, Equation (17) can be used as a recursion expression for the iterative determination of A_j, exactly analogous to the use of the simple stress-ratio expression for determining a fully stressed design. Equation (17) can then be written

$$A_j^{\nu+1} = \left[\left(\frac{A_j}{\Delta^* - \Delta_o}\right) \sqrt{\frac{\sigma_j^P \sigma_j^Q}{E_j \rho_j}} \sum_{i=m+1} (A_i L_i \rho_i) \sqrt{\frac{\sigma_i^P \sigma_i^Q}{E_i \rho_i}}\right]^{\nu} \qquad (18)$$

where the superscripts ν, $\nu+1$ refer to iterations. From some arbitrary starting point, iterative application of Equation (18) will then lead to convergence in a finite number of steps, on the statically indeterminate structure of minimum weight with its

*It has been recently realized that ΣC_{ij} vanishes identically also for redundant structures.

critical displacement equal to Δ^*. As in the case of the statically determinate structure, certain members may be treated as passive non-participatory in this process based upon external criteria. The rate of convergence on the minimum weight design will be controlled by the magnitude of the omitted terms. If they are truly small and localized in effect, the convergence will be rapid. This is usually the case in large complex structures, but paradoxically in small structures the reverse frequently occurs.

Before a practical use can be made of this recursion expression, for the design of minimum weight structures with displacement constraints, a number of points must be clarified:

First, in using virtual unit load methods, the internal force or stress system corresponding to the virtual external load Q need only be statically equivalent to Q. For a determinate structure only one such system exists, but in an indeterminate structure S_i^Q is not unique and a multiple choice exists for the statically equivalent system. Since the selection of any one particular statically equivalent system means that S_i^Q will be arbitrarily set to zero in some (redundant) members, this is tantamount to an arbitrary selection of active and passive members for redesign. This is also equivalent to introducing side constraints on the design variables which in turn can result only in the increase of the attainable minimum weight. Therefore, when some members are to be excluded from design changes, the correct approach is to use the actual distribution of S_i^Q, which arises from the application of the virtual load Q to the real structure. This permits all members, potentially, to participate in redesign. It also has the virtue of not requiring a separate analysis of a degenerate structure under the virtual load but simply of including Q as an additional loading case in the finite element analysis. Of greater complexity is the division between active and passive members and the associated problems of introducing multiple loads along with displacement and other types of constraints.

The recursion expression, Equation (18) will cause convergence at a vertex in the design space formed by the intersection of the main constraint (displacement condition) and the side constraints (passive member sizes). It is clear that the source and nature of these side constraints is irrelevant to the iteration procedure. This, then, provides the mechanism for the inclusion of additional criteria, whereby other forms of constraint conditions (e.g. multiple displacement constraints, stress limits, fabricational considerations, etc.) can be permitted to specify certain member sizes while other members are sized by the primary constraint. The effect of this type of combination of criteria will be to force the design simultaneously toward all potential individual minima, resulting in convergence in a vertex formed by many active constraints.

It is, therefore, necessary to establish rules or criteria for the selection of active and passive members consistent with the above philosophy.

From the examination of Equation (18) one criterion for the selection of passive member can be observed immediately, similar to the condition used for statically determinate structures. If $(S_i^P \cdot S_i^Q)$ is negative for a member, its inclusion as an active member would require its square root to be found, introducing imaginary numbers which are computationally and physically unacceptable. Thus, one criterion for passive members is

$$(\sigma_i^P \cdot \sigma_i^Q) < 0 \qquad (19)$$

Similarly to the corollary for the statically determinate case, if the size of a member is defined dominantly by other than a displacement constraint, then that member must be treated as passive in the redesign to satisfy displacement constraints. That is, if the size required for a member to satisfy a stress or fabricational constraint is greater than the size required to satisfy a displacement constraint then that member should be considered passive in using Equation (18).

The extension of these principles to multiple loads and displacement constraints can be accomplished in a manner similar to that used in determining the appropriate stress ratio for a fully stressed design. In the stress ratio method, the ratio of the actual to the allowable stress is computed for an element for each loading case and the largest ratio selected for the redesign. With displacement constraints the largest area generated using Equation (18) is taken from the combined set resulting from all loads and displacement constraints. Thus if four loading cases are specified and three displacements are constrained, a total of twelve possible areas for one member are generated by Equation (18) and the largest selected as the dominant value. A second iteration of Equation (18) is then performed in which members designed in the first iteration by a specific load/displacement case are treated as passive except when considering that particular case.

The procedure for redesign with displacement constraints which has evolved from an intensive exploration of the numerical aspects of the problem, automatically partitions members into active and passive categories for a given redesign phase. Tests on the sign of $(\sigma^P_i \cdot \sigma^Q_i)$ are incorporated, as are checks to determine whether stress and fabricational limits dominate for any member and insure that that member is treated es being passive.

An additional feature of note is the method of treating negative displacement limits. Frequently displacement limits are expressed in the form $-\Delta^* < \Delta < +\Delta^*$. For the satisfaction of the negative limit, a negative virtual load system $-Q$ is required. This has the effect of reversing all the signs on the virtual stresses and thereby interchanging the active and passive members selected by the sign of $(\sigma^P_i \cdot \sigma^Q_i)$. The effect of this is to cause some elements to participate in the satisfaction of the positive displacement while the others (passive in the first case) then participate in the redesign for the satisfaction of the negative displacement. The fact that these two constraint conditions are apparently in direct conflict does not introduce, in practice, any complications. Other conditions of stress limits, minimum sizes and multiple loads and displacements intrude and eliminate one of the contradictory constraints.

In summary, the redesign process is as follows: As a first stage, a stress ratio redesign is effected for each member with due regard to multiple loads and minimum member sizes. For every limited displacement specified, a unit load vector is generated and the corresponding stresses σ^Q_i are computed for all members along with the actual stresses σ^P_i. A matrix of the product $(\sigma^P_i \cdot \sigma^Q_i)$ is computed for all combinations of loads and limited displacements (all P_i and Q_i) for each member.

Using Equation (18), new values of A_j are computed for every member. The partitioning into active and passive members at this stage is controlled wholly by the test on the sign of $(\sigma^P_i \cdot \sigma^Q_i)$. In the presence of both negative and positive limits on the restrained displacements, new areas will be generated for each member.

The largest area is selected for each member from all the loading/displacement combinations and a record is made of the particular combination which controlled the design of each area. The areas so generated are compared with those based upon stresses or minimum sizes and the larger values selected for each member.

Again, using Equation (18), a new evaluation of areas is made for each load/displacement case but in this case the passive members for each case consist not only of those elements for which $(\sigma^P_i \cdot \sigma^Q_i)$ is negative, but also those which

were, at the end of the first iteration, critically designed by stress limits, minimum size or by a load/displacement combination other than that being currently considered. At the end of this redesign, a comparison is again made between the stress-designed and the displacement-designed sizes. This cycle is repeated, up to a maximum of three times, until no transfer occurs between the members of the group designed by stresses and those designed by displacements. The resultant design is then reanalyzed and scaled until critical. This process can be repeated as many times as necessary until a minimum weight design is achieved.

The above derivation has been presented for a structure composed of elements for which only one stress was defined. In the case of membrane plates, three components of stress σ_x, σ_y, τ_{xy} are generated for each element. For the calculation of stress limits, a single reference stress is calculated based on the Von Mises criterion

$$\sigma_{ref} = (\sigma_x^2 + \sigma_y^2 - \sigma_x\sigma_y + 3\tau_{xy}^2)^{1/2} \tag{20}$$

A limiting value is then specified for σ_{ref}. For displacement constraints, Equations (17) and (18) are modified slightly to reflect the additional terms. Wherever the term $(\sigma_i^P \sigma_i^Q)$ appears, it is replaced by the term

$$\frac{1}{E_i \rho_i} \frac{1}{\rho_i} \left[\frac{(\sigma_x^P - \mu\sigma_y^P)\sigma_x^Q + (\sigma_y^P - \mu\sigma_x^P)\sigma_y}{E} + \frac{\tau_{xy}^P \tau_{xy}^Q}{G} \right]_i \tag{21}$$

A similar modification is introduced in the calculation of Δ. The test on the sign of $(\sigma_i^P \cdot \sigma_i^Q)$ is replaced by a test on Expression (21). Otherwise the logic is unchanged.

The above redesign procedure was coded and used for a number of very successful optimizations, results of which are presented in the following section.

Among the problems considered was a 25-bar transmission tower used as an example on a number of occasions (References 3, 5, 6, and 21). There were basically two loading conditions on the tower, but in order to maintain symmetry of the structure, 4 additional loads were introduced. Of the 25 members in the structure, the conditions of symmetry partitioned them into 5 groups containing 4 identical members, 2 groups of two members plus a single bar. Application of the program to this problem was only moderately successful and a great deal of operator intervention was necessary to maintain the expected symmetry of design.

The logic of the program was reviewed and it was finally observed that such lack of symmetry was inevitable as the program stood. The difficulty was caused by the presence of the single and the two groups of two members. In the redesign, the single member is equally designed by any of the sets of loads introduced for symmetry. As the logic exists it will be selected as being designed by the first of these cases, and in a subsequent iteration will be allocated to that load case as an active member and be treated as passive by the other three loading cases of the symmetrizing set. A similar effect occurs for the twin members and the result is a nonsymmetric design. To overcome this difficulty, a symmetry option was introduced which constrains members of any specified groups to be identical. This in turn, provides an additional criterion for the selection of active/passive members in the redesign process since only one member of each symmetric group is now independently designed. With this modification, the need for additional symmetrizing loads was eliminated in the tower problem, and excellent rapidly converging results were obtained.

III. RESULTS

In order to demonstrate the practicality of the optimality criteria approach to the design of least weight structures, a computer program was developed based upon a finite element analysis procedura.

At the heart of a finite element program is the element library. At present four simple elements are provided in the optimization program. These ara the axial force member, the triangular membrane plate, the quadrilateral shear panel and the special half web shear panel used for symmetric wing analysis (Figura 1). The triangle is a constant stress element and generates the three components of stress at its centroid, σ_x, σ_y, and τ_{xy}. The use of the half web element is necessary when performing the analysis and design of thin symmetric wing-type structures. In these structures inclusion of all degrees of freedom in both upper and lower halves usually leads to severe conditioning problems associated with the numerical disparities of out-of-plane and in-plane stiffness characteristics. The solution to this difficulty lies in considering symmetric (in-plane) and anti-symmetric (out-of-plane) behavior separately and only treating half the structure. For elements lying completely in either half, this is no problem, but a special half-web element is required to cross the centerline. This trapezoidal shear web is defined by only two node points.

The program uses a simple stress-ratio method to account for strength constraints and the recursion relationship (Equation (18)) for displacement constraints. Fabricational constraints are included and the program performs the partitioning between active and passive members automatically, according to the criteria defined previously. The results presented hera illustrate the characteristics of the iterative process and emphasize some points about the behavior which still remain to be resolved.

A large scale version of the program has been developed. This program is currantly being applied to the optimization of large scale wing structures for practical aircraft in their design stage, where stiffness and strength considerations ara both of critical importance.

A. Four Bar Pyramid (Figure 2)

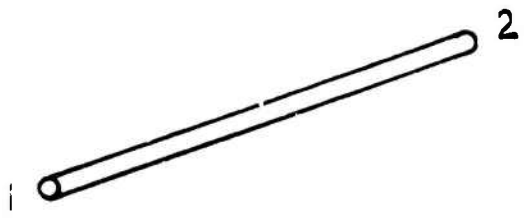
The four bar pyramid has been used extensively (Reference 22) as a test case with a variety of loading cases and constraints.

The structure is subjected to the three non-simultaneous loads at the free node

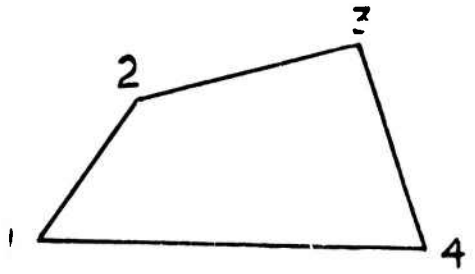
Px = 5,000 lb.
Py = 5,000 lb.
Pz = 7,500 lb.

The constraint conditions are stresses of ±25,000 psi in all members with a minimum araa of 0.1 in^2 and displacement limits of ±0.3 ins. in the y-direction and ±0.4 in. in the z-direction.

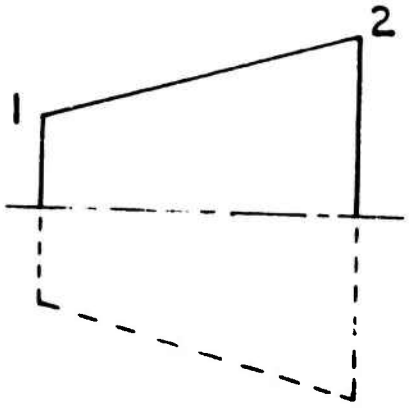
Using a uniform structure as starting point, the minimum weight design of 14.283 lb. was reached in four iterations. Table Ia lists the design history and final design. For information purposes, the program was allowed to run past the fourth iteration for another eleven cycles. Redesign continued and the weight varied but eventually returned to 14.283 lb. with slightly different values (Figure 3).



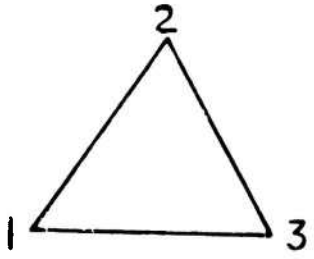
(a) Axial Force Member



(b) Quadrilateral Shear Panel



(c) Half-Web Element



(d) Triangular Membrane Plate

Figure 1. Finite Elements

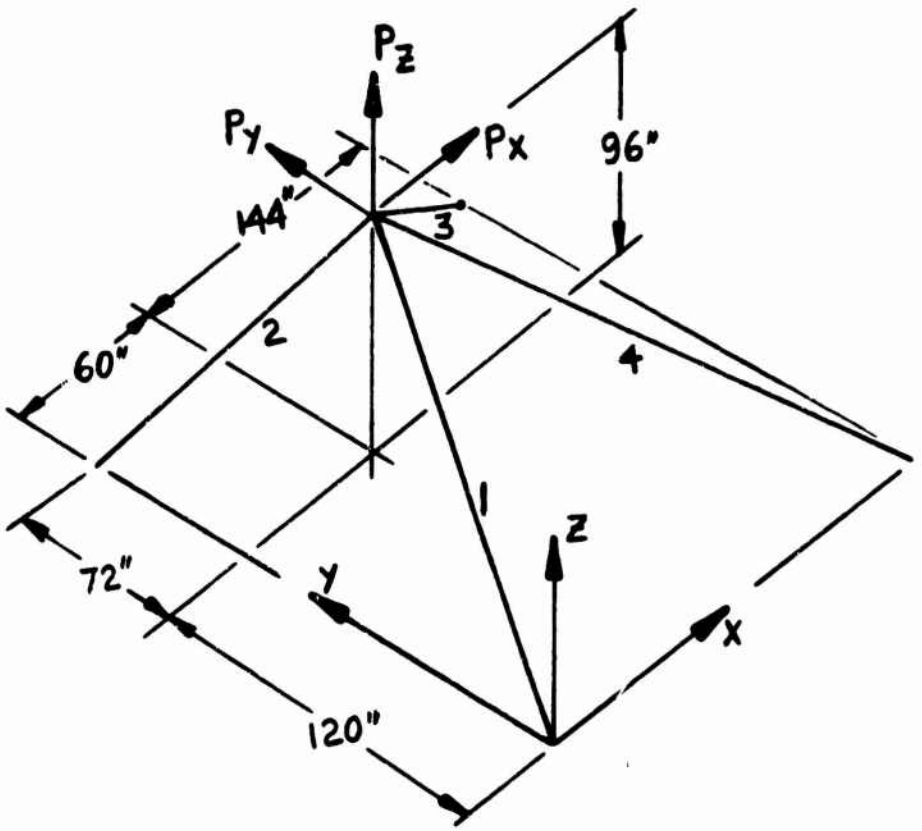


Figure 2. Four-Bar Pyramid

TABLE I

FOUR-BAR PYRAMID WITH THREE LOADS

Iteration	Weight	A_1	A_2	A_3	A_4
1	15.8407	0.2272	0.2272	0.2272	0.2272
2	14.5198	0.2714	0.2960	0.1552	0.1502
3	14.2949	0.2362	0.3226	0.1691	0.1377
4	14.2833	0.2241	0.3277	0.1756	0.1376
5	14.2861	0.2194	0.3303	0.1764	0.1391
6	14.2871	0.2172	0.3319	0.1751	0.1410
7	14.2868	0.2160	0.3329	0.1733	0.1428
8	14.2860	0.2152	0.3337	0.1714	0.1445
9	14.2853	0.2147	0.3344	0.1697	0.1460
10	14.2847	0.2143	0.3349	0.1682	0.1473
11	14.2842	0.2140	0.3353	0.1669	0.1484
12	14.2839	0.2138	0.3357	0.1658	0.1494
13	14.2836	0.2136	0.3360	0.1648	0.1502
14	14.2834	0.2134	0.3363	0.1640	0.1508

(a) Iteration History, Minimum Area = 0.1 in^2

Iteration	Weight	A_1	A_2	A_3	A_4
1	15.3821	0.3176	0.3065	0.2033	0.1054
2	15.2915	0.2181	0.3011	0.2108	0.0970
3	15.2126	0.3196	0.2964	0.2174	0.0897
4	14.3204	0.3264	0.2571	0.2571	0.0365
5	14.2986	0.3192	0.2592	0.2566	0.0351

(b) Iteration History, Minimum Area = 0.01 in^2

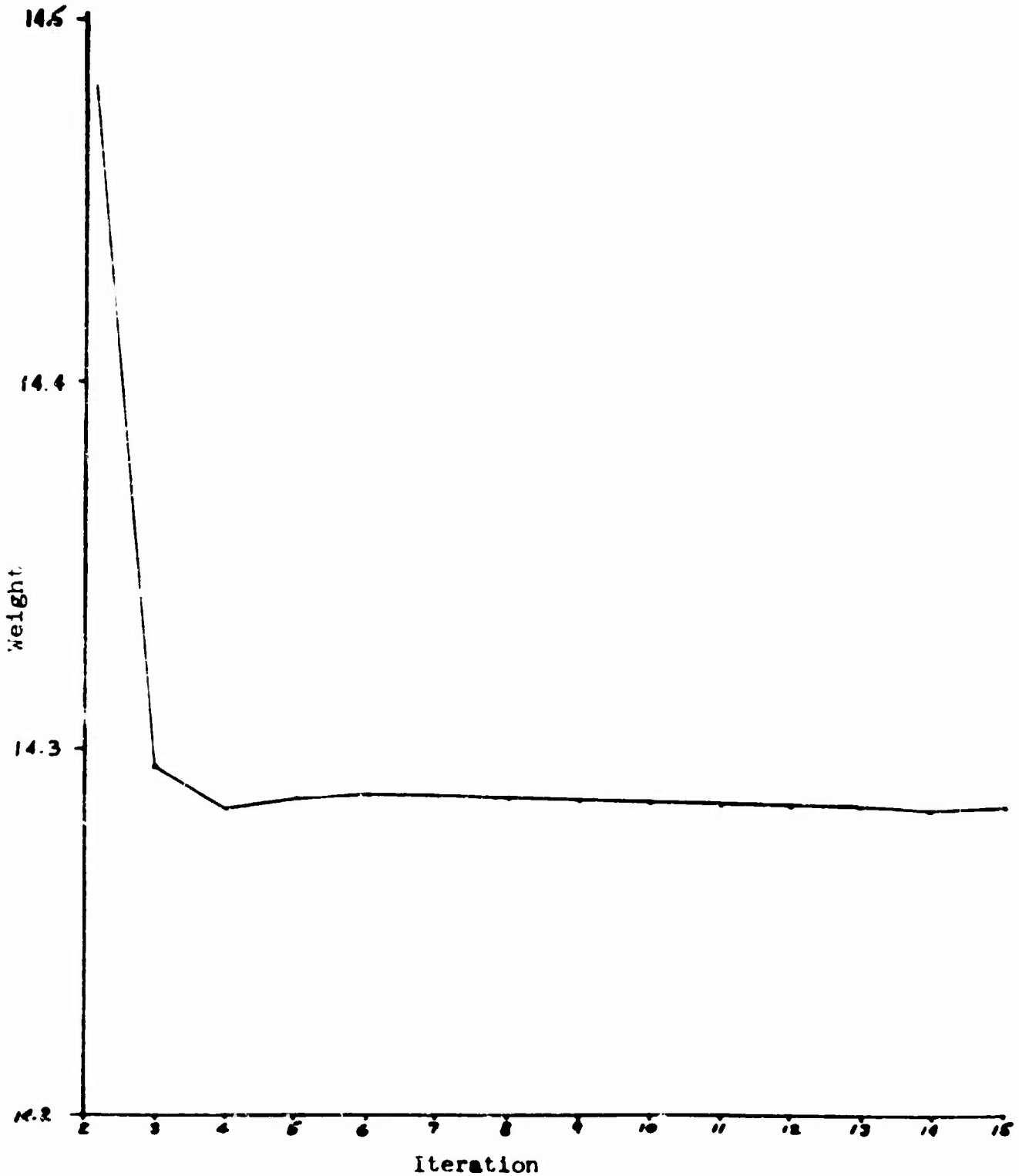


Figure 3. Iteration History for Four Bar Pyramid with Three Loads

The final designs, essentially the same as that of Reference 22, were all limited by the two displacement constraints. Neither stress nor lower side constraints (minimum member sizes) were active in the final designs. In an early development stage of the program, some difficulty was experienced in handling the side constraints effectively. Since the known least weight design was not constrained by the minimum member sizes, these were altered from 0.1 in^2 to 0.01 in^2. With this alteration a minimum weight design of 14.298 lb. was determined and is given in Table Ib. This design is considerably different from the previous design although of essentially the same weight. The indirect influence of the inactive side constraints on the final design is of interest.

B. 25-Bar Transmission Tower (Figure 4)

This structure has been used as a demonstration problem on a number of previous occasions (References 3, 5, 6, and 21). The tower, representative of a structure carrying transmission lines, has 25 axial force members. There are two loading cases (Table 3a). In spite of the directional nature of these loads, the structure is required to be doubly symmetric about the X and Y axes. The maximum displacements of the upper nodes 1 and 2 are ±0.35 in. in the X and Y directions. The bar members are all designed for a 35,000 psi tensile stress and a compressive stress based upon buckling allowables for thin wall circular tubes (Table 2b). Minimum member size is 0.01 in^2.

In some optimizations, symmetry is achieved through the use of additional loading cases but this is not necessary with the present program. Among the previous studies, slight differences are found between the final designs generated. These differences are principally associated with variations in tolerances and specified allowable stresses. An average value is 550 lb. and using a numerical search method over 100 analyses were required.

Using the current program, convergence occurred in seven iterations on a weight of 545.4 lb. Details of the history and final design are given in Tables 2c and 2d and Figure 5.

C. 72-Bar Four Level Tower (Figure 6)

The doubly symmetric tower of Figure 6 presented by Venkayya (Reference 21) was optimized using the present computer programs. Two loading cases were applied (Table 3a) and as in the case of the 25-bar transmission tower, symmetry was achieved by use of the input option. The stress limits were ±25,000 psi on all members and the displacements of the uppermost node points were limited to ±0.25 in. in the X and Y directions. Minimum member size is 0.1 in^2.

In spite of the greater size of the problem, convergence was very rapid and only eight iterations were required. Two subsequent iterations were performed to check convergence and the final design weight was 396 lb. Details of the design are given in Tables 3b, 3c, and Figure 7.

D. Cantilever Frames (Figure 8)

The 10-bar cantilever plane frame is subjected to a single loading case as indicated in Figure 8. The stress limits on all members is ±25,000 psi and limits of ±2.0 in. are also specified on the vertical displacement of the nodes at the free end. The minimum member size is 0.1 in^2, $E = 10 \times 10^6$ psi and $\rho = 0.1$ lb/in^3.

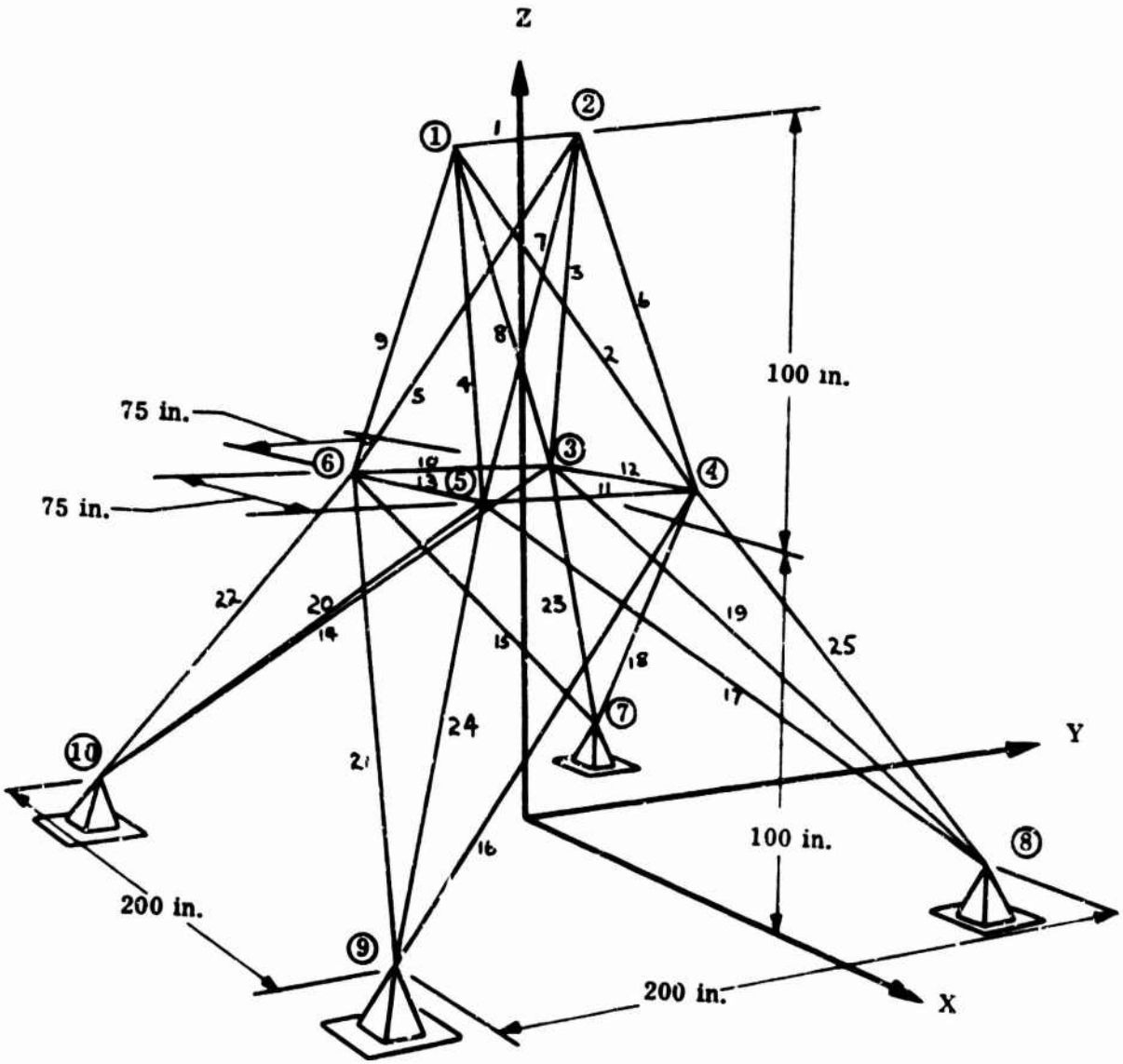


Figure 4. Twenty-five Bar Transmission Tower

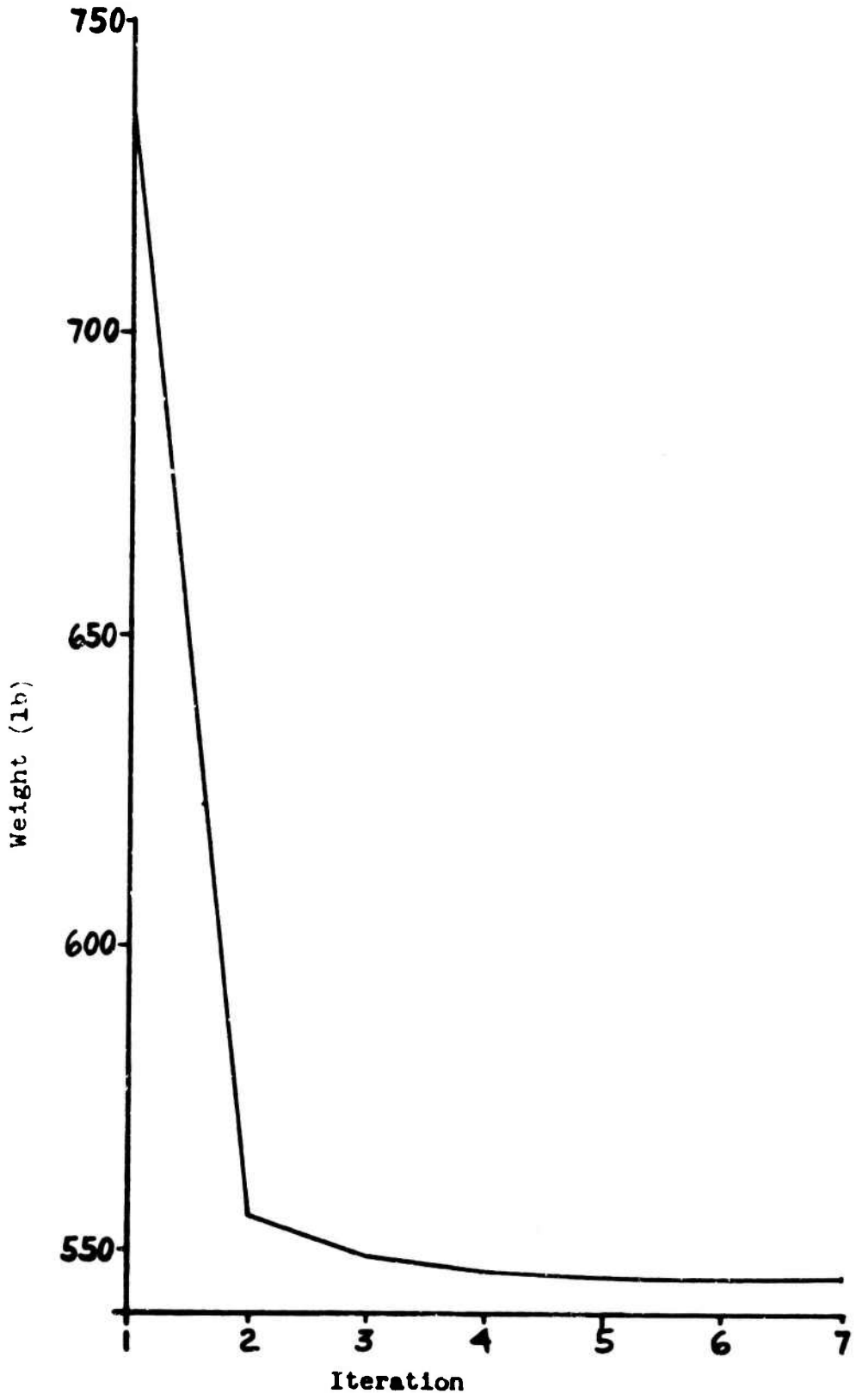


Figure 5. Iteration History for Twenty-five Bar Transmission Tower

TABLE II

25-BAR TRANSMISSION TOWER

Load Condition	Node	Direction		
		X	Y	Z
1	1	1,000	10,000	-5,000
	2	0	10,000	-5,000
	3	500	0	0
	6	500	0	0
2	5	0	20,000	-5,000
	6	0	-20,000	-5,000

a) Applied Loading Systems

Member	Allowable Stress	Member	Allowable Stress
1	-35092.	12, 13	-35092.
2, 3, 4, 5	-11590.	14, 15, 16, 17	-6759.
6, 7, 8, 9	-17305.	18, 19, 20, 21	-6959.
10, 11	-35092.	22, 23, 24, 25	-11082.

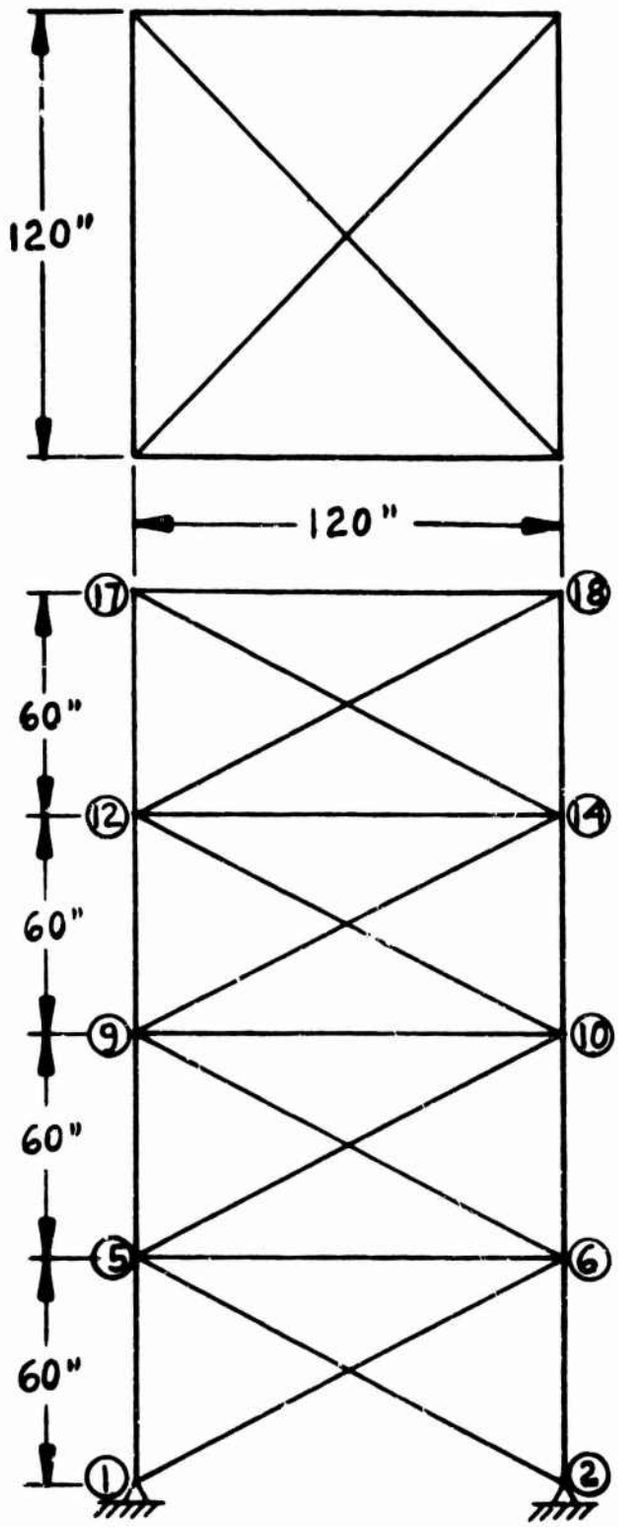
b) Allowable Compressive Stresses

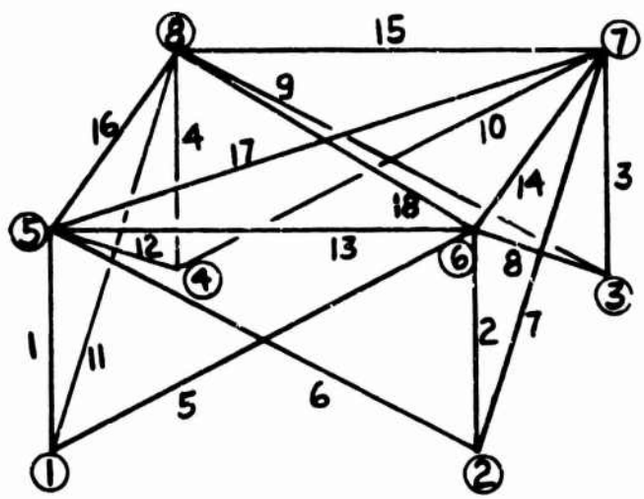
Iteration	1	2	3	4	5	6	7
Weight	734.38	555.72	549.08	546.54	545.92	545.45	545.36

c) Iteration History

Member	Area	Member	Area
1	0.01	12, 13	0.01
2, 3, 4, 5	2.0069	14, 15, 16, 17	0.6876
6, 7, 8, 9	2.9631	18, 19, 20, 21	1.6784
10, 11	0.01	22, 23, 24, 25	2.6638

d) Final Design





Typical Element Numbering for First Level

Figure 6. 72-Bar Four Level Tower

TABLE III

72-BAR TOWER

Load Condition	Node	Direction		
		X	Y	Z
1	17	5,000	5,000	-5,000
2	17	0	0	-5,000
	18	0	0	-5,000
	19	0	0	-5,000
	20	0	0	-5,000

a) Applied Loading Systems

Iteration	1	2	3	4	5	6	7	8
Weight	656.77	416.07	406.21	399.06	396.82	396.25	396.02	395.97

b) Iteration History

Member	Area	Member	Area
1, 2, 3, 4	1.4636	37, 38, 39, 40	0.5521
5, 6, 7, 8, 9, 10, 11, 12	0.5207	41, 42, 43, 44, 45, 46, 47, 48	0.6084
13, 14, 15, 16	0.1	49, 50, 51, 52	0.1
17, 18	0.1	53, 54	0.1
19, 20, 21, 22	1.0235	55, 56, 57, 58	0.1492
23, 24, 25, 26, 27, 28, 29, 30	0.5421	59, 60, 61, 62, 63, 64 65, 66	0.7733
31, 32, 33, 34	0.1	67, 68, 69, 70	0.4534
35, 36	0.1	71, 72	0.3417

c) Final Design

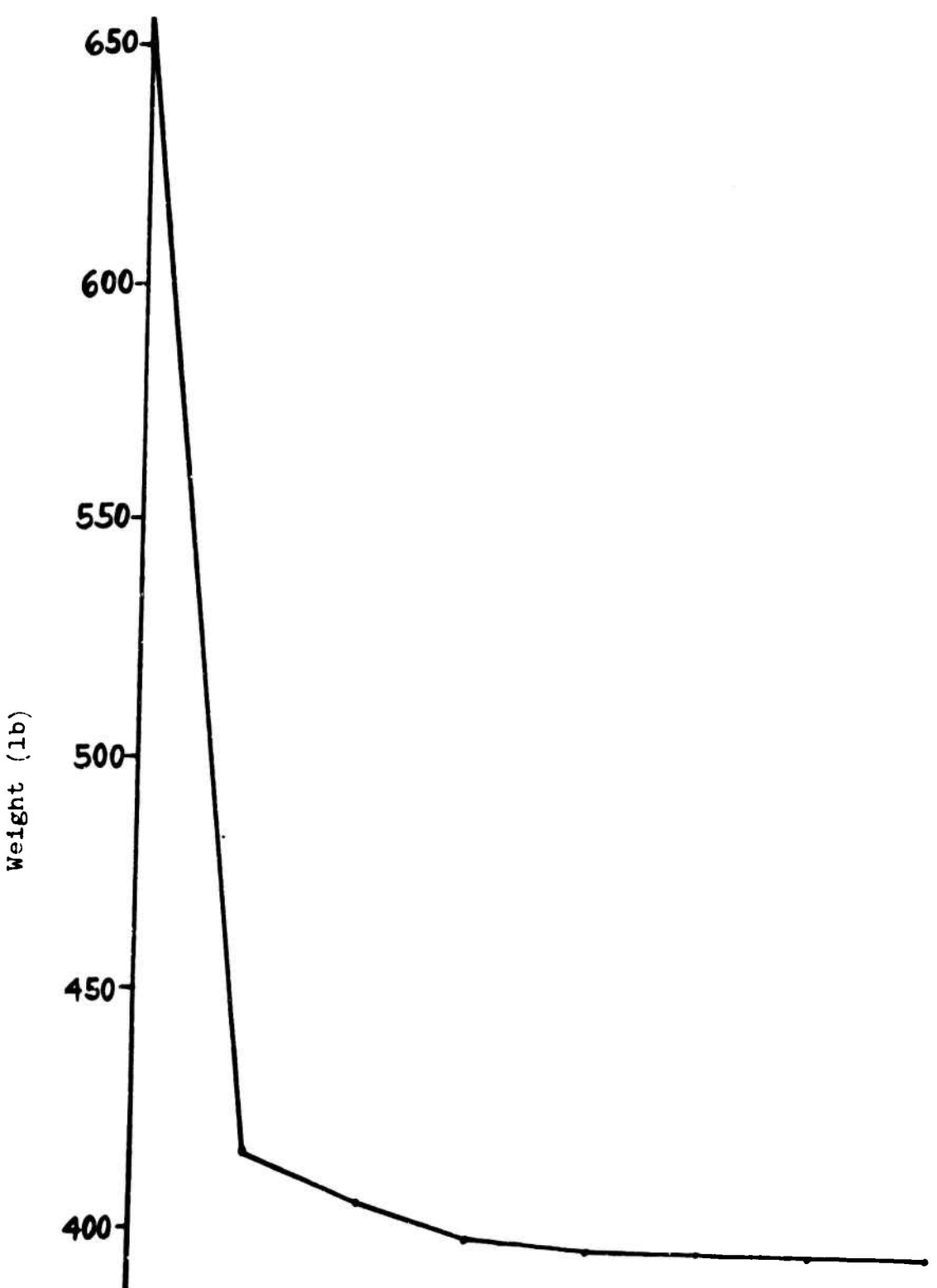


Figure 7. Iteration History for Seventy-two Bar Tower

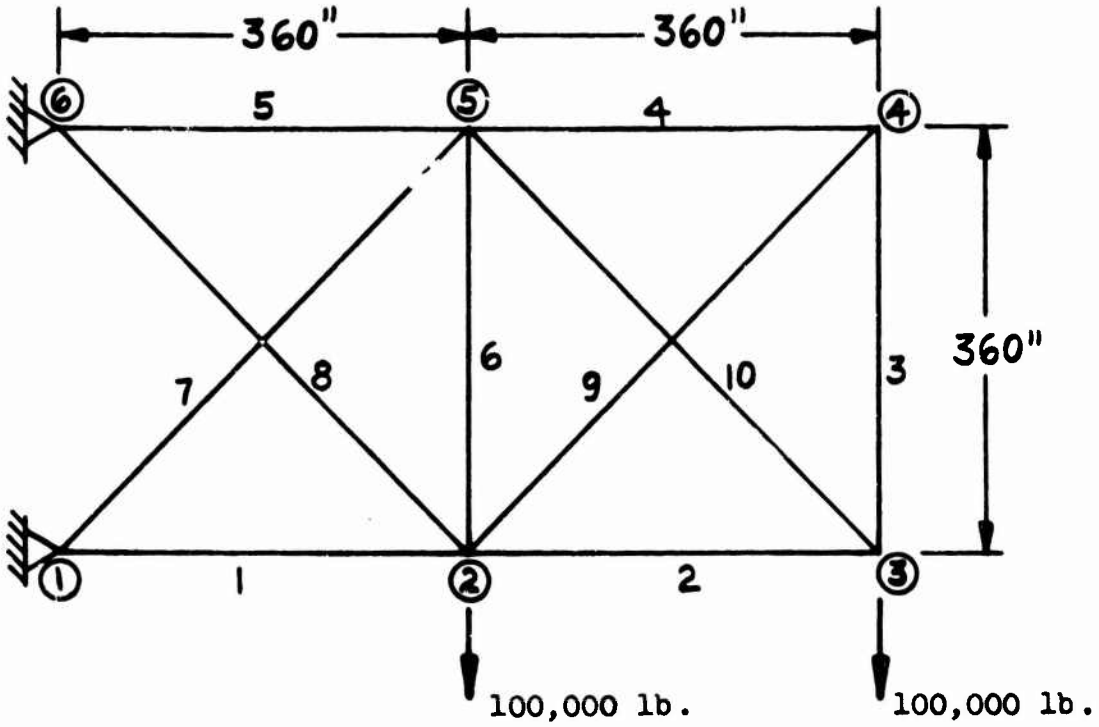


Figure 8. Cantilever Frame

This problem proved to be a particularly interesting one requiring a number of attempts for solution. In the first attempt, since program execution time is a function of the number of constraints, a limit of 2.0 in. was imposed on the vertical displacement of point 3 for which the largest displacement was anticipated. After a considerable number of iterations (using an early edition of the program) a design weight of 5233 lb was achieved, but unfortunately the vertical displacement of point 4 was 2.037 in. In view of this result, the problem was then rerun with the limit of 2.0 in. imposed on node point 4 instead of 3 After three iterations, a design weight of 4625 lb was generated, but at iteration 4, the weight jumped to 4833 lb. The program was allowed to run on and eventually converged after 29 iterations on a design of 4714 lb. Unfortunately, this design was also unacceptable having a displacement of 2.9 in. at point 3. The least weight design at 4625 lb also exceeded the acceptable displacement at point 3 by 0.833 ins.

A third run was carried out with both points 3 and 4 limited to 2.0 in. vertical displacement. The program was run for 47 iterations and the weights varied as indicated in Table 4 and Figure 9. The best design of 5112 lb. occurred at iteration 18 but subsequently the weight did go as high as 9029 lb. The process did not converge as in the previous run but was apparently settling down towards the end. Also of interest in this run is the reversal at iteration 14. The slight weight increase between iterations 13 and 14 was only of minor effect, but its presence did make the selection of an automatic termination criterion difficult.

The termination mode finally selected is a threefold test on (i) number of iterations (ii) convergence of successive iterations or (iii) if a non-convergent process generates a design which exceeds the least weight system found by a specified percentage. A typical value for this last test may be about 5% to allow the perturbation at Iterations 13 and 14, but to cut off at Iteration 19.

The dramatic rise in weight at Iteration 19 appears to have been caused by the sudden emergence of a constraint condition hitherto dormant which became associated with a major redistribution of internal forces in the frame.

This problem has been studied analytically (Reference 23) with a view to comparing the minimum weight achievable for the given configuration with an equivalent Michell structure. If members 3, 4, 6, and 9 are omitted, the theoretical minimum weight for the degenerate structure, satisfying the displacement constraint, is approximately 4970 lb. The corresponding Michell structure lying within the bounds of the present system has a weight of 4760 lb. The imposition of minimum sizes for the above omitted members necessitates some redistribution of internal strains, and prevents the ideal minimum weight being achieved by the small margin of 2.5%.

E. 18 Element Wing Box (Figure 10)

To evaluate the use of plate elements, the wing structure of Figure 10 was selected. This structure is essentially similar to that used in Reference 5, but with a small modification in the idealization. In Reference 5, one-half of the symmetric structure was idealized using five axial members, eight half-web shear panels, two quadrilateral and one triangular membrane elements. Since membrane quadrilaterals are not currently available in the new optimization program, the two quadrilaterals were replaced by four triangles as indicated in Figure 10. The change in idealization introduces differences between the two optimizations and hence results are not strictly comparable. In order to approximate the original model as closely as possible, thicknesses of the pairs of triangles forming the quadrilaterals are coupled using the program symmetry option.

TABLE IV

CANTILEVER FRAME

Iter. No.	Weight	Iter. No.	Weight	Iter. No.	Weight	Iter. No.	Weight
1	8266	13	5195	25	7508	37	5432
2	6356	14	5206	26	7239	38	5407
3	5980	15	5191	27	7050	39	5393
4	5779	16	5169	28	5818	40	5387
5	5625	17	5147	29	5793	41	5352
6	5547	18	5112	30	5788	42	5324
7	5470	19	5897	31	5767	43	5290
8	5392	20	7640	32	5707	44	5274
9	5323	21	8708	33	5677	45	5263
10	5266	22	9029	34	5595	46	5255
11	5225	23	8371	35	5564	47	5250
12	5200	24	7862	36	5488		

a) Iteration History

Member	Area	Member	Area	Member	Area
1	20.03	5	31.35	9	0.10
2	15.60	6	0.14	10	22.06
3	0.24	7	22.21		
4	0.10	8	8.35		

b) Least Weight Design (5112 lb.)

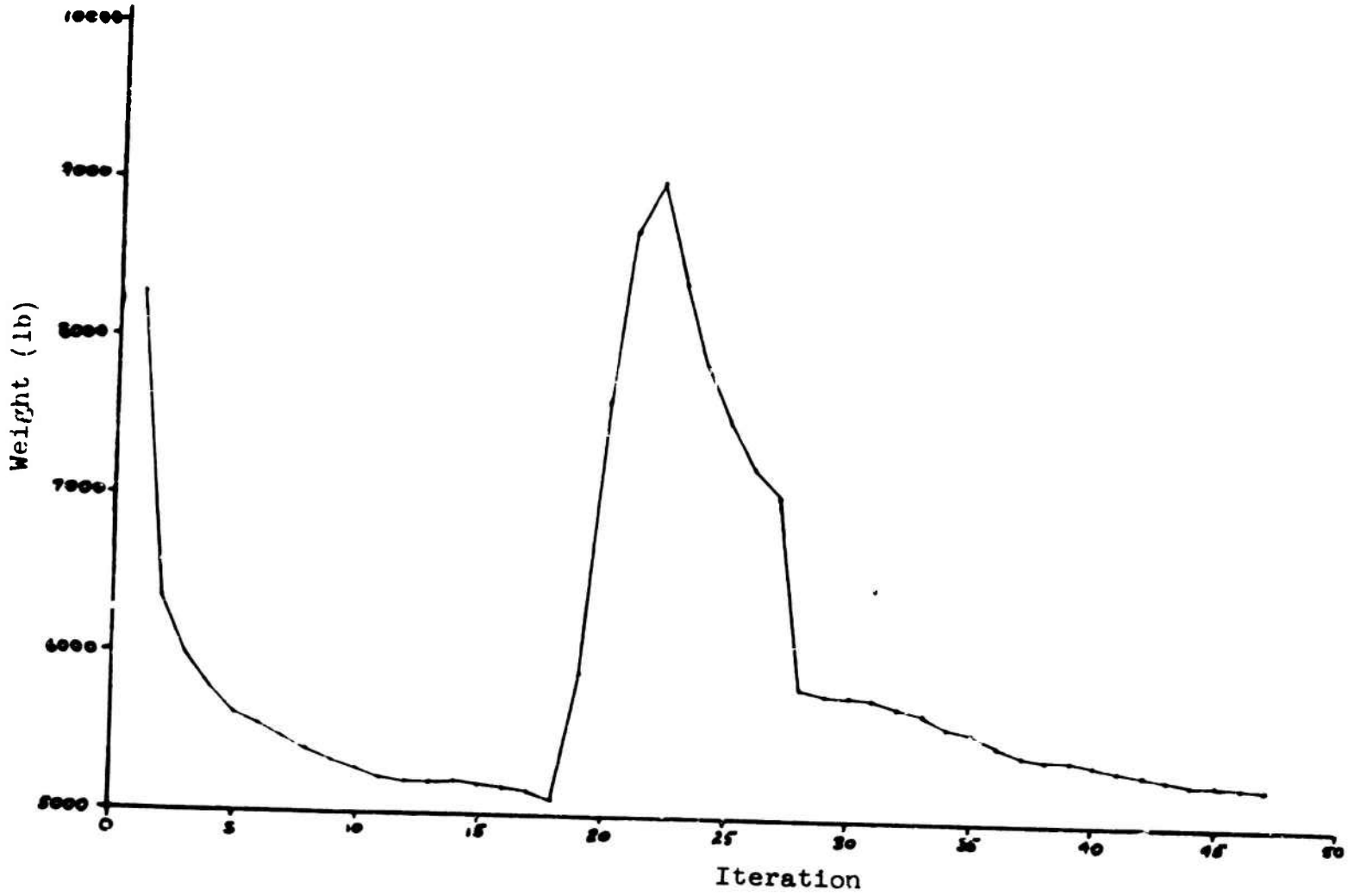
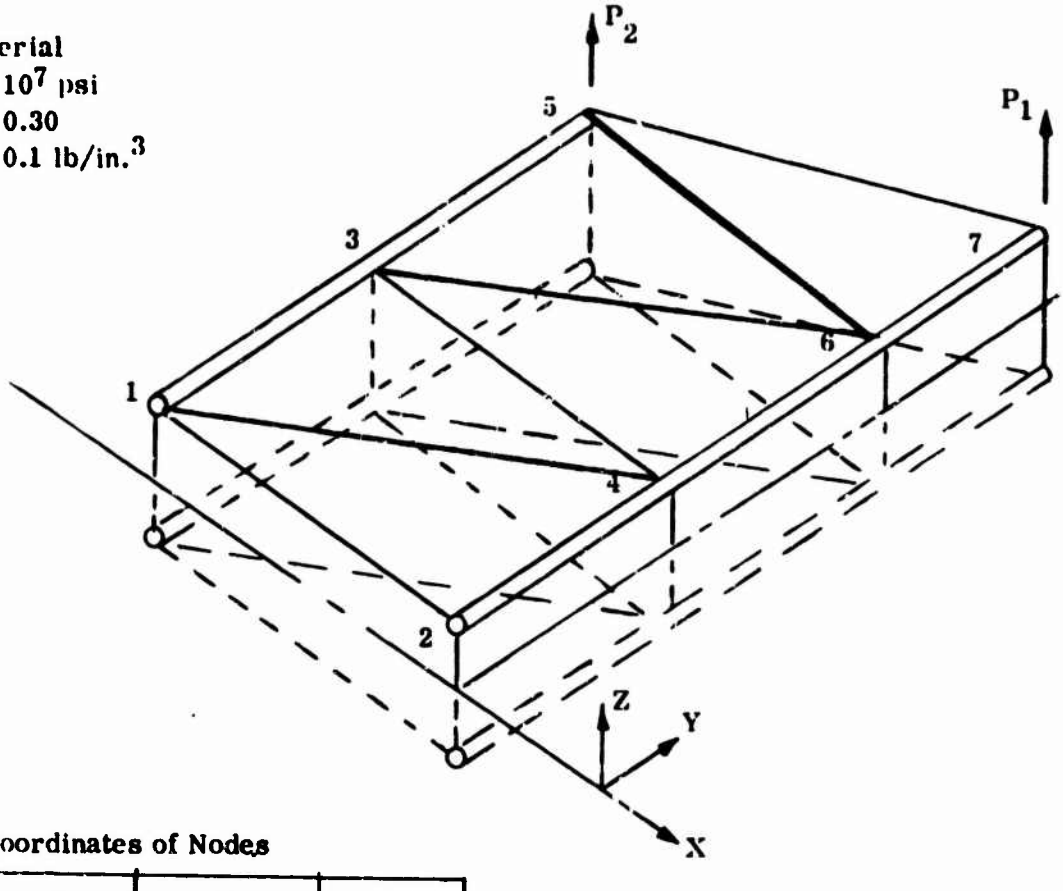


Figure 9. Iteration History for Cantilever Frame

Material
$E = 10^7$ psi
$\mu = 0.30$
$\rho = 0.1$ lb/in.^{3}



Coordinates of Nodes

	X	Y	Z
1	0.0	0.0	10.0
2	100.0	0.0	5.0
3	0.0	70.0	10.0
4	100.0	70.0	8.0
5	0.0	140.0	10.0
6	100.0	140.0	8.0
7	100.0	190.0	5.0

Boundary Conditions

Points 1 and 2 Completely Restrained

Two Loading Conditions

(1) $P_1 = 10^4$ lb

(2) $P_2 = 2 \times 10^4$ lb

All Stress Limits = $\pm 10^4$ psi

All Deflection Limits = ± 2.0 in.

Lower Limits on Axial Members = 0.10 in.^{2}

Lower Limits on Plate Members = 0.02 in.

Figure 10. Eighteen Element Wing Box

Two vertical loads are applied to the tip of the wing and the maximum allowable displacements are 2.0 in. at the tip. The maximum stresses are ±10,000 psi in all members, and the minimum sizes are 0.1 in^2 for axial members and 0.02 in. for plates.

Using the present program, a least weight design of 387.6 lb wes generated in 4 iterations. The process was relatively stable and as indicated in Table 5a and Figure 11, the weight increased slightly but steadily on subsequent iterations. 8y comparison, the numerical search approach generated a design (with quadrilateral cover plates) weighing 381.2 lb. The search procedure was relatively lengthy and required a total of 193 analyses.

This final design was used as an input to the current optimization program, which immediately indicated that the stresses in the triangular elements were in violation. Upon scaling to eliminate this violation, which arises from the different idealizations, a weight of 389.8 lb was produced. The redesign process was continued from this point and produced a design of 378.8 lb at the second iteration. These latter designs differed considerably from the former although of similar weight as indicated in Table 5b. A third starting point was selected which again generated after 22 iterations a design weighing 378.8 lb. although considerably different in member sizes from the previous design.

In the idealization of the wing both direct stress-carrying covers and spar caps are provided. The results generated tend to indicate that the wind bending stresses can be carried equally efficiently by either covers and spar caps. This implies that a large number of designs of similar weight probably exist with various distributions of materials between spar caps and covers.

Also of relevance is the difference in weights between the wing designed using two idealizations - triangles and quadrilateral membrane cover panels, indicating a considerable degree of sensitivity of the optimal design to the precise nature of the analytical model.

IV. CONCLUSIONS

The approach to the weight minimization of fixed geometry structures with displacement and stress constraints based upon the use of optimality criteria appears to offer considerable advantages over mathematical programming based methods. For comparable problems the present method reaches a similar or better design in considerably less iterations than most numerical search methods and at considerable reduced computational costs.

The most efficient optimization method apparently published so far is attributable to Venkayya (References 21 and 22). For identical problems, the numbers of iterations required to determine the minimum weight designs are essentially the same for the mathematical programming and optimality criteria approaches. A difference does reside in the computational effort since presence of active displacement constraints requires the determination of gradients of constraint functions which, as noted previously, will increase computational time considerably for large numbers of variables. With the use of optimality criteria, these calculations are avoided and the computational effort needed at each redesign stage is relatively small and approximately directly proportional to problem size.

TABLE V

WING BOX

Iteration	Weight	Iteration	Weight	Iteration	Weight
1	593.44	6	387.91	11	388.15
2	407.09	7	387.68	12	388.14
3	388.95	8	387.85	13	388.23
* 4	387.67	9	387.97	14	388.26
5	387.90	10	388.07	15	388.28

* Least Weight Design

a) Iteration History

Member		Area/Thickness	
Type	Nodes	Present Method	Ref. 14 (scaled)
Axial	13	0.6505	1.0431
	35	0.1001	0.1036
	24	0.2366	0.3508
	46	0.2352	0.3315
	67	0.1001	0.1035
Half Web	13	0.0876	0.0876
	35	0.0889	0.0895
	24	0.0808	0.0664
	46	0.0768	0.0553
	67	0.0815	0.0537
	34	0.0200	0.0219
	56	0.0200	0.0215
	57	0.0337	0.0256
Triangle	124	0.1328	0.1441
	431	0.1328	0.1441
	346	0.0702	0.0599
	653	0.0702	0.0599
	567	0.0449	0.0435

b) Final Designs

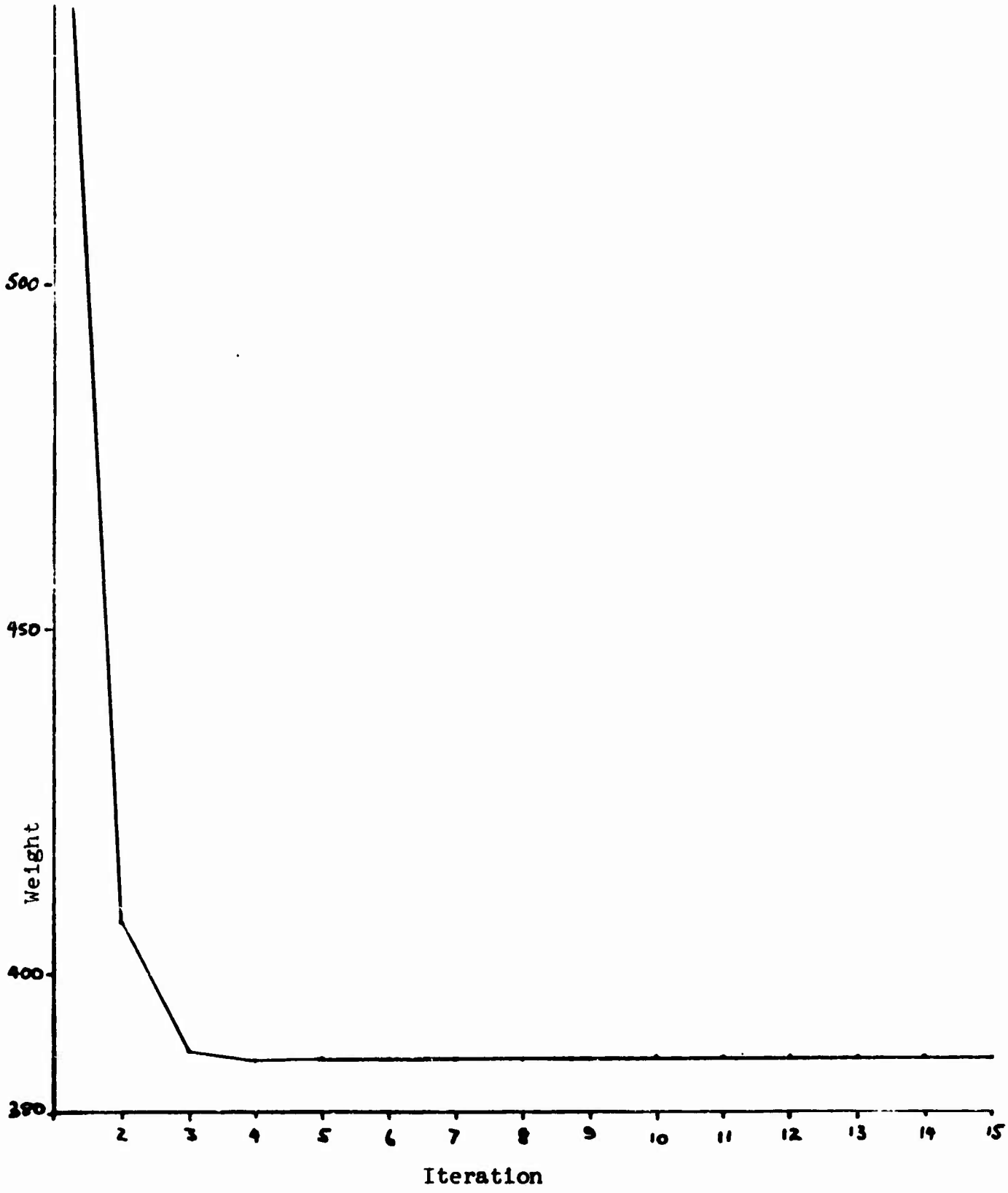


Figure 11. Iteration History for Wing Box

The results presented for the specific examples are very encouraging and indicate that some, if not all, of the difficulties encountered in large scale optimization problems can be eliminated through this type of approach. It is recognized that certain problems still remain to be resolved, particularly with regard to convergence characteristics. While the use of the omitted terms in Equation (15) might possibly affect some of the convergent behavior, their inclusion would certainly tend to negate the computational advantages indicated by the present approach.

The extension of the procedures presented above beyond stress and displacement constraint conditions can be accomplished within the general framework discussed in this paper. The criteria for the automatic partitioning between active and passive members is generally valid for other types of response phenomena. It is then only necessary to derive the criteria which must be satisfied (exactly or approximately) at the optimal design in a form analogous to Equation (15), containing provision for both variant and invariant member sizes. In such an extension, it must be recognized while the convergence can be expected to be more rapid than a nonlinear-programming approach, that there will still be a considerable level of computational complexity due to the fundamental nature of the analysis processes associated with the relevant response phenomena.

The approach presented herein represents a major step in a new direction of exploration and also provides a basis for many more developments in structural optimization.

V. REFERENCES

1. Schmit, L. A., "Structural Design by Systematic Synthesis", Proc. of 2nd National Conf. on Electronic Computation, ASCE, 1960, pp. 105–132.

2. Gellatly, R. A., Gallagher, R. H., and Luberacki, W. A., Development of a Procedure for Automated Synthesis of Minimum Weight Structures, AFFDL-TR-64-141, 1964.

3. Gellatly, R. A., Development of Procedures for Large Scale Automated Minimum Weight Structural Design, AFFDL-TR-66-180, 1966.

4. Fiacco, A. V., and McCormick, G. P., Non-linear Programming; Sequential Unconstrained Minimization Techniques, Wiley, New York, 1968.

5. Gellatly, R. A., and Marcal, P. V., "Investigation of Advanced Spacecraft Structural Design Technology," NASA Contract NASW-1344, Bell Aerosystems Report No. 2356-950001, 1967.

6. Fox, R. L., and Schmit, L. A., "Advances in the Integrated Approach to Structural Synthesis," Journal of Spacecraft and Rockets, Vol. 3, No. 6, 1966.

7. Schmit, L. A., Morrow, W. M. and Kicher, T. P., "Structural Synthesis Capability for Integrally Stiffened Cylindrical Shells," AIAA/ASME 9th Struct., Struct. Dynamics and Materials Conf., Palm Springs, California, 1968.

8. Fiacco, A. V., and McCormick, G. P., "Sequential Unconstrained Minimization (SUMT) Without Parameters," Operations Research, Vol. 9, No. 2, 1961.

9. Carroll, C. W., "The Created Response Surface Technique for Optimizing Nonlinear Restrained Systems," Operations Research, Vol. 9, No. 2, 1961.

10. Pope, G. G., and Schmit, L. A. (Ed.), "Structural Design Applications of Mathematical Programming Techniques," AGARDograph 149, NATO, 1971.

11. Gellatly, R. A. (Ed.), "Proceedings of AGARD Symposium on Structural Optimization," Istanbul, AGARD-CP-36-70, 1970.

12. Lansing, W., Dwyer, W., Emerton, R., and Ranalli, E., "Application of Fully Stressed Design Procedures to Wing Empennage Structures," AIAA/ASME 11th Struct., Struct. Dynamics and Materials Conference, Denver, Colorado, 1970.

13. Cilley, F. H., "The Exact Design of Statically Indeterminate Frameworks, An Exposition of Its Possibility, But Futility," Trans ASCE, Vol. 42, June 1900.

14. Schmidt, L. C., "Stressed Design of Elastic Redundant Trusses Under Alternative Load Systems," Austrailian Journal of Applied Science, Vol. 9, December 1958.

15. Razani, R., "The Behavior of the Fully-Stressed Design of Structures and Its Relationship to Minimum-Weight Design," AIAA 2nd Aerospace Sciences Meeting, New York, January, 1965.

16. Sved, G., and Ginos, Z., "Structural Optimization under Multiple Loading," Int. Journal of Mech. Science, Vol. 10, 1968.

17. Cox, H. L., The Design of Structures of Least Weight, Pergamon Press, Oxford, 1965.

18. Barnett, R. L., and Hermann, P. C., "High Performance Structures," NASA CR-1038, 1968.

19. Berke, L., An Efficient Approach to the Minimum Weight Design of Deflection Limited Structures, AFFDL-TM-70-4-FDTR, May, 1970.

20. Gellatly, R. A., and Berke, L., "Optimal Structural Design," AFFDL-TR-70-165, February, 1971.

21. Venkayya, V. B., Khot, N. S., and Reddy, V. S., Energy Distribution in an Optimum Structural Design, AFFDL-TR-68-156.

22. Venkayya, V. B., 'Design of Optimum Structures," Conference on Computer Oriented Analysis of Shell Structures, Palo Alto, California, August, 1970.

23. Cox, H. L., Private Communications, June–August 1971.

STRUCTURAL OPTIMIZATION VIA STEEPEST DESCENT AND INTERACTIVE COMPUTATION

Haug, Jr., E. J.,[1] and Arora, J. S.[2]
The University of Iowa

In this paper a hybrid method of optimal structural design is developed and applied, using a steepest descent optimization technique in an interactive mode of computing. A method of computing design sensitivity is developed as part of a steepest descent optimization method. A technique is presented for utilization of this data by an experienced designer, interacting with the iterative computer algorithm. The designer is allowed to alter the structural configuration to seek a global, rather than just a local, optimum. Two examples are solved with a number of loading conditions and design constraints to illustrate the flexibility of the technique. Results for these problems are of interest in their own right.

Introduction

Structural optimization methods developed during the past decade generally fall into one of two categories. The first consists of a number of optimization techniques which seek to determine an optimum design, within a well defined mathematical structure, by purely mathematical techniques. The second approach consists of providing the designer with an interactive computing tool, such as computer graphics, with which he can try nominal designs, get rapid analysis feedback, and alter his initial design based on his knowledge of structural behavior. Both methods have been used with varying degress of success on a variety of design problems. In general, the first approach has been used on smaller scale structures with well defined optimality criteria, such as minimum weight or maximum stiffness. The second approach has been used to aid designers in large scale structural design problems, primarily airframe design, such as the Air Force C-5 transport aircraft.

The possibility of utilizing a combination of these two methods for structural design has been the subject of a recent paper [1]. The present paper presents the specifics of application of an optimization technique with, designer interaction. This hybrid approach is appealing from a number of points of view. First, the problem of topological design, i.e., determination of optimum structural configuration, has been addressed with very limited success from an analytical point of view [Kato]. Topological design, in practice, is done by experienced structural designers, occasionally with the aid of interactive computation. Combined analytical and interactive computing methods appear to be essential for this important class of problems. A second problem area arises due to the difficulty in formulating a single optimality condition and mathematically precise design constraints. Often, conflicting design constraints and objectives arise during design which require experienced judgment and defy

---

[1]Chief of Systems Analysis, U. S. Army Weapons Command, Rock Island Arsenal, Rock Island, Illinois.

[2]Post-Doctoral Fellow, Department of Mechanics and Hydraulics, The University of Iowa, Iowa City, Iowa.

apriori mathematical formulation. Such problems appear to require an interactive computing capability but should profit from analytical methods which are used in automated structural optimization.

In this paper, an iterative steepest descent optimal design method is used interactively to develop a hybrid method which allows the experienced designer to make key design decisions with the aid of analytical tools which have proved to be quite successful in automated structural optimization. Several structural optimization problems are solved to illustrate the applicability of the method. Prior to developing the hybrid method referred to above, a brief review of automated design optimization methods and interactive computing techniques will be presented to establish the rationale for the methods chosen.

Many numerical methods have been applied to solve structural optimization problems. The method of linear programming has been a very popular tool with many researchers in the area of optimal structural design. Moses [2], Romstad and Wang [3], Grierson and Cohn [4] and many others have used Kelley's [5] cutting plane method for nonlinear programming. In a recent paper, Johnson and Brotten [6] incorporated the force method of structural analysis with the linear programming technique to solve some redundant trusses. They have considered only stress constraints in their problems. Venkayya, et. al. [7] have used a scheme based on the study of distribution of strain energy in the structure to optimize trusses. Structures subject to stress and displacement constraints are considered. A simple recurrence relationship is used to resize the members when only stress constraints are present. However, in the presence of displacement constraints, the problem is handled in two stages.

In recent years, many authors have used SUMT, developed by Fiacco and McCormick [8], in optimal structural design. Kavlie and Moe [9, 10] have used an interior penalty technique to solve some frame and grillage foundation problems and have reported good success with the method. They have also extended the technique so that the initial designs may be infeasible. In order to compare linear programming and SUMT, LaPay and Goble [11] solved some truss problems. They have found the nonlinear formulation to be numerically superior and more efficient. A slightly different formulation of SUMT is suggested by Schmit and Fox [12]. The unconstrained function Ψ, that is constructed contains, among other quantities, an estimate of the optimum weight of the structure. The function Ψ is so constructed that, when $\Psi = 0$, all constraints and equations of equilibrium and compatability are satisfied.

The methods cited above are based primarily on mathematical optimization ideas which yield a minimizing sequence. These techniques, however, do not give the designer much useful design information at intermediate steps which would allow him to interact with the computer algorithm. There are techniques, however, which do give design trend information at intermediate steps of the computational algorithm. Most of these methods are based on gradient projection or steepest-descent ideas.

Schmit and his co-workers [13] have used a steepest-descent alternate-step procedure for many years. Gellatly and Gallagher [14] have applied this technique and have given it a theoretical basis. Felton and Hofmeister [15] and Ridha and Wright [16] have also used it to solve certain structural design problems. A variation of the method of steepest-descent is Zoutendijk's method of usable-feasible directions. In this method, the steepest-descent mode is also followed until some constraints are encountered, at which point the direction of

travel lies between the gradients of the objective and the constraint functions. Fox and Kapoor [17], Pope [18], and Moses and Onoda [19] used this technique in conjunction with linear programming to solve some structural design problems. They have found the method to be quite efficient.

Best [20] was probably the first to suggest the use of Rosen's gradient projection method [21] to solve optimal structural design problems. Brown and Aug [22], and Seaburg and Salmon [23] have used this method quite successfully. Recently, a method of constrained steepest-descent with state equations has been used by Haug [24, 25]. This method is a variation of Rosen's gradient projection method and uses Kuhn-Tucker necessary conditions along with Lagrange multipliers. Bartel and Rim [26] have compared this method with SUMT for two and three member space frames and have found the method to be superior to SUMT in these cases.

The second element of the interactive optimal design technique is the mode of designer-computer interaction. There are several types of interactive equipment available to the designer today, the most promising of which appears to be interactive graphics. For discussions of graphics systems previously considered in structural design the reader is referred to [1, 27, 28]. The interactive system intended for use in this study consisted of a PDP-8/I computer and a Tektronic 4001A graphics terminal located remotely to an IBM 360-65 computer which operates in an interactive mode. Due to a computer center decision to delay implementation of large scale, interactive system, for this study had to be simulated. Instructions were prepared and computations were run in the batch mode. Output data was then displayed and analyzed just as it would be in the interactive mode and instructions for recomputation were given by the designer and the process repeated. The delay in designer interaction is felt to degrade performance somewhat over true interactive computing since the designer tends to forget pertinent detailed data during the time delay. For this reason, the results of this study should be a conservative estimate of the designers performance in a truly interactive mode.

Design Sensitivity Analysis and Optimization

A first step in providing information for structural design and optimization is mathematical formulation of the design problem and determination of the effect of design changes on structural performance. In this paper, a finite element formulation is utilized which allows design variables to be member size, material parameters, or structure dimensions. A distinction is made between design variables $b = [b_1, \ldots, b_m]^T$ and state variables $q = q_1, \ldots, q_n]^T$ which are generalized displacements. To allow treatment of buckling and natural frequency considerations, an additional state variable $v = [v_1, \ldots, v_p]^T$ is defined along with the associated eigenvalue ζ.

As a simple example to illustrate the notation and problem formulation, consider the three bar truss of Figure 1. The design variables b_1, b_2, and b_3 are cross-sectional areas of the members. The displacements q_1 and q_2 of the free joint constitute the state variables. If a frequency constraint is involved then v_1 and v_2 would be the same two displacements which specify the mode of vibration. This simple problem will be used later as an illustrative example in design for minimum weight with constraints on stress, displacement, buckling, and natural frequency.

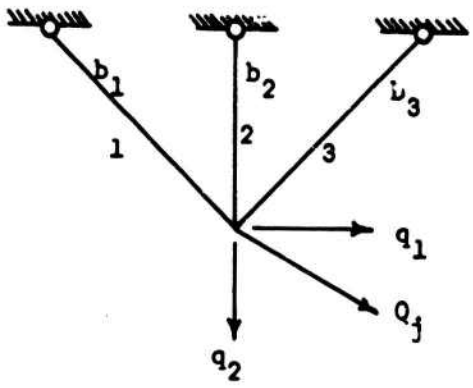


Figure 1. Three Bar Truss

In this section a steepest-descent technique is outlined for a broad class of optimal structural design problems. This technique has been used successfully in the batch, or noninteractive, mode and provides useful data for the experienced designer when it is employed in the interactive mode.

The mathematical model of the problem to be solved is to find b, q, v, and ζ which minimize a cost functional $J(b,q,\zeta)$, satisfy the equations

$$A(b)q = Q(b) \tag{1}$$

$$K(b)v = \zeta M(b)v, \tag{2}$$

and satisfy the constraints

$$\phi(b,q,\zeta) \leq 0. \tag{3}$$

In this formulation, (1) represents the structural response to a load Q. The matrix A(b) plays the role of stiffness matrix whose elements depend on member size b. Likewise, (2) is the eigenvalue problem for natural frequency, where K(b) is a stiffness matrix and M(b) is a mass matrix. The inequality constraint (3) represents stress, displacement, eigenvalue, and member size constraints, where $\phi(b,q,\zeta)$ is a vector of functions and the inequality applies to each component of the vector.

Since analytic methods of finding a solution to the above problem are unwieldy, a numerical procedure, developed in [25] will be summarized here. In the derivation of this procedure, an initial estimate of the design variable vector $b = b^{(0)}$ is made and a small change, δb, in the design variable vector, b is sought such that $b^{(1)} = b^{(0)} + \delta b$ is an improved design in some sense. The procedure is repeated with $b^{(1)}$ as the estimated design point until a convergence criterion is satisfied. In seeking the small change, δb the equations of the above problem are approximated and the theory of nonlinear programming is used to obtain an explicit expression for δb. For example, the change in $J(b,q,\zeta)$ due to changes δb, δq, and $\delta\zeta$ is approximated by

$$\delta J = \frac{\partial J}{\partial b}\,\delta b + \frac{\partial J}{\partial q}\,\delta q + \frac{\partial J}{\partial \zeta}\,\delta\zeta$$

where matrix calculus notation is used freely. Note that in the linearized form of the objective function, $J(b,q,\zeta)$ and the constraint function $\phi(b,q,\zeta)$, small changes δq, δv and $\delta\zeta$ in q, v, and ζ respectively, will be appearing. These variables are eliminated from δJ and $\delta\phi$ by the use of linearized versions of the state equations. From this procedure one obtains [24, 25]:

$$\delta J = \ell^{J^T}\delta b \tag{4}$$

$$\delta\phi = \Lambda^T\delta b, \tag{5}$$

where

$$\ell^J = \frac{\partial J^T}{\partial b}(b,q,\zeta) + \left[\frac{\partial Q(b)}{\partial b} - \frac{\partial}{\partial b}\{A(b)q\}\right]^T\lambda^J$$

$$+ \frac{\dfrac{\partial J(b,q,\zeta)}{\partial\zeta}}{v^T M(b)v}\left[\frac{\partial}{\partial b}\{K(b)v\} - \zeta\frac{\partial}{\partial b}\{M(b)v\}\right]^T v,$$

$$A = \frac{\partial \tilde{\phi}^T}{\partial b}(b,q,\zeta) + \left[\frac{\partial Q(b)}{\partial b} - \frac{\partial}{\partial b}\{A(b)q\}\right]^T \lambda^{\tilde{\phi}}$$

$$+ \frac{1}{v^T M(b) v}\left[\frac{\partial}{\partial b}\{K(b)v\} - \zeta \frac{\partial}{\partial b}\{M(b)v\}\right]^T v \frac{\partial \tilde{\phi}^T}{\partial \zeta},$$

$$\tilde{\phi} = [\phi_i(b,q,\zeta);\ \text{for each } i \text{ such that } \phi_i(b,q,\zeta) \geq 0], \tag{6}$$

$$\Delta\tilde{\phi}_i = -\tilde{\phi}_i(b,q,\zeta), \tag{7}$$

and λ^J and $\lambda^{\tilde{\phi}}$ are solution of

$$A^T(b)\,\lambda^J = \frac{\partial J^T}{\partial q}(b,q,\zeta) \tag{8}$$

$$A^T(b)\,\lambda^{\tilde{\phi}} = \frac{\partial \tilde{\phi}}{\partial q}(b,q,\zeta). \tag{9}$$

In the above, $\tilde{\phi}$ consists simply of those elements of ϕ which are zero or positive at the nominal design. The vector $\Delta\tilde{\phi}$ defined by (7) is the desired change in the violated constraints $\tilde{\phi}$. To verify that one obtains (4) and (5) by the above calculations, the equations (1) and (2) may be linearized about the nominal design and the resulting equations along with (8) and (9) used to eliminate explicit dependence on δq, δv, and $\delta\zeta$. These calculations are carried out in detail in [25].

To assure that δb is small, as required by the linear approximations, introduce a quadratic constraint on δb as follows:

$$\delta b^T W \delta b = \xi^2 \tag{10}$$

where ξ is small and W is a positive definite weighting matrix, chosen by the designer.

The design improvement problem reduces to finding δb to minimize δJ of (4) subject to the constraints (10) and

$$\delta\tilde{\phi} = A^T \delta b \leq \Delta\tilde{\phi}\ .$$

The Kuhn-Tucker necessary conditions for this problem yield:

$$\delta b = -\frac{1}{2\nu} W^{-1}(\ell^J + A\mu),$$

where μ is a Lagrange multiplier vector and ν is a scalar multiplier. By the arguments explained in [25], the value of μ can be found from the following equation:

$$B\mu = -[2\nu\Delta\tilde{\phi} + A^T W^{-1} \ell^J]\ , \tag{11}$$

where

$$B = A^T W^{-1} A\ .$$

Substitution of μ in the expression for δb yields the following equation

$$\delta b = -\frac{1}{2\nu} W^{-1}[I - AB^{-1}A^T W^{-1}]\ell^J + W^{-1}AB^{-1}\Delta\tilde{\phi}$$

Define $\eta = \frac{1}{2\nu}$

$$\delta b^1 = W^{-1}[I - AB^{-1}A^T W^{-1}]\ell^J \tag{12}$$

and

$$\delta b^2 = W^{-1}AB^{-1}\Delta\tilde{\phi}. \tag{13}$$

Thus, δb may be written as

$$\delta b = -\eta\delta b^1 + \delta b^2, \tag{14}$$

where η is an arbitrarily small quantity, interpreted as a step size, which must be chosen by the designer. The following scheme for choosing η has been tried and has worked well. When all the constraints are satisfied, so that $\Delta\tilde{\phi} = 0$, η is chosen to yield a few percent reduction in J. One first chooses the desired change $\Delta J < 0$ in J. Then assuming $\Delta\tilde{\phi} = 0$, ΔJ is given by $\Delta J = -\eta\ell^{J^T}\delta b^1$, from which one obtains

$$\eta = -\Delta J/[\ell^{J^T}\delta b^1]\ . \tag{15}$$

Once η is chosen δb is determined by (14) and an improved design is obtained by adding δb to the preceding design estimate. The procedure is then repeated to form an iterative optimal design algorithm. Details of the algorithm are given in [25]. Computational results may be found in [25] and [29].

In order to implement the computational algorithm, two principal matrices ℓ^J and A must be assembled. These matrices are composed of various matrices such as,

$$\frac{\partial}{\partial b}[A(b)q],\ \frac{\partial}{\partial b}[K(b)v],\ \frac{\partial}{\partial b}[M(b)v],\ \frac{\partial\tilde{\phi}}{\partial q},\ \frac{\partial\tilde{\phi}}{\partial b},\ \frac{\partial\tilde{\phi}}{\partial\zeta},\ \text{etc.}$$

The matrices, $\frac{\partial}{\partial b}[A(b)q]$, $\frac{\partial}{\partial b}[K(b)v]$, and $\frac{\partial}{\partial b}[M(b)v]$ can be computed explicitly from the state equations of the problem. The displacement method of structural analysis is used to obtain the state equations of the problem and the various derivative matrices are computed from them. The matrices such as

$$\frac{\partial\tilde{\phi}}{\partial q}\ ,\ \frac{\partial\tilde{\phi}}{\partial b}$$

etc. can also be computed explicitly by considering the various types of constraints one by one. The procedure of computing these matrices can be automated. For details of these computations the reader is referred to [29].

Steepest-Descent Computational Algorithm

The above procedure of successively improving the best available design can be put in a computational form as follows:

Step 1. Obtain the best available engineering estimate of the optimum

design variable vector $b^{(j)}$ and solve for $q^{(j)}$ from (1). Also compute the lowest eigenvalue $\zeta^{(j)}$ and the corresponding eigenvector $v^{(j)}$ from (2).

Step 2. Check the constraints (3) and form the constraint function $\tilde{\phi}$ of (6). Also choose the constraint error correction vector, $\Delta\tilde{\phi}$ of (7).

Step 3. Compute various matrices such as, $\frac{\partial\tilde{\phi}}{\partial q}$, $\frac{\partial\tilde{\phi}}{\partial b}$, $\frac{\partial\tilde{\phi}}{\partial\zeta}$, $\frac{\partial}{\partial b}[K(b)v]$, etc.

Step 4. Solve for λ^J and for $\lambda^{\tilde{\phi}}$ from (8) and (9).

Step 5. Assemble matrices ℓ^J and A of (4) and (5), respectively.

Step 6. Choose η from (15) and calculate the Lagrange multiplier vector μ from (11).

Step 7. Check the algebraic sign of each component of μ. If some components of μ are negative, remove the corresponding columns from A matrix. Also delete the corresponding elements from the vector $\Delta\tilde{\phi}$ and return to Step 6.

Step 8. Compute δb from (14) and put

$$b^{(j+1)} = b^{(j)} + \delta b$$

Step 9. Make a check for convergence. If $||\delta b^1||$ is sufficiently small, all constraints are satisfied and no further reduction in cost function is possible, terminate the process or return to Step 1 with $b^{(j+1)}$ as the best available estimate of the optimum design.

Interactive Structural Design Using Sensitivity Data

The steepest descent optimization method outlined in the preceding section has been used to solve a number of relatively large scale structural optimization problems with good success [25,29]. All these problems, however, have been well formulated mathematically and have involved structures with a predetermined form. Difficulties have occurred when certain structural elements tend toward zero cross section. Further, no universal method has been found to determine the best step-size η in the optimization algorithm. These and other inherent difficulties in automated optimization lead one to consider interjecting an experienced designer into the computational, optimization algorithm. The result is a hybrid structural optimization technique.

Reconsidering the design improvement step of the optimization algorithm, given by (14), one might draw a vector picture in design space, as is depicted in Figure 2. Here, $-\eta\delta b^1$ is the direction which will yield the greatest reduction in J subject to the required constraints and δb^2 is the design change required to give the desired constraint error correction. While useful in this form, there is a better display of this data for use by the experienced structural designer. The scalar components of δb^1 and δb^2 tell the designer whether he should increase or decrease his individual design variables to obtain desirable changes in overall structural response. Further, relative importance of design variable changes is given. For this reason, δb^1 may be interpreted as a vector of design sensitivity coefficients which relate individual design parameter changes to overall structural characteristics. It is extremely important to note, at this point, that these sensitivity coefficients account for constraints

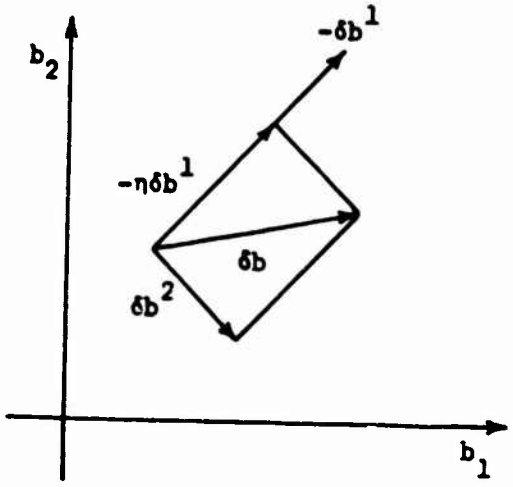


Figure 2. Vector Change in Design Space

implicitly. That is, the direction of change indicated in the design parameters will not cause significant violation in specified performance constraints such as stress limits, deflection limits, etc.

To illustrate these ideas, consider again the simple structural design problem in Figure 1. The cost function here is structural weight. If, for example, the stress in member 1 is at its allowable limit under one of the loads, then the indicated changes in design ($-\delta b_1^1$, $-\delta b_2^1$, $-\delta b_3^1$) will not increase the stress in member 1. To make the design sensitivity data of maximum use to the designer, consider the graphical display in Figure 3. In this display, σ_i are the stresses in the various members. This display gives the experienced designer a clear picture of the manner in which he should change his design parameters to reduce total weight, subject to stress constraints. He can now choose the desired reduction ΔJ in weight and take the resulting design change δb given by (14). Or, if he wishes, he can input modified design changes through an iteractive computer terminal.

There are a number of other respects in which this mode of designer interaction with the computer algorithm is beneficial. First, it often happens in the automated use of the algorithm that oscillation of admissible designs occurs because too large a design improvement has been requested. Such oscillation can often be identified by the designer after only a few iterations and the step size can be reduced to prevent loss of computer time, which can be significant in large scale problems. Conversely, if an estimate quite far from the optimum is chosen to initiate the algorithm, it often happens that the designer chooses far too small a step size. The result is a very small improvement in the design which can be sensed by the designer and improved before excessive computation time is expended.

A second important benefit from designer interaction with the algorithm arises due to the occurrence of local minima and singularities in the analytical formulation of the design problem. The problem of local minima is illustrated by Figure 4. Virtually all optimization methods seek local optima and do not solve the global optimization problem. It is easy for an optimization technique to get hung up at point B and not get to point A, which is the global minima, so the designer must try different starting points to obtain the global solution. This is a very time consuming and indefinite technique with very few analytical aids to the designer. Part of the difficulty here arises because Figure 4 is the wrong display for the designer, in that it does not utilize his knowledge and experience with structures.

A much better approach for the designer is to look at a display such as Figure 3. He can use his experience to restart the optimization algorithm at a meaningful distribution of design variables which may be quite different from the design which resulted from previous calculations. His experience, thus, aids him in starting with different trial designs.

Perhaps even more important than trying various distributions of design variables, the designer can utilize the display of Figure 3 to change the configuration of the structure based on information he accumulates during iterative design and based on his experience. For example, he might try taking member number two out of the structure and optimize based on the modified configuration. Very often, significant gains are made in this manner. Precisely this behavior

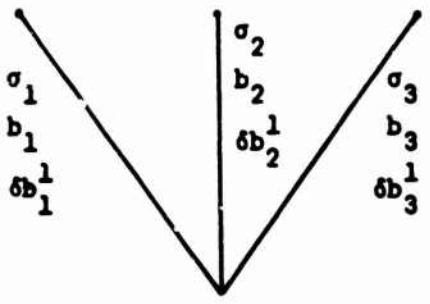


Figure 3. Display of Design Sensitivity Data

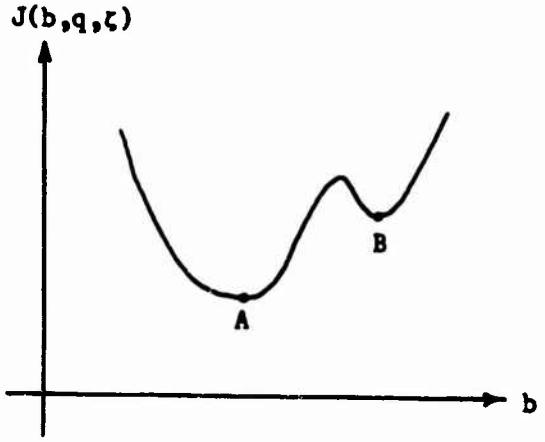


Figure 4. Local Optima

occurs in the three member truss being considered, as will be seen later when this problem is solved in detail.

There are actually compelling mathematical reasons for allowing the designer to make changes in configuration as outlined above. There are no general optimization methods, to date, which will remove a member during iterative design. The reason is that as a member cross section goes toward zero, as is required to remove a member, the equations of structural mechanics and stress constraints become singular. This sort of behavior is typical when the configuration of a system is changed and a different set of equations is required to describe the behavior. At the present time, allowing the designer to make changes in configuration appears to be the most feasible approach, which requires that he play an active role in the iterative optimization algorithm.

Example 1. A Three Member Truss

As an illustrative example of the technique presented above, an elementary optimal design problem will be solved under a number of loading conditions and a variety of constraints. The effect of designer-computer interaction on rate of convergence is examined as well as the effect of changing structural configuration. Results for a wide variety of loading and constraints are compared with solutions of similar problems solved in the literature.

Figure 5 shows the geometry and dimensions of the structure being considered. This structure has been studied by Schmit [30], Sved and Ginos [31] and Corcoran [32]. Three independent loading conditions are applied to the structure. These are as follows: 40K at 45°; 30K at 90°; 20K at 135°. The allowable stress level for members one and three is ± 5 KSI and for member two it is ± 20 KSI. The density of the material is taken as 0.10 pounds/cu. in., and Young's Modulus as 10^4 KSI. Starting from the feasible solution, $b_1 = 8.0$, $b_2 = 2.4$, $b_3 = 3.2$, Schmit [30] arrives at the solution $b_1 = 7.099$, $b_2 = 1.849$, $b_3 = 2.897$, for which $J = 15.986$ pounds. Sved and Ginos [31] have shown that this is only a local minima and by omitting member three, they obtained the solution as $b_1 = 8.5$, $b_2 = 1.5$ with $W = 12.812$ pounds. They have also shown that it is impossible to reach this minimum by an iterative optimization method unless member three is omitted from the calculations by the designer.

In the present work, considerable experimentation was done with this problem. Starting from a feasible point $b_1 = 10$, $b_2 = 5$, $b_3 = 5$, the solution obtained without interaction was $b_1 = 7.064$, $b_2 = 1.971$, $b_3 = 2.835$ and the minimum was $J = 15.97$ pounds. The variation of weight with respect to iteration number is shown by Curve 1, Figure 6. Next, by adjusting the step size in interactive computing, the solution was obtained in only five iterations. This is shown by Curve 2, Figure 6. It was observed that member two never reached its allowable stress level. As a second starting point, the area of member two was initially chosen to bring its stress to the allowable limit. The minimum reached in this case was the same as before (Curve 3, Figure 6). Another solution was obtained by starting for an infeasible point $b_1 = 5.0$, $b_2 = 1.5$, $b_3 = 0.10$. The solution in this case was $b_1 = 6.98$, $b_2 = 2.30$, $b_3 = 2.68$ with $J = 15.97$ (Curve 4, Figure 6).

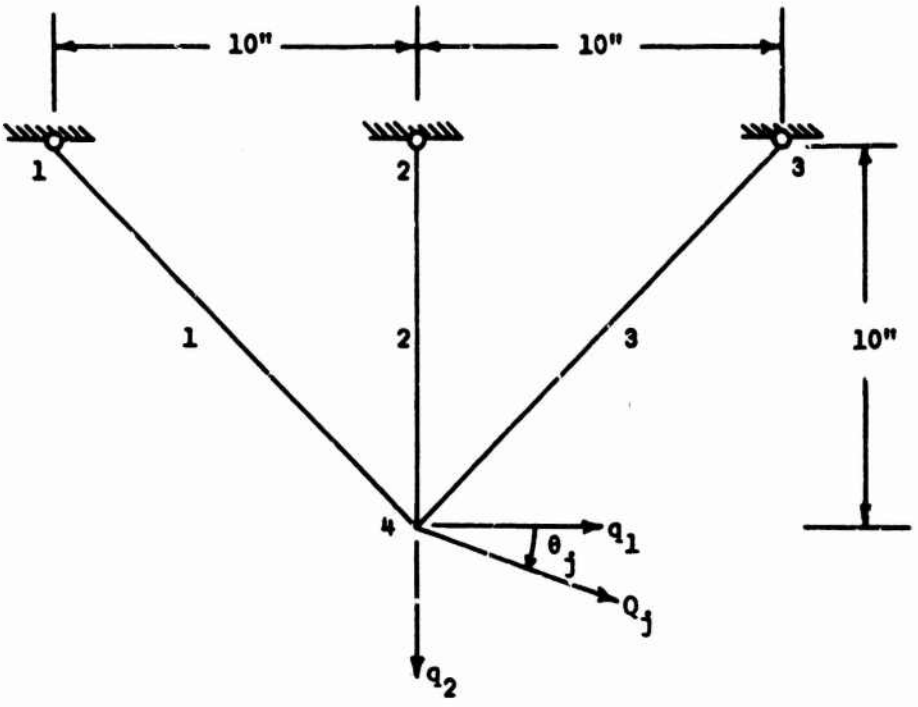


Figure 5 Three-Bar Truss

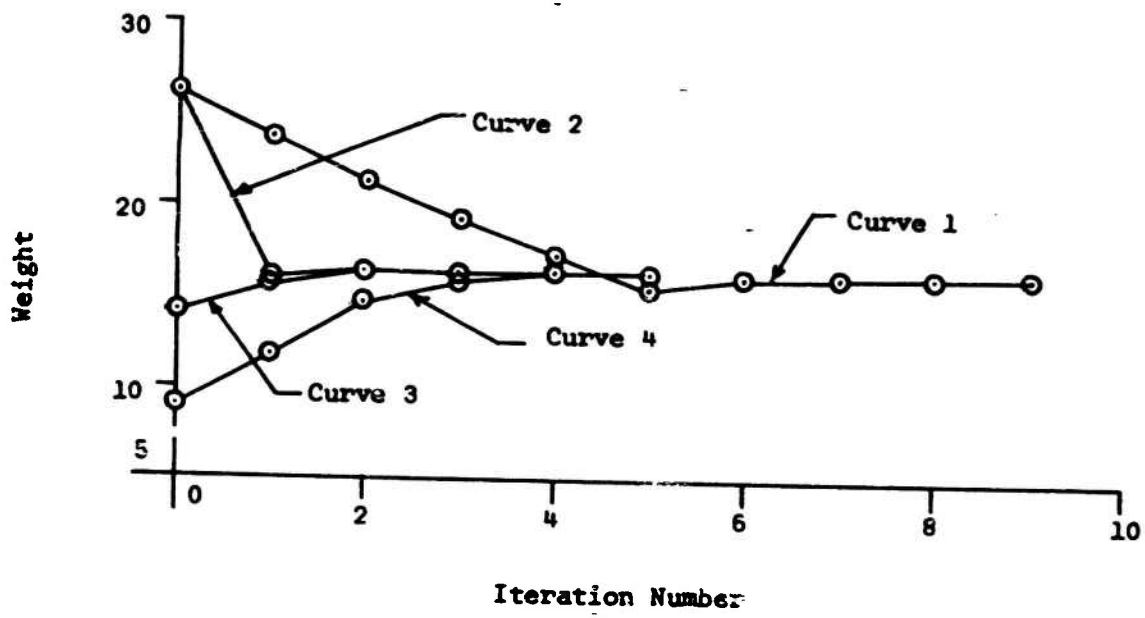


Figure 6. Weight Versus Iteration Curves for Three-Bar Truss (Stress Constraints)

Next, member three was omitted from the structure. Starting from a point $b_1 = 10$, $b_2 = 5$, the solution obtained was $b_1 = 8.0$, $b_2 = 1.5$ with $J = 12.812$ pounds (Curve 1, Figure 7) which is same as reported in [31]. At this point an interesting observation was made. The maximum horizontal and the vertical deflections of node four were as follows: with three bars, $q_1 = 0.689 \times 10^{-2}$ $q_2 = 0.595 \times 10^{-2}$ in.; with two bars, $q_1 = 0.239 \times 10^{-1}$ in., and $q_2 = 0.20 \times 10^{-1}$ in. Thus, although the optimum weight obtained by omitting member three is approximately 24 percent lower than the weight obtained by including member three, the deflections of node four in the former case were approximately four times greater than in the latter.

One might be led to believe that if deflection or frequency constraints were enforced, then the optimum structure might not be statically determinate. To investigate this possibility, displacement as well as buckling and natural frequency constraints were imposed. The deflection limits were taken as, $q_1 = \pm 0.005$ in. and $q_2 = \pm 0.005$, and the lower limit on natural frequency was taken as 3830 cps. With the starting point $b_1 = 10$, $b_2 = 5$, $b_3 = 5$, the solution obtained was $b_1 = 9.18$, $b_2 = 2.16$, $b_3 = 3.85$, and $J = 20.59$ pounds (Curves 2 and 3, Figure 7). When member three was omitted, the starting point was taken as $b_1 = 10$, $b_2 = 10$ (Curve 4, Figure 7) and as $b_1 = 18$, $b_2 = 10$ (Curve 5, Figure 7). The solution obtained in this case was $b_1 = 16.0$, $b_2 = 11.31$, and $J = 33.94$ pounds. Thus, the optimum weight obtained for the statically determinate case is approximately 70 percent higher than the optimum weight obtained for the statically indeterminate case.

It was found that interactive computing yielded convergence more rapidly than was the case in the batch mode. It is expected that even more significant reduction in computing time will occur in large scale problems.

The key point in the solution is that the configuration of the optimum design is not obvious from analytical considerations. A designers experience and insight is required to select candidate configurations and then obtain the optimum design analytically. The global solution in this case must be chosen by comparing relative minima. It may be expected, in structures with greater redundancy, that certain members may be reduced during interactive computation when they are observed to approach their allowable lower limits.

An interesting point, illustrated by Table 1, is that a statically determinate truss is optimum when only stress constraints are imposed. Quite the contrary, when the full range of constraints are imposed, a statically indeterminate truss is optimum.

Example 2. Transmission Tower

Figure 8 shows the geometry and dimensions of the transmission tower to be studied. This problem has been considered by Venkayya and others [7, other references are discussed in this report]. The tower has 25 members, 10 joints, 18 degrees of freedom and is designed for six loading conditions. The structure is indeterminate with a degree of indeterminacy of seven.

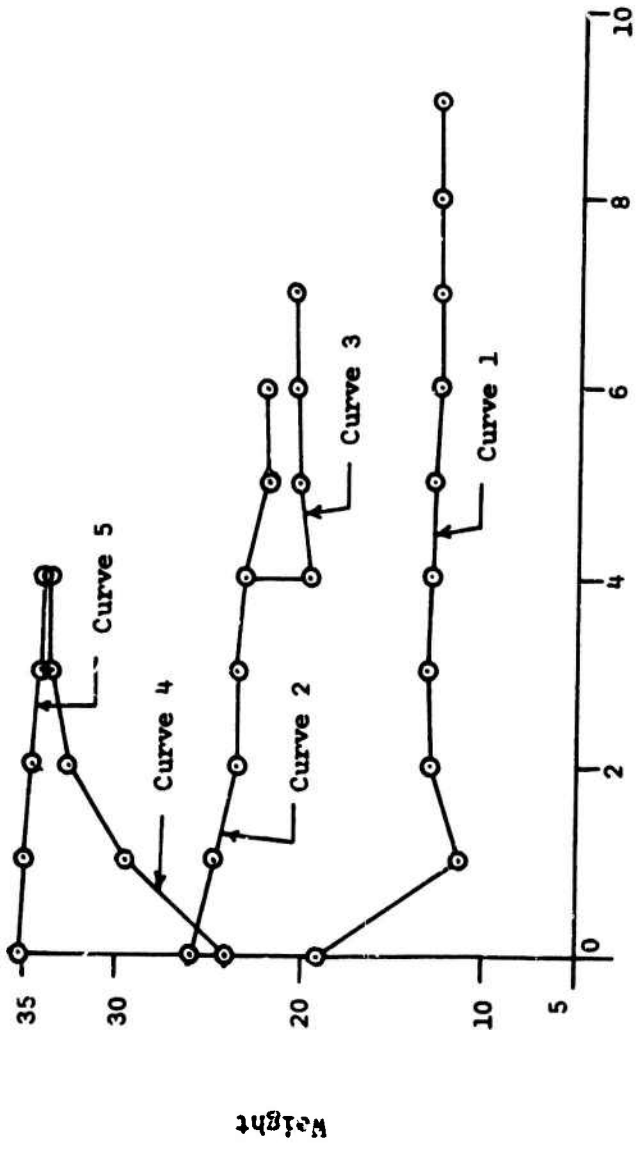


Figure 7. Weight Versus Iteration Curves for Three-Bar Truss (All Constraints)

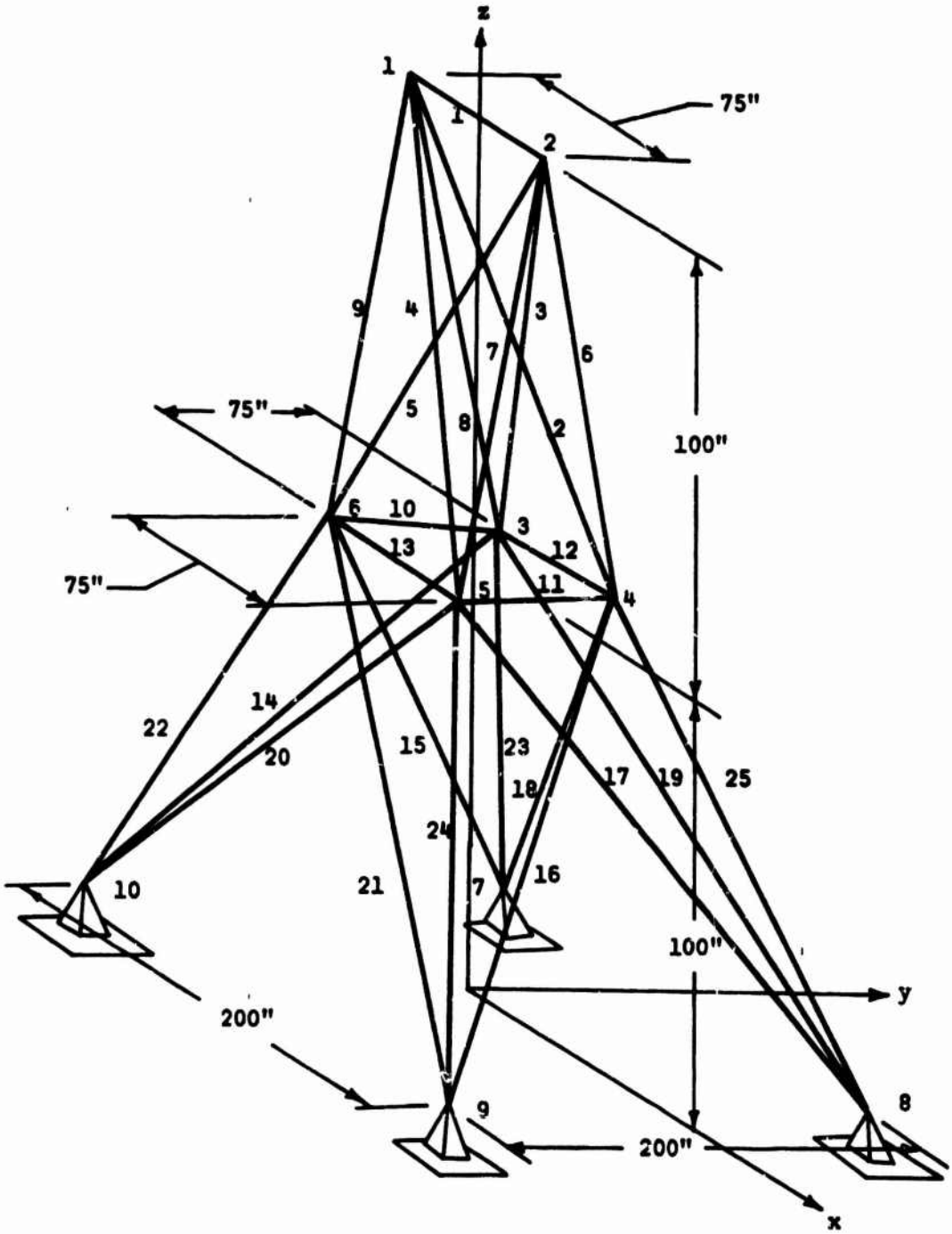


Figure 8 Transmission Tower

Table 1 Optimum Three Member Trusses

El. No.	With Stress Constraints Only		With All Constraints	
	Final Area in Sq. In.		Final Area in Sq. In.	
1	7.064	8.500	9.180	16.000
2	1.971	1.500	2.160	11.310
3	2.835	- -	3.850	- -
Wt. in lbs.	15.970	12.812	20.59	33.94
Max. Defl.	0.00689	0.02390	0.005	0.005

The tower was designed by first imposing only stress constraints, and then by imposing stress, displacement, buckling, and natural frequency constraints. Design information is given in Table 2 and the final results obtained are shown in Table 3. In this table, Columns 1, 2, 3, and 4 give the optimal designs when only stress constraints were considered and Columns 5, 6, 7, and 8 give results when all the constraints were imposed. For results given in Column 1 of Table 3, all the members of tower were included in the computation and the Curve 1 of Figure 9 shows the variation of cost function with the number of iterations. The computations of this case were monitored to determine which cross-sections went to their lower bounds.

One set of members which attained their lower limits of cross-sectional area were numbers 10, 11, 12, and 13. It was observed that these members carried small forces and could be removed without causing collapse of the tower, so they were removed from the tower. The final values of areas of cross section of the resulting structure are given in Column 2 of Table 3. Curve 2 of Figure 9 shows the variation of cost function with respect to the design cycle. The final weight in this case was slightly less than the previous case.

The next member that reached its lower limit was number one. So, it was also removed from the structure. The results of this case are given in Column 3 of Table 3 and Curve 3 of Figure 9. The final weight in this case was 86.94 pounds which is given slightly less than the previous case. Finally, members 14, 15, 16, and 17 were at their lower limits of cross-sectional area. Removal of any of these members, however, would cause collapse of the structure. Members 2 and 5 or 3 and 4 could be removed to make the structure determinate. The results for a statically determinate structure, obtained by removing members 2 and 5, are shown in Column 4 of Table 3. The final weight in this case was 106.07 pounds. It may be noted that this statically determinate structure yielded only a local optimum(Curve 4, Figure 9).

All these tower configurations were also optimized by imposing all constraints, i.e. stress, displacement, buckling, and natural frequency. The results of these cases are given in Columns 5, 6, 7, and 8 of Table 3 and the Curves 1, 2, 3, and 4 of Figure 10, respectively. It can be observed from the results of Table 3 that, for the case in which all constraints were imposed, the optimum weight of the tower increased as more redundant members were removed from the structure.

Table 2 Design Information for Transmission Tower

For each member of the structure, the modulus of elasticity, E_i, the specific weight, ρ_i, the constant, α_i (moment of inertia of ith member, $I_i = \alpha_i b_i^2$) and the stress limits are 10^4 kips/sq. in., 0.10 lbs/cu. in., 1.0 and $\pm$ 40.0 kips/sq. in. respectively. The lower limit on the area of cross section of each member is 0.10 sq. in. for the case with stress constraints only and 0.01 sq. in. for other cases. There is no upper limit on the member sizes. The resonant frequency for the structure is 173.92 cps and the displacement limits are 0.35 in. on all nodes and in all directions. There are six loading conditions and they are as follows (all loads are in kips):

Load Cond.	Node Node	Direction of Load x	y	z	Load Cond.	Node Node	Direction of Load x	y	z
1	1	1.0	10.00	-5.0	2	1	0	10.0	-5.0
	2	0	10.0	-5.0		2	-1.0	10.0	-5.0
	3	0.5	0	0		4	-0.5	0	0
	6	0.5	0	0		5	-0.5	0	0
3	1	1.0	-10.0	-5.0	4	1	0	-10.0	-5.0
	2	0	-10.0	-5.0		2	-1.0	-10.0	-5.0
	3	0.5	0	0		4	-0.5	0	0
	6	0.5	0	0		5	-0.5	0	0
5	1	0	20.0	-5.0	6	1	0	-20.0	-5.0
	2	0	-20.0	-5.0		2	0	20.0	-5.0

Table 3 Optimum Transmission Towers

EL. NO.	With Stress Constraints only — Final Areas in sq. in.				With All Constraints — Final Areas in sq. in.			
	1	2	3	4	5	6	7	8
1	0.100	0.100	- -	- -	0.010	0.010	- -	- -
2	0.376	0.377	0.346	- -	2.092	2.339	2.393	- -
3	0.376	0.377	0.346	0.100	2.075	2.386	2.404	0.548
4	0.376	0.377	0.346	0.100	2.095	2.339	2.393	0.548
5	0.376	0.377	0.346	- -	2.083	2.385	2.404	- -
6	0.471	0.470	0.494	0.779	2.357	2.085	2.076	7.132
7	0.471	0.470	0.494	0.779	2.354	2.084	2.076	6.857
8	0.471	0.470	0.494	0.779	2.350	2.113	2.083	6.895
9	0.471	0.470	0.494	0.779	2.335	2.112	2.082	7.101
10	0.100	- -	- -	- -	0.035	- -	- -	- -
11	0.100	- -	- -	- -	0.035	- -	- -	- -
12	0.100	- -	- -	- -	0.087	- -	- -	- -
13	0.100	- -	- -	- -	0.084	- -	- -	- -
14	0.100	0.100	0.100	0.165	1.113	1.114	1.139	1.785
15	0.100	0.100	0.100	0.165	1.113	1.114	1.139	1.735
16	0.100	0.100	0.100	0.165	1.112	1.117	1.146	1.727
17	0.100	0.100	0.100	0.165	1.112	1.117	1.146	1.798
18	0.277	0.279	0.292	0.413	2.056	2.047	2.027	4.317
19	0.277	0.279	0.292	0.413	2.058	2.034	2.022	4.390
20	0.277	0.279	0.292	0.413	2.046	2.047	2.027	4.400
21	0.277	0.279	0.292	0.413	2.058	2.034	2.022	4.328
22	0.380	0.374	0.363	0.547	2.822	2.878	2.896	5.655
23	0.380	0.374	0.363	0.547	2.808	2.878	2.886	5.730
24	0.380	0.374	0.363	0.547	2.803	2.926	2.895	5.743
25	0.380	0.374	0.363	0.547	2.785	2.926	2.895	5.648
Wt. in lbs.	91.13	87.90	86.94	106.97	590.32	596.64	597.82	1060.6
Max. Defl. in in.	2.288	2.305	2.311	3.489	0.350	0.350	0.350	0.350

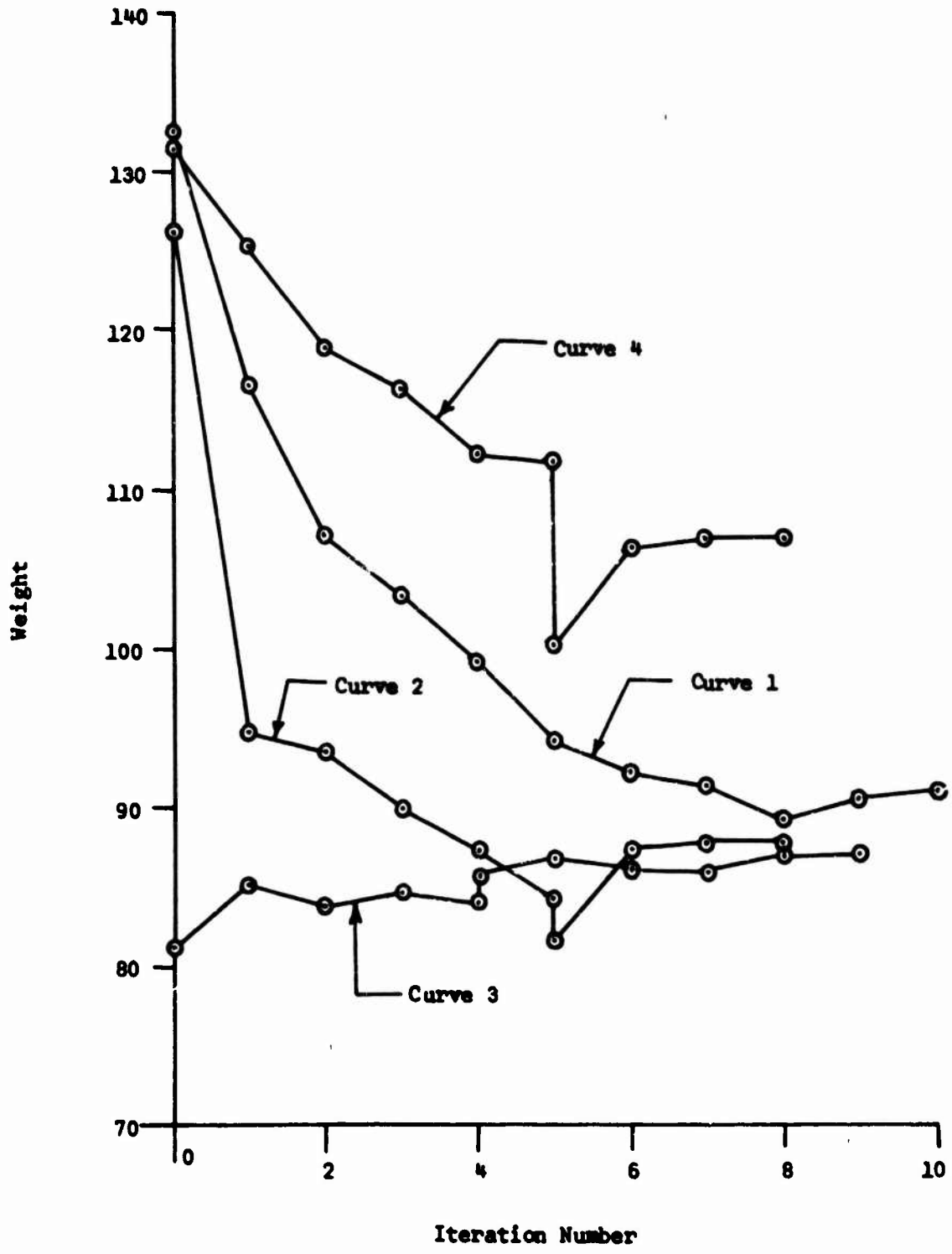


Figure 9. Weight Versus Iteration Curves for Transmission Tower (Stress Constraints)

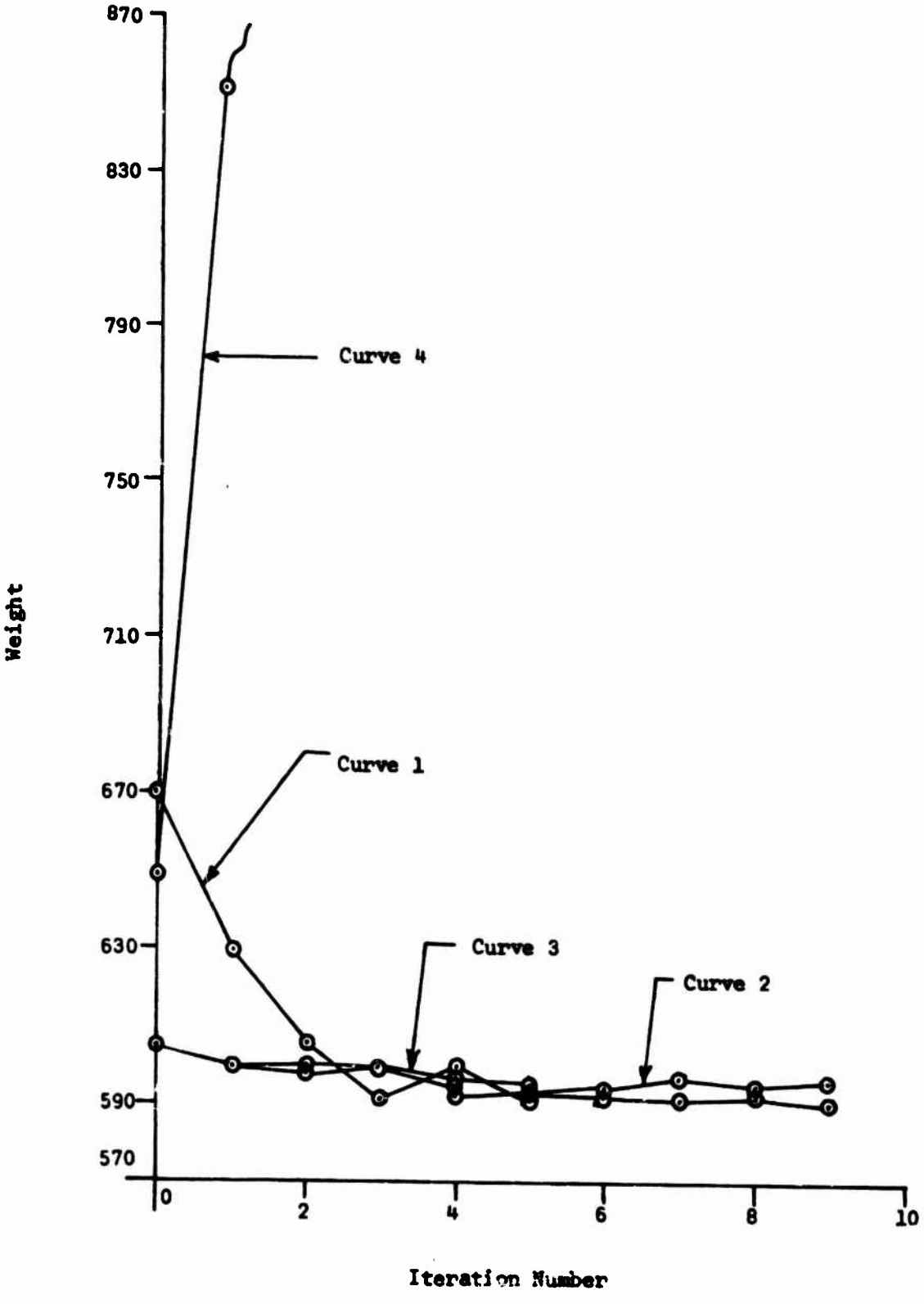


Figure 10. Weight Versus Iteration Curves for Transmission Tower (All Constraints)

Conclusions

Use of the steepest descent technique in the interactive mode, as developed in this paper has proved to be very effective from two points of view. First, computing times are considerably shorter than had been experienced when the same problems were solved in the batch mode. Second, and probably more significant, interactive computing allows the designer to alter the structural configuration in a systematic way to seek the global optimum design. This is not to say that a mathematically precise method of obtaining a global optimum has been found, for no such method is known to the authors. It appears, however, that the technique presented here makes strong use of the designers knowledge and intuition and gives him a tool with which to seek a global optimum in an organized way.

The results presented for the two examples solved in this paper are of interest in their own right. For the case when only stress constraints are imposed, results of Table 1 indicate that minimum weight designs for trusses with multiple loading may be statically determinate. However, the results of the second example given in Columns 1, 2, 3, and 4 of Table 3 indicate that all statically determinate trusses may not be lighter than the indeterminate trusses.

For the case when all constraints are imposed, results of Table 1 and 3 show that statically indeterminate trusses are lighter than the determinate trusses. Therefore, one is lead to believe that, in this case, redundancy is advantageous.

REFERENCES

[1] Douty, R., and Shore, S., "Technique for Interactive Computer Graphics in Design", Journal of the Structural Division, ASCE, Vol. 97, No. ST1, January, 1971, pp. 273-288.

[2] Moses, F., "Optimum Structural Design Using Linear Programming", Journal of the Structural Division, Proc. ASCE, Vol. 90, No. ST6, December, 1964, pp. 89-104.

[3] Romstad, K. M. and Wang, C. K., "Optimum Design of Framed Structures", Journal of the Structural Division, Proc. ASCE, Vol. 94, No ST12, December, 1968, pp. 2817-2845.

[4] Grierson, D. E. and Cohn, M. Z., "A General Formulation of the Optimal Frame Problem", Journal of Applied Mechanics, Trans. ASME, Vol. 37, No. 2, June, 1970, pp. 356-360.

[5] Kelley, J. E., "The Cutting-Plane Method for Solving Convex Programs", SIAM Journal on Applied Math., Vol. 8, No. 4, December, 1960, pp. 703-712.

[6] Johnson, D. and Brotton, D. M., "Optimum Elastic Design of Redundant Trusses", Journal of the Structural Division, Proc. ASCE, Vol. 95, No. ST12, December, 1969, pp. 2589-2610.

[7] Venkayya, V. B., Khot, N. S. and Reddy, V. S., "Energy Distribution in an Optimum Structural Design", Technical Report AFFDL-TR-68-158, Wright-Patterson Air Force Base, Ohio 45433, March, 1969.

[8] Fiacco, A. V. and McCormick, G. P., Nonlinear Programming: Sequential Unconstrained Minimization Techniques, John Wiley, New York, 1968.

[9] Kavlie, D. and Moe, J., "Application of Nonlinear Programming to Optimum Grillage Design with Nonconvex Sets of Variables", International Journal for Numerical Methods in Engineering, Vol. 1, No. 4, 1969, pp. 351-378.

[10] Kavlie, D. and Moe, J., "Autmated Design of Frame Structures," Journal of Structural Division, Proc. ASCE, Vol. 97, No. ST1, January, 1971, pp. 33-62.

[11] LaPay, W. S. and Goble, G. G., "Optimum Design of Trusses for Ultimate Loads", Journal of the Structural Division, Proc. ASCE, Vol. 97, No. ST1, January, 1971, pp. 157-174.

[12] Schmit, L. A. and Fox, R. L., "An Integrated Approach to Structural Synthesis and Analysis", AIAA Journal, Vol. 3, No. 6, June, 1965, pp. 1104-1112.

[13] Schmit, L. A., Kicher, T. P. and Morrow, W. M., "Structural Synthesis Capability for Integrally Stiffened Waffle Plates", AIAA Journal, Vol. 1, No. 12, December, 1963, p. 2820.

[14] Gellatly, R. A. and Gallagher, R. H., "A Procedure for Automated Minimum Weight Structural Design, Part I: Theoretical Basis; Part II: Applications", Aeronautical Quarterly, Vol. 17, 1966, pp. 216-230 and pp. 332-342.

[15] Felton, L. P. and Hofmeister, L. D., "Optimized Components in Truss Synthesis", AIAA Journal, Vol. 6, No. 12, December, 1968, pp. 2434-2436.

[16] Ridha, R. A. and Wright, R. N., "Minimum Cost Design of Frames," Journal of the Structural Division, Proc. ASCE, Vol. 93, No. ST4, August, 1967, pp. 165-183.

[17] Fox, R. L. and Kapoor, M. P., "Structural Optimization in Dynamics Response Regime: A Computational Approach", AIAA Journal, Vol. 8, No. 10, October, 1970, pp. 1798-1804.

[18] Pope, G. G., "The Design of Optimum Structures of Specified Basic Configuration", International Journal of Mechanical Sciences, Vol. 10, October, 1968, pp. 251-263.

[19] Moses, F. and Onoda, S., "Minimum Weight Design of Structures With Application to Elastic Grillages", International Journal for Numerical Methods in Engineering, Vol. 1, No. 4, 1969, pp. 311-331.

[20] Best, G., "A Method of Structural Weight Minimization Suitable for High Speed Digital Computers", AIAA Journal, Vol. 1, No. 2, February, 1963, pp. 478-479.

[21] Rosen, J. B., "The Gradient Projection Method for Nonlinear Programming Part I - Linear Constraints", SIAM Journal on Applied Math., Vol. 8, No. 1, March, 1960, pp. 181-217; Part II - Nonlinear Constraints, Vol. 9, No. 4, December, 1961, pp. 514-532.

[22] Brown, D. M. and Ang, A. H. S., "Structural Optimization by Nonlinear Programming", Journal of the Structural Division, Proc. ASCE, Vol. 92, No. ST6, December, 1966, pp. 319-340 and Vol. 93, No. ST5, October, 1967, pp. 618-619.

[23] Seaburg, P. A. and Salmon, C. G., "Minimum Weight Design of Light Gage Steel Members", Journal of the Structural Division, Proc. ASCE, Vol. 97, No. ST1, January, 1971, pp. 203-222.

[24] Haug, Jr., E. J., "Optimal Design of Mechanical Systems", Unpublished class notes on Optimization of Structural Systems, University of Iowa, Iowa City, Iowa 52240, 1970-71.

[25] Haug, Jr., E. J., Pan, K. C. and Streeter, T. D., "A Computational Method for Optimal Structural Design: Piecewise Uniform Structures", to be published in the International Journal for Numerical Methods in Engineering, 1971.

[26] Bartel, D. L. and Rim, K., "Optimum Design of Spatial Structures", Project Themis Report, No. 12, University of Iowa, Iowa City, Iowa 52240, August, 1969.

[27] Bates, W. F., Jr., and Cox, W. M., "Structural Analysis by Computer Graphics", Journal of the Structural Division, ASCE, Vol. 95, No. ST11, November, 1969, pp. 2433-2448.

[28] Schiffman, R. L., and Stein, J. R., "Interactive Graphics: Settlement and Slope Stability", Journal of the Structural Division, ASCE, Vol. 97, No. ST1, January, 1971, pp. 253-272.

[29] Arora, J. S., "Optimal Design of Elastic Structures Under Multiple Constraint Conditions", Ph. D. Dissertation, Department of Mechanics and Hydraulics, University of Iowa, Iowa City, Iowa, August, 1971.

[30] Schmit, L. A., "Structural Design by Systematic Synthesis", Second Conference on Electronic Computation, Structural Division of ASCE, September, 1960, pp. 105-132.

[31] Sved, G. and Ginos, Z., "Structural Optimization Under Multiple Loading", International Journal of Mechanical Sciences, Vol. 10, 1968, pp. 803-805.

[32] Corcoran, P., "Configurational Optimization of Structures", International Journal of Mechanical Sciences, Vol. 12, 1970, pp. 459-462.

[33] Young, Jr., W. J. and Christiansen, H. N., "Synthesis of a Space Truss Based on Dynamic Criteria", Journal of the Structural Division, Proc. ASCE, Vol. 92, No. ST6, December, 1966, pp. 425-442.

[34] Reinschmidt, K. F., Cornell, C. A. and Brotchie, J. F., "Iterative Design and Structural Optimization", Journal of the Structural Division, Proc. ASCE, Vol. 92, No. ST6, December, 1966, pp. 281-318.

[35] Dorn, W. S., Gomory, R. E. and Greenberg, H. J., "Automatic Design of Optimal Structures", Journal de Mecanique, Vol. 3, No. 1, March, 1964, pp. 25-53.

[36] Dobbs, M. W. and Felton, L. P., "Optimization of Truss Geometry", Journal of the Structural Division, Proc. ASCE, Vol. 95, No. ST10, October, 1969, pp. 2105-2118.

[37] Rubin, C. P., "Minimum Weight Design of Complex Structures Subject to Frequency Constraints", AIAA Journal, Vol. 8, No. 5, May, 1970, pp. 923-927.

[38] Zarchamee, M. S., "Optimum Frequency Structures", AIAA Journal, Vol. 6, No. 4, April, 1968, pp. 749-750.

DESIGN OF OPTIMUM STRUCTURES FOR DYNAMIC LOADS

V. B. Venkayya**
N. S. Khot*
V. A. Tischler**
R. F. Taylor***

Air Force Flight Dynamics Laboratory
Wright-Patterson Air Force Base, Ohio 45433

An optimization method for the minimum weight design of structures subjected to dynamic loads is presented. The method is called designing in the dynamic mode. The dynamic mode may be a single natural mode of the structure or a linear combination of a set of natural modes depending on the spatial distribution and dynamic characteristics of the forcing function. An optimality criterion for minimum weight structures in the dynamic mode is derived. It is based on the strain energy and kinetic energy of the individual elements. A recursion relation for attaining the optimality criterion is also derived. The natural frequencies are obtained by using the Sturm Sequence property in conjunction with a bisection procedure. The normal modes are determined by inverse iteration. A step by step procedure for the optimum design under dynamic response constraints is presented. It includes a method for determining the modes that predominate in the dynamic response. Some examples of design are presented to illustrate the method.

1. INTRODUCTION

Optimization and the automated design of structures is the subject of intense study in recent years. The result of these studies is the emergence of a number of successful programs for the design of practical structures. In 1968, a combined approach based on an optimality criterion and a numerical search was presented for the optimal design of structures subjected to static loads. (Reference 1). This method demonstrated the feasibility of optimizing structures of several hundred degrees of freedom and design variables. Design conditions included multiple loading conditions and constraints on stresses, displacements and sizes of the elements. A more detailed study of the convergence characteristics of this method and an extension to frame and plate elements were presented in Reference 2. This reference also contained a sparse matrix scheme for the solution and response gradient calculations. Subsequently an alternate method for static load, also based on an optimality criterion, was developed by Gellatly and Berke in Reference 3. A number of bar and plate structures were successfully designed by this approach. Dwyer, Emerton and Ojalvo developed a modified stress ratio method coupled with a numerical search procedure for the automated design of realistic airframe components in Reference 4. Their effort included the development of a large scale computer program for the optimum design of structures subjected to static loads.

---

* Aerospace Engineers, Structural Synthesis Group
** Mathematician, Structural Synthesis Group
*** Aerospace Engineer, Aeroelasticity Group

The basis for the foregoing developments was the vast body of literature available in the area of both static and dynamic structural optimization. The excellent papers by Maxwell were the earliest to consider the problems of structural optimization (Reference 5). He discussed the fully stressed design of statically determinate and indeterminate bar structures. Later, several investigators examined the limitations of a fully stressed design when applied to different configurations and design conditions. The celebrated paper by Michell dealt with the development of an optimum structural configuration for a given loading (Reference 6). This configuration consisted of an orthogonal network of tension and compression members. Later Hemp (Reference 7), Cox (Reference 8) and Chan (Reference 9) applied the concept of Michell structures to many problems of structural optimization. The more recent papers by Prager and his associates dealt extensively with the derivation of optimality criteria for a variety of design conditions (References 10, 11 and 12). These papers firmly established a place for the optimality criteria approaches in the design of optimum structures. In the same vein, Turner (Reference 13), Taylor (Reference 14), Masur (Reference 15), McIntosh and Eastep (Reference 16) and others investigated specific problems of structural optimization under dynamic and stability constraints. Barnett and Hermen formulated an interesting optimality criterion based on virtual work for displacement contrained problems (Reference 17). Melosh presented an extensive discussion of the validity of the fully stressed design for the design of indeterminate structures (Reference 18). Young and Christiansen used the fully stressed design criterion for designing in the lowest normal mode (Reference 19).

A parallel but a very significant development in the area of structural optimization was initiated by Schmit and Klein (References 20 and 21). Their formulation involved the application of numerical search methods to structural optimization problems. This approach provided a better understanding of structural optimization under different design conditions. Gellatly and Gallagher (Reference 22) and Gellatly (Reference 23) have further developed these ideas and applied them to the design of prectical aerospace structures. Fox and Kapoor (Reference 24), Rubin (Reference 25), McCart, Haug and Streeter (Reference 26) have successfully applied search procedures to the design of structures under dynamic load conditions. The recent dissertations of Salinas (Reference 27), Loomis (Reference 28) and Arora (Reference 29) are excellent contributions to the further understanding of structural optimization.

The list of references cited in the foregoing discussion is at best incomplete. A more detailed survey of contributions to structural optimization can be found in papers by Wasiutynski and Brandt (Reference 30), Sheu and Prager (Reference 31), and Pope and Schmit (Reference 32).

The purpose of this paper is to present an efficient method based on an optimality criterion and a numerical search for the design of structures subjected to dynamic loads. In this paper the effect of damping is not considered, but the authors hope to discuss this problem at a later date.

2. ENERGIES OF MOTION AND EQUATIONS OF DYNAMIC ANALYSIS

The interest of the present paper is primarily in built-up structures and the continuum is approximated by a discretized finite element model. The selection of the displacement method of finite element analysis is prompted by the intent to automate the analysis block. However, the basic principles derived in this paper are generally valid for a continuum as well as other discretized models. In an

iteretive approach a large part of the effort is expended in the repeated analysis of the structure. The basic equations of analysis are presented here as a ready reference for the development that follows. For details the reader is referred to References 33 and 34.

In a finite element scheme the force displacement relations of the individual elements are derived by an energy formulation with the assumption of exact or approximate displacement functions. The energy expressions for the i^{th} element in motion may be written as

$$u_i = \tfrac{1}{2}\int_{V_i} \sigma_i^t \, e_i \, dV_i \tag{1}$$

$$\tau_i = \tfrac{1}{2}\int_{V_i} \gamma_i \, \dot{w}_i^t \, \dot{w}_i \, dV_i \tag{2}$$

where u_i is the strain energy and τ_i the kinetic energy of the element, σ_i and e_i are the stress and strain matrices, respectively, w_i is the displacement vector, and γ_i is the mass density. The dot ($\dot{}$) represents a time derivative throughout this development.

The displacement vector may be represented by a set of discrete generelized coordinates of the element in the following form,

$$w_i = \phi_i \, v_i \tag{3}$$

where v_i is the vector of generalized coordinates for the element and ϕ_i is a rectangular matrix whose elements are functions of the spatial coordinates. The strain in the i^{th} element may be written as

$$e_i = \phi_i' \, v_i \tag{4}$$

The prime in the above equation indicates a derivative with respect to spatial coordinates. For an element made of a linearly elastic material, stresses and strains are related by

$$\sigma_i = G_i \, e_i \tag{5}$$

Substituting Equations 3, 4 and 5 in 1 and 2, the expressions for the strain and kinetic energies of the element are written as

$$u_i = \tfrac{1}{2}\int_{V_i} v_i^t \, \phi'^t \, G_i \, \phi_i' \, v_i \, dV_i \tag{6}$$

$$\tau_i = \tfrac{1}{2}\int_{V_i} \gamma_i \, \dot{v}_i^t \, \phi_i^t \, \phi_i \, \dot{v}_i \, dV_i \tag{7}$$

Since the generalized coordinates vector $\mathbf{v}_i$ is independent of the spatial coordinates, it can be taken out of the integration and the energy expressions may be written as

$$u_i = \tfrac{1}{2}\mathbf{v}_i^t \mathbf{k}_i \mathbf{v}_i \tag{8}$$

$$\tau_i = \tfrac{1}{2}\mathbf{v}_i^t \mathbf{m}_i \mathbf{v}_i \tag{9}$$

Where $\mathbf{k}_i$ and $\mathbf{m}_i$ are the element generalized stiffness and mass matrices respectively, and they are given by

$$\mathbf{k}_i = \int_{V_i} \boldsymbol{\phi}_i'^t \mathbf{G}_i \boldsymbol{\phi}_i' \, dV_i \tag{10}$$

$$\mathbf{m}_i = \int_{V_i} \gamma_i \boldsymbol{\phi}_i^t \boldsymbol{\phi}_i \, dV_i \tag{11}$$

The expressions for the total energies of the structure are obtained by summing the component energies,

$$U = \tfrac{1}{2}\sum_{i=1}^{m} \mathbf{v}_i^t \mathbf{k}_i \mathbf{v}_i \tag{12}$$

$$T = \tfrac{1}{2}\sum_{i=1}^{m} \mathbf{v}_i^t \mathbf{m}_i \mathbf{v}_i \tag{13}$$

where the scalar m represents the total number of elements in the structure.

The generalized coordinates of the elements and the structure are related by

$$\mathbf{v}_i = \mathbf{a}_i \mathbf{r} \tag{14}$$

where $\mathbf{r}$ is the vector of system generalized coordinates, and $\mathbf{a}_i$ is the compatibility matrix. The invariance of the energy with a coordinate transformation permits writing the energy expressions in the following form,

$$U = \tfrac{1}{2}\mathbf{r}^t \mathbf{K} \mathbf{r} \tag{15}$$

$$T = \tfrac{1}{2}\dot{\mathbf{r}}^t \mathbf{M} \dot{\mathbf{r}} \tag{16}$$

where $\mathbf{K}$ and $\mathbf{M}$ are the generalized stiffness and mass matrices respectively of the system and are given by

$$\mathbf{K} = \sum_{i=1}^{m} \mathbf{a}_i^t \mathbf{k}_i \mathbf{a}_i \tag{17}$$

$$\mathbf{M} = \sum_{i=1}^{m} \mathbf{a}_i^t \mathbf{m}_i \mathbf{a}_i \tag{18}$$

The j^{th} Lagrange's equation of the system is given by

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{r}_j} - \frac{\partial T}{\partial r_j} + \frac{\partial U}{\partial r_j} = R_j \tag{19}$$

Substituting Equations 15 and 16 into 19 gives the equations of motion of the system in the following form

$$\mathbf{M}\ddot{\mathbf{r}} + \mathbf{K}\mathbf{r} = \mathbf{R} \tag{20}$$

where **R** is the generalized force matrix which includes the externally applied dynamic and static forces. These forces may be concentrated or distributed. In the case of distributed and concentrated forces applied at points and directions other than those of the generalized coordinates, **R** can be derived by the principle of virtual work. **R** can also include damping and other internal elastic forces.

In deriving Equation 20 only structural mass is considered. However, any non-structural mass such as fuel tanks in an airplane wing can be included in the generalized mass matrix by superposition. The magnitude of the nonstructural mass is added to the diagonal elements of **M** corresponding to their points of application. If they are lumped masses without specific geometric dimensions, only the translational degrees of freedom will be effected. When these masses are not located at the points of the generalized coordinates, then the equivalent effect on the generalized mass can be derived by D'Alembert's principle and the principle of virtual work

$$\mathbf{M} = \mathbf{M}_{\text{structural}} + \mathbf{M}_{\text{nonstructural}} \tag{21}$$

When the system is at rest and the rate of loading is sufficiently slow, Equation 20 reduces to the familiar force displacement relation of the static case

$$\mathbf{R} = \mathbf{K}\mathbf{r} \tag{22}$$

In such a case the internal forces and stresses of the elements can be determined by

$$s_i = k_i v_i \tag{23}$$

$$\sigma_i = \beta_i s_i \tag{24}$$

In the dynamic case the solution of Equation 20 consists of a homogeneous part and a particular solution. The homogeneous part corresponds to free vibration and is obtained by solving

$$\omega^2 \mathbf{M}\mathbf{r} = \mathbf{K}\mathbf{r} \tag{25}$$

where ω represents the frequency of vibration.

Equation 25 represents a standard eigenvalue problem and its solution amounts to obtaining eigenvalues, ω_κ (κ=1,2,...,n) and their corresponding eigenvectors. The matrices **M** and **K** possess symmetry and bandedness properties. The details of a method for obtaining the eigenvalues and eigenvectors of a banded symmetric matrix are presented in Section 5.

There are several approximate methods in the literature for obtaining the particular solution of the matrix Equation 20. One of these is the normal mode method which is discussed in Section 7.

3. ENERGY CRITERION FOR AN OPTIMUM DESIGN IN FREE VIBRATION

The response of a structure is largely governed by its dynamic characteristics and the nature of the forcing function. The natural frequencies and modes of free vibration are the primary dynamic charecteristics of the structure. A design method should have the ability to manipulate the dynamic characteristics to obtain an optimum response for the given forcing function. An optimality criterion for a minimum weight structure in free vibration is derived here in terms of discrete variables. This derivation is similar to the optimality criterion established in References 1 and 2 for static loading conditions.

The dynamic equation governing the free vibration of the system is given by the integral

$$I = \int_{t_1}^{t_2} (T - U)\, dt = \int_{t_1}^{t_2} L\, dt \tag{26}$$

where L is the Lagrangian. The expressions for the strain energy and kinetic energy of the system, U and T are given by Equations 12 and 13. From Hamilton's principle the natural mode of vibration is the one for which the integral I attains a stationary value in the time interval t_1 to t_2.

Suppose A and A' are two designs in the neighborhood of the minimum weight design. The weights of the structure corresponding to the two designs are proportional to W and W' which are defined as

$$W = \sum_{i=1}^{m} A_i\, l_i \tag{27}$$

$$W' = \sum_{i=1}^{m} A'_i\, l_i \tag{28}$$

For one dimensional elements A_i and l_i are areas and lengths respectively. For plate elements these parameters are defined in Reference 2.

The difference in the strain energy and kinetic energy of the i^{th} element in free vibration in a finite time interval is given by

$$\mu_i = \tfrac{1}{2}[\mathbf{s}_i^t\, \mathbf{k}_i\, \mathbf{s}_i - \omega^2 \mathbf{s}_i^t\, \mathbf{m}_i\, \mathbf{s}_i] \tag{29}$$

$$\mu'_i = \tfrac{1}{2}[\mathbf{s}'^t_i\, \mathbf{k}'_i\, \mathbf{s}'_i - \omega'^2\, \mathbf{s}'^t_i\, \mathbf{m}'_i\, \mathbf{s}'_i] \tag{30}$$

where the vector $\mathbf{s}_i$ is the displacement of the i^{th} element when the structure is in the natural mode of vibration. Here prime refers to the second design. If ρ_i and ρ'_i are defined as μ_i and μ'_i, per unit volume (density), then

$$V = \sum_{i=1}^{m} \mu_i = \sum_{i=1}^{m} A_i \, l_i \, \rho_i \tag{31}$$

$$V' = \sum_{i=1}^{m} \mu'_i = \sum_{i=1}^{m} A'_i \, l_i \, \rho'_i \tag{32}$$

where V and V' are the difference in the strain energy and the kinetic energy of the total system in the two cases.

Since the two designs are assumed to be in the neighborhood of each other, a limiting condition can be written in the following form

$$\lim_{\Delta A \to 0} \left[\sum_{i=1}^{m} A_i \, l_i \, \rho_i - \sum_{i=1}^{m} A'_i \, l_i \, \rho'_i \right] = 0 \tag{33}$$

Let **S** and **S'** be the respective eigenmodes of the two designs. Since the geometric configuration of the structure is the same in both cases, eigenvector **S** of the first design is kinematically admissible for the second design and vice versa. If the second design is forced to vibrate in the mode **S** with eigenvalue ω, then from Rayleigh's principle one can write

$$\sum_{i=1}^{m} A'_i \, l_i \, \rho_i \geq \sum_{i=1}^{m} A'_i \, l_i \, \rho'_i \tag{34}$$

The quantity ρ_i on the left side of the inequality is valid for the following reasons. In the case of bar structures, ρ_i depends on the mode shape only and not on the sizes of the elements. In the case of beam elements, ρ_i is independent of the areas of the elements, provided the radius of gyration of each element in the first design is the same as that in the second design. It is not necessary that all elements have the same radius of gyration. Since the two designs are assumed to be in the neighborhood of each other, this assumption is not unreasonable.

Invoking the limiting condition stated in Equation 33 , the inequality 34 can be written as

$$\sum_{i=1}^{m} (A'_i - A_i) \, l_i \, \rho_i \geq 0 \tag{35}$$

If the first design has constant ρ_i for all its elements, then the weight of the second design, in which this condition is not satisfied, is greater than the first design, i.e.,

$$\sum_{i=1}^{m} A'_i \, l_i \geq \sum_{i=1}^{m} A_i \, l_i \tag{36}$$

or

$$W \geq W \tag{37}$$

Now the optimality criterion is stated as follows:

A structure with a given natural frequency will be of minimum weight design when the difference in strain energy density and kinetic energy density is a constant for all its elements while vibrating in its natural mode. For the static case, the above optimality criterion is identical to the one stated in References 1 and 2.

An iterative algorithm that yields a structure satisfying the optimality criterion is derived in the next section.

4. RECURSION RELATION BASED ON OPTIMALITY CRITERION

The optimality criterion stated in the previous section can be attained by iteration only. This requires a recursion relation for iteration and an assurance that iteration using this relation converges to a design satisfying the optimality criterion. The recursion relation derived in this section is similar to the one given in References 1 and 2 except for the modification necessary to account for the kinetic energy.

The volume of the i^{th} element may be written as

$$V_i = \alpha_i \Lambda l_i \tag{38}$$

where the quantity l_i is defined as

$$l_i = \frac{V_i}{\Lambda \alpha_i} \tag{39}$$

α_i is the relative design variable, and Λ is an arbitrary normalization parameter for the design vector. In the case of bar elements l_i is simply the length of the element, and $\Lambda \alpha_i$ is the area of the element. In the case of elements in bending and torsion, appropriate variable definitions are given in Section 6 of Reference 2. When the system is vibrating in the normal mode, the difference in strain energy and kinetic energy of the i^{th} element is a finite time interval is given by

$$\mu_i = \frac{1}{2\Lambda} [s_i^t k_i s_i' - \omega_p^2 s_i^t m_i s_i'] \tag{40}$$

where ω_p is the frequency of vibration in the p^{th} mode, and s_i' is given by

$$s_i' = \Lambda s_i \tag{41}$$

If the quantity u_i per unit volume is assumed to be constant for each element, the following relation can be written

$$\Lambda^2 = c^2 \frac{\mu_i'}{V_i'} \tag{42}$$

where c is the constant of proportionality, and μ_i' and V_i' are given by

$$\mu_i' = \tfrac{1}{2}[s_i^t k_i s_i' - \omega_p^2 s_i^t m_i s_i'] \tag{43}$$

$$V_i' = \alpha_i l_i \tag{44}$$

Multiplying both sides of Equation 42 by α_i^2 and taking the square root yields

$$\alpha_i \Lambda = c \alpha_i \frac{\mu_i'}{V_i'} \tag{45}$$

where $\alpha_i \Lambda$ is the i^{th} design variable which is expressed as a function of α_i, the relative variable. The form of Equation 45 suggests the following recursion relation for determining the design variable in each cycle.

$$(\alpha_i \Lambda)_{\nu+1} = c(\alpha_i)_\nu \left.\frac{\mu_i'}{V_i'}\right|_\nu \tag{46}$$

where ν refers to the cycle of iteration.

The procedure for using Equation 46 is as follows:

1. With an assumed relative design vector (such as $\alpha_1 = \alpha_2 = \ldots = \alpha_m = 1.0$) the desired normal mode and the corresponding eigenvalue ω_p are determined.

2. Then the relative strain energy and kinetic energy of each element in this mode are determined.

3. Each design variable for the next iteration is determined by substituting the relative energies in Equation 46 .

4. The design vector is now normalized with respect to the largest variable. This normalization eliminates the need for determining the constant of proportionality.

5. The procedure is repeated until the optimality criterion is attained or the desired frequency is obtained.

The optimality criterion and the iterative algorithm are derived in the context of designing the structure in one of the natural modes. However, the same procedure will be used for designing the structure in the dynamic mode. The dynamic mode may consist of a single natural mode or a linear combination of a set of natural modes depending on the spatial distribution and dynamic characteristics of the forcing function. A more detailed discussion is given in Section 7.

5. EIGENVALUES AND EIGENVECTORS OF BANDED SYMMETRIC MATRICES

The algorithm derived in the last section presupposes the existence of an economical means of determining the eigenvalues and eigenvectors of the equation

$$\omega^2 \mathbf{M} r = \mathbf{K} r \tag{47}$$

The generalized mass and stiffness matrices **M** and **K** are both symmetric and are, in general, positive definite. In addition, when **M** and **K** are derived by a displacement formulation (Equations 17 and 18), they are sparsely populated and have similar patterns of nonzero element distribution. For the successful application of the optimization algorithm derived in Section 4, the method of solution of Equation 47 should possess the following properties:

1. It should allow for the determination of any desired frequencies and the corresponding normal modes without determining any of the remaining frequencies and modes.

2. The method should have a provision for taking advantage of the sparseness characteristics of the mass and stiffness matrices.

A method of solution that possess these two properties was described by Gupta (Reference 35) and Peters and Wilkinson (Reference 36). The method is based on the Sturm Sequence property for the equation

$$\mathbf{AX} = \lambda \mathbf{IX} \tag{48}$$

where λ is the eigenvalue, $\mathbf{X}$ is the eigenvector, and $\mathbf{I}$ is the identity matrix. When $\mathbf{A}$ is real and symmetric, the leading principal minors of $(\mathbf{A} - \lambda \mathbf{I})$ form a Sturm sequence; i.e., the number of eigenvalues greater than λ is equal to the number of agreements in sign between consecutive members of the sequence p_r $(r = 0, 1, \ldots, n)$ which is given by

$$p_r = \det (\mathbf{A}_r - \lambda \mathbf{I}) \tag{49}$$

where $\mathbf{A}_r$ is the r^{th} leading principal submatrix of $\mathbf{A}$. The $\det (\mathbf{A}_0 - \lambda \mathbf{I})$ is assumed to be equal to 1.

Peters and Wilkinson have shown (Reference 36) that the Sturm Sequence property is valid for the problem

$$\mathbf{AX} = \lambda \mathbf{BX} \tag{50}$$

when the matrix $\mathbf{B}$ is symmetric and positive definite. Under these conditions the sign of $\det (\mathbf{A}_r - \lambda \mathbf{B}_r)$ is the same as that of $\det (\mathbf{A}_r - \lambda \mathbf{I})$.

The Sturm Sequence property in conjunction with a simple bisection procedure allows for the determination of any eigenvalue, without determining any of the remaining eigenvalues, provided there is a means of evaluating the signs of the leading principal minors of $(\mathbf{A} - \lambda \mathbf{B})$.

Wilkinson presented (Reference 37) a simple method of evaluating the leading principal minors of a square matrix. It is a variation of Gaussian elimination with partial pivoting. The method consists of $(n - 1)$ major steps, and each major step, r, consists of $(n - r)$ minor steps, where n is the order of the matrix. The r^{th} major step in the elimination only involves up to $(r + 1)$ rows of the matrix and consists of the following (p238, Reference 37):

For each value of i from 1 to r

a. Compare a_{ii} and $a_{r+1,i}$ i.e. If $|a_{r+1,i}| > |a_{ii}|$, interchange $a_{r+1,j}$ and a_{ij} $(j = 1, \ldots, n)$. When an interchange takes place, it should be recorded.

b. Compute $m_{r+1,i} = \dfrac{a_{r+1,i}}{a_{ii}}$ and overwrite on $a_{r+1,i}$

c. For each value of j from (i + 1) to n: Compute $a_{r+1,j} - m_{r+1,i}a_{ij}$ and overwrite on $a_{r+1,j}$. If the cumulative sum of the number of interchanges from the beginning to the end of the r^{th} major step is k, then the principal minor p_{r+1} is given by

$$p_{r+1} = (-1)^k a_{11}\, a_{22} \cdots a_{r+1,r+1} \tag{51}$$

It should be noted that only the sign of p_{r+1} is of interest and not its magnitude. When the matrix $(A - \lambda B)$ is banded, evaluation of the principal minor involves only $(r - q)$ to r rows, i.e., a total of $(q + 1)$ rows, where $(2q + 1)$ is the maximum band width of $(A - \lambda B)$. This procedure requires only $(q + 1).(2q + 1)$ storage locations, in addition to the storage required for each of the matrices A and B. Some discussion of the storage requirements of the mass and stiffness matrices of the displacement method is presented at the end of this section.

After determining the eigenvalues by the above procedure, the eigenvectors can be determined by inverse iteration using the following recurrence relation (Reference 37).

$$(A - \lambda_i B)\, X_i^{\nu+1} = B X_i^{\nu} \tag{52}$$

where the subscript i stands for the number of the eigenvalue and mode of vibration, and ν refers to the cycle of iteration. Only two or three cycles of iteration are necessary for obtaining the eigenvector.

The selection of the initial vector X_i^0 can normally be quite arbitrary. However, in some special circumstances, the iteration may not converge to the desired vector. An example of this is when the mass and stiffness of a beam include the degrees of freedom corresponding to the axial and flexural deformation. The inverse iteration may fail to converge to the flexural modes when the initial vector contains nonzero terms only in the degrees of freedom corresponding to axial deformation. This happens because the linear theory of displacement formulation does not include the coupling terms between axial and flexural deformations. To avoid such a condition it is recommended that all the elements of the initial vector be set equal to 1.

Although Gaussian elimination with partial pivoting was used in solving for the eigenvalues as described in Reference 37, storage limitations necessitated the use of direct Gaussian elimination in the inverse iteration for determining the eigenvectors. Though Wilkinson strongly recommends partial pivoting to avoid an instability in the decomposition, our results were found to be quite satisfactory.

The following comments on the pattern of distribution of the nonzero elements of the mass and stiffness matrices will help in the efficient use of the above method (Reference 38). In a displacement formulation the mass and stiffness matrices generally have the same pattern of distribution of the nonzero elements. Thus, the comments made in reference to the stiffness matrix are valid for the mass matrix.

The distribution of the nonzero elements is dependent upon the way the nodes of the finite element model are numbered. Because of the symmetry of the stiffness matrix only the lower or upper triangular matrix is considered. For the purpose of this discussion definitions of the following terms are in order. The gross population (P_{gross}) of the stiffness matrix is defined as the total number of elements in the upper triangle of the matrix. The net population (P_{net}) is the total number of nonzero elements in the upper triangle. Zeros resulting from transformations are not

excluded from the net population. The apparent population ($P_{apparent}$) is the actual number of elements considered as nonzeros by a given solution scheme. From these definitions

$$P_{net} \quad < \quad P_{apparent} \quad < \quad P_{gross} \tag{53}$$

For a given structure P_{gross} and P_{net} are invariant and are given by

$$P_{gross} = \frac{N\ (N + 1)}{2} \tag{54}$$

and

$$P_{net} = \frac{n\ (n + 1)}{2}\ (\text{number of nodes}) + \sum_{i=1}^{m} \frac{n^2[k_i(k_i - 1)]}{2} - n^2\,(NR) \tag{55}$$

where N is the total number of degrees of freedom of the structure, n is the number of degrees of freedom of each node (all the nodes are assumed to have the same number of degrees of freedom; when this is not true the necessary modification is simple), k_i is the number of nodes to which the i^{th} element is connected, and m is the number of elements in the structure. The quantity NR is given by

$$NR = \sum_{i=1}^{p} (a_i - 1) \tag{56}$$

where a_i is the number of elements connecting the same pair of nodes and p is the total number of pairs of directly connected nodes. If the structure consists of bar and/or beam elements only, NR is always zero. For the example shown in Figure 1, the value of NR is 3.

The quantity $P_{apparent}$ is dependent on the nature of the solution scheme used. For Gaussian elimination with no pivoting (LDL^T), $P_{apparent}$ may be defined as

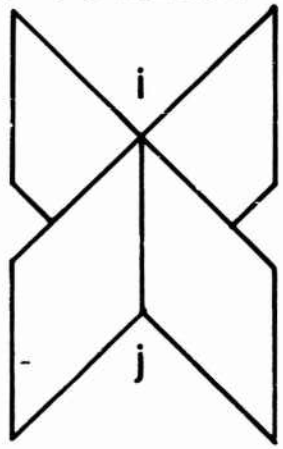
$$P_{apparent} = \sum_{i=1}^{n} Q_j \tag{57}$$

where $Q_j = j - R_j + 1$ and where R_j is the row number of the first nonzero element in the j^{th} column. The solution scheme is most efficient when $P_{apparent} = P_{net}$. However, in large practical structures this condition is difficult to attain.

The value of $P_{apparent}$ changes with the node numbering scheme of the finite element model. The example shown in Figure 2 illustrates this point. A seven node three dimensional bar structure (n = 3) is numbered in three different ways and the resulting effect on the respective stiffness matrices is shown. The non-zero elements are marked by (+). The populations for the three cases are also given in the same figure.

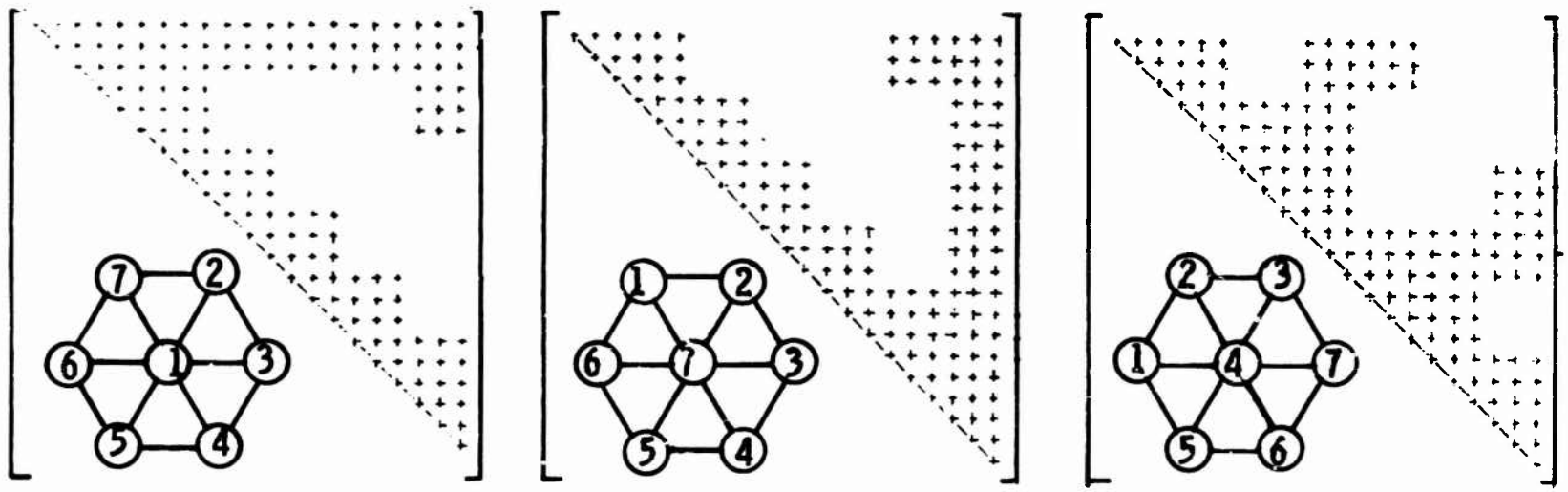
$P_{apparent}$ represents the number of storage locations required for the stiffness matrix and an equal number is required for the mass matrix. This is the procedure used for storing the stiffness and mass matrices for all the problems discussed in Section 8.

FIGURE 1. INTERSECTING PLATES



NR = 3

SCHEME NO.	P_{GROSS}	P_{NET}	$P_{APPARENT}$
1	231	150	231
2	231	150	177
3	231	150	177



SCHEME 1 SCHEME 2 SCHEME 3

FIGURE 2: DISTRIBUTION OF NONZERO ELEMENTS IN THE STIFFNESS MATRIX

6. FREQUENCY GRADIENTS AND NUMERICAL SEARCH

The recursion relation derived in Section 4, Equation 46 , is extremely rapid in effecting changes in the frequencies of the structure. If the object is simply to obtain a minimum weight structure with maximum stiffness, then Equation 46 is adequate and there is no need for determining frequency gradients. However, wher the situation demands that less than a maximum stiffness structure is required, then finer adjustments can be made by a numerical search using frequency gradients.

The expressions for the frequency gradients are derived here with the aid of the Raleigh quotient. Similar expressions are derived earlier in a number of papers (References 24, 26, 28, etc). The i^{th} frequency of the structure may be written as a function of the design variables

$$\lambda_i = \omega_i^2 = F_i\,(A_1, A_2, \ldots, A_m) \tag{58}$$

where λ_i is the i^{th} eigenvalue and $A_1, A_2, \ldots, A_m$ are the m design variables. The function F_i is actually the Raleigh quotient corresponding to the i^{th} natural mode, and it may be written as (Reference 34)

$$\lambda_i = F_i\,(A_1, A_2, \ldots, A_m) = \frac{\mathbf{X}_i^t\,\mathbf{K}\,\mathbf{X}_i}{\mathbf{X}_i^t\,\mathbf{M}\,\mathbf{X}_i} \tag{59}$$

where $\mathbf{X}_i$ is the i^{th} natural mode of the structure, and $\mathbf{K}$ and $\mathbf{M}$ are the generalized stiffness and mass matrices, respectively. If the change in λ_i due to a change in the variable A_j is $d\lambda_{ij}$, then its value can be determined by the Taylor's series expansion

$$d\lambda_\omega = \frac{\partial F_i}{\partial A_j}\,dA_j + \frac{1}{2!}\,\frac{\partial^2 F_i}{\partial A_j^2}\,dA_j^2 + \ldots \tag{60}$$

In the following derivation the terms beyond the first are neglected. From Equation 59

$$\frac{\partial F_i}{\partial A_j} = \frac{2\,\frac{\partial}{\partial A_j}\,(\mathbf{X}_i^t)\,[\mathbf{K} - \lambda_i\,\mathbf{M}]\,\mathbf{X}_i + [\mathbf{X}_i^t\,\Delta\mathbf{K}_j\,\mathbf{X}_i - \lambda_i\,\mathbf{X}_i^t\,\Delta\mathbf{M}_j\,\mathbf{X}_i]}{\mathbf{X}_i^t\,\mathbf{M}\,\mathbf{X}_i} \tag{61}$$

where $\Delta\mathbf{K}_j$ and $\Delta\mathbf{M}_j$ are the change in the stiffness and mass matrices of the structure due to a change in the variable A_j. Since a change in the size of the element effects only the stiffness and mass matrices of that particular element, $\Delta\mathbf{K}_j$ and $\Delta\mathbf{M}_j$ are the j^{th} element matrices expressed in the general structure coordinate system.

From Equation 25 the first term in the numerator of Equation 61 is zero. Then Equation 61 becomes

$$\frac{\partial F_i}{\partial A_j} = \frac{\mathbf{X}_i^t\,\Delta\mathbf{K}_j\,\mathbf{X}_i - \lambda_i\,\mathbf{X}_i^t\,\Delta\mathbf{M}_j\,\mathbf{X}_i}{\mathbf{X}_i^t\,\mathbf{M}\,\mathbf{X}_i} \tag{62}$$

Thus, to the first order of approximation, $d\lambda_{ij}$ can be written as

$$d\lambda_{ij} = \frac{[\, s_{ij}^t \, k_j \, s_{ij} - \lambda_i \, s_{ij}^t \, m_j \, s_{ij} \,]}{X_i^t M X_i} \tag{63}$$

where s_{ij} is the displacement vector of the j^{th} element when the structure is vibrating in the i^{th} mode, and k_j and m_j are the stiffness and mass matrices of the j^{th} element corresponding to the change in size dA_j, the value of dA_j is assumed to be equal to A_j in determining $d\lambda_{ij}$. The numerator in Equation 63 represents the difference in the strain energy and the kinetic energy of the i^{th} element in free vibration in a finite time interval (Equation 29).. As can be seen from Equation 63 determination of frequency gradients is computationally inexpensive when once the frequencies and modes are determined. This situation is quite contrary to the calculation of displacement gradients in the static case (Reference 2).

The numerical search algorithm using frequency gradients for the adjustment of frequencies is represented by the following recursion relation.

$$\alpha^{\nu+1} = \alpha^{\nu} \pm \Delta D^{\nu} \tag{64}$$

where α is the design variable vector, ν refers to the cycle of iteration and Δ represents the scalar step size. Each element D_j of vector D is assumed to be directly proportional to $d\lambda_{ij}$ and inversely proportional to the length l_j

$$D_j = c \frac{d\lambda_{ij}}{l_j} \tag{65}$$

The change in λ_i corresponding to the change D_j is

$$\delta\lambda_{ij} = \frac{d\lambda_{ij}}{\alpha_j} \; D_j \tag{66}$$

The total change in λ_i due to a change in the size of all the elements may be written as

$$\delta\lambda_i = \sum_{j=1}^{m} \delta\lambda_{ij} = c \sum_{j=1}^{m} \frac{1}{\alpha_j} \frac{(d\lambda_{ij})^2}{l_j} \tag{67}$$

Thus, the proportionality constant c is given by

$$c = \frac{\delta\lambda_i}{\sum_{j=1}^{m} \frac{1}{\alpha_j} \frac{(d\lambda_{ij})^2}{l_j}} \tag{68}$$

Then the element D_j is given by

$$D_j = \frac{\delta\lambda_i}{\sum_{p=1}^{m} \frac{1}{\alpha_p} \frac{(d\lambda_{ip})^2}{l_p}} \frac{d\lambda_{ij}}{l_j} \tag{69}$$

where $\delta\lambda_i$ is the desired change in the i^{th} frequency of the structure. A similar search algorithm is derived in References 1 and 2 for displacement and stress-constraint problems.

The value of Δ is set equal to 1.0. By adjusting Δ the rate of change can be further adjusted.

7. DYNAMIC RESPONSE AND OPTIMUM DESIGN

The response of a structural system subjected to a time dependent forcing function is governed by (Equation 20)

$$M\ddot{r} + Kr = R_o F(t) \tag{70}$$

where R_o is the generaiized force matrix whose elements are assumed to be functions of the spatial coordinates only. It is further assumed that the time function F(t) is the same for all the generalized forces.

If the number of degrees of freedom of the system is n, then Equation 70 represents n coupled second order differential equations.

The Sturm Sequence property in conjunction with a bisection procedure and inverse iteration permit the determination of the frequencies and normal modes economically at any range of the frequency spectrum. Then the normal mode method can be used for finding the response of the system.

In the normal mode method the response vector is approximated by a finite number of normal coordinates in the following form:

$$r = \Psi q \tag{71}$$

where each column of the matrix Ψ is a normal mode and q represents the vector of normal coordinates. If all the normal modes are included in Ψ , then Equation 71 represents the exact response of the system. However, in practice only a small number of normal modes contribute significantly to the response of the system and the remaining modes need not be included in Ψ . Which modes are significant depends on the nature of the forcing function. A procedure for determining the significant modes is discussed later in this section. Thus, the number of normal modes included in Ψ is p ($p \leq n$). Substitution of Equation 71 into 70 and premultiplication by Ψ^t gives

$$\Psi^t M \Psi q + \Psi^t K \Psi q = \Psi^t R_o F(t) \tag{72}$$

Equation 72 represents a set of uncoupled second order differential equations in normal coordinates and can be written as

$$\mathbb{M}\ddot{\mathbf{q}} + \mathbb{K}\mathbf{q} = \Psi^t \mathbf{R}_o F(t) \tag{73}$$

where $\mathbb{M}$ and $\mathbb{K}$ are diagonal matrices and their elements are given by

$$\mathbb{M}_i = \Psi_i^t \mathbf{M} \Psi_i \tag{74}$$

$$\mathbb{K}_i = \Psi_i^t \mathbf{K} \Psi_i \tag{75}$$

Now the number of uncoupled equations is equal to the number of normal modes (p) selected to approximate the response. It should be pointed out that the subscript i in Equations 74 and 75 does not refer to the i^{th} column of the nxn normal mode matrix, but instead it refers to the i^{th} column of the nxp normal mode matrix selected to express the dynamic response.

The i^{th} uncoupled equation may be written as

$$\ddot{q}_i + \omega_i^2 q_i = \frac{1}{\mathbb{M}_i} \Psi_i^t \mathbf{R}_o F(t) \tag{76}$$

where ω_i is the natural frequency corresponding to the i^{th} normal mode selected for expressing the response. With the use of the convolution integral the solution of Equation 76 can be written in the following form (Reference 34)

$$q_i(t) = \frac{1}{\omega_i^2 \mathbb{M}_i} \Psi_i^t \mathbf{R}_o D_i(t) \tag{77}$$

where $D_i(t)$ is the dynamic load factor or magnification factor and in the absence of damping is given by

$$D_i(t) = \int^t \omega_i \sin \omega_i (t - \tau) F(\tau) d\tau \tag{78}$$

where t represents the time at which the response is desired and τ is the intermediate time variable.

The complexity of evaluating the integral in Equation 78 depends on the nature of the forcing function F(t). When F(t) is periodic or aperiodic and satisfies the Dirichlet conditions in any finite interval, it can be expanded into a Fourier integral and an approximate solution to the integral can be obtained. If F(t) is obtained by experimental or other empirical data, the integral can be evaluated by any of the numerical integration schemes.

Substitution of Equation 77 in 71 gives the response of the system in the following form

$$r = \Psi \Omega D \Psi^t R_o \tag{79}$$

where Ω and D are diagonal matrices with elements given by

$$\Omega_i = \frac{1}{\omega_i^2 M_i} \tag{80}$$

$$D_i = D_i(t) \tag{81}$$

The dynamic response of the system in generalized coordinates is given by Equation 79 . The element generalized coordinates. forces and stresses can be determined by substituting Equation 79 into Equations 14 , 23 , and 24 .

If the structure is subjected to a static load vector R_o, then the dynamic load factor matrix D becomes an identity matrix and Equation 79 can be written as

$$r_o = \Psi \Omega \Psi^t R_o \tag{82}$$

In the static case the generalized masses are unity, and ω_i and Ψ are the eigenvalues and eigenvectors of the stiffness matrix.

A comparison of Equations 22 and 82 reveals that

$$K^{-1} \simeq \Psi \Omega \Psi^t \tag{83}$$

assuming K is non-singular. Equation 82 represents the static response of the structure. When the loads are time dependent, the static response is magnified by the dynamic load factor matrix D.

The dynamic response as given by Equation 79 can also be written as a linear combination of natural modes of the structure

$$r = \sum_{i=1}^{p} c_i \Psi_i \tag{84}$$

From Equations 79 and 84 c_i may be written as

$$c_i = \frac{D_i}{\omega_i M_i} \Psi_i^t R_o \tag{85}$$

where the dynamic load factor D_i is given by Equation 78 .

Each term in Equation 84 represents the contribution of one mode to the dynamic response. If all the modes of the system are included in the summation, the response given by Equation 84 would be exact to the degree of approximation

expected of a discretized model. When the forcing function is in the low frequency range compared to the actual frequencies of the structure, two or three modes at the lower end of the frequency spectrum are adequate for satisfactory representation of the dynamic response. In high frequency vibrations, however, the number of active terms are expected to be larger.

It should be recognized that the nature of the time dependence as well as the spatial distribution of the forcing function determine which of the modes predominate in the dynamic response. The design procedures that simply stiffen the structure to raise the fundamental frequency are inadequate or even produce grossly non-optimum designs for some types of forcing functions. The procedure outlined in this paper does not suffer from this deficiency, because it takes into consideration the nature of the dynamic response as represented by Equation 84 .

The modes that predominate in the dynamic response are called the critical modes, and a trial procedure for filtering these modes is presented here. It is based on a study of the virtual work of the peak dynamic forces when subjected to the dynamic response as given by Equation 84 . The virtual work of the peak forces is defined as

$$V_w = R_o^t r \tag{86}$$

Figure 3 illustrates two possible cases of the variation of the virtual work, V_w, with the number of modes included in the dynamic response. In the first case, the five modes at the lower end of the spectrum are significant. In the second case modes 7 – 11 are significant, and the remaining modes can be left out.

The recommended procedure for determining the critical modes is as follows: For a given design the dynamic response is determined by adding the effect of one mode at a time in Equation 84 . After each addition the virtual work, V_w, is determined from Equation 86 . A plot of the virtual work against the number of modes would reach a plateau when an adequate number of modes are included in the response. In general this filtering procedure need not be repeated in every optimization cycle. It would be adequate to go through this procedure for the initial design.

Thus, the step by step procedure for the minimum weight design of a dynamically loaded structure is as follows:

1. For an initial design the critical modes are determined by the procedure outlined in the foregoing discussion.

2. The peak dynamic response of the structure is determined by including all the critical modes in Equation 84 .

3. The dynamic stresses corresponding to the peak response are determined by Equations 23 and 24 .

4. By scaling the design to the constraint surface, the feasible design and its weight are determined. (See Reference 2 for the scaling procedure and other details).

5. The frequency of the forced vibration of the structure is determined by substituting the dynamic displacement vector r (Equation 84) into the Rayleigh quotient

$$\omega^2 = \frac{r^t K r}{r^t M r} \tag{87}$$

6. The structure is resized by substituting the dynamic displacement vector r (Equation 84) and ω^2 (determined in the last step) into the recurrence relation based on the optimality criterion, Equation (46).

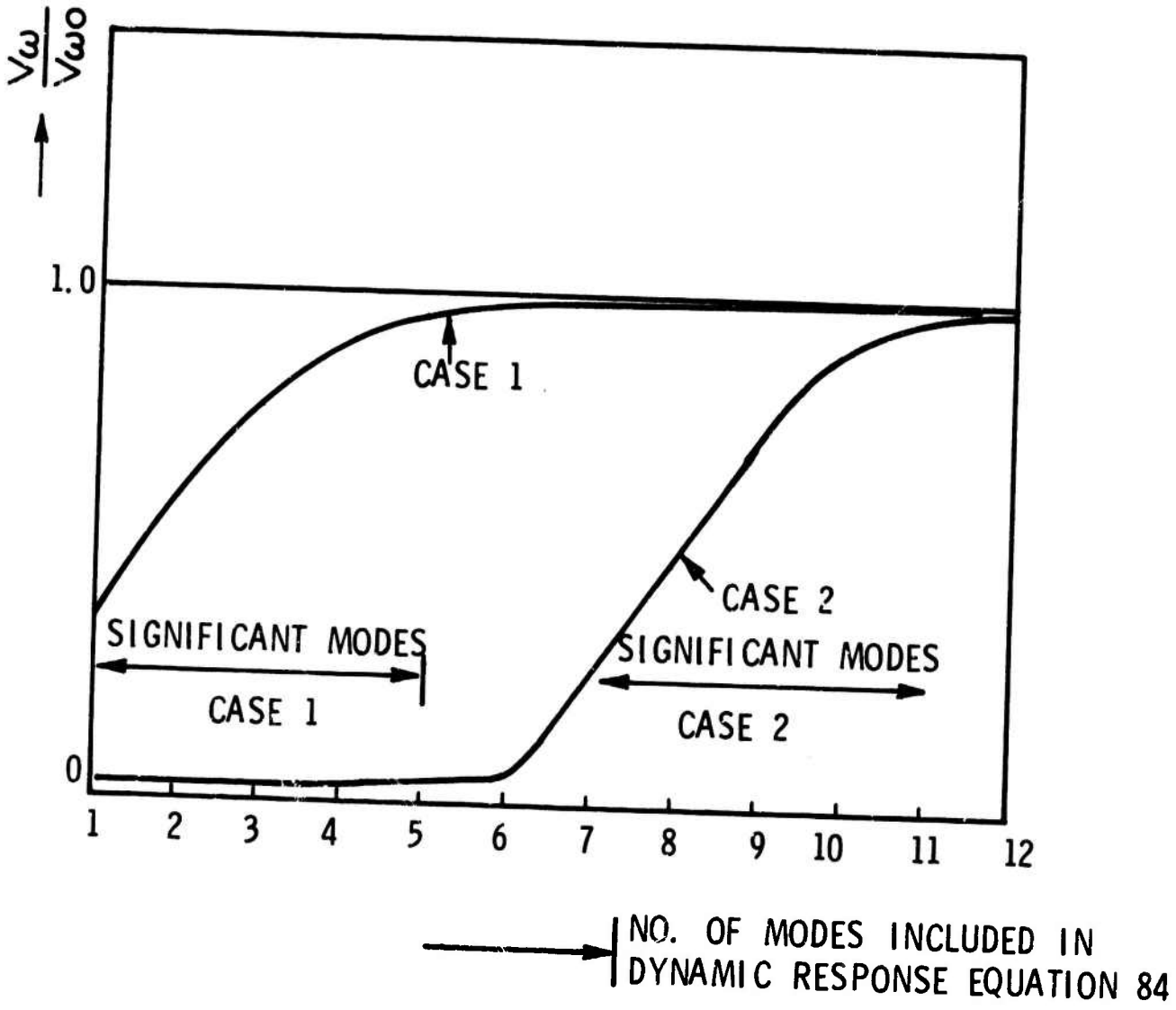

Vω/Vωo
1.0
0
CASE 1
CASE 2
SIGNIFICANT MODES
CASE 1
SIGNIFICANT MODES
CASE 2
1 2 3 4 5 6 7 8 9 10 11 12
NO. OF MODES INCLUDED IN
DYNAMIC RESPONSE EQUATION 84

FIGURE 3.

7. Steps 2 thru 6 are repeated as long as there is an improvement in the design.

The above procedure is valid for both stress and displacement constrained problems. However, in the case of displacement constraints the design can be further improved by a numerical search at the end as in the static case, Reference 2.

In the foregoing discussion it was tacitly assumed that the forcing function could be expressed by a finite number of harmonics.

8. DESIGN EXAMPLES

The design examples presented in this section may be grouped into two categories. The first group of four examples are for the purpose of studying the variation of the natural frequencies and material distribution when the structures are designed in their natural modes. Designing in one of the natural modes is tantamount to designing in the actual dynamic mode if the given forcing function activates only that natural mode. The scaling of the design to the constraint surface is the only additional step required to obtain the actual design. For instance, if a loading condition activates only the fundamental mode, then the relative distribution of the material obtained by designing in the fundamental mode would be the same as that obtained by designing in the actual dynamic mode.

The recursion relation based on the optimality criterion derived in Section 4 gives the stiffest structures when they are designed in their fundamental mode. The stiffness of the structure is measured by the ratio of the fundamental frequency to the relative weight.

In the next group of two examples the structures are designed for dynamic loads. The first example in this group is a rectangular frame with beam elements. This frame is designed for three different loading cases. These loading cases are devised to show how the spatial distribution of the forcing function activates different modes even though the dynamic characteristics are the same in all cases. The second example in this group is a wing structure idealized by bar elements. This structure is designed for a periodic forcing function. It represents a medium size problem in terms of degrees of freedom and number of elements.

In the examples involving beam elements the principal moment of inertia is assumed to be the design variable. It is further assumed for the problems solved in this paper that the radius of gyration of the beam section is the same for all the elements. A sandwich beam with constant depth and variable face sheet thicknesses would conform to this assumption. The elastic constants, mass properties, and other design data are given in Table 1 for all the examples.

Example 1: Cantilever Beam — Design in the First Mode.

The cantilever beam shown in Figure 4 is designed in its fundamental mode. The design started with the assumption of equal sizes for all the elements. The elements are resized by the iterative algorithm based on the optimality criterion (Equation 46). The first four natural frequencies, the relative weight, and the ratio of the fundamental frequency to the relative weight are given in Table 2 for

Table 1: Design Data for all Examples

Example	1,2,3	4	5	6
Elastic Modulus (10^6 Lbs/In^2)	29.000	29.000	29.000	10.000
Radius of Gyration (In)	2.646	3.426	8.731	
Mass Density (10^{-2} Slugs/In^3)	.884	.884	.888	.313
Stress Limit (10^3 Lbs/In^2)			29.000	25.000

Table 2: Cantilever Beam – Design in First Mode

$\lambda_i = \omega_i^2 \, 10^6/E$

Cycle No.	λ_1	λ_2	λ_3	λ_4	Rel. Wt.	λ_1/Wt.
1	.316	12.312	97.680	377.734	16.966	.019
2	2.353	31.270	150.098	448.047	7.413	.317
3	5.043	46.914	179.102	483.203	5.885	.857
4	7.398	58.164	200.195	504.297	5.280	1.401
5	9.086	65.195	210.742	511.328	4.991	1.820
6	10.264	70.117	221.289	518.359	4.840	2.121
7	10.967	72.578	224.805	525.391	4.754	2.307
8	11.494	75.391	228.320	525.391	4.704	2.443
9	11.846	76.797	231.836	532.422	4.673	2.535
10	12.022	76.797	231.836	532.422	4.653	2.584

Sizes of the Elements in the Final Design

Element No.	1	2	3	4	5	6	7
Rel. Size	.0002	.003	.018	.082	.250	.566	1.00

each iteration. This table also contains relative sizes of the elements in the final design. The distribution of the moment of inertia along the length of the beam is shown in Figure 4.

Example 2: Simply Supported Beam: Design in the First Mode

The simply supported beam shown in Figure 5 is designed in its fundamental mode. The first four natural frequencies, the relative weight of the beam, and the ratio of the fundamental frequency to the relative weight are given in Table 3 for each iteration. It is interesting to note that the values of the natural frequencies are reduced slightly but the reduction in weight is much more pronounced. The relative sizes of the elements in the final design are given in Table 3. The graphic representation of the material distribution is shown in Figure 5.

Example 3: Fixed – Fixed Beam: Design in the First Mode

The fixed–fixed beam shown in Figure 6 is designed in its fundamental mode. The results of this design are given in Table 4 and the graphic representation of the material distribution is shown in Figure 6.

Example 4: Rectangular Frame: Design in the First Four Modes

The rectangular frame shown in Figure 7 is designed independently in the first four modes. The natural frequencies and the relative sizes are given in Table 5. The natural modes and the material distribution of the final designs are shown qualitatively in Figure 7.

Example 5: Rectangular Frame: Design for Dynamic Loads

The rectangular frame shown in Figure 7 is designed for three different cases of dynamic loads. In all three cases, the forcing functions are periodic and have the same circular frequency. The spatial distribution of the forcing functions are shown for the three cases in Figure 8.

The dynamic load factor, $D_r(t)$, corresponding to the periodic forcing function after integration (Equation (78)) is given by

$$D_r(t) = \frac{1}{[1-(\frac{p}{\omega_r})^2]} (\sin pt - \frac{p}{\omega_r} \sin \omega_r t)$$

A plot of the dynamic load factors corresponding to the first four natural frequencies of the final design in Case 1 is given in Figure 9. The peak values of $D_r(t)$ are used in the dynamic load factor matrix in evaluating response (Equation 84) even though they do not occur at the same time. This approximation, in general, produces conservative designs. A more accurate procedure would he to evaluate the dynamic response at sufficiently small time intervals and to select the response at which the dynamic stresses are a maximum.

Case 1: Dynamic Lateral Force

In the first case the frame is subjected to a periodic lateral force. The design started with the assumption of equal sizes for all the elements. For this design a plot of the virtual work, V_w, against the number of modes in the response is given

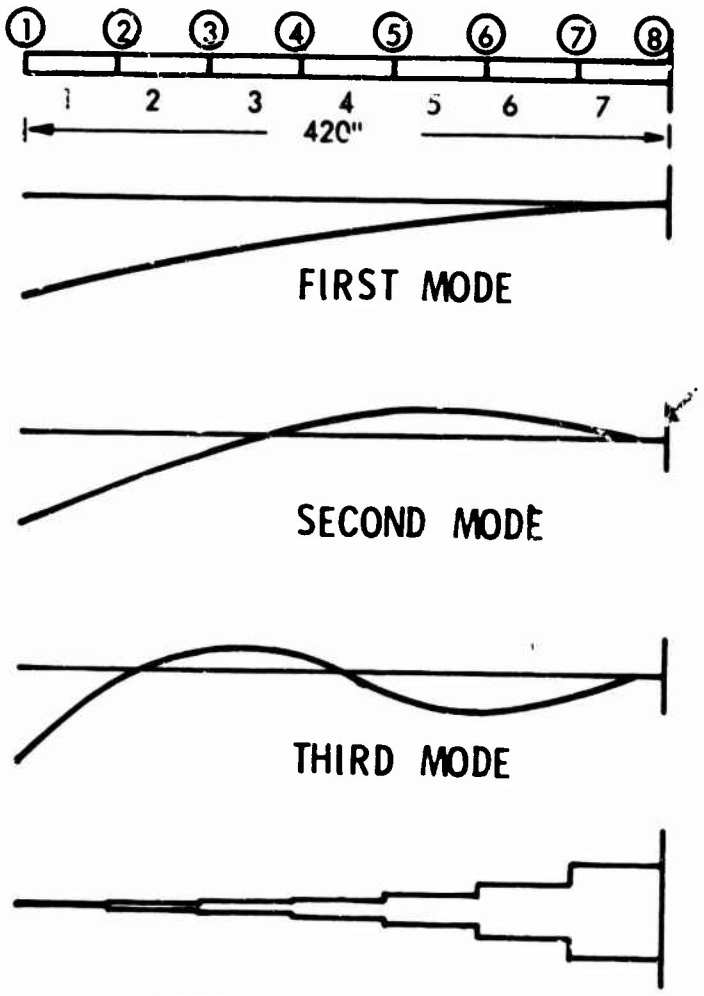


FIGURE 4. CANTILEVER BEAM

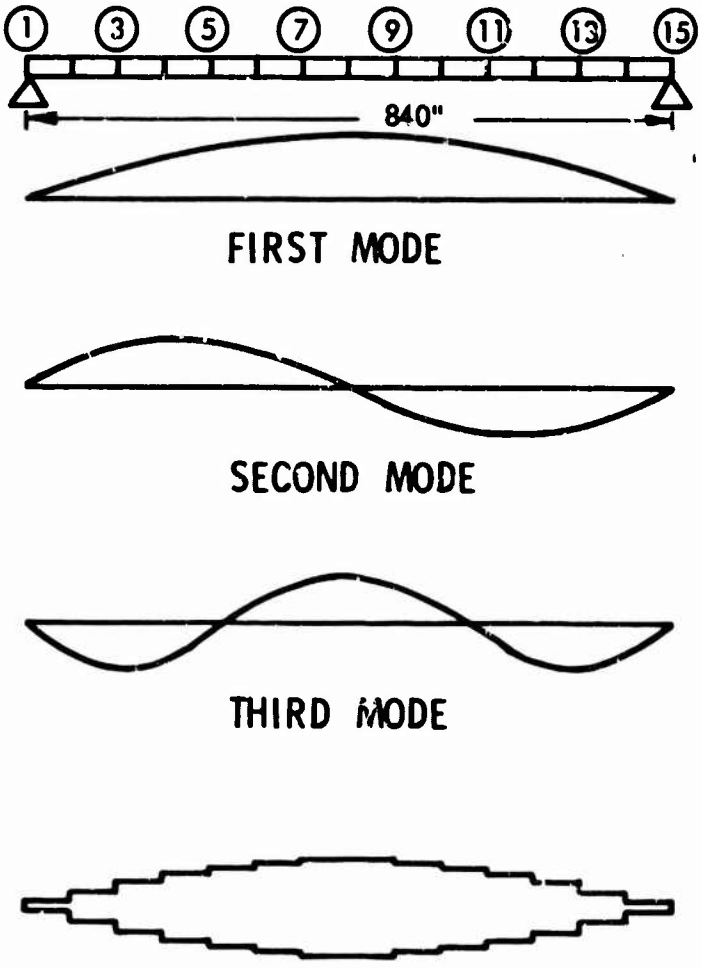


FIGURE 5. SIMPLY SUPPORTED BEAM

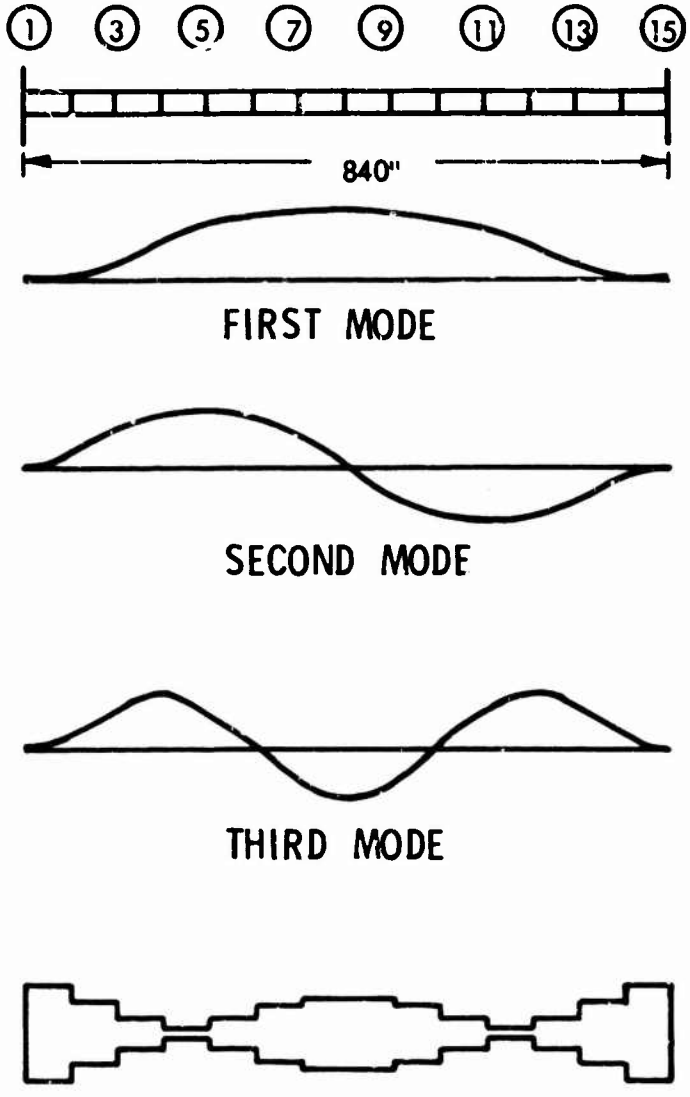


FIGURE 6. FIXED - FIXED BEAM

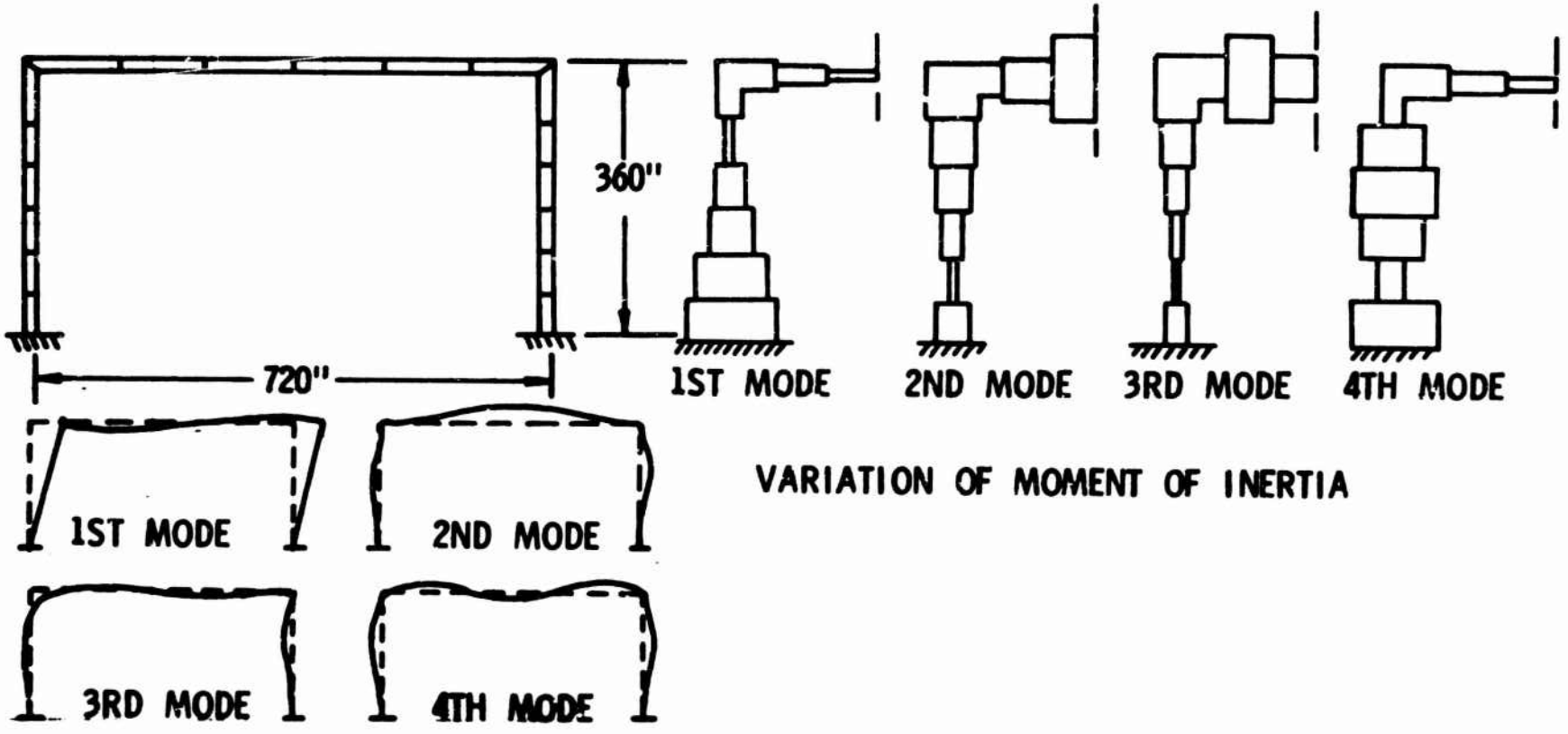
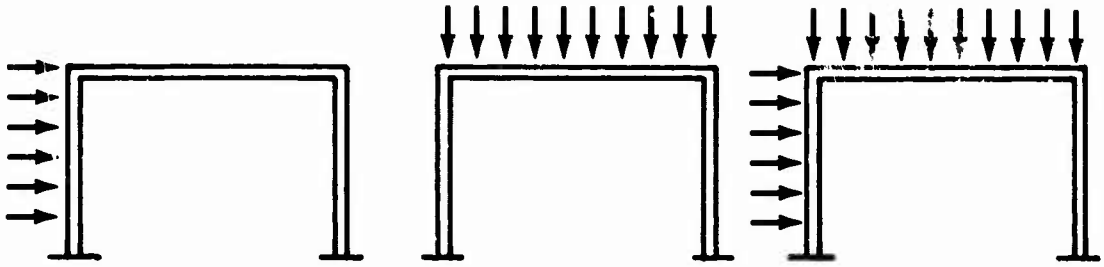


FIGURE 7 RECTANGULAR FRAME: DESIGNS IN NATURAL MODES



$\underset{\sim}{R} = \underset{\sim}{R}_0 \sin pt$

*Note: $p/\omega_r \ll 1/10$ in all cases

PEAK FORCES

CASE 1	1000 LBS AT NODES 2 TO 7 (X-DIRECTION)
CASE 2	1000 LBS AT NODES 7 TO 13 (Y-DIRECTION)
CASE 3	COMBINATION OF CASES 1 AND 2

FIGURE 8: DYNAMIC LOAD DISTRIBUTION

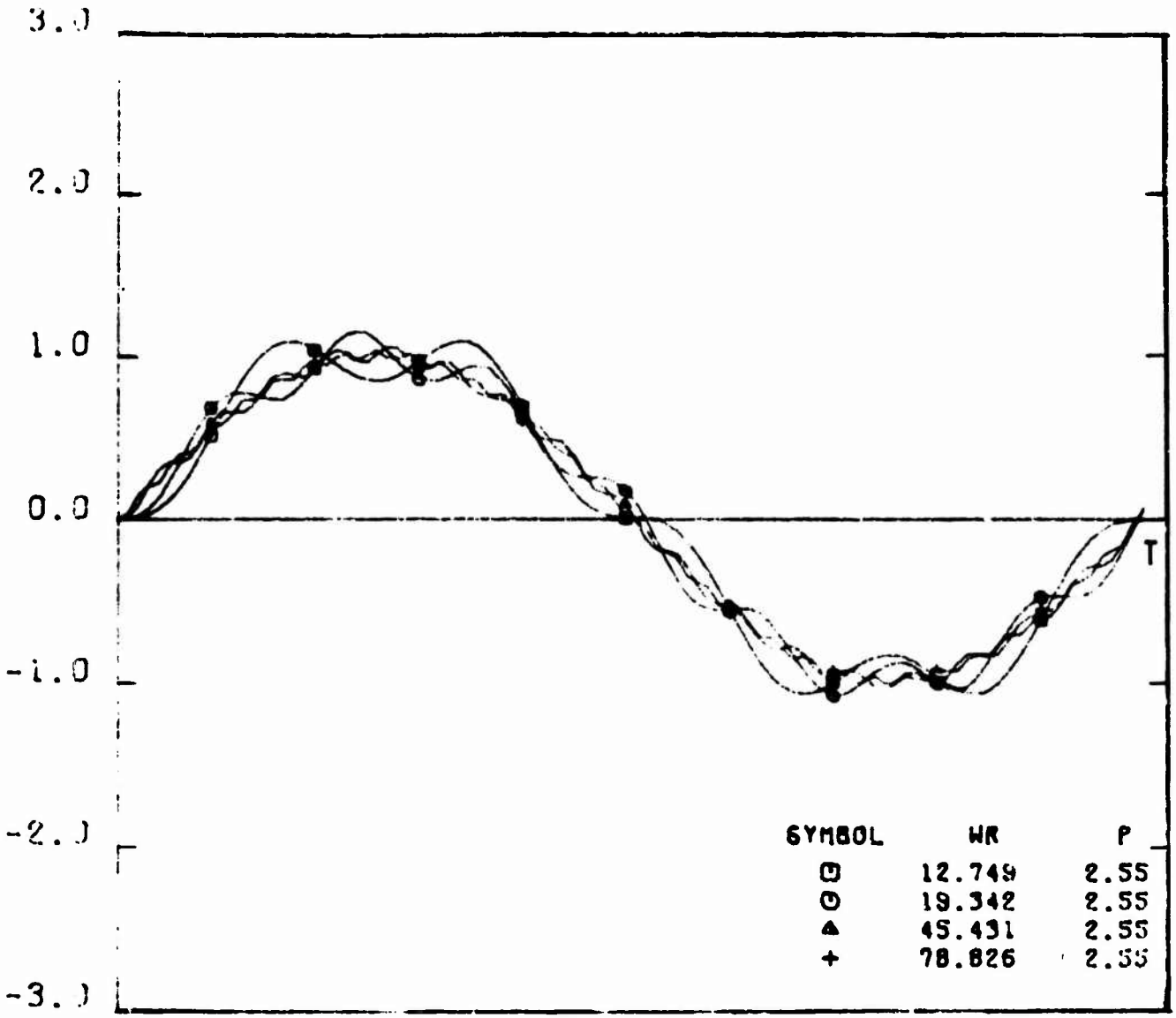


Figure 9: Variation of Elements of Dynamic Load Factor Matrix

Table 3: Simply Supported Beam – Design in First Mode

$\lambda_i = \omega_i^2 \, 10^6/E$

Cycle No.	λ_1	λ_2	λ_3	λ_4	Rel. Wt.	λ_1/Wt.
1	.155	2.494	12.549	39.883	33.932	.005
2	.145	2.248	11.846	38.477	21.921	.007
3	.143	2.213	11.846	38.477	21.402	.007
4	.143	2.213	11.846	38.477	21.365	.007
5	.143	2.213	11.846	38.477	21.362	.007

Sizes of the Elements in the Final Design

Element No.	1	2	3	4	5	6	7
Rel. Size	.119	.313	.508	.687	.836	.944	1.00

TABLE 4: FIXED – FIXED BEAM – DESIGN IN FIRST MODE

$\lambda_i = \omega_i^2 \, 10^6/E$

Cycle No.	λ_1	λ_2	λ_3	λ_4	Rel. Wt.	$\lambda_{1/Wt.}$
1	.796	6.027	23.183	63.789	33.933	.023
2	.993	8.242	33.379	76.797	18.034	.055
3	1.009	8.102	31.973	72.578	18.231	.055
4	1.044	8.102	31.621	70.117	17.879	.058
5	1.062	8.242	31.621	69.414	17.691	.060
6	1.062	8.242	31.621	69.414	17.597	.060
7	1.062	8.242	31.621	69.414	17.553	.061
8	1.062	8.242	31.270	68.711	17.530	.061
9	1.062	8.242	31.270	68.711	17.518	.061
10	1.062	8.242	31.270	68.711	17.512	.061

SIZES OF THE ELEMENTS IN THE FINAL DESIGN

Element No.	1	2	3	4	5	6	7
Rel. Size	1.00	.657	.330	.089	.307	.545	.684

TABLE 5: RECTANGULAR FRAME (DESIGNS IN FOUR MODES)

$\lambda_i = \omega_i^2 \, 10^6/E$

		λ_1	λ_2	λ_3	λ_4	Rel. Wt.
Initial Design		.413	1.343	10.440	24.590	34.700
Final Design	1st. Mode	1.062	1.642	9.508	31.270	11.227
	2nd. Mode	.175	.979	9.367	23.184	16.198
	3rd. Mode	.164	.852	8.242	26.699	15.026
	4th. Mode	.589	1.896	11.494	22.832	14.866

SIZES OF THE ELEMENTS IN THE FINAL DESIGN

Element No.		1	2	3	4	5	6	7	8	9
Rel. Size	1st. Mode	1.000	.733	.475	.234	.066	.236	.292	.198	.079
	2nd. Mode	.208	.062	.154	.316	.473	.620	.366	.518	1.000
	3rd. Mode	.183	.060	.125	.231	.288	.293	.433	1.000	.574
	4th. Mode	1.000	.267	.621	.947	.722	.219	.334	.180	.169

SYMMETRICAL ABOUT CENTER LINE

In Figure 10. This plot shows that the first mode is the most significant one, and the modes beyond the fourth have little contribution. The first four natural frequencies of the initial and the final designs for this loading case are given in Table 6. Table 6 also contains the sizes of the elements in the final design. The graphic representation of the material distribution is shown in Figure 11.

Case 2: Dynamic Vertical Force.

The plot of the virtual work, V_W, for this Case (Figure 10) shows that only the second mode contributes significantly to the dynamic response. All other modes including the first have very little contribution. The final design and the first four natural frequencies are given in Table 6. The material distribution is shown in Figure 11.

Case 3: Combination of Lateral and Vertical Forces.

For this loading case the first two are the significant modes (see Figure 10). The modes beyond the fifth have little contribution. The final design and the natural frequencies are given in Table 6.

It is evident from the results of this example that the spatial distribution of the forcing function determines primarily the modes that participate in the dynamic response.

Example 6: Wing Structure Subjected to Periodic Forces

Figure 12 shows a typical transport wing idealized as a three dimensional bar structure. It has two hundred forty degrees of freedom and four hundred forty elements. The wing is subjected to a sinusoidal forcing function. The final design and the first four natural frequencies are given in Table 7.

9. SUMMARY AND CONCLUSIONS

The optimization method presented in this paper may be called designing in the dynamic mode. The dynamic mode may be a single natural mode of the structure or a linear combination of a set of natural modes depending on the spatial distribution and dynamic characteristics of the forcing function. There are three basic steps in this optimization procedure. The first step consists of determining the number of significant modes necessary to represent the dynamic mode. In the second step, the given design is scaled to satisfy the dynamic stress and displacement constraints. After this step the weight of the feasible design can be determined for comparison with the designs in the previous iterations. In the third step, the elements of the structures are resized by the algorithm based on either the optimality criterion or the numerical search. Only the last two steps are repeated in each iteration for obtaining the optimum structure.

Each of the three major steps consist of several substeps. The details of these substeps are pointed out throughout the body of this paper. For instance, Section 2 contains a summary of the equations of dynamic analysis. They are presented in the context of the displacement method of finite element analysis. An optimality criterion for minimum weight structures is derived in Section 3. It is based on the study of the strain energy and the kinetic energy of the elements in the dynamic mode. An iterative algorithm for attaining the optimality criterion is derived in Section 4. This section also contains a step by step procedure for using the algorithm in the design of minimum weight structures. In Section 5, a procedure for the evaluation of eigenvalues and eigenvectors of banded symmetric matrices is discussed. The eigenvalues

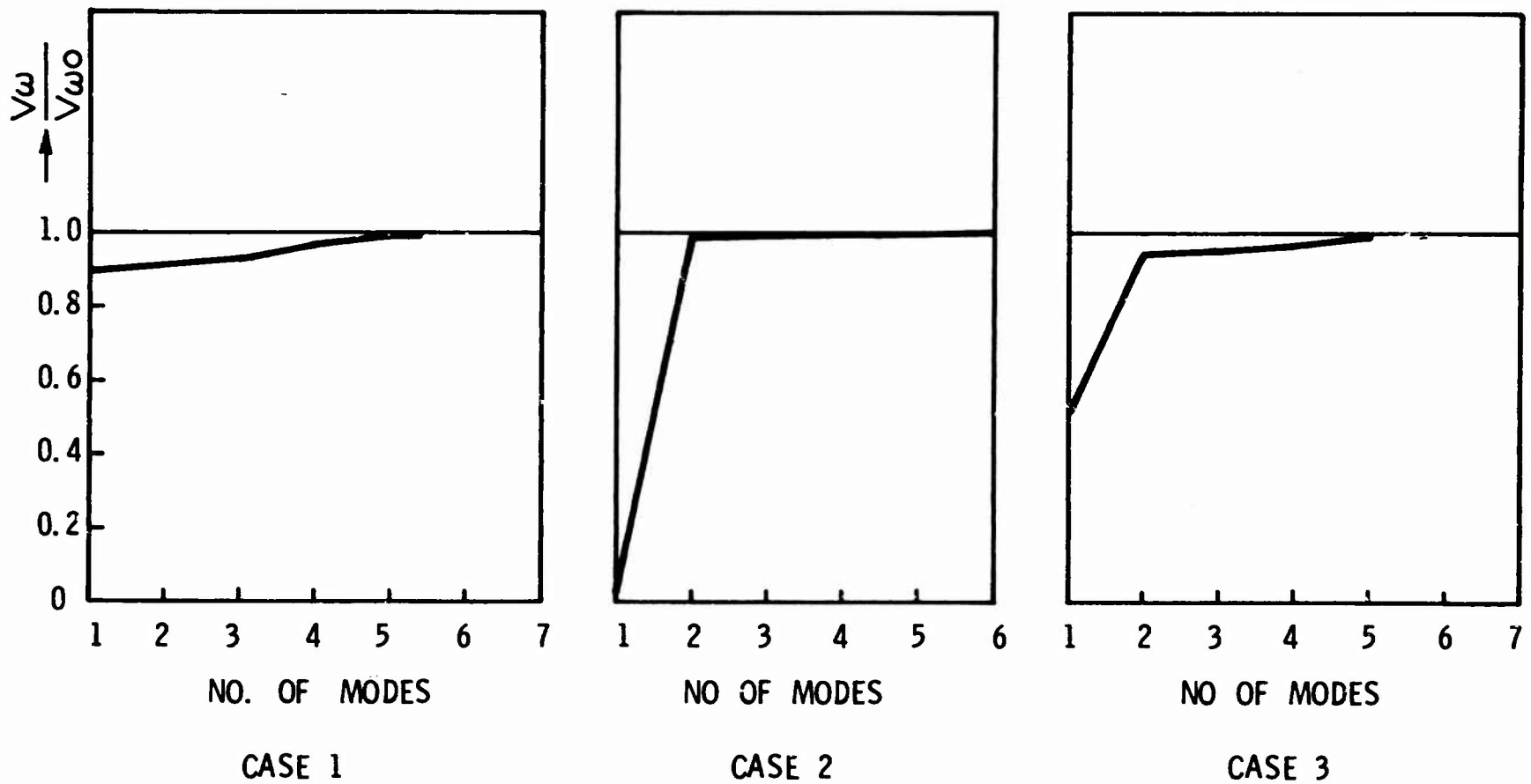
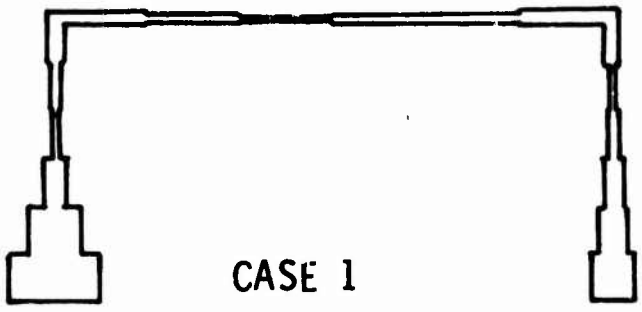
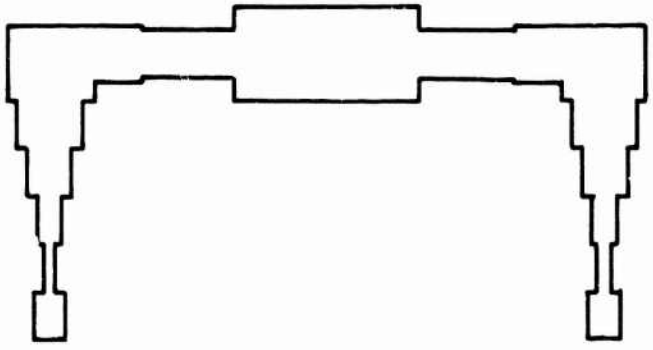


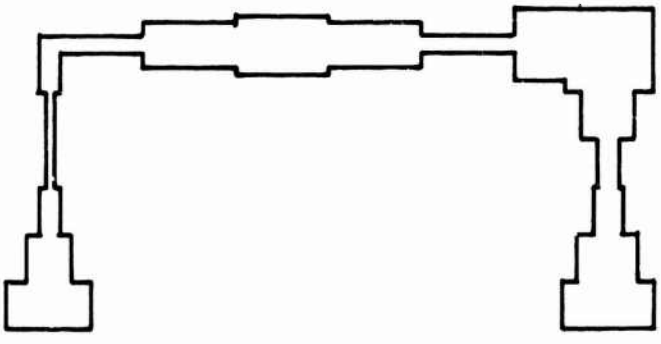
FIGURE 10. PLOT OF VIRTUAL WORK ~ NO. OF MODES IN DYNAMIC RESPONSE



CASE 1



CASE 2



CASE 3

FIGURE 11. MATERIAL DISTRIBUTION OF DYNAMICALLY LOADED FRAME

LOADS PER NODE (POUNDS)

①	②	③
500	1000	1500

RIB DIMENSIONS

RIB AT	AB	CD	EF	GH
X = 0"	5.25"	7.00"	5.25"	3.50"
X=600"	24.00"	32.00"	24.00"	16.00"

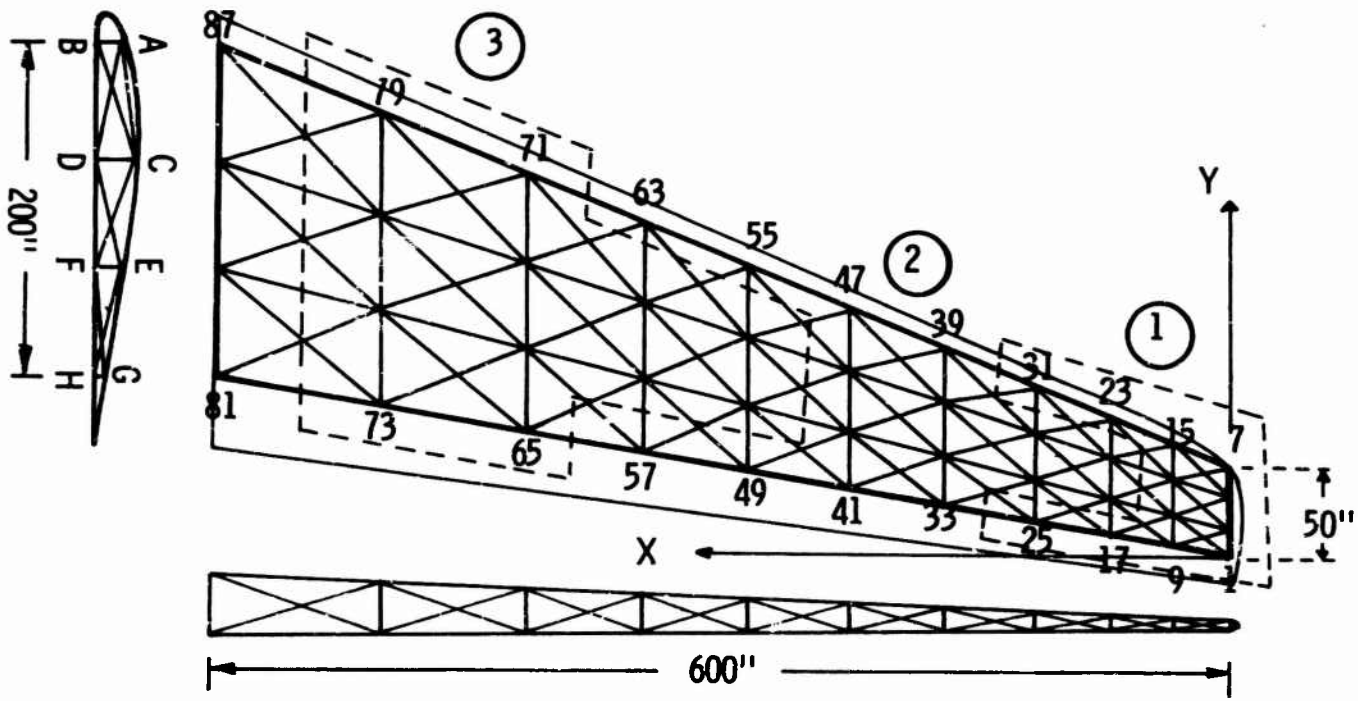


FIGURE 12. WING STRUCTURE

TABLE 6: RECTANGULAR FRAME: DESIGN FOR DYNAMIC LOADS

Case 1: Periodic Lateral Forces

$$\lambda_i = \omega_i^2 \, 10^6/E$$

Design	λ_1	λ_2	λ_3	λ_4	Weight
Initial	2.67	8.66	67.30	155.37	1080
Final	6.52	12.73	71.17	214.26	522

Cycle No.	1	2	3	4
Weight	1080	808	526	522

Elem.	1	2	3	4	5	6
Area	6.61	3.75	1.56	0.36	0.99	1.18
Inertia	503.80	285.51	119.11	27.77	75.43	90.05
Elem.	7	8	9	10	11	12
Area	0.91	0.53	0.19	0.36	0.76	1.16
Inertia	69.64	40.16	14.23	27.64	57.64	88.47
Elem.	13	14	15	16	17	18
Area	0.99	0.30	0.63	1.39	2.15	2.92
Inertia	75.77	22.54	48.20	105.64	163.89	222.61

Case 2: Periodic Vertical Forces

$$\lambda_i = \omega_i^2 \, 10^6/E$$

Design	λ_1	λ_2	λ_3	λ_4	Weight
Initial	2.67	8.66	67.31	155.37	563
Final	1.57	7.54	63.79	151.86	506

Cycle No.	1	2	3	4
Weight	563	534	517	506

Elem.	1	2	3	4	5	6
Area	0.77	0.35	0.48	0.95	1.48	2.02
Inertia	58.71	26.40	36.29	72.36	112.61	154.19
Elem.	7	8	9	10	11	12
Area	1.22	1.13	2.06	2.06	1.13	1.22
Inertia	92.68	85.76	157.17	157.17	85.76	92.68
Elem.	13	14	15	16	17	18
Area	2.02	1.48	0.95	0.48	0.35	0.77
Inertia	154.19	112.61	72.36	36.29	26.40	58.71

Case 3: Periodic Combined Forces

$$\lambda_i = \omega_i^2 \, 10^6/E$$

Design	λ_1	λ_2	λ_3	λ_4	Weight
Initial	2.67	8.66	67.31	155.37	940
Final	2.53	8.10	63.09	200.20	718

Cycle No.	1	2
Weight	940	718

Elem.	1	2	3	4	5	6
Area	3.41	1.73	0.65	0.26	0.26	0.82
Inertia	260.18	131.87	49.34	19.70	19.70	62.16
Elem.	7	8	9	10	11	12
Area	0.70	1.79	2.35	1.82	0.69	2.98
Inertia	53.40	136.49	178.89	138.87	52.90	227.35
Elem.	13	14	15	16	17	18
Area	3.83	2.30	0.90	0.98	2.41	3.91
Inertia	292.10	175.65	68.02	74.49	183.49	297.90

TABLE 7: WING STRUCTURE: DESIGN FOR DYNAMIC LOADS

Design	λ_1	λ_2	λ_3	Weight
Initial	23.887	146.582	448.047	9220.312
Final	880.469	1043.95	3337.89	2115.030

Cycle No.	1	2	3	4
Weight	9220.312	3278.977	2389.517	2115.030

are determined by the use of the Sturm Sequence property and a bisection procedure. The eigenvectors are obtained by inverse iteration. A search algorithm for frequency constraint problems is derived in Section 6. It is similar to the one derived for displacement and stress constraint problems in References 1 and 2.

Section 7 contains a step by step procedure for the design of structures with dynamic stress and displacement constraints. This section also describes a procedure for filtering the significant normal modes of a structure subjected to a given dynamic load. It is based on the virtual work of the peak forces in the dynamic mode. This is the most crucial section for the successful application of the method presented in this paper.

Section 8 contains applications of the general procedure presented in this paper to specific problems. In the first group of problems the structures are designed in their natural modes. In the second group, the structures are designed in the dynamic mode when they are subjected to periodic dynamic forces.

It should be pointed out that the method is not limited to the case of periodic forces only. Extensions of the method to the design of structures in dynamic modes such as the ones resulting from aperiodic forces, flutter modes, and static stability modes are contemplated for the future. It is assumed that these dynamic modes can be represented by a linear combination of the natural modes of the structure.

REFERENCES

1. Venkayya, V. B., Khot, N. S., and Reddy, V. S., Energy Distribution in an Optimum Structural Design, AFFDL–TR–68–156.

2. Venkayya, V. B., "Design of Optimum Structures", Conference on Computer Oriented Analysis of Shell Structures, Palo Alto, California, August 1970. Also published in Journal of Computers and Structures, Vol. 1, p. 265–309, Pergamon Press, August 1971.

3. Gellatly, R. A. and Berke, L., Optimum Structural Design, AFFDL–TR–70–165.

4. Dwyer, W. J., Emerton, R. K., and Ojalvo, I. U., An Automated Procedure for the Optimization of Practical Aerospace Structures (Volume I), AFFDL–TR–70–118.

5. Maxwell, C., "Scientific Papers II", 1869––Reprinted by Dover Publications, New York, 1952.

6. Michell, A. G. M., "The Limits of Economy of Material in Frame Structures", Phil. Mag. Series IV, Vol. 8, No. 47, London, November 1904.

7. Hemp, W. S., "Theory of Structural Design", AGARD Report 214, October 1958.

8. Cox, H. L., The Design of Structures of Least Weight, Pergamon Press, Oxford, 1965.

9. Chan, A. S. L., "The Design of Michell Optimum Structures", College of Aeronautics Report No. 142, 1960.

10. Prager, W., and Taylor, J. E., "Problems of Optimal Structural Design", Trans. Journal of Applied Mechanics, ASME, Vol. 35, Series E., No. 1, 1968.

11. Sheu, C. Y., and Prager, W., "Minimum Weight Design with Piecewise Constant Specific Stiffness", Journal of Optimization Theory and Applications, Vol. 2, No. 3, 1968, pp 179–186.

12. Prager, W., and Marcal, P. V., "Optimality Criteria in Structural Design", AFFDL–TR–70–166.

13. Turner, M. J., "Design of Minimum Mass Structures with Specified Natural Frequencies", AIAA Journal, Vol. 5, No. 3, 1967.

14. Taylor, J. E., "Maximum Strength Elastic Structural Design", Proceedings, Journal of the Engineering Mechanics Division, ASCE, Vol. 95, EM3, Proc. Paper 6617, June 1969.

15. Masur, E. F., "Optimum Stiffness and Strength of Elastic Structures", Proceedings Journal of the Engineering Mechanics Division, ASCE, Vol. 96, EM5, Proc. Paper, 7580, October 1970.

16. McIntosh, S. C., and Eastep, F. E., "Design of Minimum-mass Structures with Specified Stiffness Properties", AIAA Journal Vol. 6, No. 5, 1968.

17. Barnett, R. L., and Herman, P. C., "High Performance Structures", NASA CR–1038, 1968.

18. Melosh, R. J., Structural Analysis, Frailty Evaluation and Redesign, AFFDL–TR–70–15.

19. Young, J. W. Jr., and Christiansen, H. N., "Synthesis of a Space Truss Based on Dynamic Criteria", Journal of the Structures Division, ASCE, Vol. 92. ST6, Proc. Paper, December 1966.

20. Schmit, L. A. "Structural Design by Systematic Synthesis", Proceedings of the Second Conference on Electronic Computations, ASCE, September 1960.

21. Klein, B., "Direct Use of the Extremal Principles in Solving Certain Optimizing Problems Involving Inequalities", Journal of the Operations Research Society of America, Vol. 3, No. 2, May 1955.

22. Gellatly, R. and Gallagher, R., Development of a Procedure for Automated Synthesis of Minimum Weight Structural Design, AFFDL–TR–64–141.

23. Gellatly, R., Development of Procedures for Large Scale Automated Minimum Weight Structural Design, AFFDL–TR–66–180.

24. Fox, R. L. and Kapoor, M. P., "Structural Optimization in the Dynamics Response Regime: A Computational Approach", AIAA Dynamics and Aeroelasticity Specialists Conference, New Orleans, La., April 1969.

25. Rubin, C. P., "Dynamics Optimization of Complex Structures", AIAA Dynamics and Aeroelasticity Specialists Conference, New Orleans, La., April 1969.

26. McCart, B. R., Haug, E. J. and Streeter, T. D., "Optimal Design of Structures with Constraints on Natural Frequency", AIAA, Vol. 8, No. 6, June 1970.

27. Salinas, D., On Variational Formulations for Optimal Structural Design, Ph.D. Dissertation, University of California, Los Angeles, 1968.

28. Loomis, W. C. Optimal Design of Discrete Structural Systems, Ph.D. Dissertation, University of California, Los Angeles, 1971.

29. Arora, J. S., Optimal Design of Elastic Structures under Multiple Constraint Conditions, Ph.D. Dissertation, University of Iowa, Iowa City, 1971.

30. Wasiutynski, Z., and Brandt, A., "The Present State of Knowledge in the Field of Optimum Design of Structures", Applied Mechanics Review, 16, pp 341–350, 1963.

31. Sheu, C. Y. and Prager, W., "Recent Developments in Optimal Structural Design", Journal of Applied Mechanics, Series E, 35, pp 102–106, 1968.

32. Pope, G. G., and Schmit, L. A., Structural Design Applications of Mathematical Programming Techniques, AGARDograph No. 149, February 1971.

33. Przemieniecki, J. S., Theory of Matrix Structural Analysis, McGraw-Hill Book Company, New York, 1968.

34. Hurty, W. C., and Rubinstein, M. F., Dynamics of Structures, Prentice-Hall, Inc., Englewood Cliffs, 1964.

35. Gupta, K. K., "Free Vibrations of Single-branch Structural Systems", Journal of the Inst. of Mathematics and its Applications, Vol. 5, No. 3, 1969.

36. Peters, G. and Wilkinson, J. H., "Eigenvalues of $Ax=\lambda Bx$ with Band Symmetric A and B", Computer Journal, Vol. 12, 1969, pp 398–404.

37. Wilkinson, J. H., The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1965.

38. Venkayya, V. B., Discussion on "Efficient Solution of Load-Deflection Equations by Melosh, R. J. and Bamford, R. M.," Proceedings Journal of the Structural Division, ASCE, Vol. 96, ST1, Proc. Paper 6990, January 1970.

SESSION 5. STRUCTURAL APPLICATIONS

Session Chairman

D. S. Warren

Douglas Aircraft Co.
Long Beach, California

ELASTIC CRACK ANALYSIS BY A FINITE ELEMENT HYBRID METHOD*

T. H. H. Pian**, P. Tong*** and C. H. Luk****

Massachusetts Institute of Technology

The paper presents an efficient finite element method for evaluating the elastic stress intensity factors at the tip of a sharp crack. The method is based on a hybrid stress model for which special stress terms which represent the correct singularity behavior at the crack tip can be included. The magnitudes of such singular terms which are among the unknowns of the final matrix equations are, in fact, the stress intensity factors to be evaluated. Example solutions include plane stress cracks of both the opening type (Mode I) and in-plane shear type (Mode II) for isotropic materials and of the opening type for anisotropic materials.

---

*Work described in this paper was supported in whole by the Air Force Office of Scientific Research under Contract F44620-67-C-0019

**Professor of Aeronautics and Astronautics

***Associate Professor of Aeronautics and Astronautics

****Graduate Research Assistant

SECTION I

INTRODUCTION

The fundamental quantity to be considered in fracture mechanics is the "stress intensity factor" (Reference 1). For example, for an elastic plane stress or plane strain problem, in the immediate vicinity of the tip of a through crack the stresses varies as $1/\sqrt{r}$ where r is the radial coordinate of any point in the plane with the crack tip as the origin. The coefficient of this singular term is a measure of the strength of the singularity and is called the stress intensity factor k. It has a unit of ksi $\sqrt{in}$. The two types of crack governing the plane stress and plane strain problems are: (a) Mode I or the opening mode (b) Mode II or the in-plane shear mode (Figure 1). Another type of crack which involves out-of-plane is classified as Mode III crack.

For the two types of crack in the plane stress and plane strain problems, the singular terms of the stress distribution for isotropic materials are given respectively by:

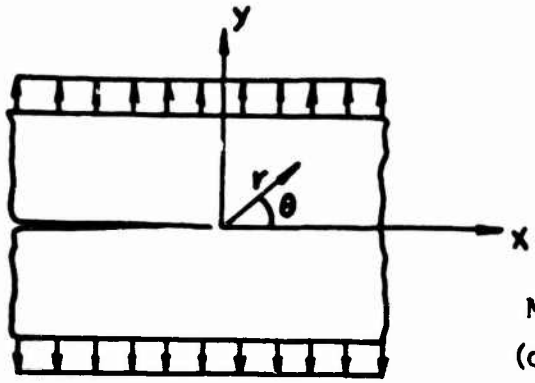
Mode I:

$$\underset{\sim}{\sigma} = \begin{bmatrix} \sigma_x \\ \sigma_y \\ \sigma_{xy} \end{bmatrix} = \frac{k_I}{\sqrt{2r}} \begin{bmatrix} \cos\frac{\theta}{2}\left(1 - \sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right) \\ \cos\frac{\theta}{2}\left(1 + \sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right) \\ \sin\frac{\theta}{2}\cos\frac{\theta}{2}\cos\frac{3\theta}{2} \end{bmatrix} \qquad (1)$$

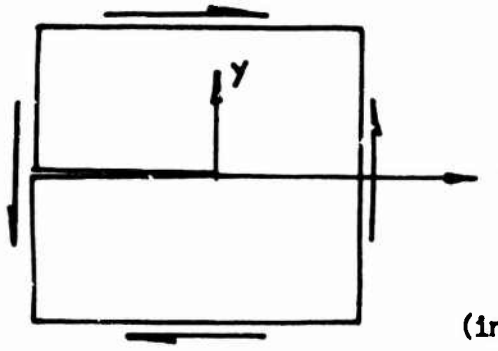
Mode II:

$$\underset{\sim}{\sigma} = \begin{bmatrix} \sigma_x \\ \sigma_y \\ \sigma_{xy} \end{bmatrix} = \frac{k_{II}}{\sqrt{2r}} \begin{bmatrix} -\sin\frac{\theta}{2}\left(2 + \cos\frac{\theta}{2}\cos\frac{3\theta}{2}\right) \\ \sin\frac{\theta}{2}\cos\frac{\theta}{2}\cos\frac{3\theta}{2} \\ \cos\frac{\theta}{2}\left(1 - \sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right) \end{bmatrix} \qquad (2)$$

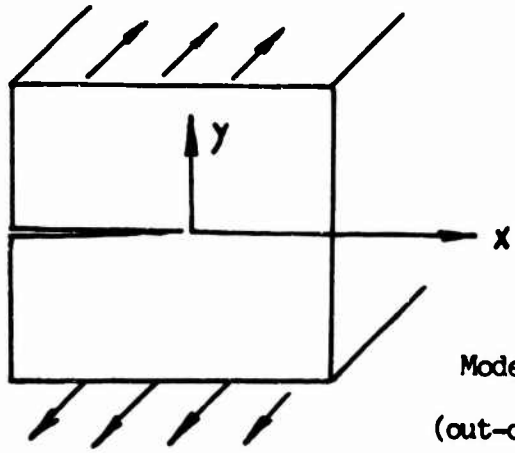
where θ is the angle measured from the x-axis, k_I and k_{II} are directly proportional to the magnitude of applied loading and is dependent on the geometry of the structure, the size and shape of the crack, and the nature of the applied loading. For an infinite sheet with a straight crack of length 2a under uniform tensile stress σ along the direction normal to the crack, the stress intensity factor k_I is equal to $\sigma\sqrt{a}$. The stress intensity factors for a single crack or for a crack array in an infinite sheet under different loading conditions can be determined directly using, for example, the complex function approach. However, for finite structures and for irregular cracks, approximate methods must be employed for evaluating the stress intensity factors. If the stress distribution near the tip of the crack can be estimated, then a fitting



Mode I crack
(opening mode)



Mode II crack
(in-plane shear mode)



Mode III crack
(out-of-plane shear
mode)

Figure 1. Three types of cracks

with a function of $1/\sqrt{r}$ shape will permit the determination of the quantity k.

The displacement field in the vicinity of the crack tip is also related to the stress intensity factors. For the two types of crack in plane stress and plane strain states, the displacement components u and v in the vicinity of the crack tip are given by,

Mode I:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{k_I (2r)^{1/2}}{8G} \begin{bmatrix} (2\kappa-1)\cos\frac{\theta}{2} - \cos\frac{3\theta}{2} \\ (2\kappa+1)\sin\frac{\theta}{2} - \sin\frac{3\theta}{2} \end{bmatrix} \tag{3}$$

Mode II:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{k_{II} (2r)^{1/2}}{8G} \begin{bmatrix} (2\kappa+3)\sin\frac{\theta}{2} + \sin\frac{3\theta}{2} \\ -(2\kappa-3)\cos\frac{\theta}{2} - \cos\frac{3\theta}{2} \end{bmatrix} \tag{4}$$

where G is the shear modulus and κ takes the value (3-4ν) for plane strain state; and, (3-ν)/(1+ν), for plane stress state, ν being the Poisson's ratio. It is seen that the stress intensity factor k can also be evaluated by examining the displacements field near the crack tip. For example, for the Mode I crack a particularly convenient displacement component to be used to determine the quantity k_I is the opening of the crack.

Irwin (Reference 1) has shown that the stress intensity factor is related to the strain-energy release rate $\mathcal{G}$, i.e., the change in strain energy in the structure per unit distance of crack extension. For the two modes in the plane stress and plane strain problems, the relationship is given by

Mode I:

$$\mathcal{G}_I = \frac{\pi(\kappa+1)k_I^2}{8G} \tag{5}$$

Mode II:

$$\mathcal{G}_{II} = \frac{\pi(\kappa+1)k_{II}^2}{8G} \tag{6}$$

It is seen that a third approach for estimating the stress intensity factor is to determine the strain energy release rate. Another quantity which has been shown to be equal to $\mathcal{G}$ is the so-called J-integral, or a line integral involving the strain energy and the displacement gradient around any closed boundary surrounding the crack tip (Reference 2).

The critical value of k for which a crack becomes unstable is designated as k_c. The critical value of k_I is sometimes referred to as the fracture toughness k_{I_c}. The fracture toughness is a material property and is a function of plastic zone behavior at the tip of crack; hence it is related to the ductility of the material. For design purpose, the fracture toughness k_{I_c} is considered as an equivalent or fictitious stress intensity factor corresponding to the point of crack extension instability and is determined by assuming the material as elastic. When the fracture toughness k_{I_c} for a given material and plate thickness has been determined by tests, the estimation of the critical state of stress at the crack extension instability then is reduced simply of the determination of the elastic stress intensity factor k_I. The use of the concept of elastic stress intensity factor in structural design is valid for high strength materials used in aerospace applications such as aluminum alloy, steel and titanium, for which the plastic zone size at the tip of the crack is small in comparison with the crack length.

Elastic stress intensity factors for a number of crack configurations have been catalogued in various texts (Reference 3 and 4). For many configurations which involve cracks in an infinite plate closed form solutions for the stress intensity factor are available. Many of the cases given in Reference 3 were treated by approximate methods such as (1) the boundary collocation of stress functions (2) the conformal mapping technique and (3) continuous dislocation arrays and singular integral equations (Reference 2).

Practical fracture mechanics analysis often calls for the estimation of stress intensity factors for structural elements of irregular geometry and design. For such problems the use of the approximate methods listed above would require considerable algebraic manipulation for the set-up of each individual problem. The finite element method which has the advantage of easy adaptation to modern computer technology for obtaining a complete analysis of complex structures with very simple input instructions certainly becomes the most logic technique for the crack analysis.

The finite element method is also versatile in many other aspects. For example, the extension of the finite element analysis of isotropic materials to that of anisotropic materials usually involves the change of only a few cards in the complete computer program. The extension of stress analysis under mechanical loading to that under thermal loading again does not require a very extensive program revision. For the analysis of the stress distribution near

a crack in a general 3-dimensional solid, the finite element method appears to be the only feasible scheme.

There have been a number of papers and reports dealing with the use of finite element methods for evaluating the elastic stress intensity factor of Mode I cracks (References 5 to 9). All of these are based on the assumed displacement approach. In the solutions of References 5, 6 and 7, the finite elements used at the tip of the crack contain no singularities in strains and stresses. As shown in Reference 10, in such a case, no matter how high the order of the polynomials is used for the assumed displacement functions, the rate of convergence of the finite element solution remains proportional to $\sqrt{\varepsilon}$ where ε is the element size. In order to obtain an accurate evaluation of the stress intensity factor, the size of the elements must be extremely small. Kobayashi and his co-authors (Reference 5), for example, indicated that in order to obtain sufficient accuracy for their assumed displacement finite element method the size of the element near the crack tip must be smaller than one-twentieth of the half-length of the crack. They, in fact, concluded that it is not reliable to obtain the stress intensity factor by fitting the approximate stress distribution obtained by the finite element method, and that, instead, the fitting of the crack surface displacement distribution is a more accurate scheme. This method of using crack surface displacement for determination of the stress intensity factor has also been followed by other authors (References 6 and 7). Anderson et. al. (Reference 6) also pointed out: the use of the energy release rate $\mathcal{G}$ for crack extension is a more efficient way of determining the elastic stress intensity factor. The energy method, however, is not applicable for the cases which involve the combination of Mode I and Mode II cracks (such as an inclined crack under uniform in-plane tensile stress or a crack in a general anisotropic material). The use of crack surface displacements is also not very convenient for problems which involve the singular behavior of both types of crack.

References 8 and 9 concern the use of special elements which contains singularities at the crack tip. Levy et. al. (Reference 9) used a special element which was derived mainly for elastic-perfectly-plastic material, hence it contains an 1/r type strain singularity. Thus, although they included in this report an elastic stress distribution solution obtained by using this special element, only a qualitative discussion of the accuracy for the stress intensity factor was given. Reference 8 involves the use of an assumed strain distribution which contains the correct singularity, but since the assumed displacement does not satisfy the interelement boundary compatibility, the convergence of the solution cannot be guaranteed.

The present paper consists of the formation of the finite element solution of the crack stress intensity problem using special finite elements derived by a procedure similar to the hybrid

stress model (References 11 and 12). In the formulation of the hybrid stress model by Pian, a stress distribution in terms of a number of undetermined stress parameters is assumed within each element, and boundary displacements which are compatible with the neighboring elements are interpolated in terms of the nodal displacements. The stress parameters within each element are independent of those of the other elements and, hence, can be eliminated by the use of the Principle of Minimum Complementary Energy, and the resulting stiffness matrix of the element can be obtained. For the crack problem, special stress terms which represent the correct stress singularity behavior at the crack tip are included in addition to the non-singular stress terms for several elements in the neighborhood of the crack tip. The parameters for the non-singular stress terms which are independent for different elements can be eliminated first but those for the singular terms are kept, together with the nodal displacements, as the unknowns of the final matrix equations. The magnitudes of these singular terms are, in fact, the stress intensity factors k to be determined. Since in this method the singular part of the solution is extracted out in the correct analytical form, the nodal displacements of the finite element analysis correspond to a solution which has no singularity. Thus, the convergence behavior of this method remains the same as that of a common elasticity problem which contains no stress singularity.

Detailed discussion of the formulation of the assumed stress hybrid model for the crack analysis problem and numberical examples for the evaluation of this method are given in the following sections.

SECTION II

FORMULATION OF HYBRID STRESS MODEL

The variational principle which governs the assumed stress hybrid model is a modified principle of minimum complementary energy (References 12 and 13) for which the functional to be varied is

$$\Pi_{mc} = \sum_n \left(\int_{V_n} \frac{1}{2} S_{ijkl}\, \sigma_{ij}\, \sigma_{kl}\, dV - \int_{\partial V_n} T_i u_i\, dS + \int_{S_{\sigma_n}} \bar{T}_i u_i\, dS \right) \quad (7)$$

where

σ_{ij} = stress tensor

S_{ijkl} = elastic compliance tensor

V = volume

S = surface

S_σ = portion of S over which the boundary tractions are prescribed

T_i = component of surface traction

u_i = boundary displacement

n = nth finite element

∂V_n = entire boundary of nth element

The stress tensor σ_{ij} satisfies the equilibrium equations

$$\sigma_{ij,j} + \bar{F}_i = 0 \tag{8}$$

where $\bar{F}_i$ is the prescribed body force.

In the present finite element formulation for which the nature of the stress singularity near the crack tip is known it is a simple matter to include the correct singular behavior of the stress distribution in the assumed approximate functions for σ_{ij} for the elements in the vicinity of the crack tip. For simplicity in the present development we assume that the body forces are not present. The assumed functions for σ_{ij} are then divided into two parts, one of which contains no singularity while the other contains the proper singular behavior. In matrix form the assumed stresses are expressed as

$$\underset{\sim}{\sigma} = \underset{\sim}{P}\,\underset{\sim}{\beta} + \underset{\sim}{P_s}\,\underset{\sim}{\beta_s} \tag{9}$$

where $\underset{\sim}{P}\underset{\sim}{\beta}$ may be simply polynomials which satisfy the homogeneous stress equilibrium equations while $\underset{\sim}{P_s}\underset{\sim}{\beta_s}$ corresponds to the stress singularity terms. For example, $\underset{\sim}{\beta_s} = \{k_I\ k_{II}\}$ are the stress intensity factcrs for the Mode I and Mode II cracks and for isotropic materials the elements in $\underset{\sim}{P_s}$ matrix are the functions for these two types of crack given in Equations 1 and 2. For anisotropic materials the behavior of the stress singularity at the crack tip can also be obtained by the complex variable method (Reference 14). As a reference the singular stress terms for the opening type crack in an orthotropic panel is given in the Appendix. It should be noted that $\underset{\sim}{P_s}\underset{\sim}{\beta_s}$ represents a stress distribution which satisfies not only the equilibrium equations but also the compatibility equations.

The surface tractions for each element are related to the assumed stress distribution and can be expressed as

$$\underset{\sim}{T} = \underset{\sim}{R}\,\underset{\sim}{\beta} + \underset{\sim}{R}_s\,\underset{\sim}{\beta}_s \tag{10}$$

For elements away from the crack tip where the stress singularity effect is no longer important only the $P\beta$ parts appear in Equations 9 and 10. Similar to the procedure used in formulating the regular hybrid stress method the element boundary displacements u_B are interpolated in terms of the nodal displacement q, i.e.

$$\underset{\sim}{u}_B = \underset{\sim}{L}\,\underset{\sim}{q} \tag{11}$$

Substituting of Equations 9, 10 and 11 into Equation 7 and realizing that the singular stress terms are included only in a certain number of elements, say p elements from n = 1 to n = p, in the vicinity of the crack tip, we obtain the following expression

$$\Pi_{mc} = \sum_{n=1}^{p} \left(\frac{1}{2}\underset{\sim}{\beta}^T \underset{\sim}{H}\underset{\sim}{\beta} + \underset{\sim}{\beta}^T \underset{\sim}{H}_s \underset{\sim}{\beta}_s + \frac{1}{2}\underset{\sim}{\beta}_s^T \underset{\sim}{H}_{ss}\underset{\sim}{\beta}_s - \underset{\sim}{\beta}^T \underset{\sim}{G}\underset{\sim}{q} - \underset{\sim}{\beta}_s^T \underset{\sim}{G}_s \underset{\sim}{q} + \underset{\sim}{\bar{Q}}^T \underset{\sim}{q} \right) \tag{12}$$

$$+ \sum_{n=p+1}^{m} \left(\frac{1}{2}\underset{\sim}{\beta}^T \underset{\sim}{H}\underset{\sim}{\beta} - \underset{\sim}{\beta}^T \underset{\sim}{G}\underset{\sim}{q} + \underset{\sim}{\bar{Q}}^T \underset{\sim}{q} \right)$$

where

$$\underset{\sim}{H} = \int_{V_n} \underset{\sim}{P}^T \underset{\sim}{S}\,\underset{\sim}{P}\, dV$$

$$\underset{\sim}{H}_s = \int_{V_n} \underset{\sim}{P}^T \underset{\sim}{S}\,\underset{\sim}{P}_s\, dV$$

$$\underset{\sim}{H}_{ss} = \int_{V_n} \underset{\sim}{P}_s^T \underset{\sim}{S}\,\underset{\sim}{P}_s\, dV$$

(continued on next page)

$$\underset{\sim}{G} = \int_{\partial V_n} \underset{\sim}{R}^T \underset{\sim}{L}\, d\ell$$

$$\underset{\sim}{G_s} = \int_{\partial V_n} \underset{\sim}{R_s}^T \underset{\sim}{L}\, d\ell$$

$$\underset{\sim}{\bar{Q}}^T = \int_{S\sigma_n} \underset{\sim}{\bar{T}}^T \underset{\sim}{L}\, d\ell \qquad (13)$$

and m is the total number of elements.

The stationary conditions of the functional given by Equation 12 with respect to variations of $\underset{\sim}{\beta}$ which are independent from one element to the other then yields

$$\underset{\sim}{H}\underset{\sim}{\beta} + \underset{\sim}{H_s}\underset{\sim}{\beta_s} - \underset{\sim}{G}\underset{\sim}{q} = 0 \qquad \text{for } n \leq p$$

$$\underset{\sim}{H}\underset{\sim}{\beta} - \underset{\sim}{G}\underset{\sim}{q} = 0 \qquad \text{for } n > p \qquad (14)$$

By solving for $\underset{\sim}{\beta}$ from Equation 14 and substituting back into Equation 12 we can express the functional π_{mc} in terms of the generalized displacement $\underset{\sim}{q}$ and the stress intensity factors $\underset{\sim}{\beta_s}$, i.e.

$$\pi_{mc} = \sum_{n=1}^{p} \left(-\frac{1}{2}\underset{\sim}{q}^T\underset{\sim}{k}\underset{\sim}{q} - \underset{\sim}{\beta_s}^T\underset{\sim}{m}\underset{\sim}{q} + \frac{1}{2}\underset{\sim}{\beta_s}^T\underset{\sim}{n}\underset{\sim}{\beta_s} + \underset{\sim}{\bar{Q}}^T\underset{\sim}{q} \right) + \sum_{n=p+1}^{m} \left(-\frac{1}{2}\underset{\sim}{q}^T\underset{\sim}{k}\underset{\sim}{q} + \underset{\sim}{\bar{Q}}^T\underset{\sim}{q} \right) \qquad (15)$$

where

$$\underset{\sim}{k} = \underset{\sim}{G}^T\underset{\sim}{H}^{-1}\underset{\sim}{G}$$

$$\underset{\sim}{m} = \underset{\sim}{G_s} - \underset{\sim}{H_s}^T\underset{\sim}{H}^{-1}\underset{\sim}{G}$$

$$\underset{\sim}{n} = \underset{\sim}{H_{ss}} - \underset{\sim}{H_s}^T\underset{\sim}{H}^{-1}\underset{\sim}{H_s} \qquad (16)$$

By expressing the element nodal displacements q in terms of independent generalized global displacement $\underset{\sim}{q}^*$ and by realizing that $\underset{\sim}{\beta}_s$ is in common for the p element in the neighborhood of the crack tip we can write

$$\pi_{mc} = -\frac{1}{2}\underset{\sim}{q}^* \underset{\sim}{K}\, \underset{\sim}{q}^* - \underset{\sim}{\beta}_s^T \underset{\sim}{M}\, \underset{\sim}{q}^* + \frac{1}{2}\underset{\sim}{\beta}_s^T \underset{\sim}{N}\, \underset{\sim}{\beta}_s + \bar{\underset{\sim}{Q}}^{*T} \underset{\sim}{q}^* \tag{17}$$

where $\underset{\sim}{K}$, $\underset{\sim}{M}$, $\underset{\sim}{N}$ and $\underset{\sim}{Q}^*$ are obtained by assembling the corresponding element matrices.

The stationary condition of the expression of π_{mc} with respect to $\underset{\sim}{q}^*$ and $\underset{\sim}{\beta}_s$ then yields

$$\underset{\sim}{K}\, \underset{\sim}{q} + \underset{\sim}{M}^T \underset{\sim}{\beta}_s = \bar{\underset{\sim}{Q}}^*$$

and

$$\underset{\sim}{M}\, \underset{\sim}{q}^* - \underset{\sim}{N}\, \underset{\sim}{\beta}_s = 0 \tag{18}$$

The solution of Equation 18 then yields the nodal displacements $\underset{\sim}{q}^*$ and the stress intensity factors $\underset{\sim}{\beta}_s$.

A few remarks should be made concerning this hybrid model formulation:

(1) The present formulation leads to a matrix mixed method with nodal displacements and some stress parameters as unknowns. Thus, although this formulation has the advantage of yielding the unknown stress intensity factors directly, it is not compatible with the most commonly used finite element formulation, i.e. the matrix displacement method. The formulation, however, can be easily modified by first grouping the p elements at the crack tip and then eliminating the stress parameters $\underset{\sim}{\beta}_s$ of the singular terms as explicit unknowns among these elements. An element stiffness matrix can thus be evaluated for the super finite element consisting of this group of elements. Thus except that this super element will have much larger number of degrees of freedom than that of the remaining conventional elements, the formulation now becomes a matrix displacement method.

(2) For most crack problems the crack surface is stress free. Thus, for those elements which contain the crack surface, the assumed functions $\underset{\sim}{P}_i$ should be so chosen that the stress free boundary conditions are observed. Such step has been demonstrated to offer considerable improvement in the accuracy of the finite

element solutions by the hybrid stress model particularly when the number of finite elements is small (References 15 and 16). It will be shown in the following section that this step is even more crucial for the present formulation. Since the elastic stress intensity factor refers to the stress distribution at the immediate vicinity of the crack tip it is essential to provide an accurate description of the stresses for the elements near the crack tip.

(3) According to Equations 3 and 4 the displacements u and v corresponding to the stress singularity term should be proportional to the square root of r. This suggests that an appropriate interpolation function of the boundary displacements for the elements in the vicinity of the crack tip is one which includes a $\sqrt{r}$ term.

(4) In order to obtain a converging solution when the element size becomes smaller and smaller the singular stress terms must be included in a fixed region and not in only a certain fixed number of elements. For example, if the singular stress terms is included only in the elements which contain the crack tip, then if the element size approaches zero the stress singularity effect will completely disappear and the solution will certainly not converge to a correct stress intensity factor.

(5) Since $\underset{\sim}{P}_s\underset{\sim}{\beta}_s$ satisfies both compatibility and equilibrium equations, the volume integral for $\underset{\sim}{H}_{ss}$ (Equation 13) can be transformed into a line integral

$$\underset{\sim}{H}_{ss} = \int_{\partial V_n} \underset{\sim}{R}_s \underset{\sim}{L}_s \, d\ell \tag{19}$$

where L_s is a boundary displacement matrix corresponding to $\underset{\sim}{S}\underset{\sim}{P}_s$. In fact, in the case of isotropic materials L_s can be evaluated from Equations 3 and 4. Since $\underset{\sim}{P}_s$ and $\underset{\sim}{L}_s$ are proportional to $1/\sqrt{r}$ and $\sqrt{r}$ respectively the integrand in Equation 19 involves only sine and cosine functions in θ. Thus, although in the volume integral form the integrand in H_{ss} is proportional to 1/r, in the transformed form the singularity disappears.

In the similar manner, since $\underset{\sim}{P}\underset{\sim}{\beta}$ satisfies the equilibrium equations the volume integral for $\underset{\sim}{H}_s$ can also be transformed into a surface integral,

$$\underset{\sim}{H}_s = \int_{\partial V_n} \underset{\sim}{R} \underset{\sim}{L}_s \, ds \tag{20}$$

It is seen that the integrands of $\underset{\sim}{H}_s$ and $\underset{\sim}{G}_s$ are proportional to $\sqrt{r}$ and $1/\sqrt{r}$ respectively. When the integrals in these two

matrices are evaluated numerically special care must be taken to insure the accuracy of the integration when the range of r starts from zero. In the present formulation a transformation of variable is introduced prior to the application of the Gaussian quadrature. For example, for a line segment passing through r=0 the integration is carried out in terms of a variable z defined by $r=z^2$; thus, $dr=2zdz$, $\int rdr=2z^2dz$ and $dr/\sqrt{r}=2dz$.

SECTION III

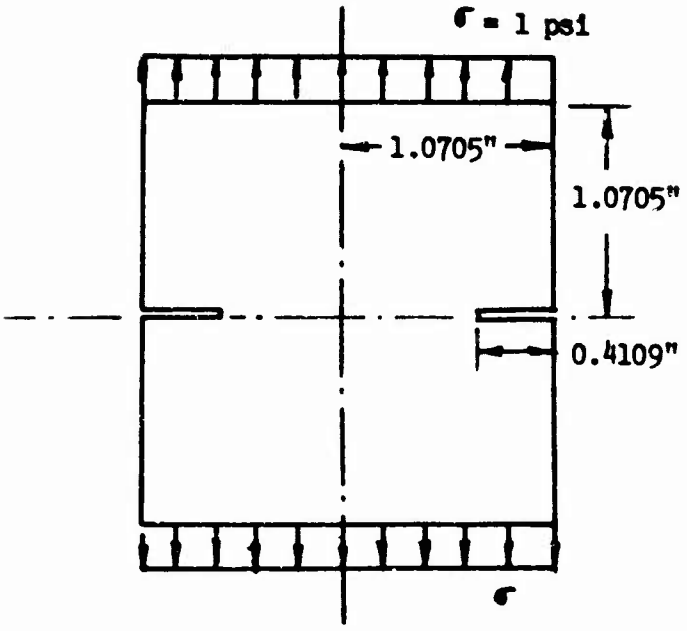
RESULTS OF DETERMINATION STRESS INTENSITY FACTORS

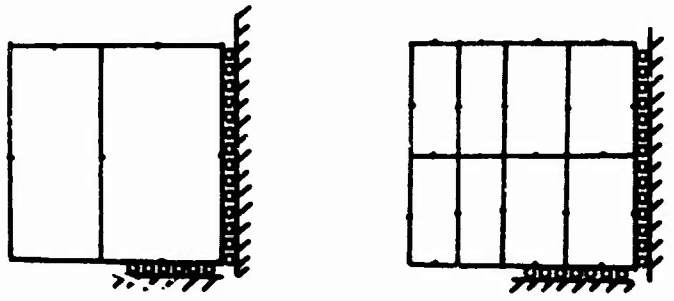
To illustrate the method outlined in this paper several numerical results are presented in this section.

1. The first problem concerns a rectangular panel with symmetric edge cracks under in-plane tension (Figure 2). This problem has been solved by Bowie (Reference 17) using the complex variable and conformal mapping technique. For the present finite element solutions the elements eight-node rectangular elements of various mesh sizes were used. Since the structure possesses double symmetry only one quarter of the panel is needed for the finite element analysis, and the only stress singularity term is the Mode I type. In all the solutions the stress singularity term is included only in the two elements at the crack tip.

In these solutions the non-singular stress terms include complete cubic functions in x and y for σ_x, σ_y and σ_{xy}. This means that 18 β's are used for those elements which do not have stress free boundary conditions or for which such conditions are not enforced. In the case that the stress free boundary conditions are enforced there remains only 10 independent β's. Since the elements used are rectangular elements with three nodes along each side, quadratic displacements are assumed for the boundaries of most of the elements. Only for the two elements at the crack tip the displacements are also assumed to be proportional to $a+b\sqrt{r}+cr$. As we have pointed out earlier, the $\sqrt{r}$ term is included here because it corresponds to the singular stress term.

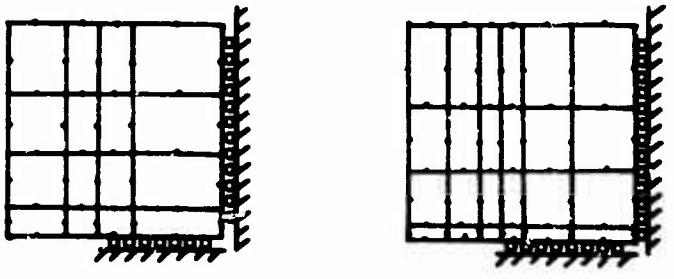
Finite element solutions have been obtained using four different element sizes and arrangements as shown in Figure 2. The values of the stress intensity factor k_I were determined from the finite element solutions by three different methods: (1) by direct calculation of the stress coefficient β_s, (2) by using Equation 3 and considering the crack opening displacement v at the lower left corner of the element at the left of the crack tip, (3) by calculating the strain energy release rate $\mathcal{J}$ from the finite element solutions of two problems with the crack lengths differ by $\Delta\ell=0.05$. Values of k_I have been evaluated for the four element sizes by the first two methods. The third method was





1 x 2

2 x 4



4 x 4

4 x 6

Finite Element Divisions

Figure 2. Tension of rectangular plate with side cracks

applied only to the case of 2x4 mesh. These values are listed in Table I. By comparing with the analytical solution by Bowie we can conclude that all three methods provide about the degree of accuracy. The first method, of course, is the most direct one. The error of k_I obtained by the first method is less than 0.5% when the 4x6 mesh is used. It is only about 6% when the 1x2 mesh is used.

Table I also includes the result of an investigation of the effects of different assumed stresses and boundary displacements on the solution. The effect of not including the $\sqrt{r}$ term in the boundary displacement is apparently not very large although the solution for the case of 4x6 mesh is rather poor when compared with the result obtained when the $\sqrt{r}$ term is included. However, if the stress force boundary conditions for the element at crack tip are not enforced the resulting k_I would be completely unacceptable.

The second example consists of a rectangular isotropic panel (2bx2c) with a center crack of length 2L and four edge cracks of length a as shown in Figure 3. The stress intensity factors k_I and k_{II} for the edge cracks due to uniform tension loading are determined by the finite element method. The effect of the length of the center crack on k_I and k_{II} is investigated. In each solution a quarter of the panel is subdivided into forty-two 8-node elements which include special elements around the tips of the crack. A plot of k_I and k_{II} versus L/c is given in Figure 3.

The third example consists of a rectangular orthotropic panel (2bx2c) with a center crack of length 2L as shown in Figure 4. The crack stress intensity--k_I due to uniform tension loading are determined by the finite element method using twenty-four 8-node rectangular elements for a quarter of the panel. In formulating the special elements at the tip of the crack the stress singularity terms are given by Equation A-3 of the Appendix. Values of k_I are obtained for different ratios of the principal Young's moduli, Ex/Ey. The results are compared with the numerical results obtained by Bowie and Freese (Reference 19) by complex variable approach and extension of the modified mapping-collocation techniques. In determining k_I for the orthotropic panels in Reference 19, the parameter η_1 in Equation A.2 is kept as unity while η_2^2 ($=E_x/E_y$) is left as a variable. For the present finite element analysis the individual orthotropic elastic constants are needed. A reference value of E_x is first chosen, E_y can then be determined by Equation A.2. The Poisson ratio ν_{xy} is then taken as 0.38 and the shear modulus G_{xy} can be calculated. It is seen that the comparison between the present finite element solutions and the complex variable solutions is, indeed, excellent.

TABLE I

STRESS INTENSITY FACTOR k_I FOR A FINITE PLATE WITH SYMMETRIC EDGE CRACKS BY USING SPECIAL EIGHT-NODE HYBRID STRESS ELEMENTS AT THE CRACK TIP

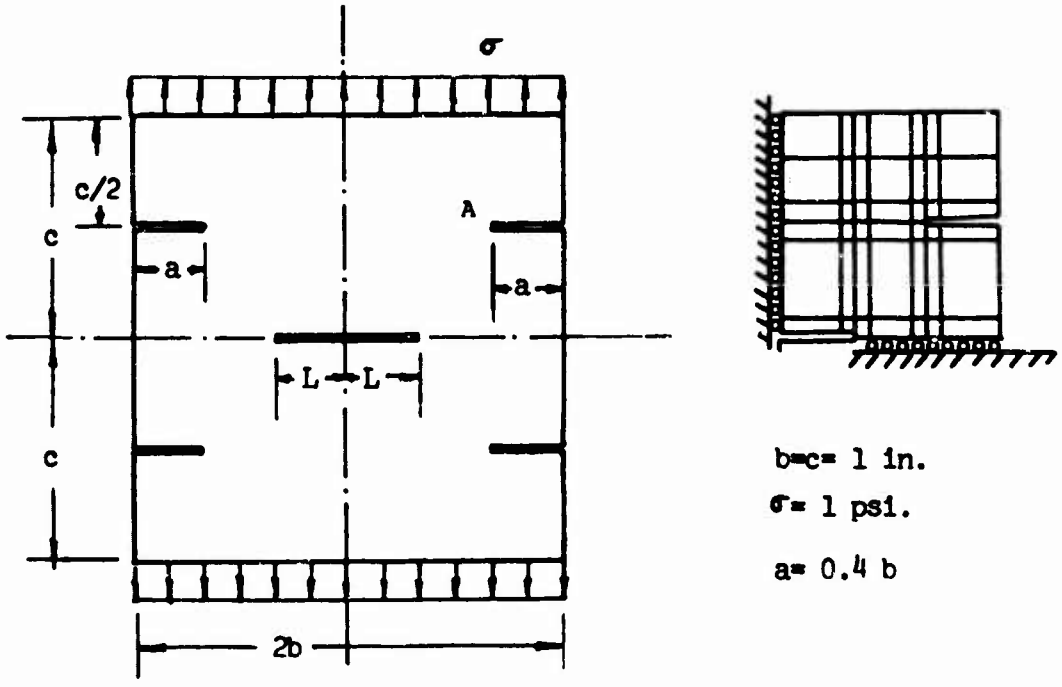
		Stress Intensity Factors k_I				
Assumed Element Stresses		Stress Free State Enforced				Stress free state not enforced
Assumed Boundary Displacements		$\sqrt{r}$ term included			$\sqrt{r}$ term not included	$\sqrt{r}$ term included
Methods for evaluating k_I*		(1)	(2)	(3)	(1)	(1)
Mesh Arrangement	Degrees of freedom					
1 x 2	24	0.840	0.904		0.823	0.508
2 x 4	72	0.822	0.797	0.780		
4 x 4	128	0.812	0.784			
4 x 6	184	0.791	0.790		0.838	0.634
Solution by complex variable method (Bowie)		0.793				

*The three methods used for evaluating k_I from the finite element solutions are:

(1) Direct calculation of stress coefficient

(2) From crack opening displacement v (at $\theta = \pi$)

(3) From strain energy release rate $\mathcal{G}$ (from two finite element solutions of two slightly different crack lengths).



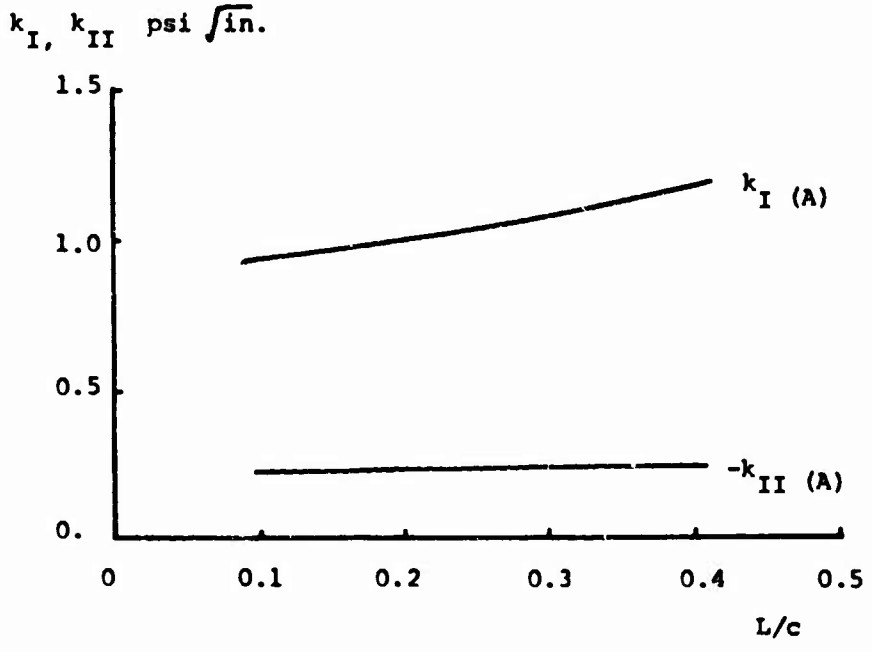


Figure 3. Effect of an adjacent crack on the stress intensity factors of a given crack

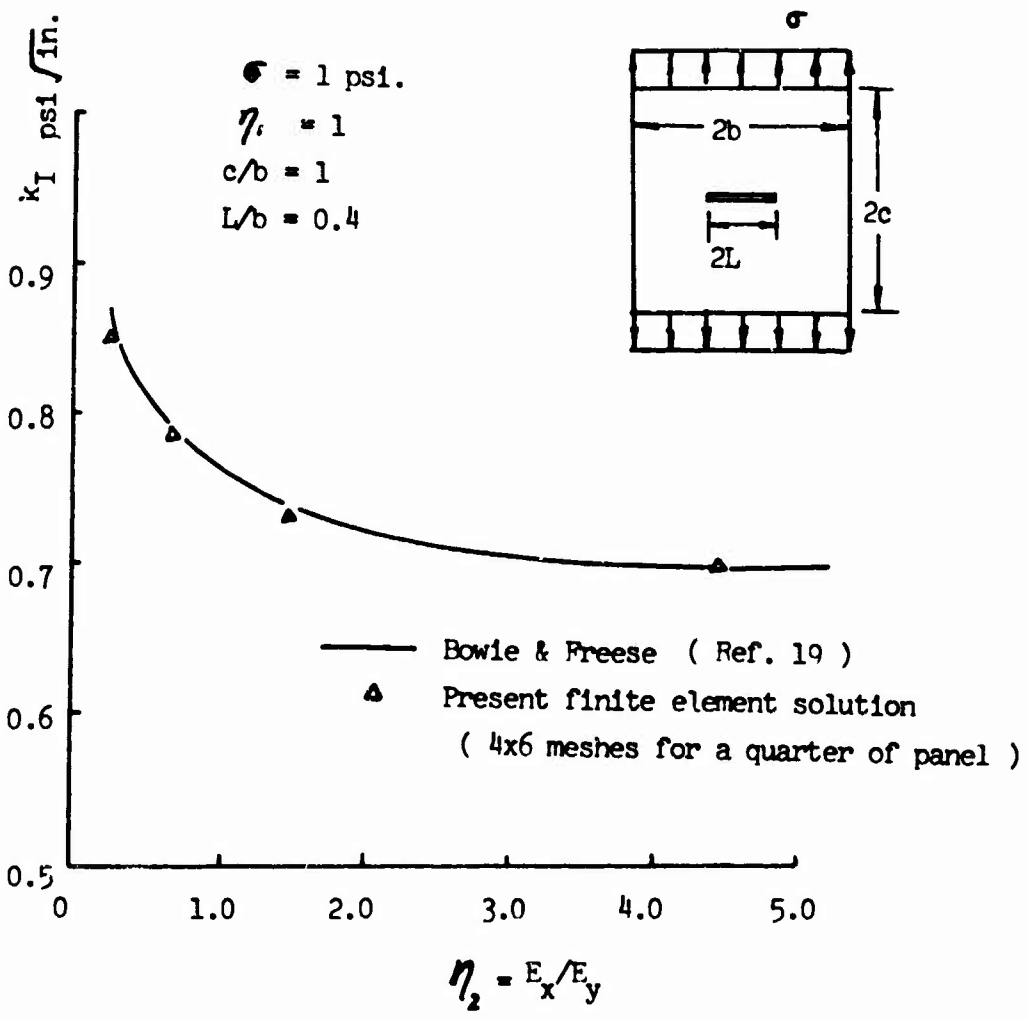


Figure 4

Stress Intensity Factor k_I of
Orthotropic Plate with a Center Crack

SECTION IV

CONCLUSIONS AND REMARKS

The following conclusions and remarks can be made for the present method for evaluating the stress intensity factors of sharp cracks:

(1) The hybrid stress model is an ideal scheme for setting up a finite element method for evaluating the stress intensity factor. It can easily take into account the stress singularity in the elements near the crack tip. It is very convenient to take the correct behavior of boundary displacements and prescribed boundary traction into account.

(2) The numerical results indicate that fairly accurate estimation of the stress intensity factors can be obtained even when the size of the eight-node rectangular element is as large as the length of the crack.

(3) The present formulation leads to a matrix mixed method, but it can be easily modified to fit the conventional finite element programs based on the matrix displacement method.

(4) The method is suitable for isotropic as well as anisotropic materials and is adaptable to problems for which behaviors of Mode I and Mode II cracks exist simultaneously.

(5) For plates and shells with through cracks there also exist stress singularities at the crack tip and corresponding stress intensity factors can be defined. Hartranft and Sih (Reference 18) have concluded that for such problems one must employ the Reissner's refined plate theory which considers the transverse shear effect instead of the ordinary Kirchhoff plate theory. Since the hybrid stress model has been shown to be a convenient method for treating the transverse shear effect (Reference 11) it is also a logic scheme for the problem of through cracks in plates and shells.

SECTION V

REFERENCES

1. Irwin, G. R. "Fracture Mechanics," Structural Mechanics, Proc. First Symposium on Naval Structural Mechanics, Edited by J. N. Goodier and N. J. Hoff, Pergamon Press, 1960, pp. 557-594.

2. Rice, J. R. "Mathematical Analysis in the Mechanics of Fracture," Fracture - An Advanced Treatise, Vol. I, Edited by H. Liebowitz, Academic Press, 1968, pp. 191-311.

3. Wilhem, D. P. "Fracture Mechanics Guidelines for Aircraft Structural Applications," AFFDL-TR-69-111, Feb. 1970.

4. Paris, P. C. and Sih, G. C. "Stress Analysis of Cracks," Fracture Toughness Testing and Its Application, ASTM STP #381 (1965), pp. 30-83.

5. Kobayashi, A. S., Maiden, O. E., Simon, B. J., and Iida, S. "Application of the Method of Finite Element Analysis to Two-Dimensional Problems in Fracture Mechanics," University of Washington, Department of Mechanical Engineering, ONR Contract Nonr-477(39), NR 064 478, TR No. 5, Oct. 1968.

6. Anderson, G. P., Ruggles, V. L. and Stibor, G. S. "Use of Finite Element Computer Programs in Fracture Mechanics," Int. J. of Frac. Mech., Vol. 7, No. 1, March 1971, pp. 63-76.

7. Chan, S. K., Tuba, I. S. and Wilson, W. K. "On the Finite Element Method in Linear Fracture Mechanics," Engineering Fracture Mechanics, 1970, Vol. 2, pp. 1-17.

8. Byskov, E. "The Calculation of Stress Intensity Factors Using the Finite Element Method with Cracked Elements," Int. J. of Frac. Mech., Vol. 6, No. 2, June 1970, pp. 159-167.

9. Levy, N., Marcal, P. V., Ostergren, W. J. and Rice, J. R. "Small Scale Yielding near a Crack in Plane Strain--A Finite Element Analysis," Brown Univ., Div. of Engineering, TR NASA NIGL 40-002-080/1, Nov. 1969.

10. Tong, P. and Pian, T. H. H. "On the Convergence of the Finite Element Methods for Problems with Stress Singularities," to be published.

11. Pian, T. H. H. "Derivation of Element Stiffness Matrices by Assumed Stress Distribution," AIAA J., Vol. 2, No. 7, 1964, pp. 1333-1336.

12. Pian, T. H. H. and Tong, P. "Rationalization in Deriving Element Stiffness Matrix by Assumed Stress Approach," Proc. Second Conference on Matrix Methods in Structural Mechanics, AFFDL-TR-68-150, 1968, pp. 441-469.

13. Pian, T. H. H. and Tong, P. "Basis of Finite Element Methods for Solid Continua," Int. J. Numerical Methods in Engineering, Vol. 1, 1969, pp. 3-28.

14. Sin, G. C. and Liebowitz, H. "Mathematical Theories of Brittle Fracture," Fracture - An Advanced Treatise, Vol. II, Edited by H. Liebowitz, Academic Press, 1968, pp. 67-190.

15. Pian, T. H. H. "Element Stiffness Matrices for Boundary Compatibility and for Prescribed Boundary Stresses," Proc. First Conference of Matrix Methods in Structural Mechanics, AFFDL-TR-66-80, pp. 457-477.

16. Yamada, Y., Nakagiri, S. and Takatsuka, K., "Analysis of Saint-Venant Torsion Problem by a Hybrid Stress Model," Seisan-Kenkyu (Monthly Journal of Institute of Industrial Science, University of Tokyo), Vol. 21, No. 11, 1969, (paper presented at Japan-U.S. Seminar on Matrix Methods of Structural Analysis and Design, Aug. 25-30, 1969, Tokyo).

17. Bowie, O. L. "Rectangular Tensile Sheet with Symmetric Edge Cracks," J. of Appl. Mech. Trans. ASME, Vol. 31, Series E, No. 2, 1964, pp. 208-212.

18. Hartranft, R. J. and Sih, G. C. "Effect of Plate Thickness on the Bending Stress Distribution Around Through Cracks," J. Math and Physics, Vol. 47, No. 3, 1968, pp. 276-291.

19. Bowie, O. L. and Freese, C. E. "Central Crack in Rectangular Sheet with Rectilinear Anisotropy" Proc. Army Symposium on Solid Mechanics, 1970 - Lightweight Structures, AMMRC MS 70-5, Army Materials and Mechanics Research Center, Watertown, Massachusetts, Dec. 1970, pp. IV-1 - IV-14.

APPENDIX

The determination of the near crack tip stress distributions and the stress intensity factors for a general anisotropic material in plane stress or plane strain states has been presented, in details, in a survey paper by Sih and Liebowitz (Reference 14). The stress distributions are functions of two complex parameters s_1 and s_2 which are related to the anisotropic elastic constants. For the orthotropic panel with E_x and E_y as the principle Young's moduli, ν_{xy}, the Poisson ratio and G_{xy}, the shear modulus, the parameters s_1 and s_2 are taken as the roots of the characteristic equation

$$\frac{1}{E_x} s_j^4 + \left(\frac{1}{G_{xy}} - \frac{2\nu_{xy}}{E_x} \right) s_j^2 + \frac{1}{E_y} = 0 \qquad \text{(A-1)}$$

The parameters s_1 and s_2 are purely imaginary, i.e. $s_1 = i\eta_1$ and $s_2 = i\eta_2$, and

$$\eta_1 \eta_2 = (E_x / E_y)^{1/2}$$

$$\eta_1 + \eta_2 = \sqrt{2} \left\{ (E_x/E_y)^{1/2} + E_x/(2G_{xy}) - \nu_{xy} \right\}^{1/2} \qquad \text{(A-2)}$$

The stress field at the tip of a opening crack which is parallel to the x-axis is given by

$$\sigma_x = \frac{k_I}{(2r)^{1/2}} Re \left[\frac{s_1 s_2}{s_1 - s_2} \left(\frac{s_2}{(\cos\theta + s_2 \sin\theta)^{1/2}} - \frac{s_1}{(\cos\theta + s_1 \sin\theta)^{1/2}} \right) \right]$$

$$\sigma_y = \frac{k_I}{(2r)^{1/2}} Re \left[\frac{1}{s_1 - s_2} \left(\frac{s_1}{(\cos\theta + s_2 \sin\theta)^{1/2}} - \frac{s_2}{(\cos\theta + s_1 \sin\theta)^{1/2}} \right) \right]$$

$$\sigma_{xy} = \frac{k_I}{(2r)^{1/2}} Re \left[\frac{s_1 s_2}{s_1 - s_2} \left(\frac{1}{(\cos\theta + s_1 \sin\theta)^{1/2}} - \frac{1}{(\cos\theta + s_2 \sin\theta)^{1/2}} \right) \right] \qquad \text{(A-3)}$$

FINITE ELEMENT ANALYSIS OF PLATE BUCKLING USING A MIXED VARIATIONAL PRINCIPLE*

by

D. J. Allman**

(Royal Aircraft Establishment, Farnborough, UK)

August 1971

A (mixed) variational principle for stresses and displacements is presented from which the governing differential equations of the bifurcational buckling of thin elastic plates are derived. Subsequent finite element analysis, using a linear approximation for the bending moments and a cubic approximation for the displacements within each element, leads to the familiar linear eigenvalue problem for determining the critical intensities of in-plane loading. Numerical solutions are given to a number of simple examples involving square plates using a recently published eigenvalue algorithm which takes advantage of the banded properties of the final matrix formulation. Comparisons are made, where possible, with results from the finite element analyses of other authors and with exact or classical Rayleigh-Ritz solutions.

**Senior Scientific Officer, Structures Department.

1. INTRODUCTION

Although the analysis of plate buckling problems using the finite element method is closely related to the analysis of plate bending it has not received the same enthusiastic attention. The displacement formulation is favoured by most authors[1-5], although their attention is often restricted to rectangular plates[1-3] and in many cases[1,2,4] the principle of minimum potential energy is applied in a manner which is not strictly correct. Convergence to the exact solution with consistent refinement of the finite element mesh size is thus unassured. Examples of the associated difficulties can be seen in the work of Anderson *et al.*[4], as reported by Clough and Felippa[5], in the context of the non-compatible triangular element of Bazeley *et al.*[6]

The development of a satisfactory triangular element is an important step in dealing with the buckling of plates of arbitrary shape. However, the possibility of using compatible triangular elements for this purpose is unattractive. The simplest displacement elements[6,7] use, as unknown parameters, the values of the transverse displacement w and its partial derivatives $\frac{\partial w}{\partial x}$ and $\frac{\partial w}{\partial y}$ at the element corners together with the normal derivatives $\frac{\partial w}{\partial n}$ at the mid-points of the element sides; a total of twelve degrees of freedom for each element. Now, in the interests of computational efficiency, it is desirable to eliminate the element mid-side nodes. This can be achieved quite simply by assuming a linear variation of normal slope along the element boundaries, but unfortunately the resulting triangular element with nine degrees of freedom produces an excessively over-stiff idealisation. In an effort to overcome this difficulty Clough and Felippa[5] derive a quadrilateral 'compound' element which is synthesized from four compatible triangular elements. Solutions to the plate buckling problem using this element show an improvement in accuracy over the former triangular element with nine degrees of freedom, but this is, to some extent, offset by the corresponding increase in computation required in the calculation of the element stiffness matrix.

In the present work, a finite element solution to the plate buckling problem is presented which is based on a mixed variational principle of the type given by Reissner[8]. This approach has also been previously considered by Cook[9] using a mixed formulation[10,11] of the constant bending moment equilibrium element[12,13], but the accuracy of the numerical results was found to be poor. The element 'stiffness' matrix used here is identical to that derived by Allman[12] for plate bending problems, where a linear banding moment field, defined inside a triangular element, is used in conjunction with a cubic displacement field on the element boundary. The so-called element 'geometric stiffness' metrix, which represents the contribution of the in-plane loading to the transverse stiffness of the plate, is calculated using the relatively simple non-compatible cubic displacement field given by Bazeley *et al.*[6] This is shown, in Appendix A, to provide a correct application of an associated mixed variational principle. The unknown parameters of the finite element model are the values of w, $\frac{\partial w}{\partial x}$ and $\frac{\partial w}{\partial y}$ at the corners of the triangular element, so that the connection properties are identical to those of the simple displacement elements[6,7] mentioned previously. However, the elimination of the mid-side nodes does not here affect the numerical accuracy to such a marked degree.

In the numerical solution of the buckling problem, advantage is taken of the banded form of the linear eigenvalue problem, as described by Peters and Wilkinson[14]. This is believed to be a new feature in the solution of plate buckling problems and a brief description of the method is included in the present paper. A number of simple numerical solutions involving uniform plane stress states is calculated to illustrate the accuracy which can be attained using the mixed finite element formulation. These are compared, where possible, to the classical solutions given by Timoshenko and Gere[15] and to results from the finite element analyses of other authors.

Although the numerical examples given in the paper apply specifically to cases of uniform plane stresses, the finite element analysis can also be used to calculate the buckling loads of plates which are subjected to more complex in-plane loading systems. For these problems, the distribution of plane stress resultants in the plate may be calculated to an accuracy which is probably adequate for the buckling analysis using a mesh of compatible triangular elements with linearly varying displacements; the same mesh being subsequently used for the eigenvalue formulation. This procedure, which gives a uniform stress field in each element, has the two-fold advantage of making the calculation of the geometric stiffness matrix a relatively simple matter and ensuring that the homogeneous equations of plane stress equilibrium are satisfied identically in each element, as required by the associated mixed variational principle for plate buckling.

2. A MIXED VARIATIONAL PRINCIPLE FOR THE BUCKLING OF THIN PLATES

The finite element analysis, which is presented in this paper, is based on a mixed variational principle of the type given by Reissner[8]. Here, the governing differential equations of plate buckling, together with their homogeneous kinematic and traction boundary conditions, are derived by considering arbitrary variations of an appropriate mixed functional.

Consider a thin elastic plate of area A, lying in the plane of the rectangular Cartesian coordinates x_i $(i = 1,2)$, which is in equilibrium under a system of applied membrane loads. In the absence of body forces, the distribution of stress resultants λN_{ij} $(i,j = 1,2)$ in the plane of the plate satisfies the homogeneous equations of equilibrium, so that, employing the usual summation convention,

$$\lambda \frac{\partial N_{ij}}{\partial x_i} = 0 \quad , \tag{2-1}$$

where the parameter λ is the ratio of a critical intensity of applied loading to the actual intensity of loading. The critical values of λ are now sought at which the plate can buckle out of the plane $x_1 - x_2$ and still remain in equilibrium under the action of the applied membrane loads.

To avoid ambiguity, the plate is assumed to be singly connected and bounded by a piecewise smooth contour $C = C_T \cup C_k$, C_T and C_k being respectively those parts of the total boundary C on which traction and kinematic boundary conditions are prescribed. The distance s around C is measured anti-clockwise starting from some suitable reference point, while n_i and t_i are the Cartesian components of the unit exterior normal n and the anti-clockwise unit tangential vector t (see Fig.1). On the boundary C the

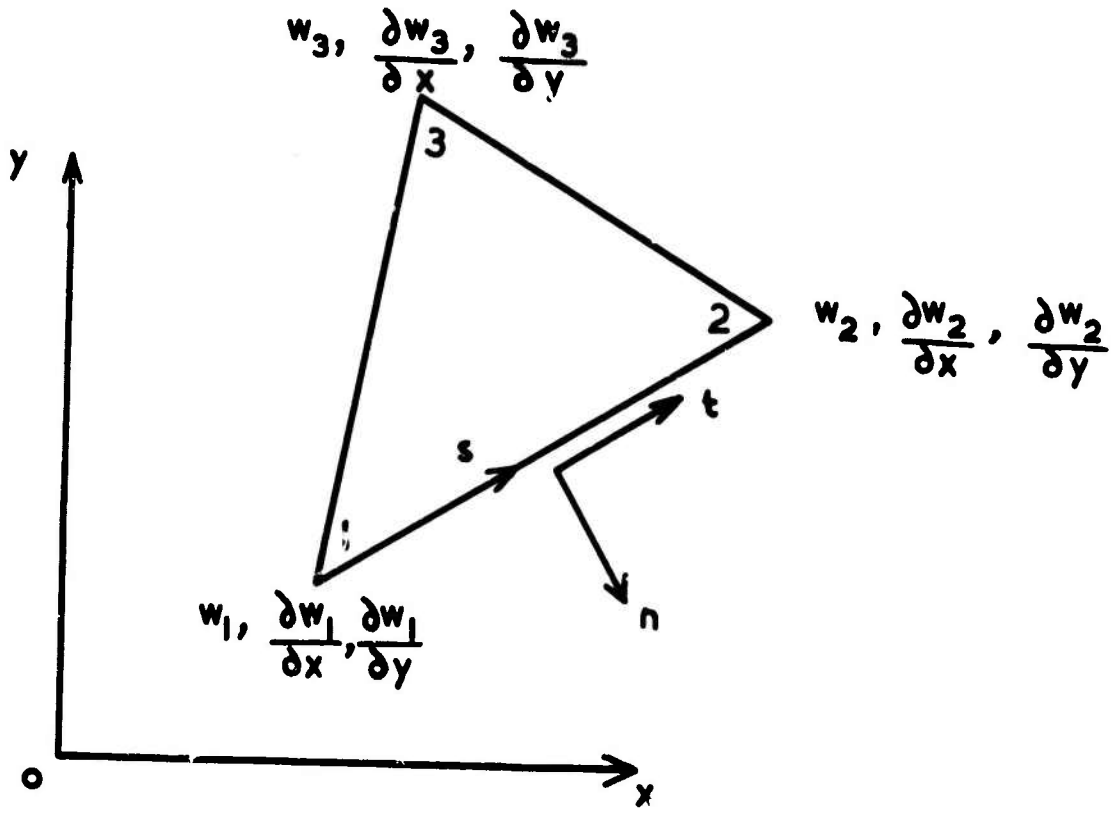


FIG.1 CONNECTION PROPERTIES OF THE MIXED TRIANGULAR ELEMENT

field of bending and twisting moments M_{ij} gives rise to a normal bending moment M_n, a twisting moment M_{ns} and a shearing force Q_n given by

$$\left.\begin{aligned} M_n &= M_{ij}n_in_j \quad , \\ M_{ns} &= M_{ij}n_it_j \quad , \\ Q_n &= \frac{\partial M_{ij}}{\partial x_i}n_j \quad . \end{aligned}\right\} \tag{2-2}$$

The transverse Kirchhoff force is given by

$$V_n = Q_n + \frac{\partial M_{ns}}{\partial s} \quad , \tag{2-3}$$

while at 'corners' N_i which occur on the piecewise smooth contour C, the twisting moment M_{ns} produces concentrated forces

$$R_N = [M_{ns}]_{s^-}^{s^+} \quad , \tag{2-4}$$

where the points s^- and s^+ immediately precede and succeed the corner point N respectively.

Furthermore, because of the rotation of the mid-surface of the plate which occurs in the change from the unbuckled to the buckled position of equilibrium, the distribution of plane stress resultants contributes to the total transverse force on the boundary C. This resultant force is defined as

$$\bar{V}_n = V_n + \lambda N_n \frac{\partial w}{\partial n} + \lambda N_{ns} \frac{\partial w}{\partial s} \quad , \tag{2-5}$$

where λN_n and λN_{ns} are the normal and tangential components of the membrane stress resultants λN_{ij}, on the boundary C, defined by

$$\left.\begin{aligned} \lambda N_n &= \lambda N_{ij}n_in_j \quad , \\ \lambda N_{ns} &= \lambda N_{ij}n_it_j \quad . \end{aligned}\right\} \tag{2-6}$$

Using the above notation, the functional of the mixed variational principle is now defined as

$$\begin{aligned} J = \tfrac{1}{2} &\iint_A \left\{ C_{ijkl}M_{ij}M_{kl} + \lambda N_{ij}\frac{\partial w}{\partial x_i}\frac{\partial w}{\partial x_j} \right\} dA \\ &+ \iint_A \left\{ \frac{\partial^2 M_{ij}}{\partial x_i \partial x_j} + \lambda N_{ij}\frac{\partial^2 w}{\partial x_i \partial x_j} \right\} w dA - \sum_{C_T} [R_N w_N] \\ &- \int_{C_T} \bar{V}_n w ds + \int_{C_T} M_n \frac{\partial w}{\partial n} ds \quad , \end{aligned} \tag{2-7}$$

where C_{ijkl} is the elastic compliance tensor which, for an isotropic plate of thickness h and Young's modulus E, is written

$$C_{ijk\ell} = \frac{1}{D(1-\nu^2)}\left[(1+\nu)\delta_{ik}\delta_{j\ell} - \delta_{ij}\delta_{k\ell}\right] \; ; \qquad (2\text{-}8)$$

ν being Poisson's ratio and $D = \dfrac{Eh^3}{12(1-\nu^2)}$ the flexural rigidity of the plete. The Kronecker delta $\delta_{ij} = 0$ for $i \neq j$, while $\delta_{ii} = 1$ (no sum on i). The functional J has no subsidiery conditions and is to be made stationary by considering arbitrary variations in both the bending moment and displacement fields.

Recelling the Green's formulae

$$\iint_A \frac{\partial^2 \delta M_{ij}}{\partial x_i \partial x_j} w dA - \iint_A \frac{\partial^2 w}{\partial x_i \partial x_j} \delta M_{ij} dA = \sum_C \left[w_N \delta R_N\right] + \oint_C w \delta V_n ds - \oint_C \frac{\partial w}{\partial n} \delta M_n ds \quad \ldots (2\text{-}9)$$

and

$$\iint_A \lambda N_{ij} \frac{\partial w}{\partial x_i} \frac{\partial \delta w}{\partial x_j} dA + \iint_A \lambda N_{ij} \frac{\partial^2 \delta w}{\partial x_i \partial x_j} w dA + \iint_A \lambda \frac{\partial N_{ij}}{\partial x_i} \frac{\partial \delta w}{\partial x_j} w dA$$
$$= \oint_C w \delta \bar{V}_n ds - \oint_C w \delta V_n ds \qquad (2\text{-}10)$$

and noting that the last term on the LHS of equation (2-10) vanishes because of the plane equilibrium equations (2-1), the variational equation

$$\delta J = 0 \qquad (2\text{-}11)$$

is the required stationary condition of equation (2-7)'. The Euler equations and natural boundary conditions of equation (2-11) are the governing equations of the plate buckling problem; these are: the equation of equilibrium of the buckled configuration

$$\frac{\partial^2 M_{ij}}{\partial x_i \partial x_j} + \lambda N_{ij} \frac{\partial^2 w}{\partial x_i \partial x_j} = 0 \; ; \qquad (2\text{-}12)$$

the three constitutive relations

$$\frac{\partial^2 w}{\partial x_i \partial x_j} + C_{ijk\ell} M_{k\ell} = 0 \; ; \qquad (2\text{-}13)$$

the homogeneous traction boundary conditions on C_T,

$$\bar{V}_n = M_n = R_N = 0 \; , \qquad (2\text{-}14)$$

and the homogeneous kinematic boundary conditions on C_k,

$$w = \frac{\partial w}{\partial n} = 0 \; . \qquad (2\text{-}15)$$

3. FINITE ELEMENT ANALYSIS

For the finite element analysis, it is convenient to transform the functional of equation (2-7) into a slightly different form. This is achieved by using the identity of equation (2-10) with δw replaced by w and requiring that the kinematic boundary conditions (equation (2-15)) are satisfied *a priori* on C_k. It is then found that, for a typical element e, the contribution J_e of the element to the total functional J is given (in an unabridged notation) by

$$J_e = \tfrac{1}{2}\iint_{\Delta} \frac{1}{D(1-\nu^2)}\left[(M_x + M_y)^2 + 2(1+\nu)(M_{xy}^2 - M_x M_y)\right] dA$$

$$+ \iint_{\Delta}\left(\frac{\partial^2 M_x}{\partial x^2} + 2\frac{\partial^2 M_{xy}}{\partial x \partial y} + \frac{\partial^2 M_y}{\partial y^2}\right) w\,dA$$

$$- \sum_{C_\Delta} [R_N w_N] - \oint_{C_\Delta} V_n w\,ds + \oint_{C_\Delta} M_n \frac{\partial w}{\partial n}\,ds$$

$$- \frac{\lambda}{2}\iint_{\Delta}\left[N_x\left(\frac{\partial w}{\partial x}\right)^2 + 2N_{xy}\frac{\partial w}{\partial x}\frac{\partial w}{\partial y} + N_y\left(\frac{\partial w}{\partial y}\right)^2\right] dA \quad , \qquad (3\text{-}1)$$

with subsidiary conditions $w = \frac{\partial w}{\partial n} = 0$ on C_k. Here, the surface integrals are calculated over the area Δ of the finite element and the summation and line integrals are evaluated at the element corners and anti-clockwise around the finite element boundary C_Δ respectively. This formulation of the functional is more convenient than that given in equation (2-7) because, as in the conventional stiffness method, the process of numerical calculation of J_e is identical for each element, while the kinematic boundary conditions are applied as a separate operation after the calculation of the global matrices is complete.

The details of a mixed finite element for plate bending have been given previously by Allman[12]. Here, that analysis is extended to deal with plate buckling and, for convenience, some of the previous results are repeated below.

A linear field of bending moments is defined by

$$\begin{bmatrix} M_x \\ M_y \\ M_{xy} \end{bmatrix} = N_1 \beta \quad , \qquad (3\text{-}2)$$

where N_1 is a (3×9) matrix which renders the individual linear fields orthonormal with respect to the strain energy metric of an element and β is a vector of nine undetermined moment parameters $\beta_1, \ldots, \beta_9$. The twelve generalised loads at the boundaries of a triangular element are given by the components of the vector

$$Q = G^T\beta \quad , \tag{3-3}$$

where G^T is a (12 × 9) matrix calculated from the linear bending moment field of equation (3-2). Twelve generalised displacements corresponding to the generalised loads are given by the components of the vector

$$q = TW \quad , \tag{3-4}$$

where T is a (12 × 9) transformation matrix calculated from a compatible displacement field defined on the element boundary in terms of a vector of nine generalised displacements, viz.

$$W^T = \left[w_1, \frac{\partial w_1}{\partial x}, \frac{\partial w_1}{\partial y}, w_2, \frac{\partial w_2}{\partial x}, \frac{\partial w_2}{\partial y}, w_3, \frac{\partial w_3}{\partial x}, \frac{\partial w_3}{\partial y} \right] \quad . \tag{3-5}$$

The subscripts 1, 2, 3 refer to the corners of the triangular element numbered in anti-clockwise order (see Fig.1).

The calculation of the element 'geometric stiffness' matrix involves the determination of the first derivatives of the displacement w inside an element. However, for a mixed finite element analysis, it is unnecessarily restrictive for this purpose to employ a displacement field with both w and $\frac{\partial w}{\partial n}$ continuous across element boundaries. In fact, it is demonstrated in Appendix A that it is adequate to use the simple expression proposed by Bazeley *et al.*[6], denoted here by $\bar{w}$, which violates slope compatibility on the element boundary. The potential energy of the in-plane loads is written, in terms of the non-compatible field $\bar{w}$, as

$$\frac{\lambda}{2} \iint_{\Delta} \left[N_x \left(\frac{\partial \bar{w}}{\partial x} \right)^2 + 2N_{xy} \frac{\partial \bar{w}}{\partial x} \frac{\partial \bar{w}}{\partial y} + N_y \left(\frac{\partial \bar{w}}{\partial y} \right)^2 \right] dA = \frac{\lambda}{2} W^T K_1 W \quad , \tag{3-6}$$

where

$$K_1 = \iint_{\Delta} \left[\frac{\partial \bar{w}}{\partial x}, \frac{\partial \bar{w}}{\partial y} \right] \begin{bmatrix} N_x & N_{xy} \\ N_{xy} & N_y \end{bmatrix} \begin{bmatrix} \frac{\partial \bar{w}}{\partial x} \\ \frac{\partial \bar{w}}{\partial y} \end{bmatrix} dA \tag{3-7}$$

is a (9 × 9) element 'geometric stiffness' matrix (otherwise known as the 'initial stress' matrix[2] or the 'stability coefficient' matrix[1]) and where W is the vector of nine generalised displacements defined previously.

The functional of equation (3-1) can now be written, in matrix notation, as

$$J_e = \tfrac{1}{2}\beta^T\beta - Q^T q - \tfrac{1}{2}\lambda W^T K_1 W \tag{3-8}$$

and, substituting from equation (3-3) and equation (3-4), we have

$$J_e = \tfrac{1}{2}\beta^T\beta - \beta^T(GT)W - \tfrac{1}{2}\lambda W^T K_1 W \quad . \tag{3-9}$$

Considering variations $\delta\beta$ and δW, the contribution of J_e to the total system of equations is

$$\delta J_e = \delta\beta^T[\beta - (GT)W] - \delta W^T[(GT)^T\beta + \lambda K_1 W] \quad , \tag{3-10}$$

and since $\delta\beta$ is quite independent of inter-element continuity,

$$\beta = (GT)W \quad . \tag{3-11}$$

Substituting for β in equation (3-10) and summing over all elements, we find

$$\sum_e \delta W^T[K_0 + \lambda K_1]W = 0 \quad , \tag{3-12}$$

where

$$K_0 = (GT)^T(GT) \tag{3-13}$$

is a (9 × 9) elastic 'stiffness' matrix for the element. Finally, noting the arbitrary nature of δW when the kinematic boundary conditions of equations (2-15) are applied, and defining global matrices by

$$\bar{K}_0 = \sum_e K_0 \; ; \; \bar{K}_1 = \sum_e K_1 \; ; \; \bar{W} = \sum_e W \quad , \tag{3-14}$$

equation (3-12) may also be written as

$$-\bar{K}_1\bar{W} = \mu\bar{K}_0\bar{W} \tag{3-15}$$

where $\mu = \frac{1}{\lambda}$. Here, $\bar{K}_0$ and $\bar{K}_1$ are real banded symmetric matrices; $\bar{K}_0$ being positive definite when the rigid body degrees of freedom are restrained (see Appendix B). The critical intensities of the applied in-plane load are now determined from the eigenvalues μ of the above system of equations where, in view of the inverse relation between μ and λ, the first critical value of λ corresponds to the maximum positive value of μ.

4. DETERMINATION OF EIGENVALUES OF BAND SYMMETRIC MATRICES

The determination of the buckling loads from the finite element formulation presented previously involves the calculation of the eigenvalues of the linear matrix equation (3-15). In general, both of the global matrices $\bar{K}_0$ and $\bar{K}_1$ are real banded symmetric matrices with $\bar{K}_0$ positive definite. However, if equation (3-15) is reduced to the standard symmetric eigenvalue problem by using the Cholesky decomposition of $\bar{K}_0$ given by

$$\bar{K}_0 = LL^T \quad , \tag{4-1}$$

where L is a lower triangular matrix, we obtain

$$-(L^{-1}\bar{K}_1L^{-T})(L^T\bar{W}) = \mu(L^T\bar{W}) \quad , \tag{4-2}$$

where the matrix $(L^{-1}\bar{K}_1L^{-T})$ is full and the advantage of the banded property of the matrices $\bar{K}_0$ and $\bar{K}_1$ is lost.

In a recent paper, Peters and Wilkinson[14] describe how the eigenvalues of equation (3-15) can be calculated in a way which takes account of the intrinsic banded nature of $\bar{K}_0$ and $\bar{K}_1$. Their method is based on the fact that the leading principal minors of the matrix $(\bar{K}_1 + \mu\bar{K}_0)$ form a Sturm sequence. This implies that the number of eigenvalues which is greater than μ is equal to the number of agreements in sign between consecutive members of the sequence of leading principal minors

$$\det\,(\bar{K}_1 + \mu\bar{K}_0)_r \quad , \quad (r = 0,1,\ldots,n) \quad , \tag{4-3}$$

where n is the order of $\bar{K}_0$ and $\bar{K}_1$ and, by definition,

$$\det\,(\bar{K}_1 + \mu\bar{K}_0)_0 = 1 \quad . \tag{4-4}$$

In the computational algorithm, the leading principal minors are not determined explicitly but only the sign of $\det\,(\bar{K}_1 + \mu\bar{K}_0)_{r+1}$ in terms of the sign of $\det\,(\bar{K}_1 + \mu\bar{K}_0)_r$. This is done by comparing the signs of the pivotal elements (the products of which give the minors) of the triangularisation of the matrix $(\bar{K}_1 + \mu\bar{K}_0)$ given by Martin and Wilkinson[16] which is specially designed for the band symmetric case.

The algorithm is, perhaps, best described by considering its use for the calculation of all the eigenvalues μ_i $(i = 1,2,\ldots,M)$ in the interval $a < \mu_i \leq b$. The Sturm sequence counts S_a and S_b are first determined at a and b. This shows that $M = S_b - S_a$ eigenvalues lie in the given range and that a and b can be taken as lower and upper bounds for all of them. To determine each eigenvalue, repeated bisection of the interval combined with the Sturm sequence count is used until upper and lower bounds are found which contain just one eigenvalue. Every time a Sturm sequence count is made, the upper and lower bounds for each eigenvalue are updated. When a single eigenvalue is isolated in this way, an alternative technique with a higher convergence rate is used to locate that eigenvalue to the required accuracy. For this purpose, Peters and Wilkinson[14] recommend the use of a method of successive linear interpolation, and an effective procedure is described in their paper. Notice that this means all multiple eigenvalues are determined entirely by bisection, since an interval is never attained which contains only one such eigenvalue.

In the present work the kinematic boundary conditions are applied in such a way (see Appendix B) that the corresponding eigenvalues are determined as $\mu = -1$. Hence only the positive eigenvalues of equation (3-15) represent a solution to our plate buckling problem, and the initial lower limit can be conveniently chosen as $a = 0$. The choice of a suitable initial upper limit is, however, not straightforward, but a convenient procedure is described by Peters and Wilkinson[14] as follows. The quantity ρ is calculated which is given by

$$\rho = \frac{\|\bar{K}_1\|_\infty}{\|\bar{K}_0\|_\infty} \quad , \tag{4-5}$$

where the ∞-norms of $\bar{K}_0$ and $\bar{K}_1$ are defined as

$$\left.\begin{aligned} \|\bar{K}_0\|_\infty &= \max_i \sum_j |(\bar{k}_0)_{ij}| \ , \\ \|\bar{K}_1\|_\infty &= \max_i \sum_j |(\bar{k}_1)_{ij}| \end{aligned}\right\} \qquad (4\text{-}6)$$

The Sturm sequence is then obtained for

$$\mu = \rho, 4\rho, 4^2\rho, \ldots . \qquad (4\text{-}7)$$

until an upper bound is obtained to the largest eigenvalue (i.e. when $S_b = 0$).

Usually, in practical problems, it is necessary to calculate only the lowest intensity of loading λ at which a plate buckles; in consequence the maximum value of μ is to be determined. In this case the lower limit a is continuously updated to the value of the previous upper limit b during the process of calculating the Sturm sequence of equation (4-7). It is then a matter of simple programming to arrange for the largest eigenvalue to be located by bisection before transferring to the more rapidly convergent interpolation scheme. This process can be extended in an obvious way so that, if desired, several of the lower order critical intensities can be calculated.

5. NUMERICAL RESULTS

The results of a number of simple examples, computed with the mixed element, are given in Tables 1 and 2 for the cases of isotropic square plates subjected to various simple uniform stress distributions along their boundaries. The kinematic boundary conditions are assumed to be either all simply supported or all clamped edges and the finite element solutions are compared with the classical solutions quoted by Timoshenko and Gere[15]. All the examples are calculated using both a (4×4) mesh and an (8×8) mesh over the whole plate, taking advantage of symmetry as shown in Figs.2a, 2b and 2c. Poissor's ratio is taken as $\nu = 0.3$.

Where possible, the results are also compared to those given by Clough and Felippa[5] for their compatible element and to the results quoted by Anderson *et al.*[4] using the non-compatible displacement element of Bazeley *et al.*[6] Although all of these solutions involve the same final number of unknown parameters, it must be remembered that those given by Clough and Felippa involve an additional partial solution to build up their compound element. It is apparent that the results from the mixed element are, for the cases given, mainly superior to the other two. Furthermore, the results obtained from the non-compatible element appear to converge very slowly.

A comparison is also made with the solutions obtained using the more restrictive rectangular elements given by Kapur and Hartz[1], Dawe[2] and Carson and Newton[3]; of these, only the last is a correct application of the principle of minimum potential energy. The results given by Carson and Newton converge more rapidly than those given by Kapur and Hartz and Dawe and are also more accurate than those obtained from the present mixed element. However, their element involves an extra degree of freedom $\left(\text{namely the twist } \frac{\partial^2 w}{\partial x \partial y}\right)$ at each of the element corners.

Table 1

CRITICAL LOADS OF SIMPLY SUPPORTED SQUARE PLATE UNDER UNIFORM PLANE STRESSES

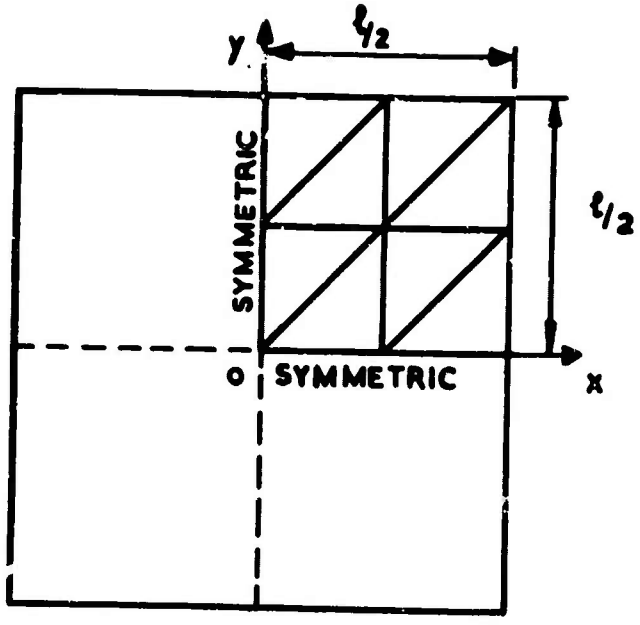
Loading case	Mesh size	Triangular elements			Rectangular elements			Classical solution
		Present analysis	Clough and Felippa	Anderson *et al.*	Kapur and Hartz	Dawe	Carson and Newton	
Uniform uni-axial compression	4 × 4	4.031	4.126	3.72*	3.770	3.978	4.001	4.00
	8 × 8	4.006	4.031	3.94*	3.933	3.993	4.000	
Uniform bi-axial compression	4 × 4	2.016				1.989		2.00
	8 × 8	2.003				1.997		
Uniform shear	4 × 4	10.131				9.481	9.418	9.34
	8 × 8	9.468				-	-	
Multiplier = $\pi^2 D/\ell^2$ for all cases.								

*Results communicated privately by Dr. R. G. Anderson.

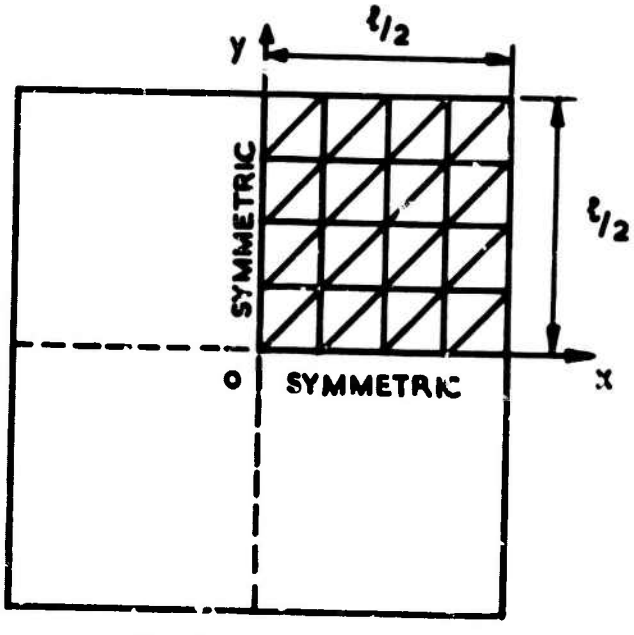
Table 2

CRITICAL LOADS OF CLAMPED SQUARE PLATE UNDER UNIFORM PLANE STRESSES

Loading case	Mesh size	Triangular elements			Rectangular elements			Classical solution
		Present analysis	Clough and Felippa	Anderson *et al.*	Kapur and Hartz	Dawe	Carson and Newton	
Uniform uni-axial compression	4 × 4	10.990		9.30	9.284	10.147		10.07
	8 × 8	10.252		-	9.782	10.065		
Uniform bi-axial compression	4 × 4	5.602	5.625	5.043	4.975		5.327	5.31
	8 × 8	5.356	5.399	5.119	5.160		5.305	
Uniform shear	4 × 4	17.382					15.043	14.71
	8 × 8	15.172					-	
Multiplier = $\pi^2 D/t^2$ for all cases.								

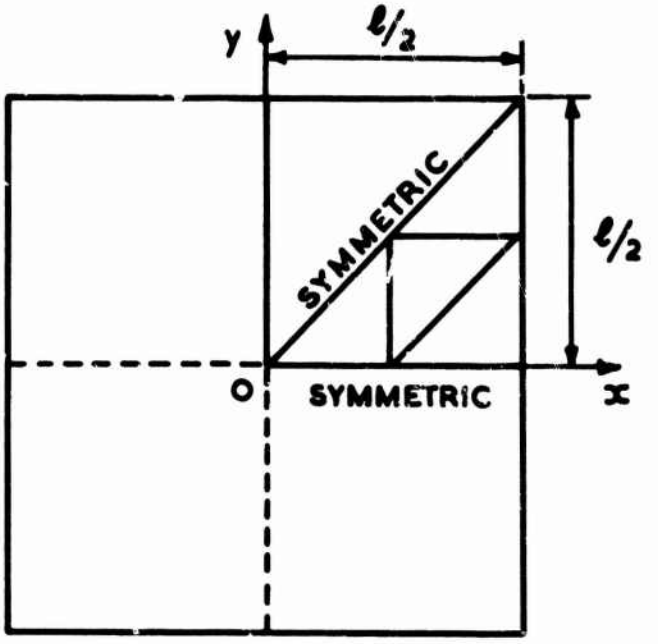


(4×4) MESH

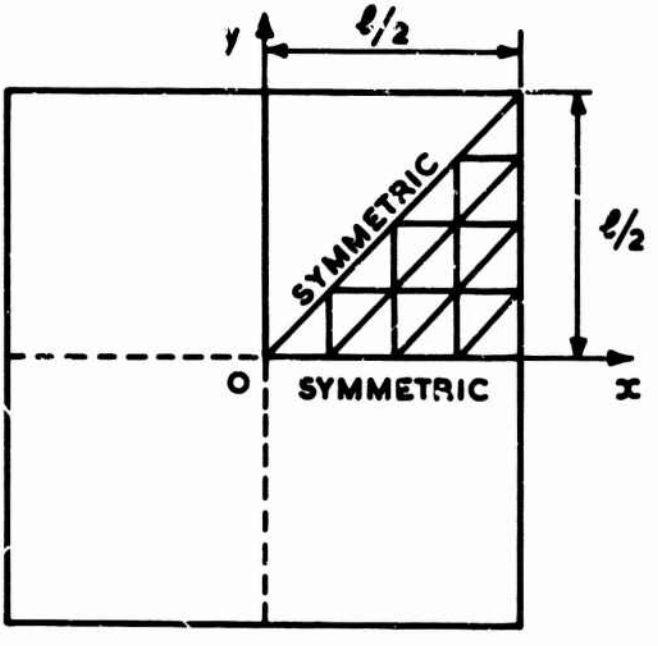


(8×8) MESH

FIG. 2a MESH ARRANGEMENT FOR UNIFORM UNI-AXIAL COMPRESSION OF A SQUARE PLATE.

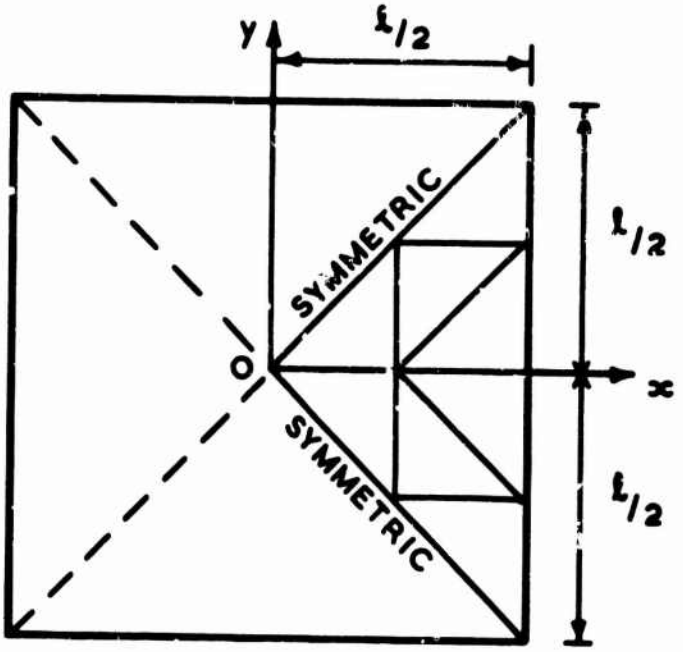


(4x4) MESH

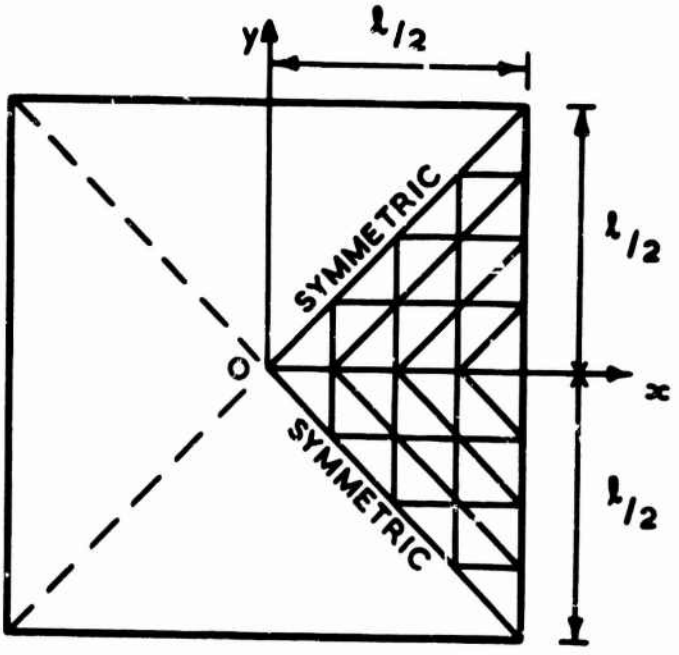


(8x8) MESH

FIG. 2b MESH ARRANGEMENT FOR UNIFORM BI-AXIAL COMPRESSION OF A SQUARE PLATE



(4 x 4) MESH



(8 x 8) MESH

FIG. 2c MESH ARRANGEMENT FOR UNIFORM SHEAR OF A SQUARE PLATE

6. CONCLUSIONS

It is shown that a variational principle for stresses and displacements can be used to derive the governing equations of the plate buckling problem. Solutions to simple numerical examples, using a finite element scheme based on this mixed variational principle, indicate that very satisfactory accuracy can be achieved for the critical states of in-plane loading. In the solution of the final eigenvalue problem use is made of a recently published algorithm which takes account of the banded nature of the final matrix formulation.

REFERENCES

No.	Author(s)	Title, etc.
1	K.K. Kapur B.J. Hartz	Stability of plates using the finite element method. J. Eng. Mech. Div., Proc. ASCE, Vol.92, No.EM2, April 1966, pp.177-195
2	D.J. Dawe	Application of the discrete element method to the buckling analysis of rectangular plates under arbitrary membrane loading. Aero. Quart., Vol.XX, May 1969, pp.114-128
3	W.G. Carson R.E. Newton	Plate buckling analysis using a fully compatible finite element. AIAA Jnl., Vol.7, No.3, March 1969, pp.527-529
4	R.G. Anderson B.M. Irons O.C. Zienkiewicz	Vibration and stability of plates using finite elements. Int. J. Solids and Structs, Vol.4, No.10, October 1968, pp.1031-1055
5	R.W. Clough C.A. Felippa	A refined quadrilateral element for analysis of plate bending. Proceedings of the Second Conference on Matrix Methods in Structural Mechanics, AFFDL-TR-68-150, December 1969, pp.399-440
6	G.P. Bazeley Y.K. Cheung B.M. Irons O.C. Zienkiewicz	Triangular elements in plate bending - conforming and non-conforming solutions. Proceedings of the First Conference on Matrix Methods in Structural Mechanics, AFFDL-TR-66-80, November 1966, pp.547-576
7	R.W. Clough J.L. Tocher	Finite element stiffness matrices for analysis of plate bending. Proceedings of the First Conference on Matrix Methods in Structural Mechanics, AFFDL-TR-66-80, November 1966, pp.515-545
8	E. Reissner	On a variational theorem in elasticity. J. Maths and Phys., Vol.29, 1950, pp.90-95
9	R.D. Cook	Eigenvalue problems with a 'mixed' plate element. AIAA Jnl., Vol.7, No.5, May 1969, pp.982-983
10	L.R. Herrmann	Finite element bending analysis for plates. J. Eng. Mech. Div., Proc. ASCE, Vol.93, No.EM5, October 1967, pp.13-26
11	K. Hellan	Analysis of elastic plates in flexure by a simplified finite element method. Acta Polytechnica Scandinavica, Civ. Eng. and Bldg. Constr. Series, No.46, 1967

REFERENCES (Contd)

No.	Author(s)	Title, etc.
12	D.J. Allman	Triangular finite elements for plate bending with constant and linearly varying bending moments. IUTAM Symposium on High Speed Computing of Elastic Structures, Liège, 1970
13	L.S.D. Morley	The constant bending moment plate bending element. J. Strain Analysis, Vol.6, No.1, January 1971, pp.20-24
14	G. Peters J.H. Wilkinson	Eigenvalues of $Ax = \lambda Bx$ with band symmetric A and B. Computer Jnl., Vol.12, 1969, pp.398-404
15	S.P. Timoshenko J.M. Gere	Theory of elastic stability. McGraw-Hill, New York (second edition 1961), Chapter 9, pp.348-439
16	R.S. Martin J.H. Wilkinson	Solution of symmetric and unsymmetric band equations and the calculation of eigenvectors of band matrices. Num. Math., Vol.9, 1967, pp.279-301
17	L.S.D. Morley	Bending of a square plate with a central square hole. RAE Technical Report 69031 (1969)

Appendix A

CALCULATION OF THE GEOMETRIC STIFFNESS MATRIX USING A NON-COMPATIBLE DISPLACEMENT FIELD

In section 3 it is found convenient to calculate the element geometric stiffness matrix using the simple displacement field proposed by Bazeley *et al.*[6] which satisfies continuity of only the transverse displacement on the element boundaries. It is shown below that the use of this approximation still represents a valid application of a mixed variational principle.

Let the non-compatible displacement field inside an element be denoted by $\bar{w}$, while on the element boundary a second displacement field, denoted by w, is defined which satisfies the kinematic requirements of both displacement and normal slope continuity. The contribution J_e of an element e to the total functional J is now assumed to be given by (using the summation convention with $i,j = 1,2$)

$$J_e = \iint_\Delta \left\{ \tfrac{1}{2} C_{ijk\ell} M_{ij} M_{k\ell} + \frac{\partial^2 M_{ij}}{\partial x_i \partial x_j} \bar{w} \right\} dA - \sum_{C_\Delta} [R_N w_N]$$

$$- \oint_{C_\Delta} V_n w ds + \oint_{C_\Delta} M_n \frac{\partial w}{\partial n} ds - \frac{\lambda}{2} \iint_\Delta N_{ij} \frac{\partial \bar{w}}{\partial x_i} \frac{\partial \bar{w}}{\partial x_j} dA \quad , \qquad \text{(A-1)}$$

with subsidiary condition $w = \bar{w}$ on C_Δ. Considering variations in both the bending moment and displacement fields it is found that

$$\delta J_e = \iint_\Delta \left\{ \left(\frac{\partial^2 \bar{w}}{\partial x_i \partial x_j} + C_{ijk\ell} M_{k\ell} \right) \delta M_{ij} + \left(\frac{\partial^2 M_{ij}}{\partial x_i \partial x_j} + \lambda N_{ij} \frac{\partial^2 \bar{w}}{\partial x_i \partial x_j} \right) \delta \bar{w} \right\} dA$$

$$- \sum_{C_\Delta} [R_N \delta w_N] - \oint_{C_\Delta} \left(V_n + \lambda N_n \frac{\partial \bar{w}}{\partial n} + \lambda N_{ns} \frac{\partial \bar{w}}{\partial s} \right) \delta w ds$$

$$+ \oint_{C_\Delta} M_n \delta \left(\frac{\partial w}{\partial n} \right) ds + \oint_{C_\Delta} \left(\frac{\partial w}{\partial n} - \frac{\partial \bar{w}}{\partial n} \right) \delta M_n ds \quad . \qquad \text{(A-2)}$$

The stationary condition $\sum_e \delta J_e = 0$ thus provides the governing equations of the buckling problem and, in particular, the last term of equation (A-2) restores the condition of slope continuity $\frac{\partial w}{\partial n} = \frac{\partial \bar{w}}{\partial n}$ on the element boundary.

Appendix B

APPLICATION OF HOMOGENEOUS LINEAR CONSTRAINTS TO BANDED MATRIX EQUATIONS

A general and convenient procedure for applying non-homogeneous linear constraints to banded matrix equations is given by Morley[17]. This procedure can be simplified considerably for the plate buckling problem, which is formulated in terms of the banded global matrices $\bar{K}_0$ and $\bar{K}_1$, where the linear constraints correspond to the application of homogeneous kinematic boundary conditions given by equation (2-15).

It is required to solve the eigenvalue problem

$$\delta\bar{W}^T(\bar{K}_0 + \mu\bar{K}_1)\bar{W} = 0 \quad , \tag{B-1}$$

where the variations $\delta\bar{W}$ are not arbitrary, but are subject to homogeneous linear constraints imposed by the kinematic boundary conditions

$$P\bar{W} = \Gamma\bar{W} \quad . \tag{B-2}$$

It is always possible, by suitable row and column interchanges, to arrange that P is a non-singular tri-diagonal matrix (see Morley[17]), whereas Γ is a diagonal matrix with unit and zero elements on the leading diagonal; the zero elements corresponding to the boundary values. As an example, equation (B-2) is written out in full for typical homogeneous linear constraints, viz.

$$\begin{bmatrix} \ddots & & \\ & P_{i,i} & P_{i,i+1} \\ & P_{i+1,i} & P_{i+1,i+1} \\ & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ \bar{W}_i \\ \bar{W}_{i+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & & \\ & 0 & \\ & & 0 \\ & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ \bar{W}_i \\ \bar{W}_{i+1} \\ \vdots \end{bmatrix} \quad . \tag{B-3}$$

Incorporation of equation (B-2) into equation (B-1) leads to the result

$$\delta\bar{W}^T(P^{-1}\Gamma)^T(\bar{K}_0 + \mu\bar{K}_1)(P^{-1}\Gamma)\bar{W} = 0 \quad , \tag{B-4}$$

where $\delta\bar{W}$ is now an arbitrary quantity which may be deleted from equation (B-4). It is apparent, from the form of equation (B-3), that the calculation of the inverse matrix P^{-1} requires the inversion of submatrices of order 2 at most and that P^{-1} is, moreover, a tri-diagonal matrix. Consequently, the matrix operations of pre-multiplication by $(P^{-1}\Gamma)^T$ and post-multiplication by $(P^{-1}\Gamma)$ neither increase the bandwidth nor disturb the symmetry of the global matrices $\bar{K}_0$ and $\bar{K}_1$, and it is possible to take advantage of the banded

symmetric form of equation (B-4) in computer programs. The effect of these matrix operations is to replace the elements of certain rows (and corresponding columns) in equation (B-1) by zero elements. This makes the matrix $\bar{K}_0$ singular. In the present work, in order that the eigenvalue algorithm of Peters and Wilkinson[14] may be used, the matrix $\bar{K}_0$ is made positive definite by substituting a unit element in the leading diagonal position of a zero row (or column); an identical substitution is also made in $\bar{K}_1$. All the eigenvalues of equation (B-4) which correspond to the homogeneous linear constraints of equation (B-2) now take the values $\mu = -1$. Thus, in our plate buckling problem, only the positive eigenvalues of equation (B-4) correspond to the critical intensities of applied in-plane loading. These are determined as described in section 4.

APPLICATION OF FINITE ELEMENT METHOD FOR CONTINUUM MECHANICS PROBLEMS

Arup Chattopadhyay* and A.V. Setlur**

General nonlinear equations for a continuum are formulated such that they are convenient for discretization by the finite element method. The resulting set of nonlinear algebraic or ordinary differential equations are linearized and solved by the parametric differentiation method. The scope of application of the procedure is very wide. Numerical examples are presented to indicate the scope. The results, in general, show excellent agreement with known solutions for one step or short range processes. For long range solutions, the incremental step becomes a critical factor.

*Post Doctoral Fellow, Department of Civil Engineering, Laval University, Quebec, Canada, formerly, Research Scholar, Department of Civil Engineering, Indian Institute of Technology, Kanpur, India

**Assistant Professor, Department of Civil Engineering, Indian Institute of Technology, Kanpur, India

INTRODUCTION

The major tasks in the development of a procedure for solving problems of a wide class may be listed as:

a) The formulation of the basic equations for this wide class of problems in a variational, differential or integro-differential form. The form of the equations should be convenient for the envisaged method of solution.

b) The evaluation of physical constants appearing in the field equations. Even though general theories being rapidly put forward, their use is inhibited by the lack of available experimental values. Also, experimental procedures have to keep pace with sophisticated theories so that the experimental errors do not nullify the accuracy of the formulation.

c) The development of efficient mathematical and numerical tools for solving the equations.

In this paper, an attempt is made to solve a wide class of problems in the field of continuum mechanics. Attention is focused on tasks (a) and (c). It is assumed that values of physical constants would be available in some form.

The general field equations of continuum mechanics has been derived from thermodynamic principles and the hypothesis of material frame indifference[1]. The procedure for formulation follows the concepts awarded by Green and Rivlin[2]. Monopolar and dipolar stress fields and temperature field are considered.

Constitutive equations which should follow certain basic axioms[3] form the most difficult part of the general formulation. These equations bring out the particular behavior of a given material under prescribed environmental conditions. In this study, these equations are kept in a general form and, later, in the illustrative examples, particular forms are specified.

Finite element method is employed to annihilate the space dependency and express the field equations and constitutive relations in terms of chosen field variables at discrete points. The field equations are put in a form convenient for finite element idealization. It should be noted that it is not necessary to have the equations in variational form for use of finite element method[4]. The discretized algebraic or differential equations, depending on whether or not time derivatives are involved, are, in general, simultaneous and nonlinear. This obviates the use of any direct method of solution. Parametric differentiation method[5,6,7] is employed to obtain a stepwise solution.

Two numerical examples are presented to indicate the generality of the approach.

GOVERNING EQUATIONS

The complete mathematical statement describing the state of an ideal deformable medium can be expressed through

a) law of conservation of mass

b) thermodynamic laws

c) constitutive relations

d) environmental stipulations expressed as boundary and initial conditions on loads, displacements and heat flow.

The variables commonly employed to describe these changes of state are: stress and displacement components, temperature, entropy and heat flow. A complete description covering a large class of problems and materials would be rather cumbersome to handle besides being intractable even for some of the simpler cases. Hence, in this paper, simplifying assumptions are made which would yet be general for a fairly wide class of problems.

Assumptions

1. The hypothesis of material frame indifference is applicable.
2. Concept of force defined through work rate principle is valid.
3. Displacement derivatives up to second order will be active for the internal energy expression (Grade II material)
4. Supply of non mechanical energy is only due to thermal effects.
5. Multipolar kinetic energy and heat flux has been neglected.
6. Antisymmetric parts of dipolar stresses i.e. $T^{|kj|i}$ are non-existent.

7. Only monopolar and dipolar stress fields are considered.

Under these restrictions, the foregoing laws would be invoked to arrive at the final set of equations.

Since the derivation of these equations is rather lengthy, only a few salient steps would be indicated. The detailed derivations could be found in Reference 8.

All quantities are referred in terms of the undeformed curvilinear coordinate system X^i or the initial position vector $\bar{X}$ (at time $t = 0$). Any point X with position vector $\bar{X}$ will be denoted after deformation as x and its position vector as $\bar{x}$ where

$$\bar{x} = \bar{x}(\bar{X},t) \tag{1}$$

The deformed position vector $\bar{x}$ can also be expressed in terms of the deformation vector $\bar{u}$, such that

$$\bar{x} = \bar{X} + \bar{u} \tag{2}$$

where $\bar{u} = \bar{u}(\bar{X},t)$ and $\bar{u}(\bar{X},o) = 0$

Material Frame Indifference[1,2]:

Any event has to be specified with respect to a reference frame for space as well as time. This specification is not unique and depends upon different observers. But the fundamental measurable quantities such as distances, angles and time intervals are independent of the observer. Consequently, any change in frame must preserve them along with the temporal orders of all events. The most general relation for such a change of frame may be expressed as,

$$\bar{x}^*(\bar{X},t^*) = \bar{R}(t)\ \bar{x}(\bar{X},t) + \bar{D}(t)\ ;\ t^* = t - a \tag{3}$$

where $(\bar{x}^*,t^*)$ and $(\bar{x},t)$ represent the same event in different frames, $\bar{D}(t)$ is a constant vector, $\bar{R}(t)$ is a proper orthogonal tensor and a is a real number. Two motions of a given medium related by Eq.(3) are said to be equivalent.

For a change of frame,

(a) a scalar remains unchanged

(b) a vector $\bar{V}$ transforms to $\bar{V}^*(\bar{X},t^*) = \bar{R}(t)\ \bar{V}(\bar{X},t)$ (4)

(c) a tensor $\bar{\bar{S}}$ transforms to $\bar{\bar{S}}^*(\bar{X},t^*) = \bar{\bar{k}}(t)\ \bar{\bar{S}}(\bar{X},t)\ \bar{\bar{R}}^{-t}(t)$

(5)

Functions whose values are scalars or tensors will be called frame indifferent, if both dependent and independent variables transform according to the above laws.

Definition of Forces:

The quantitative definition of forces is not unique nor straight forward, though it can be conveniently derived through the concept of work rate. The basic idea is that if P^i is a component of vector, and V_i is an arbitrary velocity component and if the scalar P^iV_i is the rate of work corresponding to the velocity vector V, then P^i may be defined as force in the direction of X^i. In continuum mechanics, this idea may be readily extended to define forces of more general nature[2,8]. Considering only the monopolar and dipolar cases, the work rate for body forces F^i and F^{ij} per unit mass, surface forces P^i and P^{ij} per unit area and stresses T^{ij} and T^{ijk} per unit area normal to axis X^i are given respectively as,

$$\rho_o(F^iV_i + F^{ij}V_{j|i}),\ P^iV_i + P^{ij}V_{j|i} \text{ and } T^{ij}V_j + T^{ijk}V_{k|j} \tag{6}$$

where ρ_o is the initial mass density and bar indicates covariant differentiation.

Basic Laws:

An element of a continuum will obey the following three laws through any stage of evolution[1,9].

a) Conservation of mass: The mathematical expression for conservation of mass may be expressed as

$$\rho(\bar{x},t) = \rho_o(\bar{X},0) \ /\ \left|\frac{\partial x^k}{\partial X^L}\right| \tag{7}$$

This relation may be taken as a definition of density, $\rho(\bar{x},t)$, at time t and for the point $\bar{x}$.

b) Second law of thermodynamics: This law asserts the existence of a quantity called entropy, η, such that

$$\eta = \eta(e,\ v_\sigma,\ \bar{\bar{T}}) \tag{8}$$

where e is the internal energy density and ν_σ are a certain number of mechanical parameters influencing the motion.

The absolute temperature T is defined through entropy as

$$\frac{1}{T} = \frac{d\eta}{de} \tag{9}$$

If dQ^* is the increase of heat energy at temperature T, then the increase of entropy can be written as,

$$d\eta = d\eta_i + \frac{dQ^*}{T} \tag{10}$$

where $d\eta_i$ is the production of entropy inside the system.

The second law asserts that for any process in nature, the production of entropy inside the system cannot be negative, i.e.

$$d\eta_i \geq 0 \tag{11}$$

where the equality stands for reversible process and the inequality for irreversible process.

The second law does not provide any additional equation but restricts the form of the constitutive relations.

c) Balance of energy: This law is expressed as

$$K^o + E^o = P + F \tag{12}$$

where K and E are the kinetic and internal energies respectively, P and F are the power supplied by external mechanical forces and the total non-mechanical power, in the present study, due to heat. Eq. (12) may also be termed as the first law of thermodynamics.

Considering an arbitrary material volume, v, of a continuum bounded by a surface, s, having a unit normal, $\bar{n}$, in the reference configuration, Eq. (12) can be written as

$$\rho_o(V^iV_i + e)dv = \rho_o(q + F^iV_i + F^{ij}V_{j|i})dv + (h + P^iV_i + P^{ij}V_{j|i})ds \tag{13}$$

Letting the arbitrary internal volume shrink to a point, and then applying the frame indifference principle, it can be shown that Eq. (13) reduces to

$$P^i = n_j T^{ji} \; ; \; p^{ji} = n_k T^{kji} \text{ and } h = n_i Q^i \tag{14}$$

Substituting relations (14) into Eq. (13) and transforming surface integrals to volume integrals and again applying the same principle, it is found that,

$$T^{ki}{}_{|k} + P_o F^i - \rho_o V^i = 0 \tag{15}$$

$$t^{*(ji)} = 0 \tag{16}$$

$$\rho_o \ell = t^{*(ji)} V_{i|j} + T^{i(kj)} V_{i|jk} + \rho_o q + Q^i{}_{|i} \tag{17}$$

where $t^{*ji} = T^{kji}{}_{|k} + \rho_o F^{ji} + T^{ji}$

In Eqs. (15), (16) and (17), the volume integral has been dropped, since these relations are true for any arbitrary volume. Eqs. (15) and (16) are the Cauchy's laws of motion whereas Eq. (17) has to be satisfied by constitutive relations for a medium.

If sufficient smoothness requirements are assumed, Eq. (18) may be inverted to give

$$e = e(\eta, \nu_\sigma, \bar{X}) \tag{18}$$

Since Eq. (17) involves terms upto second derivative of the mechanical parameters, expression (18) can be prescribed in the form

$$e = e(\eta, E_{ij}, E_{ijk}, \bar{X}) \tag{19}$$

with the restriction on E_{ij} and E_{ijk} that the scalar function e should be invariant under rigid body rotation.

One of the several ways is to prescribe the classical strain tensor for E_{ij} and Toupin's strain gradient[10] for E_{ijk}, i.e.

$$E_{ij} = u_{i|j} + u_{j|i} + u_{k|i} u^k{}_{|j}$$

$$\text{and} \quad E_{ijk} = E_{ij|k} + E_{ik|j} - E_{jk|i} \tag{20}$$

The dependence of the density function e on $\bar{X}$ accounts for the non-homogeneity of the material.

Comparison of the time derivatives of e in Eqs. (17) and (18) yields,

$$(t^{*(ij)} - \hat{T}^{(ij)})\, A_{ij}\, dv = 0 \tag{21}$$

$$(T^{(kj)i} - \hat{T}^{(kj)i})\, V_{i|jk}\, dv = 0 \tag{22}$$

$$\rho T\dot{n}^{o} = {}_{d}\hat{T}^{(ij)}\, A_{ij} + {}_{d}\hat{T}^{(kj)i}\, V_{i|jk} + \rho_o q + Q^{i}{}_{|i} \tag{23}$$

where, $A_{ij} = (V_{i|j} + V_{j|i})/2$,

the dissipative part of the stresses,

$${}_{d}\hat{T}^{(ij)} = t^{*(ij)} - {}_{e}\hat{T}^{(ij)} \; ; \; {}_{d}\hat{T}^{(kj)i} = T^{(kj)i} - {}_{e}\hat{T}^{(kj)i}$$

and the thermodynamic tensions,

$${}_{e}\hat{T}^{ij} = \rho_o \frac{\partial e}{\partial u_{j|i}} \; ; \; {}_{e}\hat{T}^{(kj)i} = \rho_o \frac{\partial e}{\partial u_{i|jk}}$$

Here, prefixes e and d are used to denote recoverable and irrecoverable parts of the stresses. Eqs. (15), (16), (21), (22) and (23) constitute a set of 40 equations in 31 unknowns and hence do not form a determinate set. However, the part of the stress $T^{|kj|i}$ does not contribute to the equations of motion nor produces any work[8], although it may have an important role for the boundary conditions[2]. For the sake of simplicity, this part will be assumed to be non existent within the body or on the bounding surfaces. This makes the set determinate.

Eq. (23) may be put in a slightly different form as:

$$(\rho_o \overset{o}{T}\overset{o}{T} \frac{d}{dt}(\frac{\partial \psi}{\partial T}) + \rho_o \overset{o}{T} q - \overset{o}{T}_{|i} Q^{i} + W_d \overset{o}{T})dv + \eta_i \overset{o}{T} Q^{i}\, ds = 0 \tag{24}$$

where $W_d = {}_{d}\hat{T}^{(ji)} A_{ij} + {}_{d}\hat{T}^{(kj)i}\, V_{i|jk}$ and $\psi = \eta T$ representing Helmholtz free energy. Eq. (24) furnishes the thermal equation of state.

Constitutive Equations:

The general physical laws in themselves do not suffice to determine the deformation of a body subjected to given loading. A vast amount of literature is available on the mathematical derivations of constitutive equations based on the concepts of irreversible thermodynamics[11,12] as well as those obtained from considering the differential type[13] and the axiom of fading memory[14]. Any discussion on their general nature is beyond the scope of the present paper.

To solve any problem, it is necessary to prescribe the constitutive relations for:

$$ {}_{e}\hat{T}^{(ij)},\ {}_{e}\hat{T}^{(kj)i},\ {}_{d}\hat{T}^{(ij)},\ {}_{d}\hat{T}^{(kj)i},\ \psi_{\alpha} \text{ and } Q^{i}. $$

One of the forms may be of the following type:

$$ \hat{T}^{(ij)} = \bar{T}^{(ij)} \cdot \left(\frac{\partial E_{ij}}{\partial u_{i|j}} + \frac{\partial E_{ij}}{\partial u_{j|i}} \right) /2 \tag{25} $$

$$ \hat{T}^{(kj)i} = \bar{T}^{(kj)i} \cdot \left(\frac{\partial E_{ijk}}{\partial u_{i|jk}} \right) \tag{26} $$

$$ \psi = \bar{\psi}(E_{ij},\ E_{ijk},\ A_{ij},\ A_{ijk},\ T,\ \overset{o}{T},\ t^{P}) \tag{27} $$

$$ Q^{i} = \bar{Q}^{i}(E_{kl},\ E_{klm},\ T_{|k},\ t^{P}) \tag{28} $$

where $\bar{T}^{(ij)}$ and $\bar{T}^{(kj)i}$ are Piola type stresses,

A^{i}s are the Rivlin-Eriksen tensors[3]

and t^{P} denotes the dependence of the function on their history of their arguments for $t \geq t^{P} \geq -\infty$ where t is the present value of time.

The two illustrative examples presented at the end use particular forms of Eqs. (25) to (28).

Summary:

Eq. (14) may be combined with Eq. (15) and (16) to yield:

$$(T^{ji}{}_{|j} + \rho_o F^i - \rho_o \overset{o}{V}{}^i) V_i dv + (P^i - \eta_j T^{ji}) V_i ds = 0 \tag{29}$$

$$(T^{ji} - t^{*(ji)} + T^{(kj)i}{}_{|k} + \rho_o F^{ji}) V_{i|j} dv$$

$$- (P^{ji} - \eta_k T^{(kj)i}) V_{i|j} ds = 0 \tag{30}$$

Eqs. (24) to (30) form the final set of equations. The final form is quite convenient for discretization by the finite element method.

The surface integrals in Eqs. (24), (29) and (30) would include the work done by the unbalanced surface stresses or heat flux vector. Thus, the discontinuity in stresses or heat flux vector would be taken into account in 'the mean'.

SOLUTION TECHNIQUE

Once the general equations are obtained, it is a matter of technique to get a solution out of them for a definite problem. Theoretically, it is possible, but the amount of complexities encountered for the exact solution are virtually prohibitive. Hence, finite element method will be employed here, since it may be considered as one of the most powerful discretization techniques available among the approximate techniques. Application of this method annihilates the space dependency and expresses the field equations in terms of the chosen field variables at discrete points. The resulting equations, in general, will be nonlinear and simultaneous which obviate the use of any direct method of solution. Hence, a parametric differentiation scheme[5,6,7,8] will be applied to obtain a stepwise solution.

Parametric differentiation technique is a systematic and straightforward 'marching' approach. The method consists of differentiating the governing equations with respect to a chosen parameter, λ, (say load) reducing the original set of nonlinear equations to linear equations in the differentials. The reduced set can be solved straightaway using any integration technique for a small step increment through a quadrature formula. This way the solution will march out from λ to the required value of the parameter, say λ_1. In essence, the nonlinearity of the problem is restricted only in the quadrature manipulation which virtually poses no difficulty. The method can be best demonstrated by taking a simple example Say it is required to solve the nonlinear differential equation,

$$\dot{C}\dot{U} + L = 0 \qquad \text{(A)}$$

with boundary condition

$$U = U_{t_o} \quad \text{at} \quad t = t_o \qquad \text{(B)}$$

where, $C = C(U,\dot{U})$, $L = L(t)$, $U = U(t)$ and $\dot{U} = dU/dt$

Choose the forcing parameter λ such that

$$\bar{L} = \bar{L}(t,\lambda) \text{ where } L(t) = \bar{L}(t, \lambda_1)$$

and $$\bar{U} = \bar{U}(t,\lambda) \text{ where } U(t) = \bar{U}(t,\lambda_1)$$

Consequently, it is possible to construct

$$\bar{C} = \bar{C}(\bar{U}(t,\lambda), \dot{\bar{U}}(t,\lambda))$$

where, $C(t) = \bar{C}(\bar{U}(t,\lambda_1), \dot{\bar{U}}(t,\lambda_1))$

Hence, the new differential equation can be written as,

$$\bar{C}\,\dot{\bar{U}} + \bar{L} = 0 \qquad \text{(C)}$$

with boundary condition,

$$\bar{U}(t_o,\lambda) = U_{t_o} \qquad \text{(D)}$$

The function $\bar{U}$ has to be chosen in such a way that it satisfies (C) and (D) at $\lambda = \lambda_o$. Now differentiation of (C) and (D) with respect to λ yields,

$$\frac{\partial \bar{C}}{\partial \bar{U}}\,\bar{U}^* + \left(\frac{\partial \bar{C}}{\partial \dot{\bar{U}}} + \bar{C} \right) \dot{\bar{U}}^* + \bar{L}^* = 0 \qquad \text{(E)}$$

and $$\bar{U}^*(t_o,\lambda) = 0 \qquad \text{(F)}$$

where, $\bar{U}^* = \frac{\partial \bar{U}}{\partial \lambda}$ and $\bar{L}^* = \frac{\partial \bar{L}}{\partial \lambda}$

The modified equation (E) and (F) is linear with respect to $\bar{U}^*$ and with variable coefficients. It can be solved, at least numerically, and the solution will be:

$$\bar{U}^*(t,\lambda_o).$$

From this, $\bar{U}(t,\lambda_o + \Delta\lambda)$ can be easily obtained by the integration,

$$\bar{U}(t,\lambda_o + \Delta\lambda) = \bar{U}(t,\lambda_o) + \int_{\lambda_o}^{\lambda_o+\Delta\lambda} \bar{U}^*(t,\lambda_o)\,d\lambda \qquad \text{(G)}$$

In similar way, the solution can be marched upto $\lambda = \lambda_1$ when $\bar{U}$ will be the same as U which is the solution of (A) and (B). The advantages of this method are that it allows the solution to be obtained throughout the range of λ and no iterative technique is needed.

For efficient execution of the quadrature in (G) the following two forward-integration methods are recommended:

1. Improved Euler-Cauchy method
2. Runge-Kutta-Gill method.

Both schemes are easily available in any text of numerical analysis.

APPLICATIONS

To show the application of the above general procedure, two simple problems have been illustrated with the assumptions that acceleration terms are neglected and dipolar stress field is not existent. Modified Euler method has been used for integration in the following problems.

Problem 1: The first problem has been selected as the heat conduction across a half space (second Danilovskaya) for a coupled thermoelastic illustration. This is solved considering rectangular finite element model having linear distribution of variables in the coordinate directions. The variables are two displacements, three stresses and temperature. For simplicity, Fourier law of heat conduction and only large rotation but small strains are considered here.

The finite element solution and the values as given by Nickell and Sackman[15] have been compared in Fig. 1. For the comparison, the following values are assumed:

$$P = 0.008 \ , \ \sigma = 15.3 \times 10^{-6} \ , \ PC_v = 0.96 \ , \ k = 0.04 \ ,$$

$$\lambda = 42.0 \times 10^{4} \ , \ u = 4.0 \times 10^{4}$$

where P, σ, C_v, k, λ and u are respectively mass density, linear thermal expansion coefficient, specific heat per unit mass, thermal conductivity and isothermal Lame's coefficients. The convective heat transfer across the surface is assumed to as, $Q^i = b(T-T_o)$ where, direction i is normal to the boundary surface, b is the film constant which is taken as 0.78, T is the temperature of the surface and T_o is the outside ambient temperature, which is 300.0. The boundary surface temperature is prescribed to vary linearly from 300 at t = 0 to 600 at $t = 1.45 \times 10^{-10}$ and then it is kept constant.

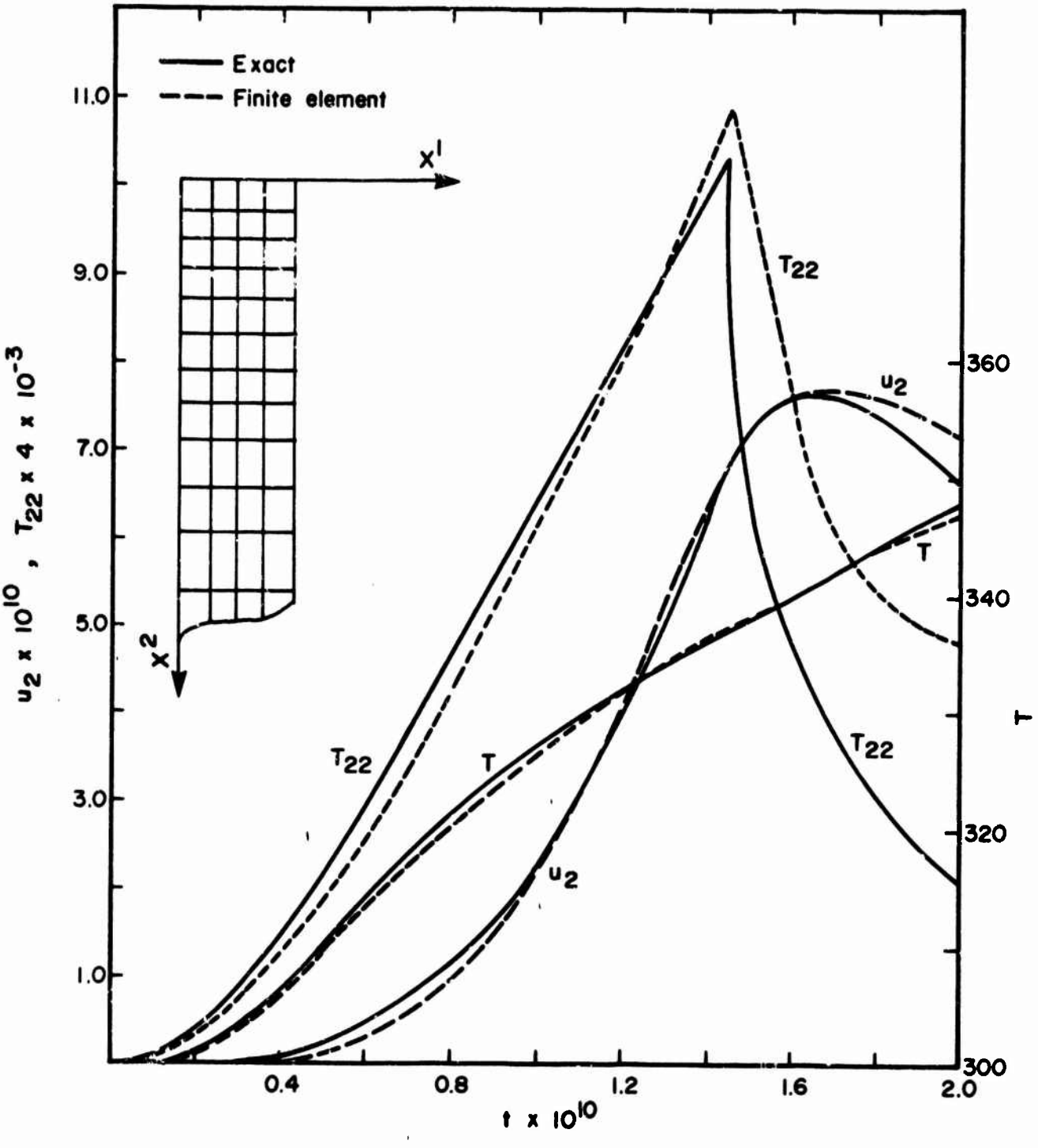


Fig. 1 - THERMAL LOADING ON AN ELASTIC HALF SPACE.

It is seen that u_2 and T agree satisfactorily up to $t = 1.60 \times 10^{-10}$. After that, they start deviating rather sharply. But, even before $t = 1.45 \times 10^{-10}$, the accuracy of stress is rather poor, the maximum error being about 8%.

Problem 2: A viscoelastic cylinder stiffened by an external elastic encasing with ablating inner surface has been chosen as the second example. This problem is analyzed using ring shaped element having triangular cross section with linear distribution of displacements in cylindrical polar coordinate system. The Young's modulus and the poisson's ratio of the elastic encasing are respectively 10^6 and 0.25. The ablating rate for the inner surface has been taken as [16] $a(t) = 25/(1-3t/4)$, where a(t) is the inner radius at time t. The pressure variation with respect to time is assumed to be $p(t) = 1 - e^{-5t}$. The constitutive relations for the viscoelastic material are prescribed as,

$$(T_{11}, T_{22}) = \int_0^t J(t-t') \quad \dot{E}_{11} + \dot{E}_{22} \quad dt'$$

$$+ \int_0^t K(t-t') \quad \dot{E}_{11}, \dot{E}_{22} \quad dt'$$

and $$T_{33} = \int_0^t J(t-t') \quad \dot{E}_{11} + \dot{E}_{22} \quad dt'$$

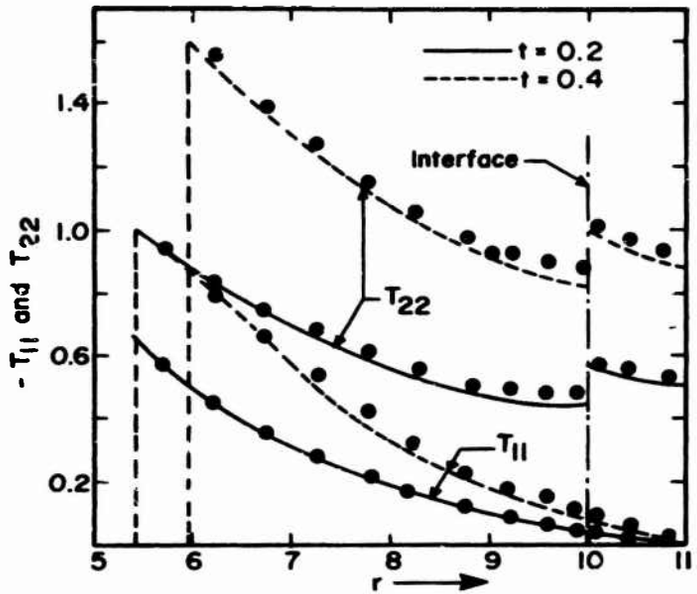
where K(t) and J(t) for this problem are taken as,

$$K(t) = 82.0 + 9282.0\, e^{-1.126t} \quad \text{and}$$

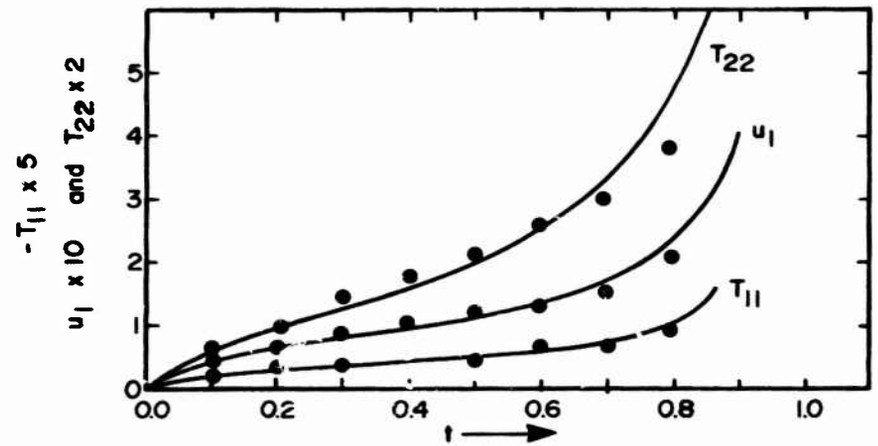
$$J(t) = 12472.7 - 3094.0\, e^{-1.126t}$$

Due to ablation of inner surface, at every new integration step, the arrangement of the element has to be modified. This is done by reducing the inner most column of elements and interpolating the previous values of strains at the centre of gravity of new reduced elements. The integrations necessary for obtaining the material coefficients have been achieved by trapezoidal rule.

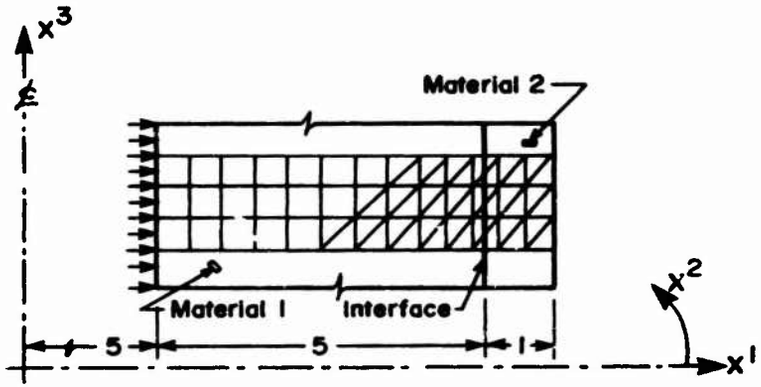
The results obtained using finite element method are compared in Fig. 2 with analytical solution presented by Ting[16]. It may be observed that stresses and displacement at the interface are agreeable up to t = 0.6. Also, the stress distributions are quite satisfactory except near the interface.



B. STRESS DISTRIBUTION IN CYLINDER



C. HISTORY AT INTERFACE



A. INITIAL ARRANGEMENT OF ELEMENTS

FOR V.E. MATERIAL I AT t = 0

Y.M. = 1.25×10^4, P.R.=0.25

FOR ELASTIC ENCASING

Y.M. = 10.0×10^4, P.R.=0.25

● FINITE ELEMENT SOLUTION

——— EXACT

Fig. 2 - ANALYSIS OF VISCOELASTIC ENCASED CYLINDER

CONCLUSION

In this paper, following the concept forwarded by Green and Rivlin[2], the general nonlinear equations for a continuum are derived. These equations are directly applied to finite element discretization of space variable and parametric differentiation for linearization.

The scope of application of these equations is very wide. With minor modifications, they can be applied to any particular case of continuum problems, for example multipolar cases, viscoelasticity, viscoplasticity, coupled thermoelasticity with dissipative properties etc. The application to problems of nonlinear stability or complicated fluid flow with thermal effects may also be possible.

In this study, utilizing this general technique, two particular cases of simpler types have been solved, though many more have been given in Reference 8. The results, in general, show excellent agreement for short range processes. For long range solutions, the incremental step length becomes a critical factor. Unfortunately, due to the limitation in available computer time, no further refinement in this direction has been possible. From the limited experimentation it appears that the approach offers no more complexities than the numerical solution of nonlinear initial value problems. Moreover, Davidenko[6] has shown that the convergency of the process can be improved considerably by employing more efficient integration technique, such as Runge-Kutta procedure. It may be concluded that the finite element discretization coupled with parametric differentiation procedure for nonlinear equations forms a powerful tool for solving complex continuum mechanics problems.

NOMENCLATURES

All quantities in this study are defined with respect to the coordinate system at undeformed state.

A_{ij}, A_{ijk}	Rivlin-Ericksen tensors for mono and dipolar fields
E_{ij}, E_{ijk}	Strains
e	Internal energy density
F^i, F^{ij}	Body forces
G^{ij}, g^{ij}	Metric tensors for undeformed and deformed state
h	Heat flux
K^{ij}	Conductivity tensor
n_k	Components of unit outward normal
P^i, P^{ij}	Surface loads
Q^i	Heat flux vector
q	Distributed energy sources inside the body per unit mass
T^{ij}, T^{ijk}	Stresses
$_dT^{ij}$, $_dT^{ijk}$	Dissipative parts of the stresses
T	Temperature
$T^{(ij)}$ etc.	Functionals for stresses
$\hat{T}^{(ij)}$ $T^{\|ij\|}$	Symmetric and antisymmetric parts of T^{ij}
t	Time
u_i, V_i	Displacements and velocities
$V_{i\|j}$	Velocity gradient where bar denotes covariant derivative

X^i, $\bar{X}$	Curvilinear coordinates and position vector
x^i, $\bar{x}$	Coordinates and position vector for deformed state
α	Indices
η	Entropy
ν_α	Thermodynamic affinities
ρ_o, ρ	Mass densities at undeformed and deformed state
τ^α	Thermodynamic stresses
ψ	Helmholtz free energy

REFERENCES

1. Truesdell, C.A. and Toupin, R.A., "The Classical Field Theories", Vol. 3, Part 1, Encyclopedia of Physics, Springer, Berlin.

2. Green, A.E. and Rivlir, R.S., "Simple Force and Stress Multipoles", Arch. Rat. Mech. Anal. 16, 1964, 325-353.

3. Eringen, A.C., Mechanics of Continua, Wiley, New York, 1967.

4. Oden, J.T. and Kross, D.A., "Analysis of General Coupled Thermo-elasticity Problems by the Finite Element Method", Proc. Conf. on Matrix Methods in Structural Mechanics, Wright Patterson A.F.B., Ohio, 1968.

5. Yakolev, M.N., "The Solution of Systems of Nonlinear Equations by a Method of Differentiation with Respect to a Parameter", U.S.S.R. Compt. Math. and Math. Phys., Part 1, 4, 1964, 198-203.

6. Davidenko, D.F., "An Application of the Method of Variation of Parameters to the Construction of Iterative Formulas of Increased Accuracy for Numerical Solution of Nonlinear Integral Equations", Soviet Math., Doklady 6, 3, 1965, 702-706.

7. Rubbert, P.E. and Landahl, M.T., "Solutions of Transonic Airflow Problem through Parametric Differentiation", A.I.A.A. Journal, 5, 1967, 470-479.

8. Chattopadhyay, A., "Application of Finite Element Method for Continuum Mechanics Problems", Ph.D. Dissertation, Dept. of Civil Engineering, I.I.T. Kanpur, India, 1970.

9. De Groot, S.R., Thermodynamics of Irreversible Processes, North Holland Pub. Co., Amsterdam, 1951.

10. Toupin, R.A., "Theories of Elasticity with Couple Stress", Arch. Rat. Mech. Anal. 17, 1964, 85-112.

11. Biot, M.A., "Linear Thermodynamics and Mechanics of Solids", 3rd U.S. Nat'l Cong. of Appl. Mech., 1958, 401-408.

12. Zeigler, H., "An Attempt to Generalize Onsager's Principle and its Significance for Theological Problems", ZAMP, 9, 1958, 748-769.

13. Coleman, B.D. and Mizel, V.J., "Existence of Caloric Equations of State in Thermodynamics", J. Chem. Phys., 40, 1964, 1116-1125.

14. Green, A.E. and Rivlin, R.S., "Mechanics of Nonlinear Material with Memory", Arch. Rat. Mech. Anal. 1, 1, 1957.

15. Nickell, R.E. and Sackman, J.J., "Approximate Solutions in Linear Coupled Thermoelasticity", Journal of Applied Mechanics, 35, 255-266, June 1968.

16. Ting, E.C., "Stress Analysis of Linear Viscoelastic Cylinders", A.I.A.A. 8, 1, 18-22, 1970.

FINITE ELEMENT ANALYSIS OF THERMOMECHANICAL PROBLEMS

J.H. Argyris, H. Balmer, J.St. Doltsinis, K.J. Willam

Institut für Statik und Dynamik der Luft- und Raumfahrtkonstruktionen, Universität Stuttgart
Imperial College of Science and Technology, University of London

The rate formulation for the quasistatic boundary value problem is rephrased after a brief discussion of "weak coupling" between thermal and mechanical processes. This description of the mechanical problem under non-isothermal conditions forms the basis for the subsequent finite element discretisation (matrix displacement method). First, a number of two dimensional finite element displacement models are investigated in regard to their strain approximation, bearing in mind that no thermal stresses should arise from linear temperature distributions. Subsequently, various possibilities are presented to approximate the spatial distribution of plastic work within higher order displacement models.

After this review of approximation problems, constitutive relations in thermal environments are discussed with special consideration of temperature dependent material properties.

First, the linear theory of uncoupled thermoelasticity is restated and extended to non-isothermal problems. For temperature dependent material properties there exists a fundamental difference between the hyper- and hypoelastic description of the material behaviour. Both formulations are discussed and applied to solve the thermoelastic problem of a cylinder with a spherical cavity under transient thermal conditions exhibiting temperature dependent material properties.

Subsequently, the elasto-plastic constitutive law is reformulated to incorporate a temperature dependent loading function. This general material relation is then restricted to the von Mises yield criterion and to the Prandtl-Reuss hardening rule in order to obtain quantitative thermo-elasto-plastic stress-strain relations for temperature dependent material properties. Moreover, the formulation is extended to incorporate creep effects, for which the prescribed thermal strains are supplemented by initial creep strains at the beginning of each time interval. The application of this material law to a realistic engineering example is illustrated on a pressure vessel nozzle under steady state conditions exhibiting temperature dependent material properties.

I INTRODUCTION

Highly stressed components of aircrafts, jet engines, turbines and nuclear reactors ore required to perform o complex pattern of operating schedules. A common problem orises thereby due to operating schemes which give rise to rapid changes of temperature and loading and hence to locul plostic deformation and associated residual stresses. Under full operation these stresses are further redistributed due to creep ond reloxation.

The objective of this paper is o realistic stress analysis of structural companents under steady state or transient temperature ond mechanical loading conditions. At elevoted temperatures most materials exhibit strain-rate sensitivity, creep, relaxotion and plostic deformations at stress levels much smaller than ot room temperature. These phenomena and prior defarmation history have considerable effect on subsequent deformation behaviour, all of which should be considered when stresses are computed throughout the structure as function of time. Some of these ospects are accounted for within the theary of thermoelasticity, thermoplasticity ond thermal creep which form the scope of this paper.

Under the assumption of quasistotic mation and weak thermomechanical coupling the solution of the thermal stress analysis reduces to two separate tasks: First, at each instont of time the temperature distribution of the structural component is determined independently of the mechanical state. Second, for these thermal conditions the stress analysis is corried out for prescribed temperature ond mechanical loading conditions occounting for temperature dependence of the material properties. For the mechanical response analysis it is ossumed that no distinction need to be made between deformed und undeformed configuration ond that the linear strain displocement relotionships pravide o proper measure of deformation. For quasi-static motion it is inferred that as a result of heat conduction temperatures change so gradually that the velocities remain small.

The finite element method (matrix displocement method) provides a tool reducing the solution of complex initial volue and boundary value problems of heat conduction ond stress distribution to the mere application of computer programs. In the present paper this method is first applied to the determination of the tronsient temperature distribution for a given thermal environment which then forms the input for the subsequent thermomechanical response analysis. The main advantage of treating both problems with the finite element method is due to the fact that the same spatiol discretisation can be used for the heat conduction and the associated stress analysis, reducing the preparation of input data to a minimum.

II THERMOMECHANICAL COUPLING AND HEAT CONDUCTION

The present investigation is restricted to thermomechanical processes of small deformations and slow temperature changes, in other words, only uncoupled problems ore considered in which inertio effects remain negligible.

The general heat balonce equation is briefly reviewed hereafter to discuss in short the assumptions in regard to thermomechanicol coupling. For non-isothermal processes recourse has to be token to the lows of thermodynamics. If the equations for conservation of mass, momentum ond energy ore supplemented by the proper constitutive equations of e.g. thermoelosticity the local heat balonce equation may be expressed as [see 1,2],

$$\mathrm{Div}\ (\underset{\sim}{k} \cdot \nabla \underset{\sim}{T}) = -T\ \left(\frac{\partial^2 \Psi}{\partial \underset{\sim}{\varepsilon}\, \partial T} : \dot{\underset{\sim}{\varepsilon}} + \frac{\partial^2 \Psi}{\partial T^2}\, \dot{T}\right) - \rho_0 S \qquad (1)^*$$

where T denotes the absolute temperature, $\nabla \underset{\sim}{T}$ the temperature gradient, $\underset{\sim}{\varepsilon}$ the measure of total deformation, S the distributed heat source, $\underset{\sim}{k}$ the material conductivity, ρ_0 the mass density before deformation and Ψ the free energy function also called after Helmholtz. Two sources of coupling may be distinguished:

a) Material Coupling:

Conductivity	$\underset{\sim}{k} = \underset{\sim}{k}\ (\underset{\sim}{\varepsilon}, T, \nabla \underset{\sim}{T})$	from the principle of
Free energy	$\Psi = \Psi\ (\underset{\sim}{\varepsilon}, T)$	equipresence of state variables

b) Physical Coupling: Presence of $\dot{\underset{\sim}{\varepsilon}}$ in equation (1).

There is very little experimental evidence indicating that the conductivity $\underset{\sim}{k}$ is a function of deformation; in most cases even the dependence on temperature gradients must be neglected due to lack of data. Normally, the free energy Ψ is a function of both, the deformation and the temperature, whereby the second term of the righthand side of equations (1) is solely dependent on temperature implying that material coupling is negligible if physical coupling does not occur. The importance of the physical coupling term involving $\dot{\underset{\sim}{\varepsilon}}$ becomes clearer if one considers an adiabatic process for which

$$\frac{\partial^2 \Psi}{\partial \underset{\sim}{\varepsilon}\, \partial T} : \dot{\underset{\sim}{\varepsilon}} = -\frac{\partial^2 \Psi}{\partial T^2}\, \dot{T} \qquad (2)$$

Equation (2) suggests that physical coupling disappears for quasistatic motions, where the body moves slowly through a sequence of equilibrium positions without exhibiting inertia effects. Vice versa for the gradual process of heat conduction, where temperatures vary slowly, the velocities remain so small that inertia effects can be neglected. From these observations it can be concluded that thermomechanical coupling should be considered in the case of e.g. thermoelasto dynamics, but not in the case of thermoelasto statics. For the solution of certain restricted classes of coupled thermoelastic field problems using the finite element method, the reader may consult for example references [3], [4] and [5].

Since the present investigation deals exclusively with the thermomechanical response analysis, the heat conduction analysis is summarized only briefly. Among the first presentations of the finite element discretisation of the uncoupled heat balance equation one should mention references [6, 7, 8, 9, 10], in which the latter three also include the transient case. The discretised form of the heat balance equation may be written in matrix notation as

$$\mathbf{C}\dot{\mathbf{T}} + \mathbf{K}\mathbf{T} = \mathbf{Q} \qquad (3)$$

$\mathbf{C}$ denotes the consistent heat capacity matrix, $\mathbf{K}$ the heat conductivity matrix and $\mathbf{Q}$ the thermal flux vector. This coupled set of first order differential equations is solved for the temperature distribution either by direct time integration or by spectral decomposition (mode superposition) similarly to procedures used for the dynamic response analysis. Explicit and implicit techniques can be applied for the solution of the parabolic problem. Normally only single step algorithms are feasible because of the excessive storage and stability requirements of multistep

---

* Invariant notation is used to formulate the equations of continuum mechanics. The symbols $\sim$, : , $\cdot$ denote vectorfields, tensor contraction and inner products.

techniques. The Wilson-Nickell algorithm [7], a mid-step algorithm similar to the Crank-Nicholson procedure, is utilized in the present investigation. Reference [10] gives a more detailed account of the finite element technique employed to solve the accompanying temperature problem.

III QUASISTATIC FINITE ELEMENT FORMULATION

In the context of non-isothermal thermomechanical problems there are two possibilities to describe the initial value problem, either by differential or by integral formulation. In extension of the standard incremental description this investigation is based on the rate (differential) formulation, as the thermomechanical constitutive relations are defined in the same form. The time dependence may not be omitted in this context, since the stress-strain relations contain, ingeneral, non-homogeneous expressions of stress- and strain-rates, e.g. for creep and relaxation.

1. Formulation of Initial Value Problem

The governing field equations can be separated into static and kinematic relations, which take in invariant notation the following form

Quasistatic Equilibrium :

$$\nabla \cdot \dot{\underset{\sim}{\sigma}} + \dot{\underset{\sim}{f}} = 0 \qquad \text{in } V$$

$$\dot{\underset{\sim}{\sigma}}^t = \dot{\underset{\sim}{\sigma}} \qquad \text{in } V$$

prescribed $\qquad \underset{\sim}{n} \cdot \dot{\underset{\sim}{\sigma}} = \dot{\underset{\sim}{p}} \qquad \text{on } S_\sigma$ (4)

Kinematic Relations :

$$\dot{\underset{\sim}{\varepsilon}} = \tfrac{1}{2} \left(\nabla \dot{\underset{\sim}{u}} + \nabla \dot{\underset{\sim}{u}}^t \right) \qquad \text{in } V$$

prescribed $\qquad \dot{\underset{\sim}{u}} = \dot{\underset{\sim}{u}} \qquad \text{on } S_u$ (5)

An appropriate generalisation of the principle of virtual work provides an equivalent statement to equations (4) and (5). Expressing the virtual work in virtual velocities and equating it to zero one obtains the following relation

$$\int_V \delta \dot{\underset{\sim}{\varepsilon}}^t \cdot \dot{\underset{\sim}{\sigma}} \, dV = \int_V \delta \dot{\underset{\sim}{u}}^t \cdot \dot{\underset{\sim}{f}} \, dV + \int_S \delta \dot{\underset{\sim}{u}}^t \cdot \dot{\underset{\sim}{p}} \, dS \qquad (6)$$

Equation (6) may be interpreted in the following way: If the virtual work is zero for any arbitrary infinitesimal virtual velocity field satisfying the kinematic relations (5), the mechanical system is in quasistatic equilibrium. Hence, the principle can be used to construct the quasistatic relations (4), if stress-strain rate relations are introduced. This procedure forms the basis for the finite element approximation method, using displacement models. For problems where the virtual work expression remains homogeneous in the rate terms the time derivatives can be omitted yielding the well-known incremental formulation for the solution of non-linear problems.

It should be emphasized that the principle of virtual work can be readily applied to non-conservative loading and that the constitutive relations need not be derivable from potential functions. The generalisation of the virtual work formulation is hence more flexible than alternative descriptions of the initial value problem using extensions of stationary variational principles.

For the complete formulation of the quasistatic problem equations (4) and (5) need to be supplemented by appropriate initial conditions at time $t = 0$

Initial static conditions:

$$\left.\begin{aligned} \mathbf{n}\cdot\dot{\boldsymbol{\sigma}} &= \dot{\mathbf{p}}_o \\ \mathbf{n}\cdot\boldsymbol{\sigma} &= \mathbf{p}_o \end{aligned}\right\} \text{"natural conditions"} \qquad (7)$$

Initial kinematic conditions:

$$\begin{aligned} \dot{\mathbf{u}} &= \dot{\mathbf{u}}_o \\ \mathbf{u} &= \mathbf{u}_o \end{aligned} \qquad (8)$$

2. Finite Element Discretisation

The virtual work equation (6) is now approximated with the help of finite element displacement models to construct a discrete form of the quasistatic relations. The finite element approximations may be written in matrix notation* as follows:

Velocity field $$\dot{\mathbf{u}} = \Phi_u \dot{\varrho}_u \qquad (9)$$

Initial strain-rate field $$\dot{\eta} = \Phi_\eta \dot{\varrho}_\eta \qquad (10)$$

Initial stress-rate field $$\dot{\tau} = \Phi_\tau \dot{\varrho}_\tau \qquad (11)$$

Total strain-rate field $$\dot{\gamma} = \dot{e} + \dot{\eta} = \nabla\Phi_u \dot{\varrho}_u \qquad (12)$$

Effective stress-rate field

$$\dot{\sigma} = E\dot{e} = E(\dot{\gamma} - \dot{\eta}) = E\dot{\gamma} + \dot{\tau}$$

$$= E(\nabla\Phi_u \dot{\varrho}_u - \Phi_\eta \dot{\varrho}_\eta) \qquad (13)$$

$$= E\nabla\Phi_u \dot{\varrho}_u + \Phi_\tau \dot{\varrho}_\tau \qquad (14)$$

The displacement interpolation Φ_u forms the basis for the spatial approximation of the displacement field $\mathbf{u}$ with $\dot{\varrho}_u$ denoting the column vector of nodal velocity values. The spatial distribution of elastically suppressed initial strains or equivalent initial stresses are described by the interpalation function Φ_η and Φ_τ. From a comparison of equations (13) and (14) it shauld be noted that both initial load approximations must satisfy the relation

$$\Phi_\tau \dot{\varrho}_\tau = -E\,\Phi_\eta \dot{\varrho}_\eta \qquad (15)$$

The initial stress and strain concept was introduced already in 1954 by the senior author [11] within the context of matrix (finite element) analysis of structures. It provides an ideal mechanism to account for elastically suppressed deformations arising from environmental conditions, such as temperature, moisture, irradiation, as well as mechanical conditions, such as nonlinearity, e.g. plasticity and creep. Within the frame of the generalised principle of virtual work all initial loads have to be considered prescribed with the exception of the instantaneous plastic strains or stresses which should be derived from the total strain-rate field $\dot{\gamma}$.

Substituting the field approximations into the virtual work equation one obtains the well-known quasistatic equilibrium equations for a typical finite element

$$k\dot{\varrho}_u = \dot{P} + \dot{J} \qquad (16)$$

---

* Bold founts denote matrices and vectors

The element quantities are defined as follows

Stiffness
$$k = \int_V \nabla \Phi_u^t \, E \, \nabla \Phi_u \, dV \tag{17}$$

Nodal forces
$$\dot{P} = \int_V \Phi_u^t \, \Phi_f \, dV \, \dot{\varrho}_f + \int_S \Phi_u^t \, \Phi_p \, dS \, \dot{\varrho}_p \tag{18}$$

Initial loads
$$\dot{J} = \int_V \nabla \Phi_u^t \, E \, \Phi_\tau \, dV \, \dot{\varrho}_\eta$$
$$= -\int_V \nabla \Phi_u^t \, \Phi_\tau \, dV \, \dot{\varrho}_\tau \tag{19}$$

Subsequently, the individual element quantities are assembled with the help of the displacement method furnishing the desired quasistatic equilibrium equations of the structure

$$K \dot{r} = \dot{R} + \dot{R}_J \tag{20}$$

3. Certain Remarks on Finite Element Approximations

In the previous section it was mentioned that two types of initial stresses and strains should be distinguished, those whose spatial variation is prescribed e.g. due to given temperature distribution and those whose variation should be derived from the displacement approximation, e.g. in the case of plasticity. The first part of this section deals with the occurrence of residual stresses which arise in the case of certain finite element approximations although the continuum is stress-free. The second part is concerned with additional approximations involving higher order elements where the spatial distribution of plastic work is represented by various interpolation schemes, instead of being derived from the variation of the total strain field.

a) Representation of Linear Strain States

Using the compatibility equations of elasticity [1] it can be shown that linear strain distributions satisfying internal compatibility do not give rise to effective stresses within a three-dimensional continuum, if no constraints are enforced due to kinematic boundary conditions. Within the context of thermal problems this implies that no effective temperature stresses may occur for linearly varying temperature fields. Hence, for a finite element discretisation of the continuum the selected total strain variation γ should contain linear strain states within each element. From equation (13) the linearly varying temperature distribution is described by initial strains as follows

$$T = a + b x \quad \rightarrow \quad \eta_\alpha = \Phi_1 \alpha \, T$$

where a and b are scalars and Φ_1 denotes a linear interpolation scheme. If the effective stresses should remain zero

$$\sigma = E \, (\nabla \Phi_u \varrho_u - \Phi_1 \alpha \, T) = 0$$

hence

$$\nabla \Phi_u \supseteq \Phi_1 \tag{21}$$

Note, that equation (21) holds only symbolically, implying that the order of variation of the displacement gradients $\nabla \Phi_u$ should contain linear strain states. General finite element theory states that the finite element approximation must include rigid-body modes as well as constant strain states to ensure convergence in the energy sense (completeness requirement, see e.g. [12]). Equ. (21) represents a further restriction on the finite element approximation, if artificial residual stresses due to linear initial strains are to be avoided. Although

this requirement is not necessary for energy convergence, the exact reproduction of linear strain variation is a highly desirable feature for practical applications.

The question arises now about the additional restrictions on the finite element approximation, which are necessary to retain linear strain states. To simplify the algebra, only the u-displacement component is considered together with its variation in the x-direction. The other components of the vector field follow the same argument. It is obvious that an element contains linear strain states, if the approximated displacement field is able to vary with the square of the coordinates at any point of the element. In other words, if the nodal displacements are defined by

$$\varrho_i = a x_i^2 \qquad \text{where } a \text{ is a scalar,} \tag{22}$$

then $u = ax^2$ should be satisfied everywhere. Substituting these nodal displacements into the finite element approximation of equation (9), one obtains

$$u = \Phi_{ui} \varrho_i = \Phi_{ui} a x_i^2 = a \Phi_{ui} x_i^2 \tag{23}$$

If the element geometry is described by the interpolation scheme Φ_x, then

$$x = \bar{\Phi}_{xj} x_j \tag{24}$$

and

$$u = ax^2 = a \left(\bar{\Phi}_{xj} x_j \right)^2 \tag{25}$$

The linear strain condition is satisfied if equations (23) and (25) yield identical results at every point of the element leading to the following condition

$$\Phi_{ui} x_i^2 \geq \left(\bar{\Phi}_{xj} x_j \right)^2 \tag{26}$$

We recall that Φ_u^i denotes the approximation of the displacement field defined by i nodes, while $\bar{\Phi}_x^j$ denotes the description of the element geometry which may be defined by j nodes. Condition (26) demonstrates clearly that in isoparametric elements where $\Phi_x = \Phi_u$ linear strain states are not contained.

Isoparametric elements: $\Phi_x = \Phi_u = \Phi$ with $i = j$ (27)

Linear strain condition: $\bar{\Phi}_i x_i^2 \geq \left(\bar{\Phi}_i x_i \right)^2$ (28)

Equality in equation (28) cannot be achieved due to the presence of the mixed terms on the right-hand side. The linear strain condition (26) suggests that subparametric elements where $\Phi_u \supset \Phi_x$ can reproduce linear strain states if the order of the displacement approximation contains the square of the geometric approximation. For example, consider a linear interpolation scheme for the geometry and a quadratic one for the displacements

with $\bar{\Phi}_x = \bar{\Phi}_1$ and $\Phi_u = \Phi_2$ (29)

One may expect satisfaction of the linear strain condition if

$$\Phi_{2i} x_i^2 \geq \left(\bar{\Phi}_{1j} x_j \right)^2 \tag{30}$$

Since algebra is extremely tedious, it is considered preferable to investigate the linear strain condition for a number of well established displacement models using numerical experiments. For simplicity, a square membrane subjected to linear temperature variation is chosen to assess the appearance and order of magnitude of residual stresses. Fig. 1 illustrates the geometrical and material properties of the problem together with the boundary conditions and temperature loading. The plane membrane is discretised by 25 nodes, each with two in-plane degrees of freedom, using Lagrangian interpolation. The effect of two parameters is investigated, the order of geometric approximation and the order of displacement approximation.

Fig. 2 illustrates the finite element lay-out of Phase I, where only square elements or isosceles triangles are used. This discretisation implies that $\Phi_x = \Phi_0$, which is synonymous to a constant "Jacobian". It also shows the element types used for the analysis, the linear displacement models like TRIM3 and QUAM4 with $\Phi_u = \Phi_1$, and the quadratic displacement models like TRIM6, [14] , QUAM8 [40] , [41] and QUAM9 [42] , having $\Phi_u = \Phi_2$ with the exception of QUAM8 where $\Phi_u \subset \Phi_2$.

Fig. 3 illustrates the finite element lay-out of Phase II, where general quadrilateral and triangular elements are used. This discretisation implies that $\Phi_x = \Phi_1$, which is synonymous to a linear variation of the Jacobian.

Finally, Fig. 4 shows the curvilinear finite element lay-out of Phase III, where isoparametric quadrilateral and triangular elements are used [42] , [43] . This discretisation implies that $\Phi_x = \Phi_2$, which is synonymous to a quadratic variation of the Jacobian. The associated curved elements are denoted by QUAMC 9 and TRIMC 6.

The results of this investigation are summarised in Table 1, where values for the residual stresses are given which arise solely from the finite element approximation. The mean square root of all nodal octaheder normal stresses and shear stresses are presented in percent of the initial thermal stresses after averaging first the nodal values.

$E = 10$ $\nu = 00$ $\alpha = 10$ $h = 10$		TRIM 3 $\Phi_u = \Phi_1$	QUAM 4 $\Phi_u = \Phi_1$	TRIM 6 $\Phi_u = \Phi_2$	QUAM 8 $\Phi_u = \Phi_2$	QUAM 9 $\Phi_u = \Phi_2$	TRIMC 6 $\Phi_u = \Phi_2$	QUAMC 9 $\Phi_u = \Phi_2$
Phase I $\Phi_x = \Phi_0$	$\bar{\sigma}_0$	1.85	3.88	0	0	0		
	$\bar{\tau}_0$	1.88	1.79	0	0	0		
Phase II $\Phi_x = \Phi_1$	$\bar{\sigma}_0$	2.15	3.34	0	2.37	0	-	
	$\bar{\tau}_0$	2.37	1.86	0	0.54	0		
Phase III $\Phi_x = \Phi_2$	$\bar{\sigma}_0$	-		-			12.50	30.67
	$\bar{\tau}_0$						14.17	34.34

Table 1 Residual Stresses in Percent of Prescribed Temperature Stresses

The results for TRIM6 and QUAM9 clearly confirm that elements with a quadratic or biquadratic displacement field and a linear or bilinear geometric description do contain exactly the linear strain states, which must be expected from equation (30). All other elements give rise to residual stresses due to linear variation of initial thermal strains. This has to be expected for the constant strain triangle TRIM3 and the associated quadriloteral QUAM4, since they simply cannot reproduce linear strain states due to the low order of the displacement approximation. The residual stresses from QUAM4 are consistently worse than those of TRIM3, because simple averaging of the linear initial strains is only possible for the TRIM3. It is interesting to note that the element QUAM8 does not contain the linear strain states for a linearly varying Jacobian, since it is based on an incomplete quadratic displacement interpolation. Furthermore, the isoparametric elements TRIMC6 and QUAMC9 create considerable residual stresses due to their "geometric distortion". It should be remarked that a more severe distortion of the elements gives rise to even larger residual stresses, an eccentric positioning of the midside nodes being particularly harmful to the quality of stresses.

The variation of some residual stresses is illustrated in form of contourline plots on Figs. 5 to 8. There the distribution of octaheder shear, which is normalised in respect to the initial thermal stress, is shown for Phase II and III using the isoparametric elements TRIM3, QUAM4, TRIMC6 and QUAMC9. It should be mentioned that some resulting stress components are of an order higher than the octaheder stresses which are proportional to the first and second stress invariant. Moreover, the contour line plots are based on average nodal values which are used as base points for a linear interpolation scheme; hence they can not reflect the staggering discontinuities of stress components which occur e.g. at the center node raising the old question of an appropriate stress evaluation. When using average nodal values, it is quite clear that the amount of data handling reduces considerably for nodes on the element boundary versus internal element nodes. Furthermore, the user is mainly interested in the values at the corner nodes which are easily identified and which also allow a very useful check of the accuracy of the finite element idealisation for nodes where the "natural" boundary conditions should be satisfied. On the other hand it is interesting to note that the stress distribution in the interior of an element seems less affected by the geometric distortion discussed above. This is particularly true in the vicinity of the element centroid suggesting the use of centroidal stress points.

Since this investigation is rather preliminary in nature, only a few observations could be made. In conclusion one can say that the subporametric elements TRIM6 and QUAM9 are preferable to other elements as they do not give rise to artificial residual stresses in case of linear initial stress or strain problems. Caution should be exercised in the case that distorted elements are used to describe the geometry of the structure. Especially the eccentricity of the midside nodes impairs the resulting stress distribution which may offset the improvement gained by a more accurate geometric representation of the real structure.

b) Plastic Work Approximation in Higher Order Elements

For finite element displacement models the kinematic assumptians form the basis for the description of internal energy within each element. Given a unique relatianship between stresses and total strains the spatial variation of internal energy is completely determined by the distribution of total strains. It is obvious that non-linear constitutive laws create additional problems for representing stresses or initial strains in higher order elements, in which the total strain distribution varies inside the element. In what follows, certain questions of approximatian are dealt with in connection with the tangential stiffness method as applied to the elasto-plastic analysis using higher order elements.

The incremental elasto-plastic stress-strain relation is described by

$$\dot{\sigma} = F\,\dot{\gamma} \tag{31}$$

For duality of the transformation laws the states of stress and strain are defined for the case of plane stress by

$$\sigma = \{\sigma_{xx}\ \sigma_{yy}\ \sqrt{2}\,\sigma_{xy}\} \qquad \gamma = \{\gamma_{xx}\ \gamma_{yy}\ \tfrac{1}{\sqrt{2}}\,\gamma_{xy}\} \tag{32}$$

For isotropic materials and the von Mises yield criterion with the Prandtl-Reuss flow rule the elasto-plastic material law can be written as follows [38]

$$F = \left(I_3 - \frac{1}{\xi + s^t E\, s}\, E\, s s^t\right) E \tag{33}$$

with

$$s = \frac{3}{2\,\bar{\sigma}}\,\sigma_D \qquad \bar{\sigma}^2 = \frac{3}{2}\,\sigma_D^t \sigma \tag{34}$$

where the subscript D denotes deviatoric components and where ξ, the plastic strain-hardening parameter, is obtained from uniaxial data

$$\dot{\eta}_p = \frac{1}{\xi}\,\dot{\bar{\sigma}} \tag{35}$$

The linear elastic material law reads for the isotropic case

$$E = 2G\left(I_3 + \frac{\nu}{1-\nu}\, e_{2,1} e_{2,1}^t\right) \tag{36}$$

with

$$e_{2,1} = \{1\ 1\ 0\} \tag{37}$$

A more detailed derivation of the elasto-plastic relationships is given in Section IV for tri-axial conditions.

The elasto-plastic element stiffness is obtained simply by replacing the linear elastic material law in the expression of the strain energy density by its elasto-plastic extension, such that

$$k_p = \int_V \Phi_\gamma^t\, F\, \Phi_\gamma \, dV \tag{38}$$

Note, that in this case the elasto-plastic stiffness refers to the straining modes φ_γ, which are identical with the so-called natural deformation modes φ_N introduced in [13]. it is obvious that the integral in (38) has to be evaluated numerically even for simple element configurations.

The question arises now about the representation of the elasto-plastic material law for deviatoric stresses which vary within the element domain. Basically, there are two possible avenues for implementing the tangential stiffness formulation using higher order elements: The direct integration route involving the construction of the elasto-plastic material law (33) at every pivot point of the numerical integration scheme, and the interpolation of the nonlinear stress-strain law from a chosen set of base point, which can be written as

$$F = \Phi_F \varrho_F \tag{39}$$

Our main goal is now to obtain quantitative results for the effect of the different schemes when forming the elasto-plastic stiffness matrix. To this end, the spectrum of eigenvalues of the difference matrix between the reference and current stiffness provides a convenient norm to measure the difference of internal energy between two associated matrices. This method has proven useful in [45] to compare stiffness matrices of an element family having identical nodal configurations. The total internal energy rate of an elasto-plastic element is given by

$$\dot{U}_p = \dot{\varrho}_\gamma^t \, k_p \, \dot{\varrho}_\gamma \tag{40}$$

This increment of internal energy can be expressed in terms of the eigenvalues of k_p in the following form

$$x^t k_p x = \lambda I \tag{41}$$

where any set of nodal strain $\overset{\text{increments}}{\dot{\varrho}_\gamma}$ can be represented as a linear combination of the set of eigenvectors x. The spectrum of eigenvalues of k represents a measure for the total internal energy rate of an element, hence this tool can also be utilized to determine the energy difference between a reference element and the element currently under investigation. It should be emphasized that only the eigenvalues of the difference matrix $k_r - k_c$ form a proper measure of this energy in contradistinction to the difference of eigenvalues of each individual stiffness matrix which in general do not correspond to the same eigenvectors. Thus

$$(\dot{U}_r - \dot{U}_c)_p = X^t (k_r - k_c)_p X = \Delta \tag{42}$$

X denotes the matrix of orthonormal eigenvectors of the difference matrix $k_r - k_c$ and Δ stands for a diagonal matrix containing all eigenvalues.

The associated spectral radius of the difference matrix forms a suitable norm to measure the "distance" between the reference stiffness and the current stiffness under investigation. The spread of eigenvalues indicates the variation of the difference between the two matrices for all possible deformation states, while the arithmetic mean yields a value illustrating the deviation of internal energy representation which can be expected in an average sense. Note that only positive or negative definite difference matrices permit a simple conclusion about the current stiffness being more flexible or more stiff than the reference stiffness.

In general the element stiffness is affected by two parameters, the material properties and the geometry. For the present considerations, attention is restricted to an equilateral linear strain triangle TRIM6, where the geometric effects are eliminated and where the total strain variation is defined by the linear interpolation scheme

$$\Phi_{\gamma} = \Phi_{,} \tag{43}$$

A representative choice for the "true" variation of the elasto-plastic material law forms the main difficulty in illustrating the evaluation of the elasto-plastic stiffness matrix using different interpolation and integration schemes. In order to avoid any bias in the representation of the material law, it should be assumed that the whole element domain plastifies instantaneously, while for simplicity the deviatoric stress components vary linearly over the element domain, as illustrated in Fig. 9. This stress assumption is valid for incipient plastic deformation only where no plastic deformation history yet distorts the linear stress distribution in the element. It should be mentioned that this condition cannot be realized since for a linear stress distribution the equivalent stresses form a second order surface over the element domain for which

$$\bar{\sigma}^2 = \frac{2}{3}\,\boldsymbol{\sigma}_D^t\boldsymbol{\sigma} \neq \text{const} \tag{44}$$

To achieve instantaneous plastification of the whole element it has to be further assumed, either that the yield limit varies over the element domain according to the shape of the equivalent stress or that the plastic deformation history is followed only at the point where $\bar{\sigma} = \bar{\sigma}_{max}$. We adopt below the second line of thought with the following material properties assuming that constant strain hardening, as illustrated in Fig. 10, prevails throughout the element and that the plastic deformation histories of other points need not be considered.

In the example, the following data were assumed:

Yield Stress:	$\bar{\sigma}_y = 19.4$	kp/mm^2
Strain Hardening:	$\beta = 0.3$	kp/mm^2
Elastic Modulus:	$E = 21000$	kp/mm^2
Poisson's Ratio:	$\nu = 0.3$	

The elasto-plastic stiffness matrix calculated with a 30 x 30 Gaussian quadrature rule is regarded as the "exact" reference stiffness for the subsequent investigation of various matrices. This implies that the elasto-plastic material law is evaluated at the 900 pivot points for the assumed distribution of deviatoric stresses and material constants. Subsequently, elasto-plastic stiffness matrices are constructed either by direct integration involving the evaluation of the elasto-plastic material law from the governing stress distribution of the pivot point considered, or by interpolation from a chosen set of base points. We then compare these elasto-plastic stiffness matrices with the reference matrix using the concept of eigenvalues of each difference matrix mentioned above.

The following numerical integration schemes are selected for comparison: 1 point, 2 point, 3 point, 4 point, 6 point and 12 point Gaussian quadrature rules. The range of eigenvalues is illustrated in Fig. 11 for each stiffness difference matrix.

$$X^t (k_{30} - k_{int}) X = \Delta \qquad (46)$$

Order of Integration: int. = 1, 2, 3, 4, 6, 12

The results are plotted in terms of the numbers of material data points which caincide in this case with the number of pivot points of the quadrature rule.

This presentation is particularly useful since it illustrates the deviation of the elasto-plastic stiffness from the reference stiffness as function af the amount af the necessary data handling.

Alternatively ta the direct integration route, five Lagrangian interpolatian schemes shown in Fig. 12 are selected to approximate the distributian of the elasta-plastic material law over the element damain. Besides the constant, linear, quadratic and cubic interpolatians, a piecewise linear interpolation scheme over triangular subregions, as proposed in [44], is chosen to represent the material law. The elasto-plastic stiffness is evaluated using a 30 x 30 Gaussian quadrature rule, but in cantradistinction ta the reference stiffness, anly the base points of the interpolatian scheme now form material data points. The range af eigenvalues is illustrated in Fig. 13 for each stiffness difference matrix. Again the results are presented in terms af the number of material data points involved in different interpolatian schemes to allow a direct camparison with the range of eigenvalues of the direct integratian methods. Nate that only linear interpolation preserves the positive definiteness af the material law within the element damain. E.g. Fig. 14 illustrates the non-positive definite material regions which are artifically introduced by cubic interpolation even though the material law at the base points is positive definite; therefore, Fig. 13 shaws twa different results for the cubic interpolatian as the elasta-plastic stiffness can be evaluated either neglecting ar cansidering the type af material law encountered at each pivot point, the secand of which may lead ta a non-positive definite stiffness matrix. Fram these observations it is quite clear that higher arder interpolatian of the material law should be avoided altogether since it changes the character af the material law instead af impraving its representatian.

Camparing Fig. 11 and Fig. 13 one can see the distortion of the "exact" elasta-plastic stiffness due ta the direct integratian method and the interpolation method far the same number af material data points involved. It seems that the interpolatian schemes yield consistently warse results than the carresponding direct integration schemes except for the linear interpolatian aver triangular subregians. This method of material representation offers a slight advantage considering that the tatal amount of data handling reduces sharply for the element assembly, where the valency of the edge nodes varies between 2 and 6 in cantradistinction ta the single valency far the interior data points af the direct integration method.

Finally, it shauld be mentioned that this element investigation using the eigenvalue technique af stiffness difference matrices is preliminary in nature and should be accampanied by similar numerical experiments on the structural level ta verify quantitatively the qualitative element results.

IV CONSTITUTIVE LAWS

In this section a number of engineering theories are considered to describe the material behavior in thermal environments. Emphasis is placed on the time rate formulation of constitutive relations fitting the scope of the quasistatic formulation. The following discussion is restricted to thermomechanically simple materials where the response at a given instant of time is determined by the histories of deformation and temperature.

A realistic material description should consider temperature dependent material properties leading in general to stress-strain relations non-linear in the temperature. Under non-isothermal conditions this temperature dependency gives rise to time variable material behaviour. The subtle point of the material ageing will be treated in the following section on thermoelasticity, where two fundamentally different material descriptions are reviewed and applied to solve a thermoelastic problem. Subsequently, the well-known elasto-plastic stress-strain relations are extended to include temperature effects including temperature dependent yielding. It will be further shown how the engineering theory of creep may be simply incorporated in the thermoelasto-plastic material law to account for viscous effects. This general constitutive law is then applied to the solution of a realistic engineering problem.

It should be emphasized that the virtual work formulation in terms of the rates permits the incorporation of very general constitutive relations as long as they define the instantaneous material response as function of deformation and temperature history. For this direct route there is no need to claim the existence of potentials from which the stress-strain relations may be derived.

1. Thermoelasticity

As introduction, the linear stress-strain relations are briefly reviewed for isothermal conditions. The three-dimensional cartesian form of Hooke's law may be stated in matrix notation as follows

$$\boldsymbol{\sigma} = \mathbf{E}\,\boldsymbol{\varepsilon} \tag{47}$$

where $\boldsymbol{\sigma}$, the stress vector, having six components, is related to the corresponding strain vector $\boldsymbol{\varepsilon}$ by the linearly elastic material law $\mathbf{E}$ involving 21 constants in case of general anisotropy. The thermal expansion due to a temperature change may be interpreted as mapping of the neighbourhood of a material point into a new configuration assuming all mechanical constraints can be temporarily neglected. Although the stress free thermal expansion may physically not be possible, this new configuration can be thought of as a reference state from which effective, stress producing, deformation is measured. Assuming that the changes of temperature remain small in regard to the reference temperature T_o, the thermal deformation may be described by the following expression which is linear in temperature

$$\boldsymbol{\varepsilon}_\Theta = \boldsymbol{\alpha}\,(T - T_o) = \boldsymbol{\alpha}\,\Theta \tag{48}$$

where for the isotropic case the vector of coefficients of thermal expansion degenerates to

$$\boldsymbol{\alpha} = \alpha\,\mathbf{e}_{3,1} \qquad \text{with} \qquad \mathbf{e}_{3,1} = \{1\ 1\ 1\ 0\ 0\ 0\} \tag{49}$$

The coefficients of thermal expansion are independent of temperature for changes of temperature $\Theta = T - T_o$ less than 40°C [2]. This linear form of the fundamental decomposition for thermomechanical deformation forms the basis of the well-known generalisation of Hooke's law also named after Duhamel-Neumann

$$\gamma = \varepsilon + \eta_\theta = e + \alpha\theta = E^{-1}\sigma + \alpha\theta \qquad (50)$$

or

$$\sigma = E(\gamma - \alpha\theta) = E\gamma + \theta\beta = E\gamma + \tau_\theta \qquad (51)$$

where the initial temperature stresses are defined by

$$\tau_\theta = \theta\beta \qquad \text{and} \qquad \beta = -E\alpha \qquad (52)$$

Under the ossumption of weak coupling the distribution of transient temperatures con be considered prescribed at each instant of time as

$$T = T(t) \qquad \text{and} \qquad \Theta = \Theta(t) \qquad (53)$$

Hence, for temperature dependent material properties the thermal stress analysis involves ageing material properties or in other words o time variable system since

$$\begin{aligned} E(T) &= E(t) \\ \beta(T) &= \beta(t) \end{aligned} \qquad (54)$$

For non-isothermal conditions there are two possibilities to describe the elastic material response, which differ fundamentally if temperature dependent material properties are considered: the classical hyperelastic formulation and the hypoelastic characterization which are summarized below [1].

a) Hyperelastic Formulation

in the classical sense an elastic material possesses a natural state in thermodynamic equilibrium at which $\theta = 0$ ond $T = T_0$ exhibiting a unique one to one correspondence between stress ond strain in an appropriate neighbourhood of this reference state. The materiol is characterized by its return to the natural state after removol of loading, or in other words the domain of the hyperelastic response functional contoins no history effects. From this postulate of elastic material behaviour (51) can be generalized, leading to the following "integral" relation

$$\sigma = E(t)\gamma + \int_{T_0}^{T} \beta(\tau)\, d\tau \qquad (55)$$

Equation (55) describes essentially an ageing viscoelastic material with infinitely short memory. It should be emphasized that this characterization is restricted by the definition of an appropriate neighbourhood from the natural state. Formal differentiation in time furnishes the equivolent rate formulotion exhibiting non-homogeneous terms in γ and θ

$$\dot{\sigma} = E(t)\dot{\gamma} + \beta(t)\dot{\theta} + \dot{E}(t)\gamma + \dot{\beta}(t)\theta \qquad (56)$$

This equation represents a total differential ensuring reversibility and path-independency of the thermoelastic process. These properties ore intrinsic prerequisites for the existence of a thermoelastic potential from which hyperelostic stress-strain relotions are usually derived.

It should be noted that due to the temperature dependence of E and β equation (56) is basicolly non-linear in temperature. In spite of this non-linearity there is no need to pursue an incremental approach for the thermal stress onalysis. Since the temperatures are prescribed in the uncoupled problem the total stress relotion of equation (55) con be used directly to determine the hyperelastic response at any instant of time. Note that in this case the material must possess a natural state to which it returns after unloading ond that on appropriate neighbourhood is defined for the excursions from this reference state.

b) Hypoelastic Formulation

In contradistinction to the hyperelastic description, the hypoelastic characterization of an elastic solid is formulated without reference to a natural state, and expresses the components of stress rate as homogeneous linear functions of the components of the total strainrate and temperature rate

$$\dot{\sigma} = E(t)\,\dot{\varepsilon} + \beta(t)\,\dot{T} \tag{57}$$

This implies that this constitutive law involves in addition to the homogeneous state variables σ, ε and T only mechanical properties characterizing the instantaneous state of the solid. These time variable moduli are independent of the speed at which changes of the state variables are affected. Thus

$$\dot{E}(t) = \dot{\beta}(t) = 0 \tag{58}$$

An equivalent integral formulation may be obtained from the principle of superposition on the rate expressions $\dot{\sigma}$, $\dot{\varepsilon}$ and $\dot{T}$. The well-known step formulation for ageing hypoelastic materials yields the following total stress expression

$$\sigma(t) = \int_{t_0}^{t} E(\tau)\,\frac{d}{d\tau}\,\{\varepsilon(\tau) - \Theta(\tau)\}\,d\tau \tag{59}$$

After integration by parts the equivalent impulse formulation furnishes

$$\sigma(t) = E(t)\,\{\varepsilon(t) - \Theta(t)\} - \int_{t_0}^{t} \frac{dE(\tau)}{d\tau}\,\{\varepsilon(\tau) - \Theta(\tau)\}\,d\tau \tag{60}$$

where the pseudo temperature Θ is defined by

$$\Theta(t) = \int_{T_0}^{T} \alpha(\tau)\,d\tau \tag{61}$$

The integral formulations clearly show that the stress at time t is a function of the entire strain and temperature history, the functional being linear in strains, but non-linear in temperature. Hence, this characterization basically implies that the elastic solid exhibits path dependency and irreversibility due to ageing effects similar to viscous phenomena. It should be remarked that the theory of hypoelasticity forms the basis for the hereditary integral representation of viscoelastic behaviour.

A comparison between the impulse formulation of equation (60) and the hyperelastic stress-strain relation (55) indicates that both material characterizations of elastic bodies coincide for time-invariable systems. Of course, the same conclusion is reached from a comparison of the analogous rate formulations, equations (56) and (57). It is obvious that, if attention is restricted to close neighbourhoods of the natural state, such that constant material properties prevail, then the hypoelastic formulation degenerates to the hyperelastic one.

c) Numerical Implementation

The finite element analysis of the transient thermoelastic problem reduces to a straight-forward application of the initial load concept, see equations (16) and (20).

The integral relationship of the hyperelastic formulation (55) yields for temperature dependent mechanical properties a time-variable set of equilibrium relations of the following form

$$K(t)\,r(t) = R(t) + R_0(t) \tag{62}$$

Note that the instantoneous stiffness $\mathbf{K}(t)$ as well os the initial loading due to temperatures $\mathbf{R}_\theta(t)$ can be evoluated directly for ony stoge of the prescribed transient thermal conditions. For temperoture independent mechanical properties equation (62) reduces to o time-invoriable system in which $\mathbf{K}(t)=\mathbf{K}$, such that transient temperature conditians involve solely the solution of a number of loading conditions.

The rate type relotionship of the hypoelostic formulotion (57) furnishes for temperature dependent materiol properties o time-variable set of incremental equilibrium relations of the following form

$$\mathbf{K}(t)\,\mathbf{r}_\Delta(t) = \mathbf{R}_\Delta(t) + \mathbf{R}_{\theta\Delta}(t) \tag{63}$$

It should be emphosized again that the hypoelastic solid requires an incrementol strategy for the thermal stress analysis, since the whole temperature history needs to be considered. Therefore, this formulation is far more laborious than the total stress analysis for the numerical implementation of hyperelastic material behaviour. In case of temperature independent mechanical properties the incremental equilibrium equation degenerates to equation (62) with a time-invariable stiffness.

d) Example

In conclusion the trcnsient thermoelastic stress distribution is determined for an "infinite" medium with a spherical cavity, the boundary of the cavity being subjected to the sudden temperature rise $\Theta_i = 600^\circ$ F. Bath, constant and temperature dependent material properties are dealt with using a hyperelastic and a hypoelostic description of the elastic behoviour. This polorly symmetric problem is selected as theoretical solutians are available for comparison with the finite element results far the case of constant material properties [16] as well as for the case of temperature dependent properties [17].

For the finite element analysis the infinite medium is idealised by a finite cylinder with a radius ten times the radius of the inner cavity and a height equal to the diameter of the cylinder. Because of symmetry only the upper half of the cylinder needs to be analysed. Bath, the transient heat canduction analysis as well as the thermoelastic stress analysis is carried out using the same finite element mesh lay-out which is illustrated in Fig. 16. 192 oxisymmetric TRIAX 6 elements are used to discretize the cylindrical structure with the cavity involving 425 nodal paints. The reader is referred to [15] far a description af the triangular ring element TRIAX 6 having 6 nodes.

Two cases oie treated, a medium with constont, temperature independent materiol properties and a temperature sensitive medium exhibiting more realistic material properties as occur in the case of SAE 1095 steel [18]. The temperature dependency of the elastic constants is described by

$$\begin{aligned}
&\text{Shear modulus} && G = 11.55 \times 10^6 \left(1 - 0.3463 \times 10^{-6}\,\theta^2\right) \\
&\text{Poisson's ratio} && \nu = 0.345 + 0.7 \times 10^{-2} \left(\theta - 565\right)^{1/6} \\
&\text{Thermal expansion} && \alpha = 0.65 \times 10^{-5} \left(1 + 0.4231 \times 10^{-3}\,\theta\right)
\end{aligned} \tag{64}$$

Fig. 17 illustrates the variation of the elastic properties with temperature. For the finite element solution of the heat conductian problem the thermal diffusivity $k/\rho c$ is assumed to remain constant. The boundary conditians of the thermal problem consist only of the prescribed temperatures $\Theta_i = 600^\circ F$ at the inner surface of the cavity and $\Theta_s = 0^\circ F$ on the outside surface.

In order to simplify the comparison with the theoretical solutian the following non-dimensional quantities are introduced

Time	$\tau = \frac{1}{a^2} \frac{k}{\rho c} t$	
Temperature	$\Theta^* = \Theta/\Theta_i$	
Radial Displacement	$u_r^* = \frac{1-\nu}{1+\nu} \frac{1}{\alpha_0 \Theta_i a} u_r$	(65)
Radial Stress	$\sigma_r^* = \frac{1-\nu}{2(1+\nu)G} \frac{1}{\alpha_0 \Theta_i} \sigma_r$	
Hoop Stress	$\sigma_t^* = \frac{1-\nu}{2(1+\nu)G_0} \frac{1}{\alpha_0 \Theta_i} \sigma_t$	

Ta define the geometry af the problem, the radius af the spherical cavity is taken as $a = 1.0$ and the thermal diffusivity is assumed to be $k/\rho c = 1.0$. Since the theoretical solutions are based on the polarly symmetric treatment af the infinite medium, the results of the problem are only presented at various radial distances af the midsection through the cavity.

Fig. 18 illustrates the radial distribution of the transient temperatures $\Theta^* = \Theta/\Theta_i$ after the sudden temperature rise $\Theta_i = 600^\circ F$ at the cavity. The results af the theoretical heat conduction analysis far the infinite medium agree very well with those of the finite element solution for the cylinder. In the vicinity af the spherical cavity there is virtually no difference between both results. The time integration af the heat balance equation was carried out with variable time steps ranging fram $\tau = 1/32$ ta $\tau = 4$.

Two parameters are investigated by means af the fallawing transient elastic stress analysis, the influence of temperature dependent material praperties and the type af elastic formulatian.

In Fig. 19, 20 and 21 displacements and stresses are compared in a temperature-sensitive and insensitive medium at different times τ after the temperature rise at the cavity. In the case of temperature-independent material propeties, the finite element results far the cylinder are shown in addition ta the theoretical solution for the infinite medium presented in [16]. The finite element results far the temperature-sensitive medium are obtained using the hyperelastic formulation; this means that at every instant of time the tangential stiffness has to be determined from the temperature dependent material praperties af Fig. 17. It should be mentioned that a linear interpolation scheme is used ta describe the variation af the elastic moduli within each TRIAX 6 element. Fig. 19 illustrates the radial distribution af displacements at various times τ. Note, that the displacements at the spherical cavity remain zero irrespective of the temperature rise at the cavity. For the temperature-insensitive properties there is very little difference between the analytic and the finite element solutians; these results deviate considerc from the temperature-sensitive results in regions af high temperature gradients. Fig. 20 presents an analogous illustration af transient radial stress components. Again, there is a significant difference between temperature-sensitive and insensitive properties in areas of

high temperature gradients. It may be of interest that the violation of the stress boundary condition $\sigma_r^* = 0$ at the cavity is very small in comparison to the large changes of σ_r^* in radial direction, hence the discretisation error seems negligible. In case of the long-term response of radial stresses the analytic solution deviates considerably from the finite element results in regions influenced by the different boundary conditions. Fig. 21 illustrates the distribution of tangential stresses at various times. There is hardly any noticeable difference between the two finite element solutions except right at the spherical cavity.

In Fig. 22, 23 and 24 transient displacements and stresses are compared in a temperature-sensitive medium, the response of which is described either by the hyperelastic or hypoelastic formulation. The finite element results are shown for the cylinder in addition to the corresponding theoretical solution of the infinite medium, presented in [17]. The finite element solutions are obtained using both, the total stress approach of the hyperelastic formulation for different instants of time, and the incremental stress approach of the hypoelastic description, proceeding through the temperature history. Poisson's ratio is assumed to remain constant at $\nu = 0.4$ since the theoretical results are based on that value. Fig. 22 illustrates the radial distribution of displacements at various times. Note that there is very little difference between the hyperelastic and hypoelastic results even in regions of high temperature gradients. The deviation of the analytic results can be explained due to the difference between the analytic and finite element temperature distributions, arising from the dependency of conductivity on the temperature in the former case [17]. The transient radial stress components, illustrated in Fig. 23, exhibit a more pronounced effect of the two possibilities to describe the elastic behaviour. Again, the long-term results of the analytic solution deviate considerably from the finite element results in regions influenced by the different boundary condition. Fig. 24 illustrates the distribution of tangential stresses at various times; these components are hardly affected by the two alternative formulations except near the spherical cavity.

In conclusion one may state that a realistic stress analysis should accommodate the temperature dependence of mechanical properties, since the assumption of temperature insensitivity underestimates considerably the structural response. The type of formulation to describe the elastic behaviour seems to have little influence onto the results. However, this statement applies only to the type of problem considered here, where the temperature sensitivity of the material properties remains relatively small. On the other hand, these quantitative results indicate that the hypoelastic formulation can be utilized if an incremental formulation is necessary, such as in plasticity and viscoelasticity. However, this costly method does not justify the expense, if applied to account for path dependency due to temperature dependent elastic properties in a transient thermal environment.

2. Thermoelastoplasticity and Creep

Following [19] and [21] the flow theory of plasticity is first reviewed in regard to non-isothermal conditions without taking recourse to thermodynamics. Subsequently, the general constitutive law is specialized to construct a set of thermo-elasto-plastic stress-strain relations which can be used to solve realistic engineering problems where high stresses of short duration occur under applied loads and temperatures. After presenting the time-independent stress-strain relation for plasticity, the formulation is extended to incorporate creep. Omitting history dependence of creep deformations, the phenomenological description of the engineering theory of creep is dealt with, which provides a simple means to account for viscous effects in metals. It is shown that the initial load techniques for the iterative solution of plasticity problems are

easily adjusted to incorporate creep. In conclusion an example is presented to illustrate the thermo-elasto- plastic and creep behaviour on a realistic engineering problem.

a) Thermo-elasta-plastic Farmulation

Consider an elasto-plastic workhardening solid where far a given state infinitesimal changes of temperatures and stresses produce a unique infinitesimal change of elastic and plastic strains. It is assumed that the constitutive law is homogeneous af arder one invalving only the time rates $\dot{\sigma}$, $\dot{\eta}_p$ and $\dot{T}$, while the material properties characterize the instantaneous state af the solid analogous ta the hypoelastic farmulatian. An additianal state variable need ta be defined ta describe the dissipation af mechanical energy via

$$k = \int_0^t \sigma^t \dot{\eta}_p \, dt \tag{66}$$

where k denotes the warkhardening parameter as a non-decreasing functian af time. It is further assumed that there exists a functian af the state variables $f(\sigma, \eta_p, T, k)$ which describes initial and subsequent yield surfaces far all plastic states. The time rate af the loading function f is utilized ta define loading between twa plastic states f and $f+df$, where

$$\dot{f} = \frac{\partial f}{\partial \sigma} \dot{\sigma} + \frac{\partial f}{\partial \eta_p} \dot{\eta}_p + \frac{\partial f}{\partial T} \dot{T} + \frac{\partial f}{\partial k} \dot{k} \tag{67}$$

The state af loading is defined by $f=0$.

and
$$\frac{\partial f}{\partial \sigma} \dot{\sigma} + \frac{\partial f}{\partial T} \dot{T} \begin{cases} < 0 & \text{Unloading} \\ = 0 & \text{Neutral Loading} \\ > 0 & \text{Plastic Loading} \end{cases} \tag{68}$$

The "normality canditian", which fallaws fram Drucker's stability postulate far inelastic materials [20], farms the basis af the associated flaw rule

$$\dot{\eta}_p = \lambda \frac{\partial f}{\partial \sigma^t} \tag{69}$$

The positive scalar λ can be evaluated fram Prager's cansistency condition $\dot{f} = 0$

$$\lambda = - \frac{\left(\frac{\partial f}{\partial \sigma} \dot{\sigma} + \frac{\partial f}{\partial T} \dot{T} \right)}{\left(\frac{\partial f}{\partial \eta_p} + \frac{\partial f}{\partial k} \sigma^t \right) \frac{\partial f}{\partial \sigma^t}} \tag{70}$$

The generality af this farmulatian leaves ample room far experimentol evidence which can be described by a variety af yield canditians and flow rules.

Ta be specific, the fallowing discussion is concerned with the extensian af the von Mises yield candition and the Prandtl-Reuss flaw rule ta nan-isothermal conditions with temperature dependent hardening in extensian of the farmulatian presented in [38]. For cyclic loading conditians, Prager's method of kinematic hardening ar Ziegler's modificatior. of this theory could be easily incorporated ta account far the Bauschinger effect. Furthermore, in case af non-metallic materials the yield canditian and flow rule could be adjusted ta include the effects of hydrostatic pressure similar ta Drucker's farmulatian far materials with internal friction and cohesion.

The formulatian below is restricted ta isotropic conditions where

$$E = 2G \left(I + \frac{\nu}{1-2\nu} e_{s,s} e^t_{s,s} \right) \tag{71}$$

For duality of stress and strain transformations the corresponding stress and strain vectors are defined as follows

$$\sigma = \{ \sigma_{xx} \;\; \sigma_{yy} \;\; \sigma_{zz} \;\; \sqrt{2}\,\sigma_{xy} \;\; \sqrt{2}\,\sigma_{yz} \;\; \sqrt{2}\,\sigma_{zx} \}$$
$$\gamma = \{ \gamma_{xx} \;\; \gamma_{yy} \;\; \gamma_{zz} \;\; \tfrac{1}{\sqrt{2}}\gamma_{xy} \;\; \tfrac{1}{\sqrt{2}}\gamma_{yz} \;\; \tfrac{1}{\sqrt{2}}\gamma_{zx} \} \qquad (72)$$

For thermo-elasto-plastic behaviour the total deformation can be decomposed into

$$\dot{\gamma} = \dot{e} + \dot{\eta}_p + \alpha \dot{T} \qquad (73)$$

For metals experimental evidence indicates that the initial yield condition and plastic deformation are not affected by hydrostatic stress and that plastic deformation takes place without volume change. Therefore, it is convenient to decompose stress and strain into hydrostatic and deviatoric components

$$\sigma = \sigma_H + \sigma_D \qquad \text{and} \qquad \gamma = \gamma_H + \gamma_D \qquad (74)$$

For isotropic materials the initial yield function may be expressed entirely in terms of the invariants of the deviatoric stresses. The most widely used yield function, also known under the name of von Mises, describes the incipience of plastic deformation by the second invariant

$$f = \tfrac{3}{2}\sigma_D^t \sigma_D - \bar{\sigma}^2 = 0 \quad \rightarrow \quad \bar{\sigma}^2 = \tfrac{3}{2}\sigma_D^t \sigma_D \qquad (75)$$

This yield condition also provides a definition for the scalar function $\bar{\sigma}$, the so-called von Mises equivalent stress, in terms of the deviatoric stress components. The associated flow rule yields then an expression for the plastic strain increment in terms of deviatoric stresses and an unknown scalar function λ

$$\dot{\eta}_p = \lambda \frac{\partial f}{\partial \sigma^t} = 3\lambda \sigma_D \qquad (76)$$

This scalar λ can be determined either with laborious manipulations from equation (70) or directly from the workhardening expression

$$\dot{k} = \sigma_D^t \dot{\eta}_p = \bar{\sigma}\,\dot{\bar{\eta}}_p \quad \rightarrow \quad \dot{\bar{\eta}}_p^2 = \tfrac{2}{3}\dot{\eta}_p^t \eta_p \qquad (77)$$

This definition of the equivalent strain $\dot{\bar{\eta}}_p$ ensures that the plastic work remains the same for uniaxial and multiaxial conditions. Substituting equation (76) into equation (77) one obtains readily

$$\lambda = \frac{1}{2\bar{\sigma}}\,\dot{\bar{\eta}}_p \qquad (78)$$

Inserting λ into the flow rule (76) leads subsequently to the following expression for the plastic strain increment

$$\dot{\eta}_p = \dot{\bar{\eta}}_p\, s \qquad \text{with} \qquad s = \frac{3}{2\bar{\sigma}}\,\sigma_D \qquad (79)$$

In thermal environments the equivalent plastic strain is a function of stress and temperature

$$\bar{\eta}_p = \bar{\eta}(\bar{\sigma}, T) \quad \rightarrow \quad \dot{\bar{\eta}}_p = \frac{1}{f}\dot{\bar{\sigma}} + g\dot{T} \qquad (80)$$

with

$$f = \frac{\partial \bar{\sigma}}{\partial \bar{\eta}_p} \qquad \text{and} \qquad g = \frac{\partial \bar{\eta}_p}{\partial T} \qquad (81)$$

The physical significance of the parameters ζ and φ is illustrated in Fig. 25. This formulation was first presented in connection with the elasto-plastic matrix displacement analysis in [32].

There are two possibilities to proceed from now on: The equivalent plastic strain-rate can be defined either in terms of the effective stress rate or directly in terms of the total strain rate, leading to the following initial load schemes.

"Initial Strain Approach" $\quad \dot{\eta}_p = \dot{\eta}_p(\mathbf{s}^t\dot{\boldsymbol{\sigma}}, \dot{T})$

This formulation describes the equivalent plastic strain-rate in terms of the increment of the elastic energy of distortion

$$\dot{\eta}_p = \frac{1}{\zeta}\dot{\bar{\sigma}} + \varphi\dot{T} = \frac{1}{\zeta}\mathbf{s}^t\dot{\boldsymbol{\sigma}} + \varphi\dot{T} \tag{82}$$

leading to the following form of the flow rule

$$\dot{\boldsymbol{\eta}}_p = \frac{1}{\zeta}\mathbf{s}\mathbf{s}^t\dot{\boldsymbol{\sigma}} + \varphi\mathbf{s}\dot{T} \tag{83}$$

The total deformation rate can now be expressed by

$$\dot{\boldsymbol{\gamma}} = \dot{\mathbf{e}} + \dot{\boldsymbol{\eta}}_p + \boldsymbol{\alpha}\dot{T} = \mathbf{E}^{-1}\dot{\boldsymbol{\sigma}} + \frac{1}{\zeta}\mathbf{s}\mathbf{s}^t\dot{\boldsymbol{\sigma}} + (\varphi\mathbf{s} + \boldsymbol{\alpha})\dot{T} \tag{84}$$

Hence, the desired strain-stress-temperature rate relationship can be written as

$$\dot{\boldsymbol{\gamma}} = \mathbf{E}^{-1}\left(\mathbf{I} + \frac{2G}{\zeta}\mathbf{s}\mathbf{s}^t\right)\dot{\boldsymbol{\sigma}} + (\boldsymbol{\alpha} + \varphi\mathbf{s})\dot{T} \tag{85}$$

Note the additional plastic loading term $\varphi\mathbf{s}\dot{T}$ arising from the temperature dependence of the hardening law, see equation (80).

"Initial Stress Approach" $\quad \dot{\eta}_p = \dot{\eta}_p(\mathbf{s}^t\dot{\boldsymbol{\gamma}}, \dot{T})$

This formulation describes the equivalent plastic strain-rate in terms of the increment of the total energy of distortion

$$\begin{aligned}\dot{\eta}_p &= \frac{1}{\zeta}\mathbf{s}^t\mathbf{E}\left(\dot{\boldsymbol{\gamma}} - \dot{\boldsymbol{\eta}}_p - \boldsymbol{\alpha}\dot{T}\right) + \varphi\dot{T} \\ &= \frac{2G}{\zeta + 3G}\mathbf{s}^t\dot{\boldsymbol{\gamma}} + \frac{\zeta}{\zeta + 3G}\varphi\dot{T}\end{aligned} \tag{86}$$

leading to the following form of the flow rule

$$\dot{\boldsymbol{\eta}}_p = \frac{2G}{\zeta + 3G}\mathbf{s}\mathbf{s}^t\dot{\boldsymbol{\gamma}} + \frac{\zeta}{\zeta + 3G}\varphi\mathbf{s}\dot{T} \tag{87}$$

This forms the basis of the desired thermo-elasto-plastic stress-strain-temperature rate law, where the effective stress rate is defined by

$$\dot{\bar{\sigma}} = \mathbf{E}\,\dot{\mathbf{e}} = \mathbf{E}\left(\dot{\boldsymbol{\gamma}} - \dot{\boldsymbol{\eta}}_p - \boldsymbol{\alpha}\dot{T}\right) \tag{88}$$

$$\dot{\bar{\sigma}} = \mathbf{F}\dot{\boldsymbol{\gamma}} + \mathbf{G}\dot{T} \tag{89}$$

The "tangential" thermo-elasto-plastic material behaviour is described by $\mathbf{F}$ and $\mathbf{G}$, where

$$F = E\left(I - \frac{2G}{\beta + 3G}\, ss^t\right) \tag{90}$$

and

$$G = -E\left(\alpha + \frac{\beta}{\beta + 3G}\, \eta s\right) \tag{91}$$

This thermo-elasto-plastic constitutive law furnishes the desired linear relationship between stress-, total strain- and temperature rates, which is strictly valid only in the differential sense.

It should be noted that the first terms on the right hand sides of equation (90) and (91) represent the linear elastic portion of the constitutive law and that the second terms constitute their corrections due to plasticity. It is easily verified that the tangential material law in equation (90) is the inverse of the first term on the right-hand side of the total strain rate relationship, equation (85) .

b) Extension to Include Thermal Creep

The following discussion is restricted to the engineering theory of creep for metals in which it is assumed that the material does not possess any memory. This characterization of creep behaviour has very much in common with the theory of plasticity. There, strain increments are considered only in respect to a monotonically increasing parameter instead of real physical time as in the case of creep. In the following it is supposed that there is no interaction between the time independent plastic deformations which are thought of as occurring instantaneously, and the creep deformations developing during a time interval. It should be emphasized that this working hypothesis of no coupling between plastic and creep deformation is open to questions. In the formulation of a unified theory [22, 23] for rate sensitive dissipative processes the uncoupled engineering theories of creep and plasticity play the role of two extreme abstractions; but at the present it is only for those that quantitative experimental data are available which can be used in structural analysis.

The total deformation rate $\dot{\gamma}$ can be decomposed into instantaneous and time dependent components

$$\dot{\gamma} = \dot{\varepsilon} + \dot{\eta}_p + \alpha \dot{T} + \dot{\eta}_c \tag{92}$$

In contradistinction to the plastic strain-rate $\dot{\eta}_p$ which is derived from the current state variables T, σ or γ and their time derivatives, the creep strain-rate $\dot{\eta}_c$ is fully defined by the state of temperature T, stress σ and a hardening parameter as shown below. The creep strain-rate forms hence a prescribed initial strain-rate similar to that due to temperature changes.

Corresponding to plasticity it is assumed that only deviatoric components contribute to the energy dissipation in case of creep in metals. Hence, the von Mises yield criterion and the Prandtl-Reuss flow rule provide a consistent tool to describe the energy dissipation of the tensor field variables by equivalent uniaxial quantities, see equations (75), (77), (79).

Equivalent stress $$\bar{\sigma}^2 = \frac{3}{2}\, \sigma_D^t \sigma_D$$

Equivalent creep strain-rate $$\dot{\bar{\eta}}_c^2 = \frac{2}{3}\, \dot{\eta}_c^t \eta_c \tag{93}$$

Flow rule $$\dot{\eta}_c = \dot{\bar{\eta}}_c\, s \qquad \text{where} \qquad s = \frac{3}{2\bar{\sigma}}\, \sigma_D$$

The question remains then how to compute the effective creep strain-rates $\dot{\bar{\eta}}_c$ for transient stresses and temperature conditions. In the following development attention is restricted to the description of uniaxial creep which can be extended to multi-axial states over the mechanism of equivalent stresses and strains.

In the present approach it is assumed that there exists a mechanical equation of state which describes the creep rate of the solid as a function of state variables [24,25,26]

$$\dot{\eta}_c = f(\sigma, T, t, \eta_c) \tag{94}$$

This equation implies that at any given time t the creep rate depends on the current state of the system and is independent of the path or previous history. However, similar to the problems arising in context with the deformation theory of plasticity there is conclusive evidence that the creep rate $\dot{\eta}_c$ does also depend on history of stress, deformation and temperature, hence the memory theory of viscoelastic-plastic solid would form a more appropriate way to describe this type of material response. The question arises then how the creep data, obtoined from constant stress and temperature experiments, could be used for the case of time-varying stresses and temperatures. Different cumulative creep theories have been proposed; however, only the well-known time-hardening and strain-hardening rules are considered below.

Time-Hardening Law

The time-hardening rule assumes that the creep rate depends only on the current stress, the current temperature and the time from the beginning of the creep process. The creep rate at any instant of time is then detined by the following equation of state

$$\dot{\eta}_c = f(\sigma, T, t) \tag{95}$$

The creep strain may be frequently written as a power function of time

$$\eta_c = f_1(\sigma)\, f_2(T)\, t^{\rho} \tag{96}$$

such that

$$\dot{\eta}_c = f_1(\sigma)\, f_2(T)\, \rho\, t^{\rho-1} \tag{97}$$

Note that for secondary creep, where $\rho = 1$, the creep rate remains constant in time.

Strain-Hardening Law

The strain-hardening rule assumes that the creep rate depends only on the current stress, the current temperature and the accumulated creep strain. The creep rate at any instant of time is then defined by the following equation of state

$$\dot{\eta}_c = f(\sigma, T, \eta_c) \tag{98}$$

Elimination of the time t from equation (96) and (97) leads to the creep rate expression below

$$\dot{\eta}_c = \rho \left[f_1(\sigma)\, f_2(T) \right]^{1/\rho} \eta_c^{(\rho-1)/\rho} \tag{99}$$

Note that for secondary creep, the time-hardening and the strain-hardening method yield identical formulae. Fig. 26 illustrates the difference of both hardening rules for a varying stress and temperature conditions at time $t = t_1$. Although experimental data rather supports the strain-hardening rule, this method may lead to difficulties in describing creep during unloading if the accumulated creep strain is large. Hence, the simpler time-hardening method

is generally favoured for numerical calculations involving primary creep. For multi-axial conditions the rate of creep strain components are determined from equation (93) simply by substituting for $\dot{\eta}_c$ either the time-hardening or the strain-hardening laws of equations (97) or (99)

$$\dot{\eta}_c = \begin{cases} \{ f_1(\sigma)\, f_2(T)\, \rho\, t^{\rho-1} \}\, s & \text{Time-hardening} \\ \{ \rho\, [f_1(\sigma)\, f_2(T)]^{1/\rho}\, \dot{\eta}_c^{(\rho-1)/\rho} \}\, s & \text{Strain-hardening} \end{cases} \tag{100}$$

The final rate formulation of an isotropic solid exhibiting elastic, plastic, thermal and creep effects can now be written as

$$\begin{aligned} \dot{\varepsilon} &= \dot{e} + \dot{\eta}_p + w\dot{T} + \dot{\eta}_c \\ &= E^{-1}\left(I + \frac{2G}{J} s s^t\right)\dot{\sigma} + (\gamma s + \alpha)\dot{T} + \dot{\eta}_c s \end{aligned} \tag{101}$$

An extension of the inverse relationship, equation (88), yields the following expression for the effective stress rate

$$\dot{\sigma} = E\dot{e} = F\dot{\varepsilon} + G\dot{T} - \dot{\eta}_c E s \tag{102}$$

For the definition of the tangential material properties F and G see equation (90) and (91). It should be emphasized again that the creep rate components have to be considered prescribed, since they are determined from the existing state of stress, temperature and time or equivalent creep strain and not from their time derivatives.

c) Numerical Implementation

The finite element analysis of the transient thermal problem involving plasticity and creep requires the treatment of a set of coupled first order differential equations. With the help of incremental methods this problem reduces to the solution of algebraic equations. Euler's algorithm forms a very simple procedure in which the information at the beginning of each time step is used for the incremental treatment of the differential equations. There are various alternatives possible, such as the modified Euler method or the improved Euler method besides the more sophisticated multistep time integration algorithms; however, the latter are unpractical as they create considerable storage problems and require a certain degree of smoothness of the state variables with time. The following discussion is restricted to simple incremental strategies for solving the quasistatic equilibrium equations. To this end the differential stress-strain relation (102), is recast into an incremental form for the finite time interval $t, t+t_\Delta$. The resulting equilibrium equation may then be written as

$$(K + K_\Delta)(t)\, r_\Delta(t) = R_\Delta(t) + R_{p\Delta}(t) + R_{\theta\Delta}(t) + R_{c\Delta}(t) \tag{103}$$

The incremental tangent stiffness $K_\Delta(t)$ may arise from the temperature dependent material properties under transient conditions as well as from the non-linearity due to plasticity. The increments of prescribed initial loads due to temperoture $R_{\theta\Delta}(t)$ and due to creep $R_{c\Delta}(t)$ can be evaluated directly from (19) for the given thermal condition and the mechanical state at the beginning of each time interval. These prescribed initial loads are simply added to the mechanical loading. The increment of the "instantaneous" plastic loads $R_{p\Delta}(t)$ arises either due to the temperature dependence of the hardening law in case of the direct "tangential stiffness" approach for the solution of the thermo-plastic problem, or altogether from the iterative treatment of the plasticity problem using initial load techniques.

The tangential stiffness method involves the repeated formation and decomposition of the structural stiffness matrix for each time increment, a rather costly strategy. Due to the finite size of the time interval further refinement is mostly required to avoid accumulation of linearization errors, hence multistep methods or equivalent iterations need to be employed together with equilibrium corrections before the beginning of a new time step. In case of temperature-independent properties modification techniques provide a very useful tool to account for localized plastifications. They furnish directly the desired information on changes of the results due to single changes in the stiffness without requiring the new formation and decomposition of the stiffness matrix [38, 39]. In case of temperature-dependent elastic properties under transient conditions the stiffness changes are not restricted to small areas, hence the modification procedures hardly apply. In case that the elastic properties remain independent of time equation (103) can be recast into the following form

$$K\, r_{\Delta}(t) = R_{\Delta}(t) + R_{P\Delta}(t) + R_{\sigma_0\Delta}(t) + R_{\varepsilon_0\Delta}(t) \qquad (104)$$

Creep and temperature only give rise to prescribed initial loads which add to the mechanical load increments. The non-linearity due to plasticity is treated using initial loaditeration in order to avoid the new formation and decomposition of the structural stiffness. This procedure accommodates a time variable system, in which the loading accounts for the time change of the structure due to creep and temperature dependent hardening as well as for the instantaneous changes due to plasticity [29, 30, 31, 32, 33, 34, 36, 37, 38].

As mentioned above the formulations of the plastic flow rule in equations (83) and (87) lead to two alternative initial load schemes for the iterative treatment of the incremental plasticity problem.

"Initial Strain Method"

For the initial strain approach the plastic strain increments are defined in terms of the increment of the elastic energy of distortion and in terms of the temperature increments, see equation (83). These plastic strain increments $\dot{\eta}_\tau$ are then utilized to construct the initial load increment $\dot{J}_\tau$ leading to an iterative process, since the current stress increments are not known a priori [38]. Fig. 27 a illustrates this formulation for the one-dimensional case. If convergence can be assured at all, this method provides better rates of convergence than the alternative initial stress approach; on the other hand, the initial strain method breaks down in case of non-hardening materials, where $\xi = 0$.

"Initial Stress Approach"

For the initial stress approach the plastic strain increments are defined in terms of the total energy of distortion and in terms of the temperature increments, see equation (87). These plastic strain increments $\dot{\eta}_\tau$ are then utilized to construct the initial load increment $\dot{J}_\tau$ leading to an iterative process, since the current total strain increments are not known a priori. Fig. 27 b illustrates this formulation for the one-dimensional case. This procedure was presented first in form of initial stresses in [37]. Note that the expression for the flow rule (87) clearly indicates that this method also applies to non-hardening plastic material where $\xi = 0$; hence this formulation enlarges considerably the range of convergence in comparison to the associated initial strain approach. It is interesting to note that the differences between the initial stress and initial strain schemes arise solely from the alternative descriptions of $\dot{\eta}_\tau$ in equation (82) and (86).

For a detailed theoretical discussion of the numerical implications, such as range and rate of convergence, the reader is referred to [38], in which the iterative initial load techniques are investigated. Basically, three methods may be distinguished which are denoted in short by DIM, the direct incremental method, NIM, the normal iterative method, and VIM, the improved iterative method. DIM is based solely on load incrementation without iteration where the magnitudes of the current elastically suppressed plastic strains are estimated directly from the previous load increment. NIM incorporates additionally within each load increment an iterative scheme accounting for the error made due to the finite size of the load increment. VIM is a Newton-Raphson procedure applied to the plastic variables in order to accelerate the convergence properties of NIM. A schematic representation of these initial load techniques is given in Fig. 28 [35, 36, 38]. It should be mentioned that none of these methods alters the initial elastic stiffness.

In the following the results of two elasto-plastic problems are presented to illustrate these iterative schemes. The number of iterations necessary to achieve a preset accuracy of plastic strains or stresses is of particular interest, as it gives an indication of the range and rate of convergence for the individual methods. The first example deals with the elasto-plastic analysis of the statically indeterminate truss, illustrated in Fig. 29, while the second one involves the rectangular membrane with central crack, shown in Fig. 30. For both cases a modification of the Ramberg-Osgood equation describes the uniaxial stress-strain law

$$\eta = \frac{1.1\sigma_y}{mE}\left[\left(\frac{\sigma}{1.1\sigma_y}\right)^m - \left(\frac{1}{1.1}\right)^m\right] \qquad (105)$$

The following material constants are chosen to characterize the aluminium alloy 2024-T3: Young's modulus $E = 11.4 \times 10^6$ lbf/in^2, initial yield stress $\sigma_y = 34\,500$ lbf/in^2, exponent $m = 10$ and elastic Poisson's ratio $\nu = 0.3$. An elastic solution determines the loading at which local yielding occurs first while the subsequent load increments are described by the non-dimensional scalar load factor λ which is incremented by $\lambda_\Delta = 0.1$.

Discussion of Results:

Tables 2 and 3 summarize the results of the elasto-plastic analyses of both structures shown in Fig. 29 and 30. Both tables present the number of iterations necessary to obtain a relative error e less than 10^{-4} for the equivalent plastic strain-rates in equation (82) or (86), where

$$e = \frac{|\dot{\bar{\eta}}_{i+1} - \dot{\bar{\eta}}_i|}{\max(\dot{\bar{\eta}}_{i+1})} = \frac{|\Delta\dot{\bar{\eta}}_i|}{\max(\dot{\bar{\eta}}_{i+1})} < 10^{-4} \qquad (106)$$

In addition the maximum equivalent stress is given for each loading step λ. Both iterative procedures, NIM and VIM, are applied to the initial stress and initial strain formulation of the non-linear elasto-plastic problem.

Table 2 presents the results for the simple truss of Fig. 29 verifying the theoretical predictions made in [38] in regard to the convergence properties of the normal iterative method. The range of convergence of NIM is larger when the initial stress formulation is used versus the initial strain formulation, which diverges already at $\lambda = 1.5$. However, if it converges at all, the initial strain formulation converges faster. Moreover, VIM, the Newton-Raphson modification of NIM, improves considerably the rate of convergence if applied to both, the initial stress and initial strain approach. It should be noted that in case of VIM and initial strain, the range of convergence is also considerably enlarged in comparison to NIM.

Load-factor λ	Maximum Effective Stress $\bar{\sigma}_{max}$ [lbf/in²]	Initial Stress Approach		Initial Strain Approach	
		NIM	VIM	NIM	VIM
1.0	34980	-	-	-	-
1.1	38930	5	4	22	5
1.2	38980	5	4		4
1.3	43690	6	4		5
1.4	42180	6	5		5
1.5	43630	5	4		5
1.6	44570	7	5		5
1.7	45340	7	5	Divergence	5
1.8	46390	7	5		5
1.9	47170	8	5		5
2.0	47980	7	5		5

Table 3 Rectangular Membrane with Crack - Number of Iterations for Relative Error of Plastic Strains $< 10^{-4}$

Load-factor λ	Maximum Effective Stress $\bar{\sigma}_{max}$ [lbf/in²]	Initial Stress Approach		Initial Strain Approach	
		NIM	VIM	NIM	VIM
1.0	34980	-	-	-	-
1.1	37730	10	4	6	2
1.2	40740	12	3	6	3
[illegible]	[illegible]	15	3	9	2
1.4	45890	20	3	17	3
[illegible]	[illegible]	26	3		2
[illegible]	[illegible]	35	3		2
[illegible]	[illegible]	55	3		2
1.8	51620	79	3		3
1.9	57060	116	3	Divergence	3
2.0	63410	163	3		3

Table 2 Indeterminate Truss - Number of Iterations for Relative Error of Plastic Strains $< 10^{-4}$

Table 3 presents the results of the rectangular membrane with central crack illustrated in Fig. 30. The comparison for this structure agrees very well with the theoretical predictions made in [38]. NIM-initial stress now requires far less iterations than in the case of the truss. Only at relatively high loads the number of iterations can be markedly reduced by the application of VIM. Whilst NIM-initial stress converges more quickly when the number of constraints is greater, the converse is true for NIM-initial strain. Even for $\lambda = 1.1$ there are 22 iterations necessary and divergence occurs already for $\lambda = 1.2$. Note that VIM-initial strain requires in most loading steps the same number of iterations as VIM-initial stress.

In conclusion it can be stated that VIM improves considerably the range of convergence of the initial strain approach as well as the rate of convergence for structures with few kinematic constraints. In case of continuum problems the initial stress formulation seems to be suited best for automatic computations. It should be mentioned that the improved iterative method VIM is considerably more costly per iteration than NIM depending on the number of Newton-variables involved. An optimum approach in regard to computational effort is adopted if Newton variables are considered only above a certain threshold.

Two more remarks may be pertinent to the discussion of different initial load techniques: the first concerns the treatment of thermo-elasto-plastic problems under transient temperature conditions, while the second deals with the choice of time steps for the creep analysis. From equation (80) one can see that there are two sources for plastic loading in the case of transient temperature distributions, the customary mechanical source due to the change of equivalent stress $\dot{\bar{\sigma}}$ and the thermal source due to changes of temperature $\dot{T}$. Fig. 25 a illustrates the two paths which are possible for transient temperature conditions. It clearly indicates that very large increments of plastic strains are to be expected for temperature loading, path O-A, leading to numerical stability problems particularly for materials with little strain-hardening. On the other hand, unloading due to a temperature decrease from T_2 to T_1 is simply accounted for by increasing the current $\bar{\sigma}_{max}$ corresponding to point O by the value at B. It should be mentioned that a similar phenomenon occurs in the case of primary creep analysis if cyclic loading is dealt with by the strain-hardening method. There unloading is being presented by path A-O in Fig. 26 a, where O may not be reached any more if sufficient strain-hardening has preceded unloading.

The last comment pertains to the choice of time steps used in the creep analysis. Here the same stability problems occur as in the classical numerical treatment of plasticity problems with the aid of the direct incremental method, denoted above by DIM. Experience has indicated that the increments of creep strains should be less than the elastic strains to insure stability, see also [29]

$$\dot{\bar{\eta}}_c < \mu \dot{\bar{\varepsilon}} = \mu \frac{\dot{\bar{\sigma}}}{E} \qquad \text{with} \qquad 0 \leq \mu \lesssim 0.25 \tag{107}$$

For numerical computations the simplest method is to derive from the prevailing state of elastic strain an admissible increment of creep strains which can be easily transformed into a physical time step via equation (97). It should be emphasized again that there is o basic difference between plastic and creep strain increments due to the homogeneous and non-homogeneous definition of these quantities. Iterative techniques similar to NIM and VIM do not apply to the virtual work formulation of the creep problem in which it is tacitly assumed that the state of stress or total strain remains constant within a given time interval for the rate of creep or corresponding relaxation method. To this end an extension of the Hellinger-Reissner variational principle [25], [28] could provide an alternative platform to deal in a more efficient way with the creep analysis of structures.

d) Example

A realistic thermal stress analysis is performed with the help of the initial load technique on a spherical pressure vessel with radial nozzle. The inside of the vessel is subjected to a sudden temperature rise for which the steady state temperature distribution is first determined. Subsequently, the instantaneous thermo-elasto-plastic stress distribution is evaluated for the vessel accounting for temperature dependent material properties. Thereafter, the stress redistribution due to creep is traced as function of time. Finally, in addition the elasto-plastic stress analysis is carried out for internal pressure under which limit design would predict partial collapse.

The meridional section of the structure is idealised by 230 axisymmetric TRIXC6 elements resulting in 551 nodal points. The same discretisation is employed for both, the temperature and the subsequent stress analysis, see Fig. 31.

The thermal boundary value problem is defined for the steady state case as follows. The inside surface of the vessel is subjected to a prescribed temperature of 250° C, while the ambient air temperature on the outside is 20° C. The upper surface of the nozzle flange and the bottom rim of the idealised structure are insulated against heat flow.

The following thermal material properties are chosen:

Thermal conductivity $k = 0.398$ kcal/cmh °C

Specific heat $c = 0.114$ kcal/kg °C

Heat transfer coefficient between the outside surface of the vessel and the air

$h = 0.0426$ kcal/cm^2 h °C

Density $\varrho = 7.9 \times 10^{-3}$ kg/cm^3

Fig. 33 illustrates the resulting steady state temperature distribution in form of contour lines.

For the subsequent mechanical analysis of the structure the material constants are taken for mild steel where the elastic properties are defined as

$$E = 21000 \text{ kp/mm}^2 \quad \text{and} \quad \nu = 0.30$$

The coefficient of thermal expansion is assumed to be temperature dependent

$$\alpha = \alpha_{10} - (\alpha_{10} - \alpha_0)\left(1 - \frac{\bar{T}}{T_1}\right)^2 \tag{108}$$

where

$$\alpha_0 = 1.0 \times 10^{-5} \text{ 1/°C}$$

$$\alpha_{10} = 1.7 \times 10^{-5} \text{ 1/°C}$$

$$T_1 = 720 \text{ °C}$$

and its variation is accounted for during the course of incremental loading which is required for the thermo-elasto-plastic stress analysis.

In the plastic range a Ramberg-Osgood type stress-strain law as given in Equation (105) is used. The temperature dependence of the yield stress is described by

$$\sigma_y = \sigma_{y_0}\left[1 - \left(\frac{\bar{T}}{\bar{T}_0}\right)^m\right] \tag{109}$$

where

$$\sigma_{y_0} = 26 \text{ kp/mm}^2$$

$$\bar{T}_0 = 850 \text{ °}$$

$$m = 20 \qquad \text{see Equation (105)}$$

Fig. 32 shows the unaxial stress-plastic strain relationship as function of the temperature.

The incremental procedure starts from a stress free reference state with applying a fraction of the steady state temperature distribution for each loading step. The final temperature condition is reached after 13 increments. Subsequently, the stepwise calculation is continued to account for secondary creep effects. The following creep law given in [24] is chosen to compute the initial creep strain increment at each time step

$$\eta_{cA} \cdot \frac{\left[E\,[\text{kp/mm}^2]\right]^{4.5}}{2390} \frac{[\text{sec}]}{0^6} \; \mathbf{S} \; ; \quad \mathbf{S} \cdot \frac{3}{2\bar{\sigma}} \boldsymbol{\sigma}_D \tag{110}$$

The time interval was automatically adjusted such that the change of the maximum equivalent stress occurring in the structure never exceeded 5% of the instantaneous maximum equivalent stress. The procedure was terminated after 21 increments which corresponds to a time of 181 days.

Fig. 34, 35 and 36 illustrate the resulting stress distributions. One observes that the plastic zones (shaded areas) develop first at the inside of the flange. This may be explained by the relatively low temperature at the outside of the flange which prevents the expansion of the hot inner partion and thus induces large compressive hoop stresses. Plastification of the inter-

section between the nozzle and the sphere, where pronounced stress concentrations occur, starts only at a later loading stage. Note that in the case of a temperature dependent yield surface the maximum equivalent stress is not alone responsible for plastic flow, but rather a combination of the state of stress and temperature. Figs. 35 and 36 illustrate the stress relaxation due to creep causing a significant redistribution of stresses in the flange.

For comparison, the spherical shell with nozzle is also subjected to a monotonically increasing pressure whereby the temperature and creep effects are disregarded. Fig. 37 shows the equivalent stress distribution for the pressure $p = 33.4\ kp/cm^2$ at the beginning of plastification. The internal pressure is further increased to $90.2\ kp/cm^2$ for which the theory of limit design predicts a collapse mechanism at the junction between the nozzle and the sphere. The stress distribution for this limit pressure, which is attained after 17 loading steps, is shown in Fig. 37. It once again indicates that limit design procedures yield but a conservative lower bound.

In conclusion one may state that an important class of thermomechanical problems can be solved, accounting for plasticity as well as creep by the initial load technique without altering the initial stiffness. The extension of this type of analysis to transient temperature conditions may seem trivial, but experience has shown that particularly the temperature dependence of the yield surface in general leads to serious convergence problems of the iterative scheme, especially if there is little strain hardening. Particular care has to be exercised in choosing an appropriate time-step for the associated heat conduction analysis in order to avoid any spurious oscillation of the temperature evolution in time and to keep small the plastic loading term due to temperature changes.

VI ACKNOWLEDGEMENTS

This paper summarizes some results of an investigation into the finite element analysis of thermal stress problems. This research has been supported by a number of organisations, including the "Bundesministerium für Bildung und Wissenschaft" (grant SBB 4), Bonn.

The authors wish to thank all members of the ISD for their continuous support in the preparation of this paper; in particular the advice and help by Dr. K.E. Buck is gratefully acknowledged.

VII. REFERENCES

[1] Fung, Y.C., Foundations of Solid Mechanics, Englewood Cliffs, N.J.: Prentice-Hall Inc., 1965.

[2] Parkus, H., Thermoelasticity, Blaisdell Publ.Co., Waltham, Mass., 1968.

[3] Nickell, R.E. and Sackman, J.L., Approximate Solutions in Linear, Coupled Thermoelasticity, J. Appl. Mech., June 1968.

[4] Chao-Hwang Lu, Nonlinear Theory of Thermoelasticity, Ph.D. Dissertation to Division of Structural Engineering and Structural Mechanics, University of California, Berkeley, 1971.

[5] Oden, J.T., and Kross, D.A., Analysis of General Coupled Thermoelasticity Problems by the FEM, Proc. 2nd Conf. on Matrix Methods in Struct. Mech., Wright-Patterson Air Force Base, Ohio, 1968.

[6] Zienkiewicz, O.C. and Cheung, Y.K., Finite Elements in the Solution of Field Problems, The Engineer, Sept. 1965.

[7] Wilson, E.L. and Nickell, R.E., Application of the FEM to Heat Conduction Analysis, Nuclear Eng. and Design, 4, 1966.

[8] Visser, W., A FEM for the Determination of Non-stationary Temperature Distributions and Thermal Deformations, Proc. 2nd Conf. on Matrix Methods in Struct. Mech., Wright Patterson Air Force Base, Ohio, 1965.

[9] Argyris, J.H., Brönlund, O.E., Fried, I., and Spooner, J.B., The Changes in the Stress Distribution round a Rectangular Hole with Rounded Corners, Caused by a Varying Internal Pressure and Temperature, Conf. of the Int. Astron. Fed., Belgrade, 1967.

[10] Argyris, J.H. and Mareczek, G., Thermomechanical Analysis of Structures, to be presented to the 4th Conf. of the Hung. Acad. of Sciences on Dimensioning and Strength Calculations, Budapest, October 1971.

[11] Argyris, J.H., Energy Theorems and Structural Analysis, Aircraft Eng. 26, 1954 27, 1956. Also as book: Butterworths London, 1960.

[12] Melosh, R.J., Basis for Derivation of Matrices for the Direct Stiffness Method, AIAA J., 1, 1963.

[13] Argyris, J.H., Recent Advances in Matrix Methods of Structural Analysis, Pergamon Press, London, 1964.

[14] Argyris, J.H., Triangulor Elements with Linearly Varying Strain for the Matrix Displacement Method, J. of the Roy. Aeron. Soc., 69, 1965.

[15] Argyris, J.H., The TRIAXℓ Element for Axisymmetric Analysis by the Matrix Displacement Method, The Aeron. J. of the Roy. Aeron. Soc., 70, 1966.

[16] Sternberg, E., Transient Thermal Stresses in on Infinite Medium with a Spherical Covity, Koninkl. Nederl. Akad. van Wetenschappen, Amsterdam, Proceed. B. 60, 1957.

[17] Nowinski, J., Transient Thermoelastic Problem for on Infinite Medium with o Sphericol Covity Exhibiting Temperature Dependent Properties, J. of Appl. Mech., June 1962.

[18] Everett, F.L. ond Miklovitz, Poisson's Ratio ot High Temperatures, J. of Appl. Phys., 15, 1944.

[19] Prager, W., Non-Isothermal Plostic Deformation, Proceed. Koninkl. Nederl. Akad. van Wetenschappen, Amsterdam, B. 61, 1958.

[20] Drucker, D.C., A Definition of Stable Inelostic Moteriol, J. of Appl. Mech., 1959.

[21] Naghdi, P.M., Stress-Strain Relotion in Plosticity ond Thermo-Plasticity, in "Plasticity", Proc. 2nd Symp. on Naval Struct. Mech., Pergamon, 1963.

[22] Naghdi, P.M. and Murch, S.A., On the Mechanicol Behaviour of Viscoelastic/Plostic Solids, Trans. ASME, Sept. 1963.

[23] Rice, J.R., On the Structure of Stress-Stroin Relations for Time-Dependent Plastic Deformation of Metals, Trans. ASME, Sept. 1970.

[24] Hult, J., Creep in Engineering Structures, Bloisdell Publ. Co., Woltham, Mass., 1968.

[25] Rabotnov, Y.N., Creep Problems in Structurai Members, North Hollond Publ. Co., 1969.

[26] Mendelson, A., Plosticity: Theory and Application, The MacMillan Co., New York, London, 1968.

[27] Wong, A.J. ond Prager, W., Thermol and Creep Effects in Work-Hardening Elastic-Plostic Solids, J. of Aeron. Sciences, May 1954.

[28] Pion, T.H.H., On the Voriotional Theorem for Creep, J. of the Aeronaut. Sciences, November 1957.

[29] Greenbaum, G.A. and Rubinstein, M.F., Creep Analysis of Axisymmetric Bodies Using Finite Elements, Nuclear Eng. and Design, 7, 1968.

[30] Gallagher, R.H., Padlog, J. and Bijlaard, P.P., Stress Analysis of Heated Complex Shapes. J. Am. Rocket Soc.,32, 1962.

[31] Argyris, J.H., Continua and Discontinua, Proc. 1st Conf. Matrix Methods Struct. Mech., Wright-Patterson Air Force Base, Ohio , 1965.

[32] Argyris, J.H., Elasto-Plastic Matrix Displocement Analysis of Three-Dimensional Continua. J. of the Roy. Aeron. Soc., 69, 1965.

[33] Argyris, J.H., Scharpf, D.W. and Spooner, J.B., Die elastoplastische Berechnung von allgemeinen Tragwerken und Kontinua, Ingenieur Archiv, Vol. 37, 1969.

[34] Argyris, J.H., Scharpf, D.W. and Spooner, J.B., The Elasto-Plostic Calculation of General Structures and Continua, Proceedings, 3rd Conference on Dimensioning, Budapest, 1969.

[35] Argyris, J.H., Buck, K.E., Scharpf, D.W., Hilber, H.M. and Mareczek, G. , Some New Elements for the Matrix Displocement Method, Proc. 2nd Conf. Matrix Methods Struct. Mech., Wright-Patterson Air Force Base, Ohio, 1968.

[36] Scharpf, D.W., Die Frage der Konvergenz bei der Berechnung elastoplastisch deformierbarer Tragwerke und Kontinua, Dr. Ing. Thesis, University of Stuttgart, 1969.

[37] Zienkiewicz, O.C., Valliapan, S., King, I.P., Elasto-Plastic Solutions of Engineering Problems, Initial Stress, Finite Element Approach, Ing. J. Num. Meth. Eng., 1969.

[38] Argyris, J.H and Scharpf, D.W., Methods of Elastoplastic Analysis, Proc. of the ISD-ISSC Symp. on Finite Element Techniques, Institut für Statik und Dynamik der Luft-und Raumfahrtkonstruktionen, Stuttgart, 1969, also to be published in ZAMP.

[39] Roy, J., Allgemeine Modifikationsverfahren für die Lineare und Nichtlineare Berechnung von Tragwerken und Kontinua mit der Matrizenverschiebungsmethode, Dr. Ing. Thesis, to be submitted at the University of Stuttgart, 1971.

[40] Irons, B.M., Numerical Integration Applied to FEM, Int. Symp. on the Use of Electronic Digitol Computers in Structural Eng. at the University of Newcostle upon Tyne, 1967.

[41] Ergatoudis, J., Irons, B.M. and Zienkiewicz, O.C., Curved, Isoparametric, Quadrilateral Elements for Finite Element Analysis, Int. J. Solids Structures, 7, 1968.

[42] Argyris, J.H. and Fried, I., The LUMINA Element for the Matrix Displacement Method, The Aeron. J. of the Roy. Aeron. Soc., 72, 1968.

[43] Argyris, J.H., and Scharpf, D.W., The Curved Tetrahedronal and Triangular Elements TEC and TRIC for the Matrix Displacement Method. Part I Small Displacements, Part II Large Displacements, The Aeron.J. of the Roy. Aeron.Soc., 73, 1969.

[44] Felippa , C.A., Refined Finite Element Analysis of Linear and Nonlinear Two-Dimensional Structures, UC-SESM Report No.66-22, University of California, Berkeley, October 1966.

[45] Willam, K.J., "Finite Element Analysis of Cellular Structures", Ph.D. Dissertation to Division of Structural Engineering and Structural Mechanics, University of California, Berkeley, 1969.

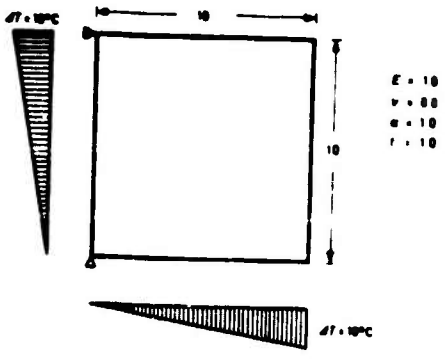


Fig 1 Quadratic Membrane
Description of "Residual Stress" Problem

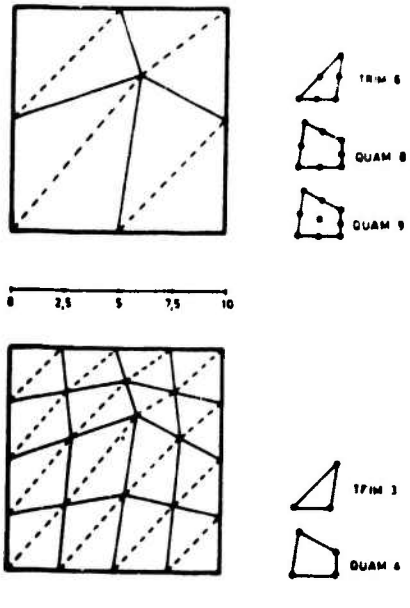


Fig.3 Discretisation of Phase II
"Linear" Jacobian

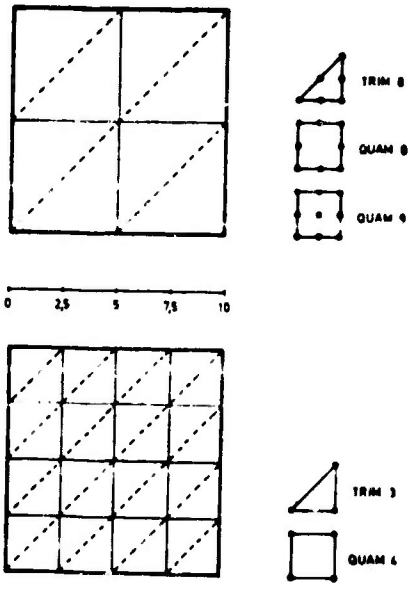


Fig. 2 Discretisation of Phase I
"Constant" Jacobian

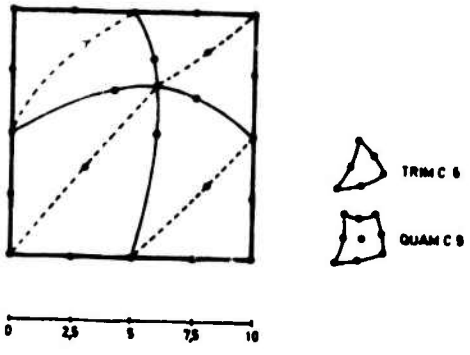
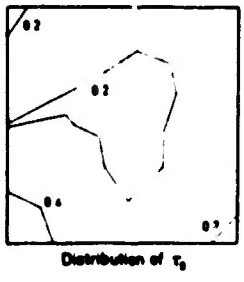
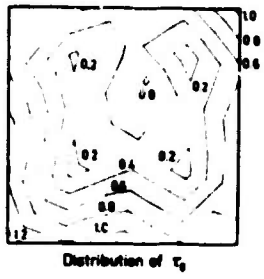


Fig. 4 Discretisation of Phase III
"Quadratic" Jacobian



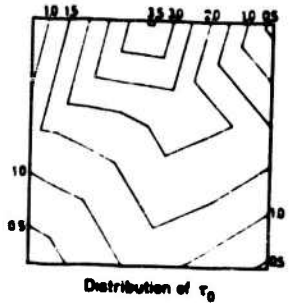
Distribution of τ_0

Fig. 5 Phase II, Element TRIM 3



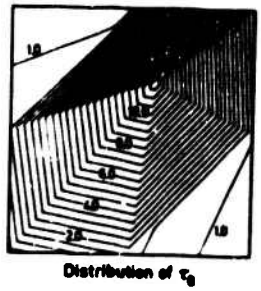
Distribution of τ_0

Fig.6 Phase II, Element QUAM 4



Distribution of τ_0

Fig. 7 Phase III, Element TRIMC 6



Distribution of τ_0

Fig. 8 Phase III, Element QUAMC 9

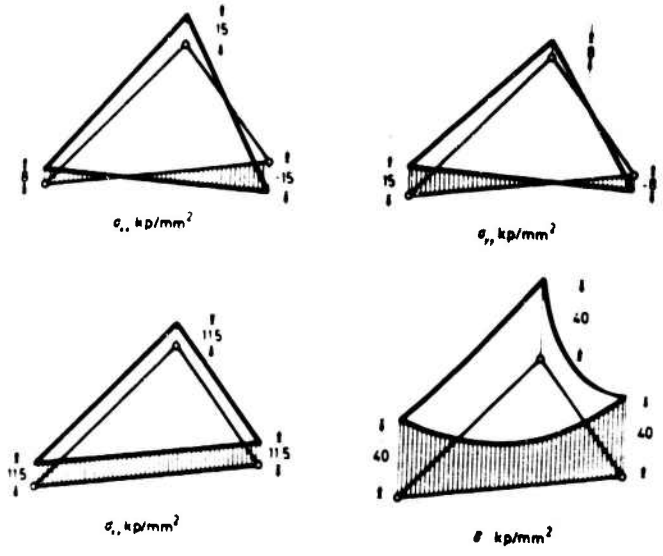


Fig 9 Stress Distribution
Deviatoric Components and Equivalent Stress

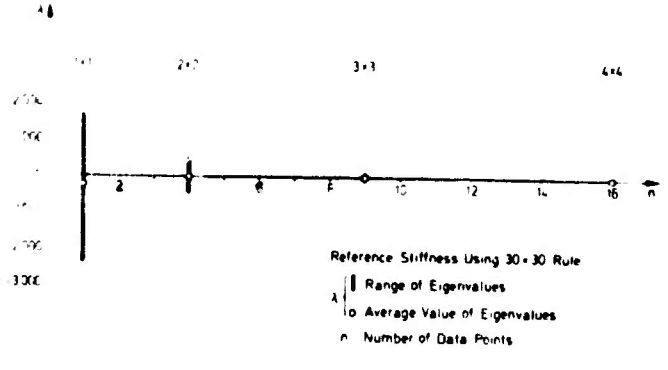


Fig 11 Comparison of Elasto-Plastic Stiffness Matrices
Different Gaussian Quadrature Rules

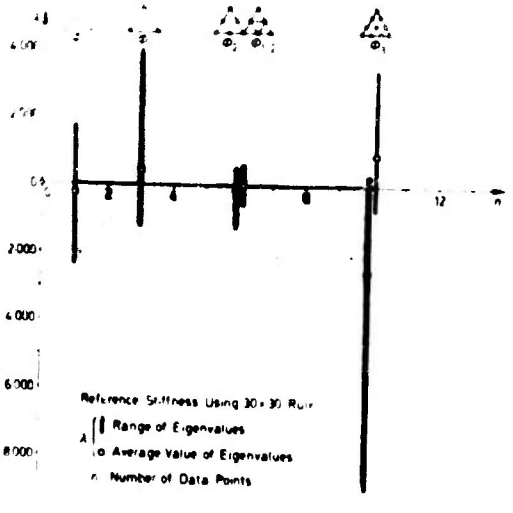


Fig 13 Comparison of Elasto-Plastic Stiffness Matrices
Different Interpolation Schemes

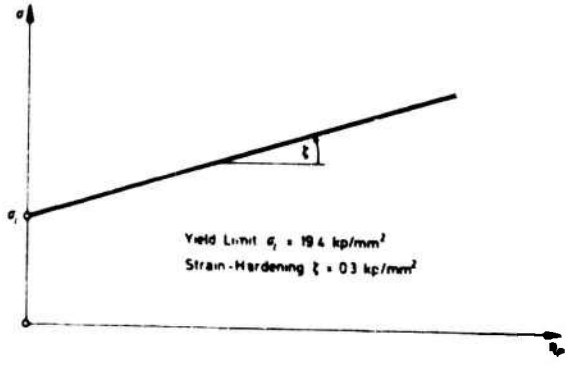


Fig 10 Stress-Plastic Strain Relationship

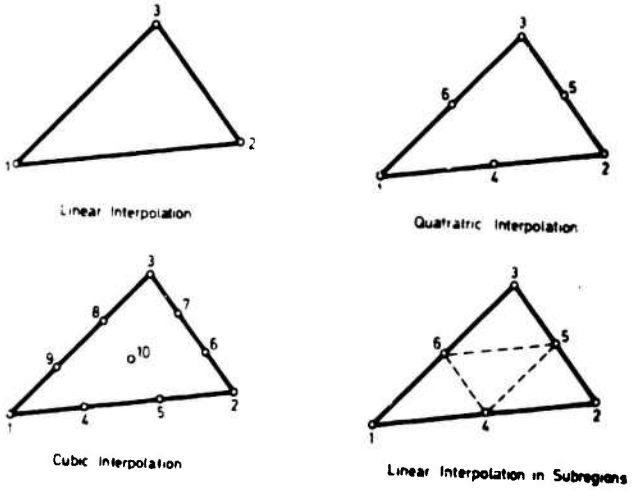


Fig 12 Representation of Elasto-Plastic Material Properties
Linear Strain Triangle TRIM 6

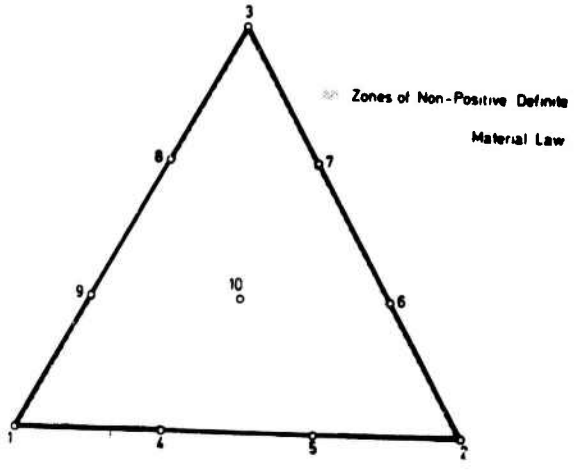
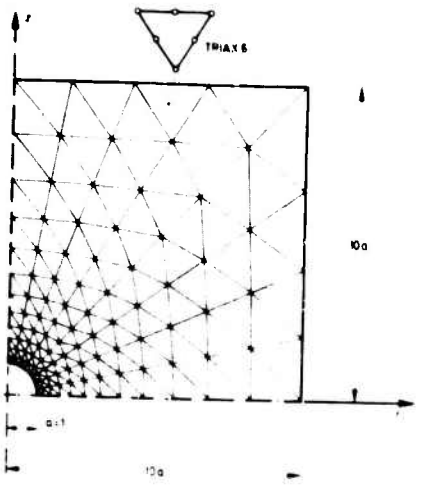


Fig 14 Cubic Interpolation Scheme



625 Nodal Points
192 TRIAX 6 Elements

Fig.16 Circular Cylinder with Spherical Cavity
Idealisation

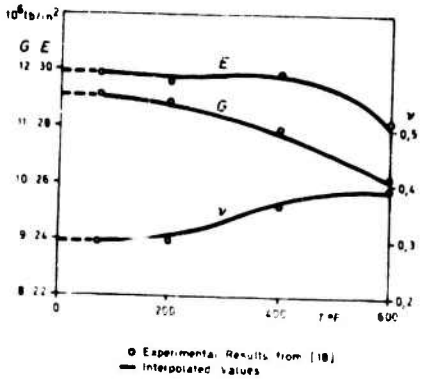


Fig.17 Temperature Dependence of Elastic Properties
Variation of E, ν and G

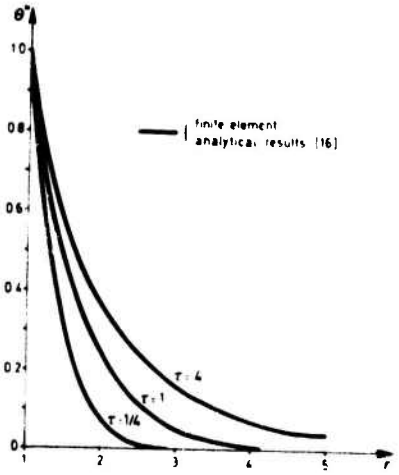


Fig.18 Transient Temperature
Radial Distribution at Different Times τ

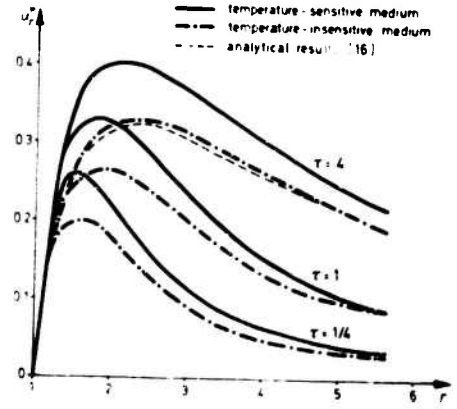


Fig.19 Radial Displacements
Radial Distribution at Different Times τ

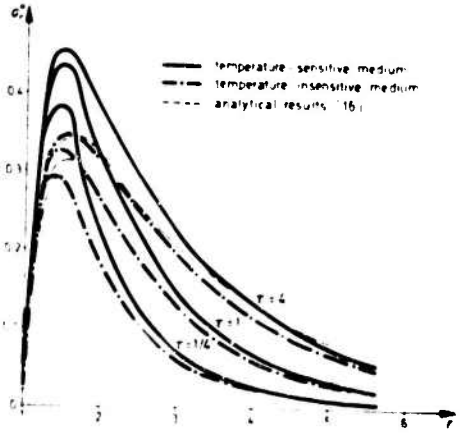


Fig.20 Radial Stresses
Radial Distribution at Different Times τ

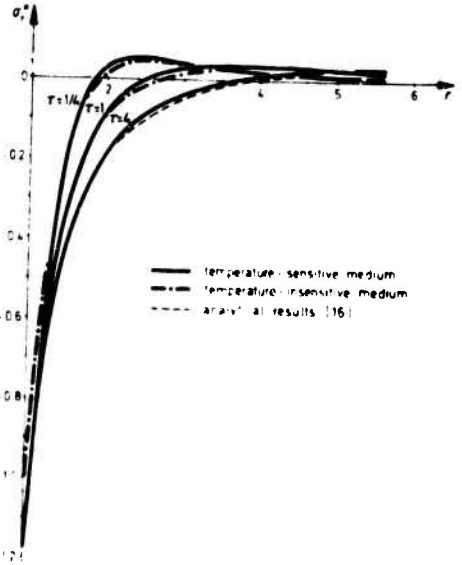


Fig.21 Tangential Stresses
Radial Distribution at Different Times τ

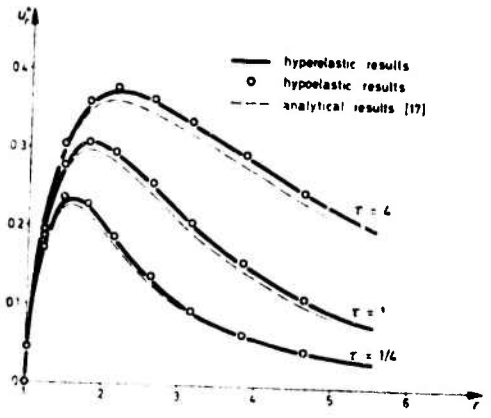


Fig 22 Radial Displacements
Radial Distribution at Different Times τ

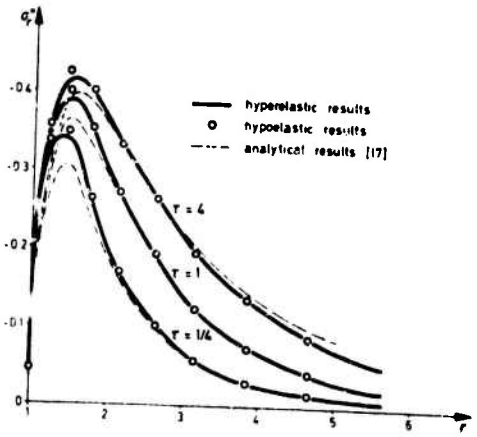


Fig 23 Radial Stresses
Radial Distribution at Different Times τ

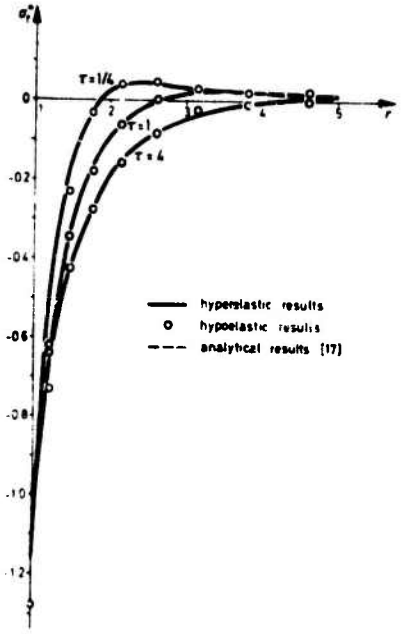
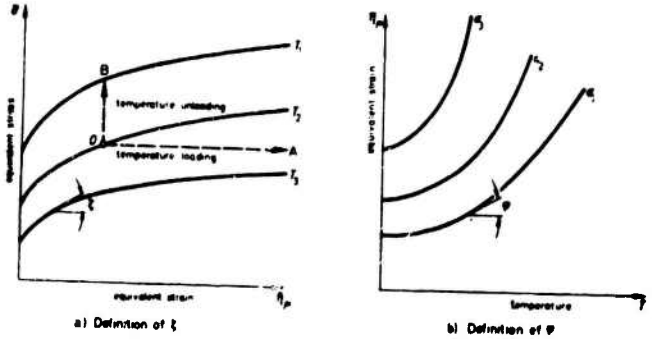
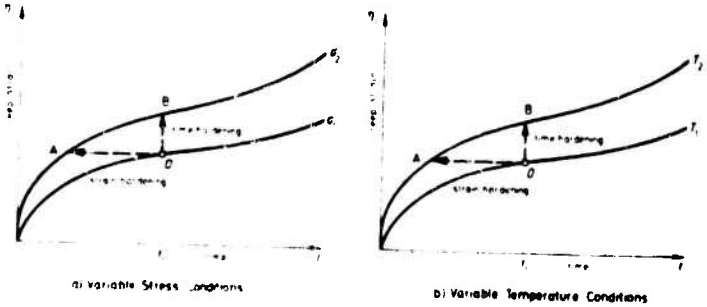


Fig 24 Tangential Stresses
Radial Distribution at Different Times τ



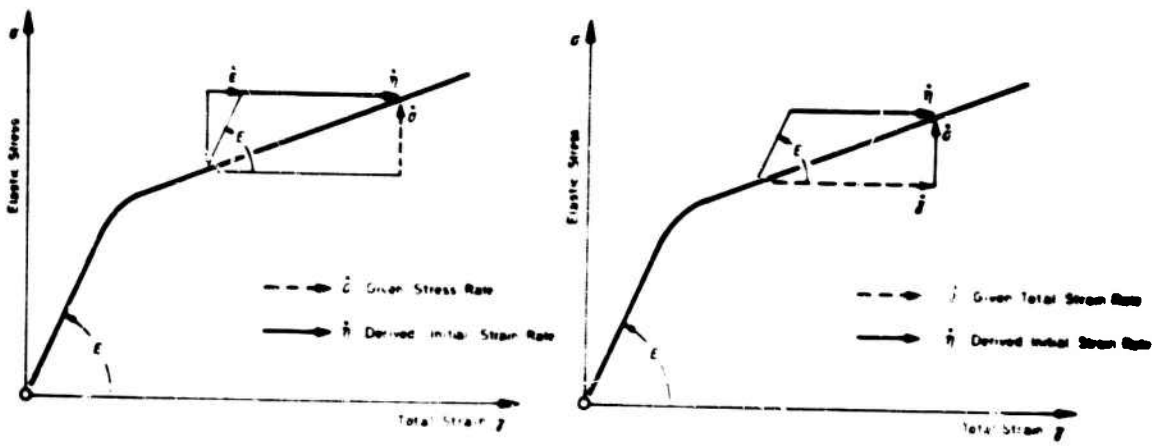
a) Definition of ξ
b) Definition of φ

Fig 25 Temperature Dependent Workhardening Parameters



a) Variable Stress Conditions
b) Variable Temperature Conditions

Fig 26 Transient Effects on Creep Rate



a) Initial Stram Approach

b) Initial Stress Approach

Fig 27 Initial Load Methods

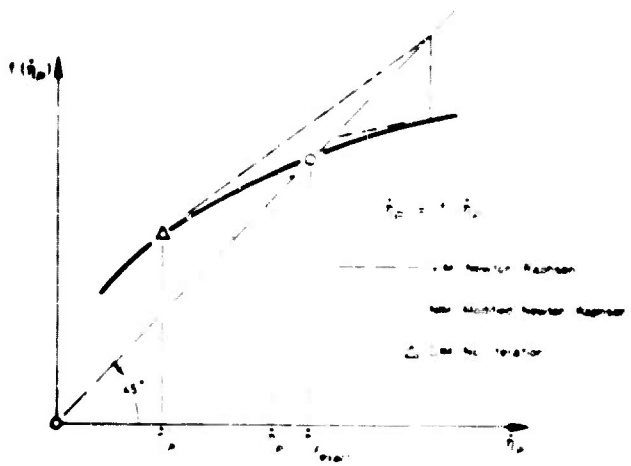


Fig 28 Incremental Initial Load Methods

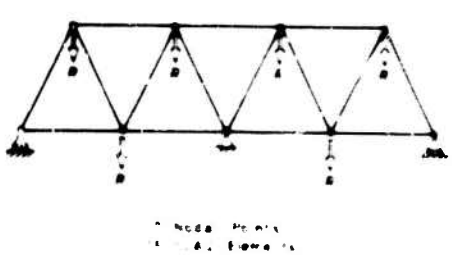


Fig 29 Indeterminate Truss

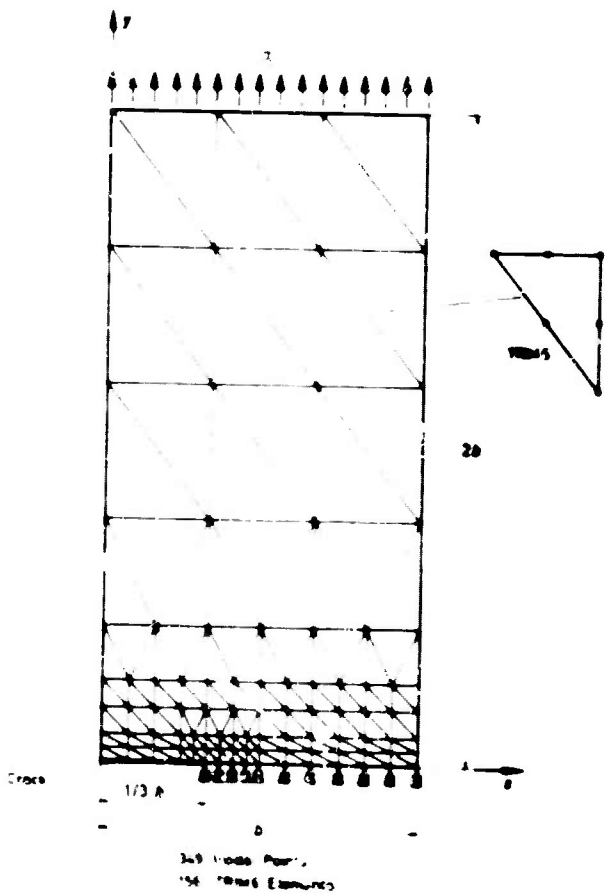


Fig 30 Rectangular Membrane with Crack

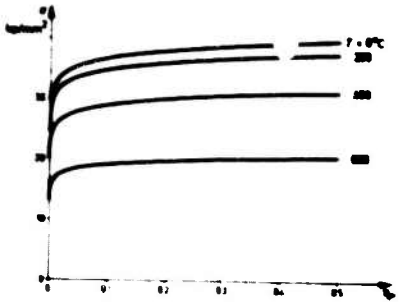


Fig 32 Uniaxial Stress - Plastic Strain Properties

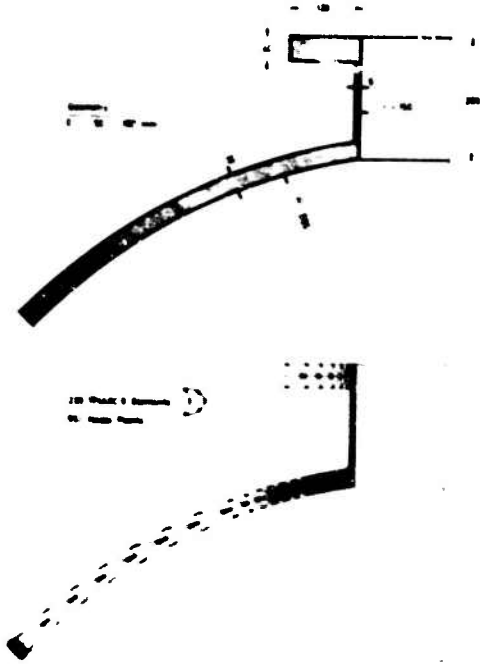
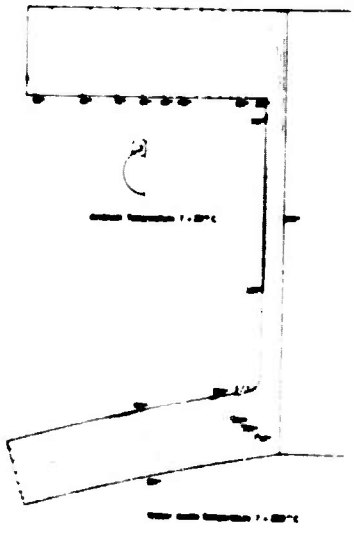
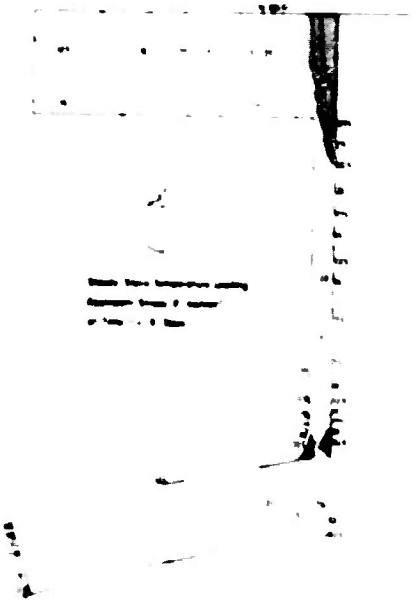


Fig 33 Spherical Shell with Nozzle
Geometry and Idealization





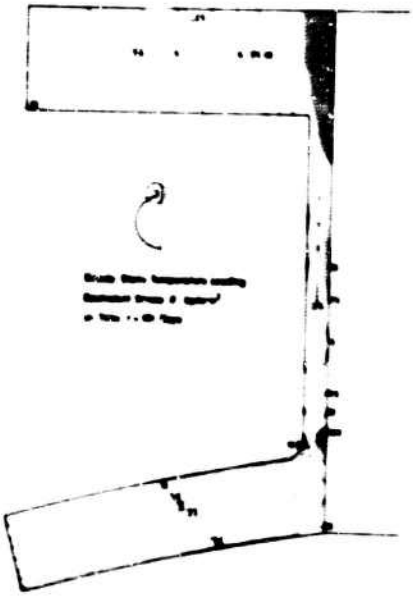


Fig 34 Spherical Shell with Nozzle

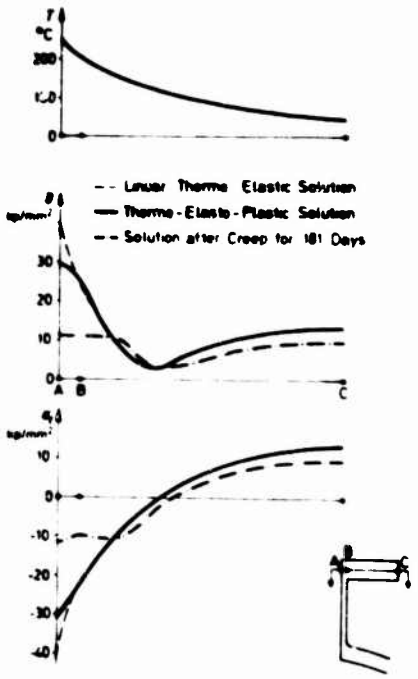


Fig 35 Spherical Shell with Nozzle
Equivalent Stresses at Section A-C

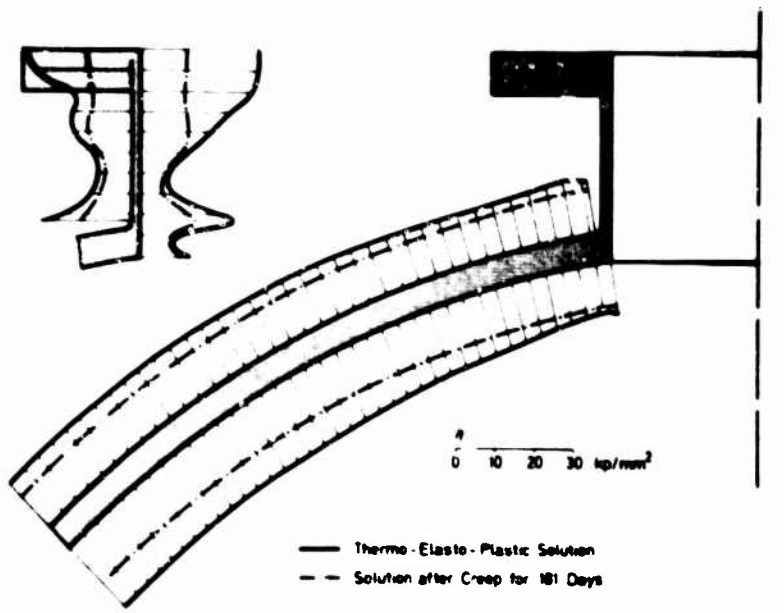


Fig 36 Spherical Shell with Nozzle
Equivalent Stress Distribution

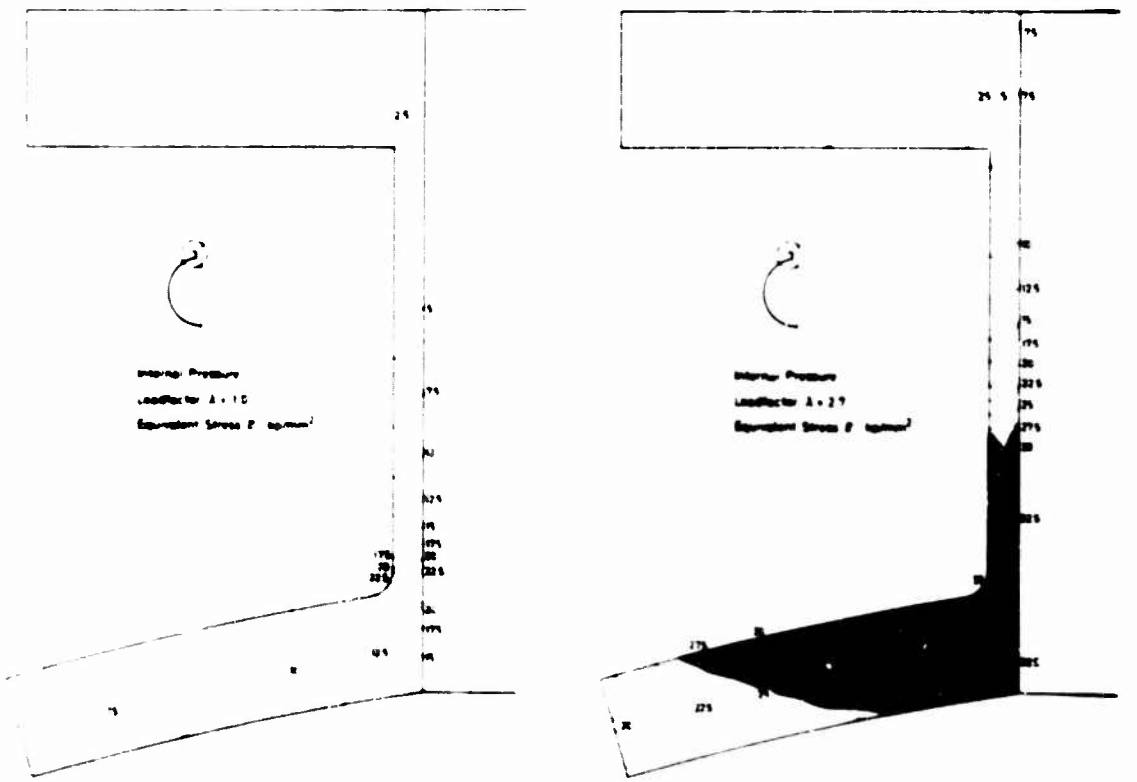


Fig 37 Spherical Shell with Nozzle
Equivalent Stress Distribution

NON-LINEAR THERMAL STRESS ANALYSIS FOR NUCLEAR POWER PLANT BY FINITE ELEMENT METHOD

WATARU MIZUMACHI

Toshiba Electric Co.,
Atomic Power Division, Tokyo, Japan

This paper presents the non-linear thermal stress analysis for the nuclear power plant by the finite element method. The heat equilibrium equation becomes the non-linear differential equation when the thermal conductivity is dependent on the temperature of the body. And the heat flow due to the radiation is in proportion to the fourth power of the absolute temperature by Stafen-Boltzmann law. The finite element method has proven very successful in analyzing these non-linear problems.

Besides, there are two types of non-linearity as far as the stress analysis is concerned. One is the material non-linearity problem and the other is the geometrical non-linearity problem. The material non-linearity problem, namely the elastic-plastic analysis, is expected in the case of designing the nuclear power plant.

The non-linear heat problems and the material non-linearity problem can be analyzed by the same nodal points and the same elements with application of the finite element method.

AFLC-WPAFB-OCT 71 550

1. INTRODUCTION

The thermal loads cause severe stress conditions in the nuclear power plant. Therefore, a precise analysis is required. In the case of the thermal stress analysis, we have to first know the temperature distribution in the structure. Then, the thermal stress can be solved.

A remarkable advantage of the finite element method lies in the fact that both analyses can be done by the same meshes. In general, the non-linear problems are not very successful from the mathematical point of view. It has become convenient to use this method to solve the non-linear problems.

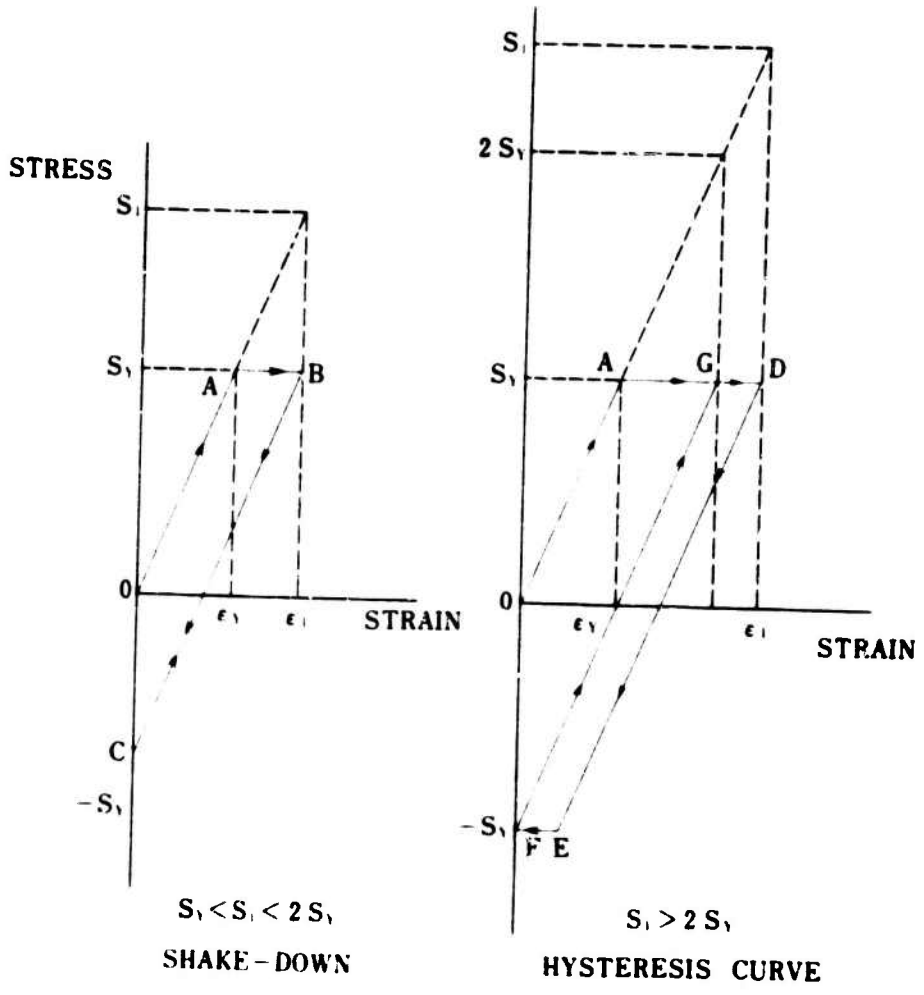
The thermal stress of the reactor pressure vessel, etc., is a self-balancing stress produced by a non-uniform distribution of temperature or by differing thermal coefficients of expansion. If the stress, neglecting stress concentrations, exceeds twice the yield strength of the material, the elastic analysis may be invalid and successive thermal cycles may produce incremental distortion. But if the stress does not exceed twice the yield point, the structure shakes down. The shakedown is the absence of a continuing cycle of plastic deformation. A structure shakes down if after a few cycles of load application, the deformation stabilizes and subsequent structural response is elastic, excluding creep effects. Therefore, not only the elastic behavior, but the plastic behavior should be analyzed.

2. FINITE ELEMENT METHOD

The finite element method was originally developed for the vibration problem of an airplane. And then it proved to be applicable to several other fields, as well as very useful in analyzing the non-linear problems.

However, the big problem is to solve the large matrix (the author calls it the Jumbo Matrix), the order of which is 10,000 to 100,000. It becomes so large because the geometrical shapes of structures in the nuclear power plant are complex and require a three-dimensional analysis. The author has used the conjugate gradient method to solve such complex systems.

Although this method has the remarkable advantages mentioned above, there exists no method without defects. The inherent disadvantage of this method lies in the fact that a lot of elements and nodal points are geometrically required to represent real structures so that reasonably accurate results are obtained. Consequently, the input data specifying the geometry and the topology are inevitably large in quantity and have to be accurate. In order to avoid the large quantiy of input data, the automatic mesh generation routine which provides for an automatic means of discretization for continuous domains is useful. Only the input data for geometrically irregular points are required. The advantages of automating the generation of input data are obvious. This facility can reduce the quantity of input data drastically, and so saves the hand-labor time and avoids the probability of human error involved in the preparation of data. Nevertheless, there might exist input errors. It is ridiculous to find the input errors after the data have been computed because the calculations to analyze the complex systems such as the nuclear power plant, take a couple of hours and need the large core memory in the computer. Therefore, any errors in geometry and topology should be previously detected by using a plotting system. This device also produces plots of the deformed structure, and stress contours.



A. S. M. E. BOILER AND PRESSURE VESSEL CODE

(SECTION – III)

"NUCLER VESSELS"

3. NON-LINEAR HEAT TRANSFER ANALYSIS

The finite element method is effecient in getting the solution for non-linear heat transfer problems because it is completely general with respect to geometry, material properties, and boundary conditions. In general, non-linear problems are idealized to linear ones in order to avoid difficult and clumsy calculations. However, non-linear problems can be solved as they are by the iteration of the same procedure by using this method.

The transient heat equilibrium equation is expressed as the first order non-linear equation.

$$C(\rho)\frac{\partial T(t)}{\partial t} + K(T)T(t) = Q(t) \qquad (3.1)$$

where

C : Heat Capacity Matrix
ρ : Density
K : Conductivity Matrix
Q : External Heat Flow
T : Temperature
t : Time

It is reasonable to assume that the flow-temperature curve consists of linear segments. Namely, the conductivity matrix is constant within each time increment. We consider equation (3.1) at time t and at time $(t + \Delta t)$. Then equation (3.1) can be written as follows:

$$C\frac{\Delta T_t - \Delta T_{t-\Delta t}}{\Delta t} + K\Delta T_t = Q_{t+\Delta t} - Q_t \qquad (3.2)$$

Then equation (3.2) becomes

$$K^* \Delta T_t = Q_t^* \qquad (3.3)$$

where

$$K^* = \frac{C}{\Delta t} + K \qquad (3.4)$$

$$Q^* = \frac{C}{\Delta t} T_{t-\Delta t} + (Q_{t+\Delta t} - Q_t) \qquad (3.5)$$

$$\Delta T_t = T_{t+\Delta t} - T_t \qquad (3.6)$$

$$\frac{\partial T_t}{\partial t} = \frac{T_t - T_{t-\Delta t}}{\Delta t} \qquad (3.7)$$

The non-linear heat equilibrium equation yields equation (3.3), which can be solved directly for the temperature increments at the end of the time increment.

In order to determine the conductivity matrix, the temperature distribution within each element is assumed to be linear to the coordinates of the nodal points. Dr. E.L. Wilson (Univ. of California, Berkeley) gives this method.

For radiation problems, the net flow transferred from one surface to another is given by Stafen-Boltzmann law.

$$Q = \sigma(T_1^4 - T_2^4) \qquad (3.8)$$

where

$\sigma = 4.88 \times 10^{-8}$ kcal/m^2 h °K^4
(Stafen-Boltzmann Coefficient)

T : Absolute Temperature

Assuming that the difference between T_1 and T_2 is small, the pseudo-conductivity matrix is given. Then the temperature distribution due to the radiation can be solved by the finite element method.

4. STRESS ANALYSIS

The element stiffness matrix of the finite element method can be written as follows by applying the principle of virtual work and the Castigliano theory.

$$[K]_e = [A^{-1}]^T \left(\iiint [B]^T [D] [B] \, dV \right) [A^{-1}] \qquad (4.1)$$

where

$[K]_e$: Element Stiffness Matrix

$[A]$: Matrix of Displacement Functions (Coordinate Matrix)

$[B]$: Strain Coefficient Matrix

$[D]$: Stress Strain Matrix (Material Matrix)

The stiffness matrix of the structure is given by assemblying the element stiffness matrices.

$$[K] = \sum_{e=1}^{N} [K]_e \qquad (4.2)$$

As far as the thermal stress is concerned, the thermal load is given as follows:

$$\{F\} = -\iiint [B]^T [D] \{\epsilon_0\} \, dV \qquad (4.3)$$

where

$$\{\epsilon_0\} = \begin{Bmatrix} \epsilon_x \\ \epsilon_y \\ \epsilon_z \\ \gamma_{xy} \\ \gamma_{yz} \\ \gamma_{zx} \end{Bmatrix} = \begin{Bmatrix} \alpha_x T \\ \alpha_y T \\ \alpha_z T \\ 0 \\ 0 \\ 0 \end{Bmatrix} \qquad (4.4)$$

$\{F\}$: Thermal Load

$\{\epsilon_0\}$: Thermal Strain

α : Coefficient of Thermal Expansion

At first displacements of each nodal point due to the thermal load can be solved.

$$\{u\} = [K]^{-1} \{F\} \qquad (4.5)$$

The stiffness matrix $[K]$ of equation (4.5) has the order of 10,000 to 100,000 to analyze structures with geometrically complicated shapes. This order is too large to apply normal methods, for example, the Gauss method. Therefore, the special method to solve equation (4.5) is expected for the finite element method.

The conjugate gradient method is successful in solving the large sparse matrix. It is inherently required to store the addresses of non-zero elements in the computer core. On the other hand, the addresses of

non-zero elements should be stored when the finite element method is applied, because this method requires knowing each number of nodal points. This fact shows that the conjugate gradient method is especially nice for solving the finite element method. Besides, there is one more big advantage. The order of the stiffness matrix becomes so large because the element stiffness matrices are assembled by equation (4.2). In the case of using the conjugate gradient method, the values and addresses of non-zero elements are needed. In other words, the element matrices need not be assembled.

The following is the procedure by which the stress analysis can be done directly by the element matrices without having the assembled stiffness matrix of a structure. Assume the initial vector to solve equation (4.5).

$$\{u\}_1 = \{\text{Random Number}\} \qquad (4.6)$$

Or unless specified, it is assumed to be equal to the thermal load vector defined by equation (4.3).

$$\{u\}_1 = \{F\} \qquad (4.7)$$

Iteration

$$\{u\}_{i+1} = \{u\}_i + \alpha_i \{S\}_i \qquad (4.8)$$

where

i : Iteration Number

The remainder $\{G\}$ is

$$\{G\}_{i+1} = \{\{F\} - [K]\{u\}_i\} \qquad (4.9)$$

The direction of movement $\{S\}$ is

$$\{S\}_{i+1} = \{G\}_{i+1} + \beta_i \{S\}_i \qquad (4.10)$$

where

$$\beta_i = \frac{\{G\}_{i+1}^T \{G\}_{i+1}}{\{G\}_i^T \{G\}_i} \qquad (4.11)$$

The step length α along the direction is

$$\alpha_i = \frac{\{S\}_i^T \{G\}_i}{\{S\}_i^T [K] \{S\}_i} \qquad (4.12)$$

The convergence condition is

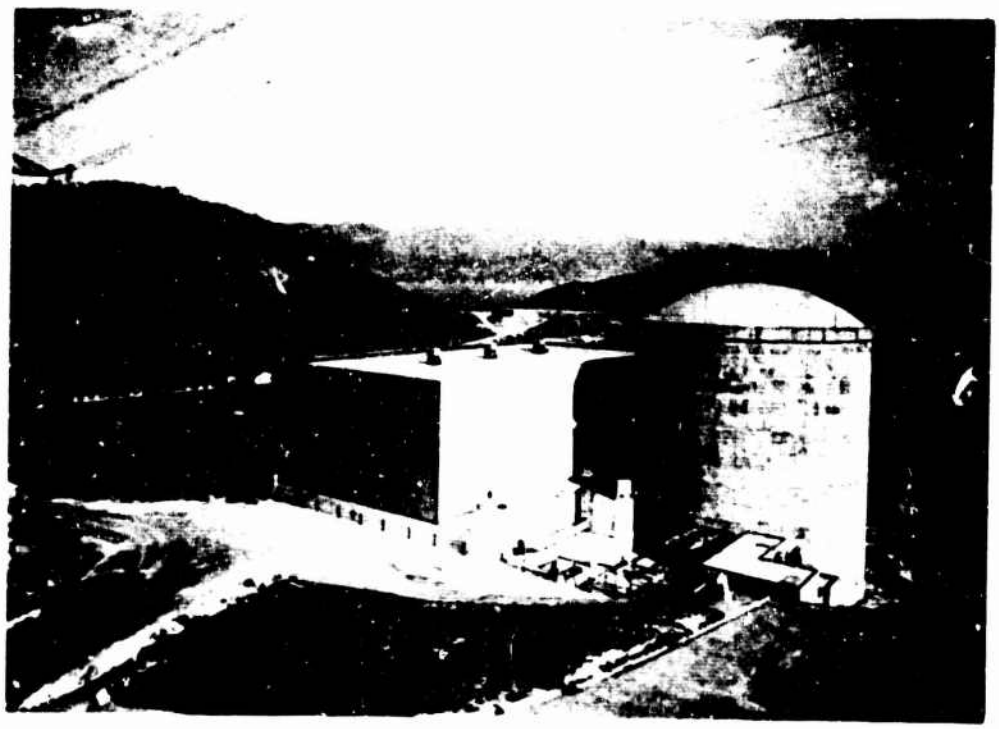
$$\frac{\sqrt{\sum_{j=1}^{n} (\{u_j\}_{i+1}^2 - \{u_j\}_i^2)}}{\sqrt{\sum_{j=1}^{n} \{u_j\}_i^2}} < \epsilon \qquad (4.13)$$

5. PLOTTER OUTPUTS OF THERMAL ANALYSIS

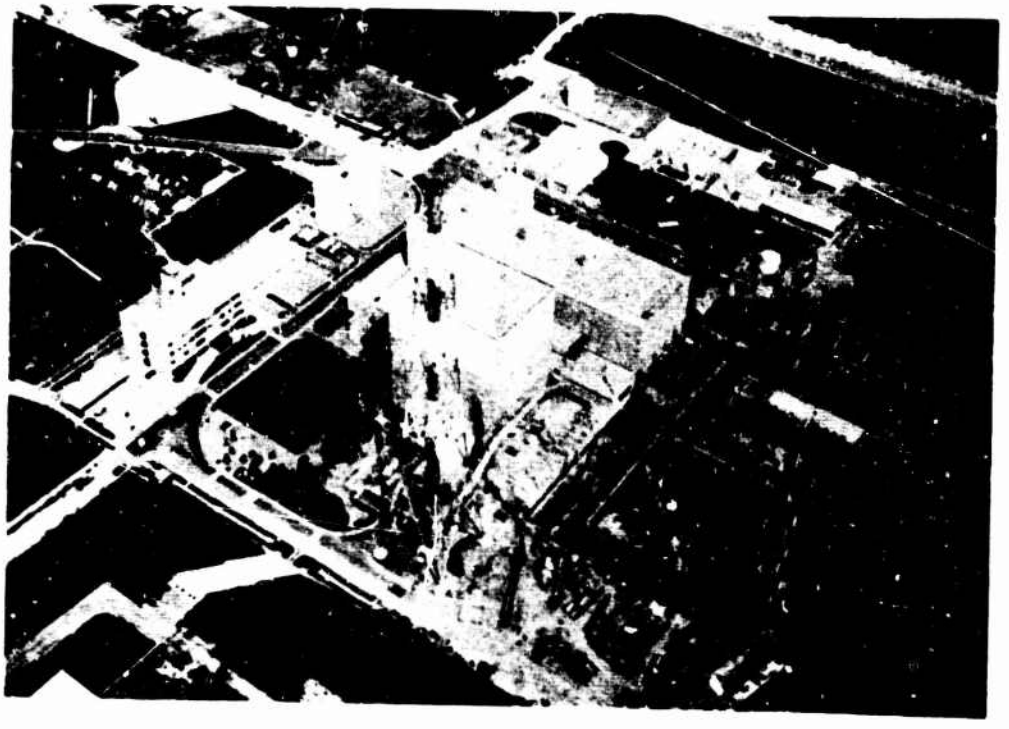
(1) Atomic Reactor Building

① TSURGA--1

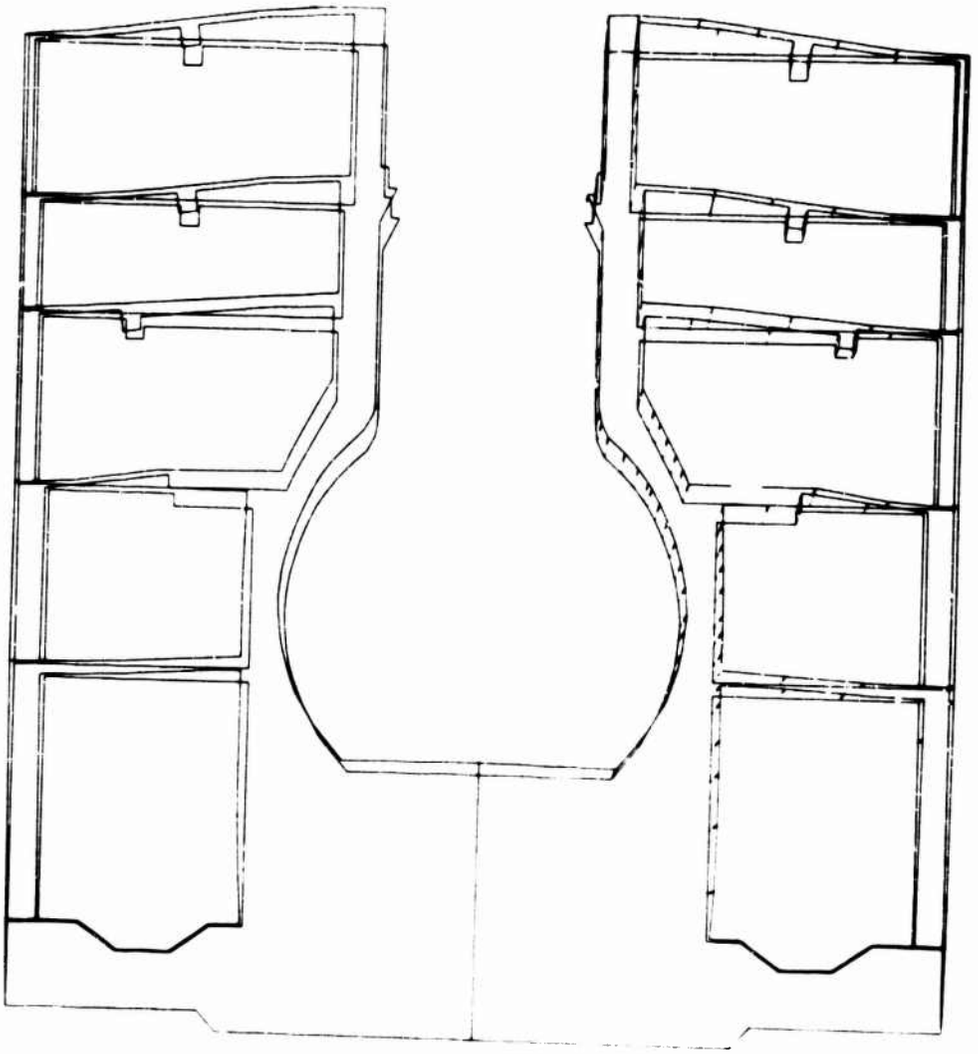
This photograph is the nuclear power station at Tsurga, site of the Japan Atomic Power Co., which provided electricity for Expo '70 held near Osaka. The type is the Boiling Water Reactor (BWR).



② TEPCO–I, II, and III (FUKUSHIMA)



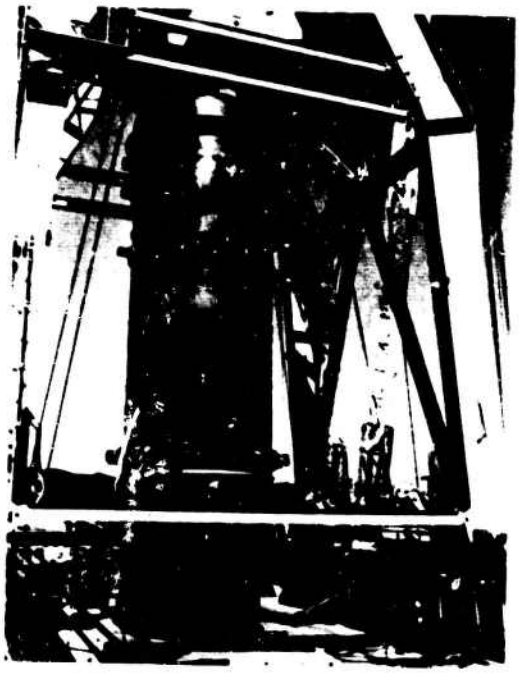
③ Thermal Stress Analysis



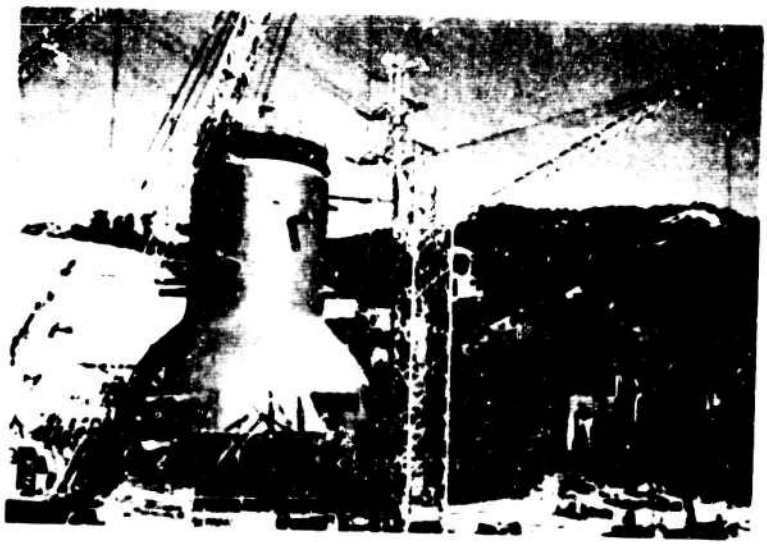
THERMAL STRESS ANALYSIS
ATOMIC REACTOR BUILDING

(2) Reactor Pressure Vessel, Primary Containment Vessel and Pedestal

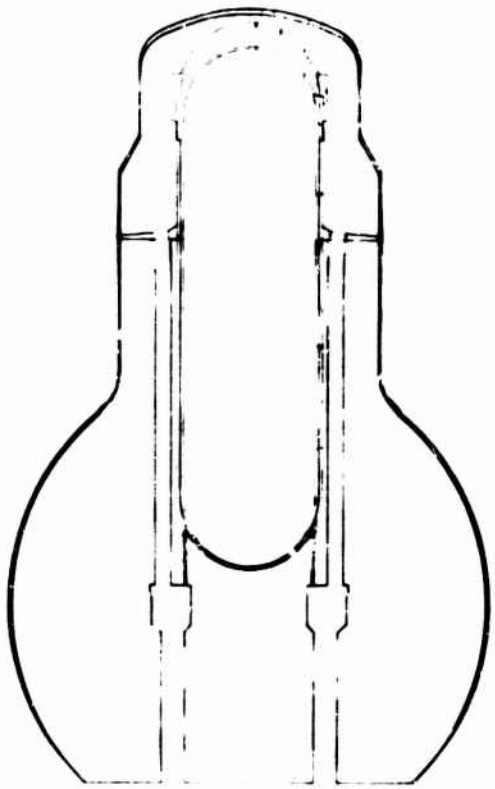
① Reactor Pressure Vessel (TEPCO–1)



② Primary Containment Vessel (TSURGA–1)



③ Thermal Stress Analysis

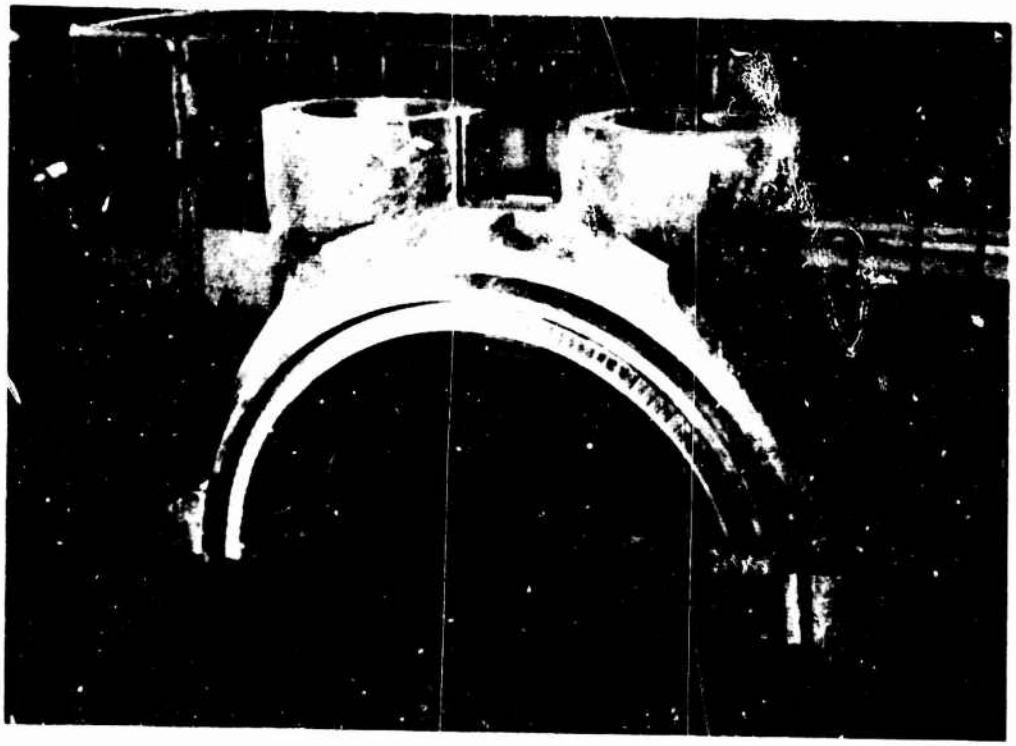


STRESS ANALYSIS
THERMAL PRESSURE GRAVITY
RPV PCV SHIELD WALL

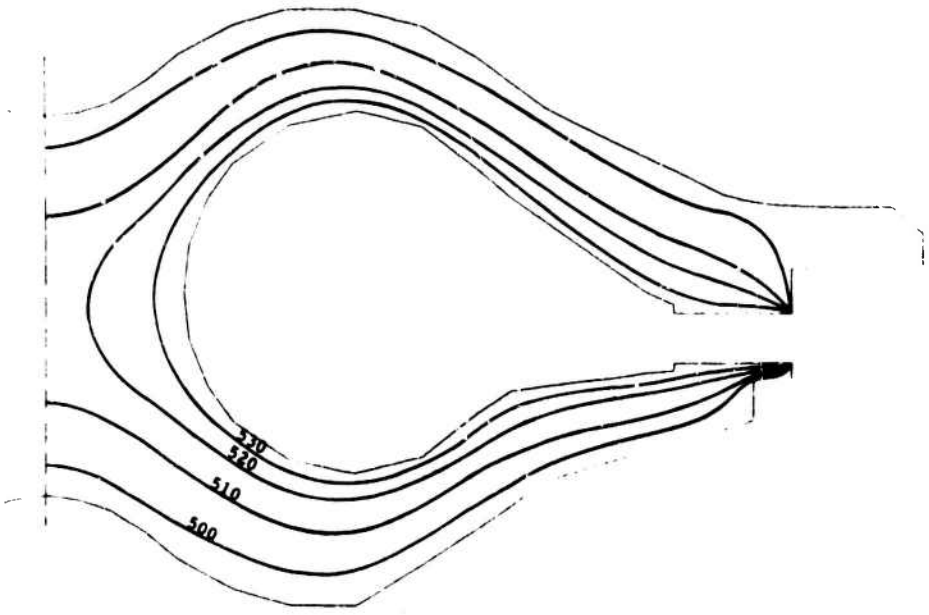
(3) Plastic Analysis for Turbine Nozzle Box

① Turbine Nozzle Box

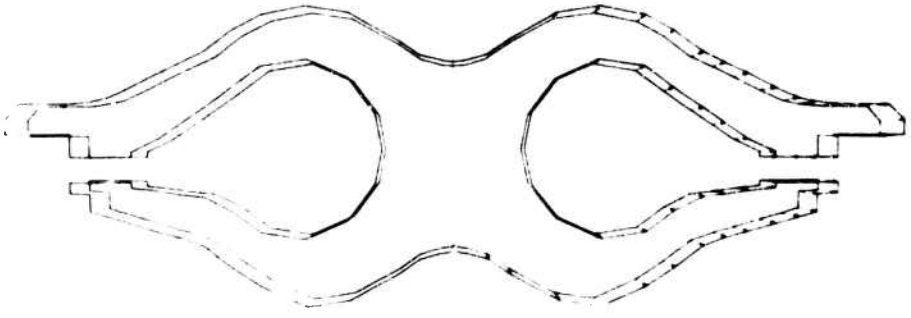
The stress of the turbine nozzle box due to the thermal load and pressure is very severe. The stress should not exceed the yield stress. However, it is important to know the plastic behavior. Following are the results of the material non-linear analysis and those of the temperature distribution.



② Temperature Distribution

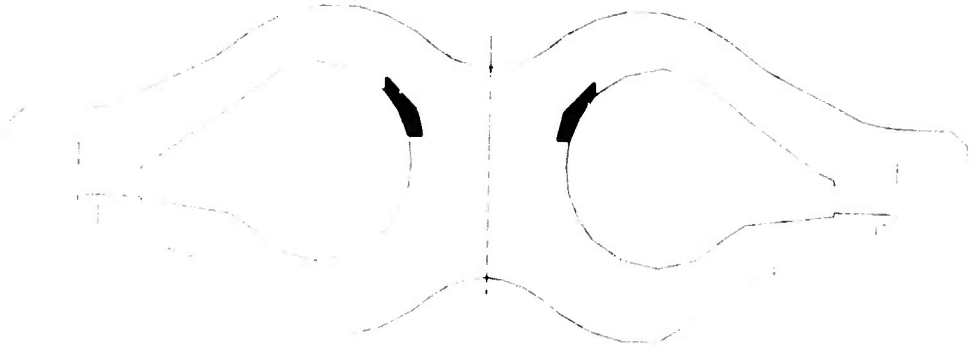


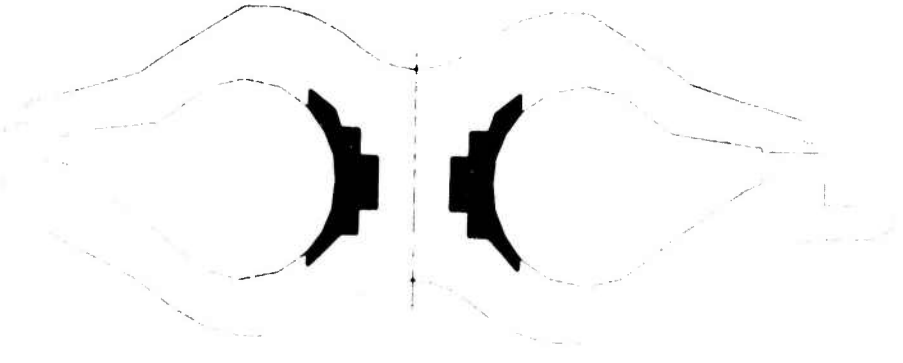
③ Elastic Thermal Stress Analysis

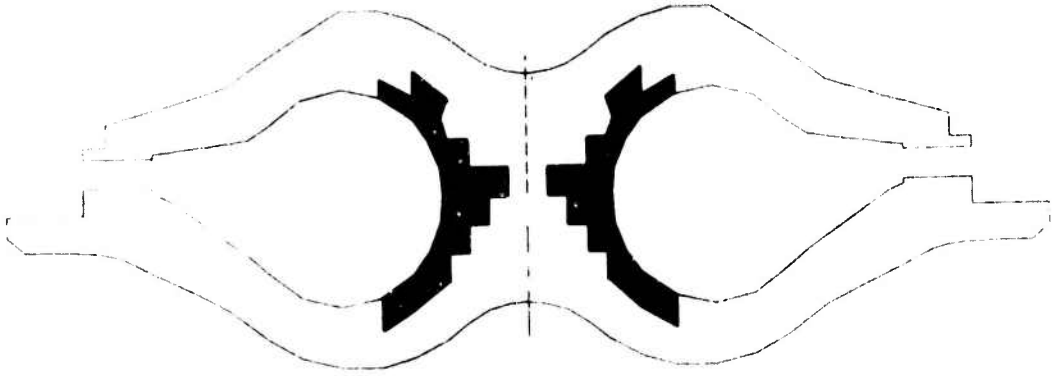


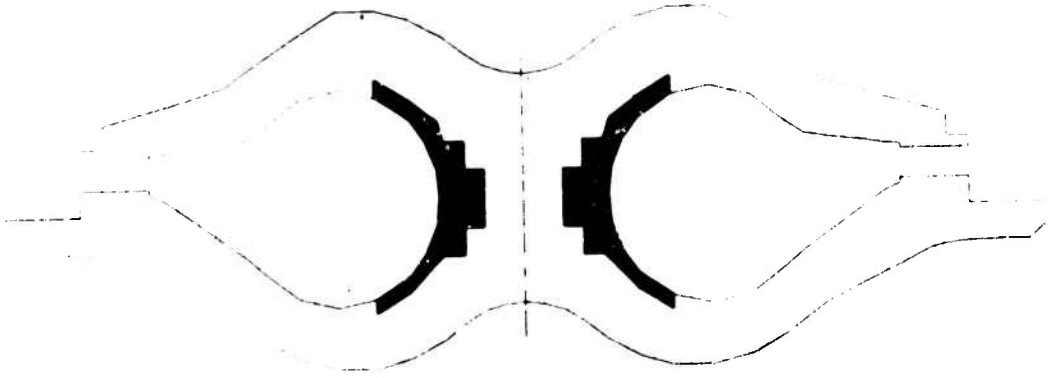
STRESS ANALYSIS
(PRESSURE;THERMAL)
TURBINE NOZZLE BOX

④ Plastic Thermal Stress Analysis
The black area exceeds the yield point.









6. CONCLUSION

Some kinds of stress analyses are expected when we design a nuclear power plant. These are the thermal stress, the stress due to pressure, and the stress due to the external force including its own weight, earthquake force and so on. Normally, the value of the thermal stress is the highest one among these kinds. Therefore, a precise thermal stress analysis should be expected.

The remarkable advantage of the finite element method lies in the fact that this method has complete generality with respect to geometry, material properties, and boundary conditions. Besides, the non-linear heat problem and the material non-linear stress analysis can be solved by the same meshes.

However, this method has two disadvantages. One is that the order of the matrix is very large. Another is that the input data specifying the geometry are large in quantity. The author uses the conjugate gradient method to treat the large matrix and develops the automatic mesh generation routine in order to avoid the large quantity of the input data.

Finally, it should be emphasized that it is not better to divide the model into more elements, even though the large matrix can be treated easily. We should select the point where the error of the geometrical shape and that of the numerical calculation balance. And also, we should select the most reasonable displacement function for each problem.

7. REFERENCES

(1) Wilson, E.L., "Non-linear Heat Transfer Analysis of Axisymmetric Solids" *Report of Structural Engineering Laboratory*, Univ. of California, Berkeley (1971)

(2) Zienkiewicz, O.C., and Cheung, Y.K., "The Finite Element Method in Structural and Continuum Mechanics" *McGraw Hill* (1967)

(3) "Japan-U.S. Seminar on Matrix Methods of Structural Analysis and Design" *Society of Steel Construction of Japan* (1969)

- ① Clough, R.W., "Analysis of Structural Vibrations and Dynamic Response"
- ② Marcal, P.V., "Finite Element Analysis with Material Non-Linearities"
- ③ Melosh, R.J., "Manipulation Errors in Finite Element Analyses"
- ④ Kawai, T., "Finite Element Analysis of Shell Structures"
- ⑤ Yamada, Y., "Recent Japanese Developments in Displacement Method for Elastic-Plastic Problems"

(4) Miyamoto, H., and Miyoshi, T., "Elastic and Plastic Analysis by Finite Element Method" *Society of Steel Construction of Japan* (1968)

(5) Akyuz, F.C., "Natural Co-ordinate Systems. An Automatic Input Data Generation Scheme for a Finite Element Method" *Nuclear Engineering and Design* (1970)

(6) Mizumachi, W., "Application of a Finite Element Method to the Atomic Power Plant" *Report of Research Institute for Mathematical Sciences*, Univ. of Kyoto (1971)

(7) Mizumachi, W., "How to Solve the Large Matrix" *Journal of the Computer Sciences & Information Processing*, Japan Computer Society (1971)

(8) Mizumachi, W., "Analyses for Atomic Power Plant by Finite Element Method" *Society of Steel Construction of Japan* (1971)

(9) Mizumachi, W., "Application of Finite Element Method to the Three-Dimensional Stress and Vibration Analysis of BWR Primary Plant Systems" *Structural Mechanics in Reactor Technology*, First International Conference, Bundesanstalt für Materialprüfung (BAM), Berlin (1971)

SESSION 6. COMPUTER GRAPHICS

Session Chairman

W. J. Batdorf

Lockheed Georgia Co.
Marietta, Georgia

DISPLAYS OF KINEMATIC AND ELASTIC SYSTEMS

by

Henry N. Christiansen*
Brigham Young University, Provo, Utah and
University of Utah, Salt Lake City, Utah

This paper reports on a continuing exploration of computer graphics techniques as they apply to the analysis of kinematic and elastic systems. The report deals with two aspects of the work; the creation of a simulator for kinematic and elastic systems, and the utilization of computer graphics software and hardware to produce continuous tone images. The simulator is a computer program which utilizes a linear finite element stiffness model to achieve exact solutions for large displacement kinematic problems of foldable plates and trusses, as well as approximate solutions for large displacement, small strain, elasticity problems. Various techniques associated with the production of continuous tone images have been utilized to aid in the display of stress, strain, and/or displacement results producible with the simulator and other finite element programs.

*Professor of Civil Engineering Science, BYU, and Research Associate Computer Science, Univ. of Utah.

SECTION 1

INTRODUCTION

The original purpose of this study was to develop a computer program which would simulate the kinematic and elastic behavior of foldable plate and truss structures. The more recent purpose has been to explore possibilities for displays of the output of the simulator and other finite element programs.

The requirement for the simulator was generated by Resch[1] who created what he terms "kinematic folded plate systems." The common property of these systems is that, by allowing only folding along the edges of a continuous line pattern, a flat sheet may be transformed into a variety of three dimensional shapes. Figure 1 illustrates a paper model of one system which is a repetition of only two non-identical plate elements. Figure 2 shows another pattern which requires four non-identical elements. These systems give promise of the mass production of non-identical structures of unusual beauty and structural efficiency.

Resch's initial investigation involved the manual folding of sheets of paper upon which the pattern had been scribed. The model could then be moved by hand to produce a variety of shell forms. When a desirable shape was obtained, the system could be stabilized by providing sufficient supports or the addition of truss and/or plate elements. These constraining systems change the character of the motion of the model from kinematics to elastic deformation.

This approach, while yielding pleasing visual effects, had several drawbacks. First, the process required repetition (the model shown in Figure 1 took 30 hours to fold) to investigate any parameter change in the "wrinkle" definition. Second the dimensions of elements added to stabilize the shell are unique to the shape chosen. Thus, each element had to be individually measured and created to fit the kinematic pattern. Inaccuracy in this process prevented duplication of a particular configuration. Finally, while the investigators were impressed with the apparent stiffness of the stabilized shells, the model was not useful in an analysis of stress and strain.

The functions of kinematic and elastic analysis were accomplished as a result of a finite element stiffness approach to the linear elasticity problem. Developments which allow this approach include the modeling of the folds as elastic hinges and the generation of applied force systems which tend to restore the original dimensions of the elements.

Upon completion of the simulator, effort was directed toward the utilization of the ability to produce continuous tone pictures to display the stress, strain and displacement results of the finite element solutions producible with the simulator and other programs. Three methods of data presentation were devised and developed. In the most primitive mode, the object is displayed under load and variation of the light intensity of the surface of the object is utilized to indicate the level of stress or strain.

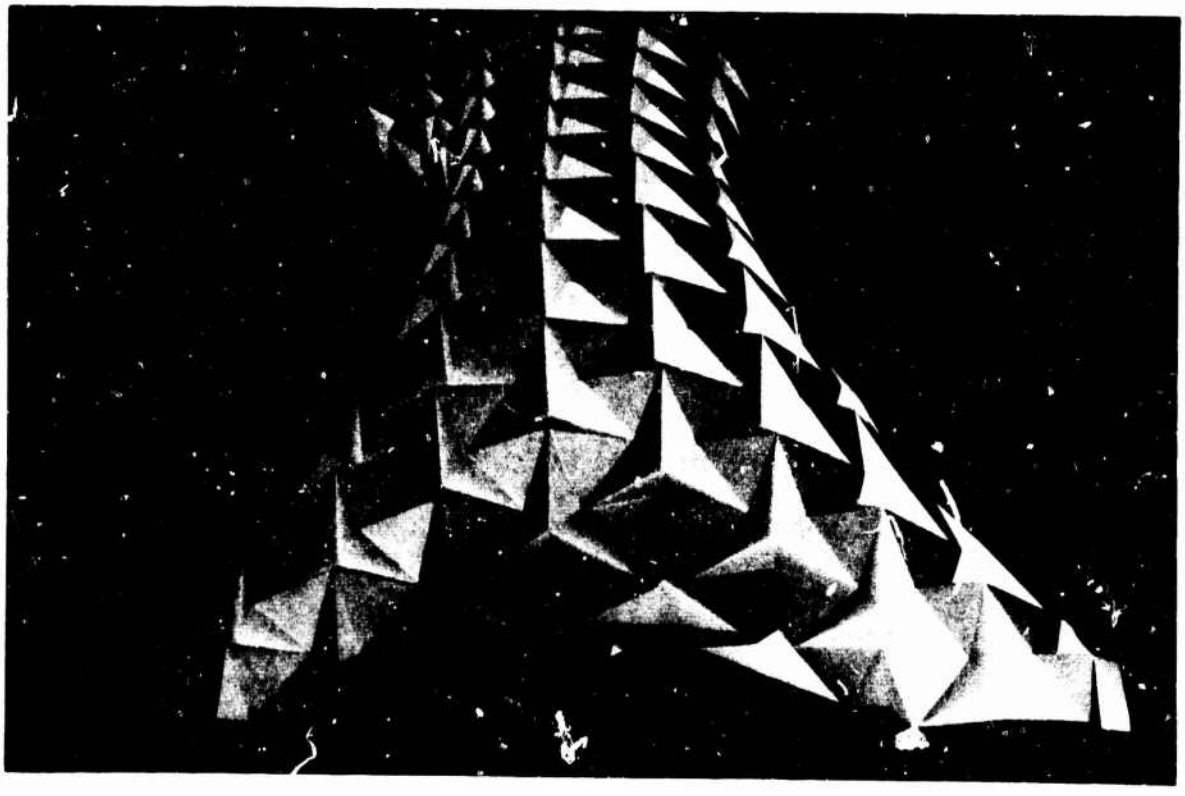


Figure 1. Folded Plate System with Two Non-Identical Elements

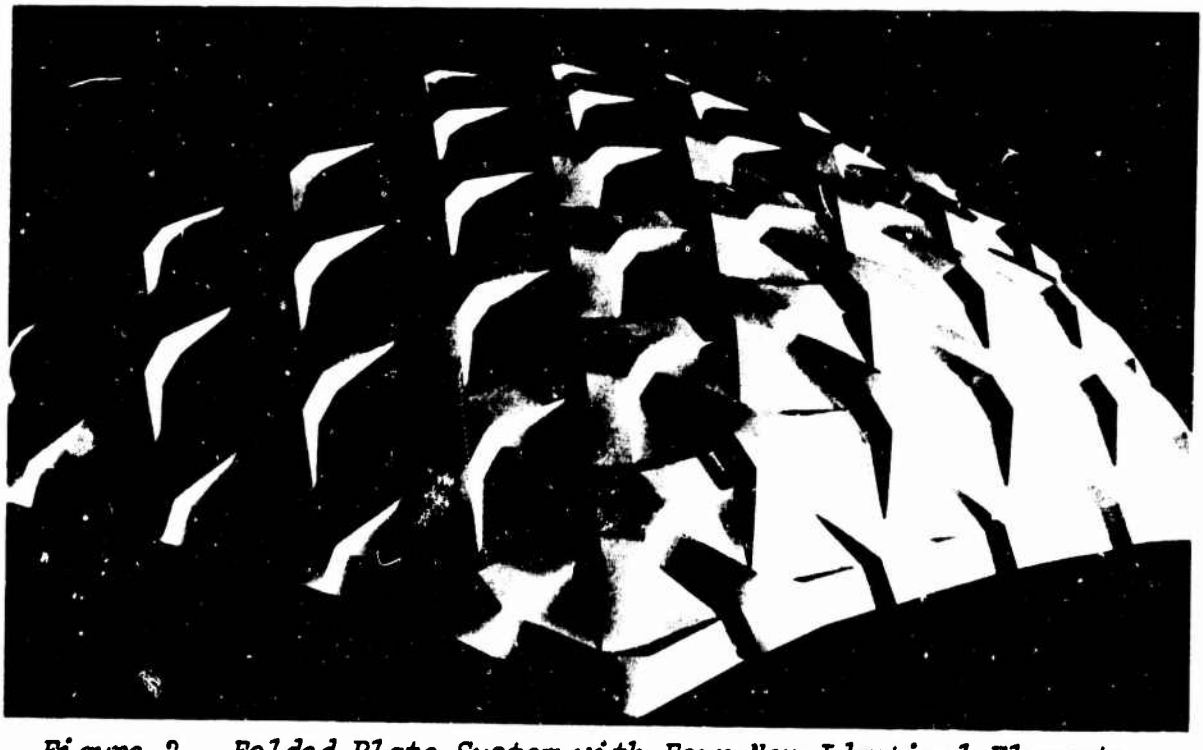


Figure 2. Folded Plate System with Four Non-Identical Elements

A variation on this scheme results from making the light intensity proportional to the square of a harmonic function of the selected stress or strain component. In this way fringe effects develop which have a character much like that produced using photo-elastic techniques. In another scheme the object is warped such that the out of plane coordinate is proportional to the stress or strain function. The resulting three dimensional shape is displayed in a conventional manner. The results of these efforts are discussed and displayed in following sections.

SECTION II

STRUCTUPAL ANALYSIS OF KINEMATIC FOLDED PLATE SYSTEMS

Two major functions are performed in the structural analysis of folded plate structures. The first function is the large displacement kinematic analysis associated with positioning the structure in the desired shape. The second function is the small displacement elastic analysis of the properly positioned and stabilized structure. While both functions are accomplished as a result of a finite element stiffness method approach to the small displacement elasticity problem, the large displacement kinematics analysis requires a repeated application of the linear theory.

Kinematic analysis is achieved by:

1) *modeling the plates as constant strain, plane stress elements and the folds as flexible springs,*

2) *performing an elastic analysis,*

3) *adding the computed displacements to the nodal coordinates,*

4) *computing nodal force systems which tend to restore the plate (or truss) element dimensions, and*

5) *repeating steps 1 thru 4 until convergence is achieved.*

When the fold element is not restored (i.e. no attempt is made to reestablish the original angles between the elements) convergence is either to a kinematically possible configuration or to the large displacement, small strain, elasticity solution. The elasticity solution results when the prescribed displacements are not possible through rigid body motions of the individual structural elements. In either case, the fold elements will be unstressed during the final iteration. If the fold elements are restored, convergence is to an elasticity solution.

After each iteration, the displacement constraints may be modified. Thus, the solution achieved is a function of the order in which the constraints are applied and removed. When the desired kinematic arrangement is achieved, the structure is stabilized with additional elements and/or displacement constraints. Additional iterations may then be performed, for specified loading systems, to obtain elastic displacements and stresses.

Much of this procedure is inherent in any finite element analysis and has been documented many times. To prevent an additional redundancy only those parts of the procedure which are thought to be at least relatively unknown are described in this paper. Two developments are judged to be in this category. The first is the definition of a stiffness matrix for a fold element and the second is the procedure by which corrective forces are generated.

STIFFNESS MATRIX FOR A FOLD ELEMENT

In this formulation, the fold between two triangular elements (see Figure 3) is assumed to have a stiffness (i.e. resistance to folding) which increases linearly with the fold length.

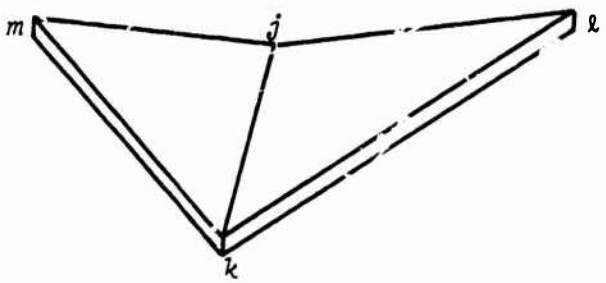


Figure 3 Fold Element jklm

The strain energy associated with this stiffness is equal to the work done by a system of force couples M in decreasing the dihedral angle between the two plate elements by the small angle θ and is thus written as

$$U = \frac{1}{2} M \theta \qquad (1)$$

The assumed stiffness suggests a generalized force-displacement relationship of the form

$$M = KL\theta \qquad (2)$$

where K and L are the stiffness per unit length and length of the fold. With this substitution, the strain energy is

$$U = \frac{1}{2} KL \begin{Bmatrix} \theta_\ell \\ \theta_m \end{Bmatrix}^T \begin{bmatrix} I \end{bmatrix} \begin{Bmatrix} \theta_\ell \\ \theta_m \end{Bmatrix} \qquad (3)$$

where θ_ℓ and θ_m are the rotations about the axis jk of elements jkℓ and jkm, and

$$\left[I \right] = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{4}$$

The next step is to express the rotations in terms of the nodal displacement components. Starting with element jkℓ and introducing a local coordinate system (X_i') as shown in Figure 4,

$$\left\{ \theta_\ell \right\} = \left[\bar{X}_\ell \right] \left\{ U_\ell' \right\} \tag{5}$$

where

$$\left\{ U_\ell' \right\} = \begin{Bmatrix} U_{3j}' \\ U_{3k}' \\ U_{3\ell}' \end{Bmatrix}, \quad \left[\bar{X}_\ell \right] = \frac{1}{LX_{2\ell}'} \left[(X_{1\ell}' - L) \quad (X_{1\ell}') \quad (L) \right],$$

and U_{3i}' is the displacement in the X_3' direction at node i.

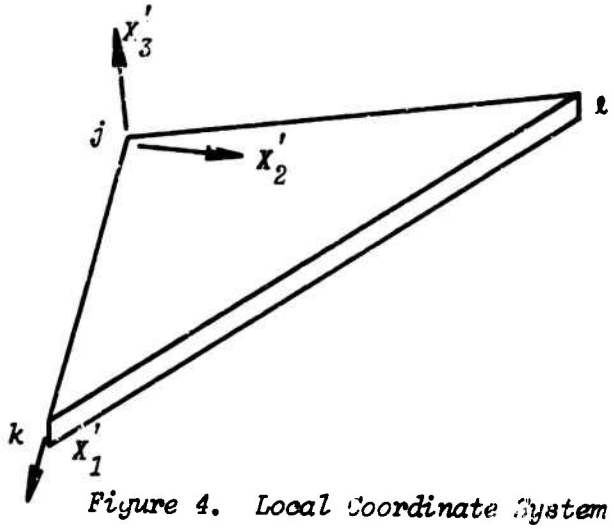


Figure 4. Local Coordinate System X_i'

In like manner (see Figure 5) for element jkm

$$\left\{ \theta_m \right\} = \left[\bar{X}_m \right] \left\{ U_m'' \right\}, \tag{6}$$

where

$$\left\{U''_m\right\} = \begin{Bmatrix} U''_{3j} \\ U''_{3k} \\ U''_{3m} \end{Bmatrix}, \quad \left[\bar{X}_m\right] = \frac{1}{LX''_{2m}} \left[(X''_{1m} - L) \; (-X''_{1m}) \; (L)\right]$$

and U''_{3i} *is the displacement in the* X''_3 *direction at node* i.

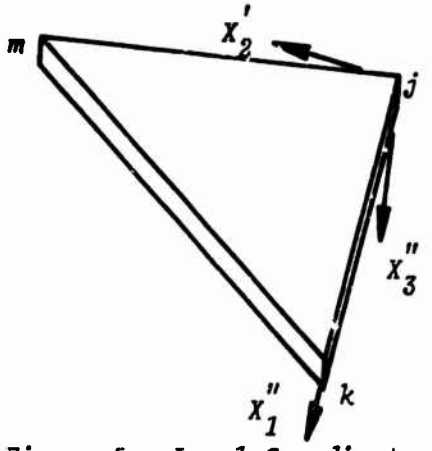


Figure 5. Local Coordinate System X''_i

Combining the rotation-node displacement relationships

$$\begin{Bmatrix} \theta_\ell \\ \theta_m \end{Bmatrix} = \begin{bmatrix} \bar{X}_\ell & \\ & \bar{X}_m \end{bmatrix} \begin{Bmatrix} U'_\ell \\ U''_m \end{Bmatrix} \tag{7}$$

Transformation to the global coordinate system displacement components is facilitated by

$$\begin{Bmatrix} U'_\ell \\ U''_m \end{Bmatrix} = \left[D\right] \left\{U\right\} \tag{8}$$

where

$$\left\{U\right\}^T = \left[U_{1j} \; U_{2j} \; U_{3j} \; U_{1k} \; U_{2k} \; U_{3k} \; U_{1\ell} \; U_{2\ell} \; U_{3\ell} \; U_{1m} \; U_{2m} \; U_{3m}\right].$$

$$[D]=\begin{bmatrix} d'_{31} & d'_{32} & d'_{33} & & & & & & \\ d''_{31} & d''_{32} & d''_{33} & & & & & & \\ & & & d'_{31} & d'_{32} & d'_{33} & & & \\ & & & d''_{31} & d''_{32} & d''_{33} & & & \\ & & & & & & d'_{31} & d'_{32} & d'_{33} \\ & & & & & & d''_{31} & d''_{32} & d''_{33} \end{bmatrix}$$

U_{in} is the displacement in the X_i (global coordinate) direction at node n, and d'_{3i} and d''_{3i} are cosines of the angles between the X_i axis and the X'_3 and X''_3 axes respectively.

With this substitution, the rotation-node displacement relationship is

$$\begin{Bmatrix} \theta_\ell \\ \theta_m \end{Bmatrix} = \begin{bmatrix} \bar{X}_\ell & \\ & \bar{X}_m \end{bmatrix} [D] \{U\}, \tag{9}$$

and the strain energy expression becomes

$$U = \frac{KL}{2} \{U\}^T [D]^T \begin{bmatrix} \bar{X}_\ell & \\ & \bar{X}_m \end{bmatrix} [I] \begin{bmatrix} \bar{X}_\ell & \\ & \bar{X}_m \end{bmatrix} [D] \{U\} \tag{10}$$

The stiffness matrix, by application of the principle of virtual displacements, is

$$[K] = KL [D]^T \begin{bmatrix} \bar{X}_\ell & \\ & \bar{X}_m \end{bmatrix}^T [I] \begin{bmatrix} \bar{X}_\ell & \\ & \bar{X}_m \end{bmatrix} [D] \tag{11}$$

CORRECTION FORCES SYSTEMS

The plane stress triangular element requires correction forces which will restore the original length of each side of the triangle. Since the columns of the element stiffness matrix are the sets of force systems required to produce individual unit nodal displacements components, by

expressing the change in length of each side of the triangle in terms of an equivalent set of nodal displacement components, a set of corrective forces may be obtained. This corrective nodal force component vector is given by the negative of the product of the element stiffness matrix and the equivalent set of nodal displacement components.

To achieve the equivalent set of nodal displacement components, displacement modes which each represent the elongation of a single side of the triangle are introduced as illustrated in Figure 6 a, b, and c.

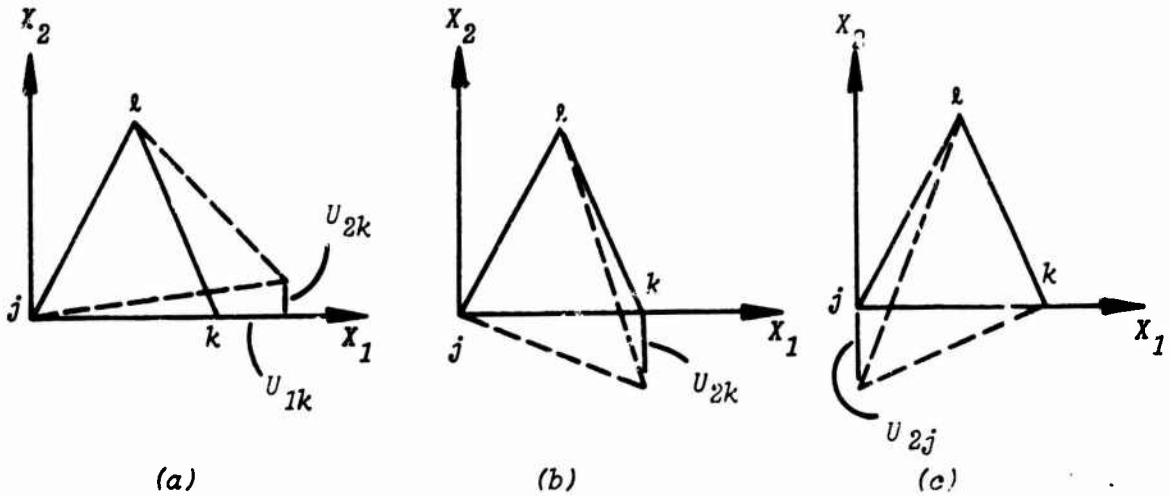


Figure 6. Side Elongation Modes for Triangular Elements

In the coordinate system shown, Figure 6a depicts a mode for which sides $j\ell$ and $k\ell$ have been preserved in length and side jk has been stretched. Figure 6b illustrates a set of nodal displacement components for which sides jk and $j\ell$ are preserved in length and side $k\ell$ is stretched. Figure 7c shows sides jk and $k\ell$ unstretched and side $j\ell$ stretched. Letting δ_{jk}, $\delta_{j\ell}$, and $\delta_{k\ell}$ represent the change in length of sides jk, $j\ell$, and $k\ell$ respectively as calculated from the current and original nodal coordinates, the relationship between the nodal displacement components and the side elongations is

$$\begin{Bmatrix} U_{1j} \\ U_{2j} \\ U_{1k} \\ U_{2k} \\ U_{1\ell} \\ U_{2\ell} \end{Bmatrix} = \begin{bmatrix} & & -L_{j\ell}/X_{2\ell} \\ 1.0 & & \\ (X_{1k} - X_{1\ell})/X_{2\ell} & & -L_{k\ell}/X_{2\ell} \end{bmatrix} \begin{Bmatrix} \delta_{jk} \\ \delta_{j\ell} \\ \delta_{k\ell} \end{Bmatrix} \qquad (12)$$

Transformation of this system of nodal displacement components to the global coordinate system is achieved by the use of the proper direction cosine matrix. The nodal force system is then computed by a premultiplication of the displacement components by the element stiffness matrix.

If the fold element is to be restored, the change in the angle of the fold is computed and a corresponding set of nodal displacements is introduced. For the coordinate systems shown in Figures 4 and 5, this set is simply $U'_{3\ell} = -\Delta\theta\ X'_{2\ell}$. A transformation to the global coordinate system and premultiplication by the fold element stiffness matrix generates a set of force couples which (if structure were otherwise unconstrained) are the small displacement approximation of the forces required to produce the observed fold opening. The correction force system is the negative of this vector.

Judgment of convergence is made by the user from a print-out of the largest absolute relative kinematic errors in the truss, triangle, and fold elements. For kinematic analysis the error associated with the fold element is ignored. Experience suggests that for rather large imposed displacements the first or second iteration solution is satisfactory for viewing purposes. Depending upon the constraints on the structure, it is normal for the kinematic error to reduce an order of magnitude for each successive iteration until limited by machine accuracy.

SECTION III

COMPUTER GENERATED DISPLAYS

Following the adage that "a picture is worth a thousand words," the past decade has seen the emergence of the field of computer graphics. The early work centered around the problems associated with the reduction of three dimensional data onto a two dimensional surface. The problem of perspective was treated by Smith[3] and Johnson[4], and a significant improvement was made by Roberts[5] with the elimination of hidden lines performed at a reasonable cost. Another important advance came as a result of the introduction of shaded pictures at the University of Utah (Romney[6], Warnock[7], and Watkins[8]), GE[9] and MAGI[10].

COMPUTER GENERATION OF CONTINUOUS TONE PICTURES

Computer generation of continuous tone images refers to the rendering of shaded objects by means of a raster driven cathode ray tube. The discussion is limited to a point source of illumination located at the eye-point so as to avoid the complication of shadows. The images produced are recorded by a camera mounted in front and pointed at the screen. As the images are being computed and displayed one line at a time (the pictures in this paper have 1024 lines), the camera, with the lens open, performs an integration function.

Hidden surfaces are removed as a result of comparisons of elements in the scene in order to determine which is in front of which. A simplification is obtained by applying a perspective transformation such that the observer is located at infinity in the direction of the Z axis, reducing the problem to a comparison of Z coordinates.

Some systems have been devised (MAGI[10], Comba[11], Weiss[12], and Mahl[13]) which have the capability to remove hidden parts for curved surfaces by restricting the class of possible surfaces or accepting long execution times. Recently at the University of Utah, Gouraud[14] perfected a method which uses a small polygon approximation of the surface to solve the hidden surface problem and then computes the shading on each polygon according to a linear function such that visual discontinuities between adjacent polygons disappear. The method (used in pictures shown in this paper) is both general and efficient.

SECTION IV

DISPLAYS OF KINEMATIC SYSTEMS

Figure 7 is a photograph of a computer generated continuous tone picture of a typical kinematic folded plate system. This hexagonal arrangement of approximately 600 triangular elements is shown in what has been termed "the flat configuration." That is, in this very regular geometry, the equilateral triangle elements all lie in a common plane.

Starting with the flat configuration, systems of constraints are applied to create various shapes. Figures 8, 9, and 10 indicate the dome like configurations achieved by holding nodes on the outer edge of the pattern and raising the center node. Characteristically, such shapes have quite uniform "star patterns" in the shallow dome configuration and non-uniform openings in the deep shell configuration.

Figure 10 shows an extension of the system which is close to the kinematic limit. That is, if the center node is raised much higher, pop-thru (a local buckling) of the star pattern will occur and/or convergence will be to an elasticity solution. Pop-thru (which is monitored by the simulator) is difficult to control and once begun often produces oscillations in the computational process.

Figures 11 and 12 show a warped effect which has been achieved by holding the center node and alternately raising and lowering the six corner nodes of the hexagonal system.

In both modes of behavior, there are six axes of geometrical symmetry which considerably reduces the computational problem for the simulator. However, as a perspective view of the structure is being created for an arbitrary point of observation, structural symmetry cannot be taken advantage of by the hidden surface routine.

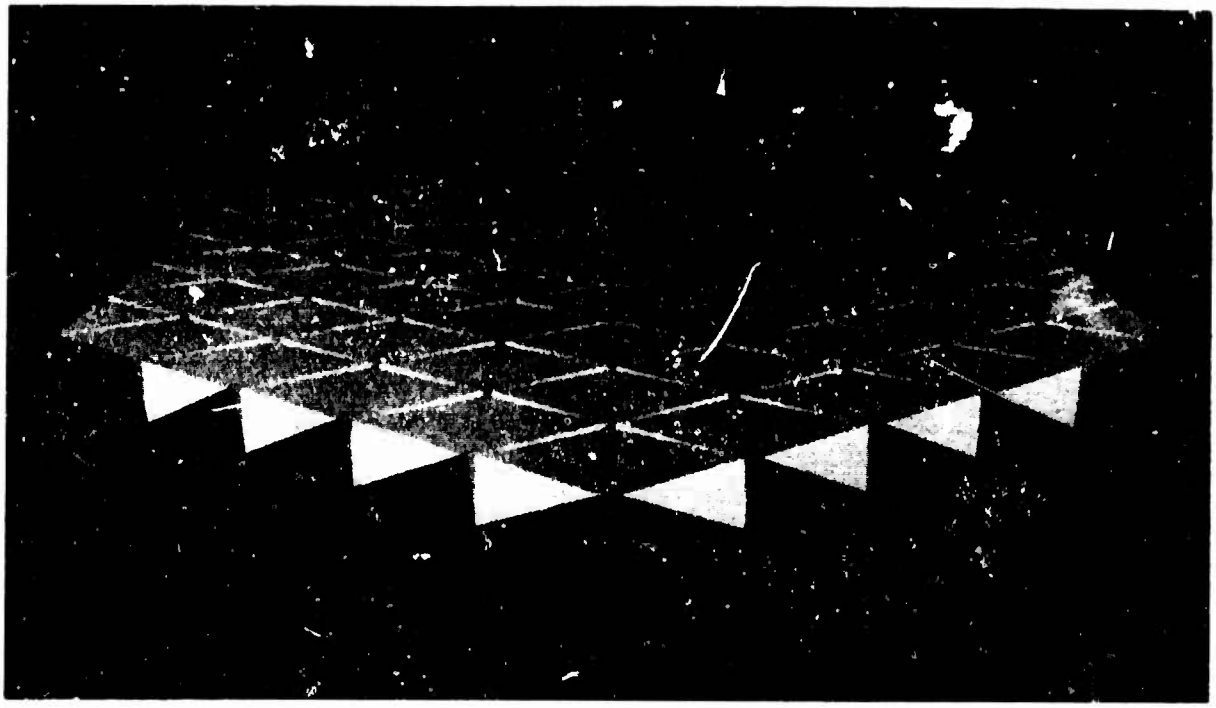


Figure 7. Flat Configuration

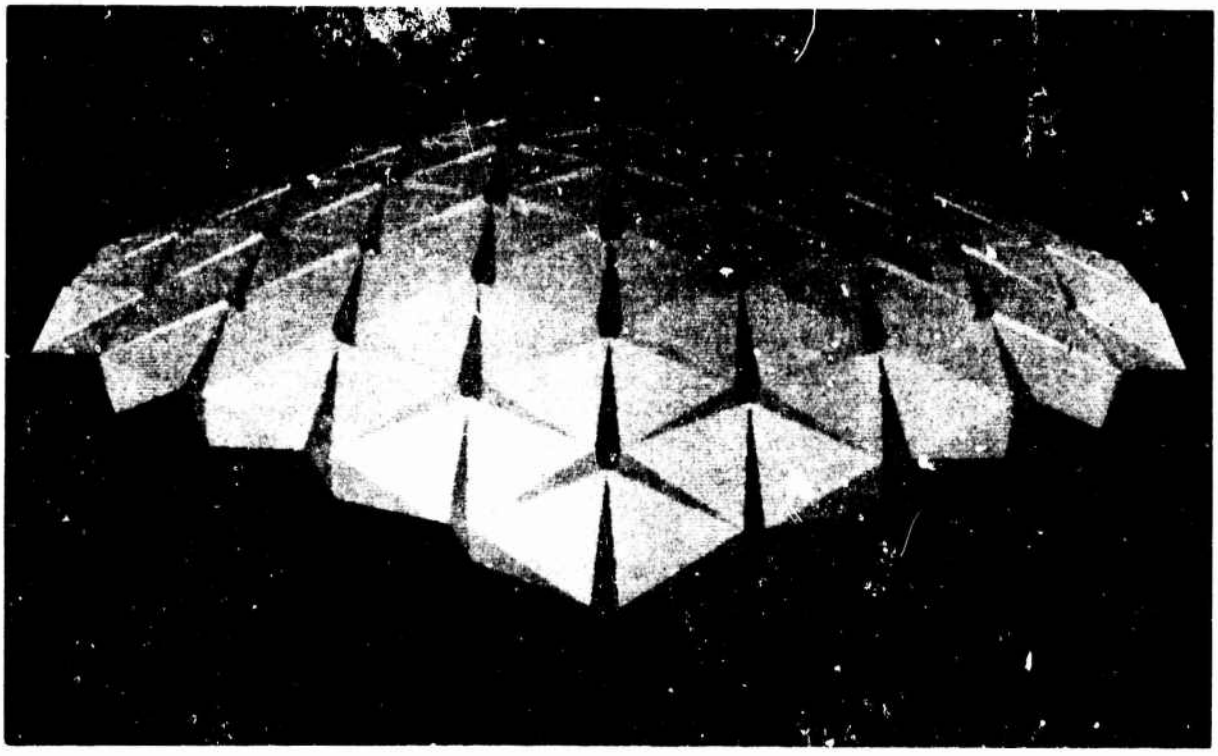


Figure 8. Shallow Dome

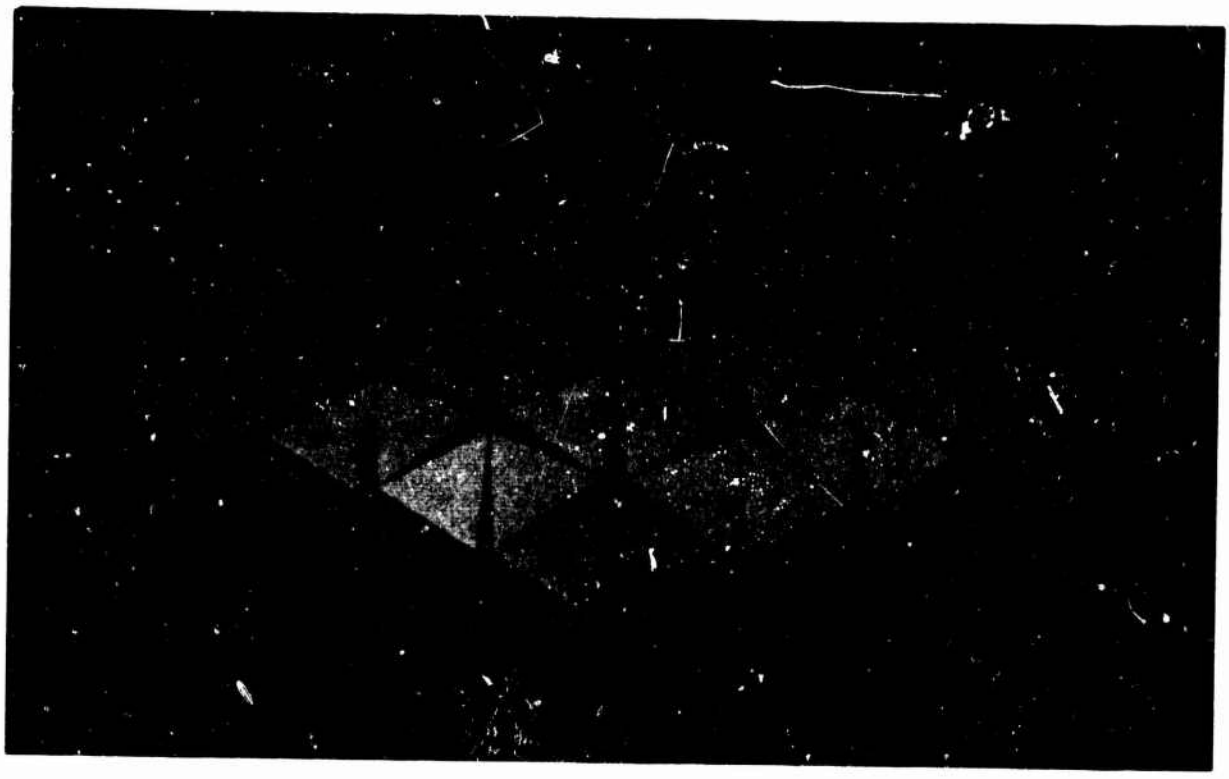


Figure 9. Intermediate Dome

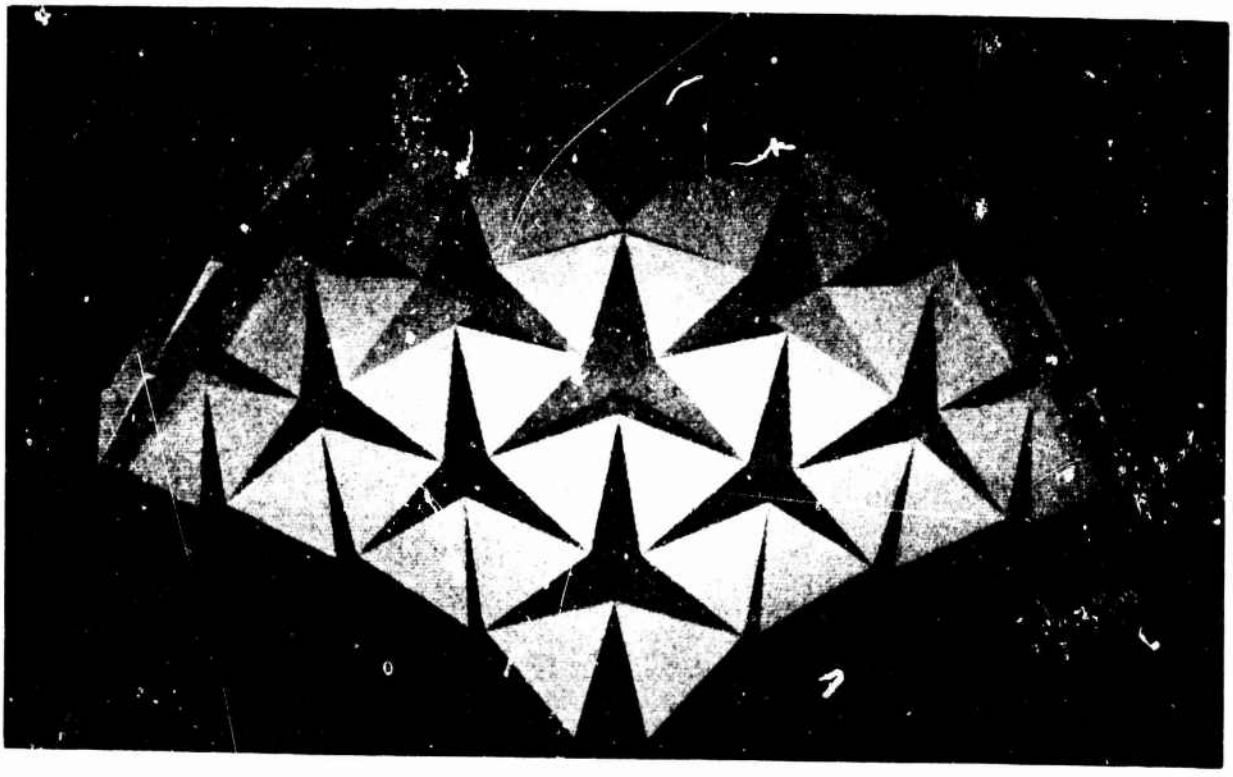


Figure 10. Deep Dome

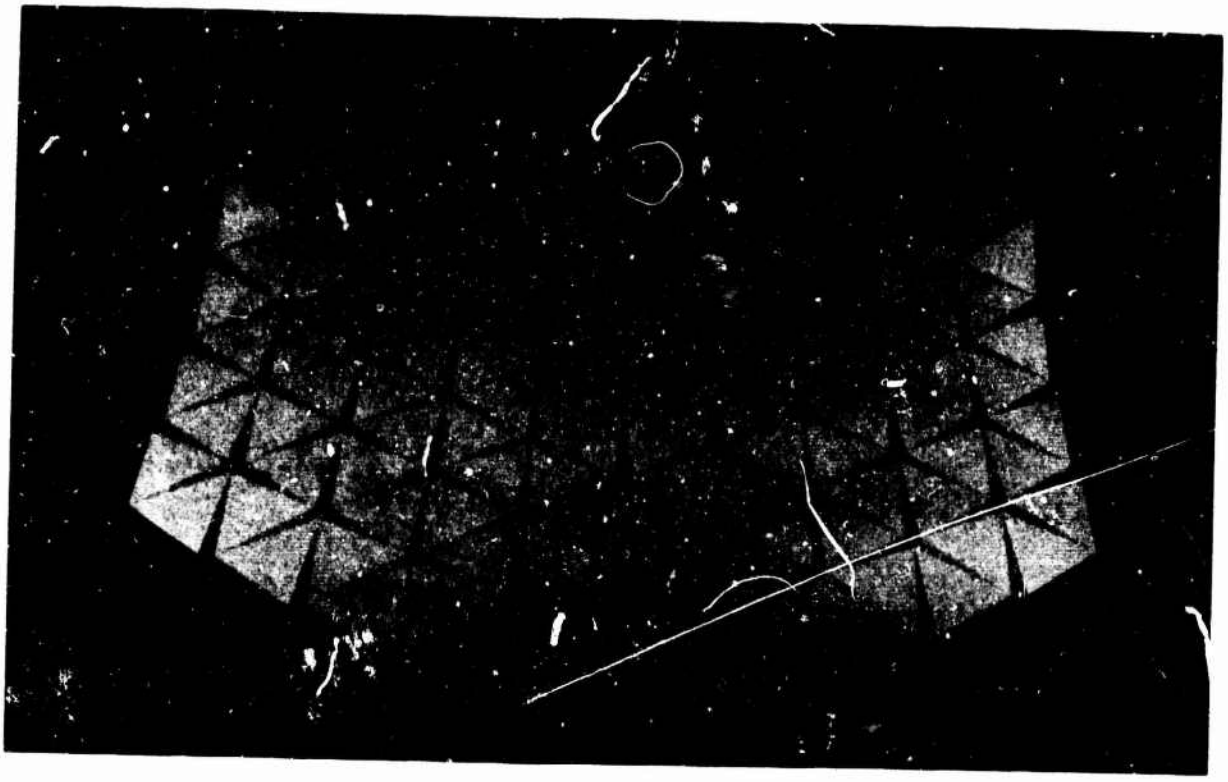


Figure 11. Slightly Warped System

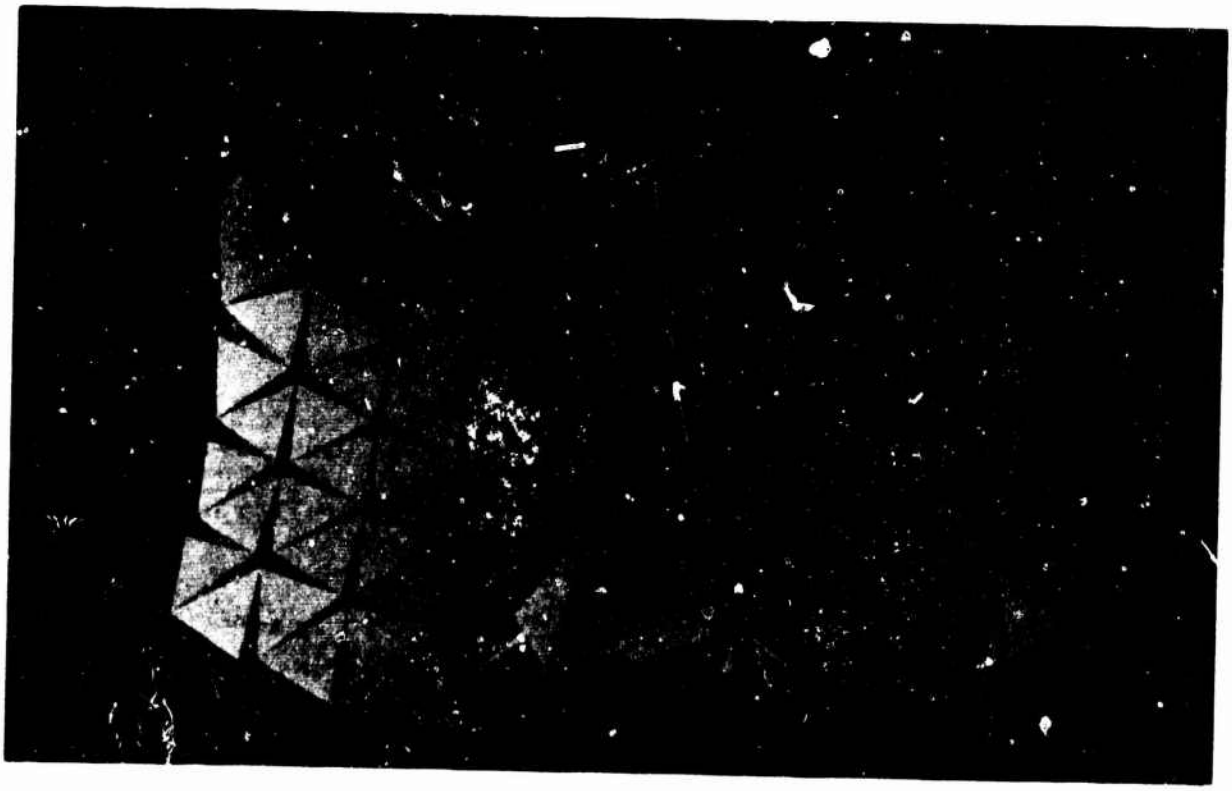


Figure 12. Highly Warped System

SECTION V

DISPLAYS OF ELASTIC SYSTEMS

INTRODUCTION

One compelling attraction of the photo-elastic technique is that it allows the analyst the opportunity to "see the stresses." With the thought that this benefit might be made available to finite element users, a study was launched into possible schemes for the presentation of stress, strain and/or displacement data available in a discrete digital format.

It was apparent that displacement data could readily be displayed in the form of a continuous tone picture of a highly distorted structure. The problem, then, centered upon ways to superimpose stress or strain information on the distorted geometry. Three methods have, so far, emerged from this effort. The first two involve the variation of the light intensity of the surface to indicate the stress or strain level. The third involves geometry modifications to accomplish the same purpose. Each method has, so far, only been applied to plane stress problems.

In each method, the data may be displayed in a "flat" or "smooth" manner. The "flat" presentation preserves the identification of the individual elements by "painting" the entire element the same average light intensity. "Smooth" shading utilizes the recently developed curved surface capability to create the illusion of continuous stress or strain variations. This is accomplished by employing linear interpolation to the light intensities computed for each corner node. In each of the pictures presented in this section, both formats are shown.

To facilitate these two formats, the elastic analysis capability of the simulator was modified in two ways. First, a quadrilateral element (composed of four constant strain triangular elements) was introduced. This element is more convenient with respect to automatic grid generation and provides relief from the grid bias problem. Second, the procedure utilized to calculate element stress and strain was changed. In order to achieve node values instead of element values, the standard procedure of premultiplying the displacement vector by stress matrices, was replaced by a least square fit routine. This procedure allows the prediction of strain components as a result of "fitting" the node displacement components with an incomplete cubic polynomial in the in-plane coordinates. The function, based upon a local coordinate system, is

$$u_i = a_{i1} + a_{i2}x + a_{i3}y + a_{i4}x^2 + a_{i5}xy + a_{i6}y^2 + a_{i7}x^2y + a_{i8}xy^2$$

The "fit" is applied to a local set of nine nodes.

BLACK TO WHITE SHADING

The initial scheme involved the shading of the elements according to the magnitude of the stress or strain component selected for presentation. Elements with a high value appear bright (i.e. highly reflective) and those with a low value appear relatively dark.

Experimentation suggests this to be a good procedure whenever high gradients are involved. For this reason, the method is useful in the display of beam problems. For these simulations, the most effective function for display is maximum shear stress (or strain). This function allows the combined display of the bending moments and shear diagrams. At the neutral-axis, the maximum shear stress is equal to the absolute value of the shear stress on the cross section and thus (at least for uniform beams) proportional to the shear diagram. At the top and bottom fibers, where shear stress on the cross section and the lateral normal stress components are both zero, the maximum shear stress is equal to one-half the absolute value of the axial stress component and, thus, monitors the bending moment diagram.

Figure 13 illustrates the "flat" and "smooth" shading of a uniform beam with "built-in" end conditions. The model, containing 504 quadrilateral elements, was assigned a modulus of elasticity of 28 million psi and Poisson's ratio of 0.25. The desired level of displacement was achieved by the specification of the displacement magnitude of the quarter span location. By scanning along the upper or lower edge of the beam, four changes in sign of the bending moment diagram can be observed. Repeating the scan along the neutral axis shows uniform shear with abrupt changes in sign at the three load locations (indicated by computer generated arrows). The "smooth" shaded picture, especially, suggests a local drop in the maximum shear stress under the one-quarter span loads. This is because the axial and lateral stress components are both approximately the same negative value, which results in a small value for the maximum shear stress. At mid-span the lateral and axial stress components are unequal and a local maximum is observed.

Figure 14 shows a beam with linearly tapered segments. This model also contains 504 elements with a modulus of elasticity of 10 million psi and Poisson's ratio of 0.15. Like the uniform beam this problem was posed in the form of specified nodal displacements and symmetry was utilized to reduce the computational effort. Note the very local compression at the tips of the beam and the interesting stress distribution above the supports.

Since this shape is not as familiar as the uniform section beam, the observer may lose track of the undistorted shape if large displacements are displayed. For example, many observers of this picture are uncertain of the linear taper for the unloaded geometry. On the other hand, the display of small displacements leaves the distortion pattern in doubt. The conflict is not easily resolved.

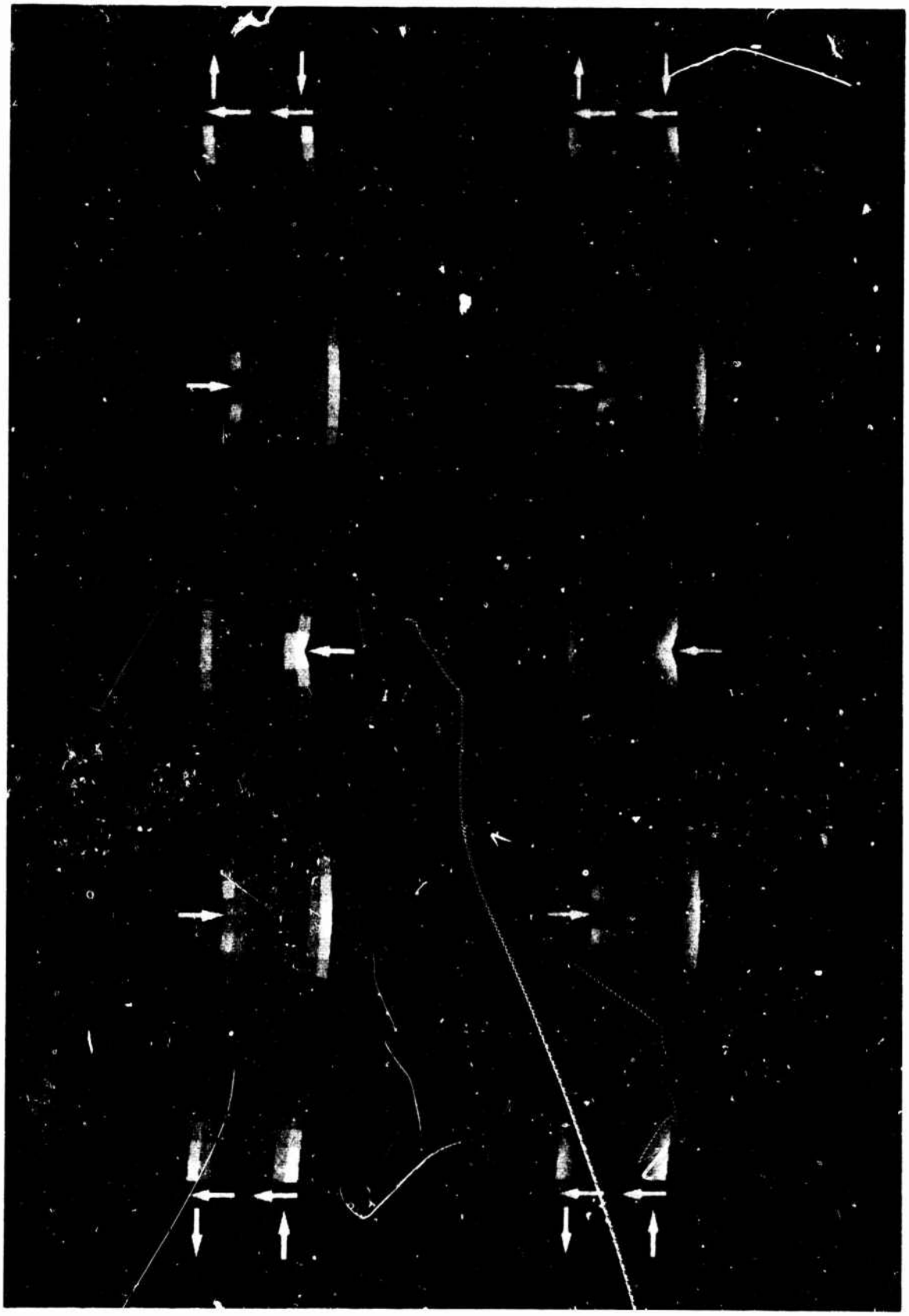


Figure 13. Uniform Beam

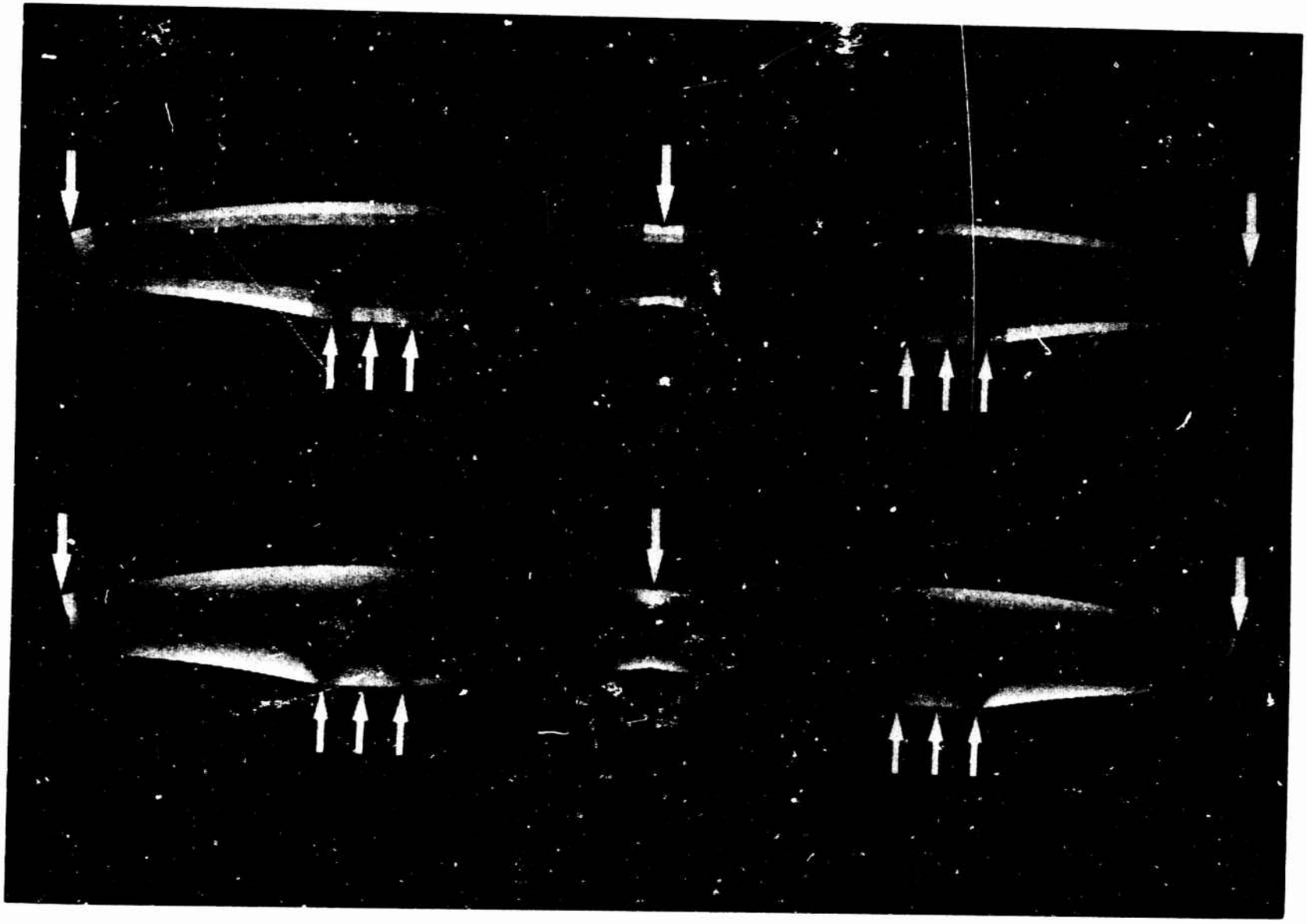


Figure 14. Tapered Beam

FRINGE PATTERNS

A modest change in the "black to white" scheme results in the display of fringe patterns. The effect is not unlike experimental stress optic patterns but has the advantage that the user may display the stress or strain component of his choice. The modification is to display a new function, ϕ, instead of the stress, σ, where,

$$\phi = \cos^2(\alpha\sigma+\beta)$$

The constants α and β are chosen so as to select a particular number of fringes and to shift the pattern.

Figures 15 and 16 illustrate this technique applied to the analysis of thin plate configurations. In both examples the modulus of elasticity is 1 million psi, Poisson's ratio is 0.45, and there are two axes of geometrical and loading symmetry. The function displayed is the maximum principal stress.

Figure 15 is a perspective representation of a rectangular (45 by 30 inch) plate with a circular hole (diameter 15 inches). The model contains 800 quadrilateral elements. The axial load varies from 66,700 psi at the center to zero at the outer edges. A complete shading cycle (white to black to white) represents a stress change of 50,000 psi.

Figure 16 illustrates a perspective view of a strip containing a circular ring section. The overall length is 8.4 inches, the width at the end is 2.4 inches and the diameter of the hole is 3.5 inches. The model contains 840 elements. The loading is a uniform axial stress of 20,000 psi. In this picture a complete shading cycle represents a stress change of 75,000 psi.

The main drawback in this format is the lack of identification of the stress level associated with any particular fringe. That is, if one scans from white to black, he may not know if the stress level has increased or decreased. Further, if he continues the scan into a new white zone, has he changed fringes? Color variation in parallel with light variation is currently being investigated as a possible solution to this problem.

A similar effect has recently been achieved by Sanford[15] who produced computer generated holographic interference patterns for known theoretical stress solutions. His patterns are shown in the undistorted geometry. An interesting contrast with Sanford's work is that, with closed form solutions, the analyst may ask for more fringes and (until the fringe width approaches the resolution of the equipment) he gets them. However, with finite element results, to ask for a large number of fringes usually results in a random pattern that looks much like a very poor quality Navajo rug.

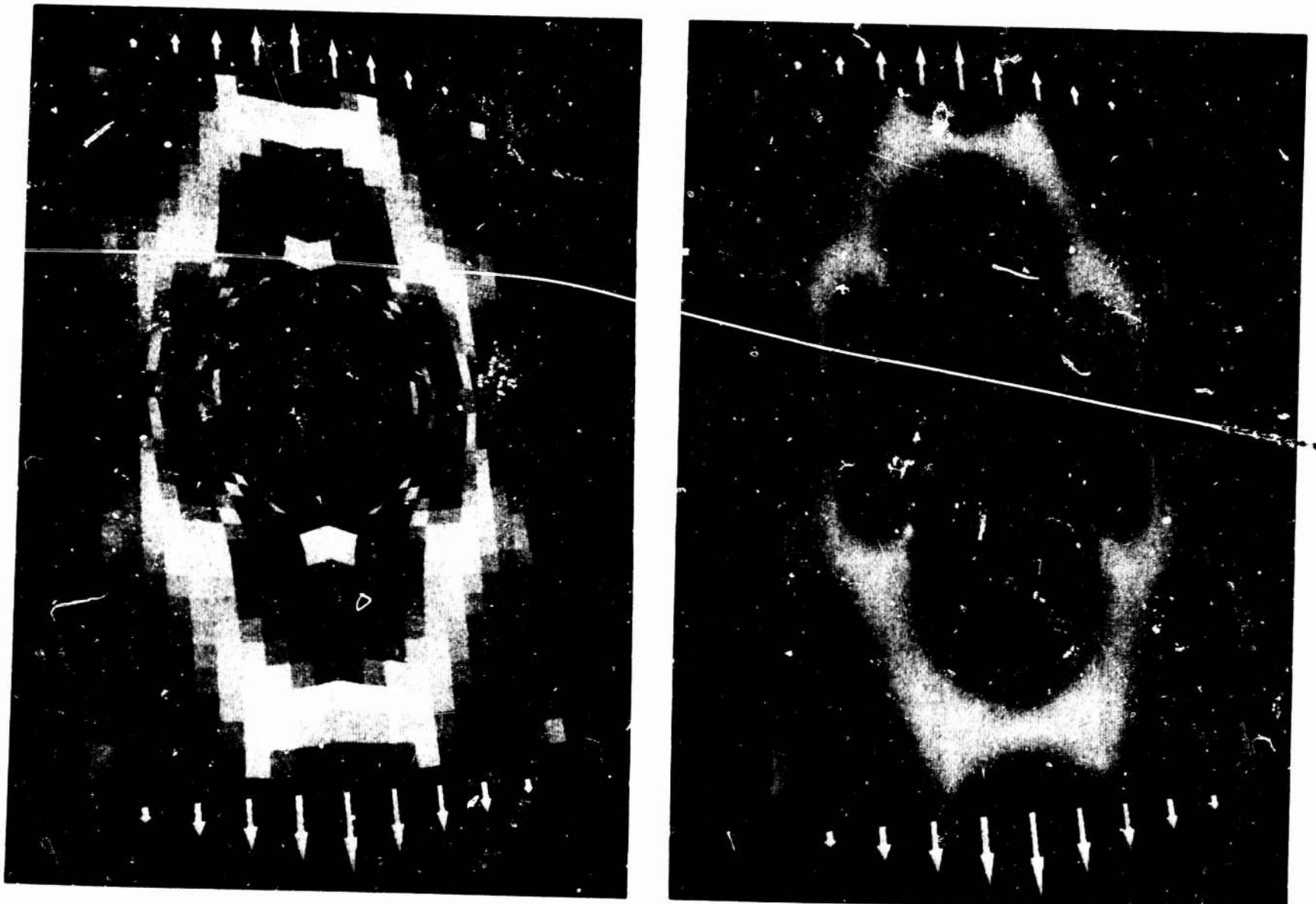


Figure 15. Plate with Circular Hole

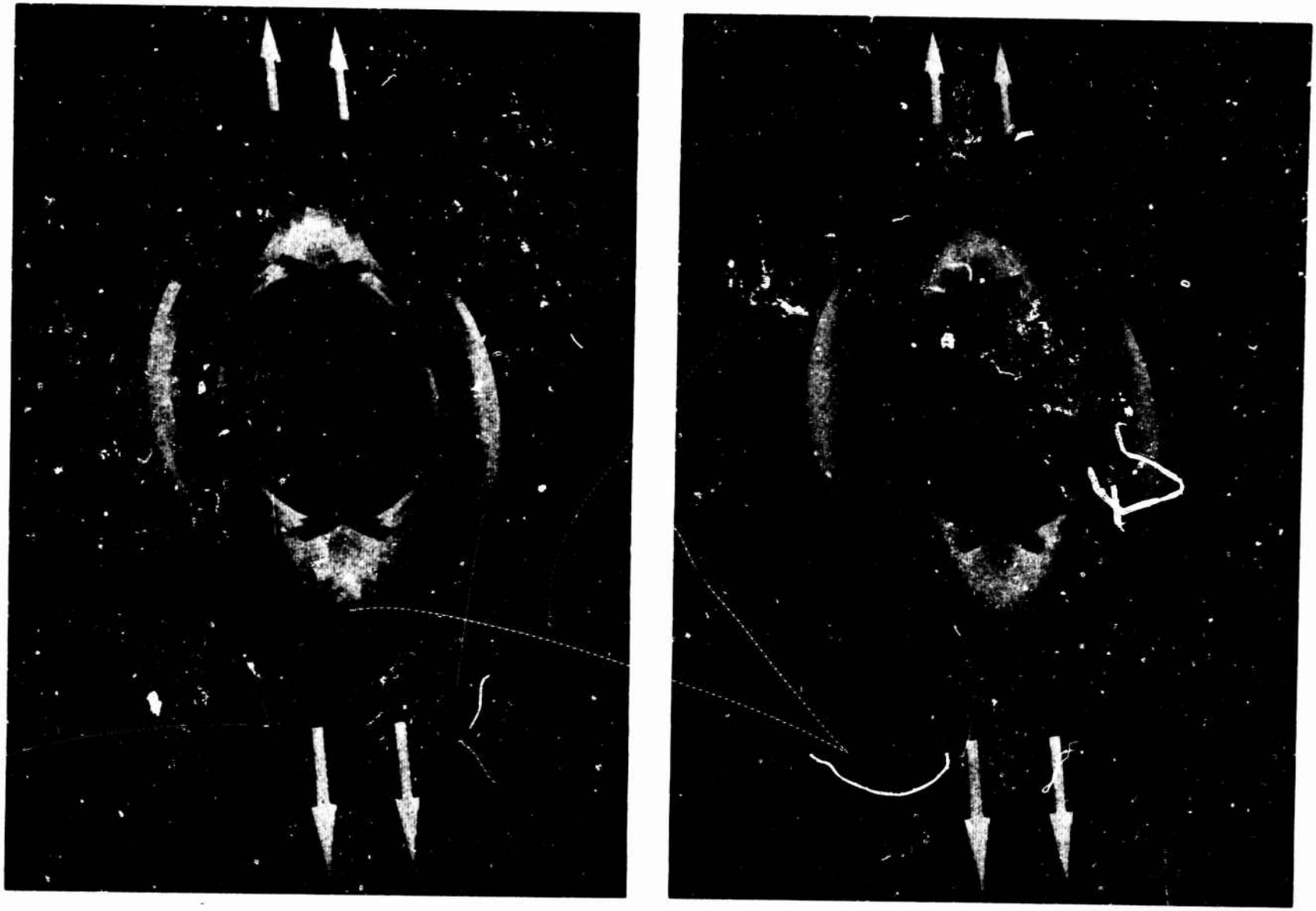


Figure 16. Strip with Circular Ring Section

WARPED SURFACES

The previously described display formats utilized variation in the light intensity of the surface to describe a chosen stress or strain function. An alternate procedure is to introduce geometry modifications in the form of computed out of plane coordinates which are proportional to the stress or strain function. This seems natural as functions of two variables are often displayed as a surface whose height above the reference plane is proportional to the function. Light variation is also utilized, but now to indicate the angle between the normal to the surface and the direction to the illumination source (the eye of the observer). For "flat" shading, the element normal is chosen as the average of the four normals computed on the basis of vector products of the adjacent sides. This average normal is also the normal obtained by the vector product of the diagonals of the quadrilateral. For the "smooth" shading version, the least square fit routine is again utilized to obtain normals at each node.

Figures 17 and 18 illustrate this display format for the problems previously discussed and displayed in a fringe effect format in figures 15 and 16.

In viewing this display form, the observer is acutely aware of the apparent stress concentrations (perhaps for the first time). While stress concentrations and sharp gradients do occur in these regions, the shapes displayed are dependent upon the local element definition and behavior of the least square fit procedure as utilized to generate strain values.

Comparing the "flat" and "smooth" shading versions, in the vicinity of the stress concentrations, suggests that very good definitions of curved surfaces may be developed from rather crude models. This is generally true. However, several drawbacks are also apparent. If this were the only display format available each picture would have to be accompanied by an undistorted picture to describe the geometry. Further it is difficult to find a satisfactory viewing angle and magnitude of distortion. The problem is somewhat compounded by the errors made by the hidden surface routine in the form of small holes in the surface. These errors apparently arise from difficulties the routine has in treating highly warped quadrilateral elements in a complex hidden surface environment.

It is expected that this form of display will come into its own when movies are made showing the transition from the undistorted to distorted geometry and allowing the observer to see the surface from many angles.

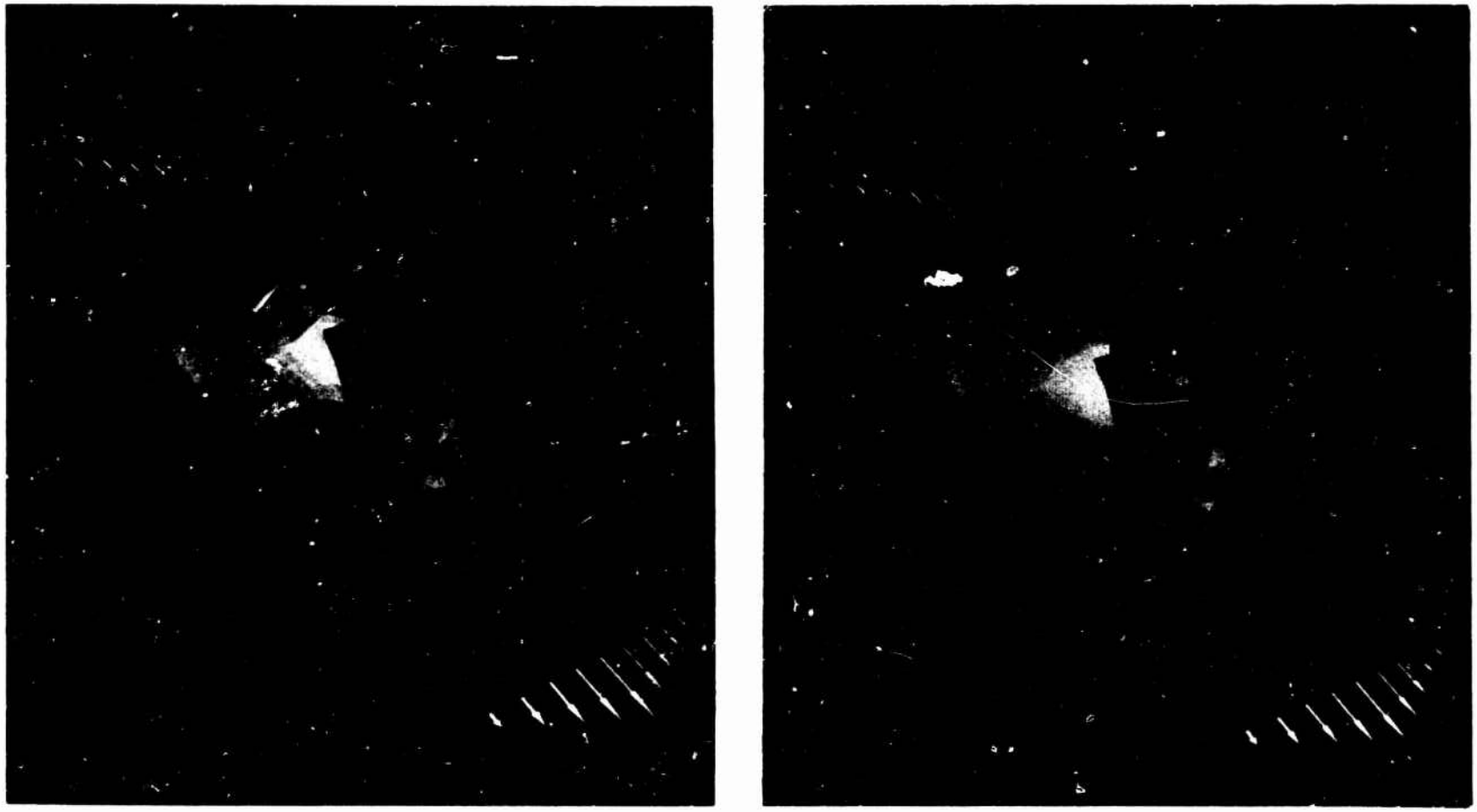


Figure 17. Flat Plate with Circular Hole

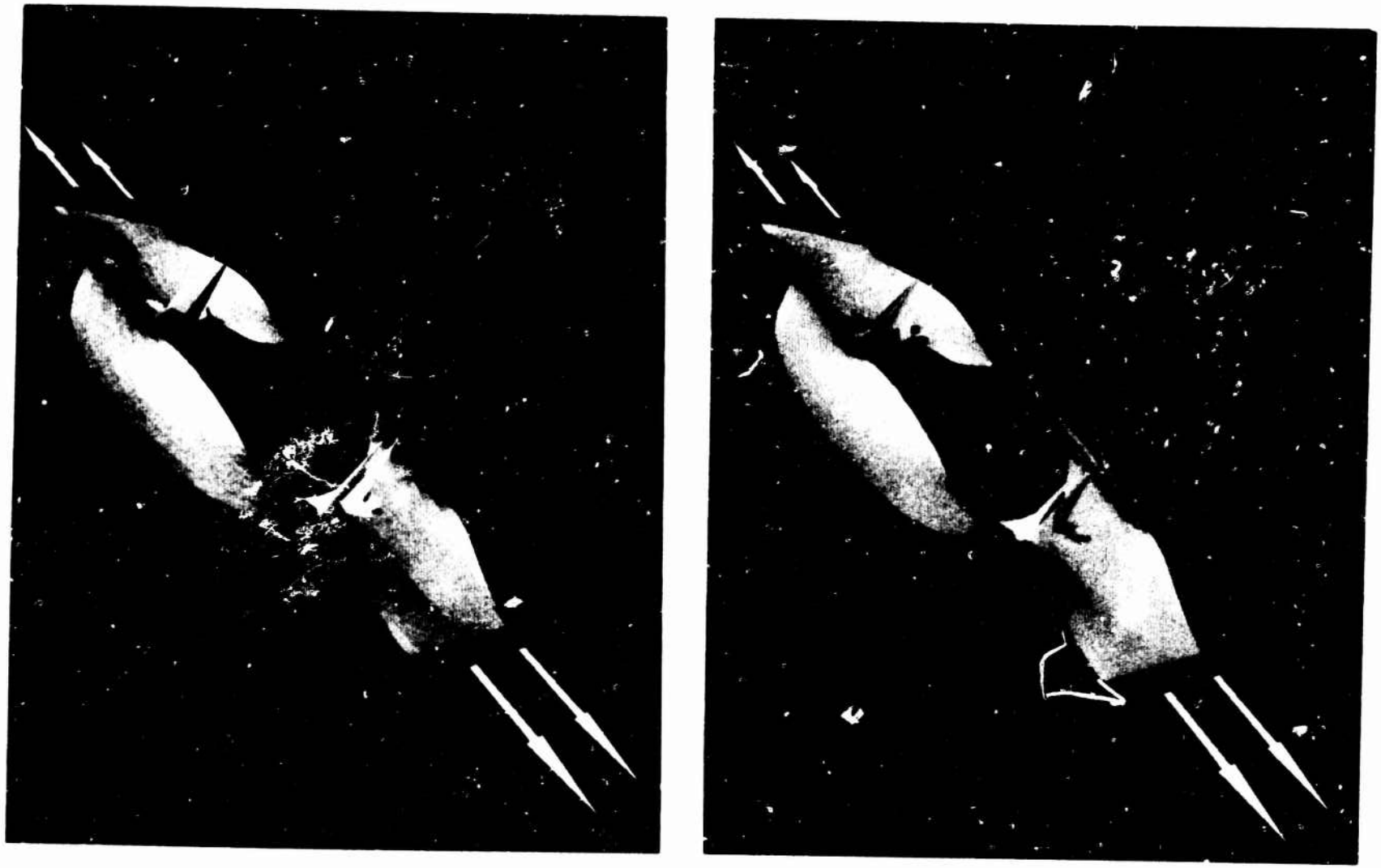


Figure 18. Strip with Circular Ring Section

SECTION VI

CONCLUDING COMMENTS

Usage to date suggests that a repeated linear analysis which employs the concept of corrective forces is an effective way to solve kinematic problems. Indeed, rapid convergence to an exact result has been the usual experience. Nevertheless, a greater potential is envisioned for the display techniques. Even here, the most likely occurrence is the development of superior methods of presentation as opposed to the adoption of these procedures.

In this regard, a current research project at the University of Utah, to design and build hardware to perform the hidden surface calculations, is very promising. Currently, the hidden surface computations require from one to two minutes of PDP-10 time for each picture. The hardware should accomplish the same task in 1/30 of a second. This will bring computer produced movies into their own, and make possible the recording of computer simulations in real time.

ACKNOWLEDGEMENT

The work presented here is part of a continuing effort at the University of Utah in graphics. This effort, sponsored by the Advanced Research Projects Agency, has produced a software/hardware system which is conducive to application studies. Pictures presented in this paper are a result of the efforts of many people in developing the basic software and hardware systems. In particular, appreciation is expressed to Ronald Resch whose program supplied the geometry of the folded plate system in the flat configuration and called the display routines for those pictures, to Mike Milochik the photographic engineer, and to Martin Yonke and Mike Dervage for assistance in making camera and scope adjustments. The author is also indebted to Dr. David C. Evans, Chairman of Computer Science, for creating the opportunity for this work to be done.

REFERENCES

1. R. D. Resch and H. N. Christiansen, "The Design and Analysis of Kinematic Folded Plate Systems," Proc. Symposium for Folded Plates and Prismatic Structures, International Association for Shell Structures, Vienna, Austria, Oct. 1970.

2. S. Timoshenko and D. H. Young, Theory of Structures, p. 231, McGraw-Hill Book Company, New York, 1945.

3. A. F. Smith, "Method for Computer Visualization," AMC Technical Report No. 8436-TM-2, Electronic Systems Laboratory, M.I.T., September 1960.

4. T. E. Johnson, "Sketchpad III: Three Dimensional Graphical Communication with a Digital Computer," Report ESL-TM-173, Electronic Systems Laboratory, M.I.T., June 1963.

5. L. G. Roberts, "Machine Perception of Three-Dimensional Solids, "Technical Report No. 315 Lincoln Laboratory, M.I.T., 22 May 1963.

6. G. W. Romney, "Computer Assisted Assembly and Rendering of Solids," Dept. Computer Science, Univ. of Utah, Salt Lake City, Tech. Rep. TR 4-20, 1970.

7. J. E. Warnock, "A Hidden Surface Algorithm for Computer Generated Halftone Pictures," Dept. Computer Science, Univ. of Utah, Salt Lake City, Tech. Rep. 4-15 June 1969.

8. G. S. Watkins, "A Real Time Visible Surface Algorithm," Dept. Computer Sci., Univ. of Utah, Salt Lake City, Tech. Rep. UTEC-CSc-70-101, July 1970.

9. R. S. Rougelot and R. Shoemaker, "G.E. Real Time Display," General Electric Co., Syracuse, N. Y., NASA Rep. NAS 9-3916.

10. MAGI, Mathematical Applications Group Inc., "3-D Simulated Graphics," Datamation, Vol. 14, Feb. 1968, p. 69.

11. P. G. Comba, "A Procedure for Detecting Intersections of Three Dimensional Objects," IBM New York Scientific Center, New York, N.Y., Rep. 39.020, Jan. 1967.

12. R. A. Weiss, "Be Vision, A Package of IBM 7090 Fortran Programs to Draw Orthographic Views of Combinations of Planes and Quadric Surfaces," Jour. Ass. Comput. Mach, Vol. 13, April 1966, pp. 194-204.

13. R. Mahl, "Visible Surface Algorithms for Quadric Patches," Dept. Computer Science, Univ. of Utah, Salt Lake City, Tech. Rep. UTEC-CSc-70-111, Dec. 1970.

14. H. Gouraud, "Computer Display of Curved Surfaces," Unpublished Ph.D. Thesis Dept. Computer Science, Univ. of Utah, Salt Lake City, June 1971.

15. R. J. Sanford and A. J. Durelli, "Interpretation of Fringes in Stress-Holo-Interferometry," Experimental Mechanics, Vol. 11, April 1971, pp. 3-8.

The Development of Computer Graphics For Large Scale Finite Element Codes

R. D. BOUSQUET*
D. N. YATES**
W. W. SABLE***
T. J. VINSON†

Missile Systems Division, Lockheed Missiles & Space Company

The development of an extensive series of computer generated graphics packages for large-scale finite element analyses is described. Suitable examples illustrating the means by which such graphics are used for eliminating input errors, output data reduction, report preparation, and radically reducing overall task time spans are presented. Programs for which such computer generated graphics have been developed include two-dimensional geometries, general structures and arbitrary three dimensional solids while special purpose versions have been used for such applications as movie generation and interference studies. Finally, Lockheed's recent activities in the development of interactive computer graphics for large scale production finite element analysis are discussed and illustrated.

* Strength Engineer, Vehicle Shell Systems/Loads, Structures and Dynamics
** Group Engineer, Vehicle Shell Systems/Loads, Structures and Dynamics
*** Structures Engineer, Vehicle Shell Systems/Loads, Structures and Dynamics
† Structures Engineer, Vehicle Shell Systems/Loads, Structures and Dynamics

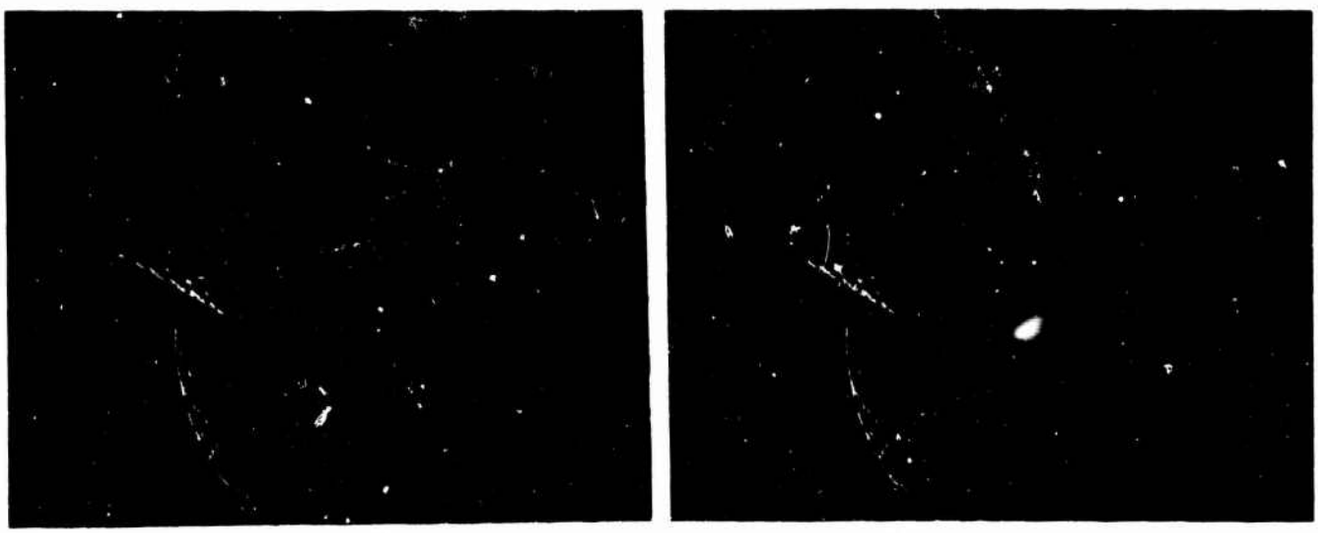
SECTION I
INTRODUCTION

THE NEED FOR GRAPHICS

The rapid development of the finite element method together with its widespread acceptance and increasing application in structural mechanics has brought with it a series of problems in addition to its many obvious advantages. Such problems as programming errors, machine accuracy and element shortcomings have been fairly commonplace. Nonetheless, the major problems have undoubtedly been concentrated in three areas: the difficulty in ensuring the input of correct mathematical models, the questionable ability to assess the validity of results, and, thirdly, the problems associated with assimilating, compressing, and organizing the voluminous output of data inherent in a finite element analysis into a usable engineering format and presentation. In all these areas it is clearly evident that the development of an extensive graphics capability simultaneously with the implementation of a finite element code would minimize and, in most cases, eliminate such problems. Over the past five years, therefore, the Missile Systems Division of Lockheed has pursued, as a major research effort, the development of a wide range of graphics packages for our library of large scale finite element programs, and these packages are now production features for all current codes.

The need for graphics was quickly established with the advent of the original Wilson axisymmetric code in 1965 (Reference 1). Reduction of the printed output from the first major analysis task required two months to complete, a timespan clearly unacceptable for a production oriented analysis group. Additional unfortunate early discoveries regarding the sensitivity of finite element analyses to such parameters as element aspect ratio and correct nodal geometry only served to emphasize the importance of graphics for input checks and determining validity of results. Finally, experience has shown that extensive graphics packages rapidly change from desirable aids to mandatory requirements as the complexity of programs and analyses increases, and as higher-order elements are developed and utilized (References 2, 3 and 4). Elements such as nonlinear isoparametric hexahedral solids (and the structural models for which they are employed) possess a degree of complexity far removed from earlier, two-dimensional constant strain formulations. The output data generated by such elements ranges from the ample to the awe-inspiring, thereby demanding that extensive graphics packages be present in order that the analyst can perform his studies with any semblance of confidence and rationality of scheduling. Hence, the true value of these graphics packages lies in their ability to

reduce errors, task time-spans, and computer and reporting costs. With these aims in mind, graphics in the batch process mode can be applied in two areas: prior to solution as a means for model debugging, and following execution as a method of display and data reduction. The debugging of a model using graphical model play-back allows the user to quickly identify and correct model errors, thereby eliminating the possibility of a faulty model being assumed correct and run for results which are obviously erroneous, or worse yet, containing hidden errors which could go undetected even through the output data reduction stage. One node out of place or the addition of an unwanted element could easily remain hidden through task completion, while destroying the entire model's validity. However, errors in input nodes and elements stand out clearly on a plot of the structure (Figure 1), particularly when the ability to automatically rotate and view the model from any position is present.



Model Showing Error

Corrected Model

FIGURE 1

DISPLAY OF INPUT ERROR

Thus, for input, time is saved by the user's ability to quickly and thoroughly debug his model, while costs are minimized by eliminating the need for reruns to obtain correct results. At the other end of the scale, computer graphics are needed as a means for reduction of output data both in an initial scanning mode and in producing a final format ready for direct insertion into a report. To accomplish these output tasks the engineer requires such graphical tools as element and nodal numbering displays, deflected figures, automatic stress and strain cuts, and isostress and isostrain contouring. For example, a deflected picture of the structure, by itself or superimposed over the undeflected figure, condenses all printed deflection output into a single visual entity. Another advantage of a deflected plot of the structure is that by magnifying the deflections and then plotting the structure, areas in which small deflections occur are exaggerated to a degree which allows the viewer to quickly determine the overall validity of the pattern of the model's deflections.

Plots of output results along any cut through the structure, together with a picture of the model defining the cut being taken, give the user a compact set of graphs showing stress and strain variation throughout any region in the model which can be compared to strain gage data (when available) along an identical cut, or inserted directly into a report. This concept compacts large quantities of output data into a few pertinent cuts allowing the user to pick out element stresses and strains at a glance. As presently developed, anywhere from one to nine of these graphs can be generated on a single page, depending on the quantity of data per graph and the amount of information to be displayed overall. Such schemes allow the engineer to significantly compress the area needed to present his analytical results.

An even larger amount of output data can be compressed into one picture through the use of contour plots. Using such plots the engineer is able to view the stress or strain pattern of an entire surface. Such patterns point out areas having high stress gradients. The first contouring program obtained by LMSC (Reference 5) contoured only data from a rectangular grid, but the same method of contour generation has now been adapted to handle models having arbitrary boundaries and deleted interior regions.

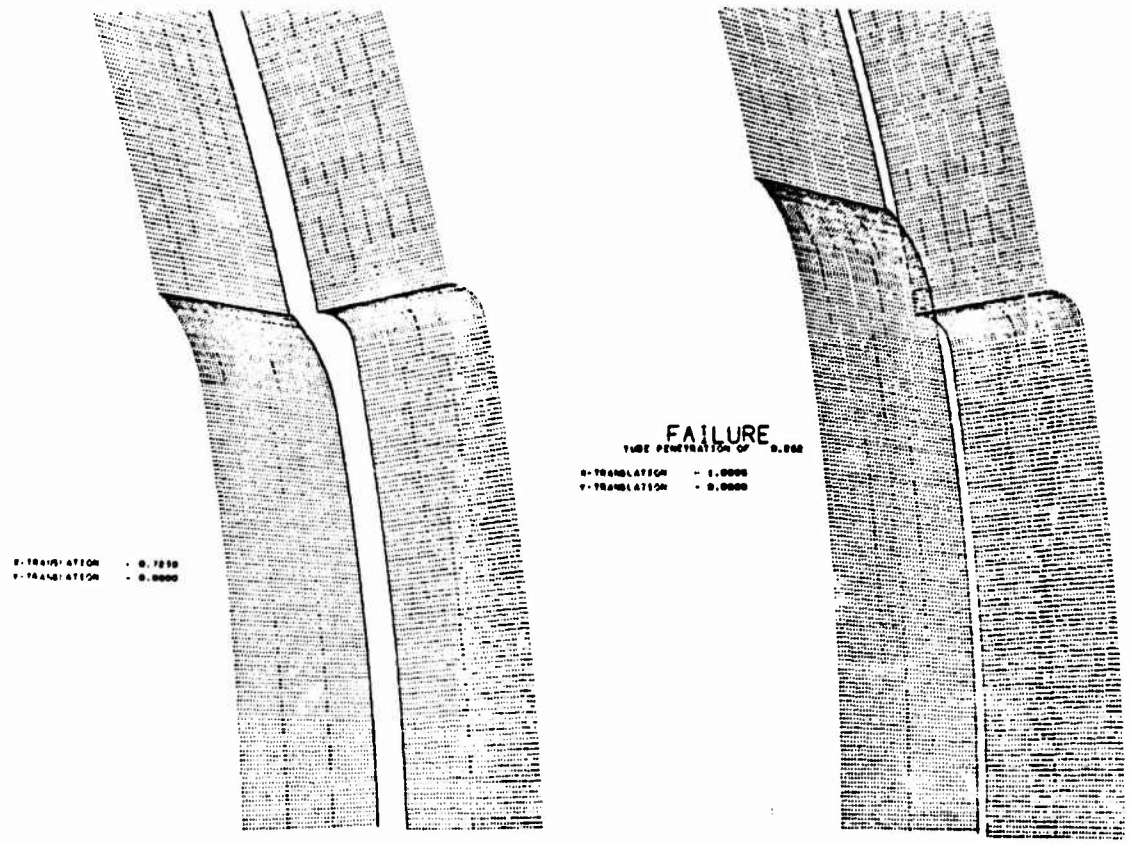


FIGURE 2

GRAPHICAL DISPLAY OF MISSILE STRUCTURE AND LAUNCH-TUBE INTERFERENCE STUDY FOR A SERIES OF POTENTIAL MISSILE ROTATIONS AND TRANSLATIONS

As noted in Reference 6, a large percentage of time on each structural analysis task is dedicated to the generation of a final report. With the use of computer generated plots this portion of a task is reduced to a minimum. The results are compact, complete, and entirely machine generated in a format which can be inserted directly into a final report, including page number, dates, names, and contract numbers. This report format is totally flexible and, hence, may be output directly as an 8-1/2 x 11 page with sufficient margins on all sides, or joined with additional data and output on an 11 x 17 page as demonstrated in Reference 7.

Neither is the use of computer generated graphics limited solely to the presentation of structural models and analytical results, for such schemes can be used to produce a display of a structure's proposed operating function or a visual representation of theoretical data such as the determination of interference failure limits for a missile and launch-tube interface for a large matrix of possible rotations and translations of the missile within the tube (Figure 2). Furthermore, it is often advantageous to present graphics data in the form of a movie to enhance visual representation. For example, a missile wing (Figure 3) was still in the preliminary design stage when a movie was completed by means of computer graphics which showed the proposed wing pivoting to its deployed position and deflecting under the maximum flight load conditions. Such movies have also been used to show the reaction of a model to a given load or to display various bending and torsional modal shapes. A model may also be positioned in space and then spun around and viewed from all sides, or an element-by-element build-up of the structure can be displayed to enable close examination of complex interactions between different sections of the model.

In the following section of this paper a detailed description of the means by which LMSC has employed such graphics packages to substantially reduce the time span and cost for both the input and output of finite element analyses is described in detail, together with recent extension of such activities into the area of interactive computer graphics.

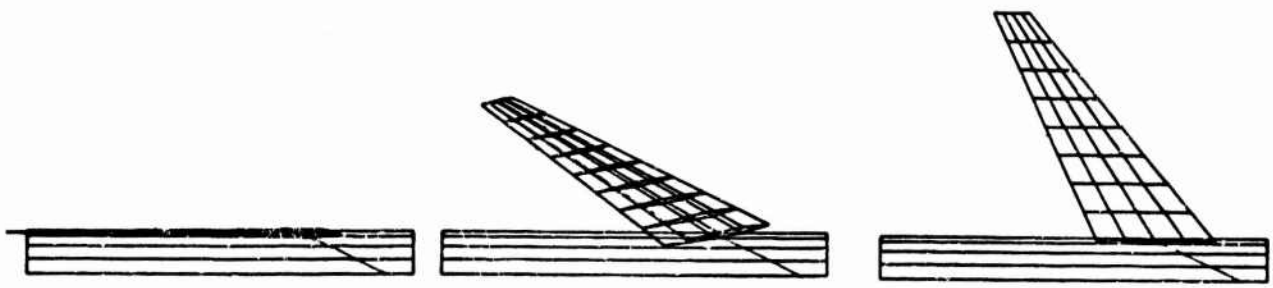


FIGURE 3

WING STRUCTURE SHOWING VARIOUS STAGES OF DEPLOYMENT

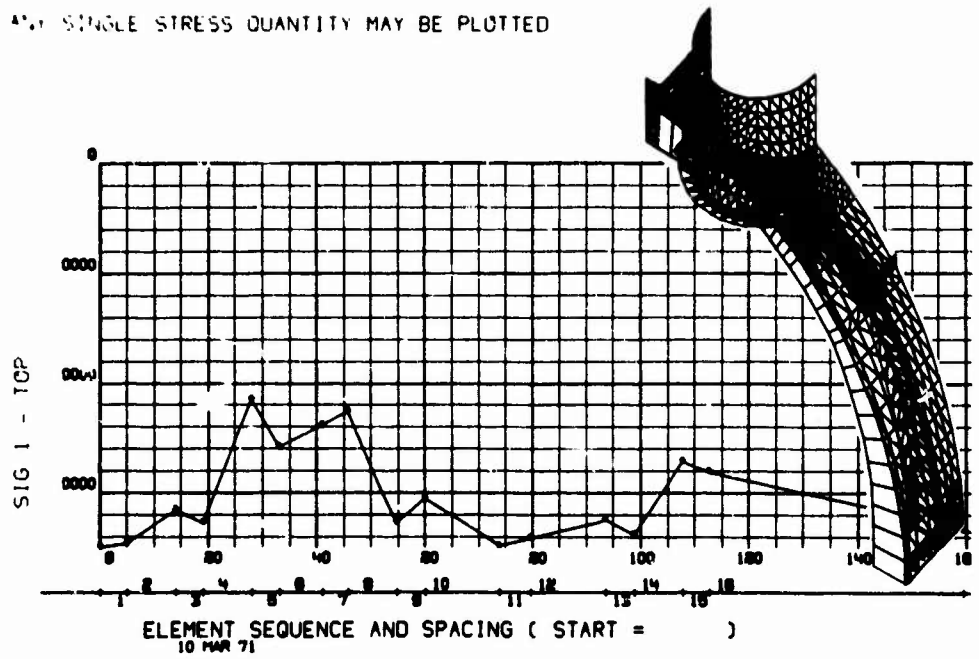
SECTION II

GRAPHICAL PRODUCTION TOOLS

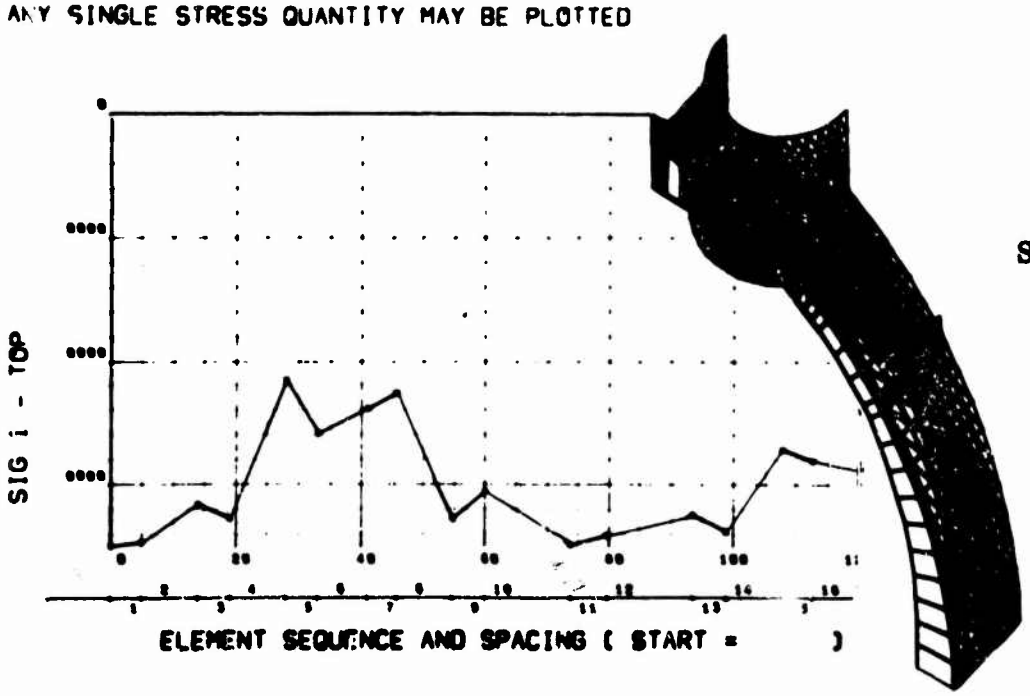
GENERATION OF GRAPHICS PACKAGES

There are a variety of machines currently available which can be used to present computer generated graphical data including both pen and CRT plotters. Lockheed's needs for a plotter with both hard copy and microfilm output, rapid plot generation, and a high degree of resolution have over the past eight years been filled through the use of a CRT plotter, specifically the Stromberg-Carlson 4020 system(generating plots in about .25 seconds and having 1024 x 1024 addressable rasters). Additional flexibility and resolution have recently been realized through the implementation of an III FR80 CRT plotter having 16,300 x 12,900 addressable rasters together with variable line width and character size. The improved resolution offered by the FR80 is illustrated by Figure 4. Every machine has different individual features (Reference 8) such as method of generation, resolution and speed, but the basic routines available to the programmer are the same. The generation of any graphical output, no matter how complex, is merely the manipulation of these few basic routines: line drawing, character display, scaling, and grid generation. The programmer uses these routines to pass his model's node and element information together with his output results from the computer to the display unit in a mode suited to his display needs. The display unit may be a hard copy or microfilm generation or any of the number of interactive display scopes, but in general the large scale batch process user relies on the hard copy option.

The structural graphics packages developed thus far fall into three main categories, depending on the finite element codes to which they are linked: 1.) Two-dimensional structures, 2.) Mixed structures, and 3.) Arbitrary solids. Each type of code varies in the form of model generated and the output quantities determined, while a similar variation exists in their accompanying graphics packages. The following section describes the current graphics schemes employed to reduce and display data for each of these three types of structural codes. A final section discusses the implementation of an interactive graphics mode for such codes, which further reduces task time span and costs.



FR-80



SC-4020

ANY SINGLE STRESS QUANTITY MAY BE PLOTTED

FIGURE 4
COMPARISON OF OUTPUT FROM TWO C.O.M. PLOTTERS

TWO DIMENSIONAL GRAPHICS

These codes analyze structures in which the finite element model can be defined in two dimensions. Examples are axisymmetric structures and plane stress structures of constant thickness. In the finite element approach the continuous structure is replaced by a system of axisymmetric elements which are interconnected at circumferential joints or nodal circles. Hence, such structures only require that graphics packages work with and display on the basis of a single XY coordinate system.

The model for an axisymmetric program makes use of mesh generation routines in which the user inputs model boundaries and a mesh pattern, with the program using this data to generate the finite element model (Figure 5).

While setting up the initial model, a user can vary the grid intensity in the critical regions to get more detailed results which may later be presented as contour plots (Figure 6), at the same time setting up definite element boundaries between different materials. A model can then be graphically displayed both in its entirety, or with any portion of the structure blown up to clarify a region with a denser grid.

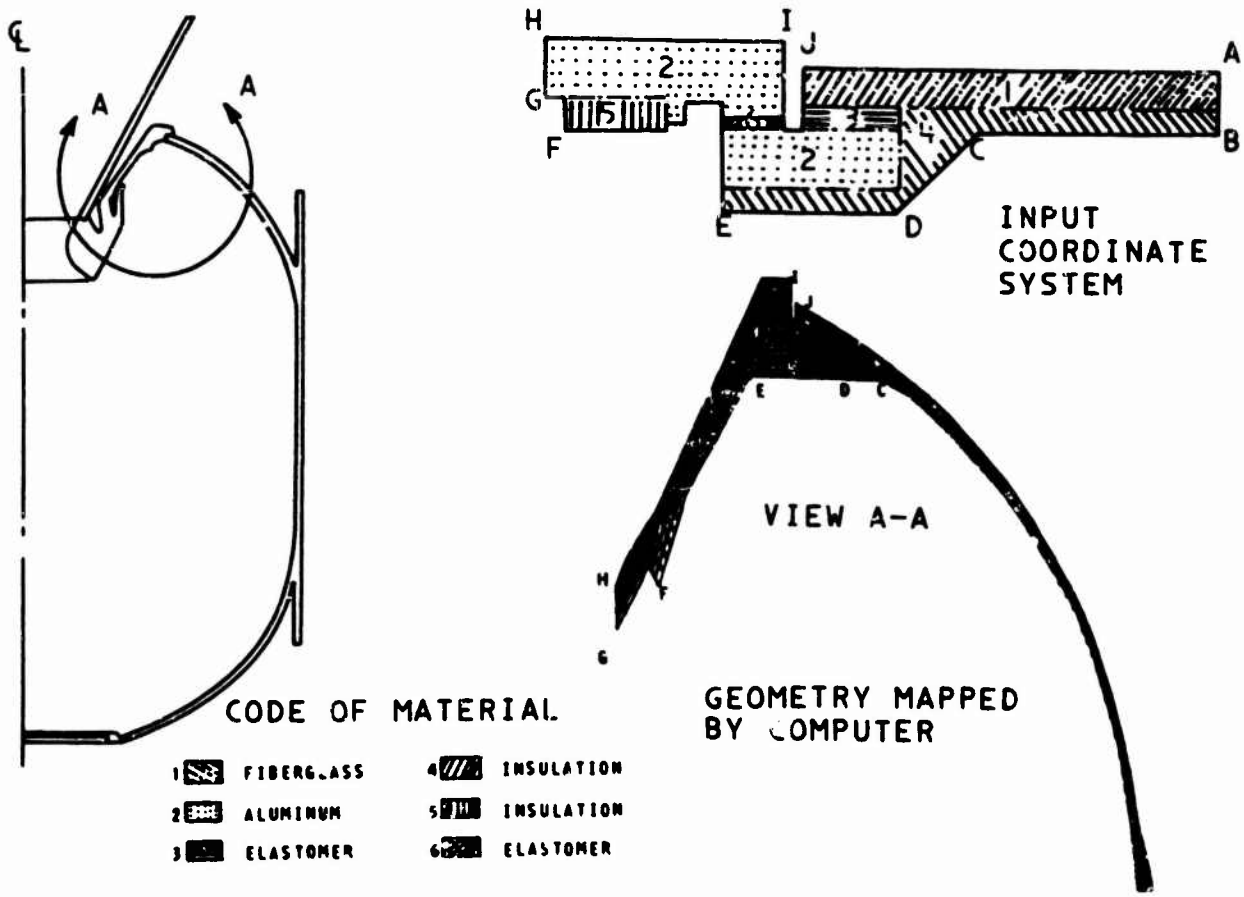


FIGURE 5

FINITE ELEMENT MODEL GENERATION

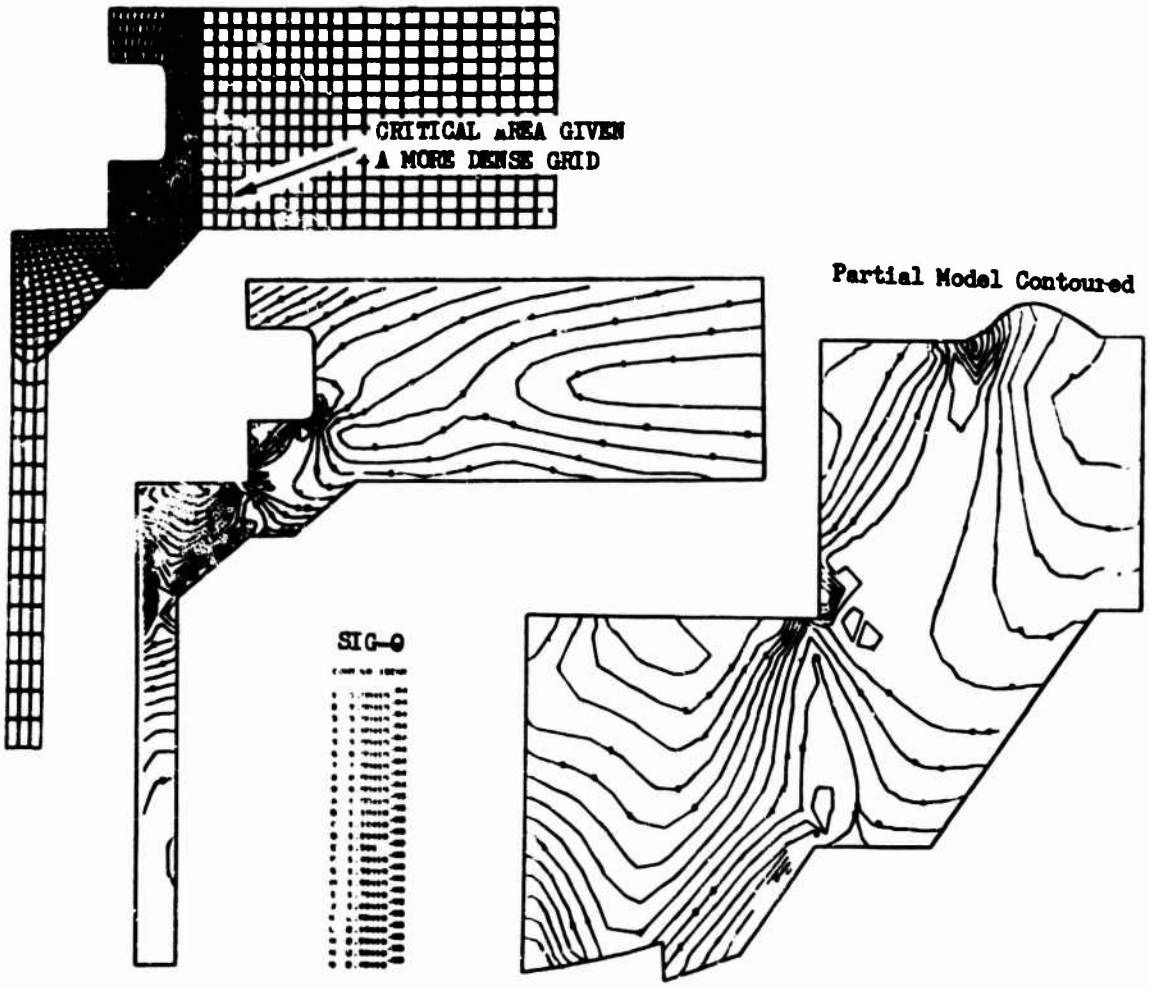


FIGURE 6
FINITE ELEMENT MESH AND RESULTANT
CONTOURS OF TANGENTIAL STRAIN

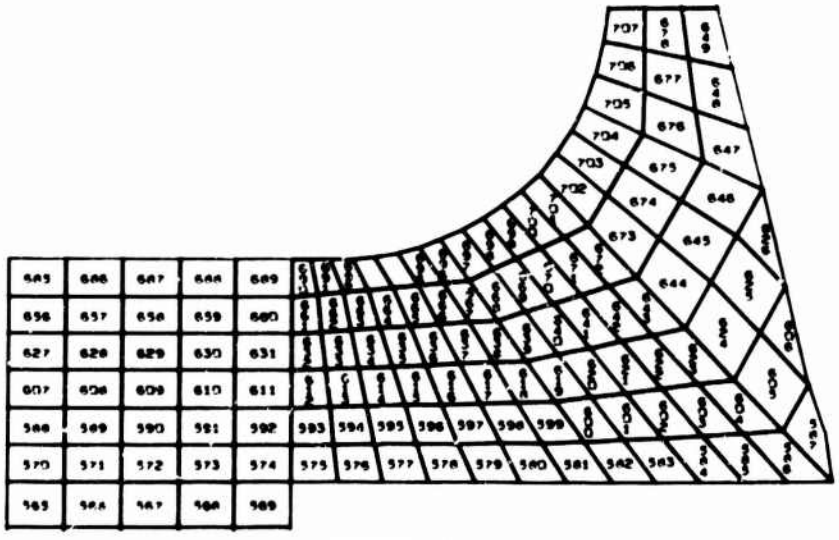


FIGURE 7
PARTIAL MODEL WITH ELEMENT NUMBERS DISPLAYED

A representation of the structure with element or nodal numbers displayed (Figure 7) is invaluable to the user when he is reducing nodal deflection data and element stress and strain outputs, or when he is attempting to detect input errors.

On symmetric structures, only a portion of the structure need be modeled for analysis, but for visualization purposes symmetric models may be "flip-flopped" to generate the entire structure (Figure 8) from a single plane of symmetry.

After the model has been analyzed the data may be reduced through the use of stress and strain cut plots (Figure 9), deflected plots (Figure 10), element state plots, contour plots, and time history plots, all of which may be output in a specific report format if desired. Deflected plots may be of the entire structure or any portion, and may show actual or magnified deflections (Figure 11). Element state plots display which elements are plastic, and which elements have been loaded and unloaded during an incremental solution. Contour plots of the full structure or any portion may be obtained with the number of contour lines being automatically selected to the desired line density (Figure 6). In the case of thermal loadings, contours of the temperature

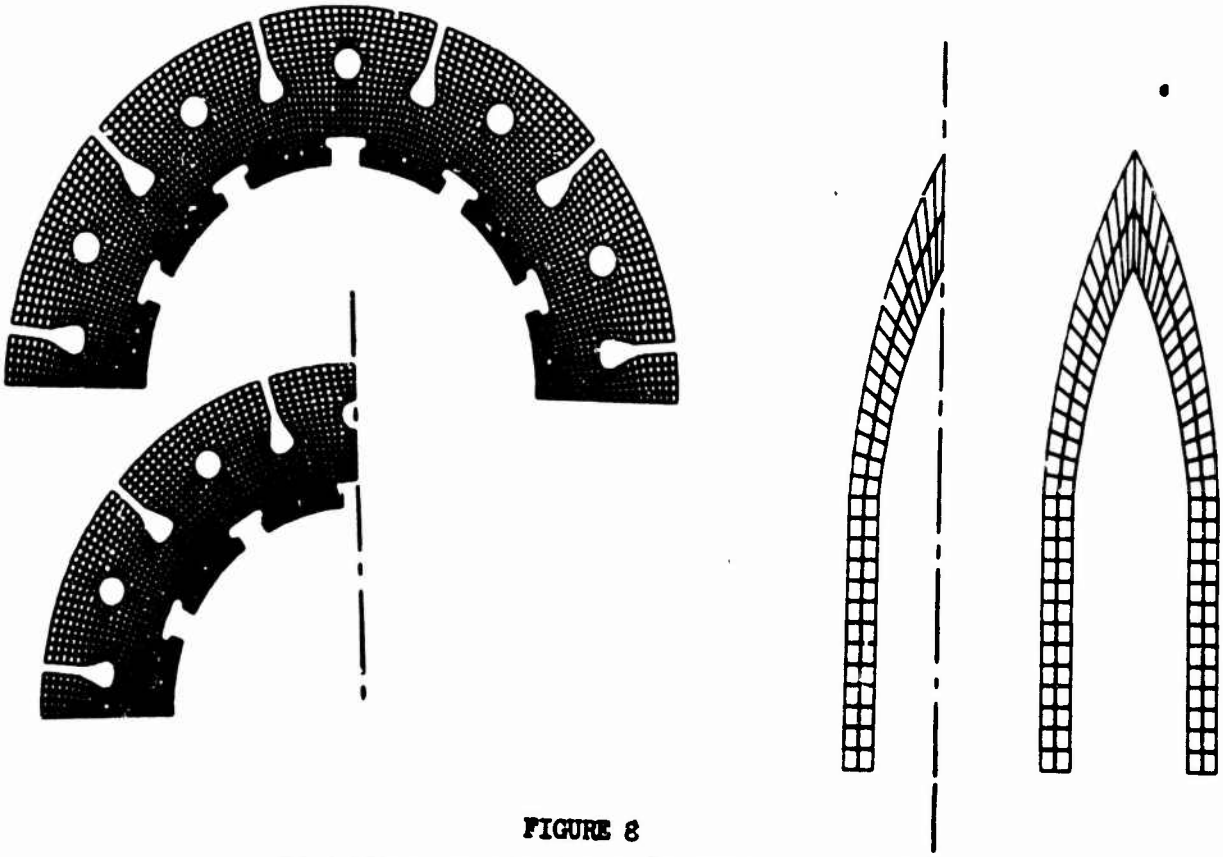


FIGURE 8

SYMMETRIC FIGURES MAY BE REVOLVED ABOUT THE AXIS OF SYMMETRY RESULTING IN A PLOT OF THE ENTIRE FIGURE

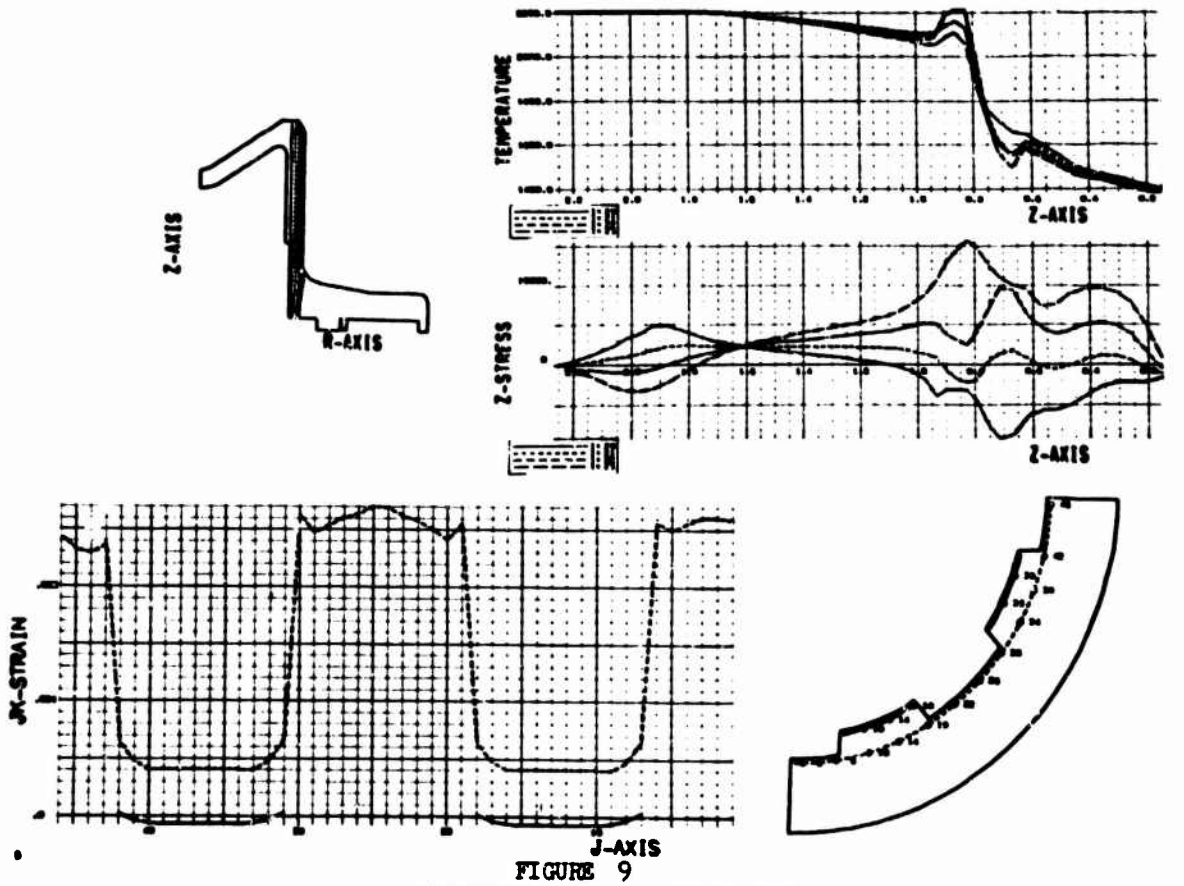


FIGURE 9
ANALYTICAL RESULTS ARE GRAPHED ALONG
ANY CUT THROUGH THE MODEL

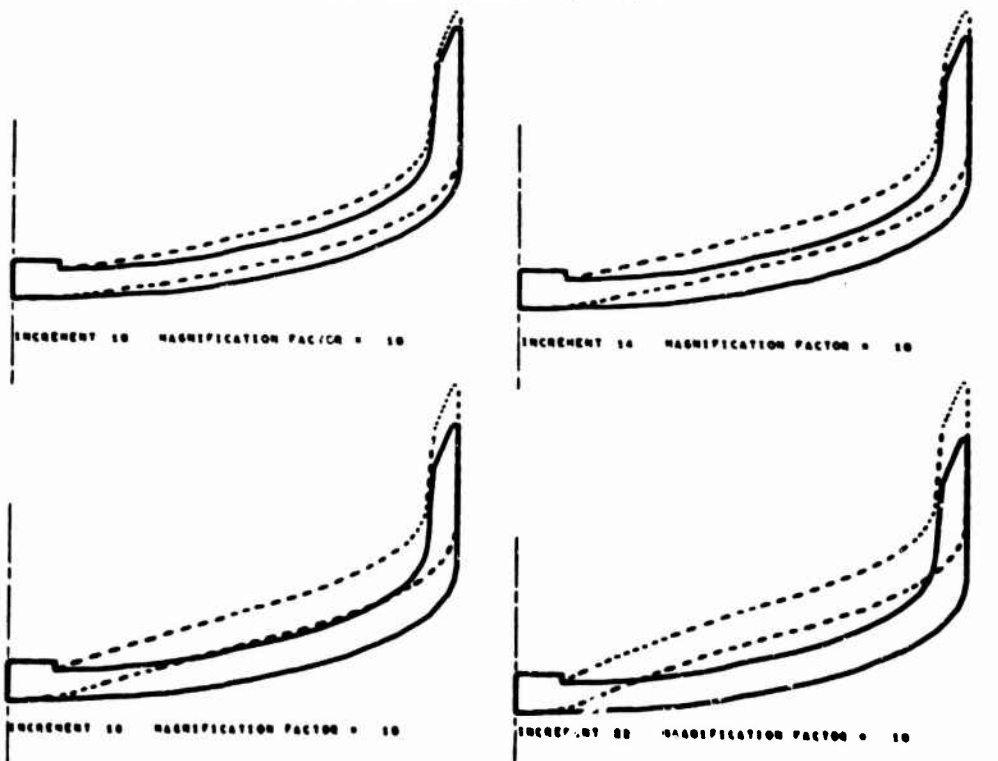


FIGURE 10 INCREMENTAL BOUNDARY DEFLECTIONS

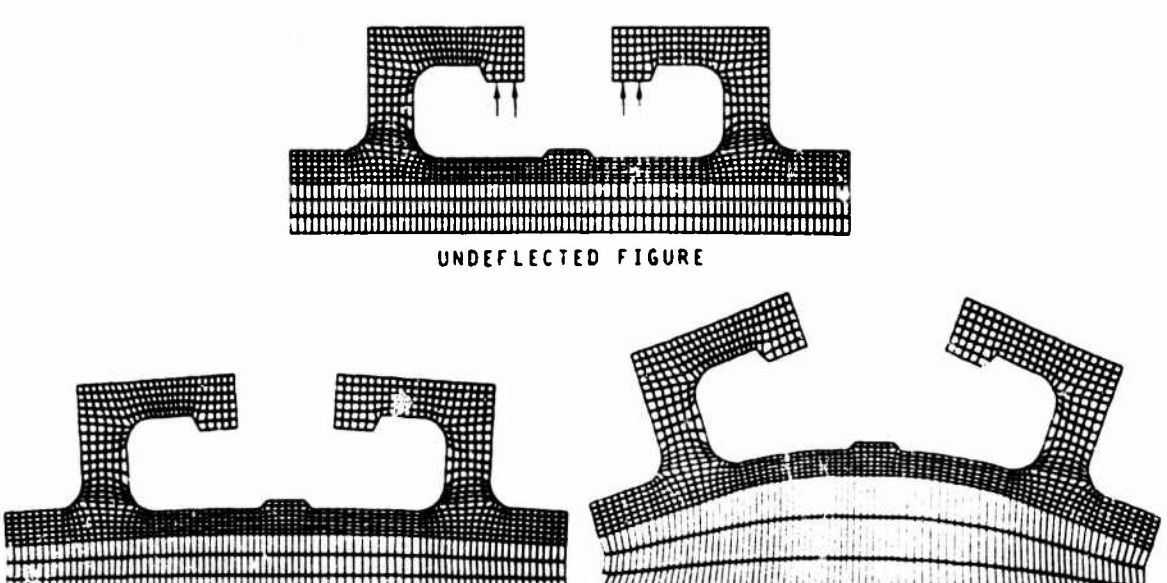
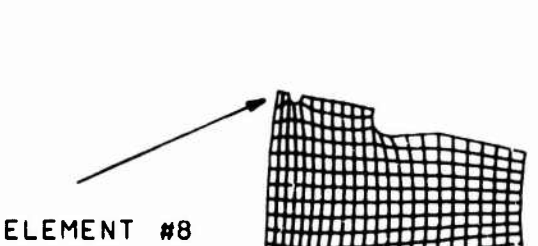
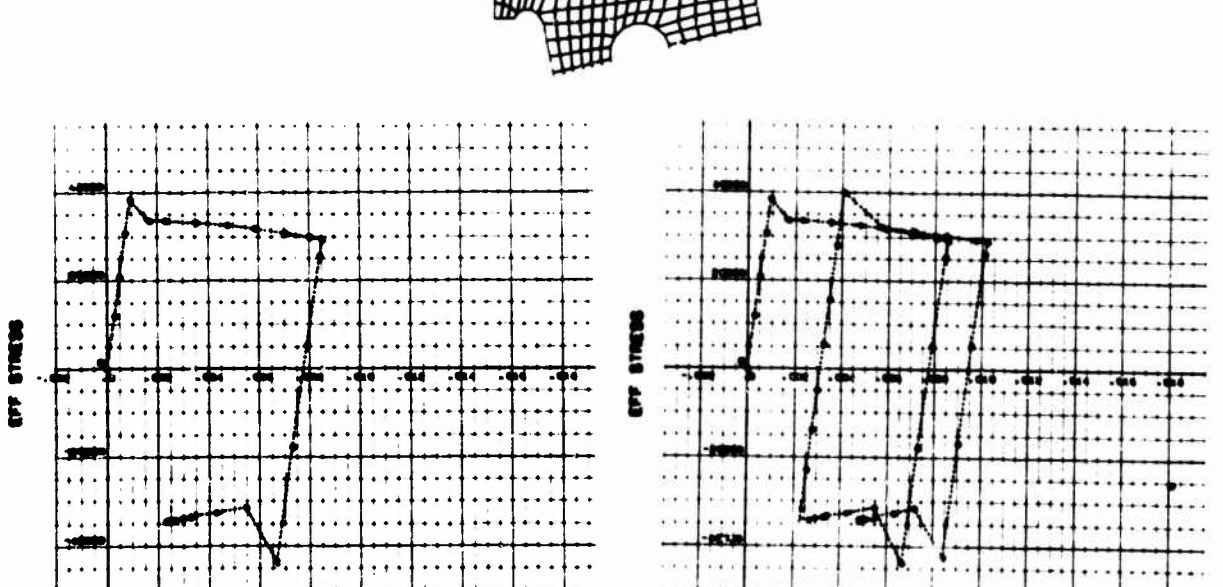


FIGURE 11 DEFLECTED FINITE ELEMENT MODEL





FIRST THERMAL CYCLE FIRST AND SECOND CYCLE

FIGURE 12 STRESS-STRAIN HISTORY (ELEMENT #8)

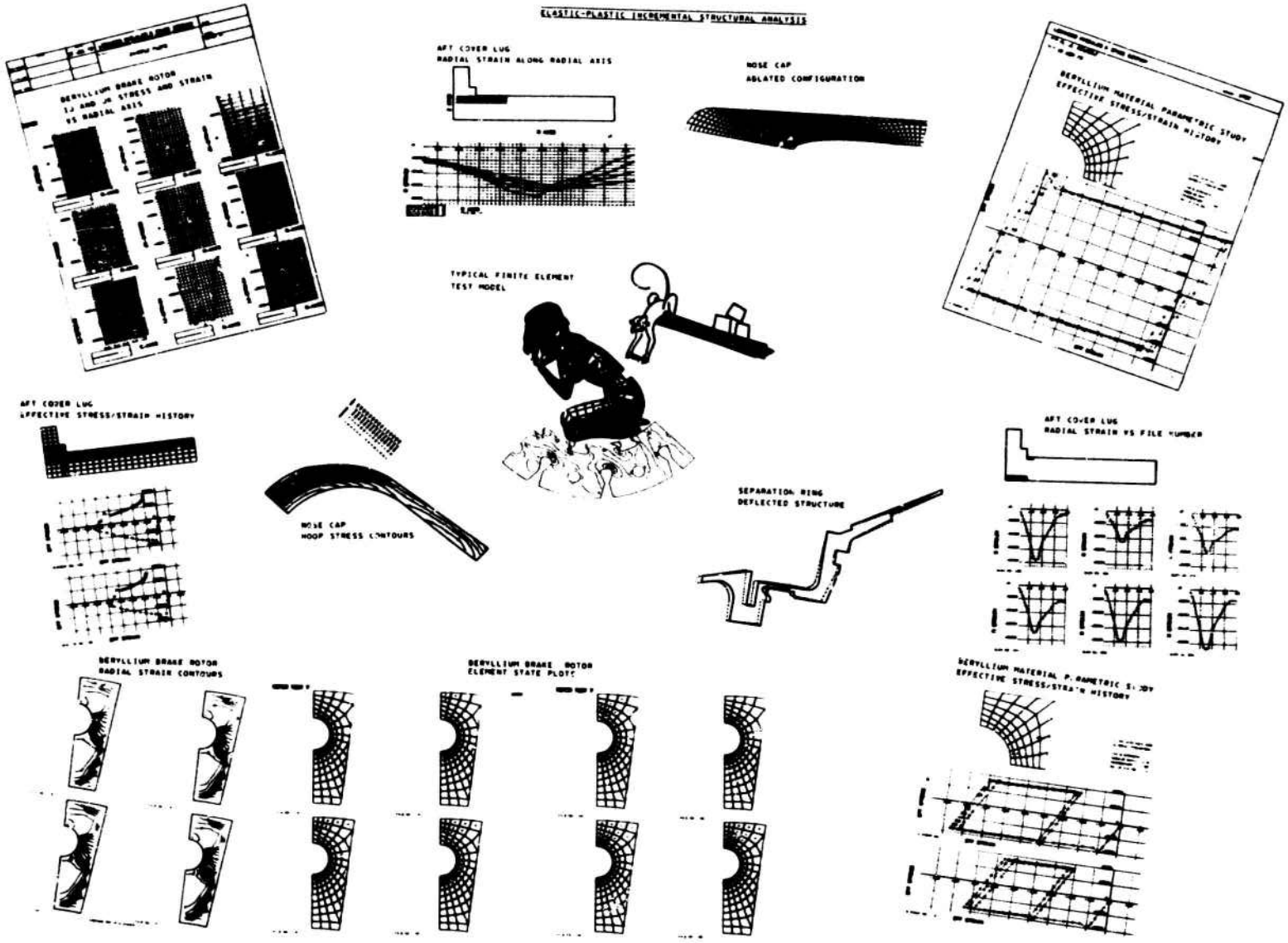


FIGURE 13
GRAPHICS PACKAGES AVAILABLE FOR
TWO-DIMENSIONAL FINITE ELEMENT MODELS

gradient applied to the structure may also be obtained. Output stresses and strains may be graphed along any cut through the structure in a variety of output modes, and in the case of incremental solutions, time history plots may be obtained showing how an element has reacted during a series of static and thermal loadings and unloadings, including plastic deformation of elements during a load increment (Figure 12). The collage presented as Figure 13 summarizes a typical graphics package currently available for two-dimensional analysis.

GRAPHICS PACKAGES FOR MIXED STRUCTURES

A mixed structure, being an assemblage of links, beams, membranes and plates, must be viewed graphically in three dimensions. A series of advanced general purpose codes in this category exists in the field of structural mechanics (References 2, 3 and 4), and these codes are capable of handling highly complex problems containing many thousands of unknowns, an example being the submarine hull launch tube model shown in Figure 14. Being three-dimensional, it is advantageous to examine the structure from a variety of viewpoints in order to clearly define a specific portion of a model in which the analyst is interested. LMSC's codes therefore have the ability to generate three orthographic views and, through a series of axes rotations, virtually any isometric view of the model. Again, as in two-dimensional packages, any portion of the structure may be expanded to clarify dense regions and similarly, element and nodal definition may be graphically represented on both partial and entire models (Figure 15).

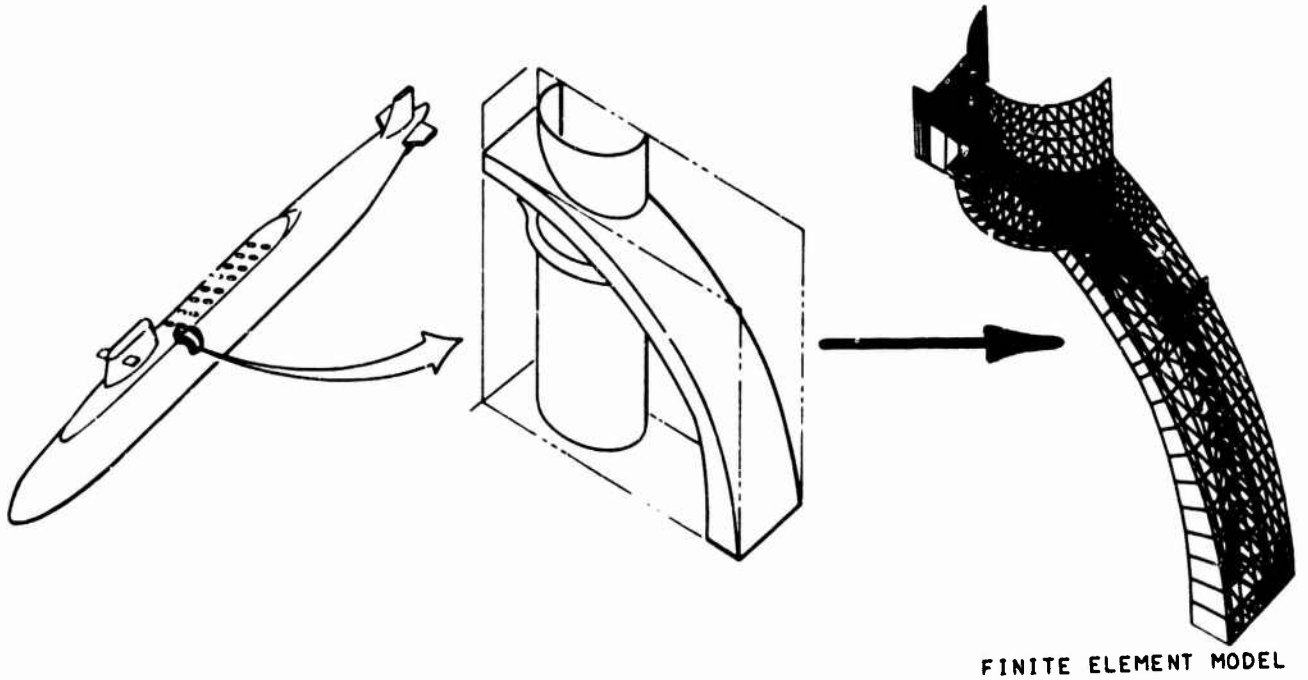


FIGURE 14

SUBMARINE HULL AND LAUNCH TUBE ANALYSIS

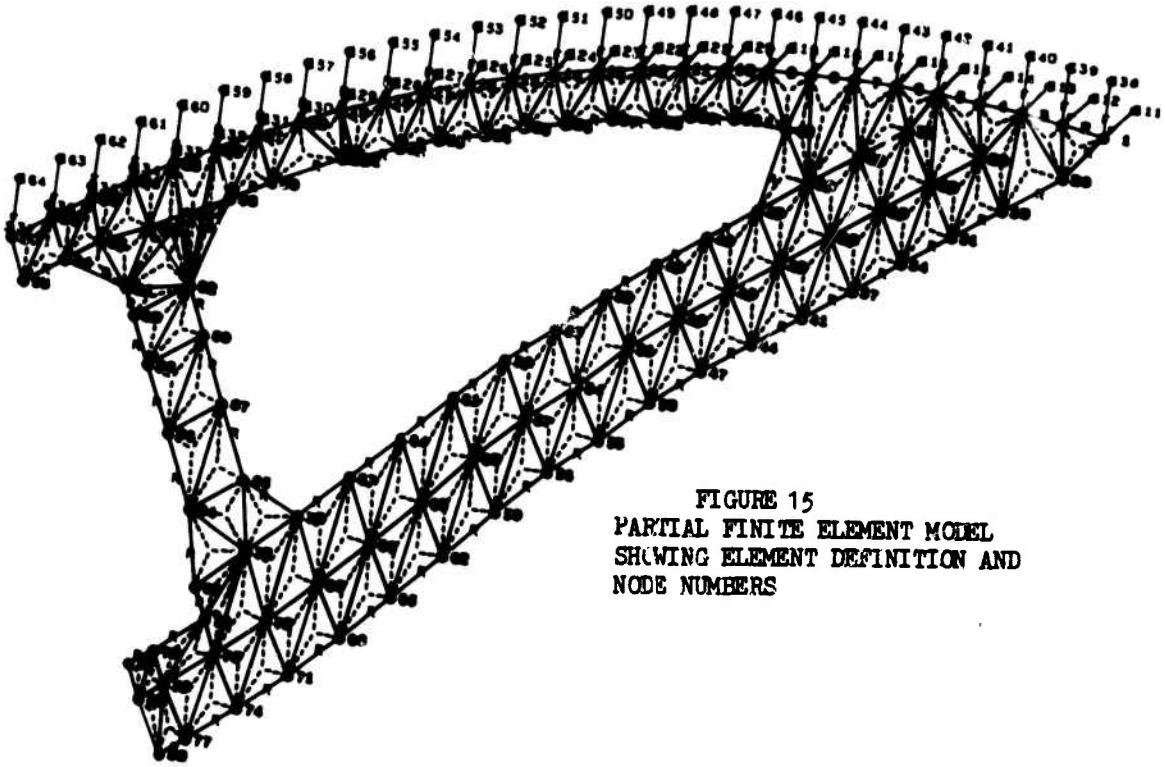


FIGURE 15
PARTIAL FINITE ELEMENT MODEL
SHOWING ELEMENT DEFINITION AND
NODE NUMBERS

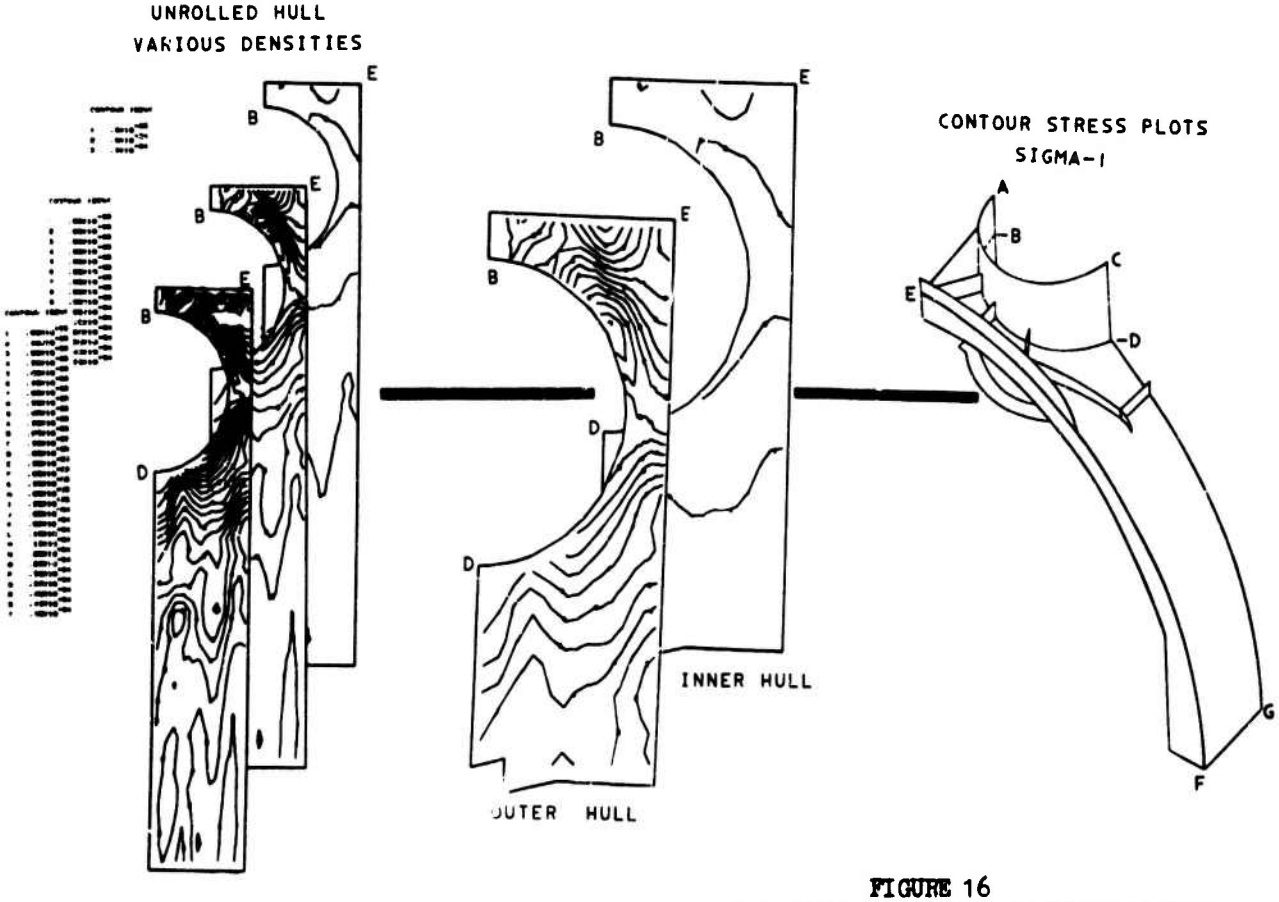
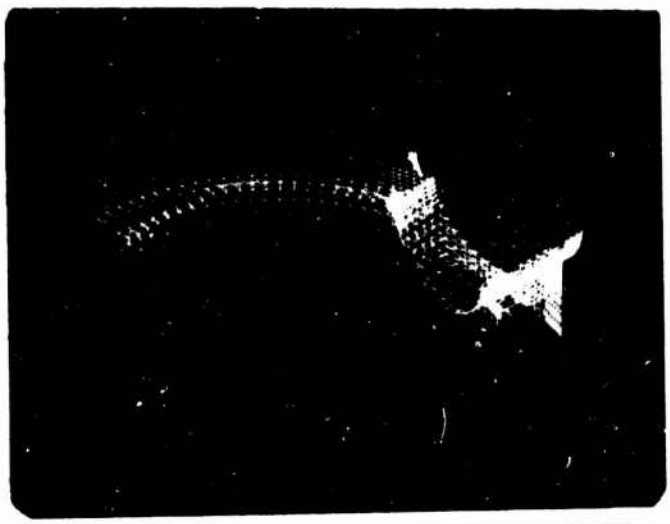


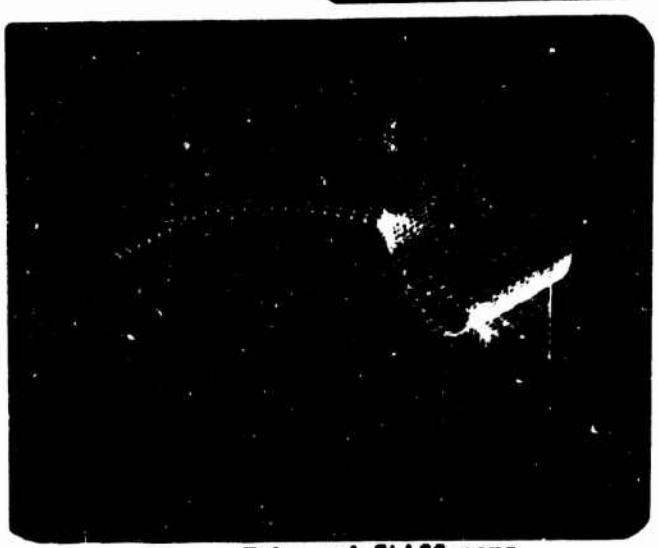
FIGURE 16
ISO-STRESS CONTOURS OF UNROLLED HULL

In mixed structures it is often necessary to overlay one element type on top of another, such as a beam along the edge of a plate, in order to approximate as closely as possible the actual structure being examined. The ability to separate each element type and plot them individually allows the user to check his model and insure its validity. The user may also generate any special combination of elements on a single plot. Thus a user can proceed with a model buildup, starting with a beam structure, next adding links, and finally generating the required number of plates and membranes. Such techniques almost eliminate the possibility of obtaining a solution containing possible hidden errors due to a faulty model.

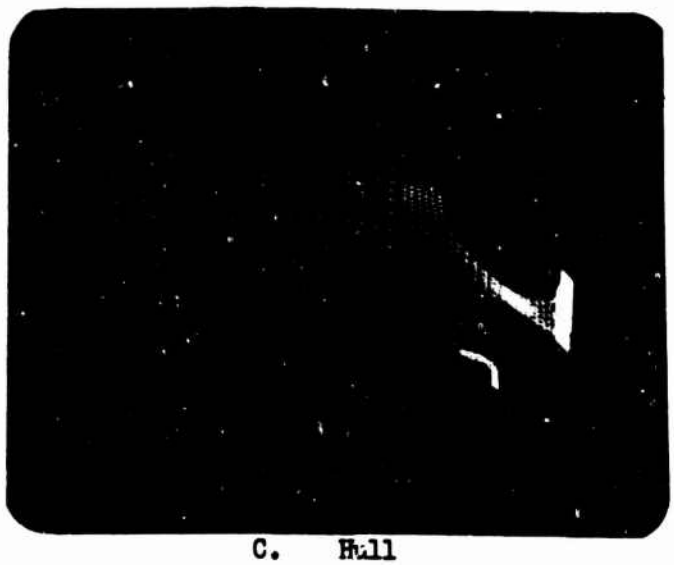
As in the case of two-dimensional analysis, models which are symmetric in nature may be "flip-flopped" to generate a larger portion of the structure, and models which are a combination of smaller components may be displayed as an exploded view giving more detail to each component individually (Figure 17).



A. Entire Model



B. Tube and Stiffeners



C. Hull

FIGURE 17
FINITE ELEMENT MODEL AND EXPLODED PARTIALS
DISPLAYED ON AN INTERACTIVE SCOPE

Structures being modeled using bending elements can be contoured, showing stresses and strains in both the local and global coordinate systems, both inside and outside the structure (Figure 16). If the surface of the model being contoured is curved, the program unrolls the surface bringing it into two dimensions for clarity of contour patterns. Again, the density of the contour lines, the limits of the contours and the contour increment are all controlled by the user. However, the program is automated to make a "best guess" on all of these quantities, if not overridden, thus saving the user from excess work, or having to "guess" at the magnitude of resultant stresses and strains prior to initial execution. The reduction of output is again aided through the use of deflected plots showing both actual and magnified deformations of the entire structure, any partial structure, or any element type. A deflected structure may be displayed either by itself or superimposed over the original undeflected shape (Figure 18). The user may also obtain a set of stress cuts along any series of elements through the structure plotting any output quantity generated. The collages presented as Figures 19 and 20 summarize graphics packages currently available for mixed structures.

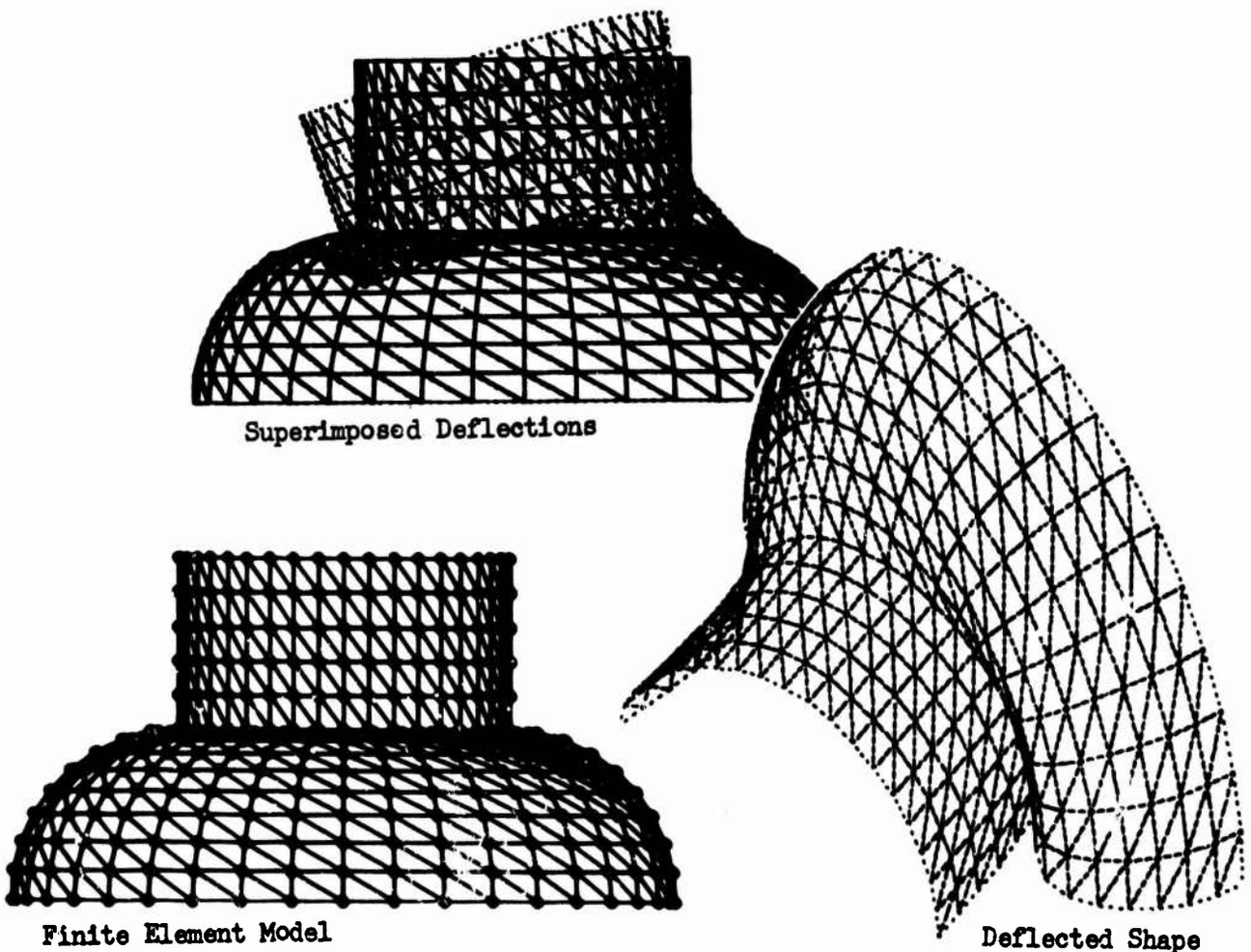


FIGURE 18

FINITE ELEMENT MODEL AND DEFLECTED SHAPE OF PRESSURE VESSEL DOME

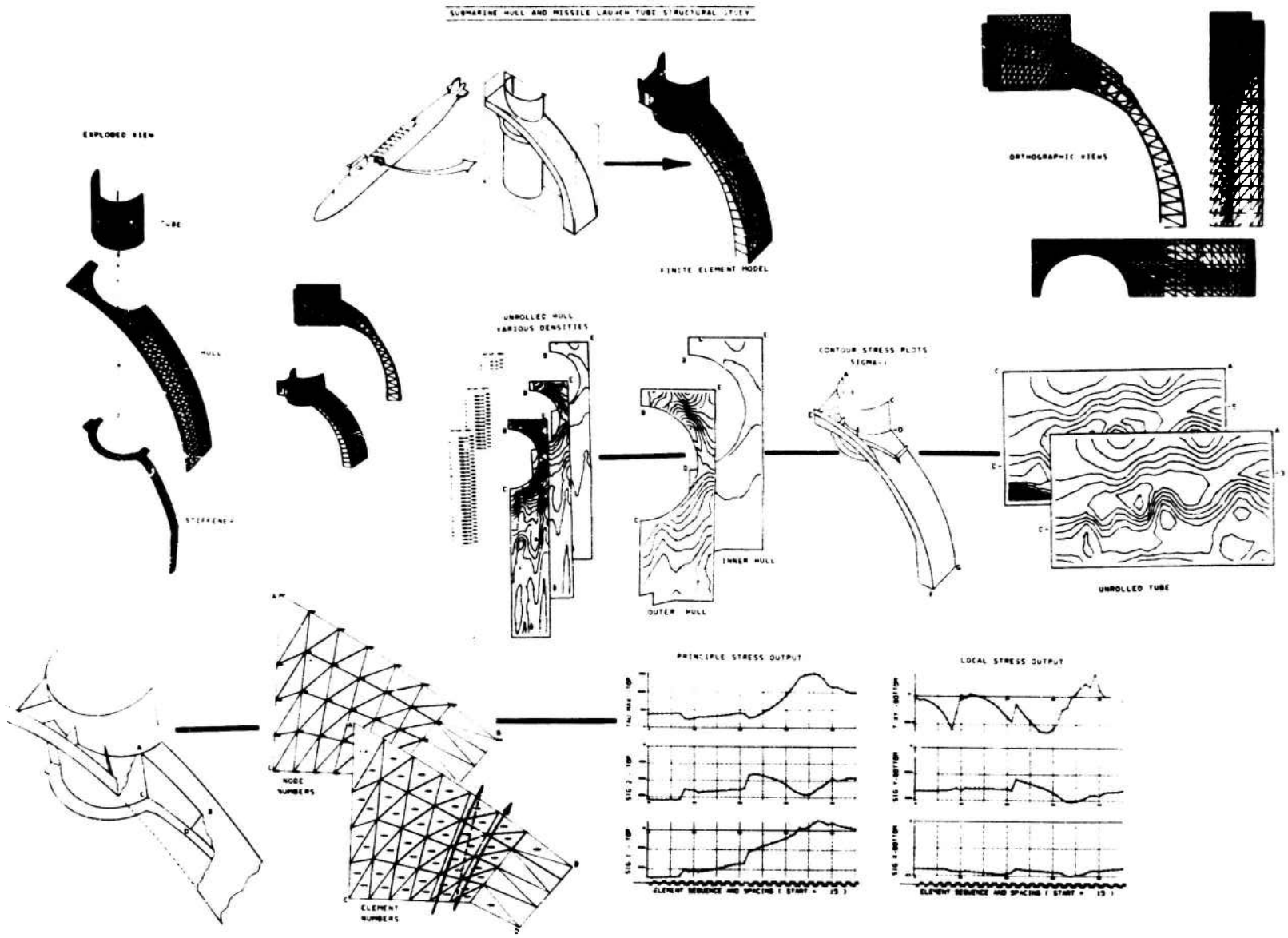
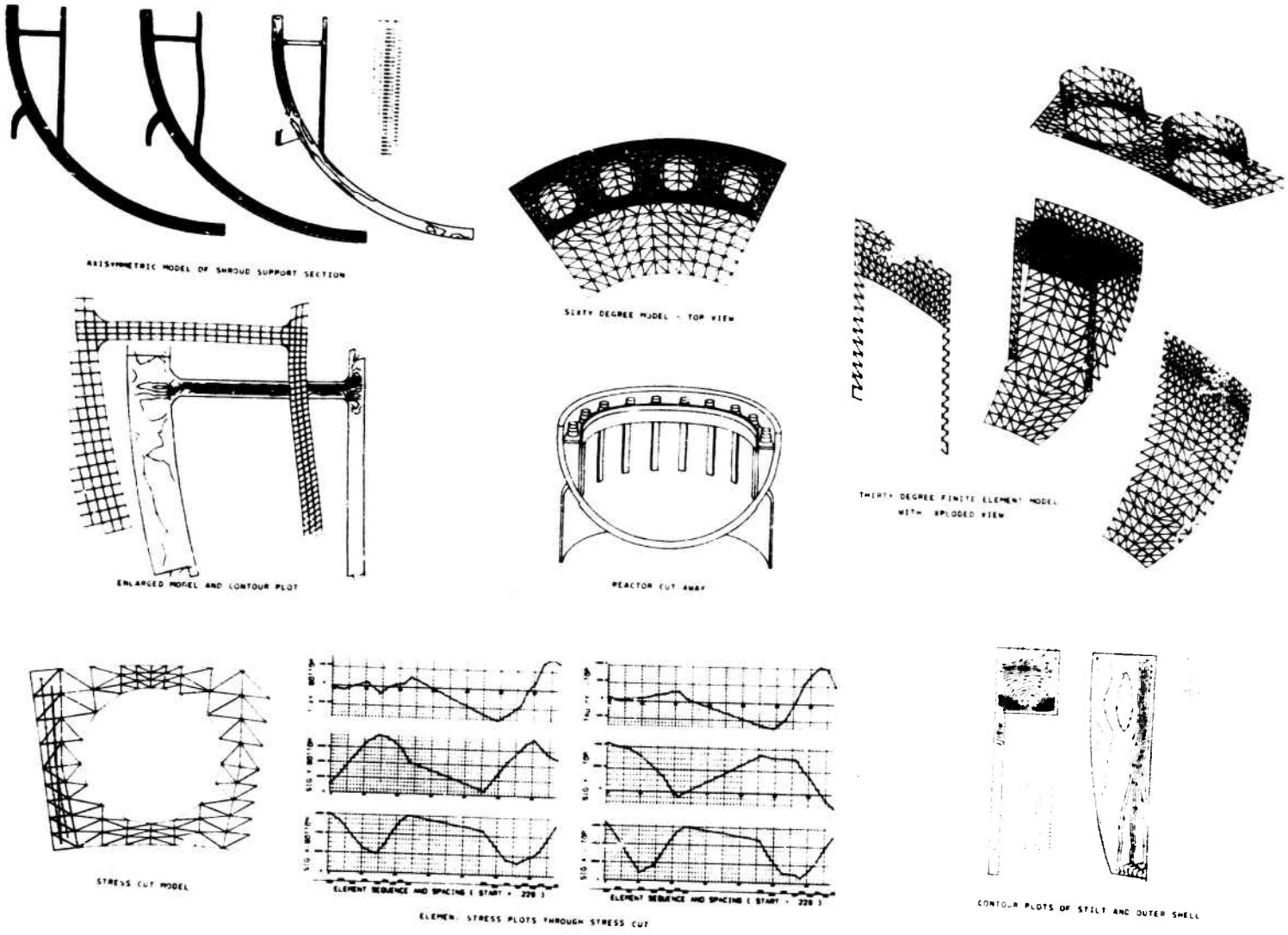


FIGURE 19
GRAPHICS PACKAGES AVAILABLE FOR
MIXED STRUCTURE FINITE ELEMENT MODELS



ELEMENT STRESS PLOTS THROUGH STRESS CUT

FIGURE 20
GRAPHICAL OUTPUT FOR REACTOR VESSEL ANALYSIS

GRAPHICS FOR ARBITRARY SOLIDS

The addition of solid elements into finite element codes (Reference 9) has further emphasized the need for more extensive display methods as discussed in References 10 and 11. Structures composed of such elements inherently contain large quantities of internal lines which tend to unduly complicate a plot of the entire structure (Figure 21). For purposes of clarity, packages have therefore been developed to show slices of elements or nodes (Figure 22) distinct from the rest of the model, with or without element and node definitions (Figure 23), with each slice being able to be rotated to any desired viewing angle. Similarly, deflected plots of portions of the structure may be displayed showing actual or magnified deflections. Such deflected figures can be displayed individually or superimposed over the undeflected structure (Figure 24). The viewer can, therefore, display any block of elements, any plane of elements through the structure, or any surface of elements. Further graphical displays allow stress and strain cuts through the structure and plotting output quantities for any element faces in addition to element centroids. Defining the stress cut visually is greatly aided by displaying the elements being graphed in reference to a plane of nodes, thus orienting the stress cut within the entire model as shown in Figure 25.

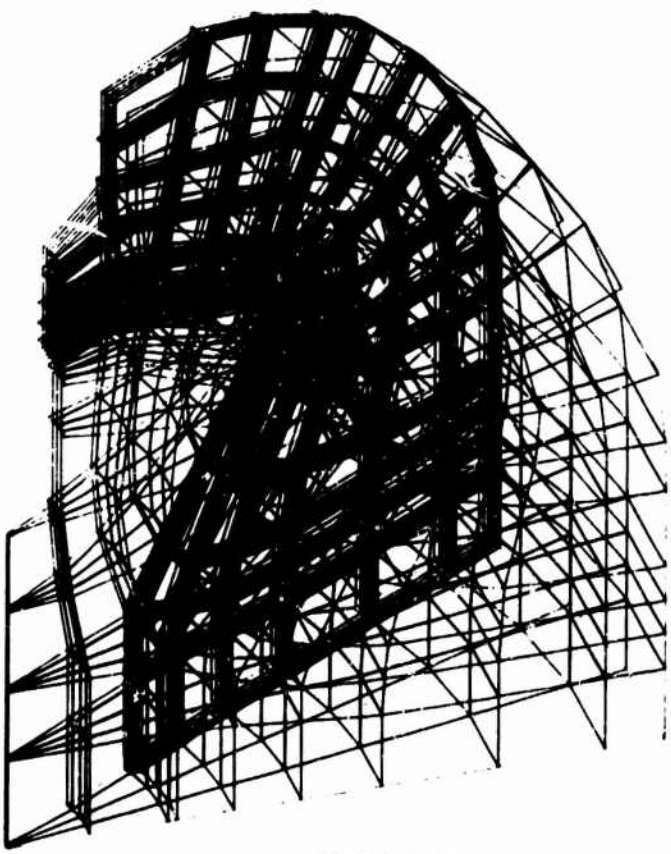


FIGURE 21

ENTIRE SOLID FINITE ELEMENT MODEL

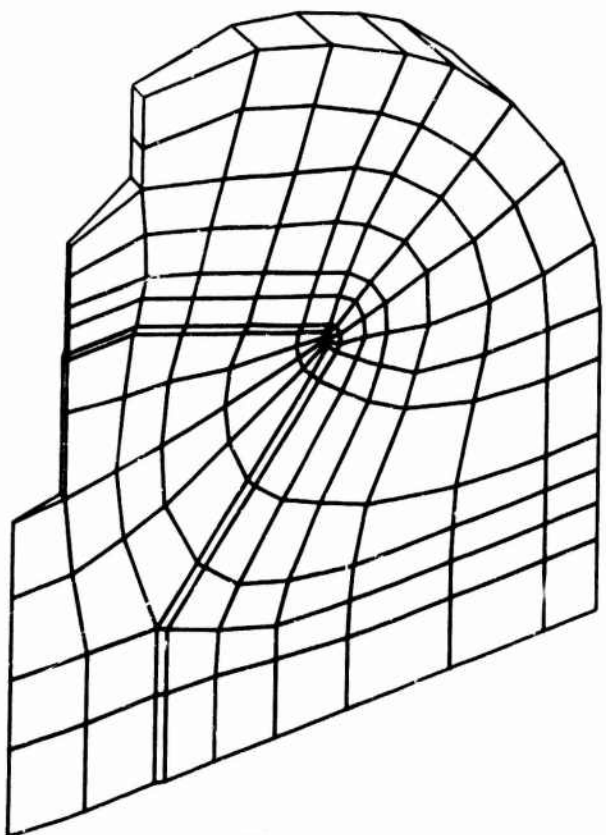


FIGURE 22

SINGLE LAYER OF ELEMENTS

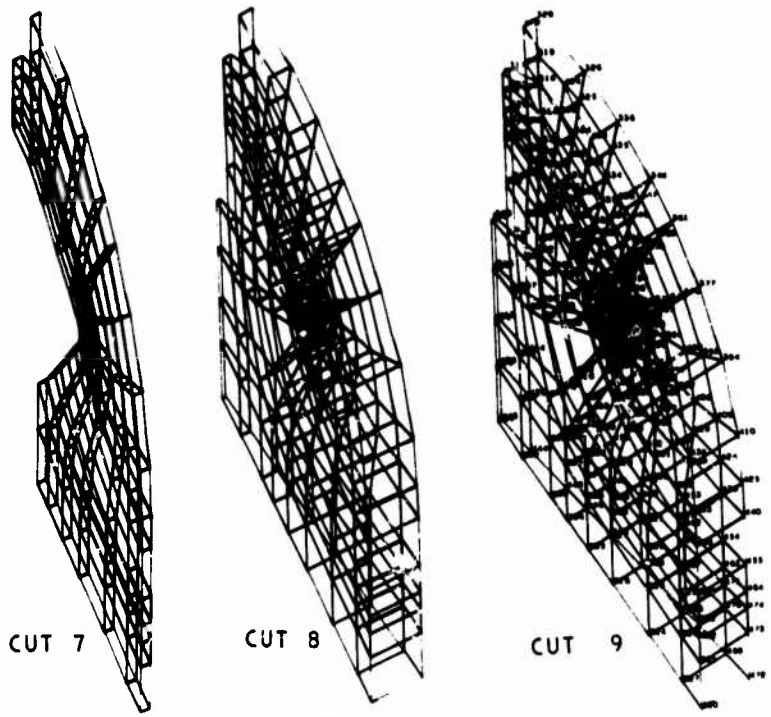


FIGURE 23
PARTIAL MODEL SHOWING SINGLE
PLANE OF ELEMENTS

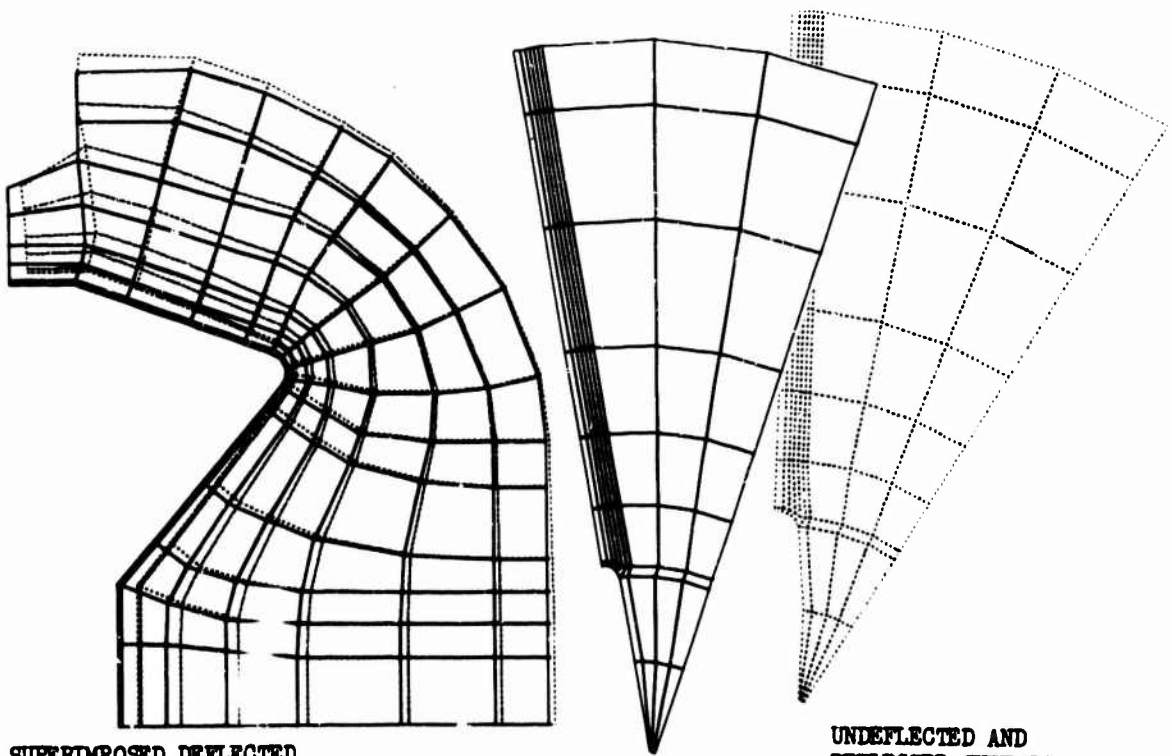


FIGURE 24
DEFLECTED STRUCTURES FOR ARBITRARY SOLIDS

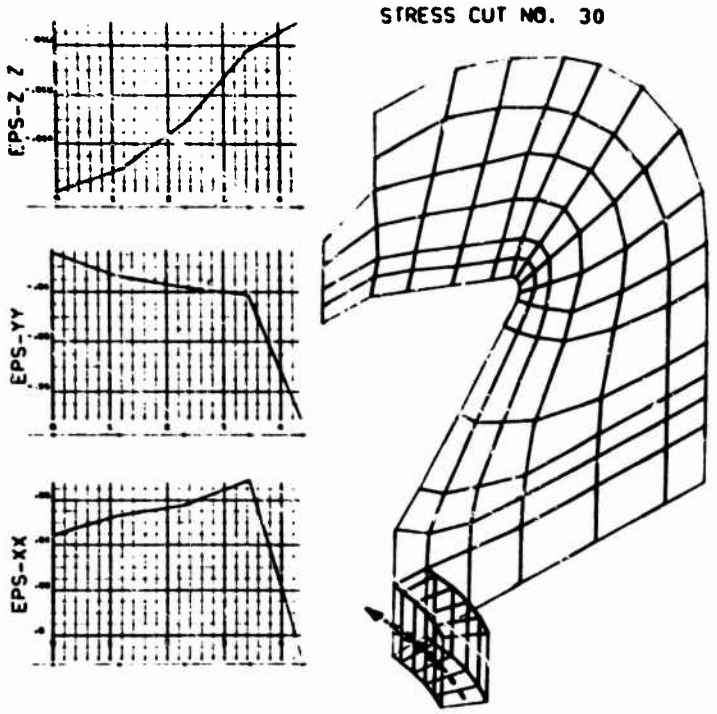


FIGURE 25
OUTPUT STRESSES AND STRAINS ARE GRAPHED
ALONG ANY SERIES OF ELEMENTS

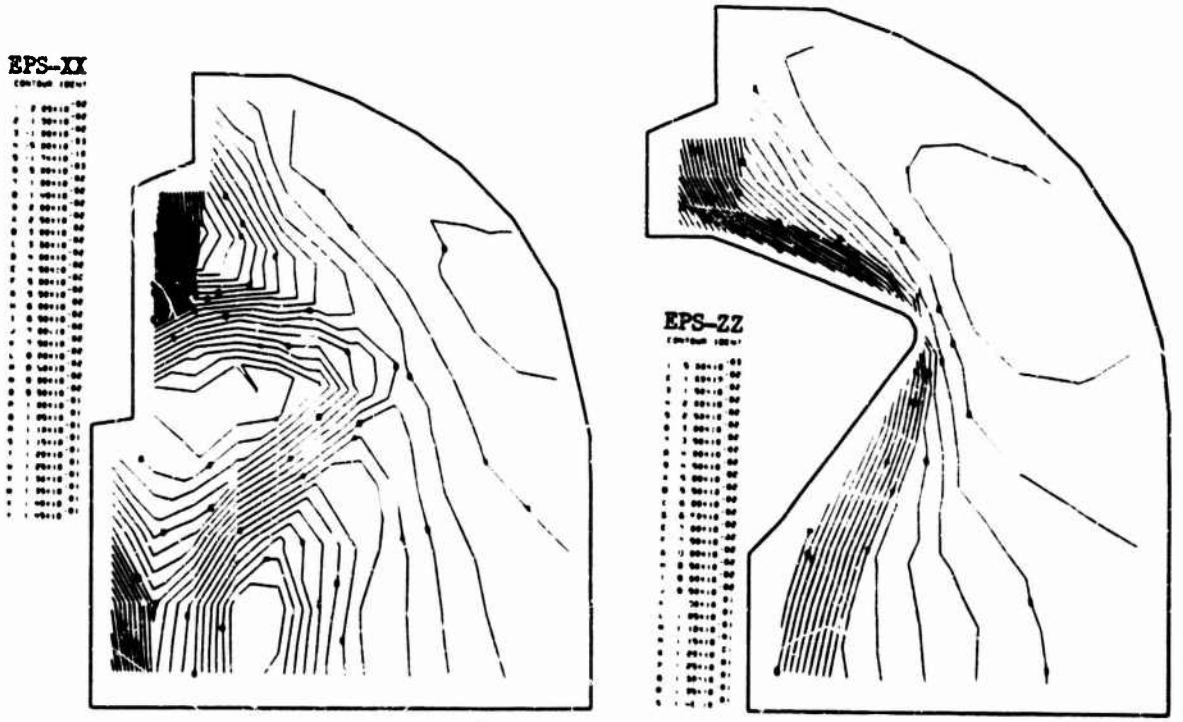


FIGURE 26
OUTPUT STRESSES AND STRAINS ARE CONTOURED
FOR A GIVEN PLANE OF ELEMENT FACES

Contour plotting in three dimensional structures can be performed on any surface of the model or on any planar slice of element faces through the model (Figure 26). In any view, output quantities for element faces as well as element centroids may be contoured to any desired degree of density.

The problem of generating a picture with hidden lines removed for any arbitrary three-dimensional structure has been solved using several different methods (References 12, 13 and 14), but the computer time and cost of generating such a picture for a structure made up of over 1900 nodes and 6300 planar surfaces cannot always be justified by the value of the final output plot. This problem is one of current interest and hidden line programs are now being tested, optimized and incorporated into both interactive and non-interactive production codes. However, through the use of graphics packages already existing in production codes the user can quickly generate a "pseudo" hidden line picture by displaying only a portion of the total model's element faces and edges (Figure 27).

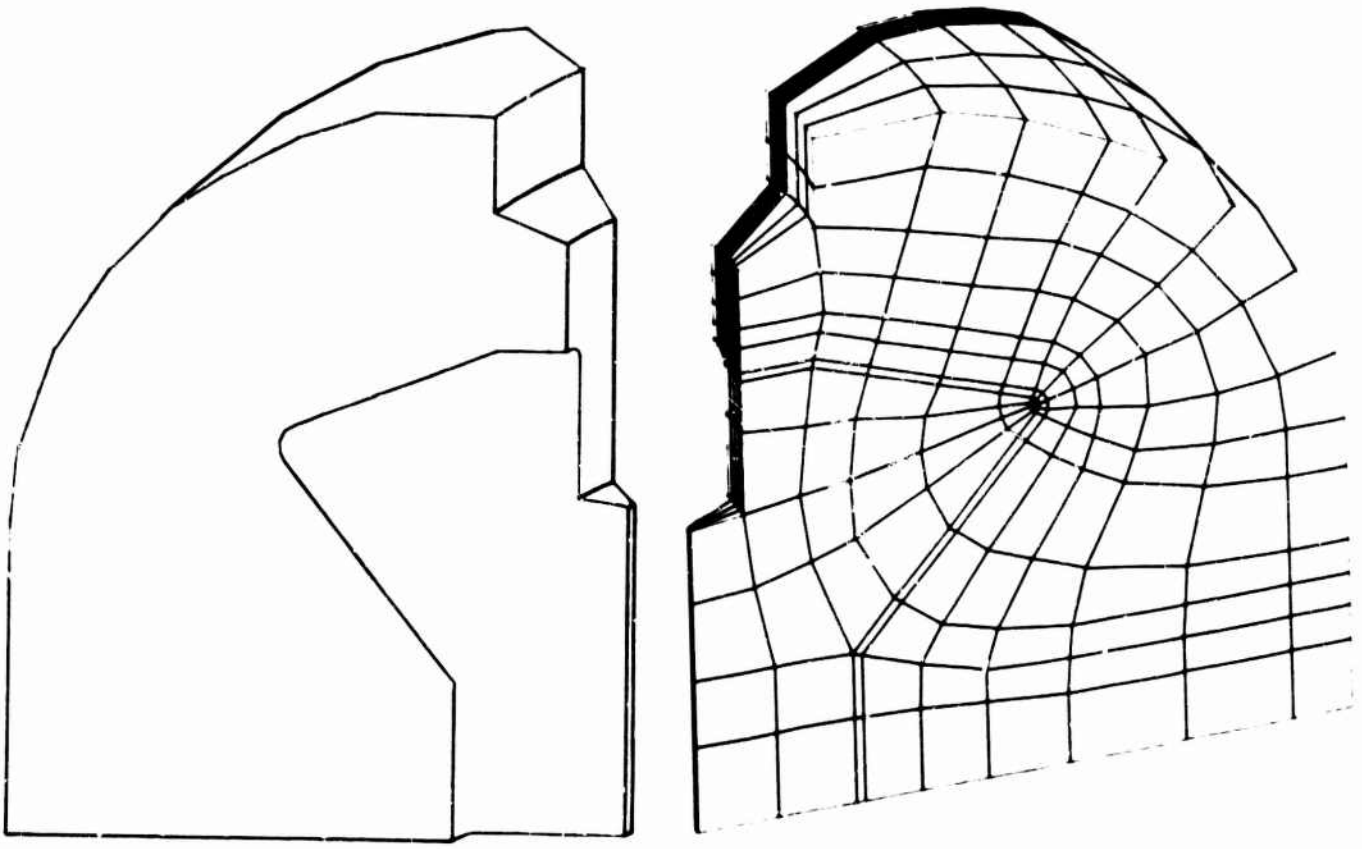


FIGURE 27

'PSEUDO' HIDDEN LINE PICTURES

SECTION III
INTERACTIVE GRAPHICS

INTERACTIVE HARDWARE

Interactive computer graphics, employing a cathode-ray-tube display terminal attached to a digital computer (Figure 28), allows the user to communicate with, and control the flow of, his own program in real time. The development of interactive graphics from a research analyst's aid to a production analysis tool has enabled the engineer to radically reduce the time span and overall cost of a given task from the levels previously required in a batch process mode. The magnitude of this savings is illustrated in Figure 29. The time necessary for a series of computer runs submitted under a batch processing mode is, of course, highly dependent on the job turnaround time of the system. Hence, while input tasks and model debugging may take days under the batch system, the same series of jobs can be completed during a single session at an interactive scope.

FIGURE 28

COMPUTEK 400 INTERACTIVE GRAPHICS TERMINAL

The heart of the interactive graphics system currently in use at LMSC is the Computek Series 400 Interactive Graphics Terminal (Reference 15) which is direct-coupled with a CDC 6400 Digital Computer and complimented by an on-line hard copy generator, joystick, digital graphics tablet, and DIS211 scope (Figure 29). The Computek is a storage tube type display retaining information it has received, as opposed to refreshed CRTs such as the IBM 2250 and the CDC 274 which are updated about 30 times per second from a memory device. Although response time is slower on a storage tube and partial screen erasing is not possible, this type of display is an order of magnitude less expensive than the refreshed type tube, while, at the same time, being more than sufficient to handle the size and type of work necessary to accompany very large finite element codes (Reference 12). In structural analysis the interactive scope is currently used to debug finite element models prior to problem solution and also as a means for selective data reduction at task completion. The solution of the problem, however, can be run on any machine, thus eliminating any machine dependence on the computer to which the CRT is coupled. It has been shown to be economically advantageous to do interactive model debugging on a smaller inexpensive computer in a time-sharing mode, subsequently running the computer model for analysis results on a machine with large core memory and mass storage devices which can handle the size and complexities of a model containing more than 6500 unknowns and finally returning to the smaller interactive computer with analysis results for data reduction and report preparation.

INTERACTIVE METHODS

The Computek, as a time-sharing terminal, allows the user to substantially compress the time scale required to accomplish a particular analysis goal due to improved response time over conventional batch processing techniques.

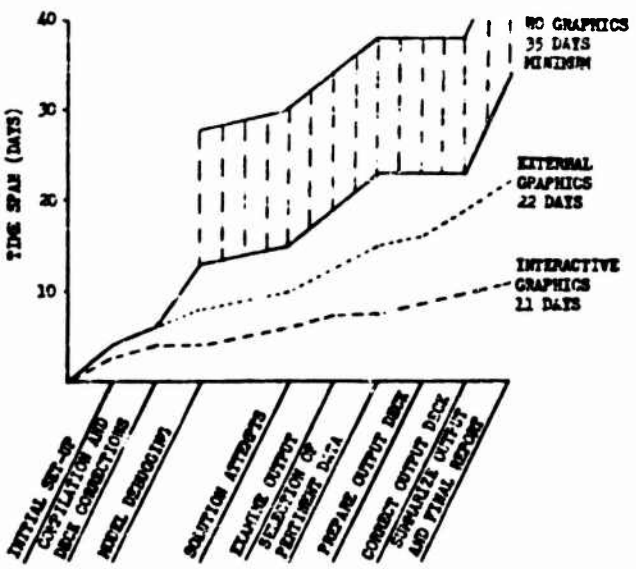


FIGURE 29

USING GRAPHICS TO REDUCE TASK TIME SPAN

In the preprocessing mode, an initial finite element model is generated using modular programming techniques and data input cards. This data is passed to the display program. The engineer, sitting at the scope terminal, has a variety of options available to him to alter and debug his model. Hence, he can display his model wholly or partially in any given view, add new nodal points and elements, change existing data, display on the screen any of the storage tables generated, and finally, recycle using these modifications to generate a corrected model (Figure 31). This pattern is repeated until the model is completely debugged, then a card deck is punched containing the model definition in a compressed form, and this punched output is fed to a larger computer to obtain analytical results. Interactive programs are currently being developed which would allow the user to reduce these results through the use of deflected pictures, output graphs, and contour plots, and thus generate a final analytical report in real time and by total reliance on computer graphics packages.

At any time during the interactive mode the user can obtain a hard copy print of the contents of the Computek screen, or store this data on a tape in a format accepted by the SC4020 and FR80 external plotters in order to obtain copies with a higher degree of resolution. Just as output from the Computek can be displayed by an FR80, so also can a tape generated for the FR80 be displayed on the Computek giving the user the ability to edit and delete results before final processing.

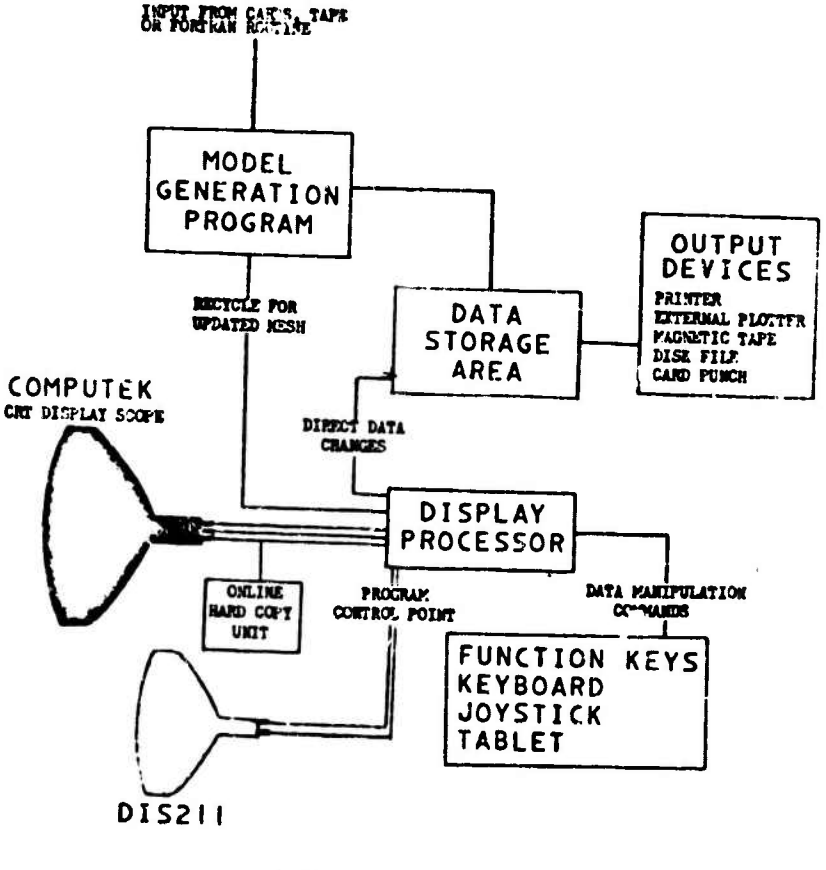


FIGURE 30

INTERACTIVE PROGRAM DISPLAY SYSTEM

A. Command Table Display

B. Data Table Display

C. Input Mesh Display

D. Finite Element Model

E. Element Number Display

FIGURE 31
INTERACTIVE DISPLAYS

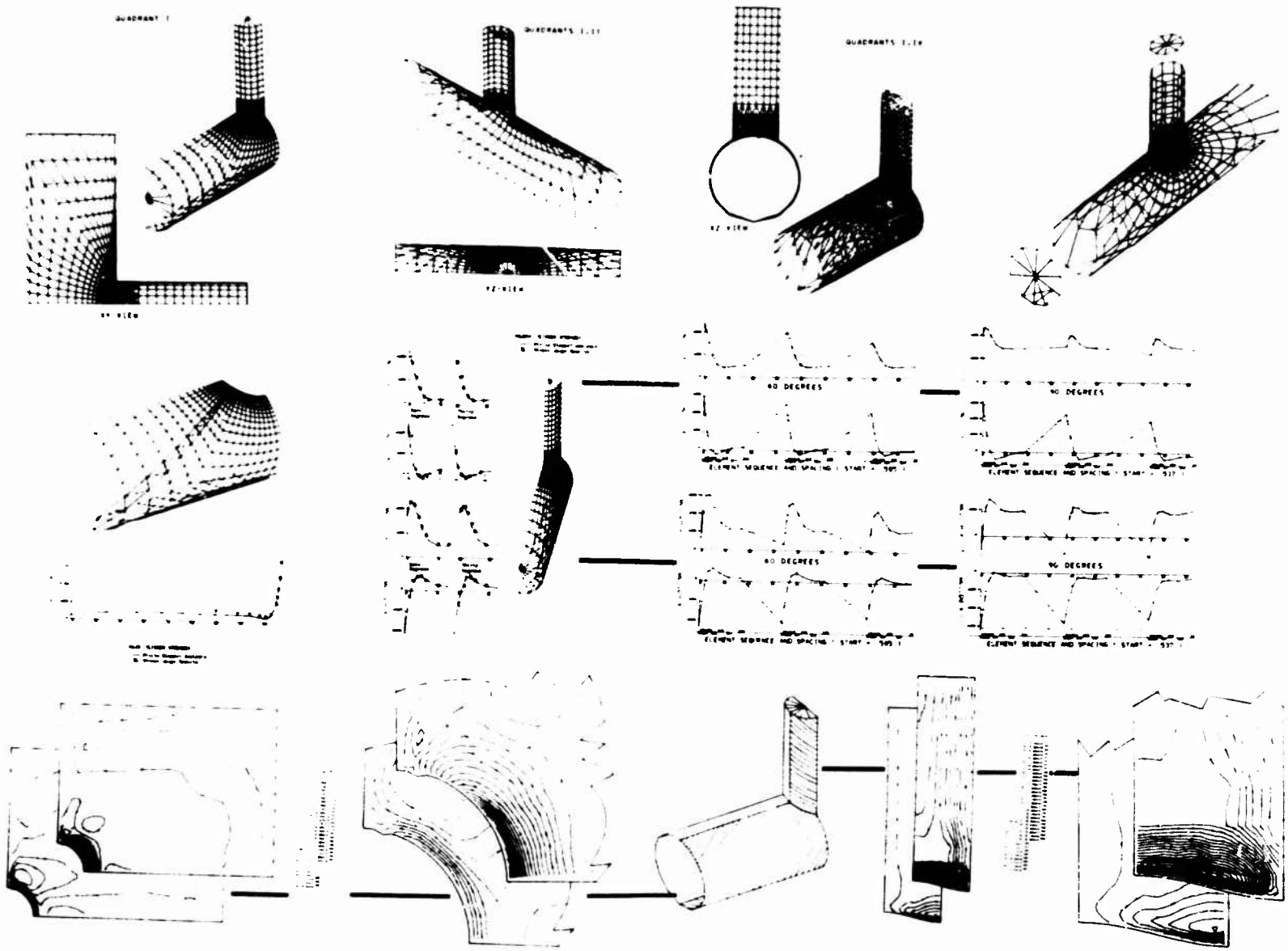
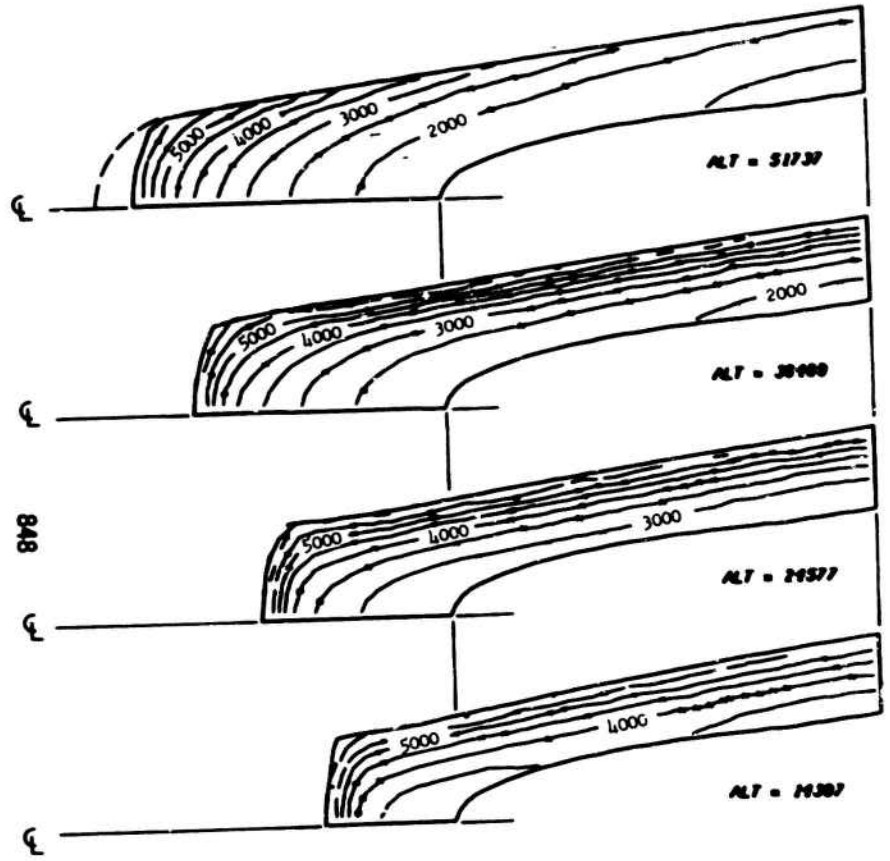
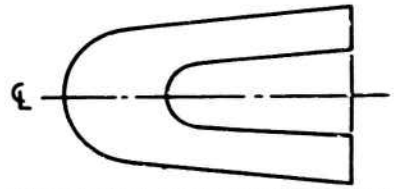


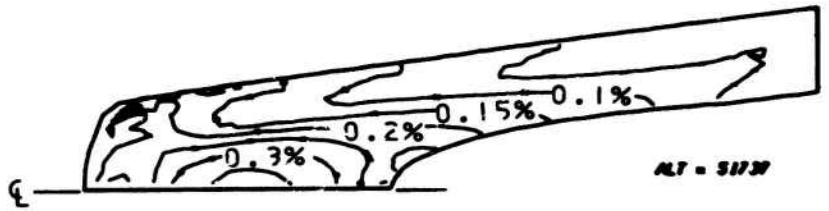
FIGURE 32
GRAPHICAL OUTPUT FOR PIPING TEE ANALYSIS



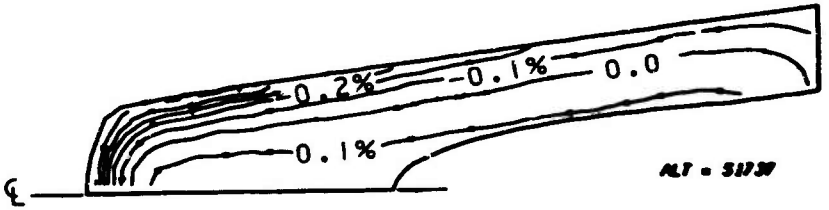
ERODED SHAPE AND ISOTHERMS



NOSETIP CONFIGURATION



CONTOUR MAP OF
MAXIMUM STRAIN



CONTOUR MAP OF
CIRCUMFERENTIAL
STRAIN

FIGURE 33
REENTRY VEHICLE NOSETIP
ERODED SHAPE AND CONTOUR MAPS

SECTION IV
PRODUCTION GRAPHICS

To demonstrate the utility of graphical tools and to emphasize the extent to which such schemes are used, a short discussion of some major analytical tasks that have recently been completed is offered here.

Figure 32 shows a three dimensional analysis of a pipe tee that was performed using the "DAISY" code. The model consisted of 886 nodes, 872 quadrilateral bending elements, 4242 unknowns, and was analyzed for five mechanical loading conditions including internal pressure. Contours of surface stresses, deflected plots, and graphs of stresses along preselected lines of the structure were used extensively. The entire reporting task required less than a week and consisted primarily of assembling graphical computer printout onto larger report format sheets.

A further value to be obtained from graphical output is shown in Figure 33, an output from a re-entry vehicle analysis. This figure shows the automatic model geometric changes as the model ablates and demonstrates the variation in the structure with altitude that is used as thermal load input. The structural model is overlayed directly onto a thermal model which has been run previously to determine thermal loads at various altitudes. The entire analysis from the time the thermal loads are obtained is accomplished in 24 hours, including stress, strain, temperature and geometric histories. All graphics are automatically prepared and output directly in a report format.

Another example of the use of computer graphics as a data reduction tool for complex structural analysis is shown in Figure 34. The problem, an advanced elastic-plastic incremental analysis of an aircraft disc brake component, required the calculation and display of stresses and strains at any point during multiple thermal cycling. The collage summarizes the computer graphics used in the presentation of output as an illustration of the ease of visual assimilation of results. This problem was quite complex and the extensive use of graphical output added immensely to the engineer "feel" and understanding of the problems.

Two final examples demonstrating the ability of graphics to bring complex finite element analysis to full production fruition which can be cited are: a thru-bulkhead initiator subjected to severe plastic deformations was modeled, analyzed and visual reporting made in less than one day; a diode subjected to four thermal load cases resulting in plastic deformation was analyzed and graphically presented in two days, the model being generated and debugged using interactive graphical procedures.

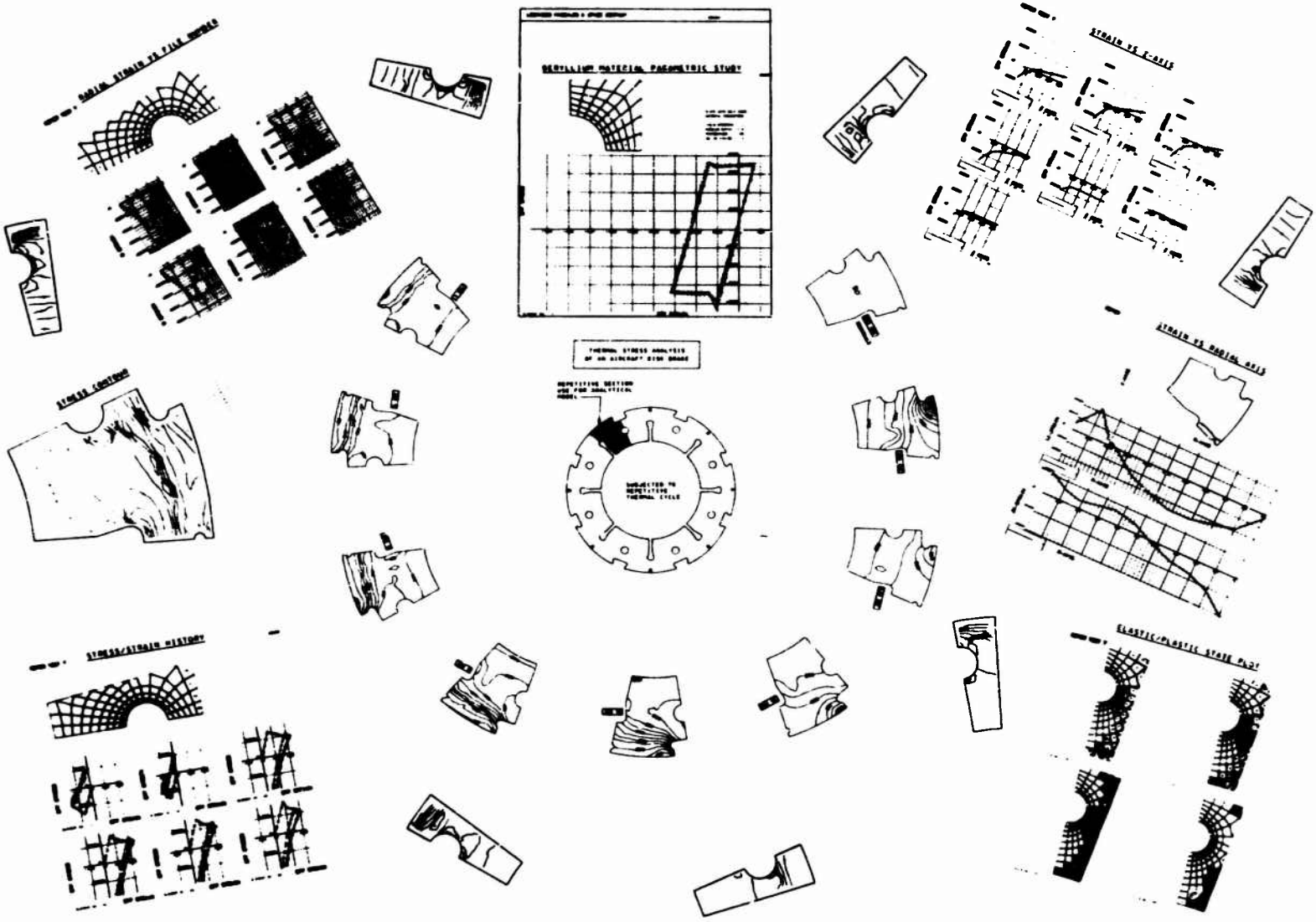


FIGURE 34.
GRAPHICAL OUTPUT FOR DISC BRAKE ANALYSIS

SECTION V
SUMMARY

While a widespread effort on all aspects of finite element development is being pursued at Lockheed including; non-linear analysis of general structures, incompressible arbitrary solids, thermal fatigue and creep, dynamics, buckling, super-elements, and high-speed equation solvers, we believe that the major factor common to, and necessary for, the implementation of all such developments into a usable production mode of operation is the availability of an extensive range of graphics packages. Such packages obviate the necessity for poring over many hundreds of pages of printed output, enables rapid checking of results and input data, and, finally, allows large scale, complex analysis tasks to be successfully completed in a minimum of time and with a minimum cost. Therefore, in addition to the rapid development of an extensive library of finite element codes we are firmly committed to the concept of advanced graphics packages, for both the interactive and batch modes, to enable such programs to truly become viable tools for large scale structural analysis problems.

REFERENCES

1. Wilson, E. L., and Jones, R. M., "Finite Element Stress Analysis of Axisymmetric Solids with Orthotropic Temperature Dependent Material Properties," Air Force Report No. BSD-TR-67-228, September 1967.

2. Kamel, H. A., "Mini-Aska Users Manual and Documentation," Research performed for Lockheed Missiles and Space Company under Contract No. POAPZ8J928OA, May 1968.

3. Loden, W., 'REXBAT 5 Users Manual" Unpublished Report, Lockheed Missiles and Space Company, Sunnyvale, California, July 1970.

4. Wilson, E. L. "SAP - A General Structural Analysis Program" UCSESM 70-20, Department of Civil Engineering, University of California, Berkeley, California, September 1970.

5. McCue, G.A., and Du Prie, H. J., "Improved Fortran IV Function Contouring Program", North American Aviation, SID 65-672, Space Sciences Laboratory, April 1965.

6. Yates, D. N., Vinson, T. J., and Sable, W. W., "The Development of Large Scale Digital Computer Codes for Production Structural Analysis", Computer-oriented Analysis of Shell Structures Conference, Palo Alto, August 1970.

7. Lambert, R. L., "Propellant Grain Three-Dimensional Strain Analysis", Lockheed Missiles and Space Company, July 1971.

8. Inman, J. S., "Evaluation Report of COM Equipment to Replace the Stromberg Data Graphix SD 4020", Unpublished Report, Lockheed Missiles and Space Company, Sunnyvale, California, December 1969.

9. Dovey, H. H., "Three Dimensional Solids Program: 8 Nodal Brick", Department of Civil Engineering, University of California, Berkeley, December 1969.

10. Sutherland, I. E., "Computer Displays", Scientific American, Vol. 222, No. 6, Pg. 56-81, June 1970.

11. Roberts, L. G., "Machine Perception of Three-Dimensional Solids", Technical Report No. 315, Lincoln Laboratory, M. I. T., Cambridge, Massachusetts, May 1963.

12. Warnock, J., "A Hidden Surface Algorithm for Computer Generated Halftone Pictures", Technical Report 4-15, Computer Science, University of Utah, Salt Lake City, Utah, June 1969.

13. Loutrel, P. P., "A Solution to the Hidden-Line Problem for Computer-Drawn Polyhedra", IEEE Transactions on Computers, C-19 3, 205, March 1970.

14. Watkins, G. S., "A Real Time Visible Surface Algorithm", UTECH-CSC-70-101, Computer Science, University of Utah, Salt Lake City, Utah, June 1970.

15. Fitzpatrick, C. E., "Users Manual - Computer Interactive Display Terminal", Unpublished Report, Lockheed Missiles and Space Company, Sunnyvale, California, January 1971.

SESSION 7. NON–LINEAR EFFECTS

Session Chairman

Colonel C. K. Grimes*

Air Force Flight Dynamics Laboratory
Wright-Patterson Air Force Base, Ohio

*retired from the Air Force, now with the Boeing Co., Wichita, Kansas

A PROCEDURE FOR FINITE ELEMENT PLATE AND SHELL PRE- AND POST-BUCKLING ANALYSIS

by Richard H. Gallagher,* S. Lien,** and S. T. Mau+

Cornell University

A procedure for finite element analysis of geometrically nonlinear problems, extending over the pre-buckling and initial post-buckling regimes, snap-through buckling, and accounting for initial imperfections, is described. The computation of the nonlinear pre-buckling path is accomplished by direct iterative solution. The bifurcation point is established by interpolation of solution points of the pre-buckling and immediate post-buckling analyses. A static perturbation method is then developed for determination of the post-buckling path of the bifurcating structure or the limit point of a structure with initial imperfections. Three numerical examples, involving an arch, flat plate and shallow shell, are presented in illustration of the procedure and in comparison with alternative approaches.

INTRODUCTION

The analysis of instability phenomena of complicated thin shells has drawn intensified interest due, in part, to the development of finite element analysis procedures for such structures.[1] Structures of this class may collapse at load levels which are less than those predicted by linear instability theory because of the role played by initial imperfections and geometric nonlinearities. The extensive efforts in the development of theories to cope with this problem have been surveyed by Hutchinson and Koiter,[2] Haftka, et al[3] and Bienek[4].

Although the various types of instability phenomena which might occur in the complete range of load-displacement behavior

* Professor and Chairman, Department of Structural Engineering

** Presently, Research Engineer, Westinghouse R&D Center, Pittsburgh, Pa.

+ Presently, Research Engineer, MIT Aeroelastic and Structures Research Lab.

prior to final collapse are not as yet fully understood, certain forms are known and are of considerable practical importance, especially those which occur in the earliest stages of loading. Figure 1a applies to "perfect" structures and represents the case in which the structure first displaces along the path defined by OAB (the fundamental path) and bifurcates (or branches) at the Point A to another path, AC. In contrast to a rising post-buckling path, as AC, a descending path AD (as pictured in Figure 1b) may be encountered.

When the structure possesses fabricational imperfections, the load-displacement behavior follows the paths indicated by dotted lines. The structure with a rising post-buckling path will have strength exceeding the bifurcation load. The strength of an imperfect structure with a descending post-buckling path in the perfect state will not achieve strengths as high as the bifurcation load unless the load-displacement path again rises at larger displacements. Such structures, under the appropriate load condition, are termed "imperfection sensitive" and the maximum load attained (Point E) is termed the "limit point".

A non-bifurcating load-displacement behavior may also occur for a structure assumed to be devoid of imperfections and may take the form similar in shape to the curve OE (Figure 1b) of the imperfection-sensitive structure. For this case the buckling phenomenon is of the 'snap-through' type.

A landmark development of procedures for establishing the shape of the post-buckling path and for determining the limit point for imperfection-sensitive structures is due to Koiter.[5] This approach uses the concept of perturbations from the bifurcation point and enables an efficient definition of load-displacement behavior in the immediate post-buckling range. Further contributions or alternative forms of these concepts have been presented by Budiansky and Hutchinson[6], Sewell[7], and Thompson[8,9].

Extensions of Koiter's procedure to the format of finite element analysis, as well as other finite element approaches to the same physical problem, have recently appeared [10-12]. Morin[10] applies a predictor-corrector scheme in calculation of

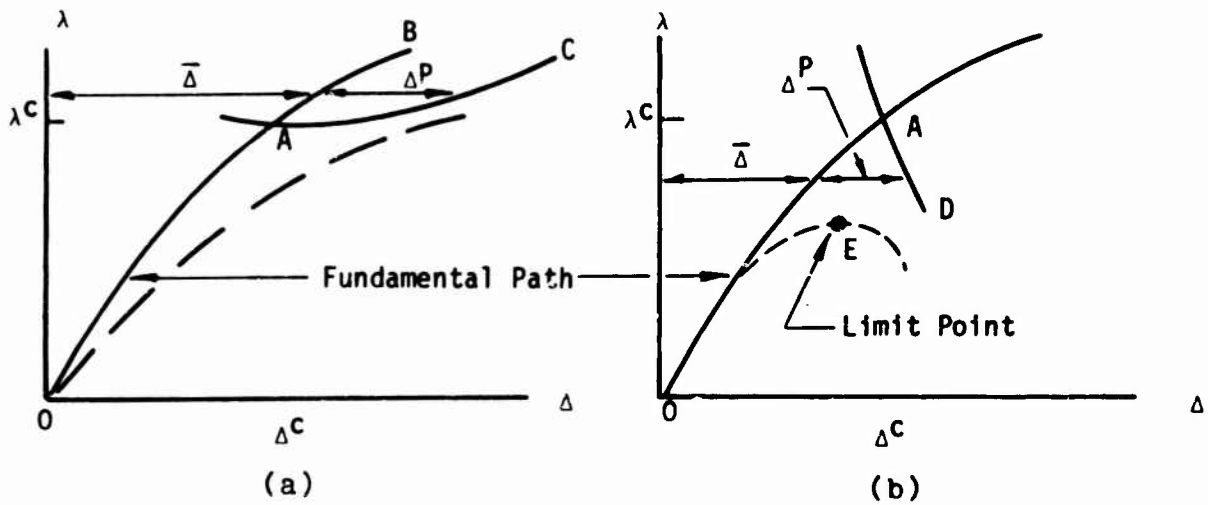


FIGURE 1. FORMS OF ELASTIC INSTABILITY

non-linear prebuckling behavior, in which a perturbation approach is employed as the predictor and Newton-Raphson iteration is employed as the corrector. The perturbation approach, in both the pre- and post-buckling computational phases, reflects earlier work by Thompson[8]. Thompson[9] has also advocated a new perturbation approach. Haftka, et al[3] propose the definition of an "equivalent structure", one in which the nonlinear terms are treated as initial imperfections, in order to exploit the concepts derived by Koiter for imperfect structures. Dupuis, et al[11], attack the solution of the nonlinear equations in an incremental-iterative manner. The work by Lang[12] is a direct adaptation of Koiter's concepts, including retention of the condition of a linear prebuckling state.

Recent analyses have shown[13] that the assumption of a linear prebuckling state may lead to inaccurate results. One of the principal aspects of the work described in this paper is the method of determination of the load and displacement state on the fundamental path, at the bifurcation point, following upon a nonlinear prebuckling state. The calculated information furnishes the necessary ingredients for a perturbation analysis of the post-buckling or limit point behavior.

The starting point of the present small strain-finite displacement formulation is the definition of element stiffness equations in the Lagrangian frame of reference. The element stiffness matrices extend to both first- and second-degree geometric nonlinearities in the element displacement parameters. Then, direct iteration is used for solution of the nonlinear algebraic equations in the prebuckling range. Unlike many widely used and seemingly more efficient procedures, direct iteration permits calculation of the fundamental path beyond the bifurcation point and definition of the latter by interpolation of the determinants of such solutions through the zero point.

For post-buckling, and for snap-through buckling for initial-imperfection situations, both displacements and loads are expanded about the bifurcation point of the perfect structure in power series in a single parameter which is related to the amplitude of the eigen-function in the deflected shape of the structure.

Upon determination of the series coefficients, the solution is a parametric representation of load vs. displacement.

Three problems are solved in verification of the present procedure and for the purpose of comparison with other methods. The first problem consists of a uniformly loaded shallow circular arch. This case, for which an exact solution is available, illustrates the calculation of the bifurcation point following upon a nonlinear fundamental path and demonstrates prediction of behavior for both perfect and imperfect structural forms. The second problem concerns the familiar case of post-buckling behavior of a perfectly flat, simply-supported rectangular plate under uniaxial compression. Finally, a solution is obtained for a hypothetical cylindrical shell structure which has been employed as the basis for numerical verification of alternative analysis methods.

A. ELEMENT AND SYSTEM FORMULATIONS

The purpose of this section is to define the general algebraic form of finite element stiffness equations for the present geometrically nonlinear analysis. This form, which applies to any given type of element, is subsequently employed in the development of procedures for pre- and post-buckling analysis of the complete structure. Presentation of the equations for specific elements is beyond the scope of this paper, although the formulative bases of certain arch, flat plate, and shell elements are outlined in Section E.

It has been shown[14] that the element stiffness equations of a "perfect" structure (no initial displacements) for small strain non-incremental finite displacement analysis, for conservative loading and a Lagrangian frame of reference, are of the general form.

$$[k]\{\Delta\} + [n_1(\Delta)]\{\Delta\} + [n_2(\Delta)]\{\Delta\} = \{F\} \qquad (1)$$

where $\{F\}$ and $\{\Delta\}$ are the element joint forces and corresponding displacements (degrees-of-freedom) respectively.

$[k]$ is the linear (small displacement theory) stiffness matrix.

$[n_1(\Delta)]$ is the first-order ("geometric") stiffness matrix, where the individual terms are linear functions of the degrees-of-freedom $\{\Delta\}$. A simplified form of this matrix permits linear stability analysis, as in Euler buckling.

$[n_2(\Delta)]$ is the second-order ("geometric") stiffness matrix, with individual terms a quadratic function of the degrees-of-freedom $\{\Delta\}$. These terms arise from the components of strain energy which are the first derivatives of w with respect to the spatial variables raised to the fourth power.

Upon assembly of the element relationships defined by Eq. (1) to form a representation of the complete structure, (global representation) the following equations are obtained.

$$[K]\{\Delta\} + [N_1(\Delta)]\{\Delta\} + [N_2(\Delta)]\{\Delta\} - \lambda\{P\} = 0 \qquad (2)$$

where the definitions of K, N_1, and N_2 for the global representation correspond to those given above for k, n_1, and n_2 for the respective elements. The "normalized" load vector $\{P\}$ represents the relative magnitude of the loads corresponding to the respective degrees-of-freedom $\{\Delta\}$; thus, the joint loads are applied in fixed proportion to one another. λ, the loading parameter is a scalar which can be adjusted to define a desired intensity of loading.

In indicial notation, Equation (2) becomes

$$K_{ij}\Delta_j + N_{ijk}\Delta_j\Delta_k + N_{ijkl}\Delta_j\Delta_k\Delta_l - \lambda P_i = 0 \qquad (2a)$$

Indicial notation is especially useful in nonlinear finite element analysis since the constants of the problem (N_{ijk} and N_{ijkl}) are readily identified and can be stored permanently, in contrast to the matrix format where $[N_1(\Delta)]$ and $[N_2(\Delta)]$ are dependent on the displacements and change continually during the numerical analysis process. The matrix(Eq. 2) and indicial (Eq. 2a) notations will be employed interchangeably throughout.

The "perfect" structure, to which the above equations apply, constitutes an analytical reference base for the study of the behavior of imperfect structures or for structures for which the applied loads deviate slightly from those which produce bifurcation, by use of the perturbation method. Analysis of such

problems requires an extension of Equations (1) and (2) to include the influence of initial displacements and the above-cited load deviations. To account for the former, we assume that the initial displacements are distributed throughout the structure in a form identical to the elastic displacements; hence, the initial displacements are properly described by joint values Δ_j^i. Also, we designate the total displacements by Δ_j^T. The system equilibrium equations now become (see Ref. 15 for a representative detailed development)

$$K_{ij}(\Delta_j^T - \Delta_j^i) + N_{ijk}\Delta_j^T\Delta_k^T - N_{ijk}\Delta_j^i\Delta_k^i + N_{ijkl}\Delta_j^T\Delta_k^T\Delta_l^T$$

$$-N_{ijkl}\Delta_j^i\Delta_k^i\Delta_l^T - \lambda P_i = 0 \qquad (3)$$

Noting now that the net displacements are $\Delta_j = \Delta_i^T - \Delta_j^i$, substituting this in Eq. (3) and collecting terms, we have

$$K_{ij}\Delta_j + N_{ijk}\Delta_j\Delta_k + N_{ijkl}\Delta_j\Delta_k\Delta_l$$

$$+ 2N_{ijk}\Delta_j^i\Delta_k + 3N_{ijkl}\Delta_j^i\Delta_k\Delta_l - \lambda P_i = 0 \qquad (4)$$

At this juncture two assumptions are made, consistent with Koiter's original development,[5] which simplify considerably the above equation. First, the term $3N_{ijkl}\Delta_j^i\Delta_k\Delta_l$ is assumed to be negligible and is discarded. Secondly, the evaluation of the term $2N_{ijk}\Delta_j^i\Delta_k$ is to be based on a <u>linearized</u> pre-buckling solution for Δ_k. Thus, we can write Δ_k as $\lambda(\Delta_k^o)^1$ where λ is the load parameter and (Δ_k^o) is the slope of the linear pre-buckling load-displacement relationship, and

$$2N_{ijk}\Delta_j^i\Delta_k = \lambda(2N_{ijk}\Delta_j^i(\Delta_k^o)^1) = -\lambda\gamma I_i \qquad (5)$$

where I_i is the component of a "load" vector representative of initial imperfections and γ is a parameter which can be adjusted to define the severity of the initial imperfections, Eq. 4 becomes

$$K_{ij}\Delta_j + N_{ijk}\Delta_j\Delta_k + N_{ijkl}\Delta_j\Delta_k\Delta_l - \lambda P_i - \lambda\gamma I_i = 0 \qquad (6)$$

It is clear that one may deal with loads which deviate slightly from those which produce bifurcation by assigning such load values directly to the vector {I}. It should also be noted that the restriction of Δ_k^o above to the linearized prebuckling

state does not exclude consideration of a nonlinear prebuckling state in the total problem; the latter is included in all subsequent operations.

B. PREBUCKLING ANALYSIS

The perturbation method is based upon an expansion about the bifurcation point of the perfect structure. In the present development it is necessary to trace the fundamental path of the perfect structure, represented by the solution of Equation (2) to points beyond bifurcation. The method chosen here is that of direct iteration.

In the basic form of the direct iterative method, assume that the solution is to be obtained for a load intensity designated by λ^q. Also, assume that solution data from a prior load level (say λ^{q-1}) is available and is designated as $\{\Delta\}^o$. Thus, the matrices $[N_1]$ and $[N_2]$ may be formed using $\{\Delta\}^o$ and we may solve Equation (2) to yield

$$\{\Delta\}^1 = [K]^{-1} \{\lambda^q \{P\} - [N_1(\Delta)^o] \{\Delta^o\} - [N_2(\Delta)^o]\{\Delta\}^o\} \quad (7)$$

where the superscript 1 on $\{\Delta\}^1$ denotes the first iteration in the solution at λ^q. We then re-form $[N_1]$ and $[N_2]$ on the basis of $\{\Delta\}^1$, so that

$$\{\Delta\}^2 = [K]^{-1} \{\lambda^q \{P\} - [N_1(\Delta)^1] \{\Delta\}^1 - [N_2(\Delta)^1]\{\Delta\}^1\} \quad (8)$$

which is now solved for $\{\Delta\}^2$. In the general, j^{th}, iterative solution

$$\{\Delta\}^j = [K]^{-1} \{\lambda^q \{P\} - [N_1(\Delta)^{j-1}] \{\Delta\}^{j-1} - [N_2(\Delta)^{j-1}] \{\Delta\}^{j-1} \quad (9)$$

The iterative sequence continues until $\{\Delta\}^j$ is within $\{\Delta\}^{j-1}$ to a specified tolerance. Note that direct iteration requires only the inversion of the linear stiffness matrix and continued re-formation of $[N_1]$ and $[N_2]$.

The knowledge of a nearby solution, as for $\{\Delta\}^o$ in Equation 7, enhances the efficiency of the iterative process. Hence, the analysis is performed at various load levels, extending from a level close to zero load through to a level somewhat beyond the bifurcation load.

Convergence difficulties are encountered when the nonlinearities are severe. Such difficulties are often manifested by continued iteration in a loop about the convergent solution. In such cases an improved procedure is to employ a higher-order iterative scheme as described in Reference 16.

C. DETERMINATION OF BIFURCATION

To determine the first branching from the fundamental path (the bifurcation point) we invoke the familiar stability condition that the second variation of the potential energy be zero at such a point. The equilibrium equation (Equation 2) represents the first variation of the potential energy and by applying the second variation one obtains the following condition at $\lambda = \lambda^c$

$$\text{Det} = | [K] + 2 [N_1(\Delta)] + 3 [N_2(\Delta)] | = 0 \qquad (10)$$

where Det symbolizes the determinant of the indicated matrix. (The factors 2 and 3 on $[N_1]$ and $[N_2]$ arise from imposition of the second variation; see Equation 2a).

To illustrate the manner in which the above condition is employed in identification of the bifurcation point Figure 2a shows a representative load-parameter-displacement (λ-Δ) plot while Figure 2b shows the corresponding variation of Det with λ. Thus, Det > 0 for $0 < \lambda < \lambda^c$ and Det < 0 for $\lambda > \lambda^c$. By establishing m solution points to either side of λ^c and by Lagrange interpolation we have

$$\lambda = \sum_{i=1}^{m} \left(\prod_{j=1,\ j=i+1}^{i-1,\ m} \frac{\text{Det}-\text{Det}_j}{\text{Det}_j-\text{Det}_i} \right) \lambda_i \qquad (11)$$

where Det_i and λ_i denote the corresponding values at the ith load level. From Eq. (11), the bifurcation load λ^c is calculated by setting Det = 0, i.e.

$$\lambda^c = \lambda \big|_{\text{Det} = 0}$$

Since the displacements and their derivatives at λ^c are needed for determination of the postbuckling path, they are also calculated by interpolation.

$$\Delta_i(\lambda) = \sum_{j=1}^{m} \left(\prod_{k=1,\ k=j+1}^{j-1,\ m} \frac{\lambda-\lambda_k}{\lambda_k-\lambda_j} \right) \Delta_i^j \qquad (12)$$

where Δ_i^j denotes the value of Δ_i at the jth load level λ_j. Then the desired displacements Δ_i^c are found by setting $\lambda=\lambda^c$. The

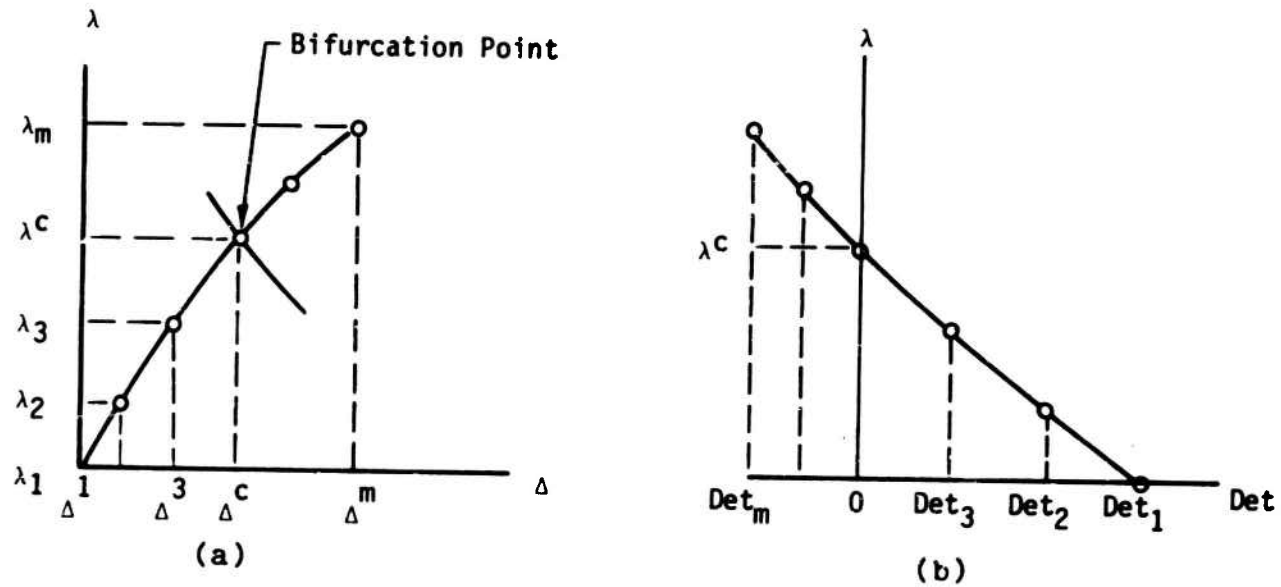


FIGURE 2. DETERMINATION OF BIFURCATION POINT VIA INTERPOLATION

derivatives of Δ_i are found by direct differentiation of Eq. (12) with respect to λ and evaluation at λ^c.

D. POSTBUCKLING ANALYSIS

Figures 1.a and 1.b illustrate, in a representative λ-Δ space, the single post-buckling path of a perfect structure emanating from the bifurcation point. We now consider the use of the perturbation method in establishment of an analytical description of this path and, simultaneously, of the path of imperfect (or differently loaded) forms of the same structure. Thus, we direct attention to the more general equilibrium relationship (Equation 6) which accounts for these imperfections and/or load components.

In initiating the analytical description of these paths, we describe the displacement state in terms of "sliding coordinates"[8]

$$\{\Delta\} = \{\bar{\Delta}\} + \{\Delta^P\} \tag{13}$$

where now $\{\bar{\Delta}\}$ describes displacements on the fundamental path of the perfect structure and $\{\Delta^P\}$ gives the displacements on the postbuckling path with the fundamental path as a reference base. Thus, a mapping of the postbuckling behavior in λ - $\{\Delta^P\}$ space can be effected with $\{\bar{\Delta}\} = 0$.

To obtain the equilibrium equation in terms of the new coordinates, we substitute Equation (13) into Equation (6). It is convenient to revert to indicial notation with $\{\Delta\} = \Delta_j$, $\{\bar{\Delta}\} = \bar{\Delta}_j$ and $\{\Delta^P\} = \Delta_j^P$, since in effecting the products $\Delta_j\Delta_k$ and $\Delta_j\Delta_k\Delta_l$ the term $\Delta_j = (\bar{\Delta}_j + \Delta_j^P)$ can be treated as a binomial in conventional manner. Thus, we have

$$K_{ij}\bar{\Delta}_j + K_{ij}\Delta_j^P + N_{ijk}(\bar{\Delta}_j + \Delta_j^P)(\bar{\Delta}_k + \Delta_k^P) + N_{ijkl}(\bar{\Delta}_j + \Delta_j^P)(\bar{\Delta}_k + \Delta_k^P)(\bar{\Delta}_l + \Delta_l^P) = \lambda P_i + \lambda\gamma I_i \tag{14}$$

and, expanding and collecting multipliers of Δ_j^P and noting that a group of the resulting terms satisfies Equation (2), we obtain the equilibrium equation in the new coordinates

$$(K_{ij} + 2N_{ijk}\,\bar{\Delta}_k + N_{ijkl}\,\bar{\Delta}_k\,\bar{\Delta}_l)\Delta_j^p$$

$$+ (N_{ijk} + 3N_{ijkl}\,\bar{\Delta}_l)\,\Delta_j^p\,\Delta_k^p + N_{ijkl}\,\Delta_j^p\,\Delta_k^p\,\Delta_l^p = \lambda\gamma I_i \quad (15)$$

An analytical representation of the fundamental path must now be established. Thus, by Taylor series expansion about the bifurcation point

$$\bar{\Delta} = \{\bar{\Delta}^c\} + (\lambda-\lambda^c)\{\bar{\Delta}'^c\} + \tfrac{1}{2}(\lambda-\lambda^c)^2\{\bar{\Delta}''^c\} + \ldots \quad (16)$$

and, by substitution into Equation (15)

$$[(K_{ij} + 2N_{ijk}\,\bar{\Delta}_k^c + 3N_{ijkl}\,\bar{\Delta}_k^c\,\Delta_l^c) + (2N_{ijk}\,\bar{\Delta}_k'^c + 6N_{ijkl}\,\bar{\Delta}_k'^c\bar{\Delta}_l^c)(\lambda-\lambda^c)$$
$$+ (N_{ijk}\,\bar{\Delta}_k''^c + 3N_{ijkl}(\bar{\Delta}_k'^c\,\bar{\Delta}_l'^c + \bar{\Delta}_k^c\,\bar{\Delta}_l''^c)\,(\lambda-\lambda^c)^2]\Delta_j^p$$
$$+ [(N_{ijk} + 3N_{ijkl}\,\bar{\Delta}^c{}_l) + (3N_{ijkl}\bar{\Delta}_l^c(\lambda-\lambda^c))]\Delta_j^p\,\Delta_k^p + N_{ijkl}\Delta_j^p\,\Delta_k^p\,\Delta_l^p = \lambda\gamma I_i \quad (17)$$

where the indicated result is obtained by truncation of the series (Equation 16) at the third term and discard of terms higher than third order in Δ_i^p or in the product Δ_i^p and $(\lambda-\lambda^c)$.

It is assumed that the postbuckling path can be described by the series

$$\begin{Bmatrix} \Delta_1^p \\ \Delta_2^p \\ \cdot \\ \cdot \\ \cdot \\ \Delta_n^p \end{Bmatrix} = \Delta_1^p \begin{Bmatrix} 1 \\ q_{2_1} \\ \cdot \\ \cdot \\ \cdot \\ q_{n_1} \end{Bmatrix} + (\Delta_1^p)^2 \begin{Bmatrix} 0 \\ q_{2_2} \\ \cdot \\ \cdot \\ \cdot \\ q_{n_2} \end{Bmatrix} + (\Delta_1^p)^3 \begin{Bmatrix} 0 \\ q_{2_3} \\ \cdot \\ \cdot \\ \cdot \\ q_{n_3} \end{Bmatrix} + \ldots \quad (18)$$

i.e., $\{\Delta^p\} = \Delta_1^p\{q_1\} + (\Delta_1^p)^2\{q_2\} + (\Delta_1^p)^3\{q_3\} + \ldots \quad (19)$

or, in indicial form

$$\Delta_i^p = \Delta_1^p\,q_{i_1} + (\Delta_1^p)^2\,q_{i_2} + (\Delta_1^p)^3\,q_{i_3} + \ldots \quad (19a)$$

where each vector ($\{q_1\}$,$\{q_2\}$, etc.) is a mode of displacement to be determined as described below, and Δ_1^p, the "path parameter",

is a pre-selected degree-of-freedom. Equation (18) designates this parameter as the first-listed degree-of-freedom. This can be done without loss of generality since it is always possible to arrange the equations so that a chosen degree-of-freedom (generally, the displacement at a prominent point) appears in the first location.

An expansion of the load parameter in the post-buckling regime is also required and is given in terms of the path parameter.

$$\lambda - \lambda^c = \Gamma_1 (\Delta_1^p) + \Gamma_2 (\Delta_1^p)^2 + \Gamma_3 (\Delta_1^p)^3 + \ldots \qquad (20)$$

where the load parameters Γ_1, Γ_2, Γ_3, .. are as yet unknown. We next substitute (19a) and (20) into (17), write the imperfection term $\lambda\gamma I_i$ as $\frac{\lambda\gamma I_i}{(\Delta_1^p)^2} \times (\Delta_1^p)^2$, and collect terms in like powers of Δ_1^p. We obtain, for $\Delta_1^p \neq 0$

$$\bar{K}_{ij}\, q_{j_1} = 0 \qquad (21a)$$

$$(\bar{K}_{ij}\, q_{j_2} + \bar{K}'_{ij}\, \Gamma_1\, q_{j_1} + \bar{N}_{ijk}\, q_{j_1}\, q_{k_1}) = \frac{\lambda\gamma I_i}{(\Delta_1^p)^2} \qquad (21b)$$

$$(\bar{K}_{ij}\, q_{j_3} + \bar{K}'_{ij}(\Gamma_2\, q_{j_1} + \Gamma_1\, q_{j_2}) + \bar{K}''_{ij}\, \Gamma_1^2\, q_{j_1} + 2\bar{N}_{ijk} q_{j_1}\, q_{k_2} + \bar{N}'_{ijk}\, \Gamma_1\, q_{j_1}\, q_{k_1} + N_{ijkl}\, q_{j_1}\, q_{k_1}\, q_{l_1}) = 0 \qquad (21c)$$

The initial imperfection term is accounted for in the manner indicated since it would have no effect if represented as a constant and would disallow the overall procedure of definition of the postbuckling path if represented as linear in Δ_1^p. Thus, a quadratic representation in Δ_1^p is chosen.

Now, from (21a) it is clear that $\{q_1\}$ is the eigenvector of $[\bar{K}_{ij}]$, normalized on the term corresponding to Δ_1^p. The value of Γ_1 is obtained by solution of an equation resulting from the premultiplication of Eq. (21b) by $\{q_1\}$, and then $\{q_2\}$ is obtained from Eq. (21b) after back-substitution of the expression for Γ_1. The values of Γ_2 and $\{q_3\}$ are determined by similar operations on Eq. (21c). If additional terms are taken in the above series the corresponding load parameters and postbuckling displacement vectors are also obtained in this manner.

Space limitations do not allow here the presentation of details of this procedure, or of the specific form of the results in terms of the basic quantities $\bar{K}_{ij}$, etc. This information is given in Reference 16. It may be of interest, however, to cite a single term of the series Eq. (20), the post-buckling load-displacement parameter relationship. Thus, for Γ_1 it is found that

$$\Gamma_1 = \frac{\bar{N}_{ijk}\, q_{i_1} q_{j_1} q_{k_1}\, (\Delta_1^p)^2 + \lambda\gamma I_{i_1}\, q_{i_1}}{\bar{K}'_{ij}\, q_{i_1} q_{j_1}\, (\Delta_1^p)^2} \qquad (22)$$

Note that when the initial imperfection is zero (represented by $\gamma = 0$), a nonzero value of Γ_1 remains, defining the postbuckling path of the perfect structure.

E. ILLUSTRATIVE EXAMPLES

The present section is devoted to a general description of numerical results for three illustrative examples, involving a shallow arch, a plate, and a curved thin shell structure, respectively. Although the problems solved are elementary from the standpoint of finite element representation, they delineate all features of the more complex situations and are among the few cases which have been studied thoroughly and for which comparison solutions or test data are available. Such comparisons are essential to a study addressed to a class of problems for which a multitude of alternative procedures have only recently been proposed.

1. Clamped Thin Shallow Circular Arch

The problem of instability of the thin shallow circular arch with clamped ends has drawn much attention in the literature of geometrically nonlinear and postbuckling analysis because it is perhaps the most sophisticated structure for which "exact" solutions have been obtained[13,17]. The objectives of this illustrative example are to verify accuracy in determination of the bifurcation point following upon a nonlinear fundamental path and to demonstrate prediction of behavior for both perfect and imperfect forms of the arch by use of the present perturbation method.

The geometry of the arch (see Figure 3) is characterized by the parameter $R\theta_o^2/h$ and this parameter also governs in part the

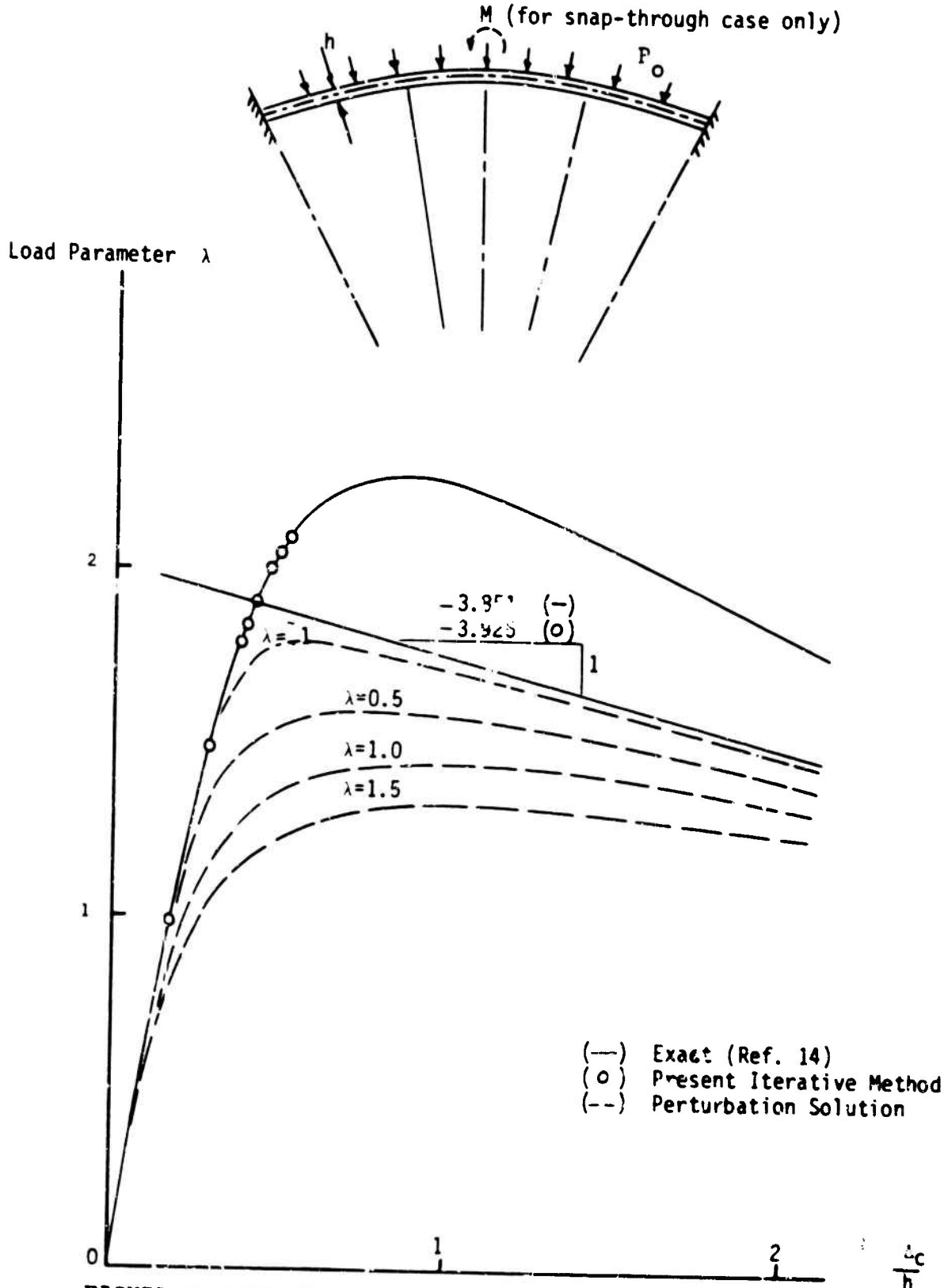


FIGURE 3. LOAD VS. CENTRAL DEFLECTION FOR PRE BUCKLING, POST BUCKLING AND IMPERFECT BEHAVIORS

form of buckling, i.e., snap-through or bifurcation. The dimensions chosen here yield a value of 10.0, the same value employed in Reference 13. The finite element representation consists of eight equally-spaced arch elements.

In the case of uniform radial loading of intensity p_o, bifurcation occurs prior to snap-through as illustrated in Figure 3 by solid lines for the classical solution. The direct iterative scheme discussed in Section C is used to yield the solution points on the fundamental path as given by the circled points. Lagrangian interpolation gives the bifurcation load $\lambda^c = 1.9075$, which is within 0.2% of the exact value. The numbers of iterative cycles to achieve convergence at each load level for various specified convergence criteria are plotted in Fig. 4. It is of interest to note that near the bifurcation point the number of cycles increases sharply, but, that monotonic convergence is observed at all load levels above or below the bifurcation point.

2. Flat Plate Post-Buckling

The next example refers to the post-buckling behavior of a flat, simply-supported rectangular plate under uniform axial compression. The element employed in these calculations is doubly-curved shell element portrayed in Figure 5, whose properties are based upon a 16-term (bicubic) expansion of each of the displacement components u, v and w. This element is a generalization of the cylindrical shell element introduced by Bogner, et al[18]. Formulation of the present representation is described in Reference 19. For this problem the principal radii of curvature are set equal to infinity.

The problem data are shown in Figure 6; due to symmetry, the analysis is performed with a single element in one quadrant. It should be noted that the analytical model permits freedom of inplane displacement; the usual restriction of classical solutions[20] to linear edge displacement states, etc. are not invoked.

The bifurcation point, calculated in a linear stability analysis (the pre-buckling state is zero in the transverse displacement) is found to be 36.3 lb./in. compared to the classical result[21] of 36.1 lb./in. (0.5% error). The postbuckling path calculated by the present perturbation method is shown via a dashed line in

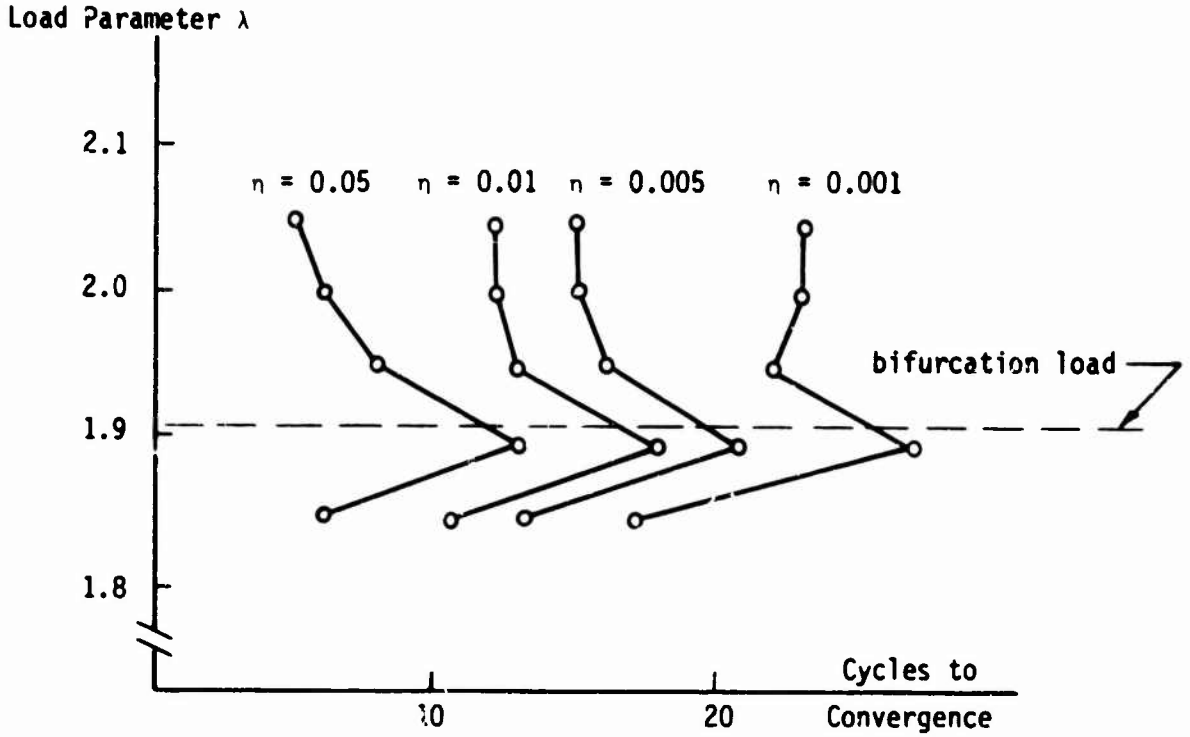


FIGURE 4. NUMBER OF ITERATIONS FOR CONVERGENCE VS. LOAD LEVEL FOR VARIOUS CONVERGENCE CRITERIA (η)

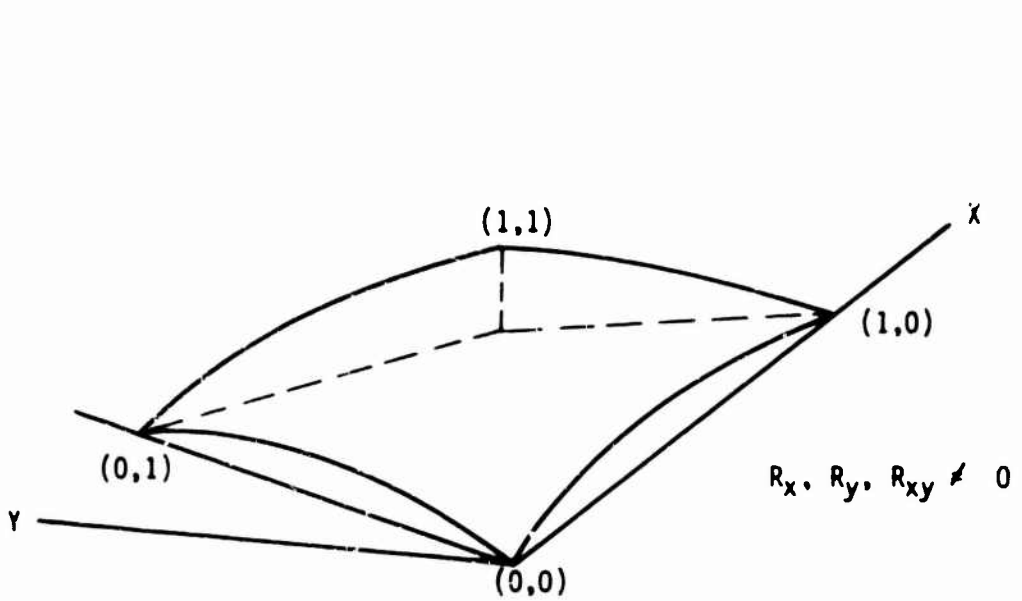
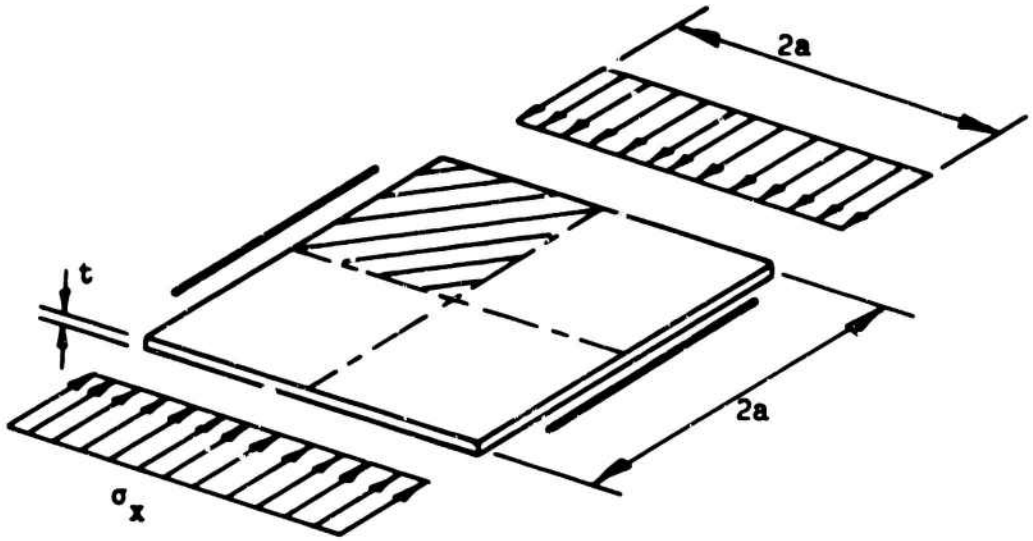


FIGURE 5. GENERAL QUADRILATERAL SHELL ELEMENT (DEFINED IN ISOPARAMETRIC COORDINATES)



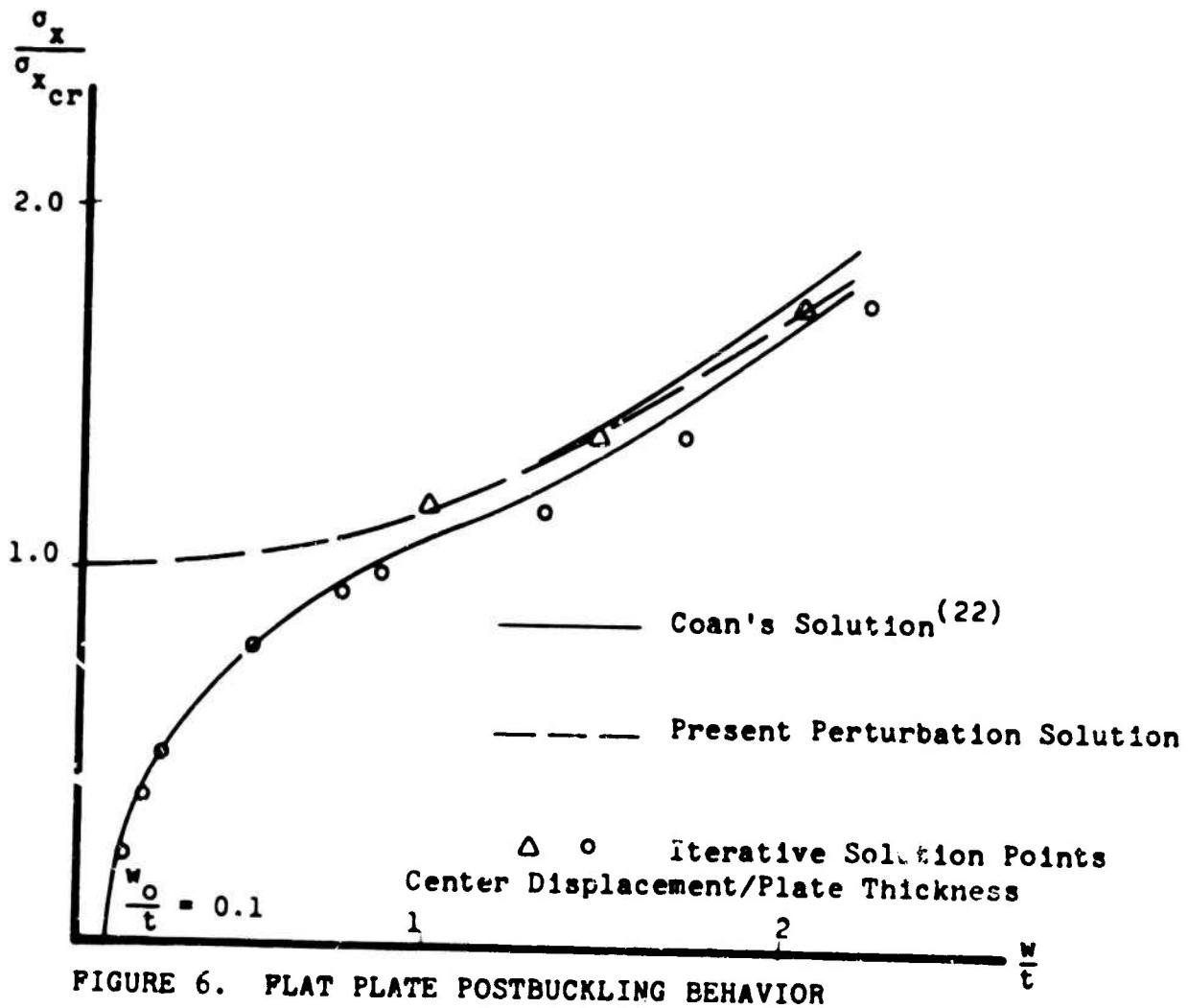


FIGURE 6. FLAT PLATE POSTBUCKLING BEHAVIOR

Fig. 6 and is compared with the solution by Coan[22] who also permits unrestricted inplane displacement. The two solutions are in close agreement. These results were checked by performance of an iterative solution of Equation 2 at applied load levels $\sigma_x/\sigma_{x_{cr}} = 1.193$, 1.385, and 1.748. These solutions, designated by triangular symbols in Fig. 6, confirm the range of validity of the perturbation solution.

The analysis of the plate with initial geometric imperfections of the form $w^1 = w^o \sin \pi\xi \sin \pi\eta$, with w^o the central displacement and $\xi = x/a$, $\eta = y/b$, was also attempted. For this case it is convenient to effect an iterative solution of Eq. 3, rather than to apply perturbation concepts. The solution points of this procedure are shown circled in Figure 6. The solution by Coan[22] for the same data is compared and good agreement is again indicated.

3. Cylindrical Shell Under Uniform Load

As a final example we consider the clamped circular cylindrical panel subjected to uniform lateral load, whose properties are described in Fig. 7. This structure has been analyzed in several studies, but only Morin[10] defines bifurcation and analyzes for higher load levels. The present computation is performed with 4 elements of the type shown in Fig. 5 (2x2 grid in a quadrant). Again, direct iterative solution is employed in preference to the perturbation method. Results are shown in Fig. 7.

The displacements of the linear analysis were chosen as the initial guess for the first load level. Subsequent initial guesses were obtained by multiplying the displacements of the previous load level by the ratio of the current to the prior load levels. It is interesting to note that the lowest determinant in the present analysis occurs at $\lambda = 0.230$ p.s.i., compared to the 0.223 p.s.i. calculated by Morin.[10] The present results and those of Refs. 23 and 24 compare well and are collectively in significant disparity with those of Ref. 10.

F. CONCLUDING REMARKS

A procedure for finite element analysis of geometrically nonlinear problems, covering the pre-buckling and initial post-

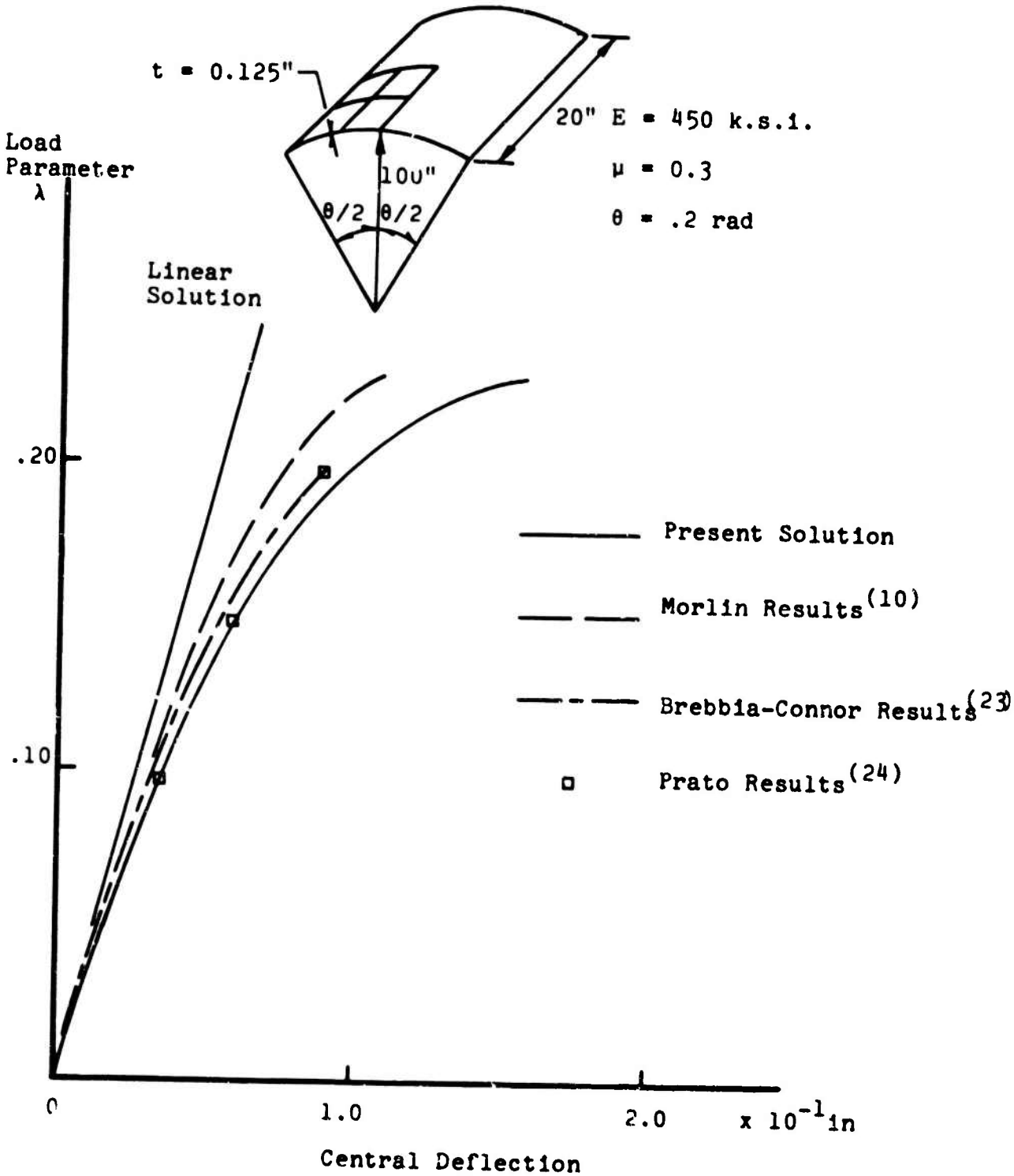


FIGURE 7. LOAD-DEFLECTION RESPONSE.
CLAMPED CIRCULAR CYLINDRICAL SHELL
UNDER UNIFORM PRESSURE LOAD.

buckling regimes, snap-through buckling, and accounting for initial imperfections, has been presented. The initial post-buckling and snap-through aspects of this procedure correspond closely to Thompson's approach[9], differing in detailed application to the limit points of imperfection-sensitive structures where the present method is believed to be computationally more efficient.

The present development applies only to the case of a single post-buckling path. Certain classical situations are known to possess multiple post-buckling paths; these stem from linear pre-buckling analyses. Even if analysis based on the more realistic representation of a nonlinear pre-buckling state discloses only a single post-buckling path at the first critical point, a serious problem may arise due to the closeness of adjacent critical points. It does not appear that the present approach is capable of dealing with this situation, due to the restriction of its validity, as an asymptotic theory, to behavior in the vicinity of the critical point.

In an alternative approach developed by the writers, and also detailed in Reference 16, none of the limitations of an asymptotic theory are present. This approach is based upon an extrapolation of determinants of the equilibrium equation and applies to limit point problems. The computational costs are higher than for a perturbation approach, however, and the reliability of an extrapolation is always in question.

The relative efficiency and accuracy of the subject procedures and various alternative approaches (e.g., references 3, 9, 10) remain open questions. Since analysts cannot agree upon the optimum procedures in the restricted area of nonlinear pre-buckling analysis (see Reference 25) it is unrealistic to expect the definition of the most appropriate approach to post-buckling analysis at the present time. These measures will be obtained only after significant experience in practical application is recorded.

ACKNOWLEDGEMENT: Work described in this paper was supported by the NASA under grant NGR-33-010-070.

REFERENCES

1. Gallagher, R. H., "Analysis of Plate and Shell Structures", Proc. of Conf. on Application of Finite Element Methods in Civil Engineering, Vanderbilt Univ., 1969.

2. Hutchinson, J. W. and Koiter, W. T., "Postbuckling Theory", Applied Mech. Reviews, Dec. 1970.

3. Haftka, R. T., Mallett, R. H. and Nachbar, W., "A Koiter-Type Method for Finite Element Analysis of Nonlinear Structural Behavior", AFFDL TR 70-130, V. 1, Nov. 1970.

4. Bienek, M., "Post-Critical Behavior", Introductory Report for Ninth Congress of IABSE, Amsterdam, May, 1972.

5. Koiter, W. T., "On the Stability of Elastic Equilibrium", Thesis, Delft, 1945.

6. Budiansky, B. and Hutchinson, J., "A Survey of Some Buckling Problems", AIAA Journal, Vol. 4, Sept. 1966, pp. 1505-1510.

7. Sewell, M. J., "A General Theory of Equilibrium Paths Through Critical Points", Proc. Royal Soc. A. 306, pp. 201-223, 1968.

8. Thompson, J. M. T., "A General Theory for the Equilibrium and Stability of Discrete Conservative Systems", ZAMP, Vol. 20, 1969.

9. Thompson, J. M. T., "A New Approach to Elastic Branching Analysis", J. Mech. Phys. Solids, V. 18, 1970.

10. Morin, N., "Nonlinear Analysis of Thin Shells", Report R70-43, Dept. of Civil Engrg., M.I.T., 1970.

11. Dupuis, G. A., Pfaffinger, D. and Marcal, P. V., "Effective Use of the Incremental Stiffness Matrices in Non-linear Geometric Analysis", IUTAM Symposium on High Speed Computing of Elastic Structures, Liege, Belgium, 1970.

12. Lang, T. E., "Post-Buckling Response of Structures Using the Finite Element Method", Ph.D. Thesis, Univ. of Washington, 1969.

13. Kerr, A. D. and Soifer, M. T., "The Linearization of the Prebuckling State and its Effect on the Determined Instability Load", Trans. ASME, Journal of Applied Mech., V. 36, pp. 775-783, 1969.

14. Mallett, R. and Marcal, P. V., "Finite Element Analysis of Nonlinear Structures", Proc. ASCE, Journal of the Structural Div., V. 94, pp. 2081-2106, 1968.

15. Vos, R., "Finite Element Analysis of Plate Buckling and Postbuckling", Ph.D. Diss., Rice Univ., Dec. 1970.

16. Mau, S-T. and Gallagher, R. H., "A Finite Element Procedure for Nonlinear Pre-Buckling and Initial Post-Buckling Analysis" NASA Contractor's Report. To be published.

17. Schreyer, H. L. and Masur, E. F., "Buckling of Shallow Arches", Proc. ASCE, Journal of the Engineering Mechanics Div., V. 92, No. EM4, pp. 1-20, Aug. 1966.

18. Bogner, F., Fox, R. and Schmit, L., "Finite Deflection Analysis Using Plate and Cylindrical Shell Discrete Elements", AIAA Journal, V. 6, No. 5, May 1968.

19. Lien, S., "Finite Element Thin Shell Pre- and Post-Buckling Analysis", Ph.D. Thesis, Structural Engineering Dept., Cornell University, 1971.

20. Yamaki, N., "Postbuckling Behavior of Rectangular Plates with Small Initial Curvature Loaded in Edge Compression", Trans. ASME, J. of Appl. Mech., V. 26, pp. 407-414.

21. Timoshenko, S. and Gere, J., "Theory of Elastic Stability", 2nd Ed., McGraw-Hill Book Co., 1961.

22. Coan, J. M., "Large Deflection Theory for Plates with Small Initial Curvature Loaded in Edge Compression", J. Appl. Mech., V. 18, 1951.

23. Brebbia, C. and Connor, J., "Geometrically Nonlinear Finite Element Analysis", Proc. ASCE, J. of the Eng. Mech. Div., V. 95 , No. EM2 Apr. 1969.

24. Prato, C., "A Mixed Finite Element for Thin Shell Analysis", Ph.D. Diss., M.I.T., Sept. 1968.

25. Haisler, W., Stricklin, J. and Stebbins, F., "Development and Evaluation of Solution Procedures for Geometrically Nonlinear Structural Analysis by the Direct Stiffness Method", Proc. AIAA/ASME 12th Structures, Structural Dynamics, and Materials Conf., Anaheim, Calif., Apr. 1971.

A GENERAL APPROACH TO PROBLEMS OF PLASTICITY AND LARGE DEFORMATION USING ISOPARAMETRIC ELEMENTS

O. C. Zienkiewicz*
G. C. Nayak**

University of Wales, Swansea, U. K.

A unified formulation of large deformation, large strain and plasticity problems is given. Lagrangian forms are preferred but an alternative Eulerian system is given. Various solution schemes are discussed and a program capable of dealing with alternatives is outlined. Isoparametric formulation is shown not only to possess merits of economy but to be highly versatile in the Lagrangian and Eulerian formulations of materially and geometrically nonlinear problems.

INTRODUCTION

Recent years have seen a growing interest in the application of the finite element method to nonlinear problems of structural mechanics. The three classes of problems which have been usually approached separately are:

(1) geometrically large deformation associated with small, elastic, strain

(2) geometrically large deformation associated with finite strains

and

(3) nonlinear material properties

Further, differing methods of formulation and of the solution of the nonlinear system have again introduced another classification. Most of these solution processes fall into two broad classes of

(1) incremental procedures where a 'marching' type of approach is used and equilibrium path is only approximately followed, with equilibrium checks occasionally introduced, and

(2) iterative procedures in which equilibrium is approached at all stages of computation.

The different approaches to the formulation and solution of nonlinear problems have been often confusing and, we believe, have led to misunderstanding on occasion. Several notable attempts at introduction of some degree of precision in the definition of formulation and solution have been made by Marcal (Reference 1 and 2)

*Professor - Head of Civil Engineering Department
**Commonwealth Research Fellow (Reader, University of Roorkee, India

Oden (References 3 and 4), Haisler et al (Reference 5) and others (References 6, 7, and 8), and similar formulation was achieved in our Institute in 1968. In this paper we shall pursue this problem again and attempt to present

(a) the essentials of the formulation which has a degree of generality and conforms to the matrix-finite element schemes

(b) a generalized solution process which can use the advantages of several possible alternatives at will

and

(c) the advantages of isoparametric elements which appear to be naturally suited to this broad class of problems

2. A general formulation of equilibrium equations in finite element context

Let

$$\mathbf{x} \equiv [x,y,z]^T \tag{1}$$

define the rectangular coordinates of a material point P in a body shown in Figure 1 before deformation. If this point is displaced by

$$\mathbf{u} \equiv [u,v,w]^T \tag{2}$$

measured relative to the fixed frame of reference, its new coordinates will be

$$\bar{\mathbf{x}} = \mathbf{x} + \mathbf{u} \tag{3}$$

Further, using the usual finite element approximation we shall take the displacements as defined by suitable shape functions $\mathbf{N}$ of coordinates x, and a set of nodal parameters $\boldsymbol{\delta}$, element by element, as

$$\mathbf{u} = \mathbf{N}\boldsymbol{\delta} \tag{4}$$

If $\bar{\mathbf{p}}$ denotes the surface forces per unit area of the deformed body, and $\mathbf{q}$ the body forces per unit mass, then a simple application of the virtual work principle yields, by equating the external and internal work, the approximate equilibrium conditions as

$$\int_{\bar{V}} \bar{\rho}\, \mathbf{q}^T d\mathbf{u}\, d\bar{V} + \int_{\bar{A}} \bar{\mathbf{p}}^T d\mathbf{u}\, d\bar{A} = \int_{\bar{V}} \bar{\boldsymbol{\sigma}}^T d\bar{\boldsymbol{\varepsilon}}\, d\bar{V} \tag{5}$$

in which $\bar{V}$, and $\bar{A}$ refer to volumes and areas of the deformed body, $\bar{\rho}$, is the density in the deformed state and $\bar{\boldsymbol{\sigma}}$ and $d\bar{\boldsymbol{\varepsilon}}$ refer to vector forms of the Eulerian (real) stress and deformation increment (see Appendix I, Equation (A15) in the distorted coordinates $\bar{\mathbf{x}}$.

Alternatively, we may rewrite Equation (5) in terms of variables referred to the original, undistorted, coordinates and now obtain

$$\int_V \rho \mathbf{q}^T d\mathbf{u}\, dV + \int_A \left[\frac{d\bar{A}}{dA}\bar{\mathbf{p}}^T\right] d\mathbf{u}\, dA = \int_V \boldsymbol{\sigma}^T d\boldsymbol{\varepsilon}\, dV \tag{7}$$

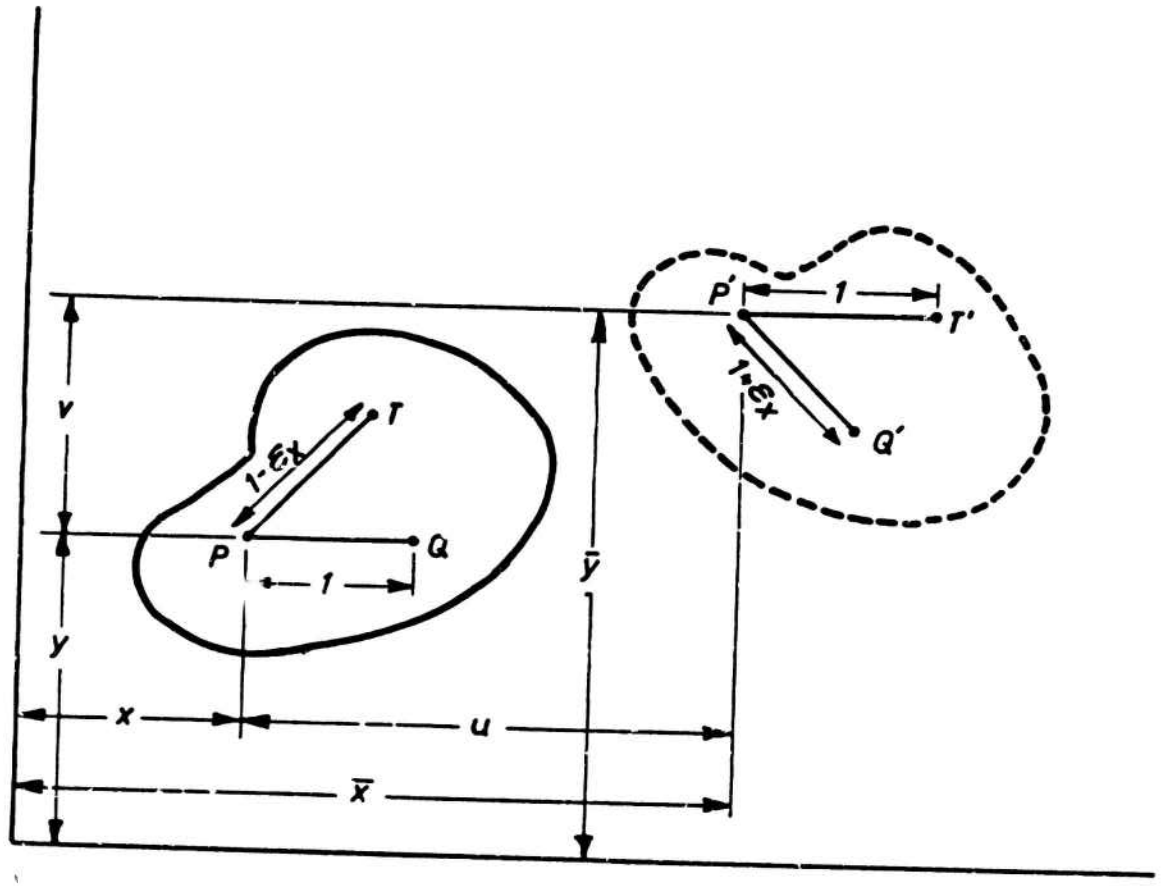


FIG. 1 DEFINITION OF LAGRANGIAN AND EULERIAN CO-ORDINATES AND STRAINS (EXACT ONLY IN INFINITESIMAL CASES)

in this σ stands for the Piola-Kirchhoff stress, and $d\epsilon$ for the increment of Green's strain, both referred to as the Lagrange formulation and both again written in vector form.

The equivalence of the two work statements and the derivation of the appropriate definitions is relegated to the Appendix where appropriate matrix forms are given. We shall content ourselves by giving here only the general statements applicable in finite element analysis. Thus, from the above definition we can write always

$$d\bar{\epsilon} = \bar{\mathbf{B}}^{0} d\delta \tag{8}$$

for the increment of Eulerian deformation using updated coordinates or

$$d\epsilon = \mathbf{B} d\delta \tag{9}$$

for the increment of Green's strain in Lagrangian (original) coordinates and obtain by substitutions in Equations 4, and 8 or 9 respectively

$$\bar{\psi} \equiv \bar{\mathbf{R}} - \int_{\bar{V}} \bar{\mathbf{B}}^{0T} \bar{\sigma} \, d\bar{V} = \mathbf{0} \tag{10}$$

or

$$\psi = \mathbf{R} - \int_{V} \mathbf{B}^{T} \sigma \, dV = \mathbf{0} \tag{11}$$

as the equilibrium equation in Eulerian and Lagrangian systems respectively. Here

$$\mathbf{R} = \bar{\mathbf{R}} = \int_{\bar{V}} \bar{\rho} \mathbf{N}^{T} \mathbf{q} \, d\bar{V} + \int_{\bar{A}} \mathbf{N}^{T} \bar{\mathbf{p}} \, d\bar{A}$$

$$= \int_{V} \rho \mathbf{N}^{T} \mathbf{q} \, dV + \int_{A} \mathbf{N}^{T} \mathbf{p} \frac{d\bar{A}}{dA} \, dA \tag{12}$$

gives the equivalent external nodal forces while the second term in Equations 10 and 11 can be interpreted as the internal force reactions.

In general both $\mathbf{R}$ and $\mathbf{B}$ depend on the displacement parameters δ and as the stress may be a nonlinear function of strain, the set of equations implied in Equations 10 and 11 is nonlinear, and special solution methods will have to be used. The solution of these equations for the displacement parameters, δ, gives the solution of the basic problem.

In both forms, Equations 10 or 11, the 'residuals' ψ or $\bar{\psi}$ can be visualized as nodal forces required to bring the assumed displacement pattern into nodal equilibrium and therefore these should be reduced to as nearly zero as possible. This physical concept is useful with calculations, which at times are so complex as to elude simple interpretation.

3. Outline of the solution of the solution of the nonlinear equations.

The Lagrangian Equation 11 or its Eulerian equivalent 10 can be solved in a variety of ways. If the loads are proportional and are represented by some proportionality parameter, λ, then we can write either of the equations in the form

$$\psi \equiv \psi(\delta, \lambda) = \mathbf{Q}(\delta).\lambda - \mathbf{F}(\delta) = \mathbf{0} \tag{13}$$

with **F** standing for 'internal forces' equilibrating the external loads and **Q** is the value of **R** for unit λ. ψ, will be called the residual force vector. For solution of Equation 13 two alternative processes can be used:

Incremental solution

The Equation 13 is considered as a function of one parameter λ and by differentiation we have

$$\frac{d\psi}{d\lambda} = -\lambda K_\lambda \frac{d\delta}{d\lambda} + Q - K_T \frac{d\delta}{d\lambda} = 0 \tag{14}$$

where

$$dQ = -K_\lambda \, d\delta \tag{15}$$

defines what may be called the 'initial load' stiffness matrix (which will be zero with conservative loading) and

$$dF = K_T d\delta \tag{16}$$

gives the well known tangential stiffness matrix.

Taking small increments of the parameter λ one can write

$$\Delta\delta = (K_T + K_\lambda)^{-1} Q \Delta\lambda \tag{17}$$

and starting from known initial equilibrium conditions solution can be incremented.

Clearly, though this process has been used by many investigators, very small increments of λ will be required not to diverge from the equilibrium condition. Such numerical refinements as the Runge–Kutta process etc. increases the accuracy of this, essentially marching, process but suffer from the same possibility of divergence.

An alternative to the incremental solution is presented by: iterative, Newton–Raphson solution.

Here the load parameter λ is considered as fixed and the solution of Equation 13 is attained iteratively. Noting that now $d\lambda \to 0$ we have by Equation 14

$$d\psi = -(K_\lambda + K_T \lambda)^{-1} \, d\delta \tag{18}$$

and if the n^{th} iterate δ_n gives a nonzero residual force ψ_n, the next iterate becomes, using the Newton–Raphson method

$$\delta_{n+1} = \delta_n + \Delta\delta_n \tag{19}$$

with

$$\Delta\delta_n = -(K_\lambda + K_T)^{-1} \psi_n \tag{20}$$

A modification frequently used to avoid repeated recalculation and inversion of a continuously varying stiffness matrix is to keep this matrix unchanged during several successive iterations. Indeed, on occasion, convergence can be combined with economy

replacing the matrix ($\mathbf{K}_\lambda + \mathbf{K}_T$) by $\mathbf{K}_0$ matrix corresponding to small deflection and elastic behaviour. Indeed, this is the basis of so called 'initial stress process' introduced by Zienkiewicz, Valliappan, and King in the context of plasticity (Reference 9).

In the Newton–Raphson process and even more so in its modifications there is no guarantee that convergence will always be achieved, or indeed that when it is achieved that the correct solution is obtained if multiple equilibrium states are possible.

It is thus most expedient to combine the various advantages of the alternative approaches in a combined algorithm.

Combined algorithm

In this the load parameter λ is incremented in several finite steps. In each step the first approximation to the solution is obtained by the incremental process using Equation 17 and the value of $\mathbf{K}_T$ pertaining to the initial conditions (or the value last found in previous step). ($\mathbf{K}_\lambda$ is not included as its value is generally small). Considering this value as the first approximation, the residual $\boldsymbol{\psi}_0$ is evaluated and iteration then commenced in one of the following ways.

(a) using Equations 18 – 20 and continuously updating $\mathbf{K}_T$

(b) proceeding as in (a) but updating the matrix once only and keeping it subsequently constant in the increment

(c) proceeding as in (a) but using the original value of the matrix at the initial conditions of the increment

(d) following the procedure (c) but replacing Equation 20 by an accelerator which gives

$$\Delta\boldsymbol{\delta}_n = \boldsymbol{\alpha}\,\mathbf{K}_T^{-1}\boldsymbol{\psi}_n \tag{21}$$

where $\boldsymbol{\alpha}$ is a diagonal matrix modifying suitably the correction by using 'accumulated experience'. Various methods of arriving at such an acceleration have been proposed and one such technique was devised by Nayak and shown to be particularly effective (References 10, 11, and 12).

A program written on the basis of providing such options is not difficult to arrive at and gives, as its extreme cases, both the simple, incremental process and the Newton–Raphson method. For each class of problem one of the options described provides the optimum solution scheme and has to be determined by experience. For a great majority of problems, (b) and (d) offer the best economy using typically 8–12 load parameter increments.

Haisler et al (Reference 5) make an assessment of nine solution combinations presented above in the context of a few numerical examples.

4. Explicit forms of Lagrangian formulation

Before a finite element solution can be attempted, a more explicit presentation of the various terms arising in the previous two sections is necessary. This will be here presented in the context of a full three dimensional solid with specialization to axisymmetric and plane problems given in Appendix II.

Strain - displacement relationships

The Green's strain in vector notation can be written using the engineering definitions as

$$\epsilon_x = \frac{\partial u}{\partial x} + \frac{1}{2}\left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial x}\right)^2\right]$$

$$\gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} + \left[\frac{\partial u}{\partial x}\cdot\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\cdot\frac{\partial v}{\partial y} + \frac{\partial w}{\partial x}\cdot\frac{\partial w}{\partial y}\right] \tag{22}$$

with

$$\epsilon = [\epsilon_x, \epsilon_y, \epsilon_z, \gamma_{yz}, \gamma_{zx}, \gamma_{xy}]^T = \epsilon^0 + \epsilon^L \tag{23}$$

where ϵ^0 is the usual linear, infinitesimal, strain vector (Reference 13) and ϵ^L is the nonlinear contribution. The nonlinear part is conveniently written as

$$\epsilon^L = \frac{1}{2}\begin{bmatrix} \theta_x^T & 0 & 0 \\ 0 & \theta_y^T & 0 \\ 0 & 0 & \theta_z^T \\ 0 & \theta_z^T & \theta_y^T \\ \theta_z^T & 0 & \theta_x^T \\ \theta_y^T & \theta_x^T & 0 \end{bmatrix} \begin{bmatrix} \theta_x \\ \theta_y \\ \theta_z \end{bmatrix} = \frac{1}{2} A\theta \tag{24}$$

(6x1) = (6x9) x (9x1)

where $\theta_x = \left[\frac{\partial u}{\partial x} \quad \frac{\partial v}{\partial x} \quad \frac{\partial w}{\partial x}\right]^T$ etc.

Rewriting, in the usual manner the approximate expression for displacements Equation 4 as

$$u = \Sigma N_i u_i, \quad v = \Sigma N_i v_i, \quad w = \Sigma N_i w_i, \tag{25}$$

we have after substituting into the expression for ϵ^0, and differentiation

$$d\epsilon^0 = B^0 \, d\delta, \quad \delta_i = [u_i v_i w_i]^T \tag{26}$$

with the well known small displacement matrix given by a typical component submatrix for node i

$$B_i^{o} = \begin{bmatrix} \frac{\partial N_i}{\partial x} & 0 & 0 \\ 0 & \frac{\partial N_i}{\partial y} & 0 \\ 0 & 0 & \frac{\partial N_i}{\partial z} \\ 0 & \frac{\partial N_i}{\partial z} & \frac{\partial N_i}{\partial y} \\ \frac{\partial N_i}{\partial z} & 0 & \frac{\partial N_i}{\partial x} \\ \frac{\partial N_i}{\partial y} & \frac{\partial N_i}{\partial x} & 0 \end{bmatrix} \tag{27}$$

Differentiation of ϵ^L yields

$$d\epsilon^L = \frac{1}{2} dA\theta + \frac{1}{2} A d\theta \tag{28}$$

which due to the structure of the matrices involved becomes simply

$$d\epsilon^L = A d\theta \tag{29}$$

The manipulation of Equation (29) is made easy by an interesting property of **A** and θ (Reference 10).

It is easy to verify that if

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

is a (9 x 1) arbitrary vector then

$$dAx = \begin{bmatrix} d\theta_x^T & 0 & 0 \\ 0 & d\theta_y^T & 0 \\ 0 & 0 & d\theta_z^T \\ 0 & d\theta_z^T & d\theta_y^T \\ d\theta_z^T & 0 & d\theta_x^T \\ d\theta_y^T & d\theta_x^T & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1^T & 0 & 0 \\ 0 & x_2^T & 0 \\ 0 & 0 & x_3^T \\ 0 & x_3^T & x_2^T \\ x_3^T & 0 & x_1^T \\ x_2^T & x_1^T & 0 \end{bmatrix} d \begin{bmatrix} \theta_x \\ \theta_y \\ \theta_z \end{bmatrix} \tag{a}$$

Thus

$$d\mathbf{A}\boldsymbol{\theta} = \mathbf{A}d\boldsymbol{\theta}$$

Similarly if a 6 x 1, vector

$$\mathbf{y} = [y_1, y_2 \ldots\ldots y_6]^T$$

then,

$$d\mathbf{A}^T\mathbf{y} = \begin{bmatrix} d\theta_x & 0 & 0 & 0 & d\theta_z & d\theta_y \\ 0 & d\theta_y & 0 & d\theta_z & 0 & d\theta_x \\ 0 & 0 & d\theta_z & d\theta_y & d\theta_x & 0 \end{bmatrix} \mathbf{y} \quad (b)$$

$$= \begin{bmatrix} y_1 I_3 & y_6 I_3 & y_5 I_3 \\ & y_2 I_3 & y_4 I_3 \\ \text{SYM} & & y_3 I_3 \end{bmatrix} d \begin{bmatrix} \theta_x \\ \theta_y \\ \theta_z \end{bmatrix}$$

Use of this second property given by (b) will be made in Equation 39 .

Now, the (9 x 1) vector $\boldsymbol{\theta}$ in Equation 24 can be written as

$$\boldsymbol{\theta} = \mathbf{G}\boldsymbol{\delta} = [G_1, \ldots G_i \ldots] \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_i \\ \vdots \end{bmatrix} \quad (30)$$

where

$$G_i = \begin{bmatrix} I_3 & \frac{\partial N_i}{\partial x} \\ I_3 & \frac{\partial N_i}{\partial y} \\ I_3 & \frac{\partial N_i}{\partial z} \end{bmatrix}$$

and, I_3, is a 3 x 3 identity matrix.

Substituting Equation 30 into Equation 29 we have finally

$$d\boldsymbol{\epsilon}^L = \mathbf{B}^L d\boldsymbol{\delta} \quad (31)$$

with

$$\mathbf{B}^L = \mathbf{A}\mathbf{G}$$

Thus the strain displacement matrix **B** of Equation 9 now becomes

$$\mathbf{B} = \mathbf{B}^{0} + \mathbf{B}^{L}$$

It can be observed that the total strain of Equation 23 can now be written, using Equations 24, 30, and 32, as

$$\epsilon = \left(\mathbf{B}^{0} + \frac{1}{2}\mathbf{B}^{L}\right)\delta \tag{33}$$

This relation was used by Oden (References 3 and 4) in obtaining a nonsymmetric, secant, stiffness matrix.

For computational purposes, it is convenient to obtain **B** explicitly by multiplying out the appropriate terms in Equation 31. The expression for full three dimensional analysis thus becomes

$$\mathbf{B}_i^{L} = \begin{bmatrix} \frac{\partial u}{\partial x}\cdot\frac{\partial N_i}{\partial x} & \frac{\partial v}{\partial x}\cdot\frac{\partial N_i}{\partial x} & \frac{\partial w}{\partial x}\cdot\frac{\partial N_i}{\partial x} \\ \frac{\partial u}{\partial y}\cdot\frac{\partial N_i}{\partial y} & \frac{\partial v}{\partial y}\cdot\frac{\partial N_i}{\partial y} & \frac{\partial w}{\partial y}\cdot\frac{\partial N_i}{\partial y} \\ \frac{\partial u}{\partial z}\cdot\frac{\partial N_i}{\partial z} & \frac{\partial v}{\partial z}\cdot\frac{\partial N_i}{\partial z} & \frac{\partial w}{\partial z}\cdot\frac{\partial N_i}{\partial z} \\ \frac{\partial u}{\partial z}\cdot\frac{\partial N_i}{\partial y} + \frac{\partial u}{\partial y}\cdot\frac{\partial N_i}{\partial z} & \frac{\partial v}{\partial z}\cdot\frac{\partial N_i}{\partial y} + \frac{\partial v}{\partial y}\cdot\frac{\partial N_i}{\partial z} & \frac{\partial w}{\partial z}\cdot\frac{\partial N_i}{\partial y} + \frac{\partial w}{\partial y}\cdot\frac{\partial N_i}{\partial z} \\ \frac{\partial u}{\partial x}\cdot\frac{\partial N_i}{\partial z} + \frac{\partial u}{\partial z}\cdot\frac{\partial N_i}{\partial x} & \frac{\partial v}{\partial x}\cdot\frac{\partial N_i}{\partial z} + \frac{\partial v}{\partial z}\cdot\frac{\partial N_i}{\partial x} & \frac{\partial w}{\partial x}\cdot\frac{\partial N_i}{\partial z} + \frac{\partial w}{\partial z}\cdot\frac{\partial N_i}{\partial x} \\ \frac{\partial u}{\partial y}\cdot\frac{\partial N_i}{\partial x} + \frac{\partial u}{\partial x}\cdot\frac{\partial N_i}{\partial y} & \frac{\partial v}{\partial y}\cdot\frac{\partial N_i}{\partial x} + \frac{\partial v}{\partial x}\cdot\frac{\partial N_i}{\partial y} & \frac{\partial w}{\partial y}\cdot\frac{\partial N_i}{\partial x} + \frac{\partial w}{\partial x}\cdot\frac{\partial N_i}{\partial y} \end{bmatrix} \tag{34}$$

To obtain the tangential stiffness matrix of Equation 16 we find the increment of the internal force term of Equation 11

$$d\left(\int_V \mathbf{B}^{T}\sigma\, dV\right) = \int_V \left(d\mathbf{B}^{T}\sigma + \mathbf{B}^{T}d\sigma\right) dV \equiv \mathbf{K}_T d\delta \tag{35}$$

As $\mathbf{B}^0$ does not vary with δ we have by taking the variation of Equation 32

$$d\mathbf{B}^T = d[\mathbf{B}^L]^T = \mathbf{G}\, d\, \mathbf{A}^T \tag{36}$$

With the notation defining the tangential material property we can write generally

$$d\sigma = \mathbf{D}_T\, d\epsilon \tag{37}$$

and therefore, after some transformation utilizing the properties of θ matrices described in Equation (b), we arrive at

$$\mathbf{K}_T = \int_V (\mathbf{G}^T\mathbf{M}\mathbf{G} + \mathbf{B}^T\mathbf{D}_T\mathbf{B})\, dV \tag{38}$$

in which

$$\mathbf{M} = \begin{bmatrix} \sigma_x\mathbf{I}_3, & \tau_{xy}\mathbf{I}_3, & \tau_{xz}\mathbf{I}_3 \\ & \sigma_y\mathbf{I}_3, & \tau_{yz}\mathbf{I}_3 \\ \text{SYM} & & \sigma_z\mathbf{I}_3 \end{bmatrix} \tag{39}$$

is a 9 x 9 matrix arising from a rearrangement of the stress terms.

In this derivation we have implied the constitutive law for the material relating directly increments of the Piola-Kirchhoff stresses with increments of the Lagrangian strain. For small strain elasticity this definition is obviously most convenient and indeed the matrix $\mathbf{D}_T$ is the familiar one of elastic constants referred to original coordinates. For large strain elasticity this formulation again can be used as in general energy expressions and are specified in terms of Lagrangian strain (viz. Mooney material) and coefficients of $\mathbf{D}_T$ now become simply the derivatives at the strain energy with respect to the various strain components.

For plasticity involving small strain, the terms of $\mathbf{D}_T$ are well identified (Reference 9) and their derivation will be given in Section 7. For large strain, plasticity constitutive relationships are more complex and the discussion of these is relegated to Appendix IV.

All the ingredients for the numerical calculations are now available for the Lagrangian approach and it will be observed, by comparison with the next section that this is generally most convenient. In particular, specification of anisotropy is easy as this is always referred to axes coinciding with the undeformed structure.

One word should be added concerning the interpretation of stresses. While these for all cases of small strain correspond with real stresses, when large strains used stresses occur in the computation are the Piola–Kirchhoff definitions and require a transformation of the true (Eulerian) stress.

5. Explicit form of Eulerian formulations

Now we continuously update the coordinates to the new system $\bar{\mathbf{x}}$ which now defines the independent variables. The displacements $\bar{\mathbf{u}}$ are still referred to the same

space directions. Strains are now determined by the derivatives of displacements with respect to the updated coordinates and are given explicitly as

$$\begin{aligned} \bar{\epsilon}_x &= \frac{\partial u}{\partial \bar{x}} - \frac{1}{2}\left[\left(\frac{\partial u}{\partial \bar{x}}\right)^2 + \left(\frac{\partial v}{\partial \bar{x}}\right)^2 + \left(\frac{\partial w}{\partial \bar{x}}\right)^2\right] \\ \bar{\gamma}_{xy} &= \frac{\partial u}{\partial \bar{y}} + \frac{\partial v}{\partial \bar{x}} - \left[\frac{\partial u}{\partial \bar{x}} \cdot \frac{\partial u}{\partial \bar{y}} + \frac{\partial v}{\partial \bar{x}} \cdot \frac{\partial v}{\partial \bar{y}} + \frac{\partial w}{\partial \bar{x}} \cdot \frac{\partial w}{\partial \bar{y}}\right] \end{aligned} \tag{40}$$

etc.

With the shape functions N_i now being given in updated coordinates, $\bar{x}$, i.e.

$$N_i \equiv N_i(\bar{x},\bar{y},\bar{z}) \tag{41}$$

we can write

$$\bar{\epsilon} = \bar{\epsilon}^0 + \bar{\epsilon}^L \tag{42}$$

with

$$\bar{\epsilon}^L = -\frac{1}{2}\bar{A}\bar{\theta} \tag{43}$$

where $\bar{A}$ and $\bar{\theta}$ have the same form as in Equation 24 but with differentiation referred now to $\bar{x}$ system of coordinates.

The strains can now be determined for all large displacements; and, if strains are small, we can find stresses as the Eulerian stress definition coincides with that of conventional stress if the material is isotropic. If anisotropy of any kind is present, however, a difficulty arises immediately as the strain directions are now those of the global axes the material properties have to be specified in rotated coordinates. Eulerian formulation is thus not particularly well suited to the study of anisotropic situations.

To determine the residual forces of Equation 10 we need to determine

$$\bar{B} \equiv \bar{B}^0 \tag{44}$$

where $\bar{B}^0$ is given by Equation 27 with derivatives of the shape function being now taken with respect to $\bar{x}$ system.

This very simple fact presents one of the advantages of the Eulerian formulation and providing $\bar{\sigma}$ is known the residuals are determined very simply. Indeed the process is identical with that of small displacement formulation providing the coordinates are adjusted.

The derivation of the tangential stiffness matrix now presents a more complex problem. Proceeding as before we find the increment of the internal force term of Equation 10 i.e.

$$dF = d\left(\int_{\bar{V}} \bar{B}^{oT}\bar{\sigma}\, d\bar{V}\right) = \int_{\bar{V}}\left(d\bar{B}^{oT}\bar{\sigma}\, d\bar{V} + \right.$$

$$\left. + \bar{B}^{oT} d\bar{\sigma}\, d\bar{V} + \bar{B}^{oT}\bar{\sigma}\, d(d\bar{V})\right) \tag{45}$$

noting that now all terms, being functions of $\bar{x}$, depend on u and hence on the nodal parameters δ. Before proceeding further we must define the stress-strain relationships. These are complicated by the fact that increments of Eulerian (Cauchy) stress and strain cannot be related by a constitutive relation. It is therefore necessary to introduce Jaumann stress increments (see Appendix I). Denoting this stress by $\bar{\sigma}_J$ we can write the relation between the two stresses as

$$d\bar{\sigma}_J = d\bar{\sigma} + d\bar{T}_\sigma\bar{\sigma} \tag{46}$$

and the constitutive relationship must be given by

$$d\bar{\sigma}_J = \bar{D}_T d\bar{\varepsilon} \tag{47}$$

where, for an isotropic material the tangential matrix of constants the same as are used in Equation 37 but in general this now referred to a direction changing with each increment.

Substituting into Relationship 45 we find that this can be written as

$$dF = (\bar{K}_\sigma + \bar{K}_o)\, d\delta \tag{48}$$

in which often the predominant term is

$$\bar{K}_o \int_{\bar{V}} \bar{B}^o\, \bar{D}_T \bar{B}^o\, d\bar{V} \tag{49}$$

The expression $\bar{K}_\sigma$ is usually very small numerically and we find can be safely omitted from the calculations. In Appendix III we discuss some of the complexities of its form but as the equilibrium check is performed exactly, we conclude that the approximate tangential matrix in its very simple form is adequate for practical use.

6. Special features of the isoparametric formulation

The formulation and numerical integration of the isoparametric element has been widely discussed and is given in detail elsewhere (References 13, 14, and 15). Summarizing succinctly, we note that the Equation 4 which represents the parametric variation of displacements is written in component form as

$$u = \Sigma N_i u_i, \quad v = \Sigma N_i v_i, \quad w = \Sigma N_i w_i \tag{50}$$

where u_i, v_i and w_i are the parameters associated with the nodes, and N_i are scalar shape functions associated with a curvilinear coordinate system ξ, η, ζ

$$N_i \equiv N_i(\xi, \eta, \zeta) \tag{51}$$

The relationships of the curvilinear and cartesian coordinates are given in an identical form with

$$\xi = \Sigma N_i \xi_i \qquad \eta = \Sigma N_i \eta_i \qquad \zeta = \Sigma N_i \zeta_i$$

and

$$x = \Sigma N_i x_i \ , \qquad y = \Sigma N_i y_i \ , \qquad z = \Sigma N_i z_i \tag{52}$$

with x_i, y_i and z_i being the nodal coordinates in the global, cartesian system.

The derivatives of N_i with respect to the two coordinate systems are related by

$$\left\{\frac{\partial N_i}{\partial x}\right\}_i = \mathbf{J}_c^{-1} \left\{\frac{\partial N_i}{\partial \xi}\right\}_i \tag{53}$$

where

$$\left\{\frac{\partial N_i}{\partial \xi}\right\}_i = \left[\frac{\partial N_i}{\partial \xi}, \quad \frac{\partial N_i}{\partial \eta}, \quad \frac{\partial N_i}{\partial \zeta}\right]^T$$

$$\left\{\frac{\partial N}{\partial x}\right\}_i \qquad \left[\frac{\partial N_i}{\partial x} \quad \frac{\partial N_i}{\partial y}, \quad \frac{\partial N_i}{\partial z}\right]^T \tag{54}$$

and

$$\mathbf{J}_c = \left[\frac{\partial (x,y,z)}{\partial (\xi,\eta,\zeta)}\right]$$

is the 3 x 3 Jacobian coordinate transformation matrix. This has to be evaluated at each integrating point and inverted to obtain the Cartesian derivatives.

It is of interest to observe the structure of this Jacobian. On substitution of Relationship 52, we observe that it can be written as

$$\mathbf{J}_c = \begin{bmatrix} \mathbf{G}_\xi^T \mathbf{X} & \mathbf{G}_\eta^T \mathbf{X} & \mathbf{G}_\zeta^T \mathbf{X} \\ \mathbf{G}_\xi^T \mathbf{Y} & \mathbf{G}_\eta^T \mathbf{Y} & \mathbf{G}_\zeta^T \mathbf{Y} \\ \mathbf{G}_\xi^T \mathbf{Z} & \mathbf{G}_\eta^T \mathbf{Z} & \mathbf{G}_\zeta^T \mathbf{Z} \end{bmatrix} \tag{56}$$

with

$$\mathbf{G}_\xi^T = \left[\frac{\partial N_1}{\partial \xi}, \quad \frac{\partial N_2}{\partial \xi}, \ldots \frac{\partial N_i}{\partial \xi} \ldots\right]$$

$$\mathbf{X}^T = [x_1, \ x_2, \ldots, \ x_i, \ldots] \tag{57}$$

etc.

This form allows the most convenient numerical evaluation and is useful to note in the context of programming.

When proceeding to evaluate the **B** matrices in the context of Lagrangian analysis, it is useful to observe that the separation into small and large strain components is not, in fact, convenient. Thus adding Expressions 27 and 34 we have on noting that $\bar{x} = x+u$, $\bar{y} = y+v$, $\bar{z} = z+w$ the form

$$\mathbf{B}_i = \begin{bmatrix} \frac{\partial \bar{x}}{\partial x}\frac{\partial N_i}{\partial x} & \frac{\partial \bar{y}}{\partial x}\frac{\partial N_i}{\partial x} & \frac{\partial \bar{z}}{\partial x}\frac{\partial N_i}{\partial x} \\ \frac{\partial \bar{x}}{\partial y}\frac{\partial N_i}{\partial y} & \frac{\partial \bar{y}}{\partial y}\frac{\partial N_i}{\partial y} & \frac{\partial \bar{z}}{\partial y}\frac{\partial N_i}{\partial y} \\ \frac{\partial \bar{x}}{\partial z}\frac{\partial N_i}{\partial z} & \frac{\partial \bar{y}}{\partial z}\frac{\partial N_i}{\partial z} & \frac{\partial \bar{z}}{\partial z}\frac{\partial N_i}{\partial z} \\ \frac{\partial \bar{x}}{\partial z}\frac{\partial N_i}{\partial y} + \frac{\partial \bar{x}}{\partial y}\frac{\partial N_i}{\partial z} & \frac{\partial \bar{y}}{\partial z}\frac{\partial N_i}{\partial y} + \frac{\partial \bar{y}}{\partial y}\frac{\partial N_i}{\partial z} & \frac{\partial \bar{z}}{\partial z}\frac{\partial N_i}{\partial y} + \frac{\partial \bar{z}}{\partial y}\frac{\partial N_i}{\partial z} \\ \frac{\partial \bar{x}}{\partial x}\frac{\partial N_i}{\partial z} + \frac{\partial \bar{x}}{\partial z}\frac{\partial N_i}{\partial x} & \frac{\partial \bar{y}}{\partial x}\frac{\partial N_i}{\partial z} + \frac{\partial \bar{y}}{\partial z}\frac{\partial N_i}{\partial x} & \frac{\partial \bar{z}}{\partial x}\frac{\partial N_i}{\partial z} + \frac{\partial \bar{z}}{\partial z}\frac{\partial N_i}{\partial x} \\ \frac{\partial \bar{x}}{\partial y}\frac{\partial N_i}{\partial x} + \frac{\partial \bar{x}}{\partial x}\frac{\partial N_i}{\partial y} & \frac{\partial \bar{y}}{\partial y}\frac{\partial N_i}{\partial x} + \frac{\partial \bar{y}}{\partial x}\frac{\partial N_i}{\partial y} & \frac{\partial \bar{z}}{\partial y}\frac{\partial N_i}{\partial x} + \frac{\partial \bar{z}}{\partial x}\frac{\partial N_i}{\partial y} \end{bmatrix} \tag{58}$$

With $\frac{\partial N_i}{\partial x}$ etc. already known, the derivatives $\frac{\partial \bar{x}}{\partial x}$ etc. have now to be found.

These again are components of a Jacobian matrix

$$\mathbf{J} \equiv \left[\frac{\partial\,(\bar{x},\bar{y},\bar{z})}{\partial\,(x,y,z)}\right] \tag{59}$$

which we shall call the deformation Jacobian matrix and which is given by a similar form to the coordinate Jacobian matrix already established.

$$\mathbf{J} = \begin{bmatrix} \mathbf{G}_x^T \bar{\mathbf{X}} & \mathbf{G}_y^T \bar{\mathbf{X}} & \mathbf{G}_z^T \bar{\mathbf{X}} \\ \mathbf{G}_x^T \bar{\mathbf{Y}} & \mathbf{G}_y^T \bar{\mathbf{Y}} & \mathbf{G}_z^T \bar{\mathbf{Y}} \\ \mathbf{G}_x^T \bar{\mathbf{Z}} & \mathbf{G}_y^T \bar{\mathbf{Z}} & \mathbf{G}_z^T \bar{\mathbf{Z}} \end{bmatrix} \tag{60}$$

where $$\mathbf{G}_x^T = \begin{bmatrix} \frac{\partial N_1}{x} & \frac{\partial N_2}{x} & \ldots & \frac{\partial N_i}{x} \end{bmatrix} \ldots \tag{61}$$

and $\mathbf{X}^T = [\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_i \ldots]$ etc.

in which $\bar{x}_i = x_i + u_i$ represent simply the updated nodal coordinates.

As the vectors $\mathbf{G}_x$ etc. given by Equation 61 have already been calculated, it is a simple matter to calculate the new coordinates and thus all the terms of the deformation; Jacobian matrix.

With the value of the **B** matrix known, all the residual forces can now be found and also the second part of the tangential matrix given by Equation 38 evaluated.

In evaluating this part of the stiffness matrix, the economics conventional in linear analysis which take account of zero terms cannot, in general, now be used (Reference 10). However, the evaluation of the first term of the tangential stiffness relation (the initial stress matrix given by Equation 38)automatically results in many such zeros, as will be noted from the Expressions 30 and 39 in which the identity matrix figures prominently. Here it is most convenient to do the multiplications explicitly. It can then be simply verified that the typical terms become simply

$$\begin{aligned} \mathbf{G}_i^T \mathbf{M} \mathbf{G}_j \equiv \mathbf{I}_3 \Big(& \sigma_x \frac{\partial N_i}{\partial x} \cdot \frac{\partial N_j}{\partial x} + \sigma_y \frac{\partial N_i}{\partial y} \cdot \frac{\partial N_j}{\partial y} \\ & + \sigma_z \frac{\partial N_i}{\partial z} \cdot \frac{\partial N_j}{\partial z} + \tau_{xy} \left(\frac{\partial N_i}{\partial x} \cdot \frac{\partial N_j}{\partial y} + \frac{\partial N_i}{\partial y} \cdot \frac{\partial N_j}{\partial x} \right) \\ & + \tau_{yz} \left(\frac{\partial N_i}{\partial y} \cdot \frac{\partial N_j}{\partial z} + \frac{\partial N_i}{\partial z} \cdot \frac{\partial N_j}{\partial y} \right) + \tau_{zx} \left(\frac{\partial N_i}{\partial z} \cdot \frac{\partial N_j}{\partial x} + \frac{\partial N_i}{\partial x} \cdot \frac{\partial N_j}{\partial z} \right) \end{aligned} \tag{62}$$

This particular structure of the initial stress matrix is of considerable importance when numerical integration is used where, for the basic element stiffness the full multiplication of submatrices has to be repeated at each integrating point. The number of operations is vastly reduced if the advantage of $\mathbf{I}_3$ in the explicit multiplication carried out in Equation 62 is used. For instance, in a solid isoparametric element, the number of operations at each Gauss point is 1830 if full multiplication is carried out, while with the explicit expression only 210 operations are needed.

For the Eulerian formulation, the calculations proceed in a parallel manner. Indeed it will be noted that the transition from one system to the other is particularly simple in the isoparametric formulation.

Now the shape functions are implied as functions of the updated coordinates $\bar{x}$ or their curvilinear equivalents ξ.

If the nodal coordinates are continuously updated the calculation follow precisely the same pattern as used in small displacements - infinitesimal strain analysis where evaluation of $\bar{B}^0$, volume elements and of the approximate stiffness matrix of Equation 49 is concerned.

The coordinate transformation Jacobian is now found from Equations 53 to 57 by replacing x_i with $\bar{x}_i$ etc., as is implied in the previous statements.

Thus, for instance

$$\bar{J}_c = \begin{bmatrix} G_\xi^T \bar{X} & G_\eta^T \bar{X} & G_\zeta^T \bar{X} \\ G_\xi^T \bar{Y} & G_\eta^T \bar{Y} & G_\zeta^T \bar{Y} \\ G_\xi^T \bar{Z} & G_\eta^T \bar{Z} & G_\zeta^T \bar{Z} \end{bmatrix} \tag{63}$$

with $\bar{X}$ etc., defined by Equation 61 .

The matrix $[B^0]$ is simply given by substitution of $\bar{X}$ coordinate with Equation 27 i.e.

$$\bar{B}_i^0 = \begin{bmatrix} \frac{\partial N_i}{\partial \bar{x}}, & 0 & 0 \\ 0 & \frac{\partial N_i}{\partial \bar{y}}, & 0 \\ 0 & 0 & \frac{\partial \bar{N}_i}{\partial \bar{z}} \\ 0 & \frac{\partial N_i}{\partial \bar{z}} & \frac{\partial N_i}{\partial \bar{y}} \\ \frac{\partial N_i}{\partial \bar{z}} & 0 & \frac{\partial N_i}{\partial \bar{x}} \\ \frac{\partial N_i}{\partial \bar{y}} & \frac{\partial N_i}{\partial \bar{x}} & 0 \end{bmatrix} \tag{64}$$

with the derivatives given similarly to Equation 53.

$$\left[\frac{\partial N_i}{\partial \bar{x}}\right]_i = \bar{J}_c^{-1} \left[\frac{\partial N_i}{\partial \xi}\right]_i \tag{65}$$

The only outstanding difference from simply pursuing a small strain analysis with continually updated coordinates, is in the determination of strain and stress components as already pointed out in the previous section where the Jaumann stress increment was introduced.

The case of a linear elastic and isotropic material with small strains but large deformation is of particular interest. Now one can establish the Eulerian stress $\bar{\sigma}$ directly from the strains by the usual relation.

$$\bar{\sigma} = \bar{\mathbf{D}}\,\bar{\epsilon} \tag{66}$$

where $\bar{\epsilon}$ is the Eulerian strain found by Equation 40 and in which $\bar{\mathbf{D}}$ is the matrix of usual elastic constants. These strains defined in Expression 40 are most conveniently found by using a new deformation Jacobian

$$\bar{\mathbf{J}} \equiv \left[\frac{\partial (x,y,z)}{\partial (\bar{x},\bar{y},\bar{z})}\right] \tag{67}$$

which is found by Expressions 60 to 61 with the bar symbols interchanged. Now vectors X listing the original coordinates are of importance. The same expressions used for evaluating Lagrangian strains are used with the above substitution to obtain Eulerian strains. i.e.

$$\bar{\mathbf{J}} = \begin{bmatrix} \bar{\mathbf{G}}_x \mathbf{x}, \; \bar{\mathbf{G}}_y \bar{\mathbf{X}}, \; \ldots\ldots \\ \ldots\ldots, \; \ldots\ldots, \; \ldots\ldots \\ \ldots\ldots, \; \ldots\ldots \; \ldots\ldots \end{bmatrix} \tag{68}$$

$$\bar{\mathbf{B}}_i = \begin{bmatrix} \dfrac{\partial x}{\partial \bar{x}} \quad \dfrac{\partial N_i}{\partial \bar{x}}, \; \ldots\ldots \\ \ldots\ldots, \; \ldots\ldots \\ \ldots\ldots, \; \ldots\ldots \\ \ldots\ldots, \; \ldots\ldots \\ \text{etc.} \\ \text{as in } 58 \end{bmatrix} \tag{69}$$

7. The tangential modulus matrix for plasticity

In previous sections we have introduced the concept of the tangential modulus matrix $\mathbf{D}_T$. For elastic problems this is well defined. For various nonlinear constitutive relations this matrix is perhaps not obvious and on occasion may not be available.

The case of plastic deformation is of particular interest and here indeed the specification of material behaviour is most conveniently given by this matrix and the corresponding deformation Equation 37 i.e.

$$d\sigma = \mathbf{D}_T\, d\epsilon \tag{37}$$

in which Lagrangian strains are raised. Although the matrix has been derived earlier (Reference 9) in context of small strain, it is convenient to present here a more compact derivation (References 10 and 12) as this will be used in the numerical examples following and is also applicable for large displacements (Reference 16).

Dividing the total increment of strain into its elastic and plastic components we can write

$$d\epsilon = d\epsilon^e + d\epsilon^p \tag{70}$$

The yield surface is generally given as

$$F(\sigma, K) = 0 \quad \text{and the plastic potential} \qquad Q(\sigma, K) = 0 \tag{71}$$

where σ now stood for the Piola–Kirchhoff stress are κ for a hardening parameter. As elastic strains associative with plasticity are always small, we can write

$$d\epsilon^e = \mathbf{D}\, d\sigma \tag{72}$$

For plastic strains the normality rule requires that

$$d\epsilon^p = d\lambda\, \mathbf{a} \tag{73}$$

where

$$\mathbf{a}^T = \left[\frac{\partial Q}{\partial \sigma_{x'}} \ldots .\right]$$

and $d\lambda$ is an undetermined positive proportionality constant.

During plastic deformation the stress remains on the yield surface and thus

$$d\,F = \mathbf{a}^T\, d\sigma - \mathbf{A}\, d\,\lambda = 0$$

where

$$\mathbf{a}^T = \left[\frac{\partial F}{\partial \sigma_{x'}} \ldots .\right]$$

and

$$A = -\frac{1}{d\lambda}\,\frac{\partial F}{\partial \kappa}\, d\kappa$$

For associative plasticity $\mathbf{a} = \dot{\epsilon}$ and if the yield surface is defined in terms of a uniaxial given stress A is equal to the slope of the uniaxial stress - plastic strain; curve.

Substituting Equation 72 and 73 with Equation 70 we have

$$d\epsilon = d\lambda \mathbf{a} + \mathbf{D}^{-1}\, d\sigma \tag{75}$$

and on premultiplying by $\mathbf{a}^T\mathbf{D}$ and using Equation 75 to eliminate $d\sigma$

$$\mathbf{a}^T\mathbf{D}\, d\sigma = A d\lambda + \mathbf{a}^T\mathbf{D}\dot{\mathbf{a}}\, d\lambda \tag{76}$$

From above $d\lambda$ can be found. Substitution of this into Equation 75 gives now the tangential matrix

$$\mathbf{D}_T = \mathbf{D} - \mathbf{D}_p \tag{77}$$

with

$$\mathbf{D}_p = \mathbf{D}\dot{\mathbf{a}}\, \mathbf{D}\mathbf{a}^T/(A + \mathbf{a}^T\mathbf{D}\dot{\mathbf{a}})$$

This matrix can be evaluated explicitly at all steps of the numerical computation providing $d\lambda$ is a positive quantity (if not purely elastic unloading takes place). It is thus necessary to compute $d\lambda$ at all stages of calculation.

In Reference 12 various yield surfaces are described in detail and the discussion of these in here is unnecessary. It is of interest to observe that the tangential modulus matrix loses its symmetry in the case of non-associative laws and therefore some additional computations or difficulties arise in such a case.

To illustrate the applicability of the processes outlined, several examples are quoted.

8.1 Thick cantilever – large elastic deformation – Figure 2

Here Lagrangian formulation is used and five parabolic isoparametric elements approximate to the beam. Linear elastic laws are used despite the appreciable strain due to the thick shape of the cantilever.

Comparison with solution with an analysis Bisshopp and Drucker (Reference 17) is given although the latter does not take the shear deformation or Poisson's ratio into account.

It is of interest to remark that two (and three) dimensional approaches of the kind outlined here have been used with success for much thinner sections and indeed, give an alternative approach to plate and shell problems (Reference 18). If thin sections are used, however, a reduction of integration order must be used to avoid spurious shear stiffness (Reference 19). Indeed, such reduced integration orders are of general applicability and improve overall performance as will be shown later.

8.2 Shallow spherical cap – large elastic deformation – Figure 3

The example is now axisymmetric and again treated as a simple solid body of revolution with 12 parabolic elements. Comparison is given with results of Haisler et al. (Reference 5) and shows good agreement with their most precise solution despite the fact this rather large increment was used here (all increments are shown in an accompanying figure).

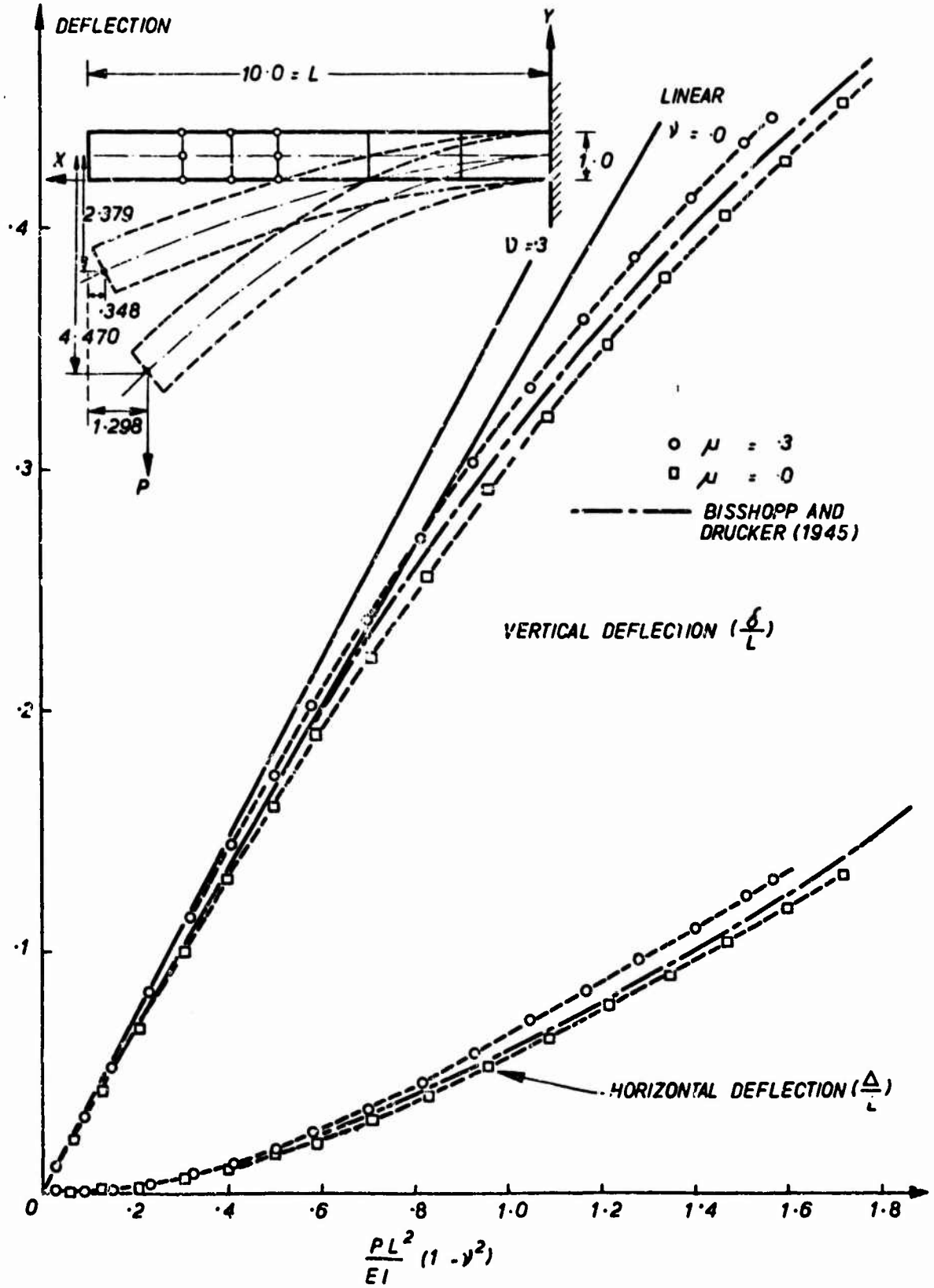


FIG. 2 LARGE DEFLECTIONS OF A CANTILEVER BEAM (ELASTICA PROBLEM)

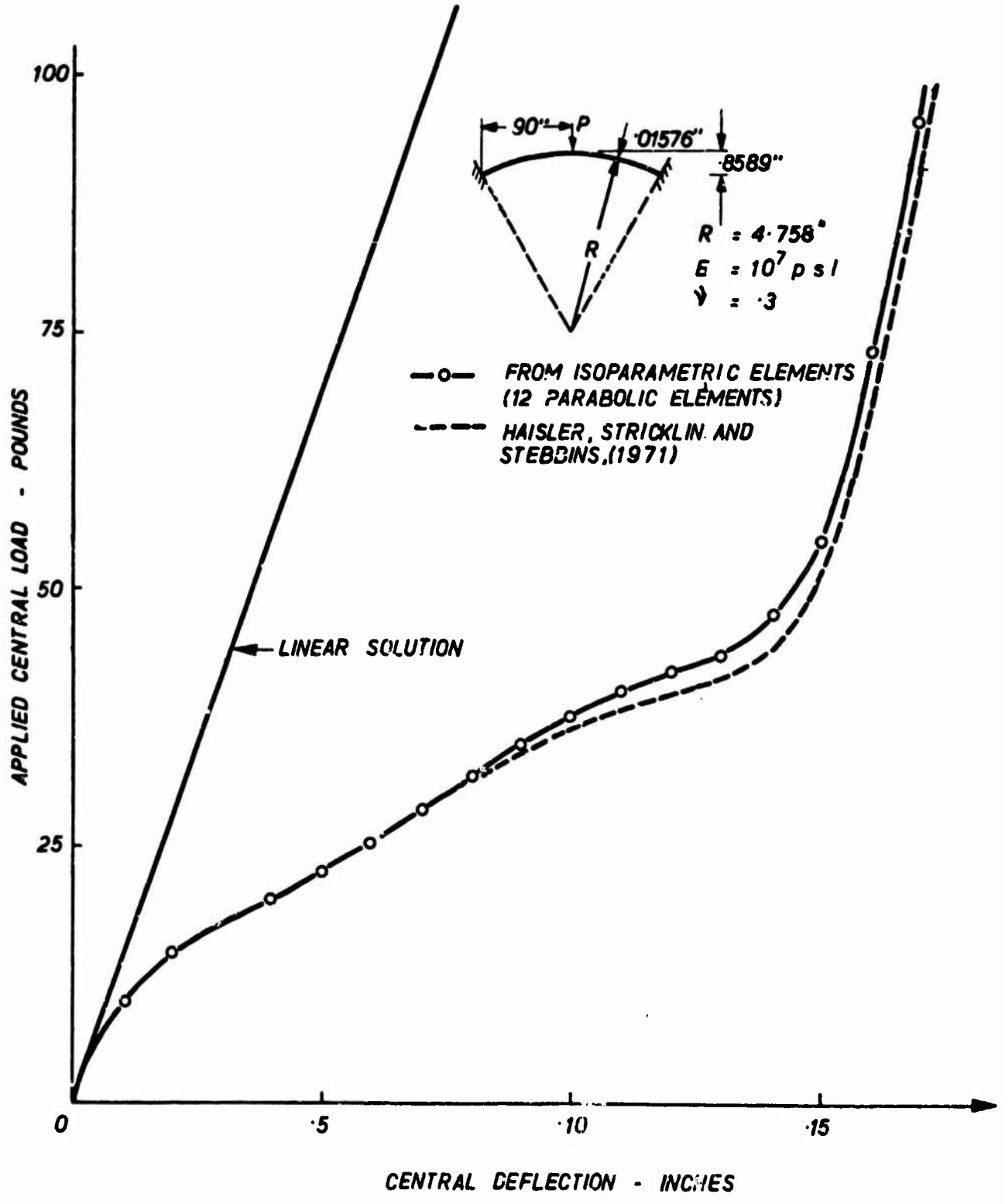


FIG. 3 LOAD DEFLECTION CURVE - SHALLOW SPHERICAL CAP

It was found in this and indeed other similar examples that the number of iterative steps in each increment is reduced considerably if displacement rather than load is incremented. This indeed is the only way of achieving solutions beyond or at the peak of load as seen in the next examples.

Displacement incrementation presents no difficulties for a single load – with several proportional loads an artifice requiring two solutions at each step must be introduced and this mitigates somewhat against the advantages of displacement increments (References 20 and 21).

8.3 Shallow arch – large elastic deformation – Figures 4 and 5

In this problem displacement of the central point was incremented to allow a study of the snap through behaviour. Excellent comparison with results quoted by Biezeno and Grammel (Reference 22) is achieved.

8.4 Bellows – Figure 6

This axisymmetric example is given to illustrate a practical application in which the nonlinear load deformation characteristics needed to be predicted.

8.5 Thick cylinder – Elastic-plastic Behaviour – Figures 7 and 8a, b, and c

This fairly trivial problem is solved to compare in some detail the efficiency of the program with every solution. Also, the differences between this application of Von Mises and Tresca yield conditions are highlighted.

Three parabolic isoparametric elements are used and integration is carried out for 4 x 4, 3 x 3, and 2 x 2 Gauss points. It is worth remarking that the latter given in this case has no detectible difference in deformation but shows a considerable improvement in stresses.

Comparison is made with results derived by Hodge and White (Reference 23).

8.6 Axisymmetric Extrusion – Plastic 'Failure' – Figures 9 and 10

An axisymmetric problem of extrusion is approached here via the solution of a fictitious elastic-plastic one of the same configuration. The development of plastic zones shown, as the displacement of the plunger is incremented, is only of academic interest. The "collapse" load at which deformation progresses without a further load increase does however coincide with the steady state extrusion process.

For this computation Johnson and Mellor (Reference 27) have obtained a solution by an approximate process. It is remarkable how closely the finite element solution approaches their predicted experimental results despite the poorness of the mesh of parabolic elements employed.

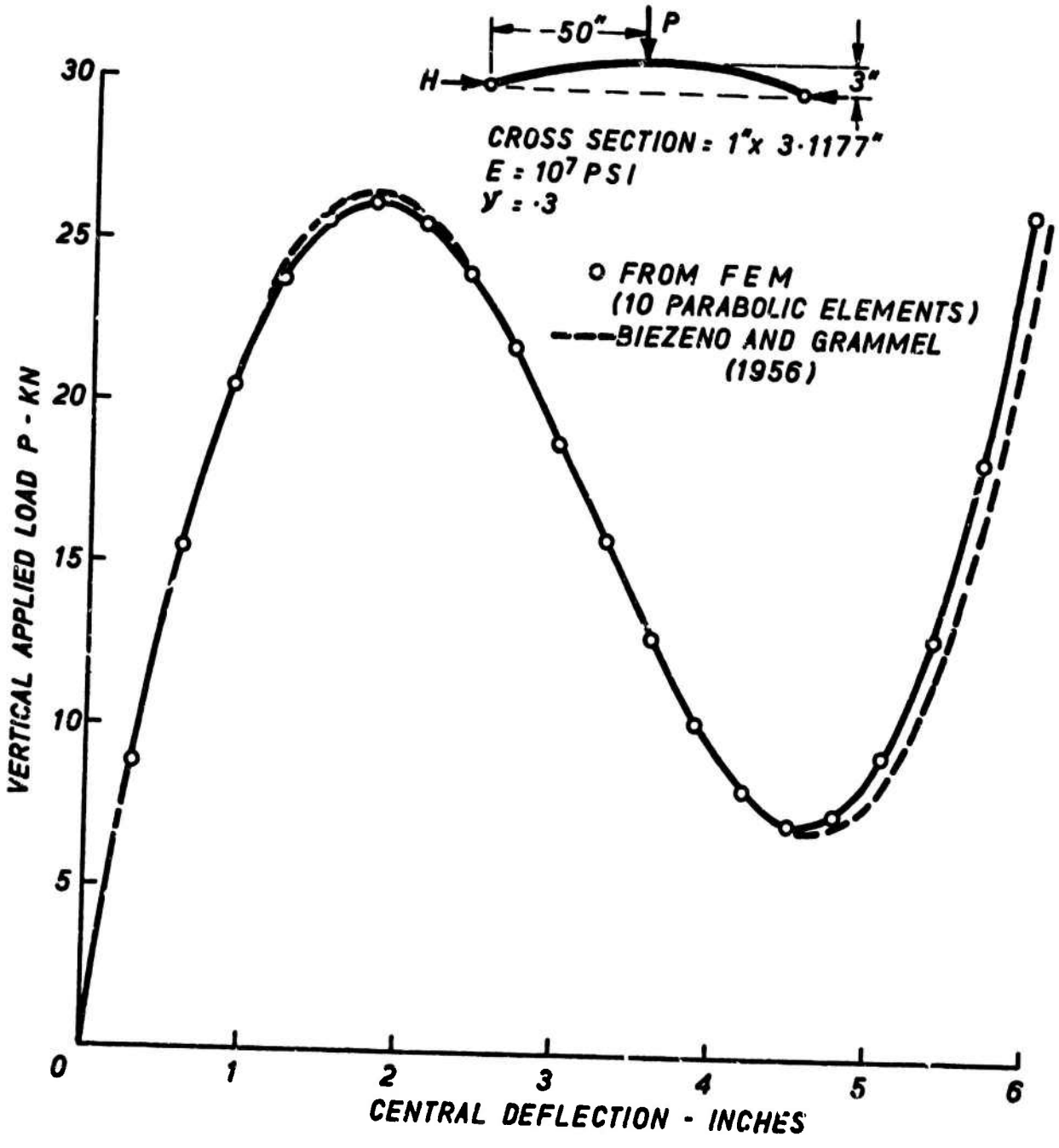


FIG. 4 A PINNED CIRCULAR SHALLOW ARCH
LOAD - CENTRAL DEFLECTION PLOT

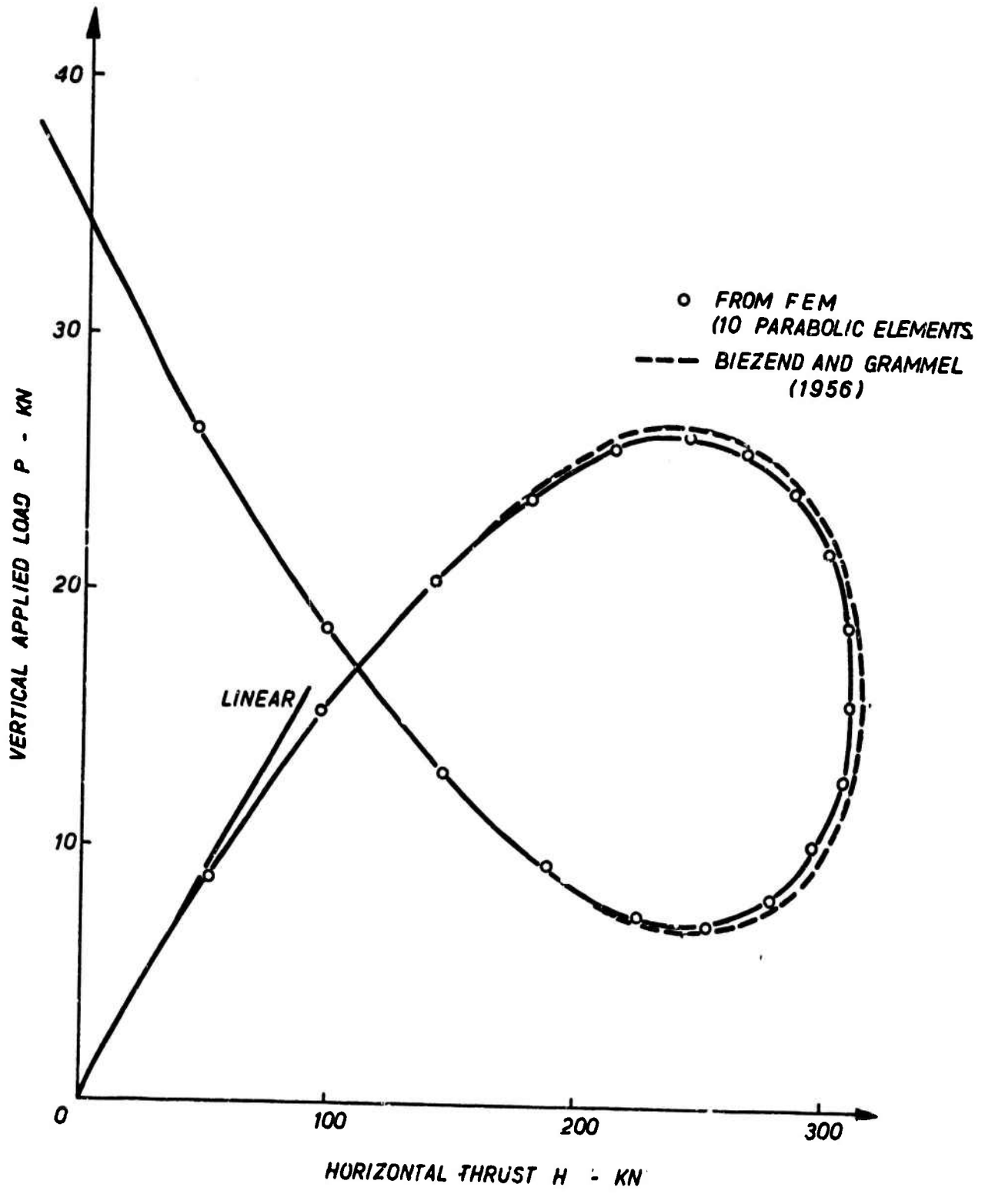


FIG. 5 A PINNED CIRCULAR SHALLOW CAP
APPLIED LOAD - HORIZONTAL THRUST CURVE

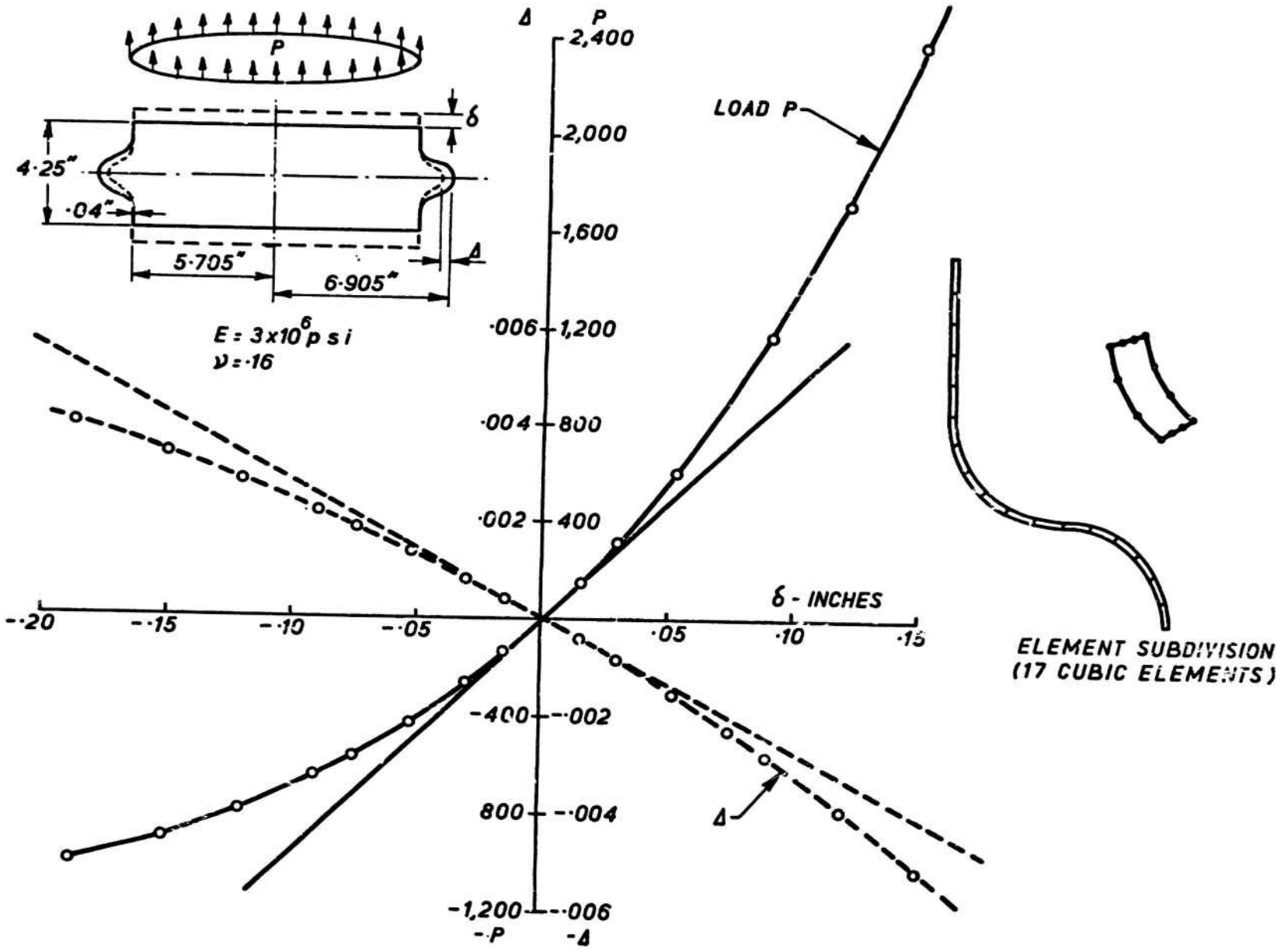


FIG 6 A BELLOWS JUNCTION. LOAD-DEFLECTION CURVES

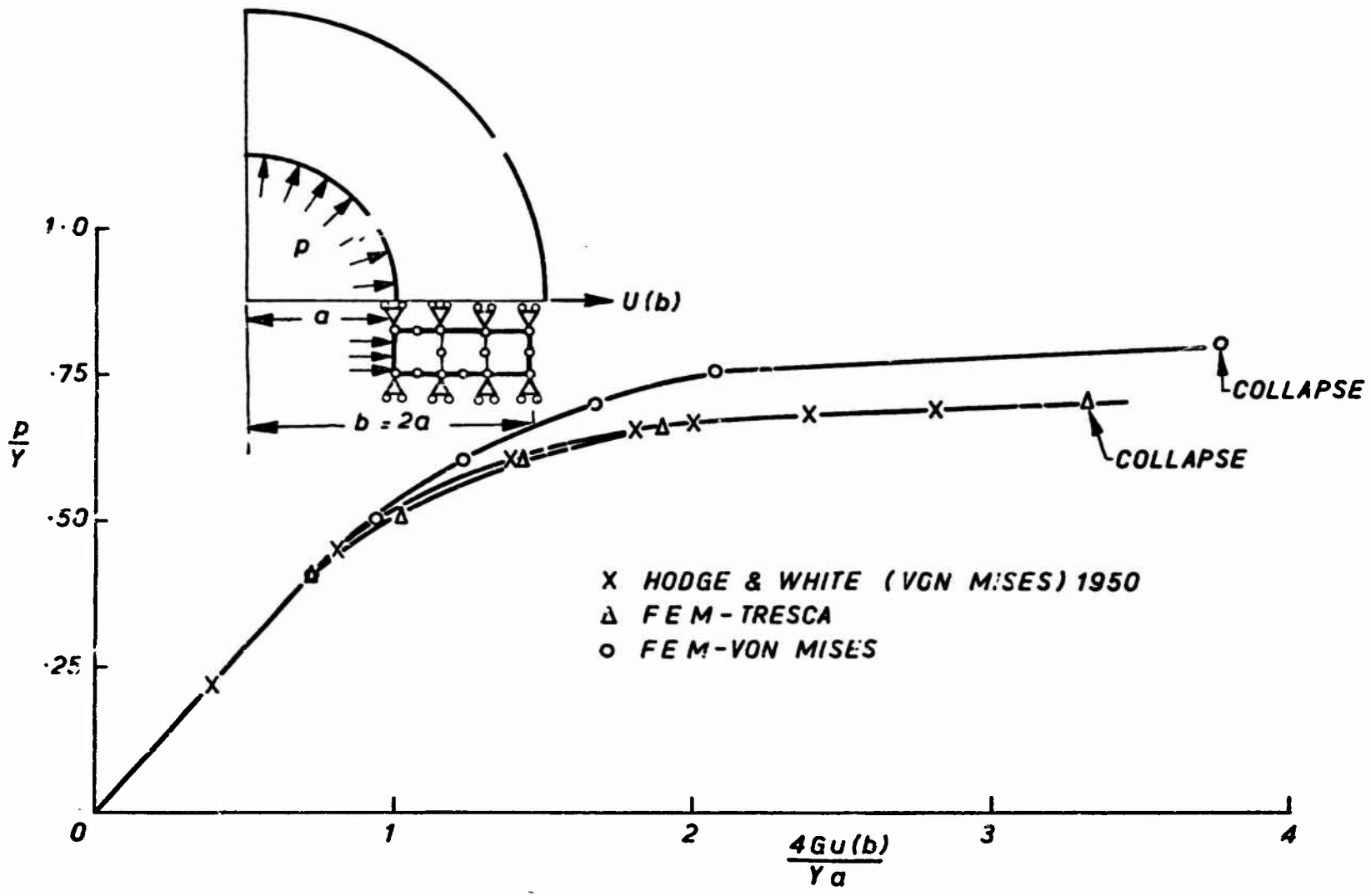


FIG. 7 THICK CYLINDER-(NO AXIAL STRAIN) NO DIFFERENCE OBTAINED FOR 2x2, 3x3 OR 4x4 GAUSS POINTS
Y-UNIAXIAL YIELD STRESS, G-SHEAR MODULUS

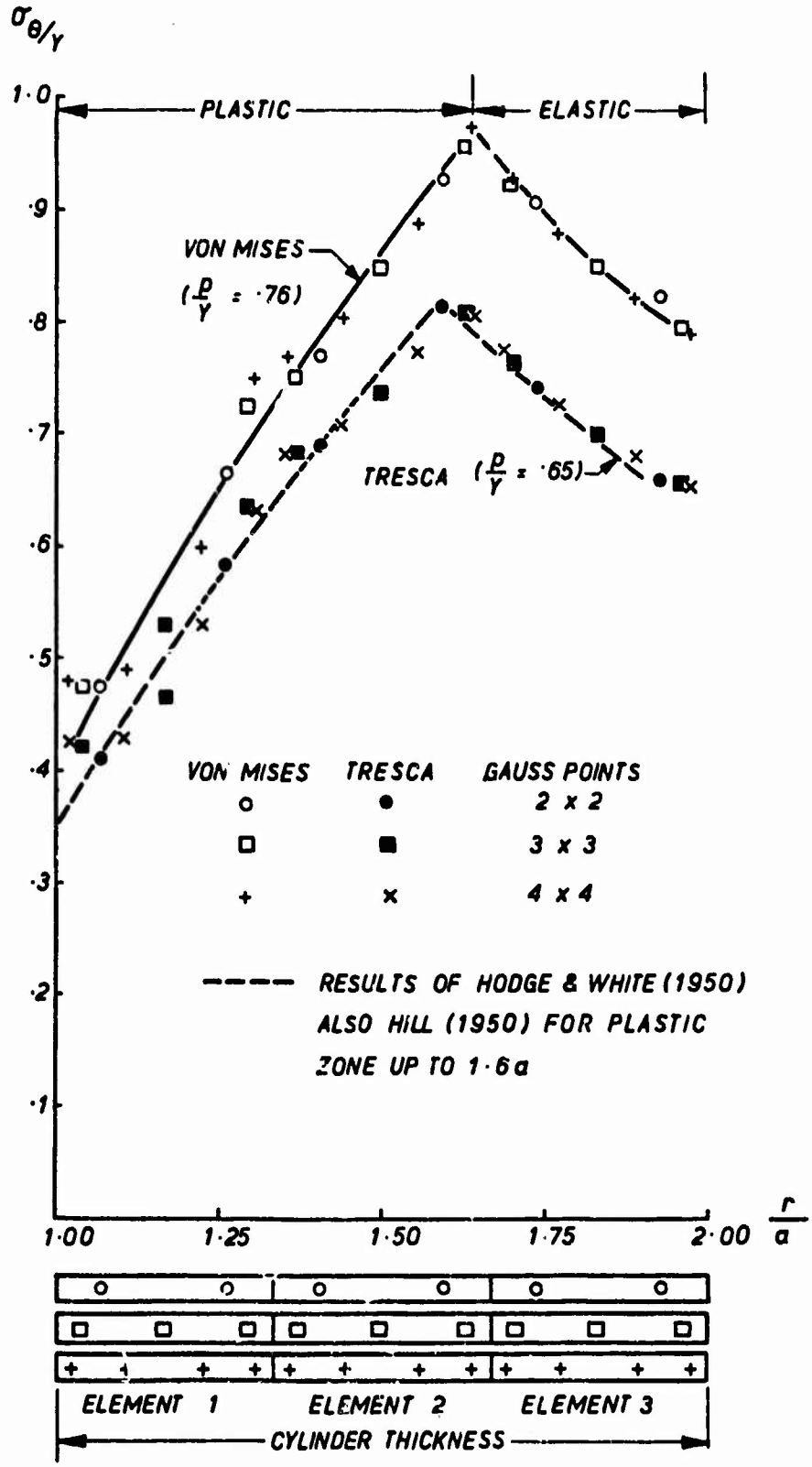


FIG. 8a THICK CYLINDER (NO AXIAL STRAIN)
HOOP STRESS DISTRIBUTION

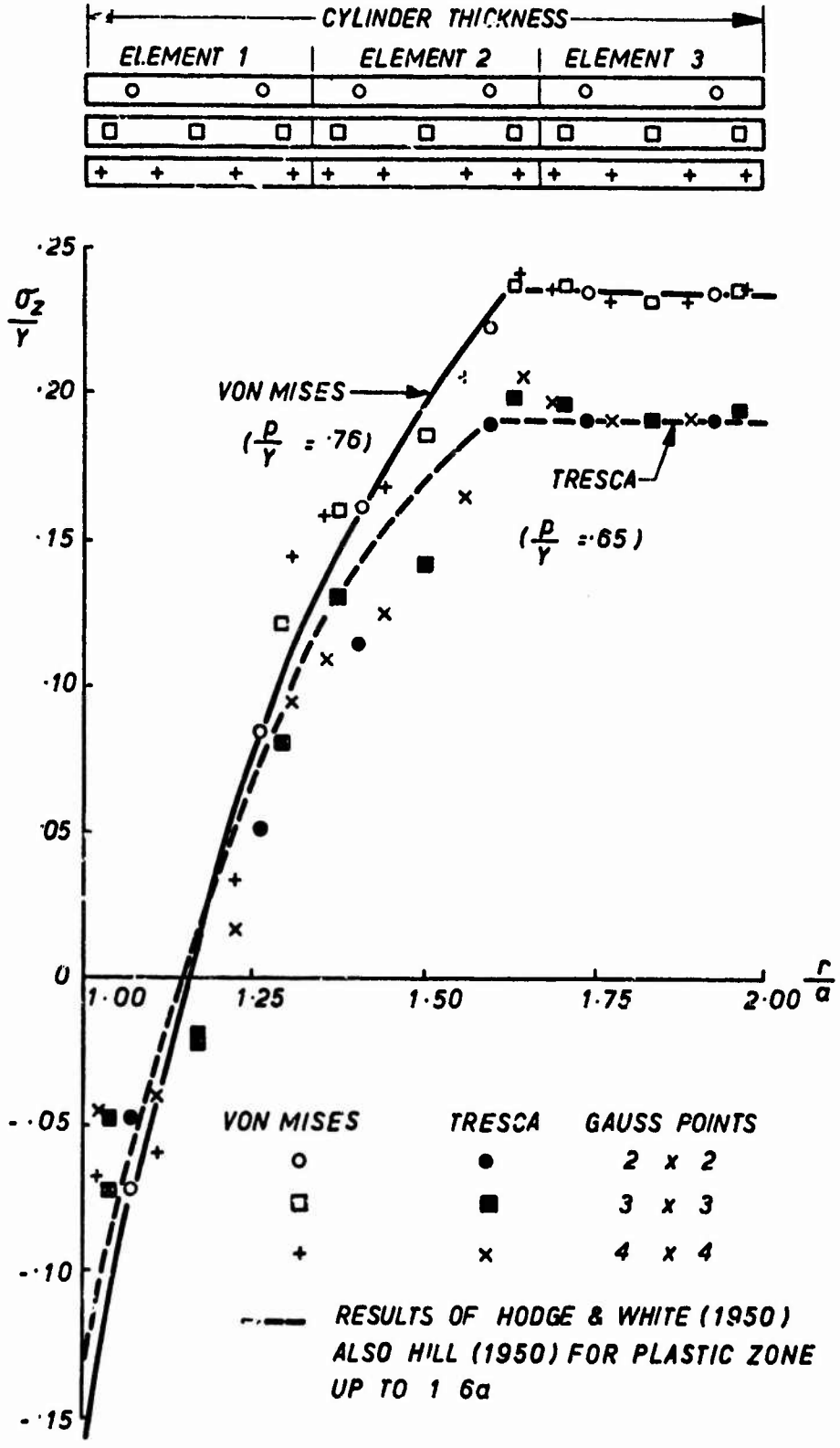


FIG. 8b THICK CYLINDER (NO AXIAL STRAIN)
AXIAL STRESS DISTRIBUTION

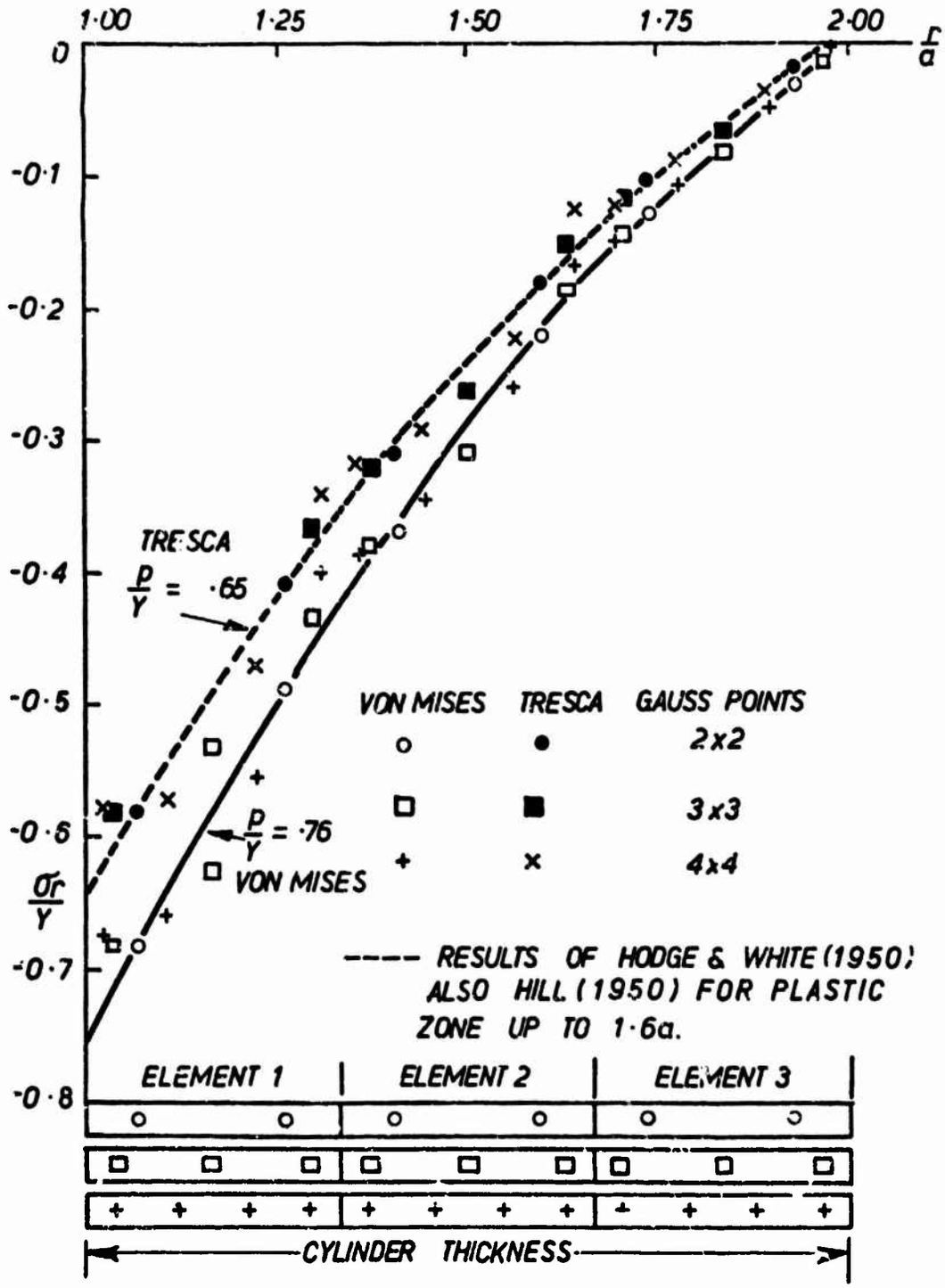


FIG. 8c THICK CYLINDER (NO AXIAL STRAIN) RADIAL STRESS DISTRIBUTION.

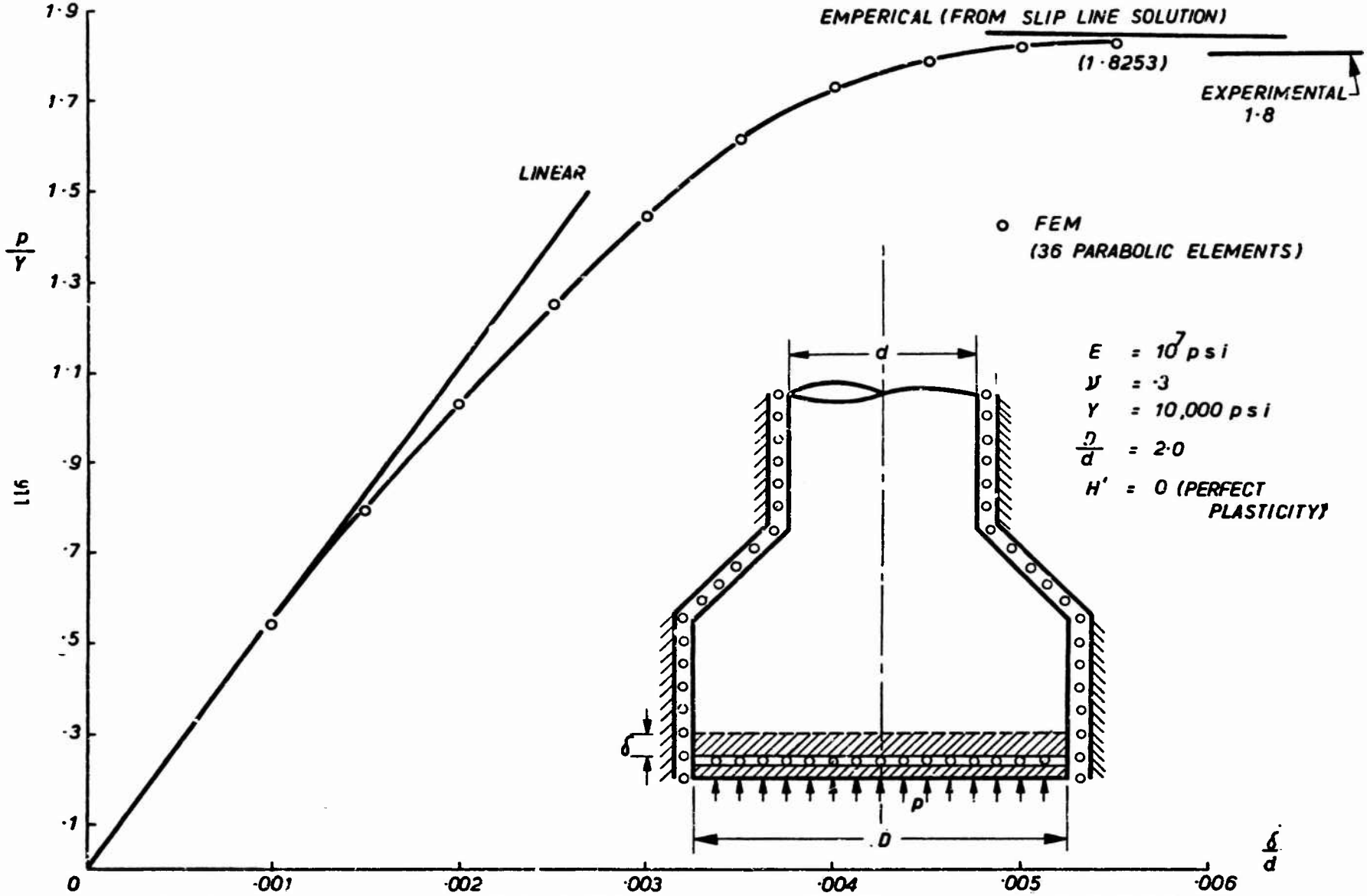


FIG. 9 AN AXISYMETRIC EXTRUSION PROBLEM, MEAN PRESSURE - DISPLACEMENT CURVE

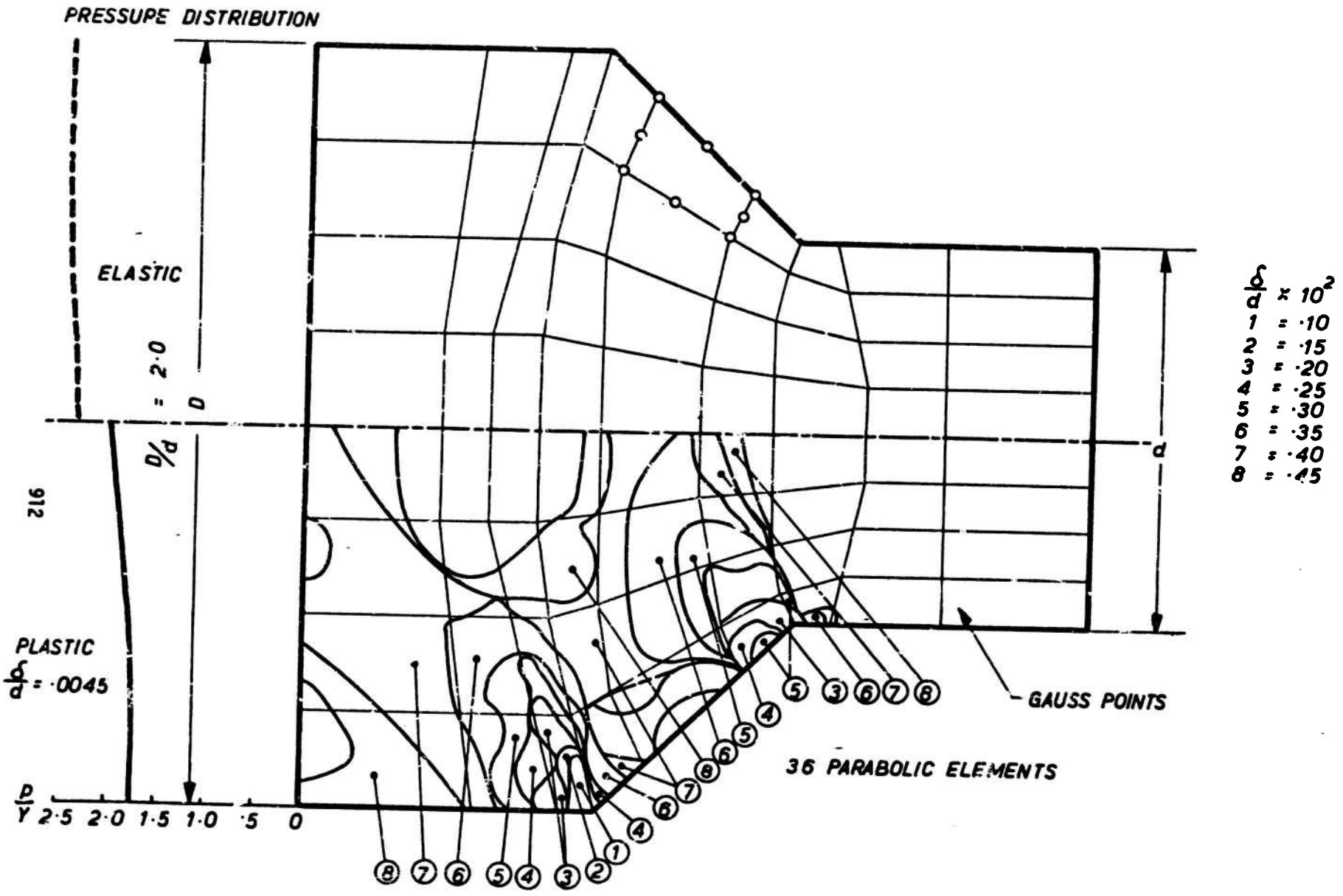


FIG. 10 AN AXISYMETRIC EXTRUSION PROBLEM SPREAD OF PLASTIC ZONES AND PRESSURE DISTRIBUTION

This example incidently indicates one of the advantages of isoparametric elements in plasticity. The "chequerband" pattern of plastic zone spread familiar in the use of constant strain triangles is substituted by a smooth progression of Gauss points at which yield is observed.

Concluding Remarks

The formulation given in this paper shows that in essence large deformations present few additional difficulties. The calculation steps in isoparametric form are standard and follow a well established pattern. The cases of infinitesimal and finite strain differ only in the form of the constitutive relation established.

While it is more conventional to describe all the steps in terms of tensor notation, the recasting of these in matrix form is computationally advantageous and presents no difficulties.

In all the problems given here, we have used a generalized continuum approach and by the use of Green or Almansi strain deformations have avoided all approximations involved in the introduction of oriented bodies and as beams, plates, or shells. However, as shown in Reference 13 the specialization to such cases presents no difficulties providing we are prepared to accept restrictions of small rotations. With this restriction, approximate relations conventionally used are introduced and the formulation in essence remains unaltered.

Several solution techniques have been introduced into one program and the choice between alternatives left to the user. It is almost impossible to determine the one which at all times is most expeditious, and therefore the choice is essential.

In plasticity problems as stated earlier (Reference 9) we find (Reference 10) that modified Newton-Raphson (initial stress) techniques are usually most economic especially if used with an efficient accelerator. Indeed, if strain-softening problems are dealt with such methods are the only ones applicable as tangential matrix may become indefinite.

In problems involving significant deformation, the same techniques may lead to very slow or non-evident convergence and here the alternative of varying the tangential matrix in each computation step is often essential.

It is thus important to "keep all options open" in nonlinear problems and indeed it appears that other solution processes as yet unexplored may in future provide further alternatives.

We have not discussed in detail here the question of deciding at what stage of iteration a solution is 'sufficiently accurate'.

In earlier work an absolute maximum magnitude of error was often specified. We find in general that a norm of residual forces or of displacement changes provides a more convenient estimate. The question as to the permissible magnitude of such norms needs further investigation and at the moment these are specified by intuitive reasoning.

The introduction of the residual force concept is, we find, an essential feature of any reasonable approach to nonlinear problems.

The merits of isoparametric element forms as the basis of nonlinear analysis have been brought out in the text. The accuracy of representation in linear elastic analysis seems to carry over to nonlinear situations.

In three dimensional problems, their use is essential to linear and nonlinear analysis alike.

Although we have shown relatively simple quadrilateral forms only in the examples given - the use of hierarchial elements (Reference 28) is worth while in complex situations.

The numerical integration used invariably makes introduction to isoparametric elements of any nonlinear formulation easy. Indeed, the main question at the present time concerns not the numerical methodology but the lack of sufficiently broad constitutive relationships to cover many existing materials.

In the formulation presented for the tangential stiffness matrix $\mathbf{K}_T$ of Equation 38 we notice that this is given in two parts — conveniently separating the initial stress stiffness matrix from the remainder. We would like to point out that it is possible to present the expression in a more compact way in which only 3 matrices are involved, i.e.

$$\mathbf{K}_T = \int_V \mathbf{G}^T \mathbf{D}_g \mathbf{G} \, dV \tag{78}$$

where $\mathbf{D}_g$ is a 9 x 9 "generalized modulus matrix". This form is more convenient in some computation and is given in Reference 10 . We have, however, retained the 'split' form which is applicable also in case of 'oriented bodies' such as bars, plates, etc. (see Ch. 19 ,Reference 13).

The program described for plasticity allows the inclusion of any generalized yield criteria (Reference 12) by a simple change of four Fortran statement cards.

REFERENCES

1. Marcal, P. V., "On General Purpose Programs for Finite Element Analysis, with Special Reference to Geometric and Material Nonlinearities". Symp. Num. Solution of Partial Differential Equations. University of Maryland, May 1970.

2. Mallett, R. H. and Marcal, P. V., "Finite Element Analysis of Nonlinear Structures". J. Struct. Div. Proc. ASCE, 94, ST9, 2081–2105, 1968.

3. Oden, J. T., "Numerical Formulation of Nonlinear Elasticity Problems". J. Struct. Div., Proc. ASCE, 93, ST3, 235–255, 1967.

4. Oden, J. T., "Finite Plane Strain of Incompressible Elastic Solids by the Finite Element Method". The Aero. Quart. 19, 254–264, 1968.

5. Haisler, W. E. and Stricklin, J. A., and Stebbins, F. J., "Development and Evaluation of Solution Procedures for Geometrically Nonlinear Structural Analysis by the Direct Stiffness Method". AIAA/ASME 12th structures, Struct. Dynamics and Materials Conference. Anaheim, California, April 1971.

6. Hartz, B. J. and Nathar, N. D., "Finite Element Formulation of Geometrically Nonlinear Problems of Elasticity". Japan – U. S. Seminar on Matrix Methods of Structural Analysis and Design, Tokyo, August 1969. Alabama University Press, 1970.

7. Martin, H. C., "On the Derivation of Stiffness Matrices for the Analysis of Large Deflection and Stability Problems". Proc. Conf. on Matrix Methods in Struct. Mech., AFIT, Wright-Patterson Air Force Base, Ohio, October 1965, 697–715, 1966.

8. Argyris, J. H., "Matrix Analysis of Three Dimensional Elastic Media – Small and Large Displacements". J. AIAA, 3, 1, 45–51, 1965.

9. Zienkiewicz, O. C., Valliappan, S., and King, I. P., "Elastic-plastic Solutions of Engs Problems, Initial Stress Finite Element Approach". Int. J. Num. Methods in Engs. 1, pp. 75–100, 1969.

10. Nayak, G. C., "Plasticity and Large Deformation Problems by the Finite Element Method". Ph.D. Thesis, University of Wales, Swansea, 1971.

11. Nayak, G. C. and Zienkiewicz, O. C., "Note on the 'Alpha' – Constant Stiffness Method for Analysis of Nonlinear Problems". To be published Int. J. Num. Meth. Eng.

12. Nayak, G. C. and Zienkiewicz, O. C., "Elasto-plastic Stress Analysis with Curved Isoparametric Elements for Various Constitutive Relations". Paper to be published.

13. Zienkiewicz, O. C., "The Finite Element Method in Engineering Science". McGraw-Hill, 1971.

14. Irons, B. M. "Engineering Applications of Numerical Integration in Stiffness Method". J. AIAA, 14, 2035–2037, 1966.

15. Ergatoudis, J. G., Irons, B. M., and Zienkiewicz, O. C., "Curved, Isoparametric, Quadrilateral Elements for Finite Element Analysis". Int. J. Solids Struct., 4, 31–42, 1968.

16. Green, A. E., and Naghdi, P. M., "A General Theory of Elastic-plastic Continuum". Arch. Pat. Mech. Anal., 18, 251–281, 1965.

17. Bisshopp, K. E., and Drucker, D. C., "Large Deflections of Cantilever Beams". Quart. Appl. Math, 3, 272–275, 1945.

18. Ahmed, S., Irons, B. M. and Zienkiewicz, O. C., "Analysis of Thick and Thin Shell Structures by Curved Finite Elements". Int. J. Num. Math. Eng. 2, 419–451, 1970.

19. Zienkiewicz, O. C., Taylor, R. L. and Too, J. A., "Reduced Integration Technique in General Analysis of Plates and Shells". Int. J. Num. Math. Eng. 3, 275–290, 1971.

20. Pian, T. H. H. and Tong Pin, "Variational Formulation of Finite Displacement Analysis". Symposium on High Speed Computing in Elasticity, I. U. T. A. M., Liege, 1970.

21. Zienkiewicz, O. C., "Incremental Displacement in Nonlinear Analysis". To be published in Int. J. Num. Math. Eng.

22. Biezeno, C. B. and Grammel, R., "Elastic Problems of Single Machine Elements". Van Nostrand, 1956.

23. Hodge, P. G., Jr., and G. N. White, Jr., "A Quantitative Comparison of Flow and Deformation Theories of Plasticity". J. Appl. Mech. 17, 180–184, 1950.

24. Johnson, W., and Mellor, P. B., "Plasticity for Mechanical Engineers, pp. 242–243. Van Nostrand, London, 1962.

25. Lee, C. H. and Koyayashi, Shiro, "Elastoplastic Analysis of Plane-strain and Axisymmetric Flat Punch Indentation by the Finite Element Method". Int. J. Mech. Sci., 12, pp. 349–370, April 1970.

26. Hill R., "A General Theory of Uniqueness and Stability in Elastic-plastic Solids". J. Mech., and Phys. Solids, 6, 236–249, 1958.

27. Lee, E. H., "Elastic-plastic Deformation at Finite Strains". J. Applied Mech., 36, 1–6, March 1969.

28. Zienkiewicz, O. C., Irons, B. M., Campbell, J., and Scott, F., "Three Dimensional Stress Analysis". Int. Union Th. Appl. Mech. Symposium on High Speed Computing in Electricity, Liege, 1970, I. U. T. A. M., Liege 1970.

29. Murnaghan, F. D., "Finite Deformation of an Elastic Solid". John Wiley, London, 1951.

30. Hill, R., "The Mathematical Theory of Plasticity". Oxford, 1950.

31. Prager, W., "Nonisothermal Plastic Deformation". Proc. Kon. Nederi. Akad. Wet., 61, 176–182, 1958.

APPENDIX I

THE MATRIX FORM OF FINITE DEFORMATION THEORY

1. Daformation and Strain

With $\mathbf{x} = [x\ y\ z]^T$ denoting undeformed rectangular coordinates, $\mathbf{u} = [u\ v\ w]^T$ the displacements measured in the same frame of reference, the coordinates of a deformed point become

$$\bar{\mathbf{x}} = [\bar{x}\ \bar{y}\ \bar{z}]^T = \mathbf{x} + \mathbf{u} \qquad \text{A1}$$

A line element ds with components $d\mathbf{x} = [dx\ dy\ dz]^T$ on deformation changes to a new length $d\bar{s}$ with components $d\bar{\mathbf{x}}$. We can write

$$d\bar{\mathbf{x}} = \mathbf{J}\, d\mathbf{x} \qquad \text{A2}$$

where

$$\mathbf{J} = \left[\frac{\partial(\bar{x},\bar{y},\bar{z})}{\partial(x,y,z)}\right] = \begin{bmatrix} \frac{\partial\bar{x}}{\partial x} & \frac{\partial\bar{x}}{\partial y} & \frac{\partial\bar{x}}{\partial z} \\ \frac{\partial\bar{y}}{\partial x} & \frac{\partial\bar{y}}{\partial y} & \frac{\partial\bar{y}}{\partial z} \\ \frac{\partial\bar{z}}{\partial x} & \frac{\partial\bar{z}}{\partial y} & \frac{\partial\bar{z}}{\partial z} \end{bmatrix} \qquad \text{A3}$$

is a Jacobian matrix defining the deformed state in some way with $[\bar{J}]$ being similarly given.

We note in passing that

$$\bar{\mathbf{J}} = \mathbf{J}^{-1} = \left[\frac{\partial(x,y,z)}{\partial(\bar{x},\bar{y},\bar{z})}\right]$$

from the definition A2 .

We can write the change of length measure as

$$\tfrac{1}{2}(ds^{-2} - ds^2) = \tfrac{1}{2}(d\bar{\mathbf{x}}^T\, d\bar{\mathbf{x}} - d\mathbf{x}^T\, d\mathbf{x})$$

$$\equiv d\mathbf{x}^T\, \boldsymbol{e}\, d\mathbf{x} \equiv d\bar{\mathbf{x}}^T\, \bar{\boldsymbol{e}}\, d\bar{\mathbf{x}} \qquad \text{A4}$$

where

$$\boldsymbol{e} = \tfrac{1}{2}(\mathbf{J}^T\mathbf{J} - \mathbf{I}) \qquad \text{A5}$$

is the Lagrangian definition of strain (Green's strain 1839–42) and

$$\epsilon = \tfrac{1}{2}(\mathbf{I} - \bar{\mathbf{J}}^{T}\bar{\mathbf{J}}) \qquad \text{A6}$$

is the Eulerian definition of strain (Almansi strain 1911) (in both $\mathbf{I}$ is an identity (unit) matrix).

The physical measuring of the strains defined above is discussed in many textbooks, but it is of importance for an engineer to realize that if strains are small (but displacements large), the first definition gives elongations and angular changes in reference to an element positioned in the undeformed body while the second refers to the same quantities of an element positioned in the deformed body.

Thus, for instance, in Figure 1 we illustrate for the small strain case the meaning of the components ϵ_x and $\bar{\epsilon}_x$.

By substituting Equation A1 into the expressions A5 and A6 the explicit engineering expressions for strain components, Equations 22 and 40 given earlier in the text are derived. In these we arrange the strains in a vector form ϵ or $\bar{\epsilon}$ in the usual manner as the strains are symmetric.

At this stage it is worth noting that the computer evaluation of the Jacobian matrices is an economic way of arriving at the strain components as described in the text.

2. Changes of Geometry

If $d\mathbf{A} = [dA_x\ dA_y\ dA_z]^T$ defines an elementary area by its three projections (on planes perpendicular to x, y, z axes) in the undeformed coordinates, and if $d\bar{\mathbf{A}} = [d\bar{A}_x\ d\bar{A}_y\ d\bar{A}_z]^T$ refers to the same area in deformed coordinates then

$$d\bar{\mathbf{A}} = |J|\ \mathbf{J}^{T-1}\, d\mathbf{A}$$

and

$$d\mathbf{A} = |J|^{-1}\ \mathbf{J}^{T}\, d\bar{\mathbf{A}} \qquad \text{A7}$$

Similarly elementary volumes dV and $d\bar{V}$ are related by

$$d\bar{V} = |J|\ dV \qquad \text{A8}$$

For elaboration on the proofs, the reader is referred to Murnaghan (Reference 29).

3. Variation of Strain

Let us consider the variation of the strain Jacobian matrix for small changes of displacement. Now

$$d\mathbf{J} = \left[\frac{\partial\ (\bar{x}+du,\ \bar{y}+dv,\ \bar{z}+dw)}{\partial\ (x,\ y,\ z)}\right] - \left[\frac{\partial\ (\bar{x},\bar{y},\bar{z})}{\partial\ (x,y,z)}\right] \qquad \text{A9}$$

which on expanding gives

$$d\mathbf{J} = \bar{\mathbf{V}}_d \mathbf{J}$$

where $$\bar{\mathbf{V}}_d \equiv \left[\frac{\partial\,(du,dv,dw)}{\partial\,(\bar{x},\,\bar{y},\,\bar{z})}\right] \qquad \text{A10}$$

is variation of deformation.

As

$$\mathbf{J}\,\bar{\mathbf{J}} = \mathbf{I}\ ,\ \text{by taking its variation}$$

we have

$$d\bar{\mathbf{J}} = -\bar{\mathbf{J}}\,\bar{\mathbf{V}}d \qquad \text{A11}$$

It is simple to show that the variation of Green's and Almansi strain matrices is given by

$$d\epsilon = \mathbf{J}^T\, d\bar{\epsilon}\, \mathbf{J} \qquad \text{A12}$$

and

$$d\bar{\epsilon} = d\bar{\epsilon} - \bar{\mathbf{V}}_d{}^T\,\bar{\epsilon} + \bar{\epsilon}\,\bar{\mathbf{V}}_d \qquad \text{A13}$$

with

$$d\bar{\epsilon} = \tfrac{1}{2}(\bar{\mathbf{V}}_d + \bar{\mathbf{V}}_d{}^T) \qquad \text{A14}$$

$$d\bar{\omega} = \tfrac{1}{2}(\mathbf{V}_d - \mathbf{V}_d{}^T) \qquad \text{A15}$$

4. Definitions of Stress

The natural definition of stresses is obviously the Eulerian one referring in the usual way to the forces per unit deformed areas.

Thus we can write

$$\bar{\sigma} = \begin{bmatrix} \bar{\sigma}_x & \bar{\tau}_{xy} & \bar{\tau}_{xz} \\ \bar{\tau}_{xy} & \bar{\sigma}_y & \bar{\tau}_{yz} \\ \bar{\tau}_{xz} & \bar{\tau}_{yz} & \bar{\sigma}_z \end{bmatrix} \qquad \text{A16}$$

in matrix notation or simply rearrange in the usual vector form as $\bar{\sigma}$.

The forces acting on an area $d\bar{\mathbf{A}}$ are given by the vector

$$d\bar{\mathbf{F}} = \bar{\sigma}\, d\bar{\mathbf{A}} \qquad \text{A17}$$

For a Lagrangian definition of stress the situation is by no means so clear and various alternatives were proposed in the literature. Here we shall use the Piola–Kirchhoff (1833, 1852) definition.

This gives an equivalent force vector acting on an original, undeformed area dA by an expression identical to Equation A17, i.e.

$$d\bar{\mathbf{F}} = \boldsymbol{\sigma}\, d\mathbf{A} \qquad \text{A18}$$

and the actual force on the deformed area

$$d\tilde{\bar{\mathbf{F}}} = \mathbf{J}\, d\mathbf{F} \qquad \text{A19}$$

From Equations A18 and A7 we have immediately the stress transformation relation

$$\boldsymbol{\sigma} = |\mathbf{J}|\, \mathbf{J}^{-1}\, \bar{\boldsymbol{\sigma}}\, \mathbf{J}^{T-1} \qquad \text{A20}$$

which shows that the above stress definition is still symmetric and can be written in as a vector $\boldsymbol{\sigma}$.

5. Rate of Work

The rate at which work is being done internally or virtual work done during a displacement $\mathbf{du}$ can be written in the usual way as in small displacement analysis

$$\int_{\bar{V}} (\bar{\boldsymbol{\sigma}}^T\, d\bar{\boldsymbol{\epsilon}})\, d\bar{V} \qquad \text{A21}$$

in which

$$d\bar{\boldsymbol{\epsilon}} = d\begin{bmatrix} \dfrac{\partial u}{\partial \bar{x}} \\ \dfrac{\partial v}{\partial \bar{y}} \\ \dfrac{\partial w}{\partial \bar{z}} \\ \dfrac{\partial w}{\partial \bar{y}} + \dfrac{\partial v}{\partial \bar{z}} \\ \dfrac{\partial u}{\partial \bar{z}} + \dfrac{\partial w}{\partial \bar{x}} \\ \dfrac{\partial v}{\partial \bar{x}} + \dfrac{\partial u}{\partial \bar{y}} \end{bmatrix} \qquad \text{A22}$$

corresponds to changes $d\bar{\boldsymbol{\epsilon}}$ and is identical to the expression for infinitesimal strains defined with respect to the instantaneous coordinate $\bar{\mathbf{x}}$.

Purely formal transformations transform the above Eulerian definition to the equivalent Lagrangian and which now is

$$\int_V (\sigma^T d\epsilon)\, dV \qquad \text{A23}$$

This form easily recognizable by similarity with small strain and displacement analysis, is oblivious of small strain but large displacements are considered and σ represents the locally directed stresses dependent on ϵ. In the more general sense of finite strain it is the merit of the Piola–Kirchhoff definition that the form is preserved.

6. Constitutive Relations and Increments of Strain / Stress

The additiveness of stress/strain changes needs to be considered in general finite element analysis.

Consider first the question of thermal strains incrementing during the deformation.

In the case of Lagrangian definition such strains can be added to the total strain irrespective of whether the material is 'anisotropic' or isotropic. In the case of Eulerian strain definition isotropic strain can be added but if Figure 1 is considered it will be seen that anisotropic initial strains need to be transformed to account for the rotation of material axes.

If small strain linear elasticity is used to define the constitutive relation of the usual form

$$\sigma = D\epsilon \qquad \text{A24}$$

again we note that correct stresses will be obtained for any material in the Lagrangian definition but, that for anisotropy a rotation of axes is necessary in defining the D matrix when the Eulerian system is used.

The same remarks apply to large deformation, elasticity where effectively the D matrix is derived by differentiation of the strain-energy density with respect to strain components.

Here again, anisotropy will favour the use of the Lagrangian definition.

For more general constitutive relations, such as may be involved in plasticity, etc., we derive relationships between changes of strain and changes of deformation.

In Lagrangian formulation we note that the stress is always associated with direction of strains and we can use the form

$$d\sigma = D_T\, d\epsilon \qquad \text{A25}$$

with elements $d\sigma$ being simply additive. In plasticity, for instance, we have to refer the straining rate to some form of actual stress changes. However, it is clear from previous discussion that, physically, changes of Eulerian stress and strain cannot be related (as changes in these will occur by pure rigid body rotation). This necessitates the introduction of another type of stress rate – due to Jaumann (1911) and

a relation of the type

$$d\sigma_J = \bar{D}_T \, d\bar{\epsilon} \qquad \text{A26}$$

These changes of the Jaumann stress can be related to the changes of Eulerian stress components (in the additive sense) by writing in vector form

$$d\bar{\sigma} = d\sigma_J + d\,T_\omega \, \bar{\sigma} \qquad \text{A27}$$

where $d\,T_\omega$ is defined by Equation A15, and is given by

$$d\,T_\omega = \begin{bmatrix} 0 & 0 & 0 & 0 & -\bar{\omega}_y & \bar{\omega}_z \\ 0 & 0 & 0 & \bar{\omega}_x & 0 & -\bar{\omega}_z \\ 0 & 0 & 0 & -\bar{\omega}_x & \bar{\omega}_y & 0 \\ 0 & -\tfrac{1}{2}\bar{\omega}_x & \tfrac{1}{2}\bar{\omega}_x & 0 & -\tfrac{1}{2}\bar{\omega}_z & \tfrac{1}{2}\omega_y \\ \tfrac{1}{2}\bar{\omega}_y & 0 & -\tfrac{1}{2}\bar{\omega}_y & \tfrac{1}{2}\omega_z & 0 & -\tfrac{1}{2}\bar{\omega}_x \\ -\tfrac{1}{2}\bar{\omega}_z & \tfrac{1}{2}\bar{\omega}_z & 0 & -\tfrac{1}{2}\omega_y & \tfrac{1}{2}\bar{\omega}_x & 0 \end{bmatrix} \qquad \text{A28}$$

where $\bar{\omega}_x = d\left(\frac{\partial v}{\partial \bar{z}} - \frac{\partial w}{\partial \bar{y}}\right) \bar{\omega}_y = d\left(\frac{\partial w}{\partial \bar{x}} - \frac{\partial u}{\partial \bar{z}}\right) \bar{\omega}_z = d\left(\frac{\partial u}{\partial \bar{y}} - \frac{\partial v}{\partial \bar{x}}\right)$

APPENDIX II

PLANE AND AXISYMMETRIC CASES

Plane Strain / Stress:

The Green strain in vector form can be written as

$$\epsilon = [\epsilon_x \; \epsilon_y \; \epsilon_{xy}]^T = \epsilon^0 + \epsilon^L \qquad \text{81}$$

where

$$\epsilon^L = \tfrac{1}{2} \begin{bmatrix} \theta_x^T & \theta_x \\ \theta_y^T & \theta_y \\ 2\theta_x^T & \theta_y \end{bmatrix} = \tfrac{1}{2} \begin{bmatrix} \theta_x^T & 0 \\ 0 & \theta_y^T \\ \theta_y^T & \theta_x^T \end{bmatrix} \begin{bmatrix} \theta_x \\ \theta_y \end{bmatrix} = \tfrac{1}{2} A\theta \qquad \text{B2}$$

3x1 3x4 4x1

The displacement gradients are arranged in vector form as

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_x \\ \boldsymbol{\theta}_y \end{bmatrix} \begin{bmatrix} \cdots & \mathbf{I}_2 & \dfrac{\partial N_i}{\partial x} & \cdots \\ \cdots & \mathbf{I}_2 & \dfrac{\partial N_i}{\partial y} & \cdots \end{bmatrix} \boldsymbol{\delta} = \mathbf{G}\boldsymbol{\delta} \qquad \text{B3}$$

$4 \times 1 \qquad\qquad 4 \times n \quad n \times 1$

where $\boldsymbol{\theta}_x = \left[\dfrac{\partial u}{\partial x} \quad \dfrac{\partial v}{\partial x}\right]^T$, $\boldsymbol{\theta}_y = \left[\dfrac{\partial u}{\partial y} \quad \dfrac{\partial v}{\partial y}\right]^T$ B4

and $\mathbf{I}_2$ is a 2x2 identity matrix.

Finally, the strain displacement relationship is described by

$$d\epsilon = \mathbf{B}\, d\,\boldsymbol{\delta}$$

where the submatrix for node i

$$\mathbf{B}_i = \left[\begin{array}{cc:cc} \dfrac{\partial \bar{x}}{\partial x} & \dfrac{\partial N_i}{\partial x} & \dfrac{\partial \bar{y}}{\partial x} & \dfrac{\partial N_i}{\partial x} \\ \dfrac{\partial \bar{x}}{\partial y} & \dfrac{\partial N_i}{\partial y} & \dfrac{\partial \bar{y}}{\partial y} & \dfrac{\partial N_i}{\partial y} \\ \dfrac{\partial \bar{x}}{\partial x} \dfrac{\partial N_i}{\partial y} + \dfrac{\partial \bar{x}}{\partial y} \dfrac{\partial N_i}{\partial x} & & \dfrac{\partial \bar{y}}{\partial y} \dfrac{\partial N_i}{\partial x} + \dfrac{\partial \bar{y}}{\partial x} \dfrac{\partial N_i}{\partial y} & \end{array}\right] \qquad \text{B5}$$

and the (4x4) $\mathbf{M}$ matrix required for 'initial stress stiffness'

$$\mathbf{K}_\sigma = \int_V \mathbf{G}^T \mathbf{M} \mathbf{G}\, dV$$

can be derived as

$$\mathbf{M} = \begin{bmatrix} \sigma_x \mathbf{I}_2 & \tau_{xy} \mathbf{I}_2 \\ \tau_{xy} \mathbf{I}_2 & \sigma_y \mathbf{I}_2 \end{bmatrix} \qquad \text{B6}$$

Axisymmetric Case:

The Green's strain in vector form in r, z, θ cylindrical coordinates is written as

$$\epsilon = [\epsilon_r \;\; \epsilon_z \;\; \gamma_{rz} \;\; \epsilon_\theta]^T = \epsilon^0 + \epsilon^L \qquad \text{B7}$$

where

$$\underset{4\times1}{\epsilon^L} = \tfrac{1}{2}\begin{bmatrix} \theta_r^T & \theta_r \\ \theta_z^T & \theta_z \\ 2\theta_r^T & \theta_z \\ \left(\frac{u}{r}\right)^2 & \end{bmatrix} \begin{bmatrix} \theta_r^T & 0 & 0 \\ 0 & \theta_z^T & 0 \\ \theta_z^T & \theta_r^T & 0 \\ 0 & 0 & \left(\frac{u}{r}\right) \end{bmatrix} \underset{4\times5}{\begin{bmatrix} \theta_r \\ \theta_z \\ \frac{u}{r} \end{bmatrix}}_{5\times1} = \tfrac{1}{2}\mathbf{A}\theta \qquad \text{B8}$$

The displacement gradient θ for an axisymmetric case now becomes a 5x1 vector and the submatrix $\mathbf{B}_i$ for node i in the strain displacement relationship can be derived as

$$\mathbf{B}_i = \left[\begin{array}{ccc:ccc} \frac{\partial\bar{r}}{\partial r} & & \frac{\partial N_i}{\partial r} & \frac{\partial\bar{z}}{\partial r} & & \frac{\partial N_i}{\partial r} \\ \frac{\partial\bar{r}}{\partial z} & & \frac{\partial N_i}{\partial z} & \frac{\partial\bar{z}}{\partial z} & & \frac{\partial N_i}{\partial z} \\ \frac{\partial\bar{r}}{\partial r} & \frac{\partial N_i}{\partial z} + \frac{\partial\bar{r}}{\partial z} & \frac{\partial N_i}{\partial r} & \frac{\partial\bar{z}}{\partial z} & \frac{\partial N_i}{\partial r} + \frac{\partial\bar{z}}{\partial r} & \frac{\partial N_i}{\partial z} \\ \left(\frac{\bar{r}}{r}\right) & \frac{N_i}{r} & & & 0 & \end{array}\right] \qquad \text{B9}$$

Finally the matrix $\mathbf{M}$ is modified to 5x5 matrix

$$\mathbf{M} = \begin{bmatrix} \sigma_r \mathbf{I}_2 & \tau_{rz}\mathbf{I}_2 & \mathbf{0} \\ \tau_{rz}\mathbf{I}_2 & \sigma_z\mathbf{I}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_\theta \end{bmatrix} \qquad \text{B10}$$

APPENDIX III

EULERIAN FORMULATION

For the derivation incremental force displacement relationship in Eulerian form, if we consider increment of the internal force term of Equation 10 we have

$$d\left(\int_{\bar{V}} \bar{\mathbf{B}}^{oT}\,\bar{\sigma}\,d\bar{V}\right) = \int_{\bar{V}} \left(d\bar{\mathbf{B}}^{oT}\,\bar{\sigma}\,d\bar{V} + \bar{\mathbf{B}}^{oT}\,d\bar{\sigma}\,d\bar{V} + \bar{\mathbf{B}}^{oT}\,\bar{\sigma}\,d\,(d\bar{V})\right) \qquad \text{C1}$$

Substituting in Equation C1

$$d\bar{\bar{\sigma}} = \bar{\bar{D}}_T \, d\bar{\bar{e}} + dT_\omega \, \bar{\bar{\sigma}} \text{ (from Equation A27)}$$

and

$$d(d\bar{V}) = (d\bar{e}_x + d\bar{e}_y + d\bar{e}_z)\, d\bar{V} \qquad C2$$

we have

$$d\left(\int_{\bar{V}} \bar{\bar{B}}^{o\,T} \bar{\bar{\sigma}}\, d\bar{V}\right) = \int_{\bar{V}} \bar{\bar{B}}^{o\,T} \bar{\bar{D}}_T \bar{\bar{B}}^{o}\, d\bar{V}\, d\delta$$

$$+ \int_{\bar{V}} (d\bar{\bar{B}}^{o\,T} + \bar{\bar{B}}^{o\,T} dT_\omega + (d\bar{e}_x + d\bar{e}_y + d\bar{e}_z)\, \bar{\bar{B}}^{o\,T} \bar{\bar{\sigma}}\, d\bar{V} \qquad C3$$

The first term on the right-hand side of the Equation C3 gives $\bar{\bar{K}}^o$ and the second term gives the initial stress stiffness matrix $\bar{\bar{K}}_\sigma$. The 6×6 matrix dT_ω appearing in second term is defined by Equation A28 and $d\bar{e}_x$, $d\bar{e}_y$ —— are given by Equation A15. Finally the variation $d\bar{\bar{B}}^{o\,T}$ can be derived by considering equation

$$\left[\frac{\partial N_i}{\partial x}\right]_i = J^T \left[\frac{\partial N_i}{\partial \bar{x}}\right] \qquad C4$$

$$d\left[\frac{\partial N_i}{\partial x}\right] = 0 = dJ^T \left[\frac{\partial N_i}{\partial \bar{x}}\right] + J^T d\left[\frac{\partial N_i}{\partial \bar{x}}\right]$$

therefore

$$d\left[\frac{\partial N_i}{\partial \bar{x}}\right] = - J^{T-1}\, dJ^T \left[\frac{\partial N_i}{\partial \bar{x}}\right]$$

and as from Equation A10 $J^{T-1} dJ^T = \bar{\bar{V}}_d^T$,

$$d\left[\frac{\partial N_i}{\partial \bar{x}}\right] = -\bar{\bar{V}}_d^T \left[\frac{\partial N_i}{\partial \bar{x}}\right] \qquad C5$$

From Equation 27 it can be easily seen that the variation $d\bar{\bar{B}}^{o\,T}$ can be easily obtained from C5 by rearranging terms. However, the rearrangement of the whole of second term on right-hand side of Equation C3 requires elaborate algebraic organization.

APPENDIX IV

LARGE STRAIN PLASTICITY

In the literature several attempts have been made to establish elastic plastic relationships at finite deformations and we shall restrict ourselves to papers by Hill (Reference 26) (1958), Green and Naghdi (Reference 16) (1965) and Lee (Reference

27) (1969). We shall classify them into two cases, (i) small strain with large rotation (References 16 and 26) and (ii) finite elastic-plastic strain (Reference 27). We shall briefly recollect some of these relationships in Lagrangian form and state the counterpoints in Eulerian form. First of all we shall consider the case of small elastic plastic strains with large rotations in which the total Green's strain can be divided into elastic and plastic components as considered by Green and Naghdi (16, section 5). Thus

$$\epsilon = \epsilon_e + \epsilon_p \qquad \text{D1(a)}$$

and so the increments of Green's strain

$$d\epsilon = d\epsilon_e + d\epsilon_p \qquad \text{D1(b)}$$

The yield criterion in terms of Piola–Kirchhoff stresses σ is expressed as (neglecting temperature effects)

$$F(\sigma, \epsilon_p, \kappa) = 0 \qquad \text{D2}$$

Where κ is a hardening parameter defining subsequent surfaces corresponding to ϵ_p and whole history. The usual restrictions for neutral states or unloading hold as in the classical plasticity theory. Also, the subsequent yield surfaces defined in general by

$$d\kappa = d\kappa(\sigma, \epsilon_p, d\sigma, d\epsilon_p)$$

is restricted to linear relation with $d\epsilon_p$ and thus normality rule is derived as

$$d\epsilon_p = d\lambda\, \mathbf{a} \qquad \text{D3}$$

with

$$\frac{\partial F}{\partial \sigma} = \mathbf{a}^T \qquad \text{D4}$$

For infinitesimal elastic strains one can write

$$d\sigma = \mathbf{D}\, d\epsilon_e \qquad \text{D5}$$

where $\mathbf{D}$ is the usual elasticity modulus matrix (Reference 13). Now the elastic plastic matrix can be derived in the same way (Reference 12).

$$d\lambda = (A + \beta)^{-1} (\mathbf{d}^T d\epsilon) \quad \text{with } \mathbf{d} = \mathbf{D}\mathbf{a},$$

$$\beta = \mathbf{a}^T \mathbf{D}\mathbf{a} \text{ and A is arbitrary hardening constant and}$$

$$d\sigma = \mathbf{D}_T d\epsilon = (\mathbf{D} - (A + \beta)^{-1} \mathbf{d}\, \mathbf{d}^T) \qquad \text{D6}$$

However, an important difference should be noticed at this stage as regards to yield criterion D2. Usually, the yield criterion is expressed in terms of true stresses $\bar{\sigma}$ and for finite strain case with isotropic hardening, following Prager (Reference 31) one can write

$$F = f(|J|\, \bar{\sigma}) - Y(\kappa) = 0 \qquad \text{D7}$$

And from Equation A13 the true stresses are related to Piola–Kirchhoff stresses

$$|J|\,\bar{\sigma} = J\,\sigma\,J^T$$

and by transforming it into vectorial form, one can easily write

$$|J|\,\bar{\sigma} = T\,\sigma \qquad \text{D8}$$

where 6x6 matrix T is function of deformation gradients. From Equations D7 and D8 one can derive

$$\frac{\partial F}{\partial \sigma} = a^T = \frac{\partial f}{\partial \bar{\sigma}} \cdot \frac{\partial \bar{\sigma}}{\partial \sigma} = \bar{a}^T T \qquad \text{D9}$$

where $\bar{a}^T = \dfrac{\partial f}{\partial \bar{\sigma}}$

the particular case of Von Mises yield criterion for isotropic material is given by equation:

$$F = \sqrt{3}\; J_2^{\frac{1}{2}} - Y(\kappa) = 0$$

where J_2 is the second invariant of deviators of stress $\bar{\sigma}$ and $Y(\kappa)$ is the yield stress from uniaxial tests. And

$$\frac{\partial F}{\partial \sigma} = a^T = \sqrt{3}\,\frac{\partial J_2^{\frac{1}{2}}}{\partial \bar{\sigma}}\,\frac{\partial \bar{\sigma}}{\partial \sigma} = \sqrt{3}\,\bar{a}_2{}^T T$$

The general derivations for various yield criteria are given in (Reference 12). And the hardening constant A can be determined from the slope of work hardening curve between Y and plastic work Wp. The plastic work Wp is obtained by integrating increments of plastic work per unit volume

$$dWp = \sigma^T d\epsilon_p \qquad \text{D10}$$

In Eulerian form one can again divide the strain rate given by Equation A13 into elastic and plastic parts:

$$d\bar{\sigma} = d\bar{\sigma}_e + d\bar{\sigma}_p \qquad \text{D11}$$

Now we shall use Equation D7 for yield criterion and restrict ourselves to isotropic materials to derive

$$d\bar{\sigma}_J = \bar{D}_T\, d\bar{\sigma} \qquad \text{D12}$$

and noting that relation with elastic component is

$$d\bar{\sigma}_J = \bar{D}\, d\bar{\sigma}_e \qquad \text{D13}$$

The essential difference with Lagrangian from Equations D5 and D6 is that we have to use Jaumann stress increments given by Equation A27 so that in the spatial form the constitutive Equations D12 and D13 are unaltered when the continuum is subjected to superimposed rigid body motions. After differentiating Equation D7 we have have to consider

$$\frac{\partial f}{\partial \bar{\sigma}} \, d\bar{\sigma}_J = \bar{a}^T \, d\bar{\sigma}_J = A d\lambda \qquad \text{D14}$$

where A is again undetermined constant depending upon hardening parameter. Writing the normality rule

$$d\bar{\epsilon}_p = d\lambda \bar{a} \qquad \text{D15}$$

we can again derive

$$d\lambda = (A + \bar{\beta})^{-1} (\bar{d}^T \, d\bar{\sigma}) \qquad \text{D16}$$

with $\bar{d} = \bar{D}\,\bar{a}$ and $\bar{\beta} = \bar{a}^T \bar{D}\,\bar{a}$

and

$$\bar{D}_T = (\bar{D} - (A + \bar{\beta})^{-1} \bar{d}\,\bar{d}^T) \qquad \text{D17}$$

It may be noted that in arriving at D17 the transformation D8 used in Lagrangian form is now no longer necessary but instead we have to use Equations D12 and A27.

Lee (Reference 27) considers the case when both elastic and plastic strains are finite and so thet Equation D1 is now no longer valid. In order to separate reversible and irreversible deformation we have to consider the problem of three configurations, initial, final and intermediate originally recognized by Sedov (1962). Let us denote the rectangular coordinates $(\bar{x}_p, \bar{y}_p, \bar{z}_p)$ of intermediate configuration corresponding to irreversible part. Then the deformation matrix.

$$J = \left[\frac{\bar{x}, \bar{y}, \bar{z}}{x, y, z}\right] = \left[\frac{\bar{x}, \bar{y}, \bar{z}}{\bar{x}_p, \bar{y}_p, \bar{z}_p}\right] \left[\frac{\bar{x}_p, \bar{y}_p, \bar{z}_p}{x, y, z}\right]$$

$$= J_e J_p \qquad \text{D18}$$

replaces Equation D1(a). Also, it will be wrong to consider Equations D1(b) and D11 as valid until one considers the elastic strain components infinitesimal i.e. $J_e \triangleq I$. The modification thus required for large strain cases is given in Reference 27 and will not be discussed any further.

ELASTO-PLASTIC ANALYSIS OF PLATES USING THE FINITE ELEMENT METHOD

Pål G. Bergan*

The Technical University of Norway, Trondheim

Ray W. Clough**

University of California, Berkeley

The formulation of a general quadrilateral finite element accounting for inelastic material behavior is described. Both flexural and membrane behavior are considered. The flow theory of plasticity is adopted, and the material is assumed to obey the von Mises yield criterion and the isotropic hardening law. Elasto-plastic problems are generally very costly in computer time, and emphasis is placed on an efficient formulation of the element stiffness. A pure incremental technique is chosen for the numerical solution procedure. Results are presented for various types of flexural and membrane plate problems. These examples demonstrate that the present method exhibits a high degree of accuracy while being highly efficient. It is concluded that the element described here is among the most versatile and efficient elasto-plastic elements yet devised.

---

*Associate Professor of Civil Engineering

**Professor of Civil Engineering

SECTION I

INTRODUCTION

Great advances were made in the field of plasticity during the 1940's and 1950's, and the development of the small deformation theory was essentially completed by the end of this period. However, solutions were then available only for a very limited number of simplified problems. The analysis of more complex problems was not possible until the advent of the electronic computer made numerical solution procedures practicable. The finite element method, which originally was applied to linear elastic systems, has proven also to provide one of the most effective numerical formulations for problems involving non-linear material behavior. Elasto-plastic finite element analyses have now been made of nearly every basic type of structural system, including plates, axisymmetric shells and solids, and general three-dimensional solids. References 1-6 are typical of the many papers describing such analyses.

The purpose of this paper is to describe the formulation and application of a highly efficient finite element for the analysis of elasto-plastic plates subjected simultaneously to in-plane and out-of-plane forces. Consideration is given here only to the incorporation of non-linear material effects in the analysis; however, the formulation could be extended without any essential difficulty to account for the non-linear effects of large displacements at the same time.

The organization of this paper follows the sequence of steps involved in the formulation of the finite element procedure. Consideration is given first to the constitutive law which expresses the non-linear relation between stresses and strains. In this work, the Prandtl-Reuss flow theory[7], which assumes the increment of plastic strains to be proportional to the corresponding deviatoric stress components, is adopted because it seems the most consistent concept both mathematically and physically.

The next step in the formulation is establishing the strain distribution. This is the essential step of the finite element method and is embodied in the selection of the element displacement interpolation functions. Two types of displacement functions are employed in the quadrilateral element developed here: (1) the transverse bending displacements are expressed by the linear curvature compatible interpolation functions derived for the Q-19 plate bending element[8], (2) the in-plane displacements are expressed by the bilinear interpolations of the plane isoparametric element family[9]. The integration procedures used in computing the element stiffness have been modified from those of the cited references, however, in order to account for the variation of material properties within the element as a function of the strains.

The last step in the analysis is the formulation and solution of the equations of equilibrium. These equations reflect the non-linearity of the material properties, of course, and are solved in this work by an incremental tangent stiffness procedure. This approach has been selected in preference to the "initial stress" or "initial strain" methods because of its general reliability and its compatibility with the incremental flow theory of plasticity. (Comparisons between the two methods are presented in References 3, 6 and 10). The efficiency of the entire solution process is demonstrated by a series of examples at the end of the paper.

SECTION II

GOVERNING MATERIAL RELATIONS

Basic Assumptions and Material Laws

A short outline of the derivations leading to the mathematical expressions for the incremental stress-strain relationship will be given in the following. Proofs and details will not be included; more complete derivations may be found in References 3, 7, 11 and 12.

This derivation is based on three major assumptions. The first assumption is that the elastic and plastic strains may be additively decomposed

$$\epsilon_{ij} = \epsilon_{ij}^{E} + \epsilon_{ij}^{P} \tag{1}$$

where E and P denote "elastic" and "plastic" respectively. Both stresses and strains are referred to a rectangular Cartesian coordinate system. The strains are assumed to be infinitesimal (in fact, Equation 1 is not valid for finite strains). The plastic part of the strains is postulated to be incompressible:

$$\epsilon_{ii}^{P} = 0 \tag{2}$$

The second major assumption is that there exists a loading function f in 9-dimensional stress-space. $f = 0$ at time t constitutes the yield criterion at that time, $f > 0$ is inadmissible. As it turns out that f is a function of σ_{ij} and ϵ_{ij}^{P}, three different loading conditions from a plastic state ($f = 0$) result

$$\begin{aligned}
&\frac{\partial f}{\partial \sigma_{ij}} \dot{\sigma}_{ij} < 0 \quad \text{(during unloading)} \\
&\frac{\partial f}{\partial \sigma_{ij}} \dot{\sigma}_{ij} = 0 \quad \text{(during neutral loading)} \\
&\frac{\partial f}{\partial \sigma_{ij}} \dot{\sigma}_{ij} > 0 \quad \text{(during loading)}
\end{aligned} \tag{3}$$

The dot denotes time differentiation.

The material is further assumed to be stable as defined by Drucker[11]. Drucker's postulate has three important implications, namely "convexity". "consistency" and "normality". Normality implies that at a regular point of the loading surface $f = 0$, the vector $d\epsilon^{P}_{ij}$ is in the direction of the outward normal to the yield surface (in stress space), so that

$$d\epsilon^{P}_{ij} = d\lambda \frac{\partial f}{\partial \sigma_{ij}} \tag{4}$$

$d\lambda$ is a non-negative scalar. When the material obeys the von Mises yield criterion (as will be assumed here), the flow rule of Eq. 4 is equivalent to the Prandtl-Reuss equation. By virtue of Eq. 4, the loading function plays the role of a plastic potential. f remains zero when going from one plastic state to another, hence

$$df = \frac{\partial f}{\partial \sigma_{ij}} \, d\sigma_{ij} + \frac{\partial f}{\partial \epsilon^{P}_{ij}} \, d\epsilon^{P}_{ij} = 0 \tag{5}$$

The initial yield criterion (von Mises) and the loading function are given by

$$f_o = \bar{\sigma} - T_o = \bar{\sigma} - k_o \sqrt{3} = 0 \tag{6}$$

where T is the initial yield stress in uniaxial tension, k_o is the yield stress in pure shear and $\bar{\sigma}$ is the equivalent stress given by

$$\bar{\sigma} = \sqrt{3J_2} = \sqrt{\frac{3}{2} s_{ij} \, s_{ij}} \tag{7}$$

in which J_2 is the second deviatoric-stress invariant and s_{ij} the deviatoric-stress components.

Further, it is assumed that the material exhibits an isotropic hardening behavior (Fig. 1). This implies a uniform expansion of the initial yield surface. The hardening is measured according to the plastic strain hypothesis (which is here equivalent to work hardening) and the yield criterion for subsequent yielding becomes

$$f = \bar{\sigma} - H(\bar{\epsilon}^{P}) = 0 \tag{8}$$

H is a function of the equivalent plastic strain $\bar{\epsilon}^{P}$ which is the integral of

$$d\bar{\epsilon}^{P} = \sqrt{\frac{2}{3}} \left\{ d\epsilon^{P}_{ij} \, d\epsilon^{P}_{ij} \right\}^{1/2} \tag{9}$$

As will be seen later, $H' = \frac{\partial H}{\partial \bar{\epsilon}^{P}}$ is of special interest and is easily obtainable from a uniaxial tension text curve, as shown in Fig. 2:

$$\frac{1}{H'} = \frac{1}{E_T} - \frac{1}{E} \tag{10}$$

E_T is the tangent modulus.

General Constitutive Equations

The following relation can be derived from Hooke's law and Equations 1, 4, 5 and 8[11, 12]:

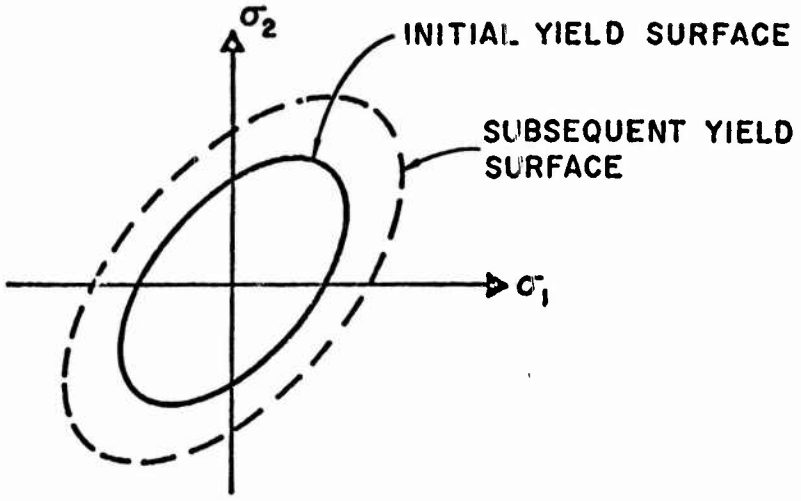


FIG. 1 ISOTROPIC HARDENING

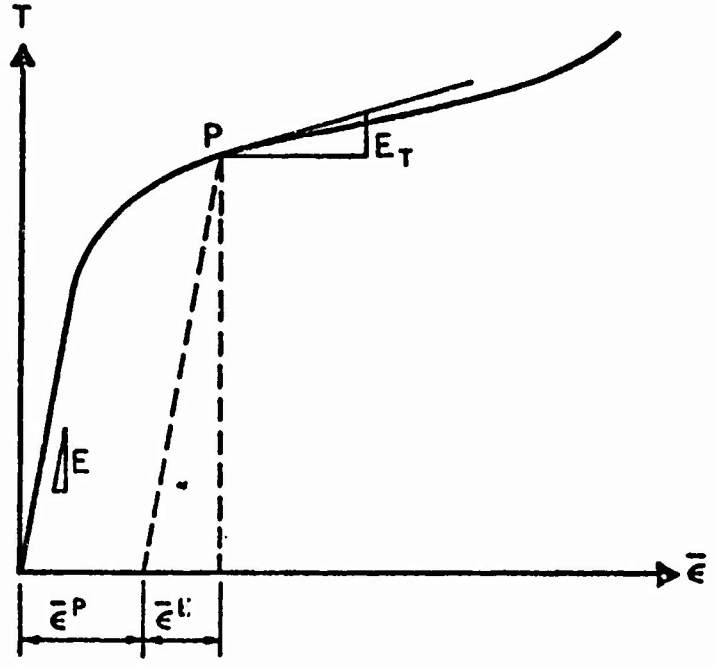


FIG. 2 EXPERIMENTAL STRESS-STRAIN CURVE

$$d\epsilon_{ij}^{P} = A_{ijkl}\, d\epsilon_{kl} \tag{11}$$

where

$$A_{ijkl} = \frac{9\, s_{ij}\, s_{kl}}{2\bar{\sigma}^{2}(H' + 3\mu)} \tag{12}$$

In Eq. 12, μ is Lame's constant given by

$$\mu = \frac{E}{2(1 + \nu)} \tag{13}$$

where ν is Poissons ratio.

The incremental stress-strain relationship is now easily obtained as follows:

$$d\sigma_{ij} = E_{ijkl}\,(d\epsilon_{kl} - d\epsilon_{kl}^{P}) = C_{ijkl}\, d\epsilon_{kl} \tag{14}$$

in which

$$C_{ijkl} = E_{ijkl} - \frac{9\mu^{2}\, s_{ij}\, s_{kl}}{\bar{\sigma}^{2}(H' + 3\mu)} \tag{15}$$

represents the incremental stress-strain relation; it is symmetric in form. E_{ijkl} represents Hooke's law for elastic deformations. Equation 15 is valid between two consecutive plastic states. According to the loading criterion Eq. 3, when

$$\bar{\sigma} < H \tag{16}$$

or

$$\bar{\sigma} = H \text{ and } s_{ij}\, d\,\epsilon_{ij} < 0 \tag{17}$$

the last term of Eq. 15 should be neglected and only Hooke's law remains.

A more general incremental relationship may be found in the book by Hill[7]. The isotropic strain hardening law that has been utilized here has some shortcomings for repeated or cyclic loading because it does not account for the so-called Bauschinger effect. However, an expression similar to Eq. 15 based on the kinematic hardening law can also be obtained without much difficulty.

Reduction to Plane Stress

Having the application to plates in mind, the special case will be considered in which

$$\sigma_{i3} = d\sigma_{i3} = 0 \qquad i = 1, 2, 3 \tag{18}$$

and also

$$d\varepsilon_{13} = d\varepsilon_{23} = 0 \tag{19}$$

The only non-vanishing terms of the incremental constitutive relation therefore become

$$\begin{Bmatrix} d\sigma_{11} \\ d\sigma_{22} \\ d\sigma_{12} \\ d\sigma_{33} \end{Bmatrix} = \begin{bmatrix} C_{1111} & C_{1122} & C_{1112} & C_{1133} \\ C_{2211} & C_{2222} & C_{2212} & C_{2233} \\ C_{1211} & C_{1222} & C_{1212} & C_{1233} \\ C_{3311} & C_{3322} & C_{3312} & C_{3333} \end{bmatrix} \begin{Bmatrix} d\varepsilon_{11} \\ d\varepsilon_{22} \\ 2d\varepsilon_{12} \\ d\varepsilon_{33} \end{Bmatrix} \tag{20}$$

Remembering that $d\sigma_{33} = 0$, a condensed relation D is obtained, the elements of which are given by

$$D_{mn} = C_{mn} - \frac{C_{m4}\, C_{4n}}{C_{44}} \qquad \begin{matrix} m = 1, 2, 3 \\ n = 1, 2, 3 \end{matrix} \tag{21}$$

where the indices of C now refer to rows and columns of the matrix defined in Eq. 20. Numerically, it is more efficient to evaluate the elements of D by Eq. 21 rather than to employ explicit expressions for each element as suggested by Yamada[13].

Now using matrix symbols, the stress-strain relation becomes

$$d\sigma = D\, d\varepsilon - E\, d\varepsilon^{E} \tag{22}$$

where E is Hooke's law for plane stress and $d\sigma$ and $d\varepsilon$ contain the three first elements of the vectors in Eq. 20.

Using Eq. 1 and solving Eq. 22 with respect to the plastic strain increment yields

$$d\varepsilon^{P} = A\, d\varepsilon \tag{23}$$

where

$$A = I - FD \tag{24}$$

Here, I is the unit matrix and F is the inverse of E. A is generally not symmetric, even though F and D are.

The matrices D and A will have to be formed repeatedly for all integration points when using the finite element method.

SECTION III

FINITE ELEMENT EQUATIONS

General Incremental Stiffness Matrix

The principle of virtual work can be derived directly from the equilibrium equations for a solid and is valid regardless of material behavior. By applying this principle to a finite element, the following work equation arises

$$du_i^T \; S_i = \int_V d\epsilon^T \sigma \, dV \tag{25}$$

where S_i and u_i are the nodal forces and the corresponding displacements. ϵ and σ are the strains and stresses over the volume of the element.

Two configurations of the body during deformation, denoted 1 and 2, will now be considered. Assuming that these two configurations are "close" to each other and writing the virtual work expression for each configuration, the following incremental equation is obtained by subtraction

$$\Delta S_i = k_I \, \Delta u_i \tag{26}$$

where

$$\Delta S_i = S_i^2 - S_i^1 \tag{27}$$

$$\Delta u_i = u_i^2 - u_i^1 \tag{28}$$

The incremental stiffness matrix in Eq. 26 is expressed as follows,

$$k_I = \int_V B^T C_\Delta B \, dV \tag{29}$$

in which the strain interpolation matrix B defines the relation between nodal displacements and internal strains

$$\epsilon = B \, u_i \tag{30}$$

Note that small displacements are assumed, so that the same strain matrix applies to both configurations. The matrix C_Δ represents the stress-strain relationship for the increment between Configurations 1 and 2. In general, the numerical value of C_Δ will not be known because it depends on the stress path from 1 to 2, thus an approximate value will have to be used in numerical computations. The simplest approximation would be the value of C at Configuration 1; better accuracy but at the cost of greater computing effort would result from a "mean value" found by a higher order computational scheme.

Incremental Plate Stiffness

By introducing the usual restrictions on the displacement field, as stated in the Kirchhoff theory for plate bending, and making use of the incremental relationship between plane stress and strain described above, the incremental stiffness matrix can be specialized for a plate element. The Kirchhoff hypothesis limits the strains to only three components, ϵ_x, ϵ_y, λ_{xy}, which vary linearly through the thickness. The coordinate system used is shown in Fig. 3.

The displacements will now be separated into two groups: the in-plane displacements defined by

$$v = \begin{Bmatrix} u \\ v \end{Bmatrix} = \phi_v \, v_i \tag{31}$$

and the out-of-plane displacement

$$w = \phi_w \, w \tag{32}$$

where ϕ_v and ϕ_w are the interpolation polynomials and v_i and w_i are associated nodal point parameters. By appropriate differentiation of these equations, the strains can now be found, expressed in terms of two strain interpolation matrices B_v and B_w, as follows.

$$\epsilon = \begin{Bmatrix} \epsilon_x \\ \epsilon_y \\ \gamma_{xy} \end{Bmatrix} = \left[\begin{array}{c|c} B_v & zB_w \end{array}\right] \begin{Bmatrix} v_i \\ \cdots \\ w_i \end{Bmatrix} \tag{33}$$

Hence, a partitioned incremental stiffness matrix for the plate is obtained by substitution into Eq. 29.

$$k_I = \left[\begin{array}{c|c} k_v & k_{vw} \\ \hline k_{vw}^T & k_w \end{array}\right] \tag{34}$$

in which the submatrices

$$k_v = \int_V B_v^T \, D_\Delta \, B_v \, dV = \int_A B_v^T \, D_{1\,1} \, B_v \, dA \tag{35}$$

$$k_{vw} = \int_V z \, B_v^T \, D_\Delta \, B_w dV = \int_A B_v^T \, D_{1\,2} \, B_w \, dA \tag{36}$$

$$k_w = \int_V z^2 \, B_w^T \, D_\Delta \, B_w dV = \int_A B_w^T \, D_{2\,2} \, B_w \, dA \tag{37}$$

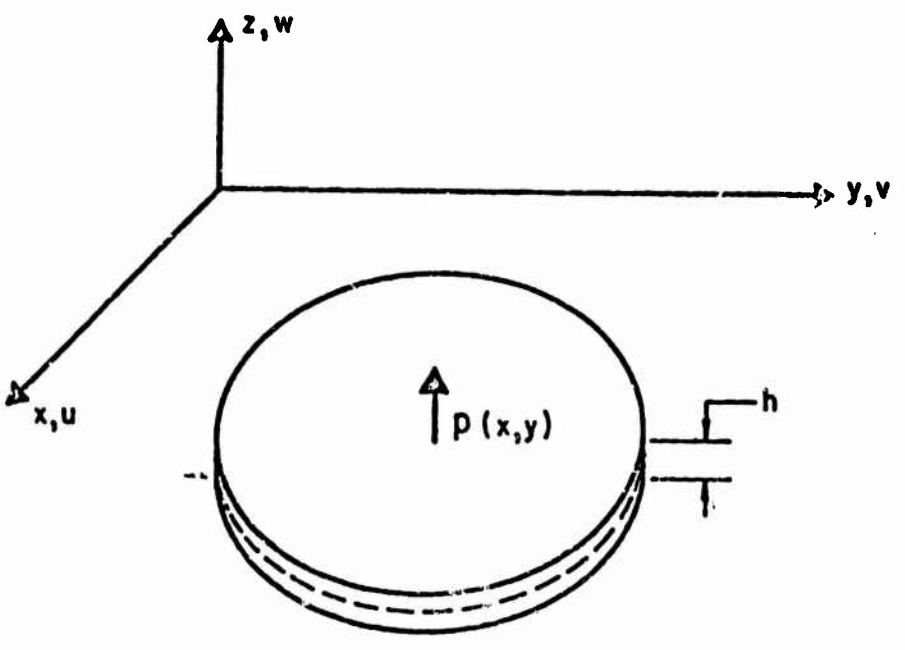


FIG. 3 DESCRIPTION OF THE PLATE IN A CARTESIAN COORDINATE SYSTEM

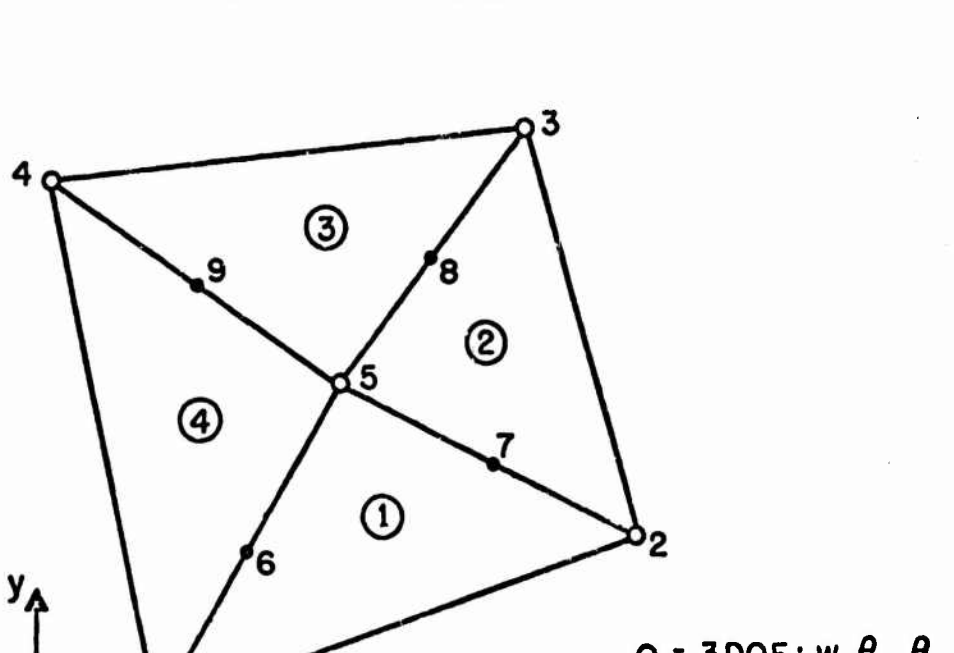


FIG. 4 THE Q19 QUADRILATERAL

represent the in-plane, coupled, and out-of-plane behavior, respectively. D_Δ is the incremental constitutive equation for plane stress defined in Eq. 21. Thus it is seen that the plate "constitutive relations" are obtained by integration through the plate thickness

$$D_{1\,1} = \int_h D_\Delta \, dz \quad , \quad D_{1\,2} = \int_h z\, D_\Delta \, dz \quad , \quad D_{2\,2} = \int_h z^2 \, D_\Delta \, dz \qquad (38)$$

Out-of-Plane Interpolation Functions

The Q-19 plate bending interpolation functions that have been adopted here already have been described completely[8], and only the main concepts of their development need be repeated for the present purpose. As is shown in Fig. 4, the Q-19 quadrilateral element is assembled from four triangle elements; the coordinates of node 5 of the quadrilateral are the averages of the coordinate of the corner nodes. An element similar to these triangles, an LCCT-12 (linear curvature compatible triangle with 12 DOF) is shown in more detail in Fig. 5. This triangle is divided into three subdomains, each having complete cubic polynomial expansions expressed by natural (triangular) coordinates for the entire triangle. By matching the deflections and rotations for all three subtriangles at the centroid 0 and enforcing continuity of slopes between these subtriangles at their midside points, all internal DOF can be eliminated. This is a tedious algebraic operation, but the resulting interpolation polynomial, which is stated completely in Reference 8, is quite simple.

The LCCT-12 bending triangle with 12 DOF can easily be reduced to an LCCT-11 triangle by prescribing a kinematic constraint such that the rotation at one midside point is set equal to the arithmetic average of the corresponding slopes at the adjacent corner nodal points. This results in a slight modification of the interpolation polynominals of the LCCT-12 element.

In-Plane Interpolation Functions

A set of in-plane displacement functions defined over a general quadrilateral domain will be needed to develop a "membrane" element consistent with the "bending" element. Only two in-plane DOF's at each corner will be used here for the sake of simplicity. An isoparametric element of this type has been described by Zienkiewicz[9]. The present element has been improved somewhat by including two internal interpolation functions associated with the displacements at the midpoint and then condensing out these internal DOF's. The displacement functions of this element are expressed in the natural quadrilateral coordinates.

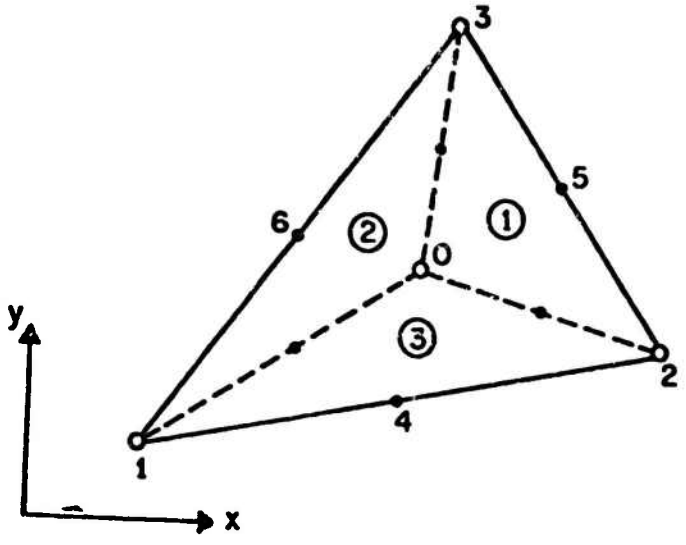


FIG. 5 THE LCCT 12 ELEMENT

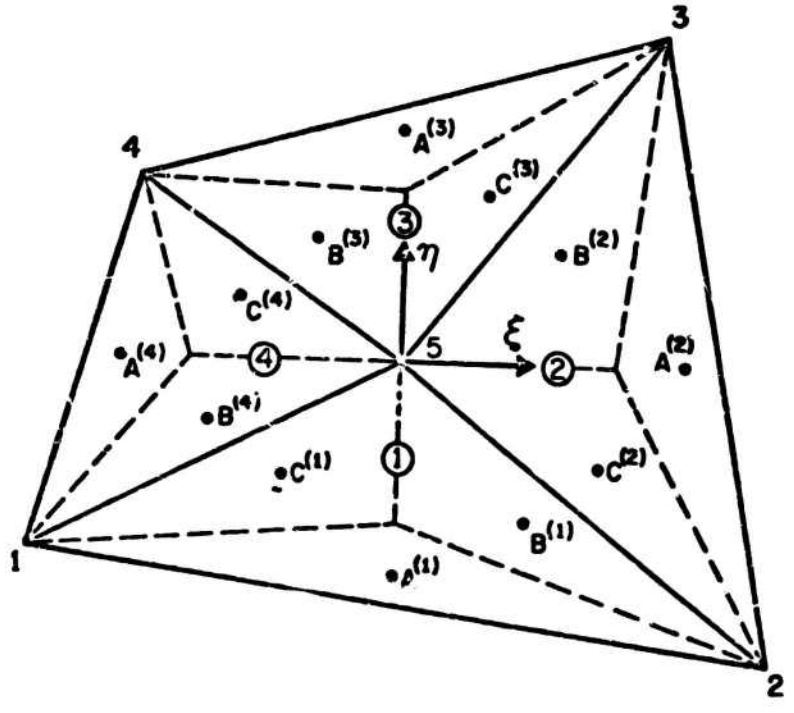


FIG. 6 INTEGRATION POINTS FOR ELASTIC-PLASTIC PLATE BENDING ELEMENT

Numerical Integration of the Element Stiffness

It will be noted in Eqs. 35-38 that the plate stiffness matrix is computed by evaluating integrals involving the constitutive relations, over the area and through the thickness of the element. The integration of the bending stiffness submatrices (Eq. 35) may be done separately for each of the four triangular areas making up the quadrilateral, because the LCCT strain interpolations are defined independently in each. Similarly, for the coupled membrane-bending stiffness submatrices (Eq. 36) the integrations may be performed for each triangle separately. However, the integration of the in-plane stiffness (Eq. 37) is performed over the entire quadrilateral because the isoparametric interpolation functions apply to the full region.

In developing the Q-19 bending stiffness for linear elastic materials, the area and thickness integrations were performed explicitly, the area integration being done in area coordinates. However, a basic problem arises in the explicit integration of stiffness matrices which represent non-linear material behavior: the material properties are not uniform throughout the element volume. In some non-linear finite element analyses, the variation of properties within the element has been expressed by additional interpolation functions, thus making explicit integration possible[14]. However, this approach would be extremely complicated if applied with the Q-19 interpolation functions described above. Moreover, the incremental stress-strain relationship frequently is a discontinuous function within the element.

In the present study, the integration over the area of each LCCT-11 bending triangle has been replaced by a simple summation of the contributions from its three subtriangles. First, the strains and incremental constitutive relations are evaluated at the centroid of each subtriangle (defined, for example, by area coordinates 4/9, 4/9, 1/9 in subtriangle 3 of Fig. 5). Then it is assumed that these local values apply over the entire subtriangle, and the stiffness of the entire LCCT-11 element is found from the following expression

$$k_{wt} = \frac{A}{3} \sum_{j=1}^{3} B_{wc}^{(j)T} D_{22}^{(j)} B_{wc}^{(j)} \qquad (39)$$

in which the superscript identifies the subtriangle and the subscript c indicates that the quantities are evaluated at its centroid. The constitutive relation matrix $D_{22}^{(j)}$ is evaluated by integration through the element thickness at the centroid of element "j", but instead of using the explicit expression of Eq. 38 the computation is performed by Gauss quadrature. Generally an 11-point scheme is used; however, only half of the points need be considered in the case of pure bending or if the material is linear elastic because of symmetry about the mid-plane.

The stiffness k_w of the complete Q-19 element is formed by assembling the four LCCT-11 triangles. If there is no coupling stiffness, the internal

DOF may be condensed out at this stage resulting in a 12 DOF quadrilateral. Although this integration scheme appears quite crude, it is surprisingly effective because each quadrilateral is represented by 12 sub-areas. As a demonstration of the element efficiency, a square simply supported plate subjected to uniform pressure was analyzed using a 2 by 2 quadrilateral mesh for one quarter of the plate. The central deflection obtained from the numerically integrated elements was only 0.4 percent greater than that given by elements integrated exactly.

The in-plane stiffness of an isoparametric quadrilateral element generally is integrated numerically by Gauss quadrature, using a 2 by 2 or a 3 by 3 system of integration points over the area. This integration scheme could be used with non-linear material properties, as well as for the linear elastic case. However, in order to evaluate the coupled stiffness submatrices in the present study, it is necessary to evaluate the in-plane strain interpolations at the same integration points associated with the bending interpolations. Therefore, the in-plane stiffness also is calculated using the same 12 integration points and the same type of summation scheme described for the Q-19 element. In this case, the contributions from the 12 subtriangles may be superposed directly, thus the in-plane stiffness for the entire quadrilateral is given by:

$$k_v = \sum_{m=1}^{12} A^{(m)} \, B_{vc}^{(m)T} \, D_{11}^{(m)} \, B_{vc}^{(m)} \qquad (40)$$

in which $A^{(m)}$ represents the area of subtriangle m, and the strain interpolations are evaluated at the subtriangle centroids. As in the bending analysis, the constitutive relation $D_{11}^{(m)}$ is evaluated numerically at the centroid by Gauss quadrature through the thickness. However in this case only a single point is needed (at mid-depth) unless both non-linear materials and combined membrane-bending conditions are involved.

One difficulty is encountered in evaluating the in-plane strains at the subtriangle centroids: the isoparametric interpolation functions are expressed in quadrilateral coordinates while the subtriangles are defined in triangular coordinates. However, noting that the midpoint of the Q-19 element is the same as the origin of the ξ-η quadrilateral coordinates, the following relations may be derived between the two natural coordinate systems on triangle 1 (see Fig. 6):

$$\xi = -\zeta_1^{(1)} + \zeta_2^{(1)} \quad ; \quad \eta = -\zeta_1^{(1)} - \zeta_2^{(1)} \qquad (41)$$

where $\zeta_1^{(1)}$ and $\zeta_2^{(1)}$ are area coordinates of triangle 1. From these, and the triangular coordinates of the subtriangle centroids, the quadrilateral coordinates of these points may be derived for triangle 1, as follows:

	Triangle			Isoparametric	
Subtriangle Centroid	$\zeta_1^{(1)}$	$\zeta_2^{(1)}$	$\zeta_3^{(1)}$	ξ	η
$A^{(1)}$	4/9	4/9	1/9	0	-8/9
$B^{(1)}$	1/9	4/9	4/9	1/3	-5/9
$C^{(1)}$	4/9	1/9	4/9	-1/3	-5/9

Similar relations are easily obtainable for triangles 2, 3, and 4.

The coupling stiffness submatrix (Eq. 36) will exist only when the element has non-linear material properties and it is subjected to both flexural and in-plane deformations. The stiffness contribution of each triangular area of the element is obtained from the membrane and bending strain interpolations evaluated at the three subtriangle centroids, using the formula

$$k_{vwt} = \frac{A}{3} \sum_{i=1}^{3} B_{vc}^{(j)T} D_{12} B_{wc}^{(j)} \qquad (43)$$

in which the notation is similar to Eq. 40. The constitutive relation matrix $D_{12}^{(j)}$ is evaluated by integrating the expression of Eq. 38 using 11 point Gauss quadrature. In the case when $D_{12}^{(j)}$ is non-zero, the corresponding expressions for $D_{11}^{(j)}$ and $D_{22}^{(j)}$ also require the use of all 11 points in their integration.

The stiffness k_{vw} of the complete quadrilateral is obtained by assembling the results given by Eq. 43 for the four triangular areas. Then the system of four submatrices shown in Eq. 34 (which has dimensions 27 x 27) may be reduced to a 20 DOF stiffness matrix by condensing out the seven internal DOF. The result represents the incremental coupled membrane-bending stiffness matrix of the quadrilateral element.

SECTION IV

EXAMPLES

Computer Program

A Fortran IV computer program was developed according to the previous derivations. The main equations associated with this new non-linear capability are those of the plasticity theory, Eqs. 1, 12, 15 and 17, and the stiffness relations, Eqs. 39, 43 and 45. Other operations are the standard ones for the finite element method.

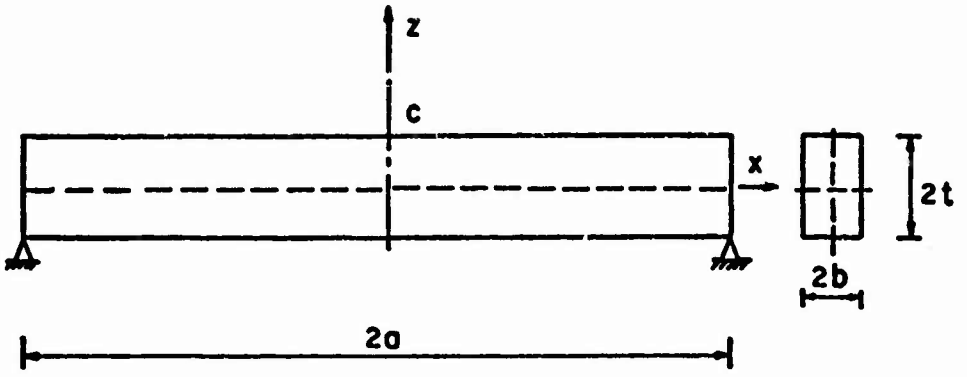
In this program the strain-displacement relationships are computed once only and retained during application of the entire sequence of incremental loads. The initial load is automatically scaled so that initial yielding is obtained at the integration point having the highest equivalent stress. The experimental stress-strain curve is identified in the input by a set of discretized values; linear interpolation is performed for intermediate points. If a new state of stress lies outside the current yield surface according to the $\bar{\sigma} - \bar{\varepsilon}$ diagram (due to the error of increment linearization), the stresses are scaled linearly back to the yield surface.

A few examples are presented in the following to illustrate the capacity of this formulation. The non-linear solutions are based on a simple Euler-Cauchy incremental technique. This method has certain advantages from a computational point of view because stress-computation and stiffness formulation can be performed simultaneously.

Simple Beam

The first example deals with the bending of a simply supported, rectangular beam subjected to uniform loading. In order to demonstrate both the bending and the membrane capabilities of this element, two different analyses were performed - one involving out-of-plane loading, the other loaded in-plane. The dimensions of the beam and the two finite element idealizations used are shown in Fig. 7.

The results of the two analyses in which the material was assumed to be elastic, perfectly plastic are plotted in Fig. 8. In this figure, the abscissa ρ is a non-dimensional loading factor for which the unit value corresponds to full yielding. The ordinate represents the ratio of the current value of mid-point deflection to the value at initial yielding. For comparative purposes, the results of a closed form solution presented by Prager and Hodge[15] also are plotted in the figure. The good elasto-plastic performance given by only four bending elements is evident. On the other hand, 20 membrane elements (neglecting symmetry) were required to obtain similar accuracy because these elements are not well adapted to representing bending. Figure 8 also shows



GEOMETRY OF BEAM

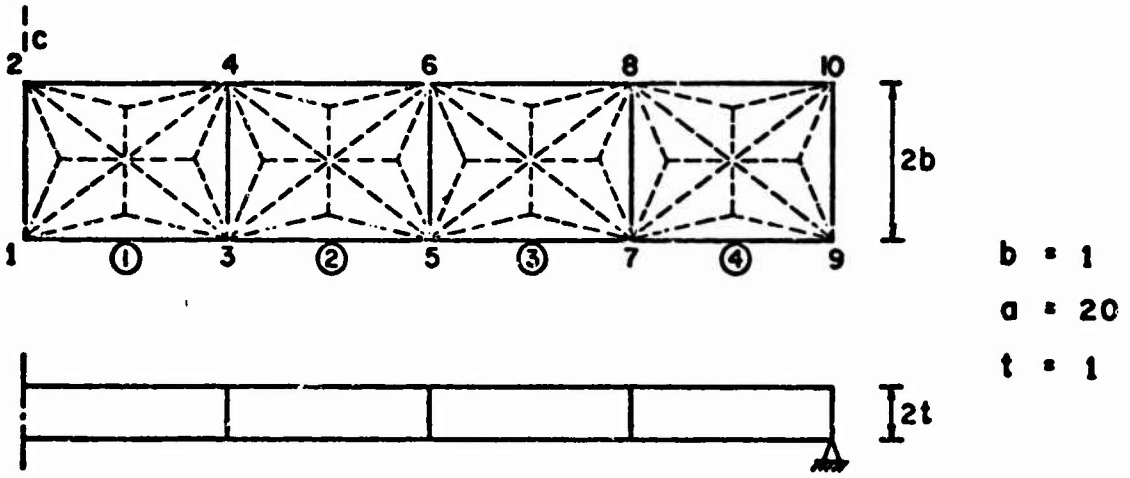
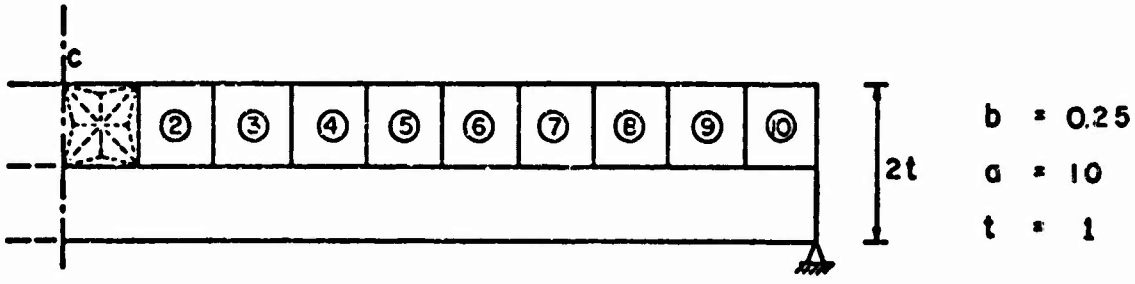
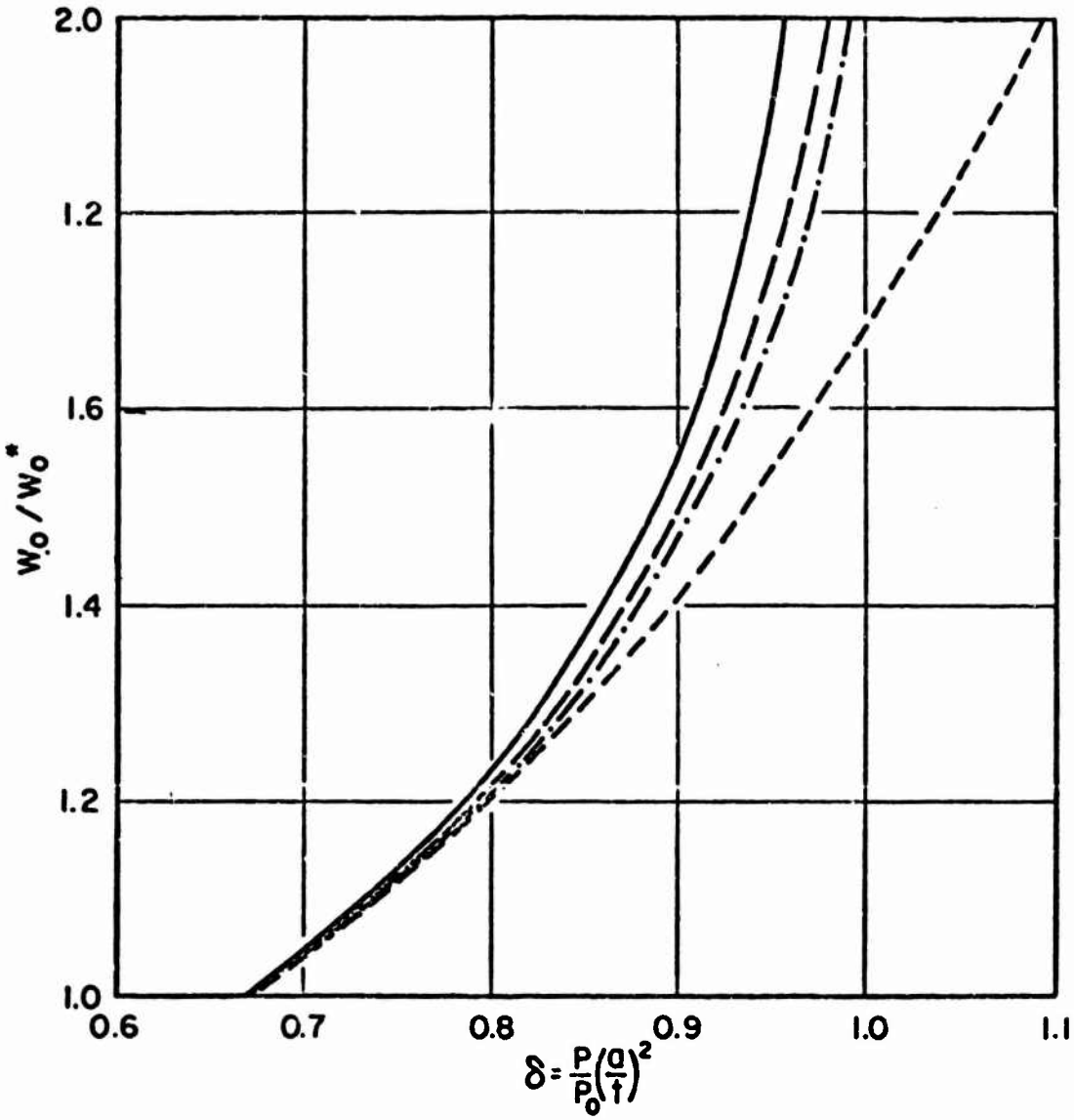


PLATE BENDING ELEMENT IDEALIZATION (1/2 BEAM)



MEMBRANE ELEMENT IDEALIZATION (1/4 BEAM)

FIG. 7 FINITE ELEMENT IDEALIZATION FOR ELASTIC PLASTIC BEAM



——— = PRAGER AND HODGE

—·—·— = 4 PLATE BENDING ELEMENTS

— — — = 10 MEMBRANE ELEMENTS

— — — — = 4 PLATE BENDING ELEMENTS, E_T = 0.25 E

FIG. 8 ELASTIC - PLASTIC BENDING OF BEAM

results obtained with the plate bending elements, when a bilinear material property is assumed for the beam for which $E_T = 1/4\ E$ after yielding.

Square Plate

Figure 9 shows a square, simply supported plate made of an elastic, perfectly plastic material. One quarter of the plate has been idealized by 2 by 2 and by 4 by 4 meshes. A graph of the mid-point deflection due to a uniform pressure also is shown in the figure. The maximum deflection obtained in this analysis can be compared with a yield line "ultimate load" analysis and an upper bound analysis given by Armen, et al[4].

Zones having three different depths of yielding through the thickness, due to a loading factor of $\rho = 0.99$, are indicated by different shading patterns in Fig. 10. The development of an incipient yield line mechanism along the diagonal of the plate is evident.

Trapezoidal Plate

The last results to be discussed are those obtained for the plate of trapezoidal shape, Fig. 11, subjected to uniform pressure. Again the material was assumed to be elastic, perfectly plastic. One half of the plate was idealized by a 4 by 4 mesh, and the pressure was applied in increments equal to 10 percent of the initial yield load. The deflection of the midpoint of of the free edge through 25 load increments is plotted in Fig. 12. The deflected shape of the free edge at two different load levels, and the yield line theory shape are sketched in Fig. 13; these curves show how the yielding process gradually changes the deflection patterns. The extent of the yielding determined at three different pressure levels are shown in Figs. 14, 15 and 16, which give some idea of the propagation of the yield zones during loading.

Further details concerning these and several other examples may be found in Ref. 12.

Computer Time Requirements

All computations in the preceding examples were made with a CDC 6400 computer (65K memory). The computer time requirements for forming the element stiffness for one quadrilateral element consisting of 12 subregions are indicated in the following tabulation.

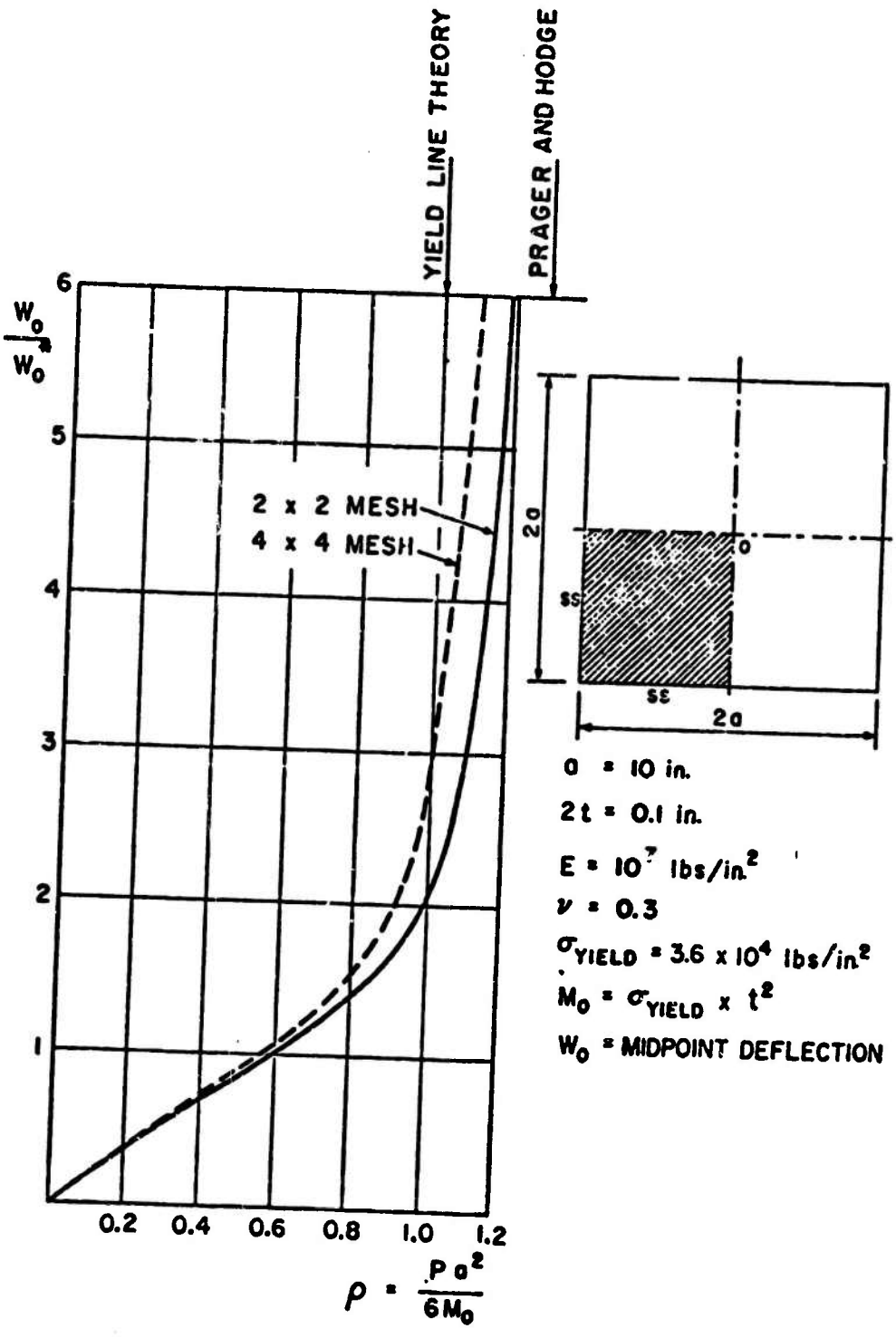


FIG. 9 LOAD DEFLECTION CURVE FOR SIMPLY SUPPORTED SQUARE PLATE

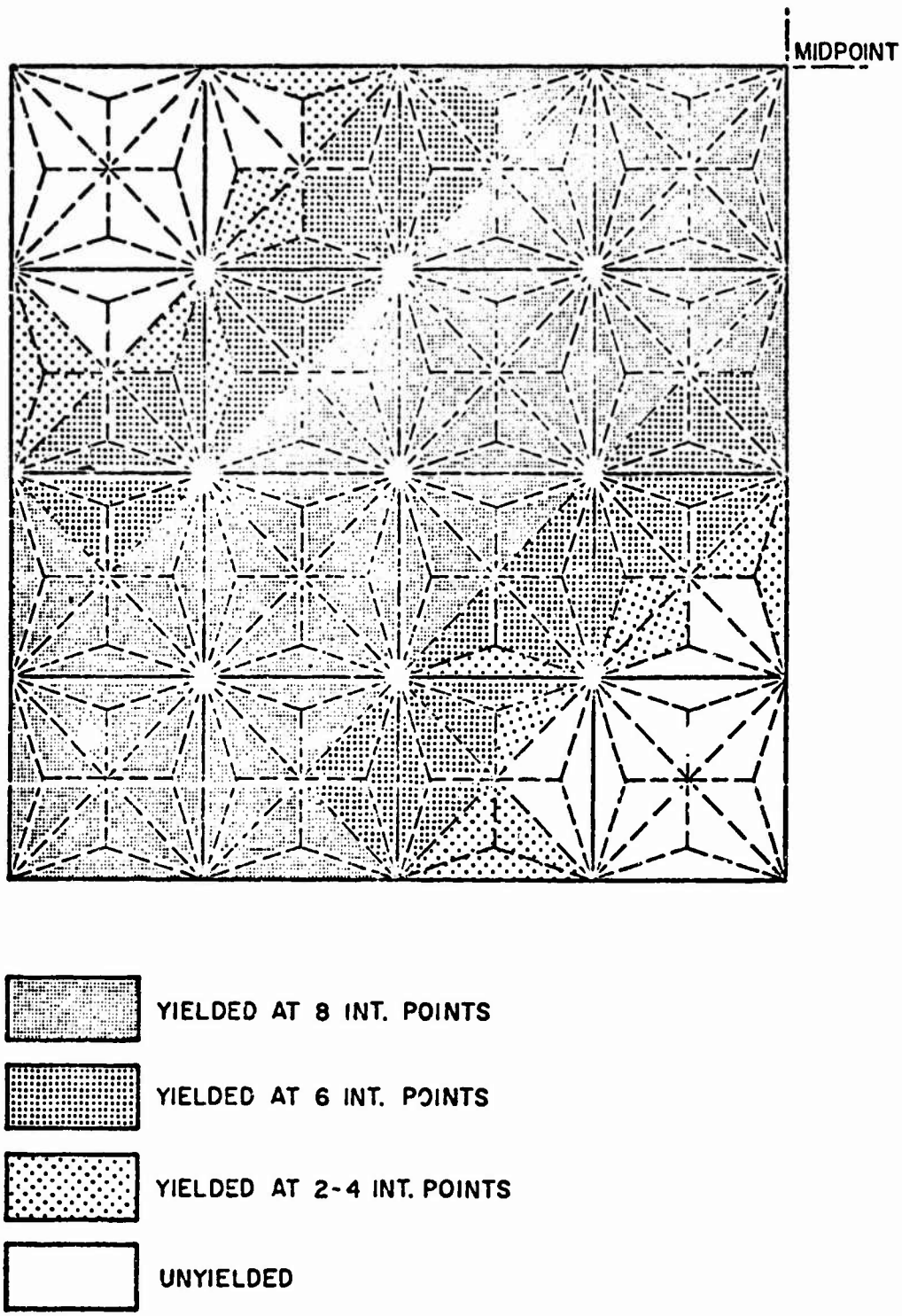


FIG. 10 EXTENSION OF YIELD ZONES FOR $\rho = 0.99$
4 by 4 MESH

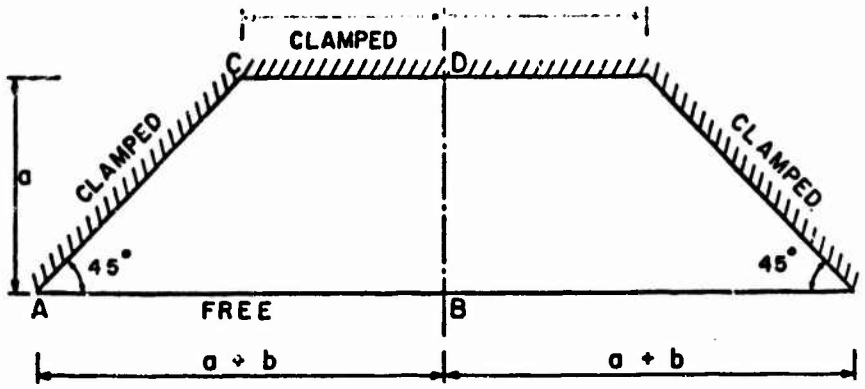


FIG. 11 TRAPEZOIDAL PLATE

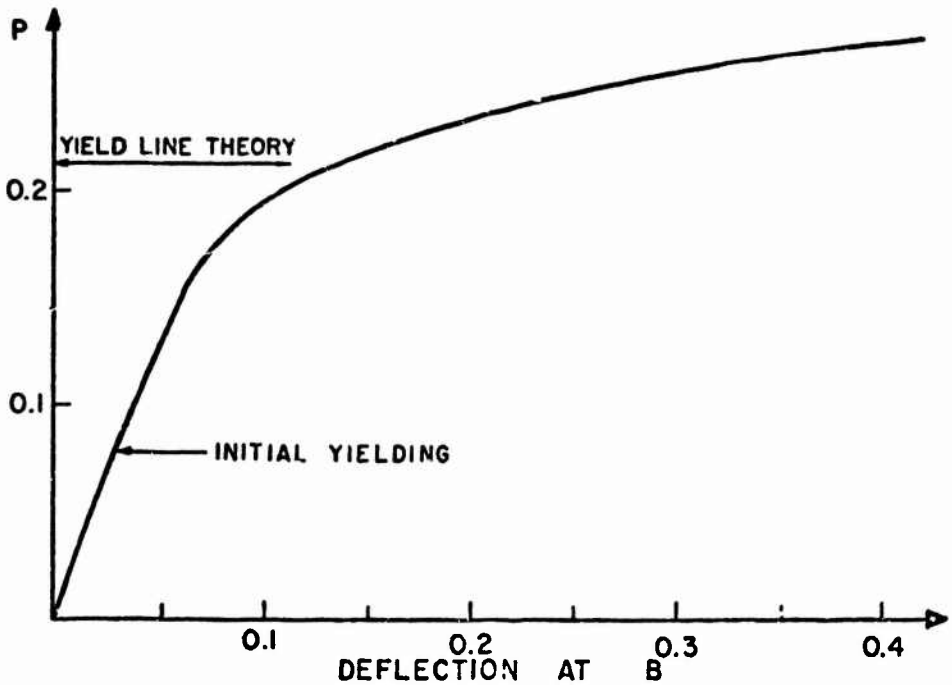


FIG. 12 LOAD-DEFLECTION CURVE FOR TRAPEZOIDAL PLATE

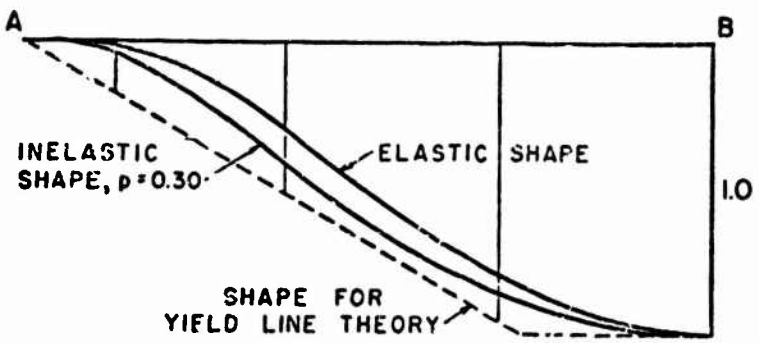


FIG. 13 CHANGES IN DEFLECTED SHAPE ALONG LINE A-B

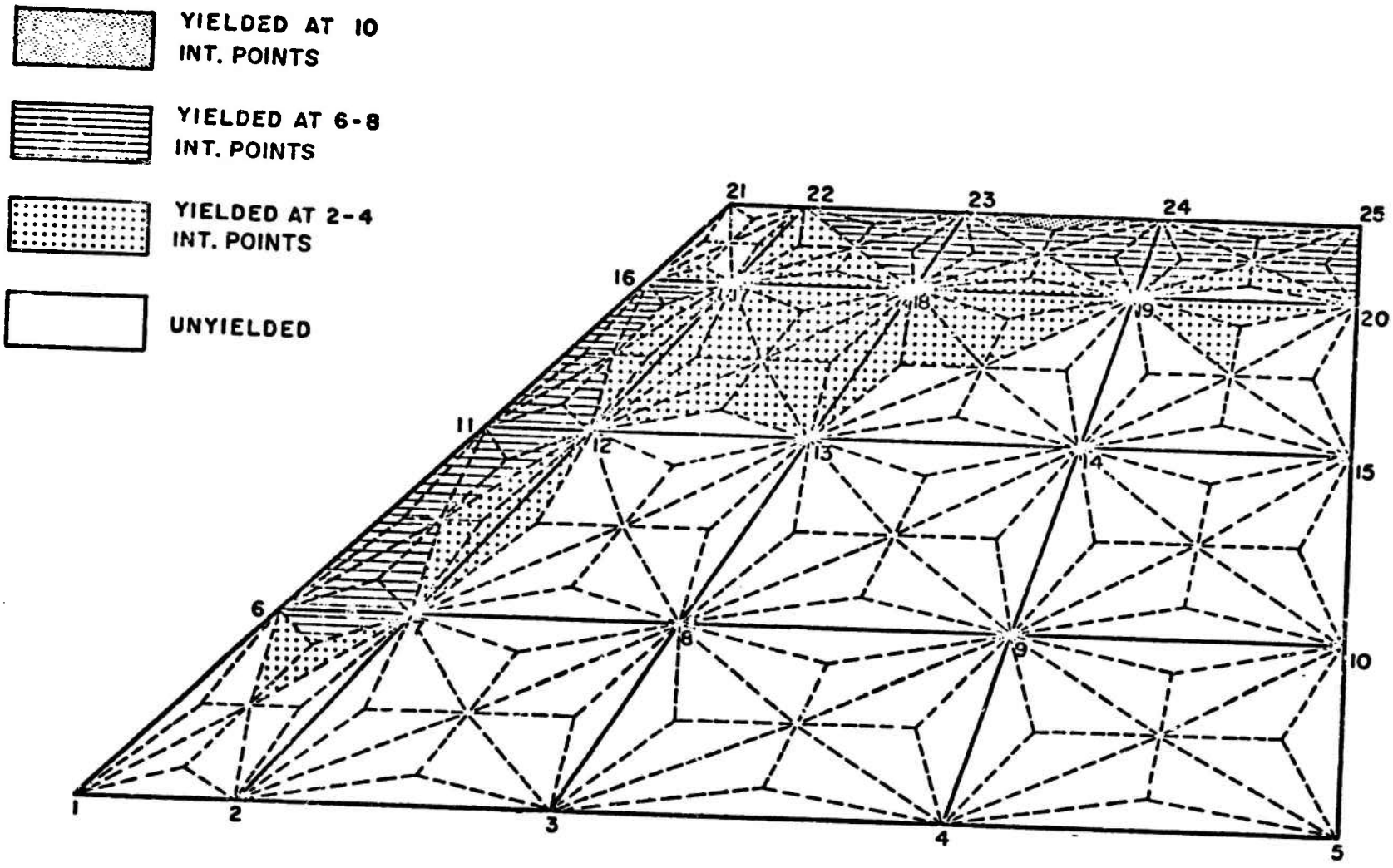


FIG. 14 EXTENSION OF YIELD ZONES FOR p = 0.173 (PSI)

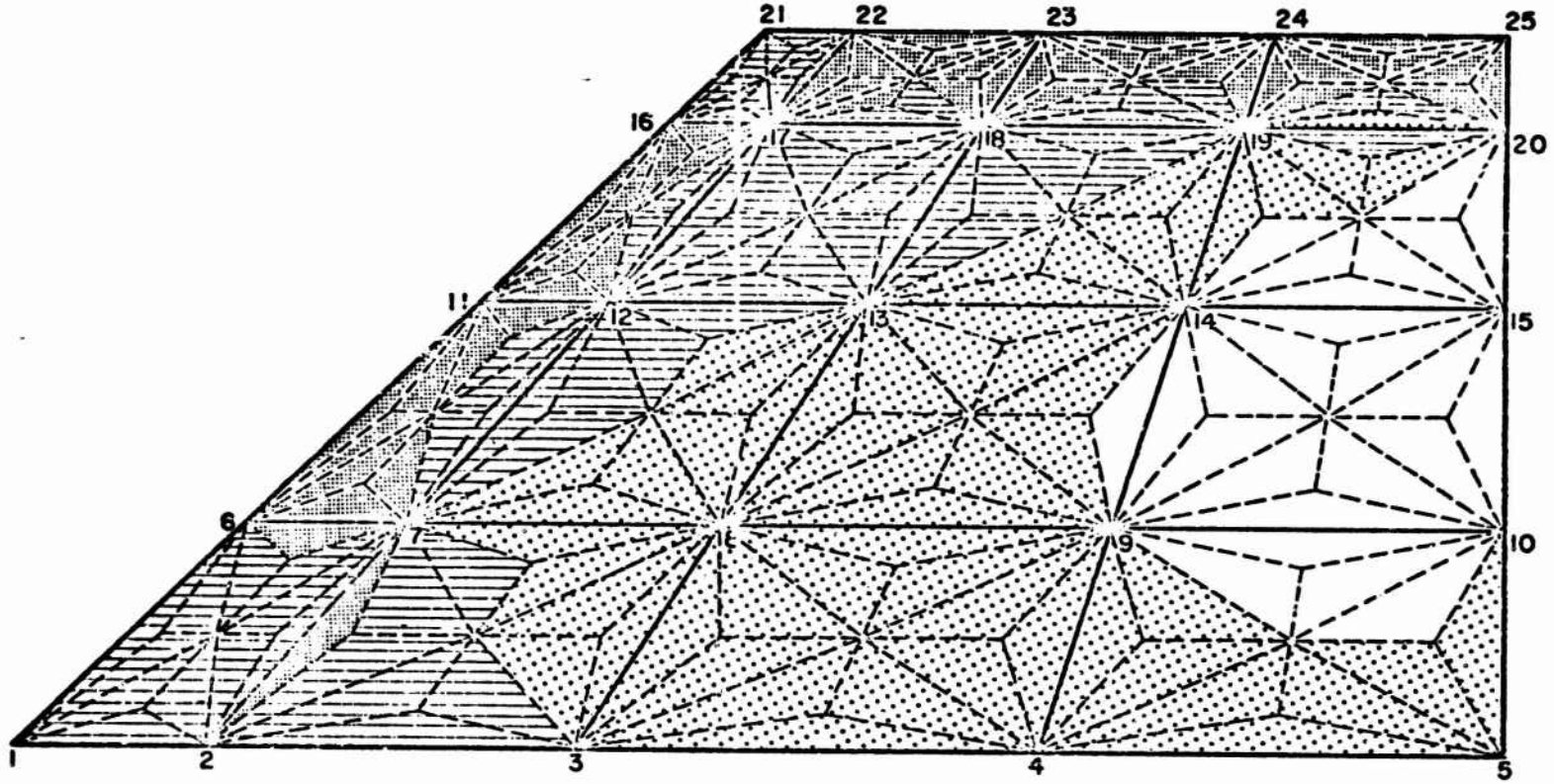


FIG. 15 EXTENSION OF YIELD ZONES FOR p = 0.221 (PSI)

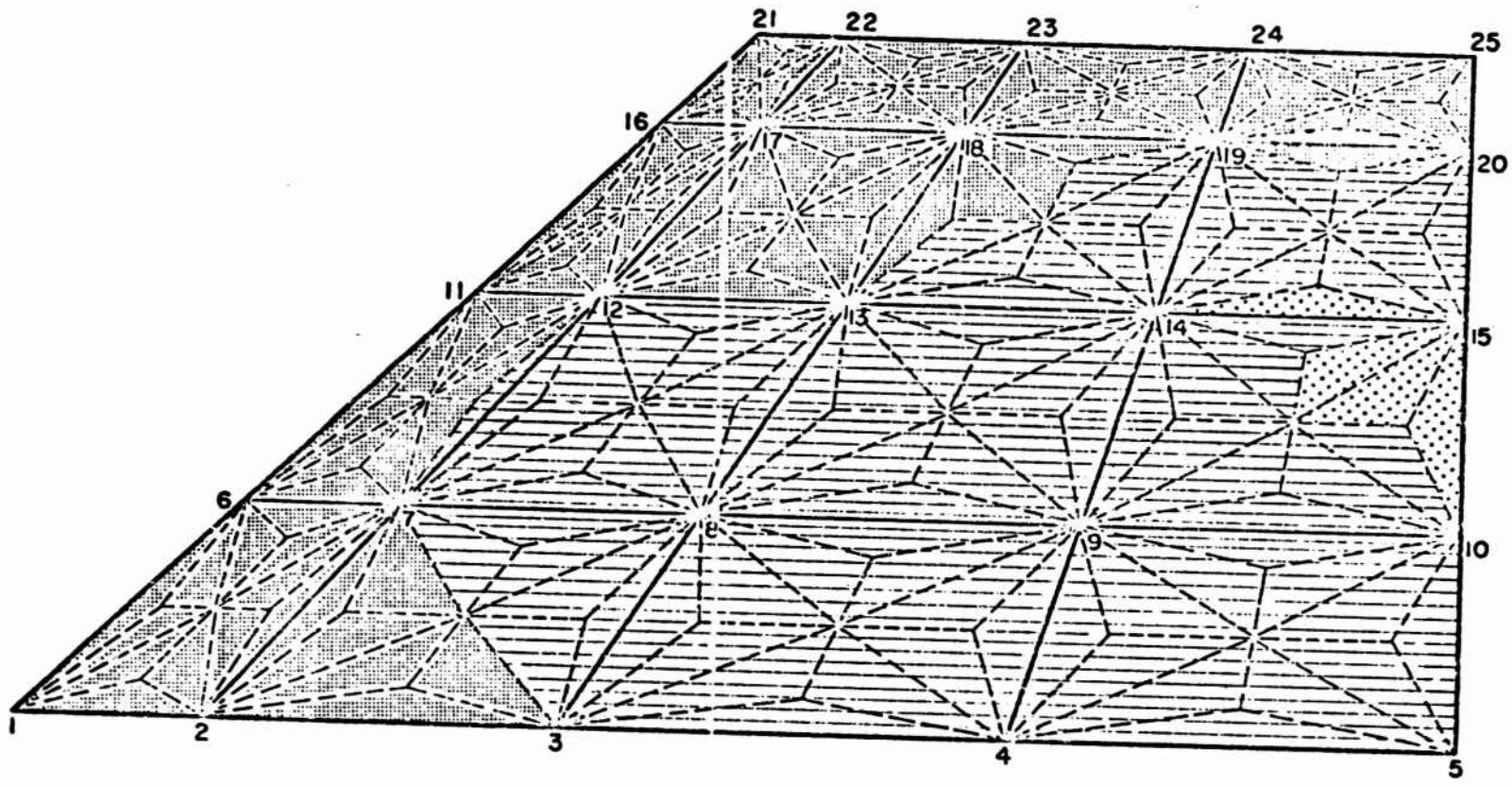


FIG. 16 EXTENSION OF YIELD ZONES FOR p = 0.268 (PSI)

Operation	Time CP secs.
Incremental stiffness for inelastic plate bending element, 11 int. points through thickness	0.170 - 0.330
Incremental stiffness for inelastic membrane element	0.125 - 0.160
Incremental stiffness for combined bending and membrane action, 11 int. points through thickness	0.540 - 0.690

The computer time requirements for performing the analyses of some of the previous examples are:

Example	No. of elements	No. of load increments	Time CP secs.
Inelastic beam, bending element	4	19	24
Inelastic beam, membrane element	10	19	35
Trapezoidal plate	16	29	171

SECTION V

CONCLUSIONS

1. The Q-19 bending element may be combined with the linear isoparametric membrane element to provide an efficient element for treating coupled membrane-bending effects.

2. The area integration required in the element stiffness evaluation may be approximated effectively by summing terms evaluated at the centroids of the 12 subtriangles of the Q-19 element. Thickness integrations are done most efficiently by Gauss quadrature.

3. Non-linear analyses may be accomplished efficiently using simple incremental load analyses based on tangent stiffness defined by the above-mentioned integration processes. Experience shows that large coupled membrane-bending problems can be handled with reasonable expenditures of computer time.

REFERENCES

1. Argyris, J. H., Kelsey, S. and Kamel, H., "Matrix Methods of Structural Analysis - A Precis of Recent Developments," Agardograph 72, pp. 1-164, Permagon Press, 1964.

2. Marcal, P. V., and King, I. P., "Elastic-Plastic Analysis of Two-Dimensional Stress Systems by the Finite Element Method," Int. J. Mech. Sci., Vol. 9, No. 3, pp. 143-145, 1967.

3. Khojasteh-Bakht, M., "Analysis of Elastic-Plastic Shells of Revolution under Axisymmetric Loading by the Finite Element Method," Ph.D. Dissertation, University of California, Berkeley, Report No. SESM 67-8, 1967.

4. Armen, H., Pifko, A., and Levine, H. S., "Finite Element Method for the Plastic Bending Analysis of Structures," Proc. Second Conf. on Matrix Methods in Struct. Mechanics, AFFDL-TR-68-150, 1968.

5. Argyris, J. H., and Scharpf, D. W., "Methods of Elasto-plastic Analysis," Proceedings ISO-ISSC Symposium on Finite Element Techniques at the ISD, Stuttgart, Germany, 1969.

6. Marcal, P. V., "Finite Element Analysis with Material Nonlinearities-Theory and Practice," Proc. Japan-U.S. Seminar on Matrix Methods of Structural Analysis and Design, Tokyo, 1969.

7. Hill, R., The Mathematical Theory of Plasticity, Oxford University Press, 1950.

8. Clough, R. W. and Felippa, C. A., "A Refined Quadrilateral Element for Analysis of Plate Bending," Proc. Second Conf. on Matrix Methods in Struct. Mech., AFFDL-TR-68-150, pp. 399-440, 1968.

9. Zienkiewicz, O. C., The Finite Element Method in Structural and Continuum Mechanics, McGraw-Hill Publ. Company, 1967.

10. Marcal, P. V., "Comparative Study of Numerical Methods of Elastic-Plastic Analysis," AIAA Journal, Vol. 6, No. 1, pp. 157-158, 1967.

11. Naghdi, P. M., "Stress-Strain Relations in Plasticity and Thermoplasticity," Plasticity, Proceedings of the Second Symposium on Naval Structural Mechanics, (editors: Lee, E. H. and Symonds, P. S.), Pergamon Press, 1960

12. Bergan, P. G., "Nonlinear Analysis of Plates Considering Geometric and Material Effects," Ph.D. Dissertation, University of California, Berkeley, March, 1971.

13. Yamada, Y., "Recent Japanese Developments in Matrix Displacement Method for Elastic-Plastic Problems," Proc. Japan-U.S. Seminar on Matrix Methods of Structural Analysis and Design, Tokyo, 1969.

14. Felippa, C. A., "Refined Finite Element Analysis of Linear and Nonlinear Two-Dimensional Structures," Ph.D. Dissertation, University of California, Berkeley; first part published as SESM Report 66-22, 1966.

15. Prager, W. and Hodge, P. G., Theory of Perfectly Plastic Solids, Dover Publications, New York, 1951.

ELASTIC AND PLASTIC INTERLAMINAR SHEAR DEFORMATION IN LAMINATED COMPOSITES UNDER GENERALIZED PLANE STRESS*

A. Levy†
H. Armen, Jr.†
J. Whiteside‡

Grumman Aerospace Corporation
Bethpage, New York 11714

Elastic and plastic interlaminar shear deformation in a laminated fibrous composite is studied by means of a finite-element method. A composite element, constructed of orthotropic membranes separated by shear-resisting media, is developed and utilized. The effect of interlaminar shear deformation at the free edges of solid laminates and laminates containing a cutout is presented.

I. INTRODUCTION

Interlaminar shear deformation, the mechanism by which load is transferred through a matrix material between two stiff laminae as the laminae tend to slide over each other, cannot be predicted by using classical plate theory. This type of deformation develops along the edges of a laminate and can be important with respect to strength predictions of composite structures, especially for a laminate with a relatively low transverse shear strength. Analogous behavior is found to exist in bonded structural joints.

Recognition of this phenomenon has generated interest in the development of analytic methods that can account for interlaminar shear deformation. Recently, this type of deformation has been described from three different viewpoints (Refs. 1-3) for the elastic behavior of a rectangular laminate. In Ref. 1 an exact

*These results were obtained in the course of work sponsored by the Air Force Flight Dynamics Laboratory under Contract F33615-70-C-1308.

†Research Scientist

‡Structural Mechanics Engineer

solution is presented for a model in which the fiber-bearing layers are replaced by homogeneous membranes with orthotropic properties and are separated by layers that develop only interlaminar shearing stresses. In Ref. 2, the exact equations of elasticity are solved by using a finite difference method. Here each lamina is assumed to be homogeneous, so that there is no physical separation of fibers and matrix through the thickness. Both solutions are applicable for the analysis of a finite laminate in regions removed from the loaded ends. In Ref. 3, a finite-element method is employed, using a model similar to that of Ref. 1, to analyze a rectangular panel of $\pm 45°$ angle ply construction.

In this paper, a finite-element method is employed to investigate the interlaminar shear behavior along straight boundaries and around circular cutouts in balanced laminates. The laminates are balanced in the sense that the laminae are symmetrically placed about the midplane so that no warping exists, that is, there is no bending generated as a result of stretching. From the previous studies (Refs. 1-3), there is sufficient evidence to indicate that the interlaminar shear stress may be of sufficient magnitude to cause plastic deformation of the matrix material. Therefore, a method to account for the effects of plasticity is included in the analysis. In addition, consideration is given to cyclic loading conditions involving reversed plastic deformation. This study is limited in that it provides for interlaminar plastic deformation only, i.e., the elements representing the in-plane behavior are assumed to remain elastic throughout the entire load history, while those elements representing the interlaminar shear behavior can assume elastic-plastic properties. While this does not provide a complete description of the coupled situation resulting from both types of plastic deformation, it does enable us to isolate the effect of interlaminar plastic deformation.

A brief description of the element and formulation of the governing matrix equation is followed by results for rectangular laminates in both the elastic and inelastic ranges. These results, obtained for solid laminates and laminates containing a circular cutout, show the effect of interlaminar shear deformation at free edges.

II. THE COMPOSITE ELEMENT

The idealized model separates the membrane and interlaminar properties of a laminated composite by using alternating orthotropic fiber-bearing segments and isotropic shear segments, as

shown in Fig. 1. The orthotropic segments carry in-plane stresses only, and may be considered to be in a state of plane stress; the shear segments carry only interlaminar shear stresses, and are in a state of pure shear. This is the same model as that used in Refs. 1 and 3.

The composite element consistent with the idealized model is shown in Fig. 2. The membrane segments are triangular orthotropic elements in which the total strains are assumed to be uniform (Ref. 4). Here the strain-displacement relation is based on a linearly varying displacement field. The stiffness properties of the interlaminar shear segments are also based on a linear displacement field, so that the shear strains may be written in terms of the nodal displacements in the following manner:

$$\gamma_{xz} = \frac{\partial u}{\partial z} = \left(u_i^{\ell-1} + u_j^{\ell-1} + u_k^{\ell-1} - u_i^{\ell} - u_j^{\ell} - u_k^{\ell}\right) / 3t$$

$$\gamma_{yz} = \frac{\partial v}{\partial z} = \left(v_i^{\ell-1} + v_j^{\ell-1} + v_k^{\ell-1} - v_i^{\ell} - v_j^{\ell} - v_k^{\ell}\right) / 3t \quad ,$$

where u and v are displacements in the x and y directions, respectively, subscripts identify element vertices, and superscripts identify element faces as shown in Fig. 2. Since the displacements vary linearly in the plane of a membrane segment, the interlaminar shear strain is computed on the basis of centroidal values of displacement, and thus the shear segment may be regarded as a shear-resisting medium connecting the centroids of adjacent membrane segments. Any number of segments may be stacked through the thickness to form the multilayered composite element.

III. FORMULATION OF GOVERNING MATRIX EQUATION

The formulation of the governing matrix equation is developed within the framework of the displacement method of finite-element analysis and follows that method as presented in Ref. 5. Accordingly, assumptions concerning the displacement field within an individual element are made in terms of discrete quantities at node points. In addition, independent assumptions may be made concerning the distribution of initial strain within an element. These assumptions concerning displacements and initial strains are used to derive the force-displacement relations for an individual element. This is accomplished by application of the principle of

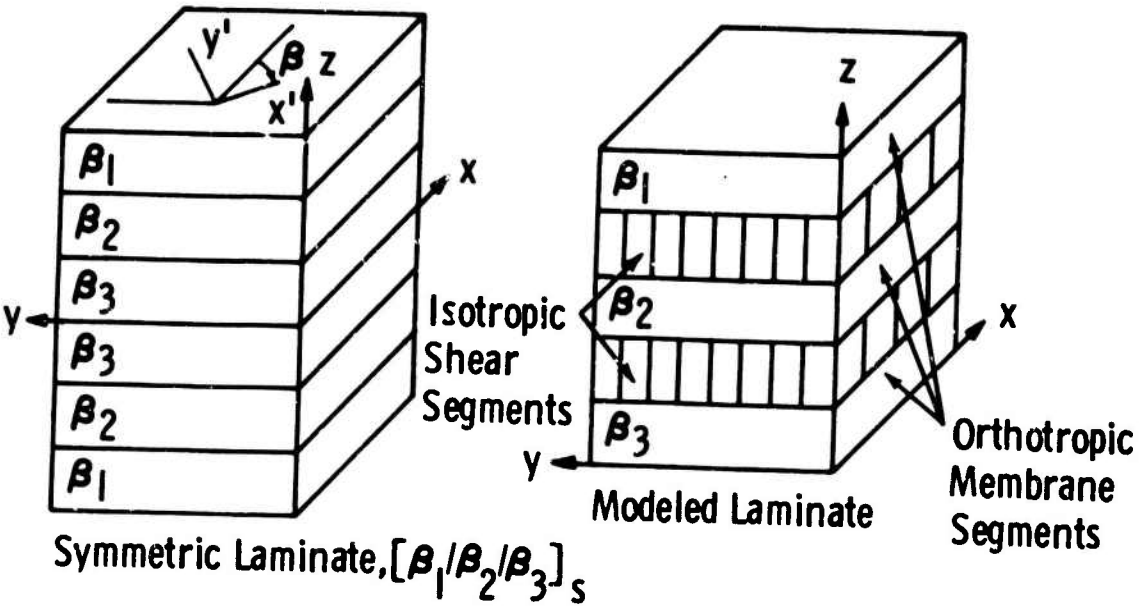


Figure 1. Laminate Model

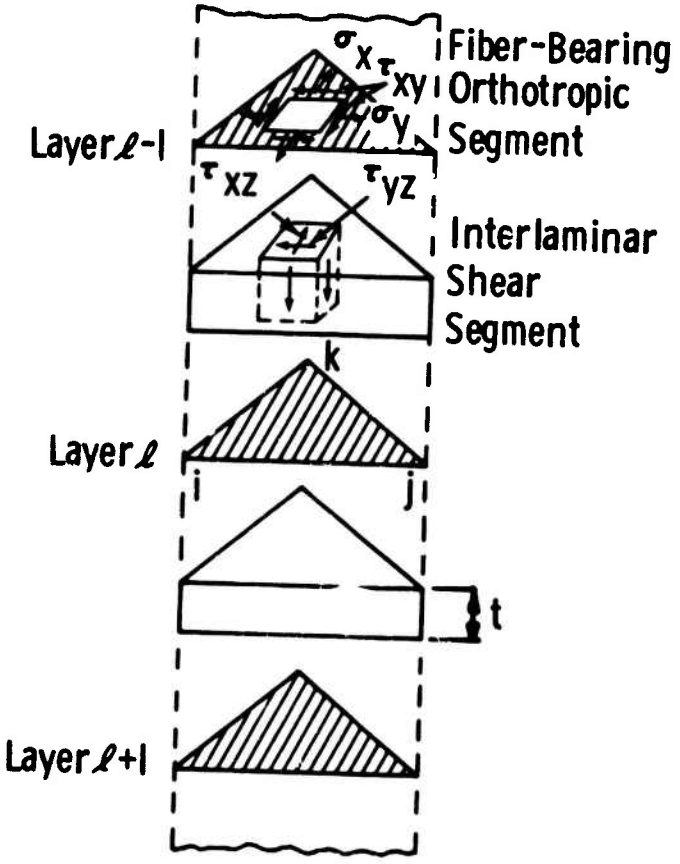


Figure 2. Composite Element

virtual work or through a consistent energy approach. The technique used to incorporate the effects of plastic behavior into a finite-element analysis is based on the initial strain concept (Ref. 5), according to which plastic strains (interlaminar shear components only) are interpreted as initial strains. At any step of the loading process, this procedure leads to the following governing linear matrix equation, in which generalized displacements, $\underset{\sim}{u}$, are related to applied loads $\underset{\sim}{f}$ and to the plastic shear strains, $\underset{\sim}{\gamma}_p$,

$$\underset{\sim}{f}^i = \underset{\sim}{k}\,\underset{\sim}{u}^i - \underset{\sim}{k}_p\underset{\sim}{\gamma}_p^i \tag{1}$$

where the superscript i refers to the current load step and where $\underset{\sim}{k}$ and $\underset{\sim}{k}_p$ represent the element stiffness and initial strain stiffness matrices, respectively. These element matrices are defined in Appendix A for a linear displacement field element in which the components of both total and plastic strains are uniform throughout.

If, in Eq. (1) the product of the initial strain stiffness matrix and the vector of plastic shear strains are considered as a vector of "effective plastic load," the equation may be rewritten by grouping together the generalized applied nodal forces and effective plastic loads, resulting in the following equation,

$$\underset{\sim}{f}^i + \underset{\sim}{q}^i = \underset{\sim}{k}\,\underset{\sim}{u}^i \tag{2}$$

where

$$\underset{\sim}{q}^i = \underset{\sim}{k}_p\underset{\sim}{\gamma}_p^i \quad .$$

A solution for the displacement $\underset{\sim}{u}^i$ in Eq. (2) requires that the value of the effective plastic load be known at the current load level, or alternatively be expressed in terms of $\underset{\sim}{u}^i$. Since the values of the plastic shear strain are not known at the current load step, and since they cannot be conveniently expressed in terms of the generalized nodal displacements, we employ a predictor procedure solution technique in which the effective plastic loads are based on values of plastic strains determined in the preceding load step. Consequently, Eq. (2) is written in the following form:

$$\underset{\sim}{f}^i + \underset{\sim}{q}^{i-1} = \underset{\sim}{k}\,\underset{\sim}{u}^i \tag{3}$$

where

$$\underset{\sim}{q}^{i-1} = \underset{\sim p}{k}\underset{\sim}{\gamma}_p^{i-1}$$

and the superscript i-1 refers to the previous load level.

The use of this type of predictor procedure results in a "drifting" of the results from an exact solution as the loading increases. However, the inaccuracies can be minimized by using small load increments. The resulting accuracy and computing times required for this procedure were found to be competitive with those that utilize an iterative procedure and/or those in which plastic stress-strain relations are introduced explicitly into the governing matrix equations, thereby requiring a reformulation of the element stiffness matrices for each increment (i.e., tangent modulus method). These conclusions are based on experiences encompassing a wide variety of problems (Refs. 5 and 6), not exclusively associated with the particular class of problems treated here. A more detailed discussion of convergence and efficiencies associated with iterative methods of plastic analysis is presented in Ref. 7.

Equation (3) is written for each element in the structural idealization, and then, by an appropriate process of assemblage, the over-all linear matrix equation for the entire structure is formed. This resulting equation is identical in form to that of Eq. (3). Thus, with the present technique, one converts the nonlinear problem into a sequence of linear problems. Material nonlinearity is accounted for by introducing subsidiary incremental stress-strain relations from an appropriate plasticity theory and by the subsequent modification of applied loading through the effective plastic load. For the results presented here, the constitutive relations in the plastic range are developed on the basis of Drucker's postulate for work-hardening materials and the Prager-Ziegler kinematic hardening theory (Refs. 8 and 9), which accounts for the Bauschinger effect in the case of reversed plastic deformation. Details associated with the implementation of kinematic hardening in the finite-element analysis are presented in Refs. 5 and 6.

IV. ELASTIC RESULTS

Rectangular Panel

A flat rectangular panel is loaded axially along two opposite edges, which are assumed to remain straight and parallel but free to strain along their length. Since we are considering layups symmetrically placed about the midplane, no interlaminar shear stresses develop in the middle surface, and we may consider only that part of the panel on one side of the middle surface. As a result of symmetric geometry and loading conditions, a rotation of 180 degrees about the z-axis results in identical stress and deformation patterns. Therefore only one quarter of the panel, as shown by the shaded area in Fig. 3, need be considered for analysis. Furthermore, for layups in which a rotation of 180 degrees of the structure on one side of the midplane about the x-axis results in identical stress and deformation patterns, viz., a $[\beta_1/\beta_2/.../-\beta_2/-\beta_1]_s$ laminate, only one-eighth of the panel need be considered. The idealization used for such a case is shown in Fig. 4.

Results for a four ply $[\pm 45]_s$ angle ply laminate of boron-epoxy construction are shown in Fig. 5. The results compare favorably with the solution of Ref. 1. The interlaminar region in which the stress fields diverge from the classical plate solution comprises a narrow region along the free edge of the laminate and its "width of influence" is approximately 1.25 laminate thicknesses (5 lamina thicknesses). This interlaminar region, being one of high stress gradient, is sensitive to the ratio t/c, the ratio of a lamina thickness to a linear in-plane dimension of an element. For a structure in which the stresses along one edge, $y = b$, are not influenced by the free edge $y = -b$, the stress field is independent of the width. For problems involving widths greater than this limiting value, the same idealization can be used to determine the effect of various values of t/c by varying the thickness, t. Figure 6 shows some typical results in which good agreement with the exact solution is obtained for values of $t/c \geq 0.4$. Values of $t/c > 1/3$, or $t/\bar{c} > 1$ may be used as a guide to the construction of a finite-element idealization, where $\bar{c}$ is the distance from the centroid of the element to the free edge. Elements whose centroidal distance from an edge are greater than one lamina thickness (1/5 "width of influence") will not in general accurately predict the interlaminar effect.

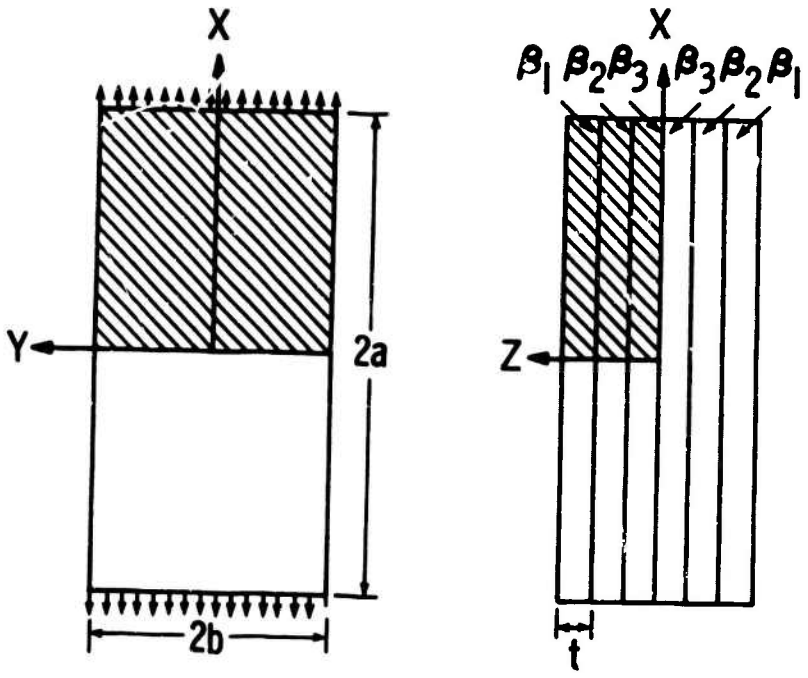


Figure 3. Laminate Geometry

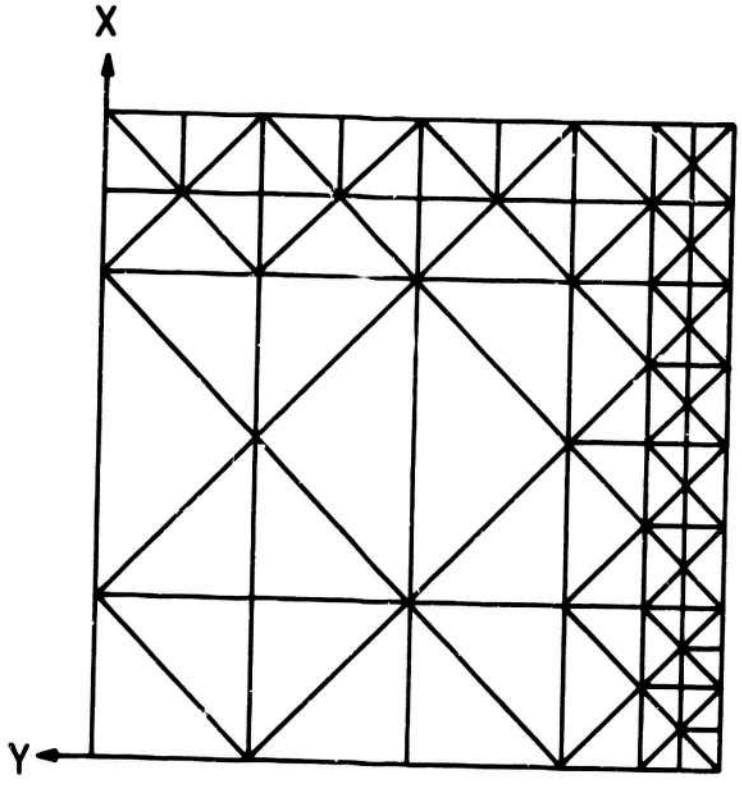


Figure. 4. Idealization of Quarter Panel

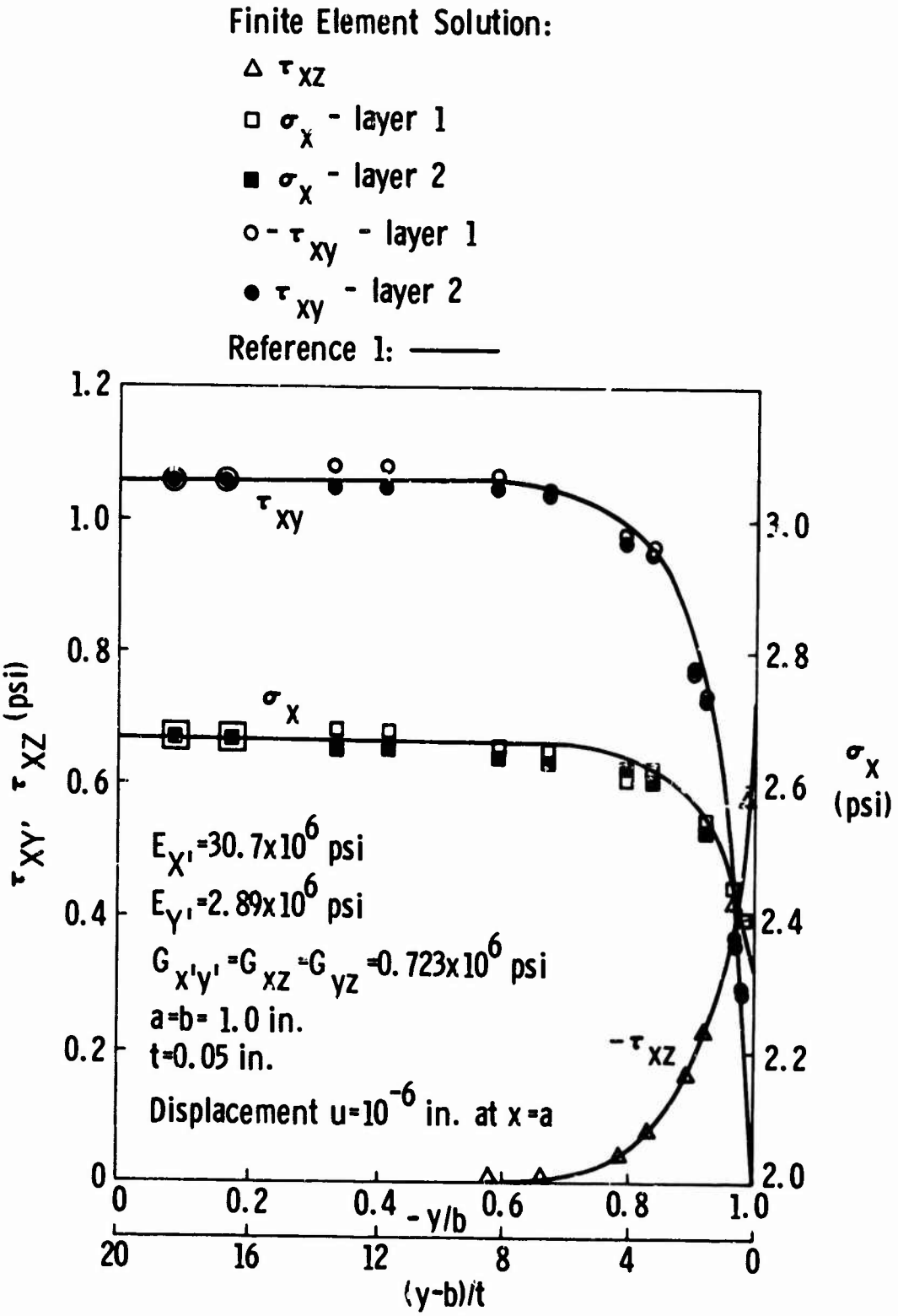


Figure 5. Stress Distributions, $[\pm45]_s$ Laminate

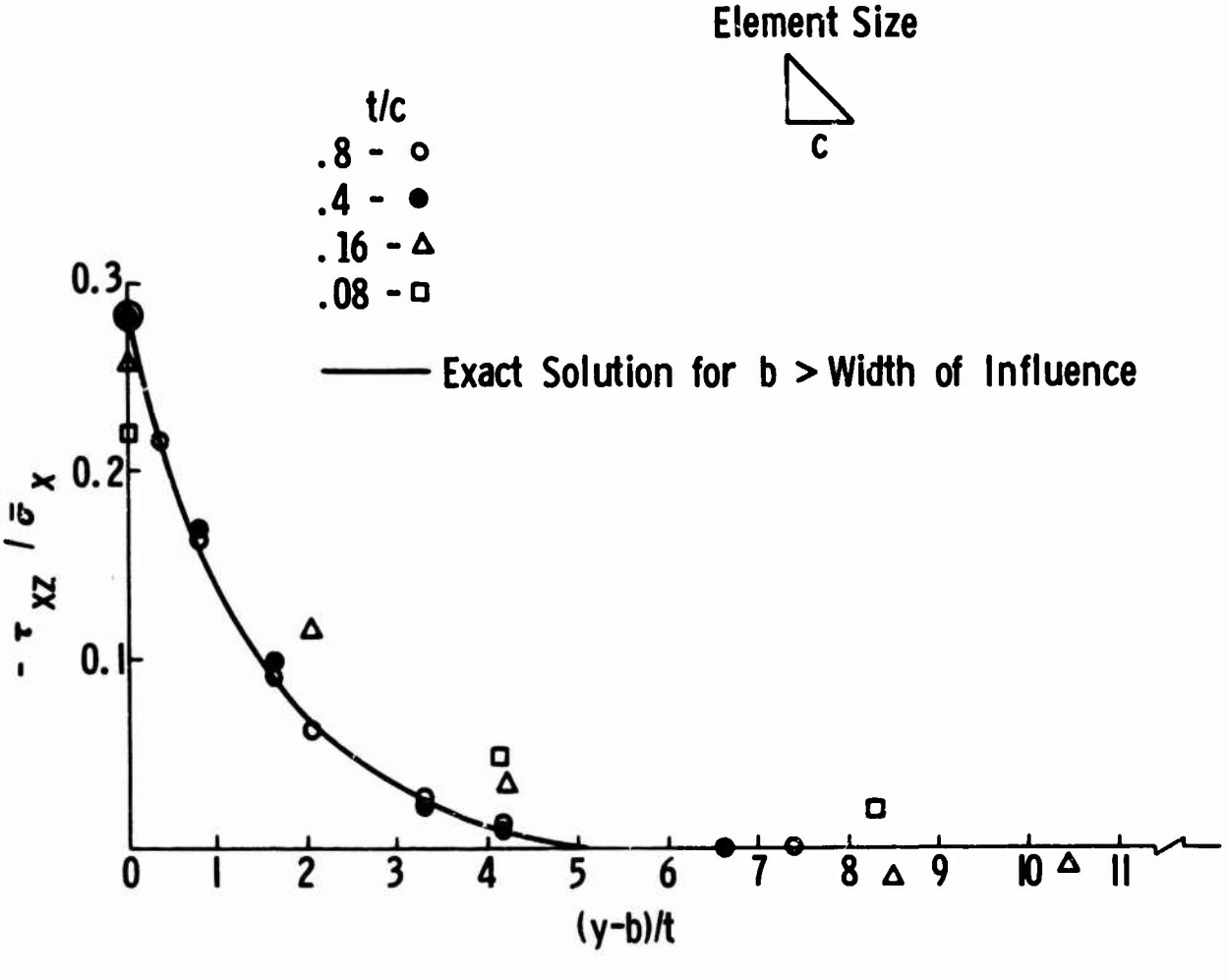
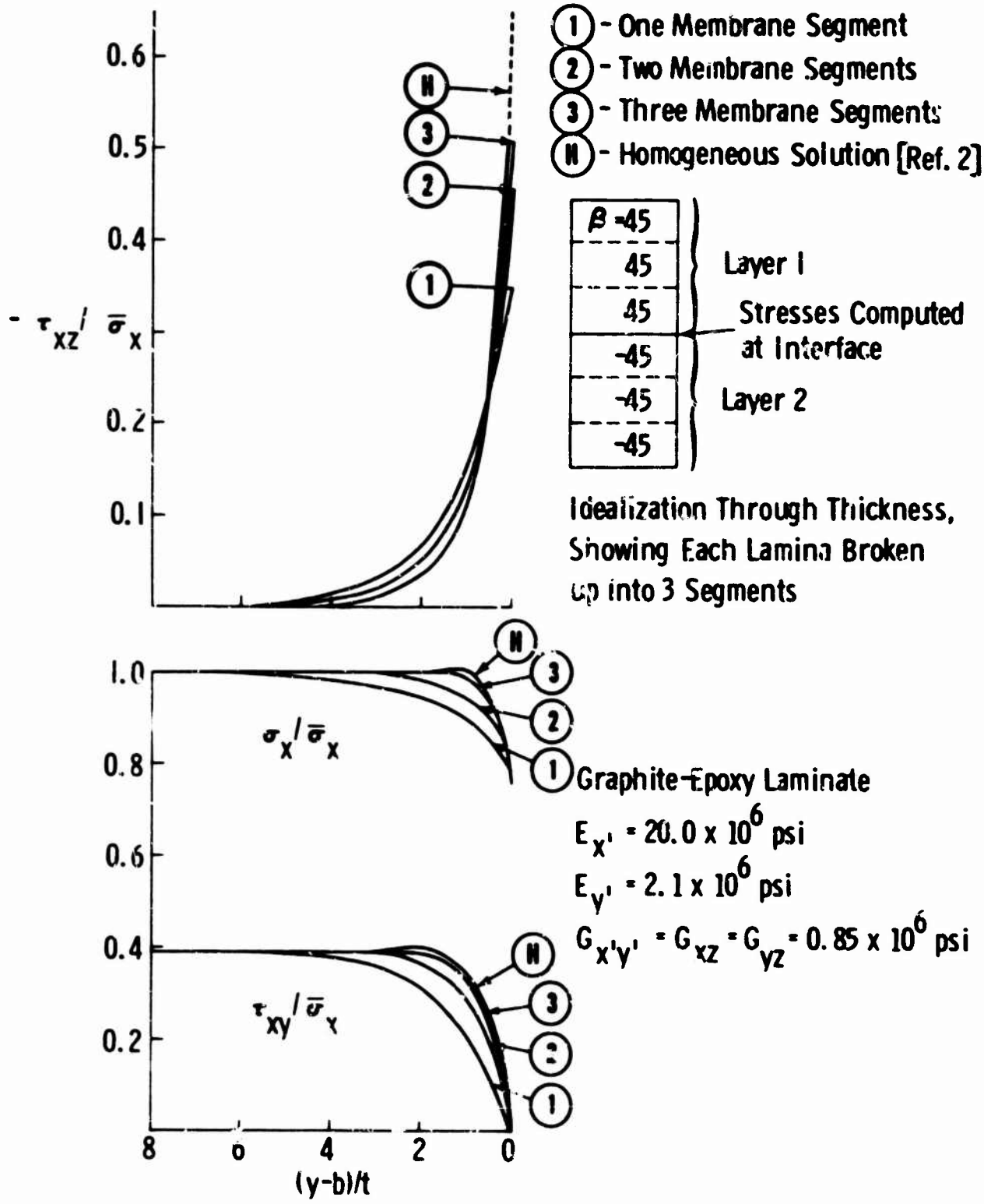


Figure 6. Computed Interlaminar Shear Stress Distribution for Various Ratios of t/c, $[\pm 45]_s$ Laminate.

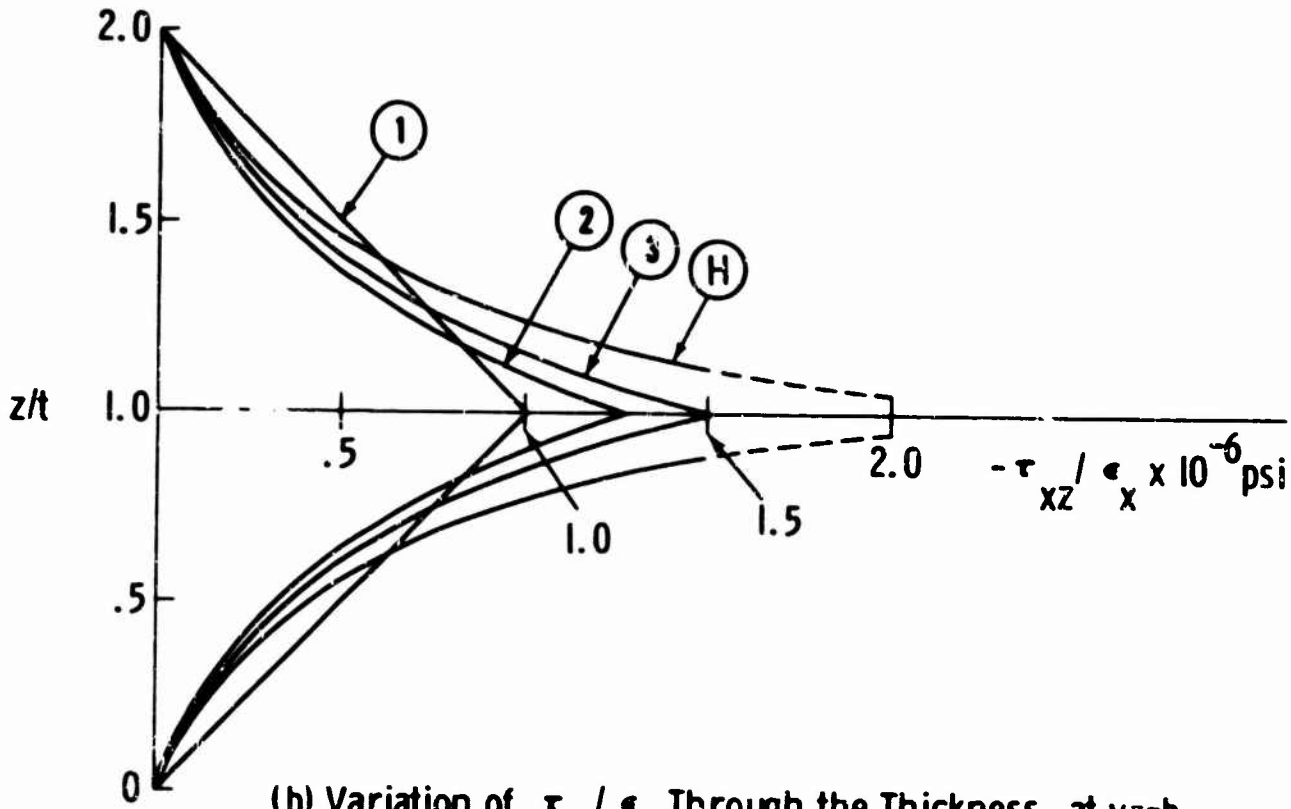
The influence of a particular construction of a finite-element idealization on the results for membrane stresses in a two dimensional problem is discussed in Ref. 10 where it is shown that deviations of stress from the true solution are caused by inaccuracies in the shear stiffness resulting from the use of certain types of finite-element patterns. The idealization shown in Fig. 4 has mixed patterns that should cause a fluctuation about the true solution rather than a consistent underprediction or overprediction of stresses. This is substantiated by an inspection of Fig. 5. From these results, however, the in-plane idealization does not appear to influence the interlaminar shear stress.

Different results are obtained when modeling a lamina as homogeneous, as in Ref. 2, as compared with the present model in which the fiber-bearing region and matrix region are considered distinct through the thickness. Using the present techniques, one can investigate an entire spectrum of models ranging from a discrete to a homogeneous one. This is done by dividing each lamina into a number of segments through the thickness, representing fiber-bearing and matrix regions alternately. As the number of segments increases, the individual laminae approach homogeneity. This is illustrated by the results shown in Fig. 7, where each layer is idealized into one, two, and three membrane segments, separated by shear segments, and compared with the results of Ref. 2 for a graphite-epoxy $[\pm 45]_s$ laminate. Figure 7a shows the variation of stresses across the width at the $\pm 45°$ laminae interface, while 7b shows the variation of interlaminar shear through the thickness. As the number of segments is increased, the present solution approaches that for a laminate composed of homogeneous laminae. Since the true structure of a composite laminate lies between a model composed of homogeneous laminae and a model composed of discrete segments, the results of the two extreme models can serve to set the bounds for the actual stress field. These results are also applicable for laminates in which identical laminae are stacked adjacent to each other. As an example, the results for a $[45/45/-45/-45]_s$ laminate, using the present model, are identical to the case in which each lamina is idealized into two membrane segments (see Fig. 7). The peak interlaminar shear increases approximately thirty percent over the $[\pm 45]_s$ laminate and has a width of influence of approximately 10 lamina thicknesses, while the peak interlaminar shear for a $[45/-45/45/-45]_s$ laminate is approximately the same as that of the four layer $[\pm 45]_s$ laminate and has a width of influence of approximately 5 lamina thicknesses. Similar results are obtained for other fiber orientations. These results indicate that the interlaminar shear can be kept to a minimum by alternating the $+\beta$



(a) Stress Distribution Across the Panel

Figure 7. Results for Various Lamina Models, $[\pm 45]_s$ Laminate.



(b) Variation of τ_{xz}/ϵ_x Through the Thickness, at y=-b

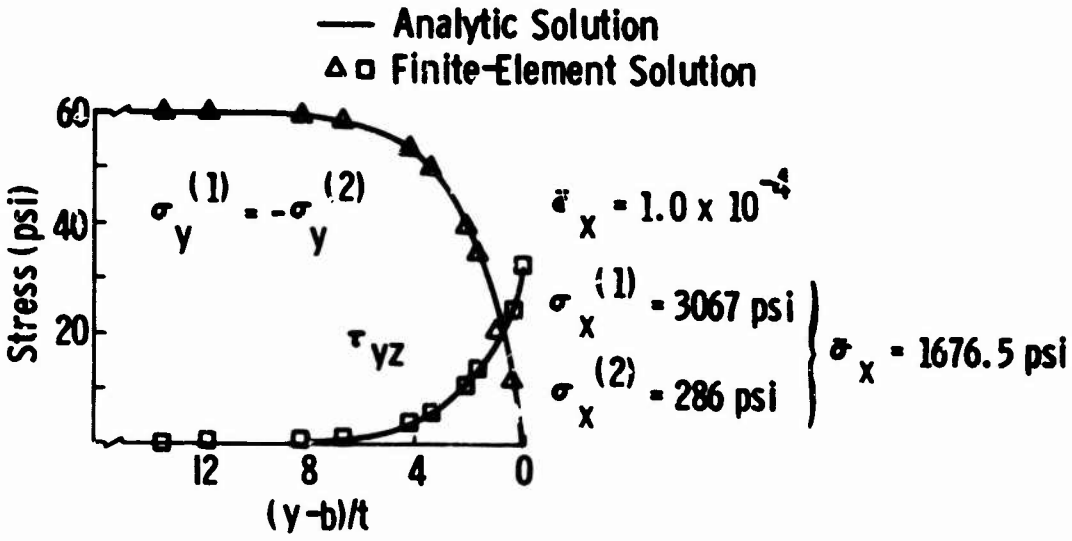
Figure 7. Results for Various Lamina Models, $[\pm 45]_s$ Laminate (Cont)

and $-\beta$ laminae rather than stacking them together. For a laminate with identical adjacent laminae, e.g., a $[45_n/-45_n]_s$ laminate, the peak interlaminar shear stress increases with increasing n, having a limiting value given by the solution of a $[45/-45]_s$ laminate in which each lamina is considered as homogeneous.

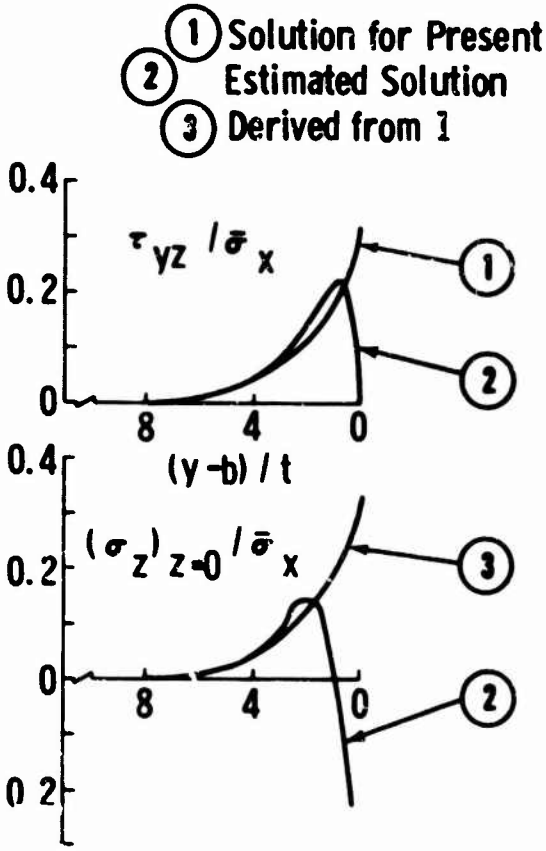
In general there will be both normal and tangential interlaminar shear stresses except in a $[\pm \beta]_s$ laminate where only a tangential shear stress appears and in a $[0/90]_s$ laminate where only a normal stress appears. In Appendix B a closed form solution is obtained for a $[0/90]_s$ laminate using the methods of Ref. 1. This solution and the finite-element solution are in good agreement as shown in Fig. 8a. The finite-element solution does not allow the shearing stress normal to the free edge, τ_{yz}, to approach zero at the free edge; however, there is reason to believe that this may be a highly localized inaccuracy. This is indicated, to some extent, in Ref. 11 in which a double lap joint is analyzed. A closed form solution in which a zero shear condition is enforced at the free edge is compared with a finite-element solution, and the results are in excellent agreement except at the free edge.

Peel Stress

The normal or "peel" stress, σ_z, is not taken into account in the present model. This may be a shortcoming from two points: 1) the importance of predicting σ_z as it relates to strength predictions of composites and 2) the effect of its inclusion on the distributions of the other stress components. Reference 12 discusses the influence of the stacking sequence on laminate strength and concludes that interlaminar normal stress is a significant factor in the delamination of composites. This is based on mathematical arguments and reference to experimental results. The experiments cited (Refs. 13 and 14) deal with the comparison of a $[\pm 45/\pm 15]_s$ with a $[\pm 15/\pm 45]_s$ boron-epoxy laminate, and a $[90/0]_s$ with a $[0/90]_s$ glass-epoxy laminate, respectively. In each case the former laminate results in a higher strength, resulting from compressive normal stress near the free edge (Ref. 12). Other investigations (Refs. 15 and 16), however, have concluded that where thermal and/or mechanical stress causes matrix cracking between fibers in the laminate plane, matrix enhancement of fiber strength (above the dry bundle strength) through transfer of load from broken to unbroken fibers depends upon load paths through the thickness. For example, both a $[0/90]_s$ and a $[90/0]_s$ laminate experience transverse residual thermal strains



(a) Solution for Present Model



(b) Estimated Stresses with Inclusion of σ_z

Figure 8. Stress Distributions, $[0/90]_s$ Laminate

sufficient to initiate such cracks after cooling from the curing temperature. Hence higher strengths would be expected with the $[90/0]_s$ laminate because load transfer from broken to unbroken fibers is possible between the inner layers. This effect is more pronounced in a boron-epoxy laminate then in a glass-epoxy laminate (Ref. 16).

The present methods, while not including σ_z in the analysis, can give some insight into its magnitude and distribution. Assuming for the time being that the introduction of σ_z does not appreciably change the remainder of the stress field, then an expression for σ_z can be obtained using the equilibrium equation

$$\sigma_{z,z} = - \tau_{yz,y} \quad .$$

Using the results in Appendix B for a $[0/90]_s$ laminate, we find

$$(\sigma_z)_{z=0} = - 2 \frac{u^*}{a} Q_{12} G \frac{(Q_{22} - Q_{11})}{Q_{11} Q_{22}} \frac{\text{ch } \sqrt{\rho}\, y}{\text{ch } \sqrt{\rho}\, b} \quad ,$$

which is shown in Fig. 9b. Along with these results, qualitative curves representing the actual response for τ_{yz} and σ_z are shown. These are based on equilibrium considerations as demonstrated in Ref. 12. The inclusion of σ_z results in a reduction of peak interlaminar shear stress as shown in Ref. 11, which discusses this very issue with respect to bonded joints in composite materials. For comparison a $[0/90]_s$ laminate is analyzed, with the inclusion of σ_z, in Ref. 17.

A qualitative determination of σ_z and its effect on the other stress components is predictable, and it appears that the need for an exact determination of σ_z, as it relates to strength predictions, is open to question.

Panel With a Cutout

A flat panel with a circular cutout is loaded along two opposite edges. The dimensions of the panel are such that the stress field around the cutout is not influenced by the external boundaries. A typical idealization is shown in Fig. 9, which contains 223 members and 138 nodes. Results for a $[\pm 45]_s$ laminate, where the ratio of hole radius to lamina thickness is 100, are shown in

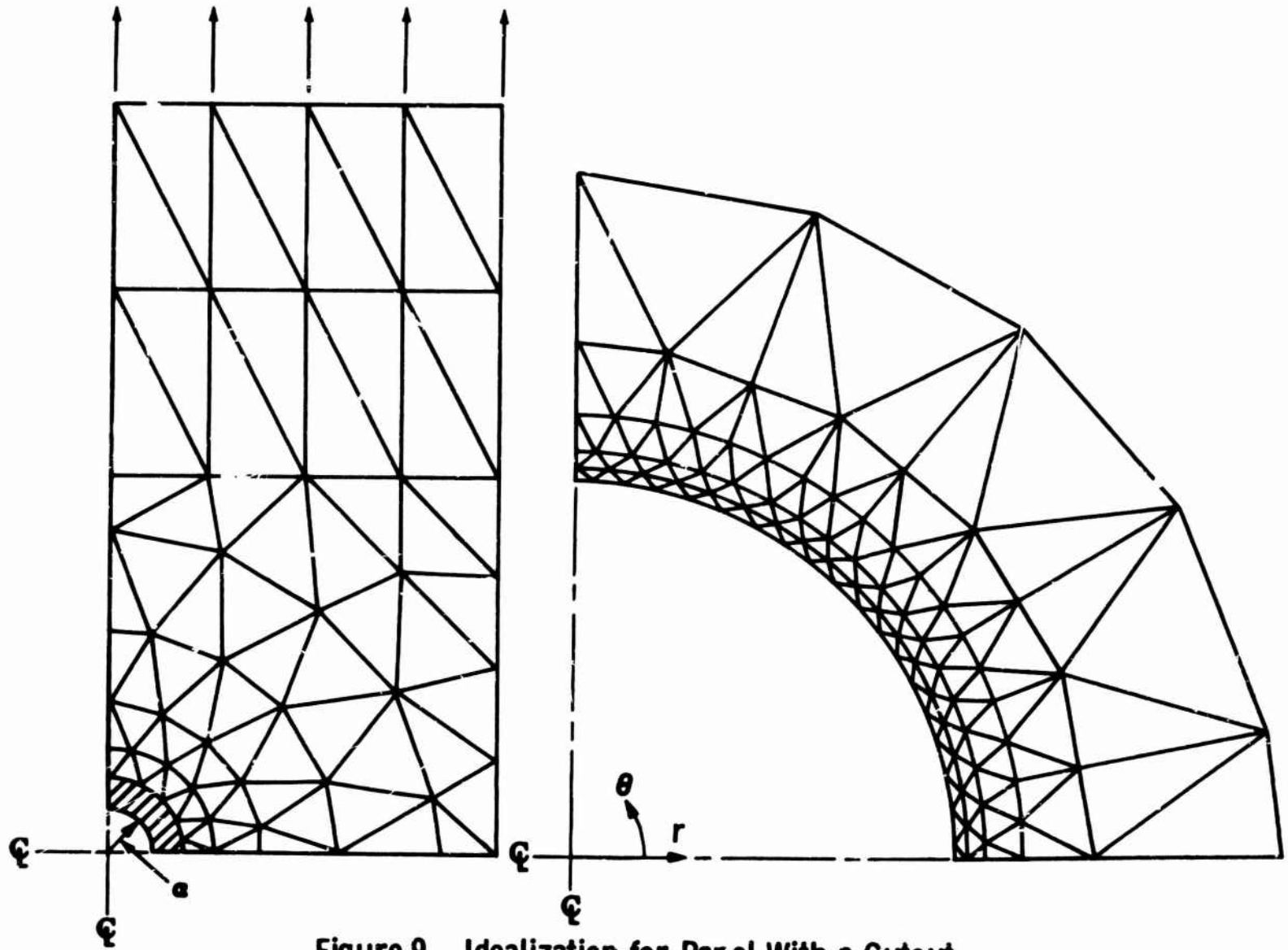
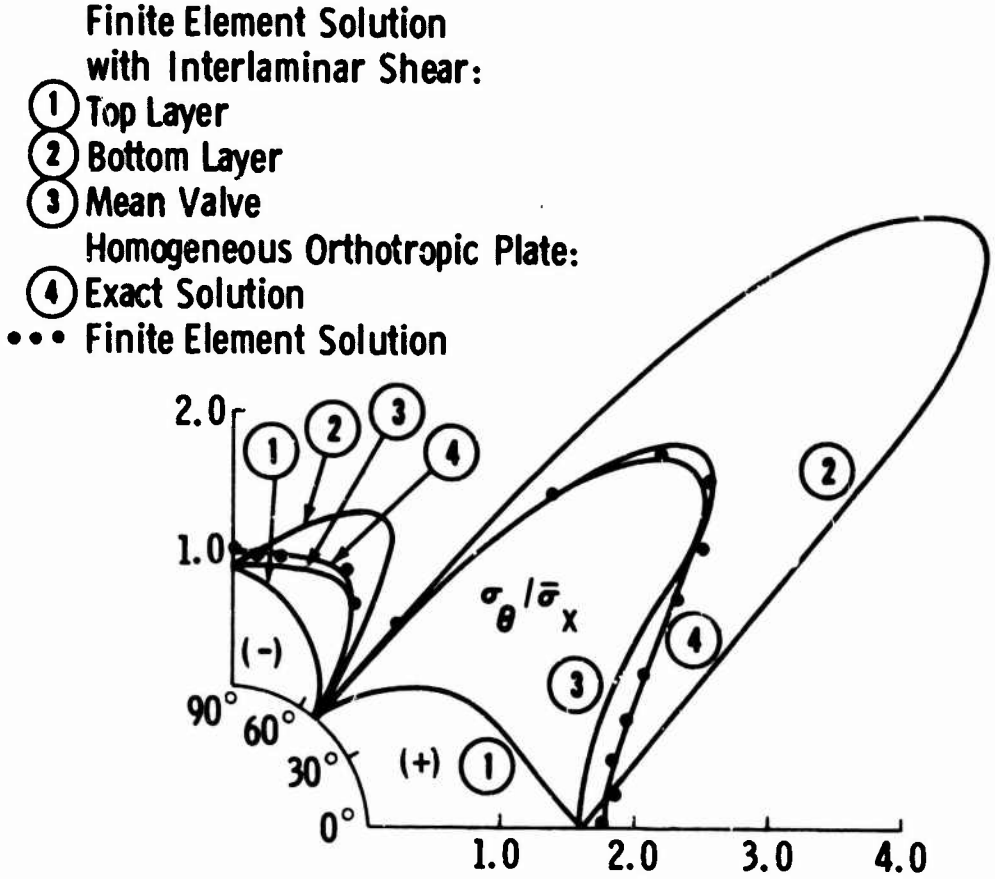


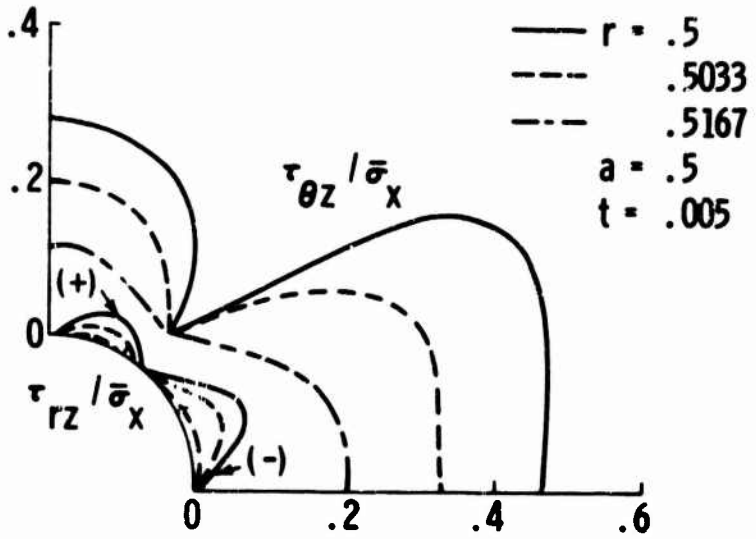
Figure 9. Idealization for Panel With a Cutout.

Fig. 10. To test the finite-element idealization, a comparison is made between the finite-element solution and exact solution of the mean circumferential stress at a distance slightly away from the cutout (Fig. 10a). These results are in excellent agreement. The inclusion of interlaminar shear causes the mean circumferential stress to deviate from the classical case, this deviation being most pronounced in the regions of high circumferential interlaminar shear, $\tau_{\theta z}$. The radial interlaminar shear stress is small compared with the circumferential interlaminar shear stress. This is true for all values of β. Along the radial coordinates $\theta = 0°$ and $\theta = 90°$, where the fiber orientations are either $\pm 45°$ or $\mp 45°$ when measured from these axes, the stress field behaves in a manner similar to that along a free edge of a $[\pm 45]_s$ rectangular panel. The straight edge solution of a $[\pm 45]_s$ laminate yields $\sigma_x/\sigma_{xa} = 0.86$ and $\tau_{xz}/\sigma_{xa} = 0.28$, where σ_{xa} is the classical normal stress at the point. The corresponding values at the edge of the cutout along $\theta = 0°$ are 0.89 and 0.26 and along $\theta = 90°$ are 0.88 and 0.27, respectively. The reason for the deviation between these results and those along the straight edge is that the interlaminar shear deformation edge effect emanates from points along a boundary and extends over a distance equal to the width of influence. Therefore a "circle of influence" exists around each point in which the stress distribution influences the stress behavior at that point. Along a straight boundary of infinite length, the stresses within the circle of influence are uniform in a direction tangent to the boundary. Along the edge of a cutout, however, the stress field is not uniform in the circumferential direction. As the radius of the hole increases, the stress field within the circle of influence approaches a uniform state, and the resulting stress field at a point approaches that of a straight edge solution. Figure 11 shows the resulting interlaminar shear for a $[\pm 45]_s$ laminate as the ratio of hole radius to laminate thickness is varied. The straight edge solution is approached at the axes within 5 percent when the ratio of hole radius to laminate thickness equals 100, and within 20 percent when the ratio equals 10. As the ratio of hole radius to lamina thickness decreases, the interlaminar shear stress flattens out as a function of θ and decreases in average value.

Figure 12 shows results for a $[\pm 30]_s$ laminate. The values of $\sigma_\theta/\sigma_{\theta a}$ and $\tau_{\theta z}/\sigma_{\theta a}$ ($\sigma_{\theta a}$ is the classical mean circumferential stress at the point) along the cutout at $\theta = 0°$ and $\theta = 90°$ for a ratio of $a/t = 100$ is similar to the corresponding values of stress along the straight edge of a $[\pm 30]_s$ and $[\pm 60]_s$ laminate, respectively. Note the significant reduction of stress



(a) Circumferential Stress at r/a = 1.007



(b) Interlaminar Shear Stresses

Figure 10. Stresses Around Cutout, $[\pm 45]_s$ Laminate

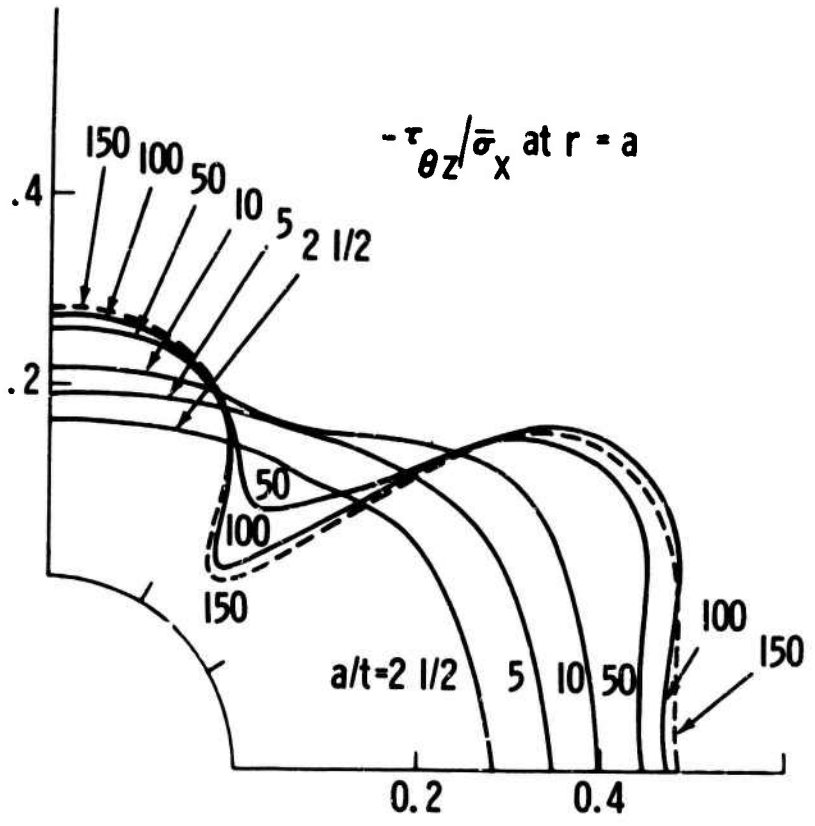
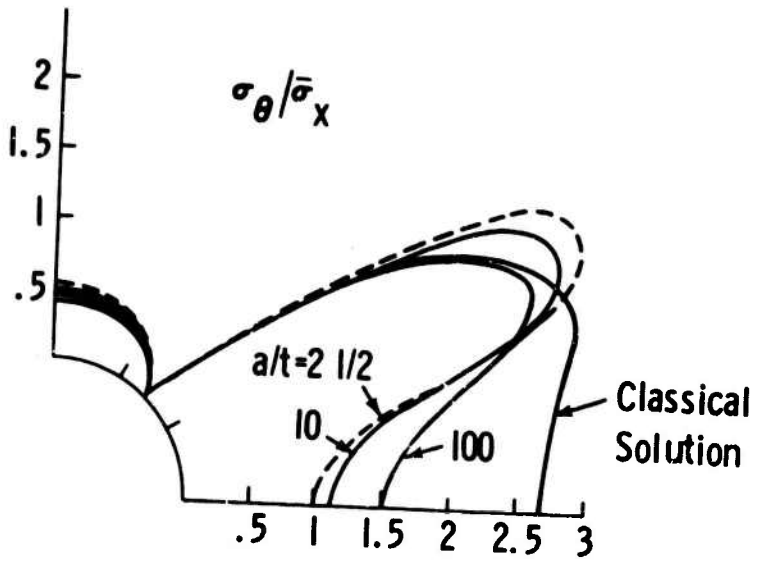
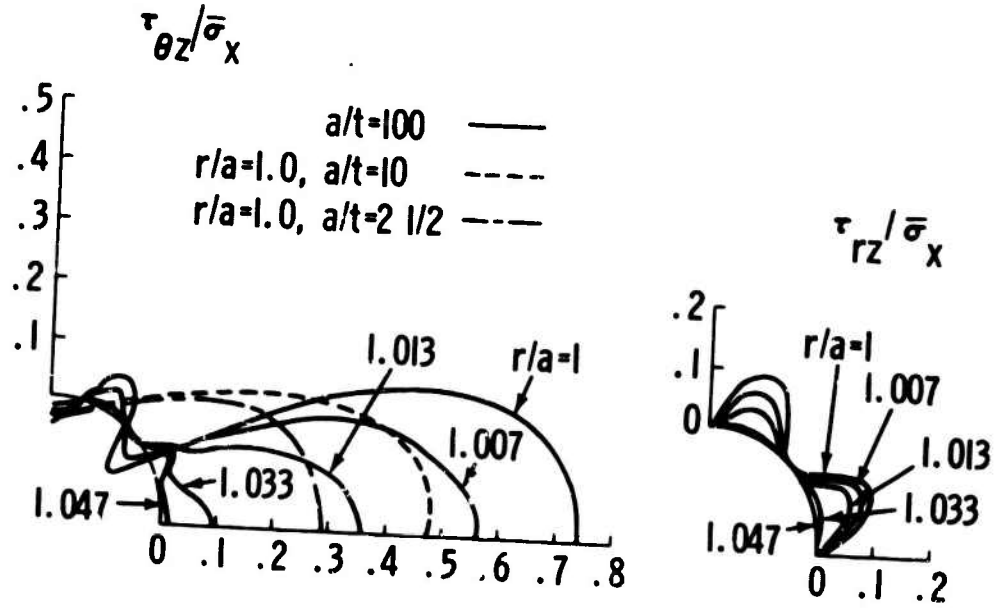


Figure 11. Interlaminar Stress Distributions for Various Radius to Thickness Ratios, $[\pm 45]_S$ Laminate.



(a) Mean Circumferential Stress at r/a=1.007



(b) Interlaminar Shear Stresses

Figure 12. Stress Distributions Around Cutout, $[\pm 30]_s$ Laminate.

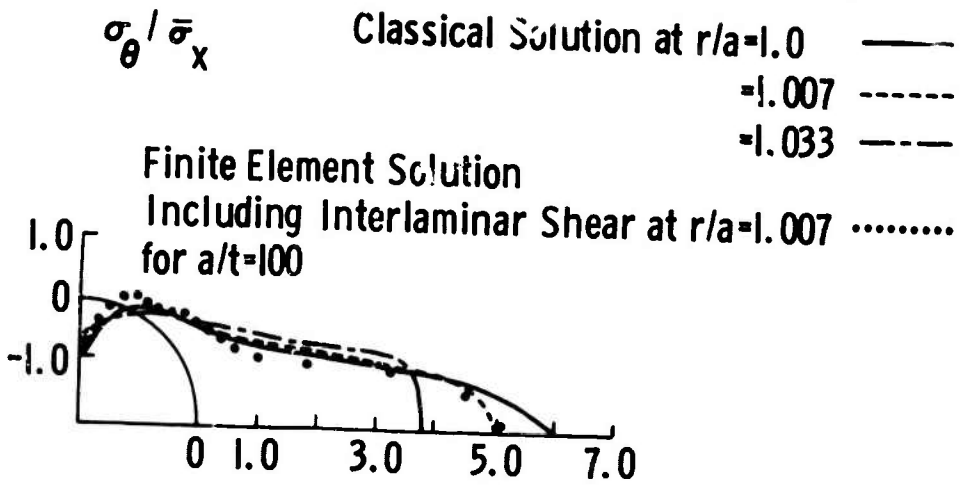
concentration in σ_θ at $\theta = 0°$ and the redistribution of $\tau_{\theta z}$ and σ_θ for the cases in which the radius to lamina thickness is reduced. The reduction in stress concentration in σ_θ at $\theta = 0°$ is predictable from the straight edge solution of a $[\pm 30]_s$ laminate.

Results for a $[0/90]_s$ laminate containing a circular cutout are shown in Fig. 13. There is little variation of σ_θ from the classical solution along the $\theta = 0°$ and $\theta = 90°$ axes. This is similar to the straight edge solution for a $[0/90]_s$ laminate where there is little diffusion of the axial stress due to interlaminar shear. The circumferential interlaminar shear is zero along $\theta = 0°$ and $\theta = 90°$, as expected, but grows rapidly away from these points, and is greater in magnitude than the maximum radial interlaminar shear stress. Unlike the behavior in a $[\pm \beta]_s$ laminate, the circumferential stress varies rapidly in the circumferential direction near $\theta = 90°$ and in both the circumferential and radial directions near $\theta = 0°$. This accounts for the difference in radial interlaminar shear stress between the results along $\theta = 0°$ and $\theta = 90°$ and those along a straight edge of a $[0/90]_s$ laminate. Of interest is the radial interlaminar shear growth along $\theta = 90°$ and change of direction along $\theta = 0°$ when the ratio of hole size to lamina thickness decreases. This is in contrast to the steady decrease of circumferential interlaminar shear around the cutout as a/t decreases.

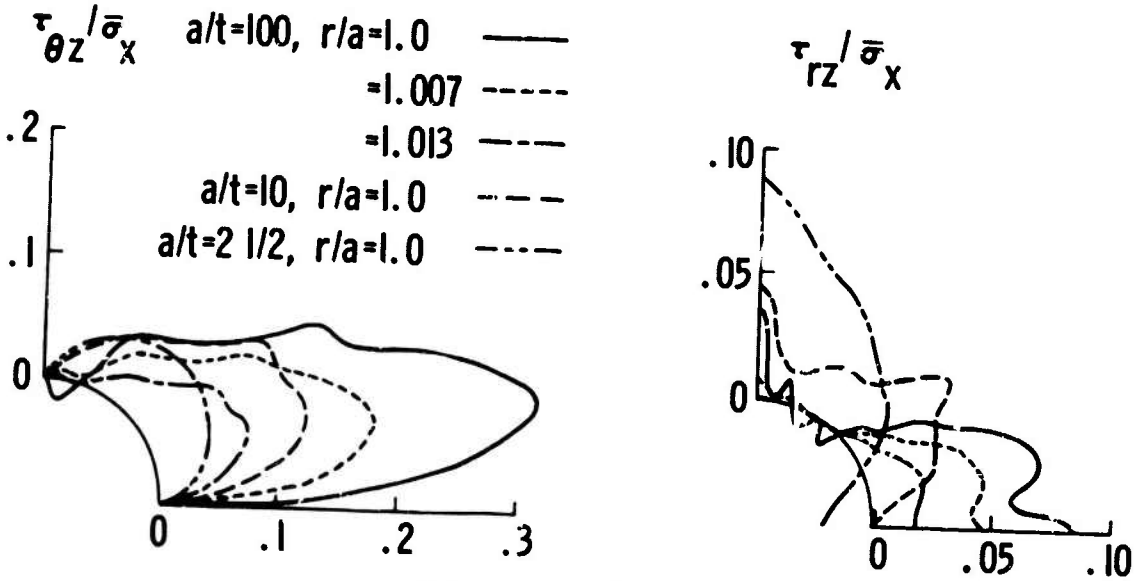
V. INELASTIC RESULTS

The method of initial strains to describe the inelastic behavior is briefly described in Section III. In this first study of the problem, it is assumed that the interlaminar matrix material can deform plastically while the in-plane material behavior is assumed to be elastic throughout the entire load history.

The interlaminar stress strain behavior of boron-epoxy is shown in Fig. 14, along with its Ramberg-Osgood parameters. Stresses are computed for a $[\pm 45]_s$ boron-epoxy laminate at a load level 4.22 times that at which yielding begins ($\tau_y = 3460$ psi). Compared to the elastic response at the same load level, the interlaminar shear stress is lower near the edge and is slightly higher at points more remote from the edge, but it does not extend over a significantly greater region. Moreover, the effect of the nonlinearity on the membrane stresses is slight (Fig. 15).



(a) Mean Circumferential Stress



(b) Interlaminar Shear Stresses

Figure 13. Stress Distribution Around Cutout, $[0/90]_s$ Laminate.

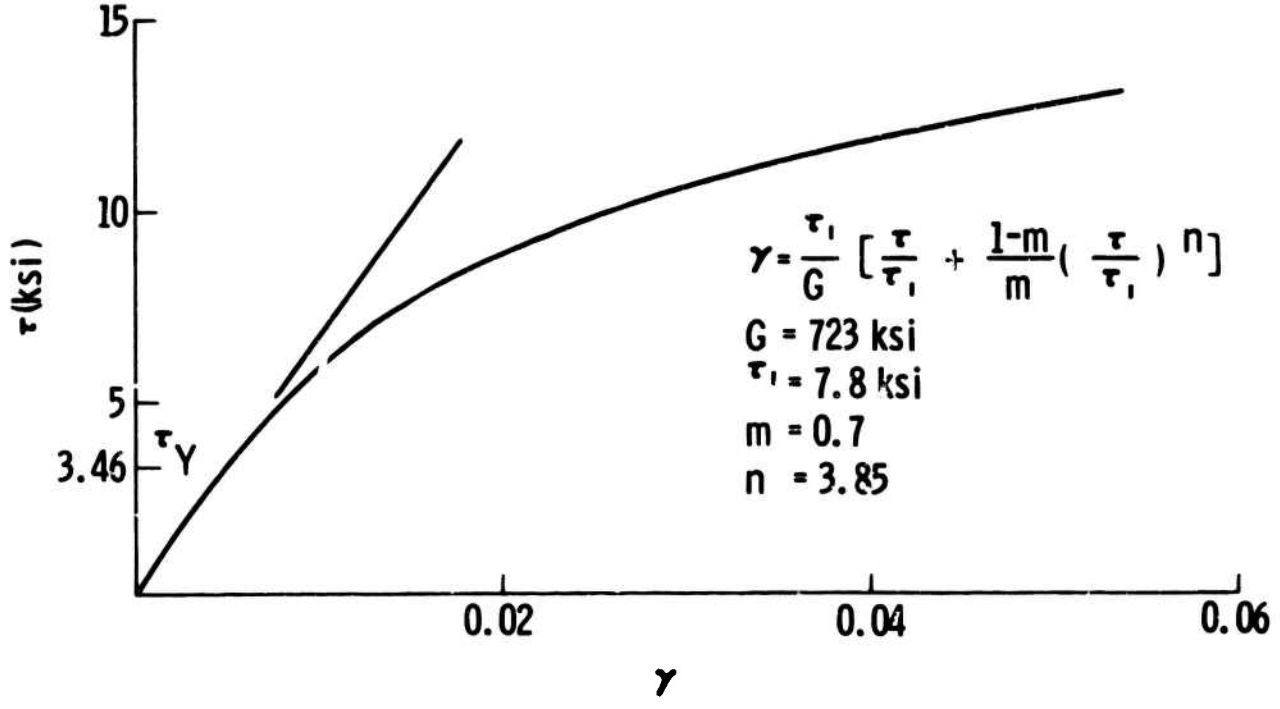


Figure 14. Interlaminar Shear Stress-Strain Curve and Ramberg-Osgood Parameters for Boron-Epoxy Laminate.

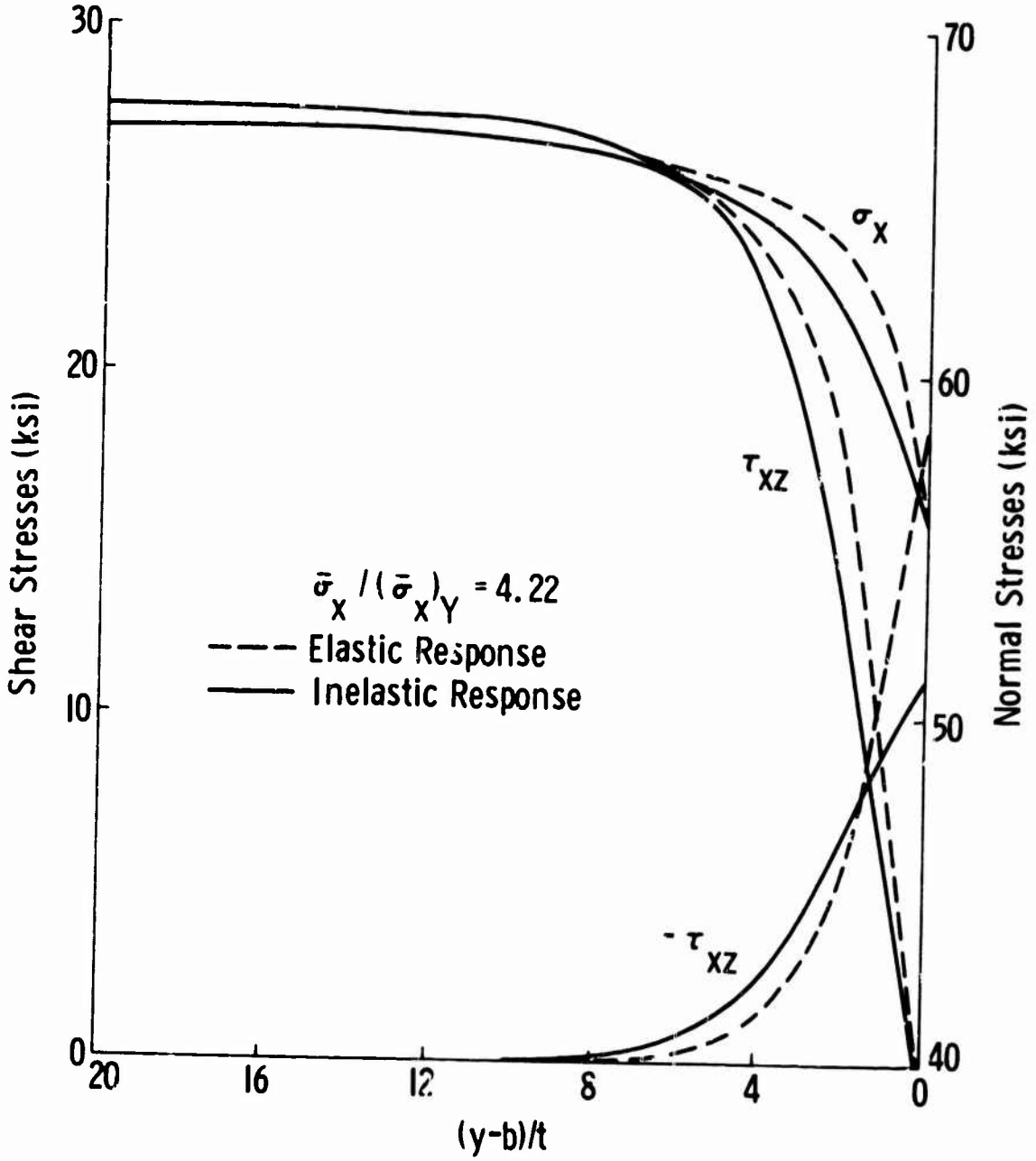


Figure 15. Comparison of Elastic and Inelastic Response, $[\pm 45]_s$ Laminate.

A multilayered $[0/\pm45/0/\bar{90}]_s$ boron-epoxy laminate is analyzed. The "loaded edge" displacements are increased until a tensile load of 112 ksi is reached (slightly greater than the expected fracture stress). The load is then decreased to a compressive load of 112 ksi and finally brought back to zero. In general, similar results are obtained during tensile loading of a $[\pm 45]_s$ laminate, i.e., a reduction in peak interlaminar shear stresses and little change in the membrane stresses are observed. The interlaminar shear components (in the layers in which they have the greatest peak magnitude) at a distance of 0.1 laminate thickness from the $y = -b$ edge are plotted in Fig. 16 for the complete loading cycle. Of interest is the buildup of residual shearing stresses that occurs for both components of interlaminar shear. The mean stress, $\bar{\sigma}_x$, varies linearly with respect to the applied displacement to within five percent, an indication that the interlaminar shear nonlinearity has little effect on the gross laminate behavior. Interlaminar shear stresses around a circular cutout in a $[\pm 45]_s$ laminate are shown in Fig. 17. High residual shearing stresses occur in the vicinity of $\theta < 45°$, and relatively low residual shearing stresses occur for $\theta > 45°$.

VI. DISCUSSION

A finite element program was developed for the study of interlaminar shear deformation in composite laminates. The main building block of this program, called COMPEL, is the COMPosite ELement, which is described in Appendix A. The displacement method is used and the solution is effected by the use of a Cholesky decomposition scheme. Typical running times are 29 sec (elastic) and 1.2 sec per increment (plastic) for an idealization containing 210 degrees of freedom with a mean bandwidth of 132.

The following observations pertain to the techniques used in analyzing a composite laminate with the use of the present composite element:

- To determine accurately the interlaminar shear stresses, values of $t/\bar{c} > 1$ must be observed in the construction of a finite-element idealization.

- Upper and lower bounds for interlaminar shear stresses can be found by varying the number of idealized segments through the thickness.

- A qualitative determination of the "peel stress," σ_z, and its effect on the other stress components is predictable.

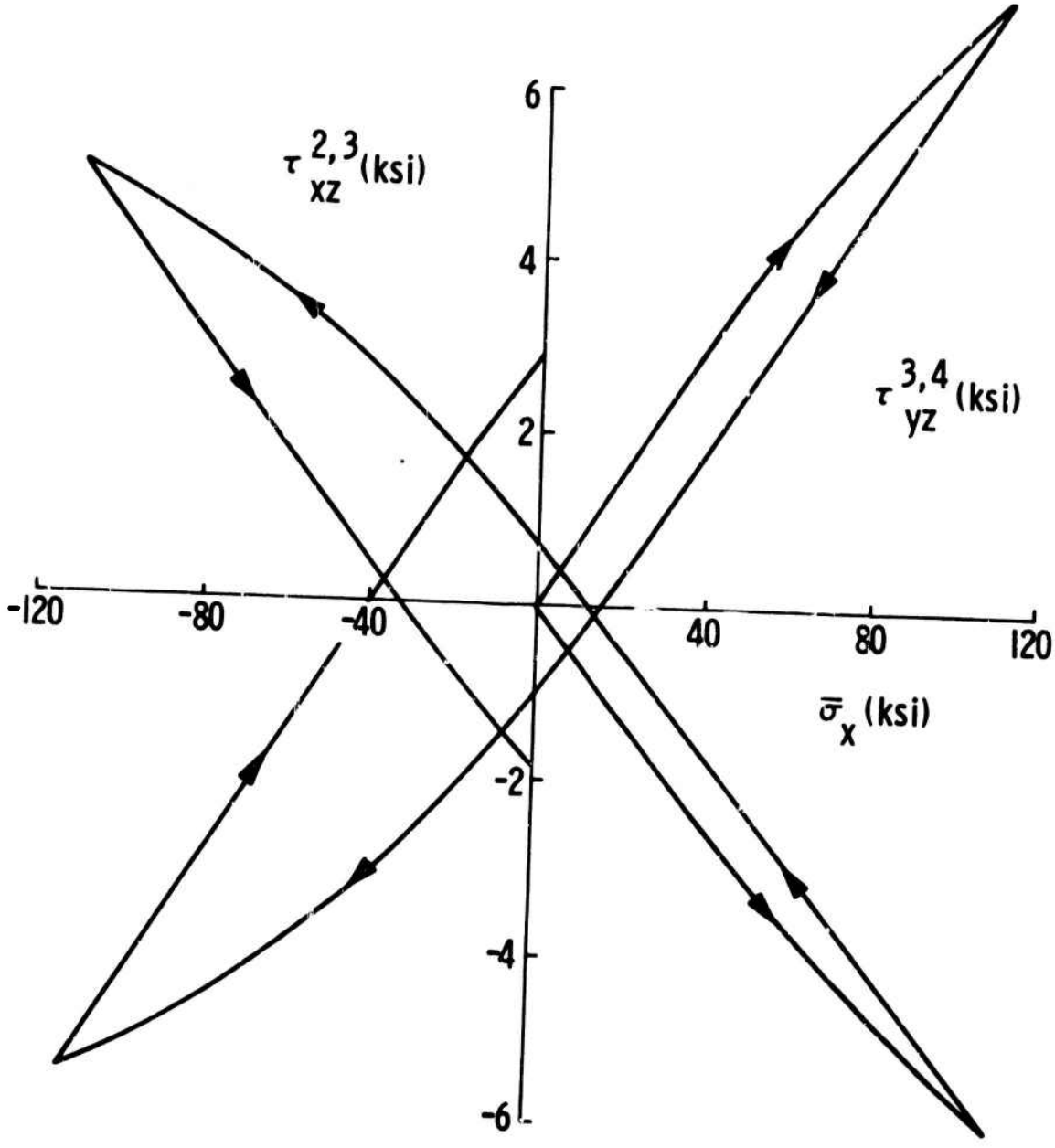


Figure 16. Interlaminar Shear Stresses at (y-b)/t=0.4 for Cyclic Loading, $[0/\pm45/0/\bar{90}]_s$ Laminate.

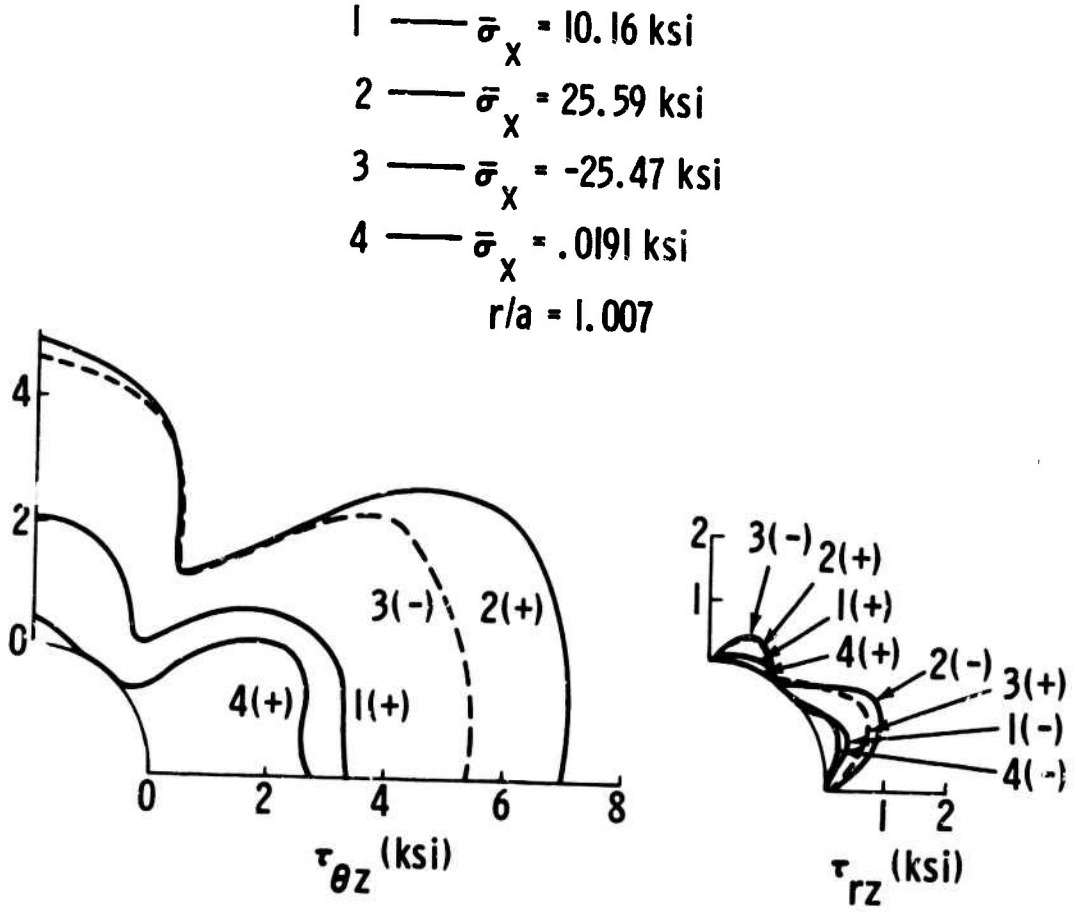


Figure 17. Interlaminar Shear Stresses Around Cutout for Cyclic Loading, $[\pm 45]_s$ Laminate.

The following results pertain to the elastic response of a composite laminate:

- The stacking sequence is the dominant factor in controlling the interlaminar shear region; e.g., alternating $+\beta$ and $-\beta$ laminae keeps the peak interlaminar shear stress and width of influence to a minimum.

- For a laminate with identical adjacent laminae, e.g., a $[45_n/-45_n]_s$ laminate, the peak interlaminar shear stress increases with increasing n, having a limiting value corresponding to a four-ply $[45/-45]_s$ laminate in which each lamina is considered as homogeneous.

- Around a circular cutout the circumferential interlaminar shear stress component tends to be of greater magnitude than the radial component for all stacking sequences, the exception being in the region between $0°$ and $90°$ laminae, where a higher radial component occurs for small values of a/t.

- For a circular cutout, as the ratio of hole radius to lamina thickness decreases, the interlaminar shear stresses flatten out, as a function of θ, and decrease in average value, except in the case noted above.

The results pertaining to the inelastic interlaminar shear behavior show:

- High residual interlaminar shear stresses develop during cyclic loading.

- Interlaminar shear nonlinearity has little effect on the gross laminate behavior.

This study accounts only for the plastic deformation of the interlaminar matrix material. Accounting for the most general state of plastic deformation requires that the analysis provide the capability to include a criterion to predict the occurrence of failure for a particular stress state in an orthotropic lamina under combined stress and a technique that can adequately represent the subsequent phenomenological behavior upon further loading. The first requirement, i.e., defining an acceptable fracture or yield criterion, has been the subject of a large number of studies (see Ref. 18 for an extensive bibliography). These approaches usually

provide an analytical expression, referred to as "interaction formulas," representing a quadratic surface in stress space, which best fits some experimental data for a given material. Although the reliability of the various criteria available is not established and in some cases is open to question, one can choose the "best" criterion for a particular material and readily incorporate such into an analysis until a "better" one becomes available. Methods to treat the behavior of the material beyond initial failure or yielding have been studied and appear in Refs. 19-21. In Ref. 19 the elastic stiffness properties of the structure are modified to reflect the development of nonlinear material behavior. In Refs. 20 and 21 prescribed conditions are imposed on the subsequent behavior of the stress components, i.e., set to zero (Ref. 20) or required to follow a unidirectional nonlinear relation (Ref. 21). As indicated in Ref. 22, these approaches to predict laminate strength represent a good beginning but require further work. Efforts toward the goal of providing a meaningful description of a general inelastic strength analysis represent a logical extension of the work here presented.

VII. REFERENCES

1. Puppo, A. H. and Evensen, H. A., "Interlaminar Shear in Laminated Composites under Generalized Plane Stress," J. Comp. Materials, Vol. 4, p. 204, 1970.

2. Pipes, R. B. and Pagano, N. J., "Interlaminar Stresses in Composite Laminates under Uniform Axial Extension," J. Comp. Materials, Vol. 4, p. 583, 1970.

3. Isakson, G. and Levy, A., "Finite-Element Analysis of Interlaminar Shear in Fibrous Composites," J. Comp. Materials, Vol. 5, p. 273, 1971.

4. Zienkiewicz, O. C., The Finite Element Method in Structural and Continuum Mechanics, McGraw-Hill, p. 32, 1967.

5. Armen, H., Pifko, A., and Levine, H. S., Finite Element Analysis of Structures in the Plastic Range, NASA Contractor's Report CR-1649, February 1971.

6. Isakson, G., Armen, H., and Pifko, A., Discrete-Element Methods for the Plastic Analysis of Structures, NASA Contractor's Report CR-803, October 1967.

7. Havner, K. S., "On Convergence of Iterative Methods in Plastic Strain Analysis," J. Solid Structures, Vol. 4, p. 491, 1968.

8. Prager, W., "A New Method of Analyzing Stress and Strains in Work-Hardening Plastic Solids," J. Appl. Mech., Vol. 23, p. 493, 1956.

9. Ziegler, H., "A Modification of Prager's Hardening Rule," Quart. Appl. Math., Vol. 17, No. 1, p. 55, 1959.

10. Walz, J. E., Fulton, R. E., and Cyrus, N. J., "Accuracy and Convergence of Finite Element Approximations," Proceedings of the Second Conference on Matrix Methods in Structural Mechanics, Wright-Patterson Air Force Base, Ohio, 15-17 October 1968.

11. Fehrle, A. C. et al., Lockheed-Georgia Company, Marietta, Georgia, Development of an Understanding of the Fatigue Phenomena of Bonded and Bolted Joints in Advanced Filamentary Composite Materials, Technical Report AFFDL-TR-71-44, Air Force Flight Dynamics Laboratory, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, June 1971.

12. Pagano, N. J. and Pipes, R. B., "The Influence of Stacking on Laminate Strength," J. Comp. Materials, Vol. 5, p. 50, January 1971.

13. Foye, R. L. and Baker, D. J., "Design of Orthotropic Laminates," presented at the 11th Annual AIAA Structures, Structural Dynamics, and Materials Conference, Denver, Colorado, April 1970.

14. Kaminski, B. E., "On the Determination of the Failure Surface for an Orthotropic Quasi-Homogeneous Material," Master's Thesis, Georgia Institute of Technology, June 1969.

15. Tsai, S. W. et al., Effect of Constituent Material Properties on the Strength of Fiber-Reinforced Composite Materials, Technical Report AFML-TR-66-190, August 1966.

16. Hadcock, R. N., "Strength of Multidirectional Boron-Epoxy Laminates," Proceedings of SPI 24th Annual Technical Conference, Soc. of the Plastics Industry, Washington, D.C., February 1969.

17. Rybicki, E. F., "Approximate Three-Dimensional Solution for Symmetric Laminates Under Inplane Loading," J. Comp. Materials, Vol. 5, p. 354, July 1971.

18. Puppo, A. H. and Evensen, H. A., "Strength of Anisotropic Materials under Combined Stresses," presented at AIAA/ASME 12th Structures, Structural Dynamics and Materials Conference, Anaheim, California, April, 1971, paper 71-368.

19. Tsai, S. W., Strength Characteristics of Composite Materials, NASA CR-224, 1965.

20. Petit, P. H. and Waddoups, M. E., "A Method of Predicting the Non-Linear Behavior of Laminated Composites," J. Comp. Materials, Vol. 3, p. 2, 1969.

21. Rosen, B. W. and Dow, N. F., "Mechanics of Failure of Fibrous Composites," in Fracture, Vol. VII, H. Liebowitz, ed., Academic Press, 1970.

22. Rosen, B. W., "Structural Composites - Design and Analysis," Second Annual Structures Design Lecture presented at AIAA/ASME 12th Structures, Structural Dynamics and Materials Conference, Anaheim, California, April, 1971.

ACKNOWLEDGMENT

The authors gratefully acknowledge the work of Joseph S. Miller in programming the analysis.

APPENDIX A

FORMULATION OF STIFFNESS MATRICES

The elastic strain energy is expressed as

$$U = \frac{1}{2}\int_V \underset{\sim}{\sigma}'\underset{\sim}{\epsilon}dV \quad . \tag{A-1}$$

The elastic strains are written in terms of the total and plastic strains as

$$\underset{\sim}{\epsilon} = \underset{\sim}{\epsilon}_t - \underset{\sim}{\epsilon}_p \quad , \tag{A-2}$$

where, for present purposes, all components of total and plastic strains are assumed constant within an element. The constant strains are represented in the following form:

$$\begin{aligned} \underset{\sim}{\epsilon}_p &= \underset{\sim}{\epsilon}_o \\ \underset{\sim}{\epsilon}_t &= \underset{\sim}{W}\,\underset{\sim}{u} \quad , \end{aligned} \tag{A-3}$$

where $\underset{\sim}{\epsilon}_o$ are the centroidal plastic strains, $\underset{\sim}{u}$ are the nodal displacements, and $\underset{\sim}{W}$ represents the matrix of coefficients associated with the strain-nodal displacement relations. By using the elastic stress-strain relation

$$\underset{\sim}{\sigma} = \underset{\sim}{C}\,\underset{\sim}{\epsilon} \tag{A-4}$$

and Eq. (A-2), Eq. (A-1) becomes

$$U = \frac{1}{2}\int_V (\underset{\sim}{\epsilon}_t - \underset{\sim}{\epsilon}_p)'\,\underset{\sim}{C}(\underset{\sim}{\epsilon}_t - \underset{\sim}{\epsilon}_p)dV \quad .$$

From a consistent energy approach, or equivalently from the application of the principle of virtual work, the following equation is arrived at:

$$\underset{\sim}{f} = \underset{\sim}{k}\,\underset{\sim}{u} - \underset{\sim}{k}_p\underset{\sim}{\epsilon}_o \tag{A-5}$$

where

$$\underset{\sim}{k} = \int_V \underset{\sim}{W}' \underset{\sim}{C} \underset{\sim}{W} dV$$

$$\underset{\sim}{k}_p = \int_V \underset{\sim}{W}' \underset{\sim}{C}\, dV \quad . \tag{A-6}$$

These general principles are applied to the specific element model outlined in the text and shown in Fig. 2, and it is found that the contribution to the stiffness matrices, expressed in Eqs. (A-6), of each of the two components of the heterogeneous material, viz., the in-plane fiber-matrix and interlaminar matrix segments, is uncoupled. The total strains, $\underset{\sim}{\epsilon}_t$, are represented in the form

$$\underset{\sim}{\epsilon}_t = (\underset{\sim}{\epsilon}_m \underset{\sim}{\gamma}) \quad ,$$

where $\underset{\sim}{\epsilon}_m$ contains the in-plane membrane strain components, ϵ_x, ϵ_y, and γ_{xy}, and $\underset{\sim}{\gamma}$ contains the interlaminar shearing strains γ_{xz}, γ_{yz}. To derive the stiffness matrices for an individual element explicitly, a notation that describes the contribution of the various segments is used. This is shown for the strain components and elsewhere, when necessary for clarity. The strain components are denoted as

$$\underset{\sim}{\epsilon}_m = (\underset{\sim}{\epsilon}_m^1 \; \underset{\sim}{\epsilon}_m^2 \; \ldots \; \underset{\sim}{\epsilon}^n)$$

$$\underset{\sim}{\gamma} = (\underset{\sim}{\gamma}^{1,2} \; \underset{\sim}{\gamma}^{2,3} \; \ldots \; \underset{\sim}{\gamma}^{n-1,n}) \quad ,$$

where n is the number of layers, and

$$\underset{\sim}{\epsilon}_m^i = (\epsilon_x^i \; \epsilon_y^i \; \gamma_{xy}^i)$$

$$\underset{\sim}{\gamma}^{i,j} = (\gamma_{zx}^{i,j} \; \gamma_{zy}^{i,j}) \quad ,$$

in which a single superscripted quantity, say ϵ_x^i, denotes an in-plane property associated with layer i, and a double superscripted

quantity, say $\gamma_{zx}^{i,j}$, denotes an interlaminar property associated with the matrix region between layers i and j.

The total strain-displacement relation, Eq. (A-3), is expressed as

$$\begin{pmatrix} \underset{\sim}{\epsilon}_m \\ \underset{\sim}{\gamma} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{W}_m \\ \underset{\sim}{W}_s \end{pmatrix} \underset{\sim}{u} \quad . \tag{A-7}$$

Conforming to the outline notation,

$$\underset{\sim}{W}_s = \begin{pmatrix} \underset{\sim}{W}_s^{1,2} & & & \\ & \underset{\sim}{W}_s^{2,3} & & \\ & & \ddots & \\ & & & \underset{\sim}{W}_s^{n-1,n} \end{pmatrix} \quad , \quad \underset{\sim}{W}_m = \begin{pmatrix} \underset{\sim}{W}_m^{1} & & & \\ & \underset{\sim}{W}_m^{2} & & \\ & & \ddots & \\ & & & \underset{\sim}{W}_m^{n} \end{pmatrix}$$

and

$$\underset{\sim}{u} = (\underset{\sim}{u}^1 \; \underset{\sim}{u}^2 \; \ldots \; \underset{\sim}{u}^n)$$

where

$$\underset{\sim}{u}^n = (u_i^n v_i^n u_j^n v_j^n u_k^n v_k^n) \quad ,$$

subscripts denoting nodal points, and,

$$\underset{\sim}{W}_s^{i,j} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & -1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & -1 & 0 & -1 & 0 & -1 \end{bmatrix} \Big/ 3t^{i,j} \quad ,$$

with $t^{i,j}$ representing the thickness. The quantities associated with the triangular in-plane orthotropic layers are available in the literature (see, e.g., Ref. 4), and will not be represented here.

The material property matrix, $\underset{\sim}{C}$ from Eq. (A-4), is expressed as

$$\underset{\sim}{C} = \begin{pmatrix} \underset{\sim}{C}_m & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{C}_s \end{pmatrix} \quad , \tag{A-8}$$

where $\underset{\sim}{C}_m$ represents the in-plane material properties, and $\underset{\sim}{C}_s$ represents the material properties of the shear segments, where $\underset{\sim}{C}_s^{i,j}$ is expressed as

$$\underset{\sim}{C}^{i,j} = G^{i,j} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ,$$

with $G^{i,j}$ denoting the interlaminar shear modulus.

The elastic stiffness matrix can now be separated into its membrane and interlaminar components by substituting Eqs. (A-7) and (A-8) into the first of Eqs. (A-6), resulting in

$$\underset{\sim}{k} = \underset{\sim}{k}_m + \underset{\sim}{k}_s ,$$

where

$$\underset{\sim}{k}_m = \int_V \underset{\sim}{W}_m' \, \underset{\sim}{C}_m \, \underset{\sim}{W}_m \, dV \quad , \quad \underset{\sim}{k}_s = \int_V \underset{\sim}{W}_s' \, \underset{\sim}{C}_s \, \underset{\sim}{W}_s \, dV$$

Written explicitly,

$$\underset{\sim}{k}_m = \begin{bmatrix} \underset{\sim}{k}_m^1 & & & \\ & \underset{\sim}{k}_m^2 & & \\ & & \ddots & \\ & & & \underset{\sim}{k}_m^n \end{bmatrix} , \quad \underset{\sim}{k}_s = \frac{A}{9} \begin{bmatrix} \underset{\sim}{p}^{1,2} & -\underset{\sim}{p}^{1,2} & & \\ -\underset{\sim}{p}^{1,2} & \underset{\sim}{p}^{1,2} + \underset{\sim}{p}^{2,3} & -\underset{\sim}{p}^{2,3} & \\ & -\underset{\sim}{p}^{2,3} & \ddots & \\ & & & \underset{\sim}{p}^{n-1,n} \end{bmatrix}$$

(A-9)

where A is the plane area of the element,

$$\underset{\sim}{p}^{i,j} = \frac{G^{i,j}}{t^{i,j}} \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} ,$$

and $\underset{\sim}{k}_m^i$ is the stiffness matrix for the i^{th} orthotropic layer. The material properties and thickness can be varied from layer to layer, as can be seen from Eqs. (A-9).

Allowing plastic deformation in the interlaminar region only,

$$\underset{\sim}{\epsilon}_p = (\underset{\sim}{0} \quad \underset{\sim}{\gamma}^p) \quad .$$

This reduces $\underset{\sim}{k}_p\underset{\sim}{\epsilon}_o$ in Eq. (A-5) to $\underset{\sim}{k}_p^*\underset{\sim}{\gamma}_o^p$, where $\underset{\sim}{k}_p^{*i,j}$ is expressed as

$$\underset{\sim}{k}_p^{*i,j} = \frac{G^{i,j}A}{3}\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & -1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & -1 & 0 & -1 & 0 & -1 \end{bmatrix}$$

APPENDIX B

NORMAL INTERLAMINAR SHEAR

A closed form solution for a $[0/90]_s$ laminate is obtainable by using the methods of Ref. 1 where, as mentioned in the text, the model is the same as that used in this paper. In the following, the notation used is that of Ref. 1. The reduced stiffness coefficients are

$$Q_{11}^{(1)} = Q_{22}^{(2)} = Q_{11}$$

$$Q_{22}^{(1)} = Q_{11}^{(2)} = Q_{22}$$

$$Q_{12}^{(1)} = Q_{12}^{(2)} = Q_{12}$$

$$Q_{33}^{(1)} = Q_{33}^{(2)} = Q_{33}$$

This leads to a set of differential equations as follows:

$$\begin{aligned} (Q_{33}\rho - k)A + kB &= 0 \\ (Q_{33}\rho - k)B + kA &= 0 \\ (Q_{22}\rho - k)C + kD &= 0 \\ (Q_{11}\rho - k)D + kC &= 0 \quad , \end{aligned} \tag{B-1}$$

where $k = G/hh_o$, and the coefficients correspond to the magnitudes of the displacements,

$$u_1 = Ae^{\rho y} \quad , \quad u_2 = Be^{\rho y} \quad , \quad V_1 = Ce^{\rho y} \quad , \quad V_2 = De^{\rho y}, \tag{B-2}$$

at the center, $y = 0$. The four roots of Eqs. (B-1), and their corresponding amplitude ratios are

$$\begin{aligned} \rho_1 &= 2k/Q_{33} && , \quad A_1 = -B_1 \\ \rho_2 &= 0 && , \quad A_2 = B_2 \\ \rho_3 &= k(Q_{11} + Q_{22})/Q_{11}Q_{22} && , \quad C_3 = -(Q_{11}/Q_{22})D_3 \\ \rho_4 &= 0 && , \quad C_4 = D_4 \quad . \end{aligned} \tag{B-3}$$

Substituting Eqs. (B-3) into the displacement representation, Eqs. (B-2), and using the stress-strain relations, we have

$$\tau_{zx} = (G/h)(u_1 - u_2) \quad , \quad \tau_{zy} = (G/h)(V_1 - V_2)$$

$$\begin{Bmatrix} \sigma_x^i \\ \sigma_y^i \\ \tau_{xy}^i \end{Bmatrix} = \begin{bmatrix} Q_{11}^{(i)} & Q_{12}^{(i)} & Q_{13}^{(i)} \\ & Q_{22}^{(i)} & Q_{23}^{(i)} \\ \text{SYMM} & & Q_{33}^{(i)} \end{bmatrix} \begin{Bmatrix} \epsilon_x^i \\ \epsilon_y^i \\ \gamma_{xy}^i \end{Bmatrix} , \qquad i = 1,2 \ ;$$

the laminate stresses are obtained,

$$\sigma_y^1 = -\sigma_y^2 = -\frac{u^*}{a} Q_{12} \frac{(Q_{22} - Q_{11})}{(Q_{11} + Q_{22})} \left(1 - \frac{\text{ch } \sqrt{\rho}\, y}{\text{ch } \sqrt{\rho}\, b}\right)$$

$$\sigma_x^1 = \frac{u^*}{a} \left[Q_{11} - \frac{2Q_{12}^2}{Q_{11} + Q_{22}} + \frac{Q_{12}^2 (Q_{22} - Q_{11})}{Q_{22}(Q_{11} + Q_{22})} \frac{\text{ch } \sqrt{\rho}\, y}{\text{ch } \sqrt{\rho}\, b} \right]$$

$$\sigma_x^2 = \frac{u^*}{a} \left[Q_{22} - \frac{2Q_{12}^2}{Q_{11} + Q_{22}} - \frac{Q_{12}^2 (Q_{22} - Q_{11})}{Q_{11}(Q_{11} + Q_{22})} \frac{\text{ch } \sqrt{\rho}\, y}{\text{ch } \sqrt{\rho}\, b} \right]$$

$$\tau_{zy} = \frac{u^*}{a} Q_{12} \frac{G(Q_{22} - Q_{11})}{hQ_{11}Q_{22}\sqrt{\rho}} \frac{\text{sh } \sqrt{\rho}\, y}{\text{ch } \sqrt{\rho}\, b} \ ,$$

where

$$\rho = k(Q_{11} + Q_{22})/Q_{11}Q_{22}hh_o \ .$$

For sufficiently wide plates the classical plate theory solution is recovered in the central region of the laminate.

SESSION 8. NON–STRUCTURAL APPLICATIONS

Session Chairman

D. Zonars

Air Force Flight Dynamics Laboratory
Wright-Patterson Air Force Base, Ohio

STRUCTURAL ANALYSIS AND THE HUMANITIES

by

John F. Abel*
Princeton University, Princeton, New Jersey

The contemporary challenge that engineers should become more sensitive to the effects of their technological decisions is in essence an exhortation to assimilate the cultural values implicit in "historic insight and social memory." A Princeton University program is described which responds to this challenge by introducing engineers to their own humanistic tradition. To illustrate how matrix analysis of structures can be taught in a humanistic context and can be applied to humanistic research, five case studies from the Princeton program are discussed. These include: the analysis of Eiffel's Maria Pia Bridge, a design study of Revell's Toronto City Hall, research of the structural performance of the ribbed vaults of Gothic cathedrals, the analysis of a Maillart concrete plane frame, and a critical study of Saarinen's Kresge Auditorium.

Humanistic Studies in Engineering

Matrix analysis of structures is being increasingly applied to humanistic studies at Princeton University. These endeavors are part of an overall educational program which is responding to the pervasive concern in American society regarding the role of technology in modern culture. The purposes of this paper are to describe briefly the Humanistic Studies in Engineering program and to discuss some of the projects from this program which entail the matrix analysis of structures.

Humanism and Engineering Technology

Much has been said in the last few years about the effects of technology on society. Engineers are certainly conscious of the criticisms levied against their profession for the failings and ill effects of engineering technology: the deterioration of the environment; the dehumanization resulting from computerization and automation; the apparent inability of public works engineering to ameliorate the blight of urbanization; and the unconscionable development of frightening and

---

* Research Associate and Lecturer, Department of Civil and Geological Engineering.

expensive weapons systems. On the other hand, the engineering profession has made acknowledged contributions to the positive aspects of the prevalent western life style, a style based upon a high standard of living with its familiar technological underpinnings: industrialization, mobility, public health services, modern utilities, etc. This superficial listing of some favorable and unfavorable effects of technology on society neglects the complex technical-political-economic interactions that characterize societal processes. Nevertheless, it is clear both that engineering technology is a major influence in modern culture and that cultural values are invoked in the discussion of technology and society.

One might define humanism as a concern with cultural values, with the "expressive, moral and contemplative aspects of living"(1)*. Thus when Lewis Mumford criticizes engineers for lacking "historic insight and social memory" in conceiving and executing projects, he is accusing the profession of a lack of humanism. This is the essence of the contemporary challenge that engineers should become ever more sensitive to the widespread effects of their technological decisions. Of course, the development of a closer relationship between technology and the humanities presents a challenge to humanists as well. They must recognize and explore the new possibilities created by technology, the "new attitudes and perceptions, new belief systems, and new ways of looking at the world."(1) To meet the goal of a closer relationship between engineering and the humanities, an experimental program in engineering education has been initiated at Princeton.

The Princeton Program(2)

The first objective of the Humanistic Studies in Engineering program is to educate engineers who, while fully competent in engineering science and design, will center their careers on a union of technology and the humanities. The principal focus here is on introducing engineers to their own humanistic tradition through a modified engineering curriculum as well as through the usual approach of elective courses in the humanities. To include humanistic considerations in their own teaching, the engineering faculty involved in the program have been studying in the humanities. In addition, humanistic scholars such as art and architectural historians have been increasingly involved in the program through collaboration with the engineering faculty and through the broadening of their own interests and course content to include aspects related to engineering.

In striving to achieve a closer relationship between engineering and the humanities, the program must develop a communication that is mutually stimulating. Therefore, the second major objective is to establish a tradition of engineering research directed toward new

---

*Numbers in parentheses indicate references listed at the end of this paper.

humanistic understanding. Significant research will encourage the involvement and collaboration of humanists in engineering education. So far, the research efforts at Princeton have centered on the history of architecture and the related arts. Throughout the research aspects of the program, a consistent effort is made to maintain a strong technical component that is absent from the research of traditional humanists. For instance, quantitative structural analysis has proven to be a decisive tool in the study of Gothic architecture.

The program is being implemented in several ways. In the classroom a humanistic emphasis has been incorporated into the context of several civil engineering courses by discussing applications and examples which have cultural and historical as well as technical significance such as James Eads' St. Louis Bridge. Undergraduate students are encouraged to undertake course projects and independent work which further explore, for example, the relationship between engineering and aesthetics. Research efforts involving graduate students and faculty in the humanities as well as in engineering are another aspect of the implementation; these efforts concentrate on history, biography, the relation between engineering and aesthetics, and the evaluation of contemporary engineering works. A natural adjunct of this research is the accumulation of documents and other source material related to humanistic studies in engineering. Publication of the research is an important element of the program; additional types of publications planned are engineering textbooks written in the context of the humanities and critical essays regarding specific engineering and architectural projects. Faculty and student interaction is furthered by periodic colloquia on humanistic research related to engineering and on value judgments encountered in engineering practice such as in the development of low-cost urban housing. Finally, to broaden the impact of the program and to stimulate similar undertakings elsewhere, public lectures are given at various campuses and other locations, research fellows from other institutions are brought to Princeton, and a national conference ("Civil Engineering: History, Heritage and the Humanities") has been held.

Structural Engineering and the Humanities

The Humanistic Studies in Engineering program is presently based in the civil engineering department. Moreover, it strongly emphasizes structural engineering, largely because a major stimulus for the development of the program was the interaction between engineering faculty and architecture students in structures courses. The interests that these students initially expressed are reflected in the nature of the current program studies: the analysis of the structures of architects such as Antonio Gaudi and Eero Saarinen, the study of the life and works of the great engineer-builders such as Pier Luigi Nervi and Eduardo Torroja, and the analysis of other structures significant in the history and theory of architecture such as Gothic cathedrals. Corresponding to the close relationship between structures, architecture, and the visual arts, the humanists from the Princeton faculty who are most actively involved in the program at present are architects and art historians.

The intermediate level courses in structural analysis and design have been broadened to include a humanistic perspective, and thus they serve as examples of the implementation of the program in the classroom. In addition to the customary rigorous consideration of engineering science and design, two emphases are present in these courses. The first is the consideration of the origins of modern structures and practice(3). The humanistic tradition in modern structural engineering is firmly rooted in the careers of the great builders of the late 19th and early 20th centuries, men like John Roebling, James Eads, Gustave Eiffel, and Robert Maillart. The discussion of the ideals and accomplishments of these men serves a dual purpose: not only do their works have a definite influence on modern architecture and art, but the technical aspects of their structures also provide worthy examples for the study of structural theory and design. The second humanistic emphasis in structures courses is the criticism of contemporary buildings. Qualitative and quantitative analyses of both outstanding and ordinary structures permit the student to develop an ability to relate visual elegance to structural efficiency.

Humanistic research in structural engineering is a natural extension of the classroom emphases, that is, the three main elements are history, biography, and criticism. Several topics are worth mentioning to illustrate the current thrust of research. Studies of Eads' St. Louis Bridge and Roebling's Brooklyn Bridge highlight the cultural and historical significance of these structures. Not only are the structures themselves symbols of the city and the era that produced them, but the historical investigations are enlightening introductions to 19th century politics. Moreover, both bridges have had an impact on the visual arts, especially the Brooklyn Bridge which is profusely represented, for example, in the paintings of Joseph Stella and the etchings of John Marin. Another research topic is a biographical study of Robert Maillart. This effort is still in an early stage wherein source documents and interviews are being obtained from Maillart's family and colleagues. An example of criticism is an examination of some of the buildings of Sir Owen Williams, a civil engineer turned architect. Finally, an investigation of the relation between engineering and aesthetics is a joint study by engineers and a social scientist of the design competition for the Washington Bridge over the Harlem River and of the planning decisions involved in subsequent changes in the urban area adjacent to the bridge.

Model studies are the primary structural analysis method in many of the projects that have already been mentioned. However, computer techniques are being increasingly applied to humanistic research, both alone and as a complement to model studies. In the remainder of this paper various case studies that utilize matrix methods of structural analysis will be discussed in further detail.

Case Studies Involving Matrix Analysis

The work reported here involves no original contributions to the technology of matrix analysis with one exception (4) which will be published in detail elsewhere. The purposes, therefore, of recounting these investigations are first to demonstrate how the matrix analysis of structures can be studied in a humanistic context and second to indicate a range of humanistic structural research to which the powerful tool of computer analysis can be applied. Consequently, in the discussions which follow emphasis will be placed upon the description of the problems and their significance rather than upon detailed quantitative results.

Eiffel's Maria Pia Bridge

The work of Gustave Eiffel has had a profound effect not only upon the practice of engineering, but also upon architecture and art.(3) His technical contributions included advances both in the design, fabrication and erection of metal structures and in the analysis and design for wind loadings. Developments in these areas can be traced in the series of iron railroad viaducts designed and constructed by Eiffel and his associates from 1864 to 1886. The culimination of Eiffel's structural work, the tower bearing his name, is a classic of metal construction. Its form is an eloquent expression of proper design of a tall structure to resist wind loading. This engineering work has become the best known symbol of France. Further evidence of the cultural significance of Eiffel's work is the fact that such influential observers as LeCorbusier(5) and S. Giedion(6) prominantly feature his viaducts and tower in their treatises on architecture.

A study project on Eiffel, presently concentrating on one of the Eiffel viaducts, the Maria Pia Bridge across the Douro River at Oporto, Portugal, is providing an opportunity for engineering learning experiences in a humanistic context. The bridge, designed and construced during 1875 to 1878, was the first two-hinged arch executed by Eiffel and with a span of 160 meters (525 feet) was the longest metal arch constructed up to that time. (See Figures 1 and 2.) An erection scheme was devised wherein the arch was erected without scaffolding from the springings upward; the two rising segments of the arch were suspended from the side-span superstructure by steel cables as successive pre-fabricated sections were lifted from barges. The overall design and erection was nearly 50% cheaper than its nearest competitor in the bidding. Historically, this structure marks the development of many of the new ideas that were later incorporated on a grander scale into Eiffel's most famous bridge, the Garabit Viaduct. Like the Garabit, the Maria Pia elegantly expresses its efficient structural function(3). For example, in plan the wider spacing of the legs of the arch at the support reflect the resistance to lateral overturning by wind forces.

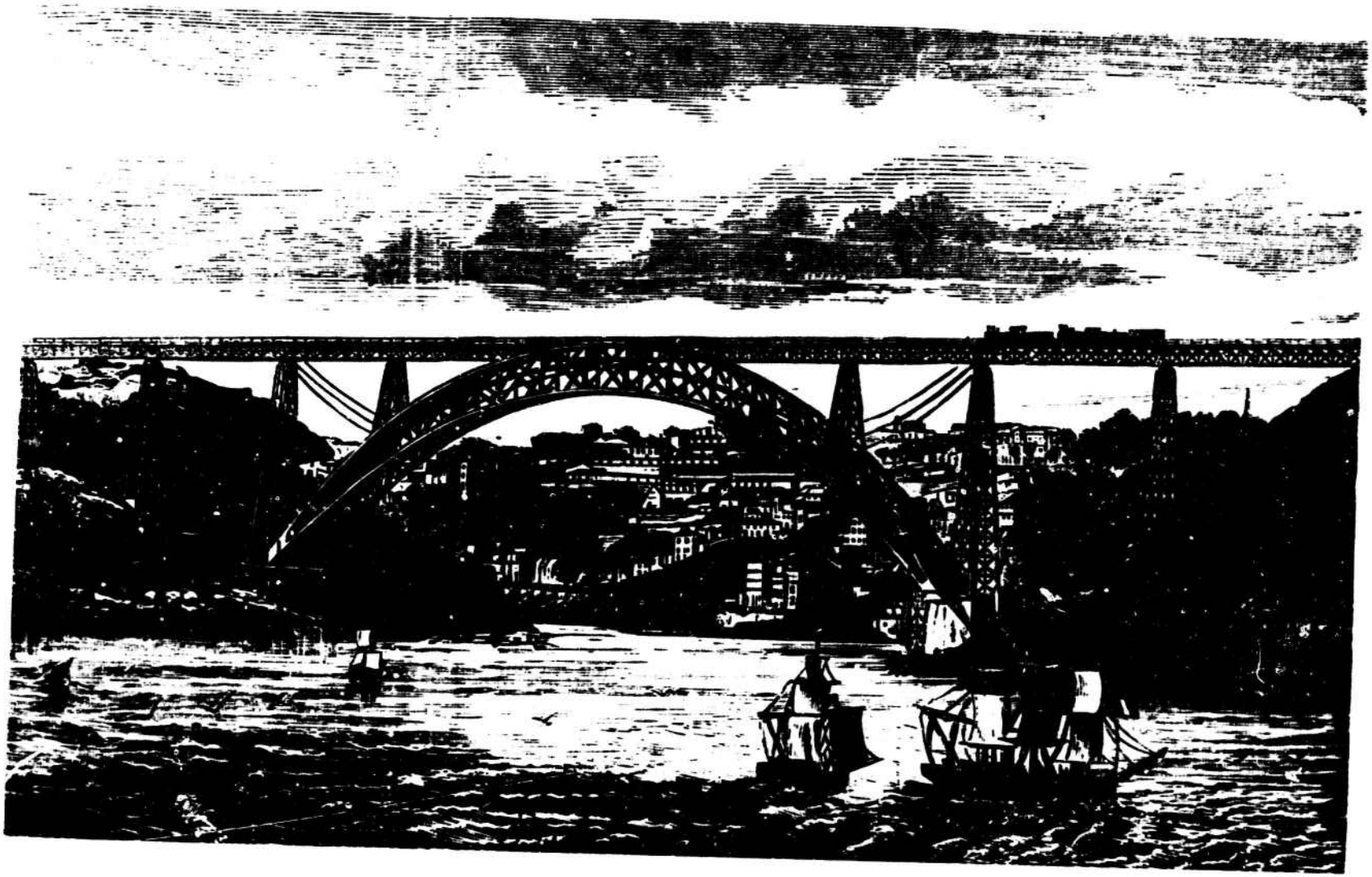


FIGURE 1. EIFFEL'S MARIA PIA BRIDGE ACROSS THE DOURO RIVER, OPORTO, PORTUGAL (from The Engineer, Vol. 45, 1878)

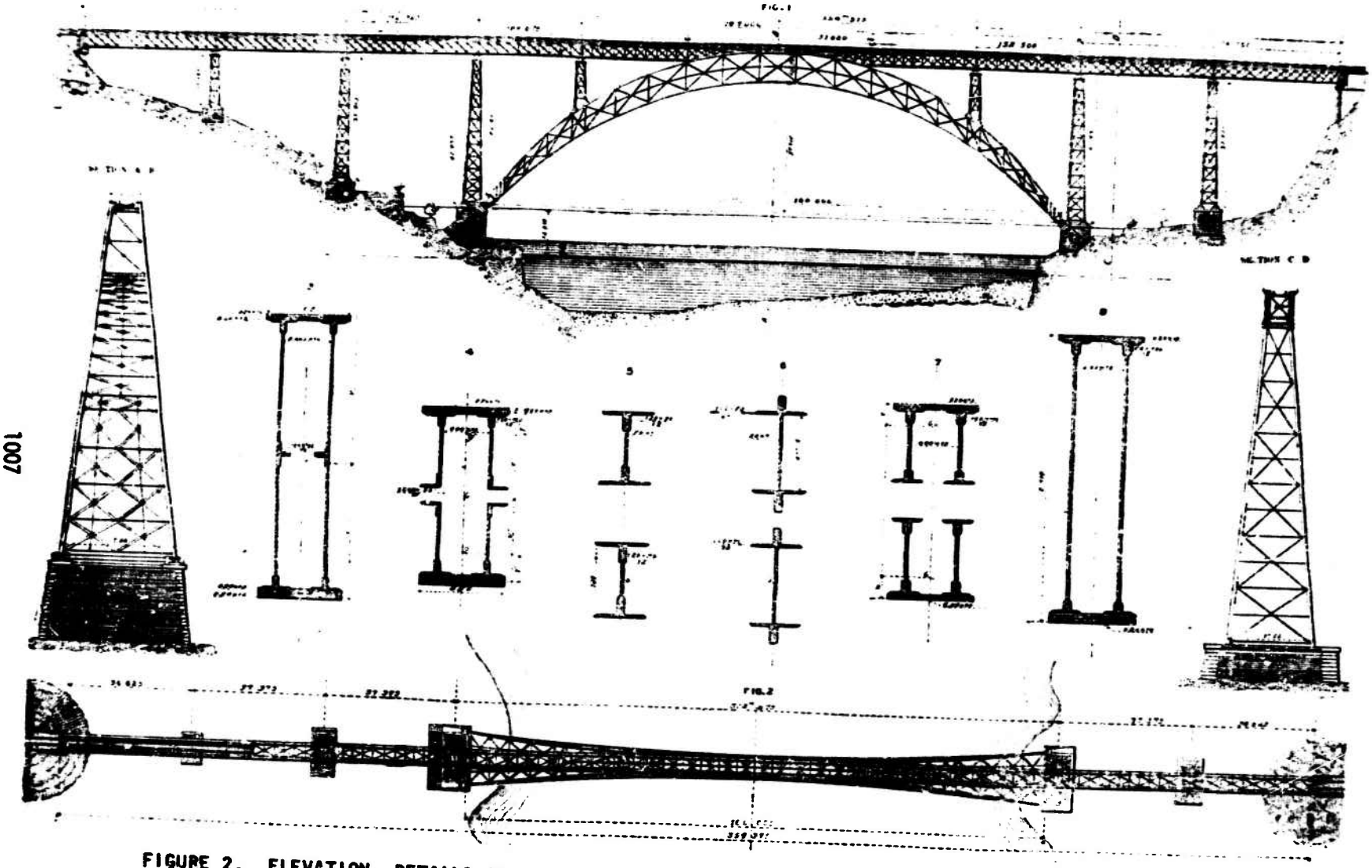


FIGURE 2. ELEVATION, DETAILS AND PLAN OF MARIA PIA BRIDGE (from The Engineer, Vol. 45,1878)

Fortunately, a paper by T. Seyrig(7), the project engineer from Eiffel's firm for the Douro bridge, gives a reasonably complete account of both the structural dimensions and the structural analysis of the bridge. Because the current investigator was seeking an exercise in the computer analysis of structures, the technical aspect of the study was the analysis of the arch for vertical loads with ICES-STRUDL II and a comparison of the modern analysis results with the original analysis. Seyrig's analysis of the internally and externally indeterminate trussed arch is interesting in itself. By using the gross areas and moments of inertia of the arch, the horizontal thrusts are computed by an approximate method* analogous to Castigliano's Second Theorem. The axial forces in the chord members are then determined from the gross moment and axial force in each panel. Apparently, the web member forces are conservatively estimated by first omitting one set of diagonal web members to obtain an internally determinate truss and then repeating the process with the other set of diagonals absent. A comparison of some typical results for horizontal reactions and chord member stresses is presented in Table 1; the agreement is quite good.

Many student exercises undertaken in current structural analysis courses are portrayed as anonymous line drawings or are simple example drawn from current practice. A supplemental purpose of studies such as the one described is to develop student exercises based on structures which are technically outstanding and historically significant. Such problems provide a natural opportunity to explore further the aesthetic and other humanistic aspects of these and associated structures. In the particular case of the Maria Pia Bridge analysis, additional opportunities accrue to the student. He can observe the application of hand methods of approximate analysis which, although now out of fashion, are still useful for preliminary analysis. On the other hand, he can see the clear advantage of computer techniques for a complete analysis when he observes the labor required for hand analysis for, say, wind loadings and assymetrical vertical loads.

Revell's Toronto City Hall

A senior thesis that exemplifies undergraduate structural engineering in a humanistic context is a design study of the Toronto City Hall(8). One component of this work is an examination of the origins, design competition, construction, and functioning of the municipal complex. This treatment includes a critical analysis of the aesthetic appeal and structural efficiency of the buildings. A second, more technical component of the study is the exploration of an alternative structural scheme for one of the buildings in the complex. Enough structural analysis and design was undertaken in this aspect to provide the student with an insight into engineering practice for concrete buildings.

---

*This method is attributed to deDion(7).

TABLE 1. COMPARISONS OF ANALYSES OF MARIA PIA BRIDGE

	Panel Number	Top or Bottom Chord	Loading 1		Loading 2		Loading 3	
			STRUDL 2	Seyrig	STRUDL 2	Seyrig	STRUDL 2	Seyrig
Horizontal Reaction, kg			140,950	140,310	95,770	95,000	70,570	70,150
Member stresses, kg/mm^2	3	top	-1.25	-1.08	0.53	0.20	-2.09	-2.20
	3	bottom	-2.06	-1.85	-2.00	-1.96	0.01	0.32
	5	top	-1.64	-1.60	-0.22	-0.16	-2.46	-2.63
	5	bottom	-1.47	-1.32	-1.84	-1.81	0.74	0.94
	7	top	-1.45	-1.40	-0.88	-0.89	-2.23	-2.23
	7	bottom	-1.61	-1.41	-1.31	-1.22	0.51	0.69
	9	top	-1.86	-1.92	-1.90	-1.94	-2.01	-2.07
	9	bottom	-0.90	-0.71	-0.04	0.09	0.69	0.84
	11	top	-1.89	-1.89	-2.10	-2.16	-0.95	-0.95
	11	bottom	-0.69	-0.51	0.34	0.62	-0.35	-0.25

Notes: Panels are numbered consecutively from the hinge. The top center panel is number 11. See Figure 2.

All loadings represent distributed loadings of intensity 4000 kg/m.
Loading 1 is a symmetrical loading of the full 160-meter arch span.
Loading 2 is a symmetrical loading of 80 meters in length.
Loading 3 is an asymmetrical loading of 80 meters in length of the half arch span. For loading 3 the stresses shown are for the members beneath the load.

In 1957 the city of Toronto launched a worldwide architectural design competition for a city hall and civic square. The city was seeking a building that would be a dramatic symbol of Toronto, a symbol of achievement and promise. The plaza accompanying the building was to be designed for the pleasure of the urban residents. From among 520 entries in the first stage of the competition, eight finalist were chosen to submit more detailed designs. Ultimately, the unanimous decision of the jury was that the design submitted by the Finnish architect Viljo Revell was the most original and that it should be implemented (Figure 3). However, the jurors were not without reservations regarding the design. There was no doubt that the strikingly curved buildings would remain a prominent feature on the Toronto skyline regardless of other buildings that might surround it in the future. Nevertheless, a minority of the jurors were disturbed that the windowless outer wall turned a blank face on the urban redevelopment area to the north. This lack of windows on one wall is also a potential shortcoming from the standpoint of internal space utilization.

The structural concept of the curved towers is of particular concern here. To avoid interior columns, Revell had envisioned the outer walls as curved shells from which the floors would be cantilevered. The jury's reaction to this scheme was that the omission of columns would unnecessarily add to the cost of the building without significantly altering the visual effect. Oddly, despite the jury's recommendations, the original structural design proceeded without columns; however, this approach was abandoned when the shell thickness reached 32 inches and still was inadequate to support the cantilevered floors(8). Radial beams and a row of columns along the midline of the buildings were added. Nevertheless, the overall towers were designed as vertical cylindrical shells made orthotropic by the horizontal floor slabs and a system of reinforcing columns(9). As built, the outer wall of the east tower is an 18-inch thick concrete monolith, horizontally and vertically reinforced.

Stimulated by reservations about the windowless wall and about the structural rationale of the shell, the student decided to investigate a structural scheme that might permit a greater architectural flexibility. If the towers were designed as frames, the shell wall could be partially or totally omitted. No changes in the basic floor plans were considered in order to limit the scope of the work so that one man could reasonably accomplish the study. The taller east tower was chosen. An initial design based on the existing framing dimensions was selected: radial beams, columns along the outer wall, and interior columns were already present, while tangential beams were added to complete the framing. ICES-STRUDL II was employed to analyze selected vertical and horizontal segments of this three-dimensional frame for dead, live, and wind loadings. It was found that by current code standards, concrete framing with the existing cross sections could be adequately reinforced to satisfy the necessary strength requirements. Although no new design for the outer facade was specifically proposed, the results of the structural study demonstrated new aesthetic potentialities.

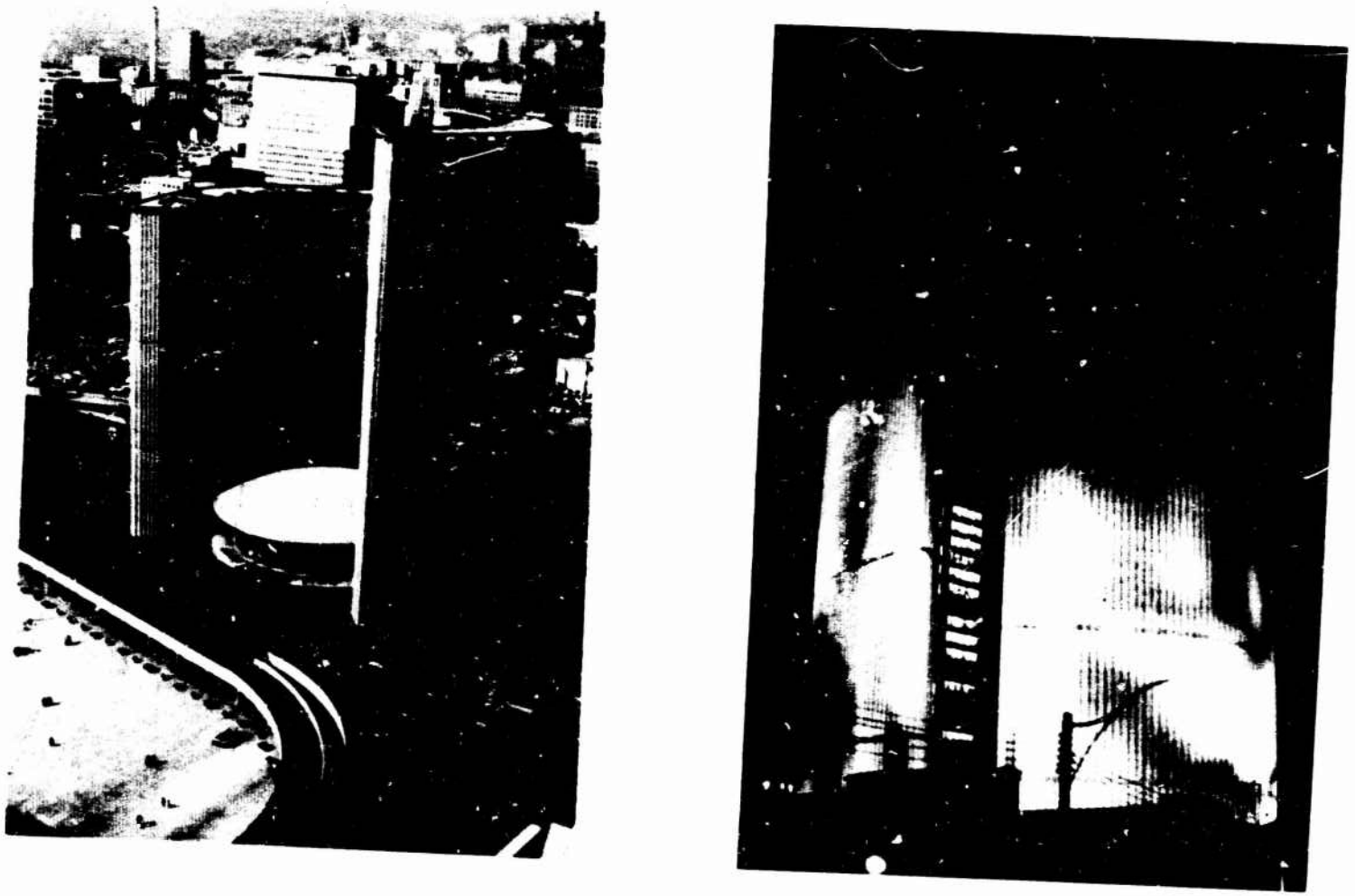


FIGURE 3. FRONT AND BACK VIEWS OF REVELL'S TORONTO CITY HALL

Gothic Cathedrals

One of the major influences on architecture and art has been the Gothic cathedral. These buildings have a manifest historical and cultural significance but in addition they have played a key role in the formulation of the architectural theory of rationalism(10). This theory was succinctly stated by Louis Sullivan as "form follows function." The first modern Gothic scholar, Eugène Viollet-le-Duc, laid the foundations for the rationalist theory in studies associated with his work of preserving and restoring the French Gothic cathedrals. He abstracted the theory from his contention that every element of the Gothic cathedral had a structural function(11), and he soon acquired a following among contemporary architects such as Sullivan. More recent inheritors of the rationalist tradition include Antonio Gaudi, Frank Lloyd Wright, and Mies Van der Rohe and engineers like Maillart, Torroja, and Nervi. The rationalist theory is more completely delineated in References 3, 4, and 10.

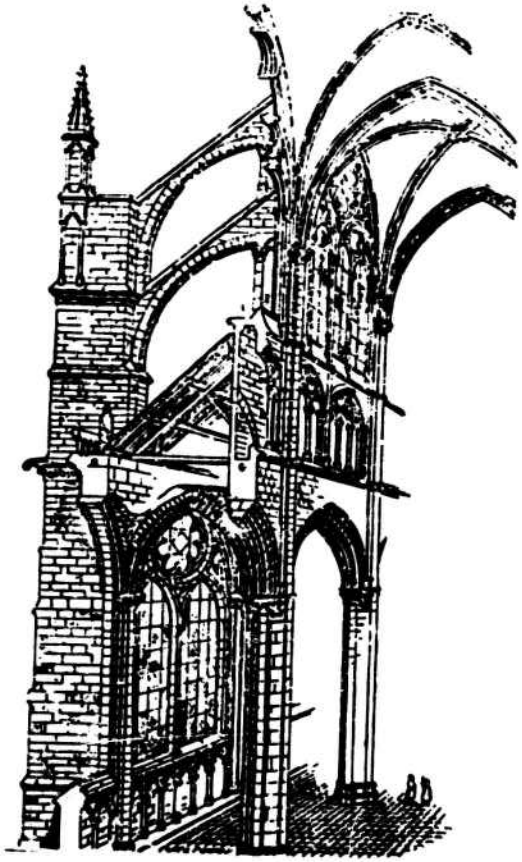
Counter to the rationalist school of thought are those who believe that architectural form is derived from purely aesthetic intentions.* Most of the debate on these opposing theories that has occurred since Viollet-le-Duc's time has involved structural theory but has consisted essentially of qualitative arguments set forth by architects and historians. There is a clear opportunity for modern structural analysts to contribute to the settlement of some of these controversies. Hence, a continuing line of faculty and graduate research at Princeton has been studies of the theory of rationalism.

Although the theory of rationalism is not specifically tied to Gothic architecture, many of the particular structural questions under debate by the theorists involve the Gothic simply because the influential Viollet-le-Duc drew most of his examples from his own broad experience with Gothic cathedrals. In Figure 4, two sectional views of the nave of Amiens Cathedral are shown. Although dimensions vary from church to church, essentially the same structural elements are present in other Gothic cathedrals. Most of the arguments about the rationality of Gothic construction are concerned with the vault and buttressing systems. In particular, the elements whose structural functions have been most disputed are the vault ribs, the flying buttresses, the outer buttress, and the pinnacles.† Numerical structural analysis in progress at

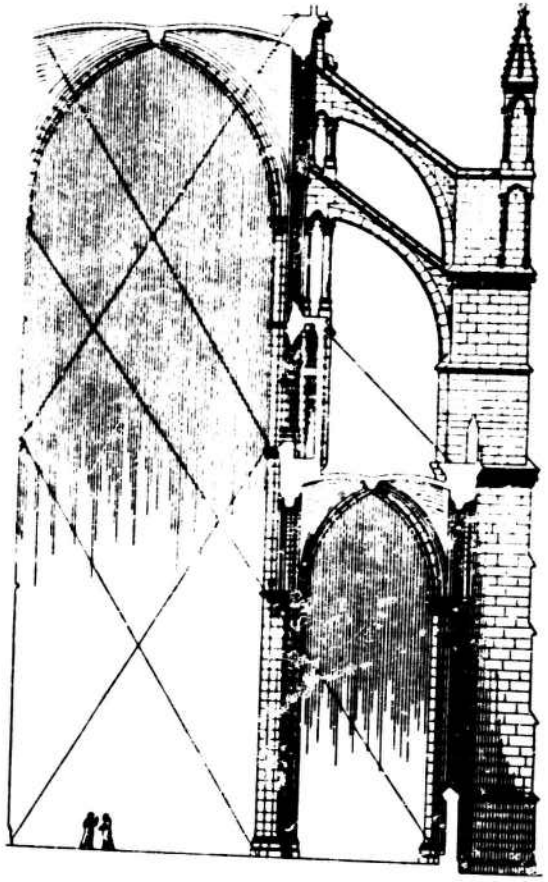
---

*As evidenced by their buildings, modern architects who are decidedly non-rationalist include Eero Saarinen and Jørn Utzon.

†Photoelastic model studies at Princeton have indicated that the pinnacle must be present and must be located as shown in Figure 4 in order to overcome tensions in the outer buttress of Amiens Cathedral.



(a) From Viollet-le-Duc



(b) From Dehio and Bezold, Die Kirchliche Baukunst des Abendlandes

FIGURE 4. NAVE SECTIONS OF AMIENS CATHEDRAL

Princeton are concerned with two aspects of the behavior of the ribbed vaults. The first project is to compute the horizontal thrusts developed by the vaulting system; these must be known in order to evaluate fully the role of the buttress system. The second study is to investigate the structural function of the ribs.

The geometry of the vault shells is quite complex. Figure 5 shows a model of two nave-vault bays of Cologne, a cathedral for which relatively complete geometric data is available. Note the double curvature of the shells, particularly along the nave-axis crown. The finite element method is being used for the numerical studies. In addition to the complex geometry, the difficult points in the analysis are the diagonal groins, the out-of-surface boundary conditions at the crown, and the introduction of the ribs. Dr. Sachs has developed a double curved shell element that accomodates the groin and the necessary boundary conditions, and he has computed the thrust(4). Further analyses are utilizing quadrilateral element assemblages of plane triangles(12). Preliminary numerical results are being verified by comparison with photoelastic model studies of the vault model shown in Figure 5. Because the project is still incomplete, it is too early to present any conclusions about the structural action of the ribs. Nevertheless, this research effort is an example of how matrix analysis of structures can contribute to architectural history scholarship.

A Concrete Frame by Maillart

An independent-study project recently undertaken by an undergraduate was the analysis of a concrete plane frame by two methods, matrix analysis and photoelastic modelling. The frame is a rather unusual one designed by Robert Maillart. The technical phase of the study enabled the student to learn more about the two methods of analysis by comparing their results. The nontechnical aspect was a critical study of the structure as it related to other work by Maillart. Hence, the general approach of this project is similar to that of the Maria Pia project discussed earlier.

Maillart (1872-1940) was a Swiss engineer-builder who, although less well-known and less influential than Eiffel, has made a considerable contribution to both engineering and the arts. In contrast to Eiffel's concentration on metal structures, Maillart was an early master in reinforced concrete construction. His arch bridges are concrete manifestations of rationalist design and visual elegance(6). Maillart's cultural impact is epitomized by a distinction few engineers receive: a sculptor, Max Bill, has assembled an architectural book entirely devoted to Maillart(13). In addition to a variety of examples of rational design, Maillart's technical contributions include the development of flat-slab floor construction with mushroom columns, the testing of large-scale concrete structures, the origination of the concept of the "shear center", and the advocacy of a theory of design for uniform safety in concrete structures(14).

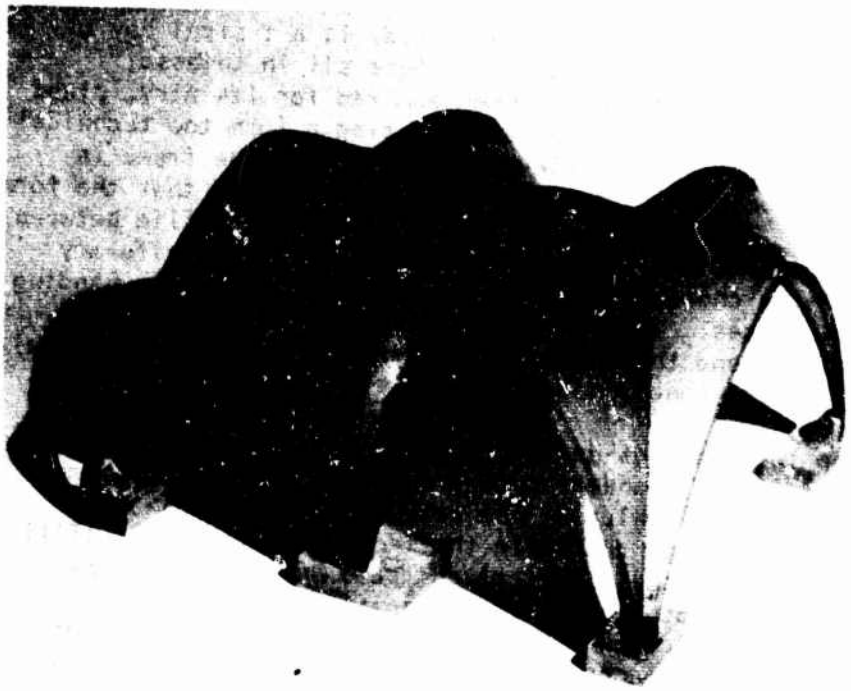


FIGURE 5. MODEL OF COLOGNE CATHEDRAL NAVE VAULTS

FIGURE 6. MAILLART'S SHED ADJACENT TO MAGAZZINI GENERALI, CHIASSO. SWITZERLAND

The Maillart frame selected for this study is a typical bay of the shed adjacent to the warehouse Magazzini Generali in Chiasso, Switzerland, Figure 6. This frame has been admired for its airy, plant-like form but also has been criticized as affecting a form too technical to be intuitively grasped(13). An elevation of half of the frame is shown in Figure 7. From this an engineer can certainly see that the form follows the function. The "truss" region with parabolic profile between the columns corresponds to an inverted moment diagram for a uniformly loaded beam; hence the chord members should tend to have little bending and equalized axial forces. Table 2 gives some results of both an ICES-STRUDL II analysis and the photoelastic model analysis. In general, both analyses confirm the functional interpretation of the form.

However, the two analyses do not agree very well for many of the individual members; it is in understanding this disagreement that much of the technical educational value of the analysis lies. Many of the members are not slender enough to enable the idealized frame to fulfill the assumptions inherent in the computerized frame analysis. On the other hand, although the photoelastic study is a full-field, two-dimensional analysis which accounts for complex behavior at the joints, it is not easy to obtain very accurate measurements of axial forces and bending moments in the members from the fringe photographs. (The forces and moments were chosen for comparison because they would be used in frame design.) Finally, it is valuable to contrast these linear elastic approaches with Maillart's ideas regarding design for uniform safety, ideas which implicitly call for an understanding of the real (cracked, nonlinear) behavior of the structure as a whole(14).

Saarinen's Kresge Auditorium

The modern implications of the theory of rationalism in architecture become evident in the critical analysis of contemporary structures. One such study undertaken at Princeton is of Eero Saarinen's Kresge Auditorium on the M.I.T. campus, Figure 8. This shell, completed in 1955, is exactly one-eighth of a sphere, supported at three points. The distance between abutments is 155 feet and the arched edges rise to 27 feet. Although the design was originally hailed by architectural critics as a rational solution, the structure has experienced enough structural shortcomings to indicate the contrary.

The first part of the study, the investigation of the particulars of the design, construction, and performance of the building, has historical significance because Kresge was one of the first large shell structures to be built in the United States. This part of the research was conducted by surveying the architectural and engineering literature and by interviewing individuals who were involved in the design and construction of the auditorium. A complete account of this study will be reported elsewhere; however, a few key findings are of interest here. For example, in the design stage, the engineers recognized the necessity of edge beams. Nevertheless, when the forms were removed from the shell, the edge deflections proved to be excessive, particularly with creep;

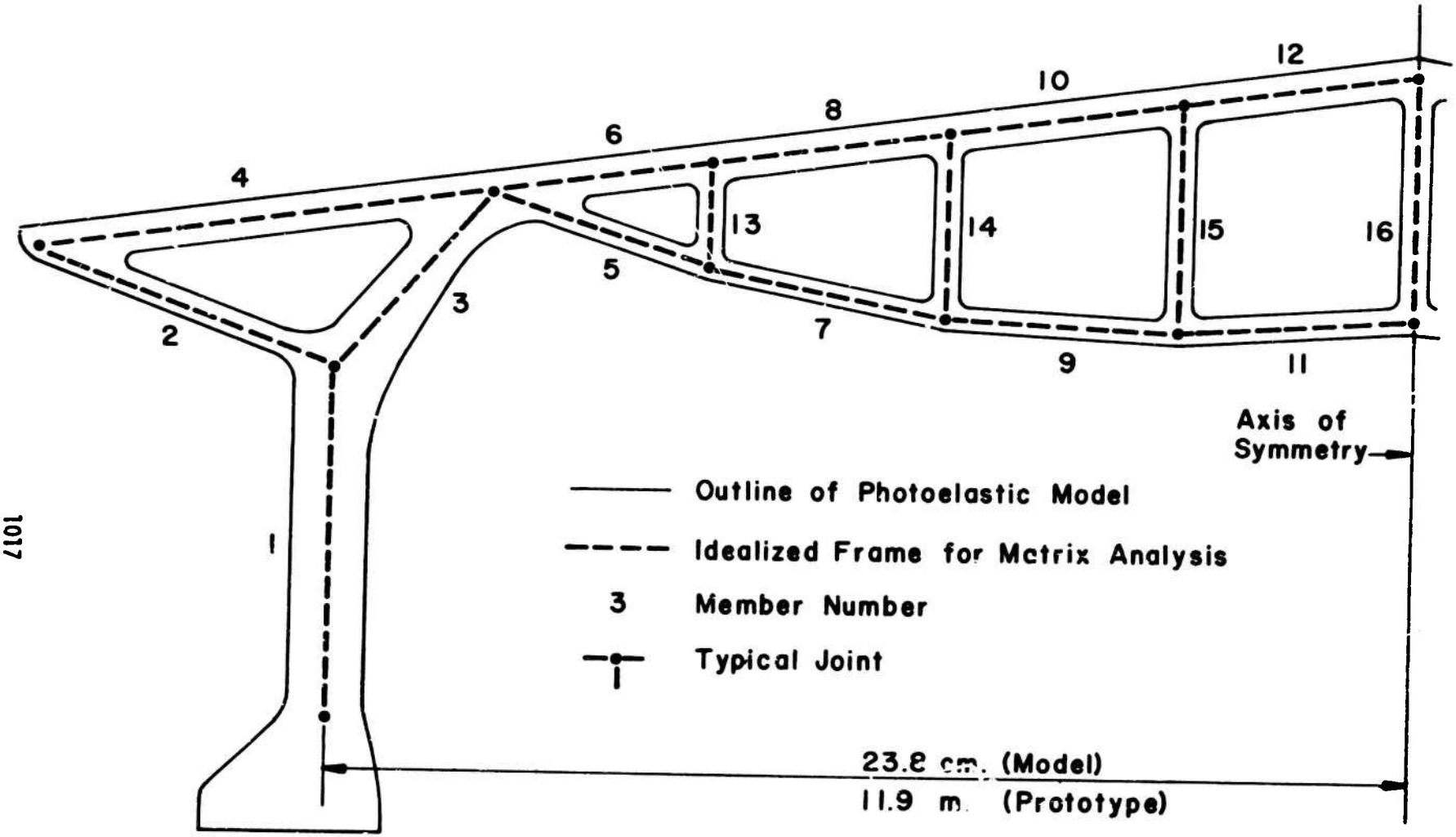


FIGURE 7. ELEVATION OF MAGAZZINI GENERALI 'SHED FRAME

TABLE 2. COMPARISONS OF ANALYSES MAGAZZINI

GENERALI SHED FRAME

Member Type	Member Number (Fig. 7)	Axial Force (kips)		Bending Moment (k-in)	
		STRUDL	Model	STRUDL	Model
Supporting Frame	1	-70	-69	750	1260
	2	-25	-24	16	21
	3	-73	-101	840	900
	4	+24	+18	75	160
Bottom Chord	7	+78	+88	24	21
	9	+78	+94	8	25
	11	+78	+88	7	12
Top Chord	8	-93	-113	98	150
	10	-96	-123	170	160
	12	-95	-102	150	150
Web Verticals	14	-12	-17	91	60
	15	-11	-21	17	5
	16	+12	+24	0	0

Notes: See Figure 7 for member numbers.

Loading is dead and snow loads only.

STRUDL analysis includes member end sizes and transverse shear deformation.

Bending moments compared are the highest quarter-point moments.

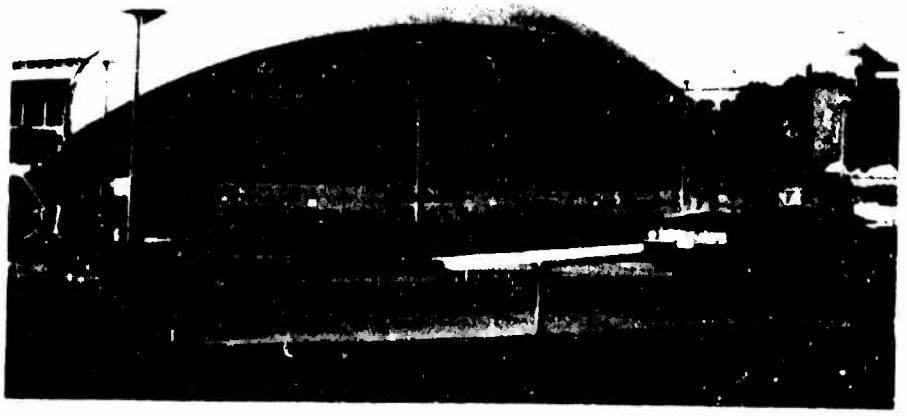


FIGURE 8. SAARINEN'S KRESGE AUDITORIUM, M. I. T.

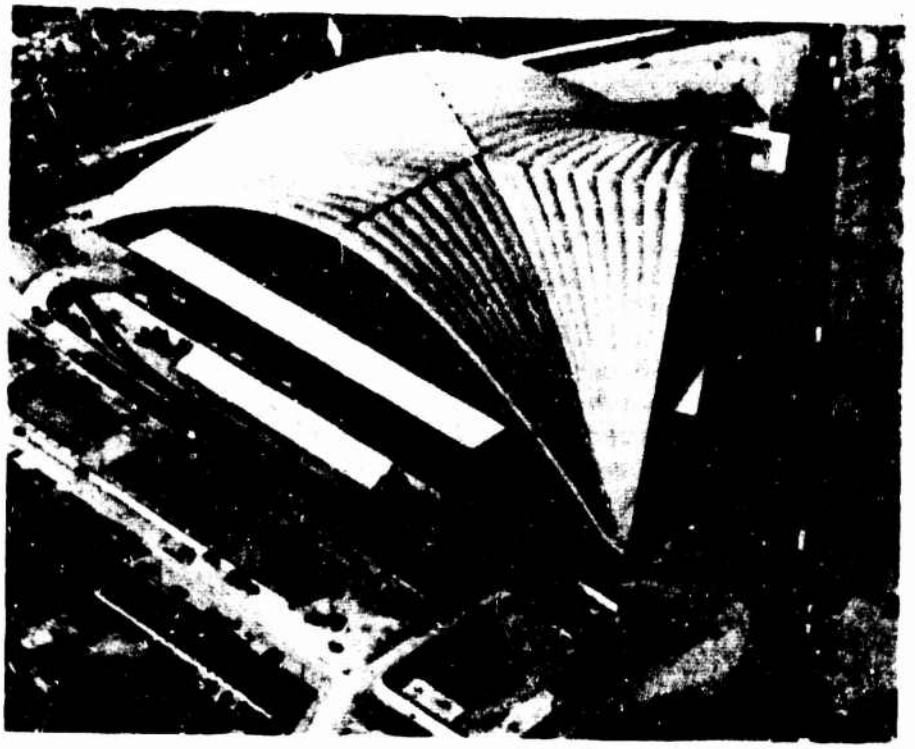


FIGURE 9. CENTRE NATIONALE DES INDUSTRIES ET DES TECHNIQUES (CNIT), PARIS

and despite aesthetic objections by Saarinen, mullions were added to the window walls to support the edge. Despite these and other efforts to strengthen and support the shell, the performance of the structure has not been satisfactory. Cracking and leaking have occurred, especially near the abutments, and the flexing of the shell under varying thermal loads has necessitated the refurbishment of the roofing at a cost of at least $250,000(15).

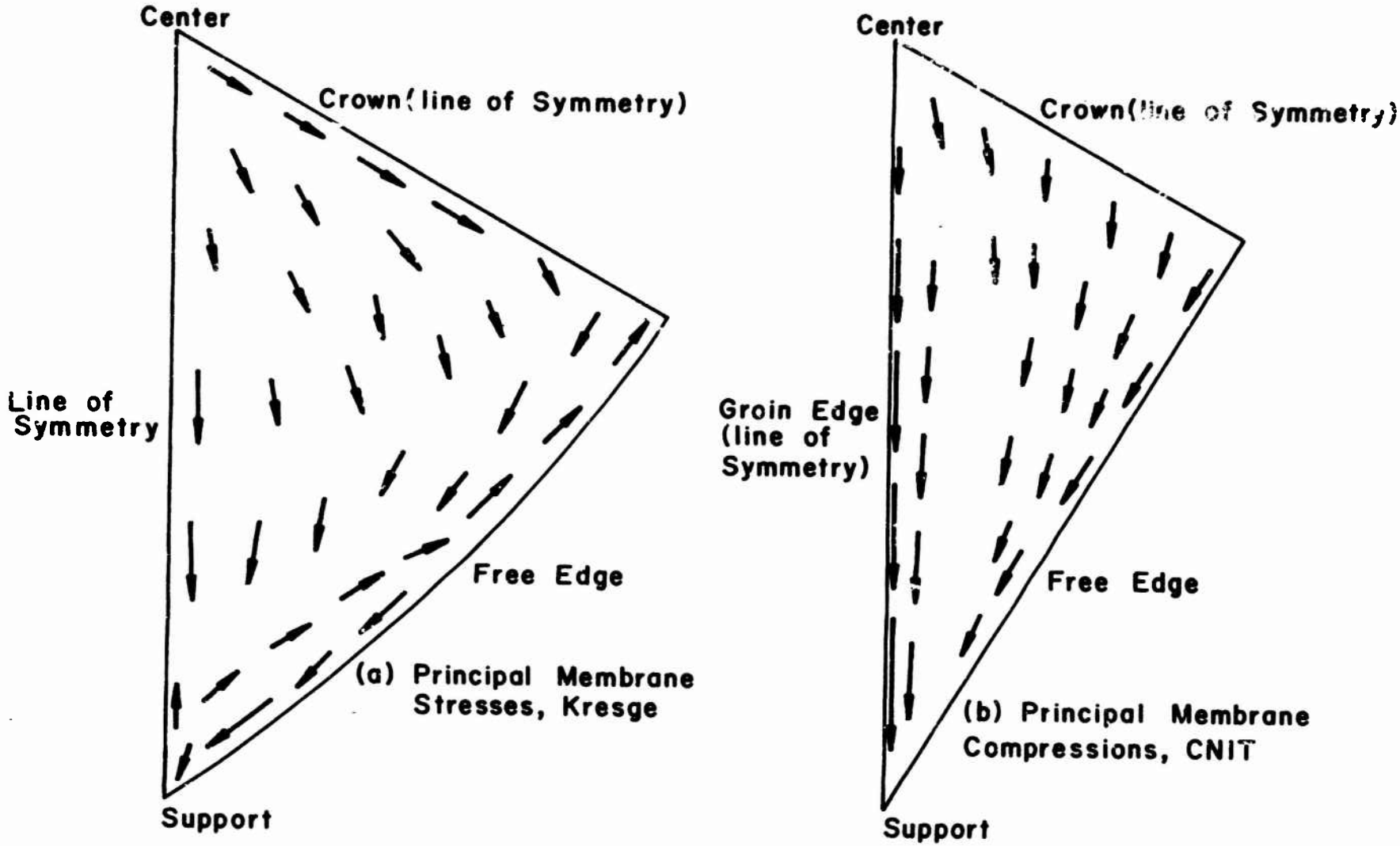
The second phase of the investigation was a comparison of Kresge with other large, point-supported shell structures that were of similar concept and function. This comparison involved not only aesthetics but also a qualitative structural evaluation. One of the buildings considered was the Centre Nationale des Industries et des Techniques (CNIT) in Paris, Figure 9. Several architects and engineers participated in the design of this structure(16); Pier Luigi Nervi, Eugene Freyssinet, and Nicolas Esquillan were the well-known engineers who were successively involved. The CNIT consists of three intersecting cylindrical segments and, like Kresge, is supported at three points. However, the building is much larger than Kresge (675 feet between abutments, 152-foot rise) and for stability considerations is a double shell.

The final aspect of the study was a numerical structural analysis of Kresge to provide a more detailed focus on the deflection and stress problems at the edges. In addition, a structure of CNIT's configuration and Kresge's span was also analyzed for contrast. The finite element method was used(12). The analysis confirmed that only the central portion of Kresge acts truly as a shell; in essence, the central spherical dome is supported by the edge arches, and the in-surface and arch bending are responsible for the cracking and excessive deflection along the edges. Some of the numerical analysis results of Kresge and CNIT are qualitatively compared in Figure 10. The upward flow of forces in Kresge represent tensions arising from the in-surface bending behavior necessary to provide adequate support for the central portion of the dome, which acts as a membrane. In addition, the edge deflections of the Kresge-sized CNIT structure were found to be one and a half orders of magnitude less than those of Kresge without mullions.

This case study provides a modern example for architectural theorists. More significant, however, is the fact that it combines criticism, one of the main characteristics of the humanistic approach, with engineering. Since architectural design from purely aesthetic or decorative considerations may have unacceptable social, functional, and monetary costs, engineers must become critics and participate in non-technical value judgments.

Concluding Remarks

The Humanistic Studies in Engineering program, as exemplified by the five structural engineering studies discussed above, is designed to enhance engineering education and ultimately professional practice. A



1021

FIGURE 10. SCHEMATIC, PLAN-VIEW COMPARISON OF THE FLOW OF MAJOR PRINCIPAL MEMBRANE FORCES IN ONE-SIXTH SEGMENTS OF KRESGE AND CNIT SHELLS (after Reference 15)

humanistic emphasis which is now generally ignored in engineering courses and texts is being integrated into the curriculum. In addition structural engineering is contributing to research in the humanities, including criticism.

Acknowledgments

The author gratefully acknowledges the efforts of various students in the Department of Civil and Geological Engineering. John V. Santosuosso, Visiting Fellow, conducted the Maria Pia Bridge study. John Matheson '71 did the senior thesis on the Toronto City Hall. Dr. Klaus Sachs, Ph.D.'71, and Kevin O'Neill, graduate student, have contributed to the investigations of Gothic vaults. James K. Chiu '71 performed the study on the Maillart frame, H. R. Werner, Jr., '70, M.S.E. '71, did the critical analysis of Kresge Auditorium. The author is also indebted to his colleagues David P. Billington and Robert Mark for their advice in preparing this paper. The Princeton Program in Humanistic Studies in Engineering is supported by grants from the National Endowment for the Humanities, the Ford Foundation, and the Rockefeller Foundation.

References

1. Mesthene, E. G., "Technology and Humanistic Values," Computers and the Humanities, Vol. 4, No. 1, September 1969.

2. Billington, D. P., and R. Mark, "Humanities in Civil Engineering," Journal of Engineering Education, Vol. 59, No. 9, May 1969.

3. Billington, D. P., "Engineering Education and the Origin of Modern Structures,"Civil Engineering, Vol. 39, No. 1, January 1969.

4. Sachs, K., "A Finite Element Analysis of a Doubly Curved Gothic Vault," Ph.D. Dissertation, Department of Civil and Geological Engineering, Princeton University, Princeton, New Jersey, June 1971.

5. Le Corbusier, Toward a New Architecture, Frederick A. Prager, New York, 1927.

6. Giedion, S., Space, Time and Architecture, Harvard University Press, Cambridge, Fourth Edition, 1963.

7. Seyrig, T., "Le Pont sur le Douro," Mémoires et Compte Rendu des Travaux de la Société des Ingénieurs Civils, September and October 1878. An English translation of the major portion of this paper appears in The Engineer, Vols. 45 and 46, 1878, and Vol. 47, 1879.

8. Matheson, J., "Aspects," B.S.E. Thesis, Department of Civil and Geological Engineering, Princeton University, Princeton, New Jersey, April 1971.

9. Roy, H. E. H., "Toronto City Hall and Civic Square," ACI Journal, Vol. 62, No. 12, December 1965.

10. Mark, R., "Gothic Cathedrals and Structural Rationalism," Transactions of the New York Academy of Sciences, to appear.

11. Viollet-le-Duc, E. E., Entretiens sur l'Architecture, Paris, 1863, and 1872. Translated into English, by H. Van Brunt, Discourses on Architecture, J. R. Osgood and Co., Boston, 1875-1881.

12. Johnson, C. P. and P. G. Smith, "A Computer Program for the Analysis of Thin Shells," Structural Engineering Laboratory Report No. SESM 69-5, University of California, Berkeley, January 1969.

13. Bill, M., Rober Maillart: Bridges and Constructions, Frederick A. Prager, New York, Third Edition, 1969.

14. Billington, D. P., "Meaning in Maillart," VIA 2, the Student Publication of Fine Arts, University of Pennsylvania, Philadelphia, to appear.

15. Werner, H. R., Jr., "Kresge Auditorium: A Numerical Analysis and Criticism," M. S. Thesis, Department of Civil and Geological Engineering, Princeton University, Princeton, New Jersey, August 1971.

16. Esquillon, N., "The Shell Vault of the Exposition Palace, Paris," Proc., ASCE, Journal Structural Div., Vol. 86, No. ST1, January 1960.

THREE-DIMENSIONAL FINITE ELEMENT MODEL FOR LAKE CIRCULATION

John W. Leonard*, Illinois Institute of Technology

David Melfi**, State University of New York at Buffalo

A solution technique is presented for the three-dimensional (3-D) circulation of water in shallow lake basins. The basis for solution is a finite element analogue to the shallow lake equations due to Eckman. Galerkin's method provides the minimization functionals for the derivation of rectangular parallelopiped element equations. Since the matrix equations include a nonlinear boundary condition linking the free-surface elevation to the interior elements, the Newton-Raphson Method is suggested as a solution method for the governing matrix equations.

* Illinois Institute of Technology

** State University of New York at Buffalo

INTRODUCTION

To date, no complete understanding of the complicated phenomenon involved in the circulation of large lakes has been obtained. In order to gain some insight into this problem, it is necessary to simplify the dynamic equations of motion. It is then possible to obtain a crude prediction of the response of the lake to various factors, thereby developing in a systematic way a better understanding of a more complicated phenomenon of interest. However, analytical solutions to even these simpler equations are rare and often crude or unduly restrictive. It is the purpose of this paper to present a formulation of a three-dimensional finite element model for the circulation in shallow lakes, e.g. Lake Erie.

Simplified dynamical equations of motion describing massive water circulations were obtained by Ekman (1)* in 1905. Analytical solutions for idealized rectangular basins have been derived by Lai and Rumer (2) based on this theory. Cheng (3,4) has presented a finite element analogue for irregular basins. However, in those works depth averaging techniques were used and the problem was formulated in terms of the average stream function. This eliminated the essential nonlinearity involved in the free surface elevation and reduced the problem to that of a two-dimensional basin. Such a formulation has no meaning near the boundaries of the basin where spurious flow across the boundaries occur. In addition, any numerical solution based on derived variables such as the stream function implies less accurate results for the real variables, i.e., velocities, because of the necessity to differentiate numerically to obtain them. For these reasons a solution technique for the three

dimensional problem was sought.

Matrix equations based on the simplified three-dimensional Navier-Stokes, continuity, and boundary equations due to Ekman (1) are obtained by the use of a combination of the methods of finite elements (5) and of weighted residuals (6). The specific method of weighted residuals chosen was Galerkins' method (7) in which the approximating function itself is taken is the weighting function.

In the succeeding sections, the method of weighted residuals as used in a finite element context, will be developed and applied to the governing differential equations for massive lake circulation. A finite element analogue to these equations for an irregular lake basin using rectangular parallelopiped elements will then be presented. Since the resulting algebraic equations are nonlinear in nature, an iterative convergence method, the Newton-Raphson method, is used to develop a solution technique for these equations. An outline of the solution process is given along with a selection technique for the initial guess required for the Newton-Raphson method.

GALERKIN, FINITE ELEMENT METHOD

The method of weighted residuals (7) consists of assuming a trial solution to the governing equations in the form of a finite series of known functions of position multiplied by undetermined parameters, and requiring that these parameters be such that the field equations and boundary conditions be satisfied in some approximate sense. The manner in which the governing equations are satisfied determines the particular subclass of the method of weighted residuals used.

The subclass chosen for the work described herein is Galerkin's method which consists of using the approximating functions as the weighting functions. Consider the set of differential equations

$$D_i(u_j) = 0 \quad \bar{x} \in V \quad i,j = 1,2 \ldots M \tag{1}$$

with associated boundary conditions

$$u_j(\bar{x}) = \bar{u}_j \quad \bar{x} \in S_u \tag{2a}$$

$$L(u_j) = 0 \quad \bar{x} \in S_L \tag{2b}$$

where $L(\cdot)$ and $D_i(\cdot)$ denote differential operators involving spatial derivatives of the dependent variables u_j, and where V is the total domain of interest bounded by $S = S_u + S_L$.

If a trial solution of the form

$$u_j^* = \sum_{k=1}^{N} c_{kj} \phi_k (\bar{x}) \tag{3}$$

is chosen, where $\phi(\bar{x})$ are specified functions of $\bar{x}$ which satisfy the boundary conditions on S_u, then the residuals, or errors, in the differntial equations and boundary conditions on S_L can be constructed as follows:

$$R_i(c_{kj}, \phi_k, \bar{x} \in V) = D_i(u_j^*) \tag{4a}$$

$$R_L(c_{kj}, \phi_k, \bar{x} \in S_L) = L(u_j^*) \tag{4b}$$

If u_j^* had been the true solution, the residuals would be identically zero. Therefore, the best choice of the c_{kj}'s for a given set of functions ϕ_k is that which makes R_o and R_i minimum. Galerkin's method consists of making the residuals orthogonal to each of the approximating functions ϕ_k,

$$\int_V \phi_k R_i \, dV = 0 \tag{5a}$$

$$\int_{S_L} \phi_k R_L \, dS = 0 \tag{5b}$$

which leads to a set of algebraic equations for c_{kj}.

The rationale behind this criterion is as follows (7): the ϕ_k's are members of a complete set (convergence) and R_L and R_i are piecewise continuous functions; a fundamental property of a complete set is that each member of the set (ϕ_k) is orthogonal to a continuous function (R_L and R_i) only if that function is identically zero; therefore, if a finite number of members of a complete set are used as approximating functions and each member is made orthogonal to the residuals, then in the limit $N \to \infty$ the residuals are identically zero.

Galerkin's method, per se, cannot be easily applied to problems with complicated domains in that it is not feasible to derive approximating functions applicable over the total region and which also satisfy all boundary conditions of the form $u(\bar{x}) = \bar{u}$, $\bar{x} \in S_u$. However, the use of Galerkin's method within the context of a finite element scheme allows one to use simple approximating functions within small regions and to replace boundary conditions with simple interface continuity conditions for interior elements by expressing the approximating solutions in terms of nodal values of the basic unknowns rather than in terms of arbitrary parameters c_{kj}.

It is possible if simple element formulations are used that not all of the interface continuity conditions are satisfied. Because of the approximate nature of the finite element method, and because of the possibility of successive refinements in the numerical modeling of the

region, it has been shown in many instances that acceptable results are still obtained despite the lack of satisfaction of all of the interface continuity conditions. Of course, it is possible to derive more complicated element formulations which satisfy a greater number of these continuity conditions through definition of additional element nodes or of additional nodal variables, e.g., nodal derivatives of the variables.

DERIVATION OF LAKE EQUATIONS

The governing differential equations for flow in a large lake are the Navier-Stokes equations in the three Cartesian directions x_i (a comma denotes differentiation with respect to the succeeding subscripts)

$$\rho \frac{Du_i}{Dt} - \sigma_{ij,j} - \rho F_i = 0 \qquad i = 1,2,3 \tag{6a}$$

the continuity equation

$$u_{j,j} = 0 \tag{6b}$$

and the constitutive relation between stresses σ_{ij} and the velocity gradients

$$\sigma_{ij} = - p \delta_{ij} + \epsilon\rho \, (u_{i,j} + u_{j,i}) \tag{6c}$$

where $\frac{Df}{Dt}$ = convective derivative = $\frac{\partial f}{\partial t} + u_j f_{,j}$

ρ = mass density, F_i, u_i = body force per unit mass and velocity components respectively in the x_i direction, p = hydrostatic pressure, and δ_{ij} = Kronecker delta and ϵ = vertical turbulent momentum transport coefficient (eddy viscosity). The use of the eddy viscosity as a crude measure of turbulence effects is justified only by the lack of a more accurate measure.

The boundary conditions for these field equations are

$$u_i = 0 \quad \text{on} \quad S_s \tag{7a}$$

$$u_j n_j = 0 \quad \text{on} \quad S_f \tag{7b}$$

$$\sigma_{ij} n_j = \tau_i \quad \text{on} \quad S_f \tag{7c}$$

$$\rho u_j n_j = q \quad \text{on} \quad S_q \tag{7d}$$

where S_s = solid boundary, S_f = free surface, S_q = inlet/outlet region, n_j = components of outward unit normal to the total boundary surface, $S = S_s + S_f + S_q$, q = flow rate per unit area of S_q, and τ_i = applied tractions on S_f.

The Rossby number, the ratio of inertial to Coriolis forces is small for currents observed in the Great Lakes. For this reason, the convective terms in Eq. 6a can be neglected. (8) Also, neglecting the horizontal components of the Coriolis force due to vertical motion, and assuming that the Coriolis parameter $f = 2\omega \sin\Phi$ is constant (ω = angular speed of rotation of the earth and Φ = mean latitude of the lake), one can write the Navier-Stokes equation as

$$\frac{1}{\rho}\sigma_{ij,j} + u_j e_{ij} f = 0 \qquad i = 1,2 \tag{8a}$$

$$\frac{1}{\rho} p_{,3} + \epsilon u_{3,jj} + \epsilon_{,3}(u_{3,j} + u_{j,3}) + g = 0 \tag{8b}$$

where e_{ij} = permutation tensor, and g = acceleration of gravity. If it is now assumed (1) vertical momentum transfer is negligible compared to the gravitational force, Eq. (8b) reduces to that of the assumption of hydrostatic pressure distribution

$$p = \rho g (\eta - x_3) \tag{8c}$$

where $\eta = \eta(x_1, x_2)$ is the free surface elevation above the datum $x_3 = 0$ on the mean surface of the lake.

In summary, the governing equations for massive circulation in large lakes are

$$\frac{1}{\rho} \sigma_{ij,j} + f\, e_{ij}\, u_j = 0 \qquad i = 1,2 \tag{9a}$$

$$u_{j,j} = 0 \tag{9b}$$

$$\frac{1}{\rho} \sigma_{ij} = - g(\eta - x_3) + \epsilon\, (u_{i,j} + u_{j,i}) \tag{9c}$$

along with the boundary conditions, Eqs. (7), of which Eq. 7b is of special interest. In terms of the free surface elevation η, Eq. 7b can be written as

which introduces a nonlinearity into the system of field equations.

An approximation to the solution of Eqs. (9) is taken in the form

$$u_i^* = \sum_k \phi_k C_{ki} \tag{10a}$$

$$\eta^* = \sum_k \phi_{ko} C_{ko} \tag{10b}$$

where ϕ_k, ϕ_{ko} are assumed functions of position, and C_{ki}, C_{ko} are unknown parameters. Equations (9) become

$$\frac{1}{\rho} \sigma^*_{ij,j} + f\, e_{ij}\, u^*_j = R_{1i} \qquad i = 1,2 \tag{11a}$$

$$u^*_{j,j} = R_{23} \tag{11b}$$

$$\frac{1}{\rho} \sigma^*_{ij} = - g(\eta^* - x_3)\delta_{ij} + \epsilon(u^*_{i,j} + u^*_{j,i}) \tag{11c}$$

$$u_1^* \eta^*_{,1} + u_2^* \eta^*_{,2} - u_3^* = R_{oo} \text{ on } x_3 = 0 \tag{11d}$$

where R_{ij} denotes the errors, or residuals, introduced by the approximations. Galerkin's criteria for this system of equations are

$$\int_v \phi_k, R_{ij}\, dV = 0 \qquad i = 1,2;\ j = 1,3 \tag{12a}$$

$$\int_{S_f} \phi_{ko}\, R_{oo}\, dS = 0 \tag{12b}$$

In order to recover the boundary conditions, Eqs. (7), from Eqs. (12), it is necessary to integrate Eqs. (12) by parts and to apply the divergence theorem (6). This yields

$$\int_v (\phi_{k,j}\, \sigma^*_{ij}/\rho - \phi_k\, f\, e_{ij}\, u_j^*)\, dv = \oint_S \phi_k\, \sigma^*_{ij}\, \eta_j\, dS \tag{13a}$$

$$\int_v \phi_{k,j}\, u^*_{,j}\, dv = \oint_S \phi_k\, u_j^*\, \eta_j\, dS \tag{13b}$$

$$\int_{S_f} \phi_{ko}\, (u_1^* \eta^*_{,i} + u_2^* \eta^*_{,2} - u_3^*) = 0 \tag{13c}$$

If Eqs. (7) and (11c) are substituted into the right hand sides of Eqs. (13), there results

$$\int_v [e\, \phi_{k,j}(u^*_{i,j} + u^*_{j,i}) - g\, \phi_{k,i}\, \eta^* - f\phi_k\, e_{jj}\, u_j^*]\, dV$$

$$= -\int_v g\, x_3\, \phi_{k,i}\, dV + \int_{S_f} \phi_k\, \tau_i/\rho\, dS + \int_{S_s+S_q} \phi_k \sigma^*_{ij} \eta_j dS \tag{14a}$$

$$\int_v \phi_{k,j}\, u_j^*\, dV = \int_{S_q} \phi_k\, q/\rho\, dS \tag{14b}$$

$$\int_{S_f} \phi_{ko}\, (u_1^* \eta^*_{,1} + u_2^* \eta^*_{,2} - u_3^*)\, dS = 0 \tag{14c}$$

where u_i^* and η^* are given by Eqs. (10) in terms of the unknown parameters C_{ki} and C_{ko}. Equations (14) constitutes a set of algebraic equations for those parameters and for the tractions $\sigma_{ij}^* \eta_j$ on the side and bottom boundaries of the lake basin.

It should be noted that the tractions $\sigma_{ij}^* \eta_j$ on $S_s + S_q$ cannot be determined from Eq. (11c) since that equation provides an inadequate description for the complex state of stress in the boundary region. Instead, as will be seen subsequently, the boundary tractions are determined directly from the finite element analogue to Eq. (14a) once the interior distribution of velocities has been determined. An analogy can be drawn to structural mechanics problems in which unknown tractions are specified at points of prescribed displacements.

FINITE ELEMENT ANALOGUE

Consider the lake to be subdivided into a finite number of rectangular parallelpiped elements interconnected at prescribed nodal locations (see Fig. 1). A two-dimensional rectangular grid of elements is superimposed at the datum $x_3 = 0$. This grid is to account for the free surface integral, Eq. 14c, and for the surface elevation η present in Eqs. (14a,b). In other words, each node of each volume element in the lake is considered to be connected through the η term with the corresponding node in the surface grid directly above the interior node.

The finite element analog to Eqs. (10) and (13) is obtained by identifying the unknown parameters C_{ki} and C_{ko} with nodal values of u_i^* and η^*, respectively. Thus

$$u_i^* = \frac{1}{b_3} [\phi][A] \quad U_i \tag{15a}$$

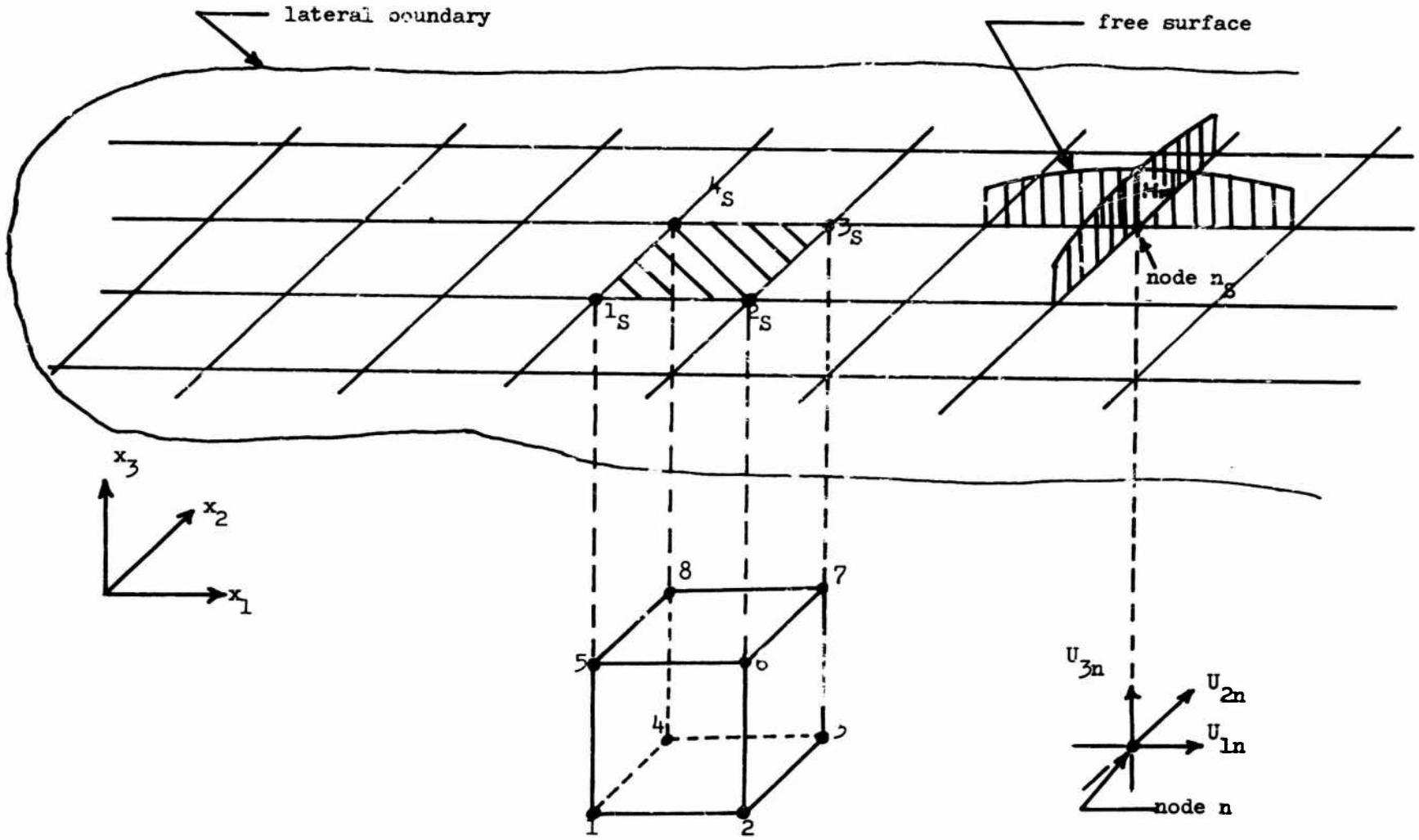


Figure 1. Typical Element in Lake Basin

$$\eta^* = \frac{1}{b_1 b_2} [\psi_o][A_o] \quad H_i \tag{15b}$$

where b_i = length of element in x_i direction, ψ = approximating polynomial over the volume element, ψ_o = approximating polynomial over the surface element, U_i = nodal values of u_i, H = nodal values of η, A and A_o = matrices of coefficients (dependent on nodal coordinates and choice of ψ) guaranteeing continuity of u_i and η across element interfaces. The products ψ A/b_3 and ψ_o $A_o/b_1 b_2$ correspond to the weighting functions ϕ and ϕ_o in Eqs. (10).

The matrix form of Eqs. (12) are

$$\sum_{e_v} \frac{1}{b_3} \int_{V_e} A^T \psi^T R_i \, dV_e = 0 \tag{16a}$$

$$\sum_{e_S} \frac{1}{b_1 b_2} \int_{S_e} A_o{}^T \psi_o{}^T R_o \, dS_e = 0 \tag{16b}$$

where e_v denotes the total number of volume elements, and e_S denotes the total number of surface elements. Therefore, after substitution of Eqs. (15) into Eqs. (16) and (14) and after regrouping of terms, the matrix form of Eqs. (14) are found to be

$$\sum_{e_v} \begin{bmatrix} D_{11} & D_{12} & D_{13} & D_{14} \\ D_{21} & D_{22} & D_{23} & D_{24} \\ D_{31} & D_{32} & D_{33} & D_{34} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ H \end{bmatrix} - \begin{bmatrix} \dot{F}_1 \\ F_2 \\ Q \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{17a}$$

$$\sum_{e_S} ([D_{41} \quad D_{42} \quad D_{43} \quad D_{44}]) \begin{bmatrix} U_{1S} \\ U_{2S} \\ U_{3S} \\ H \end{bmatrix} = 0 \tag{17b}$$

where (i = 1,2 and j = 1,3)

$$[D_{ij}] = \frac{1}{b_3^2}\, e\, [A]^T \int_{V_e} (\delta_{ij}[\psi_{,k}]^T[\psi_{,k}] + [\psi_{,j}]^T[\psi_{,i}] - \frac{fe}{e}{}_{ij}[\psi]^T[\psi])dV_e[A] \quad (18a)$$

$$[D_{3j}] = \frac{1}{b_3^2}\,[A]^T \int_{V_e} [\psi_{,j}]^T[\psi]dV_e[A] \quad (18b)$$

$$[D_{i4}] = -\frac{g[A]^T}{b_1 b_2 b_3} \int_{V_e} [\psi_{,i}]^T[\psi_o]dV_e[A_o] \quad (18c)$$

$$[D^{4i}] = [0] \quad (18d)$$

$$[D_{34}] = [0] \quad (18e)$$

$$[D_{43}] = \frac{1}{b_1^2 b_1^2}\,[A_o]^T \int_{S_e} [\psi_o]^T[\psi_o]dS_e[A_o] \quad (18f)$$

$$[D_{44}] = \frac{-1}{b_1^3 b_2^3}\,[A_o]^T \int_{S_e} [\psi_o]^T[\psi_o][A_o](\,U_{1S}\,[\psi_{o,1}] + U_{2S}\,[\psi_{o,2}])dS_e[A_o] \quad (18g)$$

$$F_i = \frac{1}{b_3}\,[A]^T \int_{S_e} [\psi]^T \frac{\tau_i^e}{\rho}\, dS_e - \frac{g}{b_3}\,[A]^T \int_{V_e} x_3\,[\psi_{,i}]^T dV_e \quad (18h)$$

$$Q = \sum_{k=1}^{\text{no. faces}} \frac{[A]^T}{b_3} \int_{S_k} [\psi]^T \frac{q_k^e}{\rho}\, dS_e \quad (18i)$$

where in Eq. (18h) $\tau_i = 0$ if the element in question is not at the surface of the lake, or τ_i = wind stress on the upper face of an element at the surface of the lake. In Eq. (18i), the summation extends over all faces of the element and $q_k = 0$ if the kth face is not an inlet/outlet region, or q_k = flow distribution across the kth face of the element at an inlet/outlet region.

Before the element matrices given above can be evaluated, particular choices of $[\psi]$ and $[A]$ must be made. In this instance a trilinear polynomial was assumed in the form

$$\Psi = [\psi_o , x_3 \psi_o] \tag{19a}$$

$$\psi_o = [1, x_1, x_1 x_2, x_2] \tag{19b}$$

and a node numbering scheme as shown in Fig. 1 was assumed. Therefore,

$$A = \frac{1}{b_1 b_2} \left[\begin{array}{c|c} x_3^+ A_o & -x_3^- A_o \\ \hline -A_o & A_o \end{array}\right] \tag{20a}$$

$$A_o = b_1 b_2 \begin{bmatrix} 1 & x_1^- & x_1^- x_2^- & x_2^- \\ & x_1^+ & x_1^+ x_2^- & x_2^- \\ 1 & x_1^+ & x_1^+ x_2^+ & x_2^+ \\ 1 & x_1^- & x_1^- x_2^+ & x_2^+ \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} x_1^+ x_2^+ & -x_1^- x_2^+ & x_1^- x_2^- & -x_1^+ x_2^- \\ -x_2^+ & x_2^+ & -x_2^- & x_2^- \\ 1 & -1 & 1 & -1 \\ -x_1^+ & x_1^- & -x_1^- & x_1^+ \end{bmatrix} \tag{20b}$$

where $x_i^{\pm}$ = x_i-coordinate ($\pm$) face. Substituting Eqs. (9) and (20) into Eqs. (18) and integrating, one obtains

$$D_{ij} = \left[\begin{array}{c|c} 2[M_{kk}\delta_{ij} + M_{ji} - (\frac{f}{e}e_{ij} - \frac{3}{b_3^2}\delta_{ij})M_{oo}] & [M_{kk}\delta_{ij} + M_{ji} - (\frac{f}{e}e_{ij} + \frac{6}{b_3^2}\delta_{ij})M_{oo}] \\ \hline [M_{kk}\delta_{ij} + M_{ji} - (\frac{f}{e}e_{ij} + \frac{6}{b_3^2}\delta_{ij})M_{oo}] & 2[M_{kk}\delta_{ij} + M_{ji} - (\frac{f}{e}e_{ij} - \frac{3}{b_3^2}\delta_{ij})M_{oo}] \end{array}\right] \tag{21a}$$

$$i,j = 1,2$$

$$D_{i3} = \frac{3}{b_3} \left[\begin{array}{c|c} -M_{oi} & -M_{oi} \\ \hline M_{oi} & M_{oi} \end{array}\right] \qquad i = 1,2 \tag{21b}$$

$$F_i = \frac{6b_1b_2}{b_3^2} \frac{\bar{\tau}_i}{\epsilon\rho} \begin{bmatrix} x_3^+ & M_o \\ -x_3^- & M_o \end{bmatrix} \tag{21c}$$

$$D_3^i = \left[\begin{array}{c|c} 2M_{io} & M_{io} \\ \hline M_{io} & 2M_{io} \end{array}\right] \qquad i = 1,2 \tag{21d}$$

$$D_{33} = \frac{3}{b_3} \left[\begin{array}{c|c} -M_{oo} & -M_{oo} \\ \hline M_{oo} & M_{oo} \end{array}\right] \tag{21e}$$

$$D_{i4} = \frac{-3g}{\epsilon} \begin{bmatrix} M_{io} \\ M_{io} \end{bmatrix} \qquad i = 1,2 \tag{21f}$$

$$D_{43} = [M_{oo}] \tag{21g}$$

$$D_{44} = \frac{1}{b_1b_2}\left[\begin{array}{c|c|c} [\,0\,] & [M_{oo}]\ U_{15} & [M_2]\ U_{15} + [M_1]\ U_{25} \end{array}\right.$$
$$\left.\begin{array}{c|c} [M_{oo}]\ U_{25} &][A_o] \end{array}\right] \tag{21h}$$

$$Q = 3b_1b_2/\rho \sum_k \bar{q}_k \left[R_1^{\pm} + R_2^{\pm}\right] \tag{21i}$$

where $\bar{\tau}_i$ and $\bar{q}$ denote average values of stress and flow over the face of the element, and M_{ij} and M_i $(i,j = 0,1,2)$ are given by

$$M_{ij} = \frac{1}{b_1b_2} [A_o]^T \int_{S_e} [\Psi_{o,i}]^T [\Psi_{o,j}] dx_1 dx_2 [A_o] \tag{22a}$$

$$M_o = \frac{1}{b_1b_2} [A_o]^T \int_{S_e} [\Psi_o]^T ds_1 ds_2 \tag{22b}$$

$$M_i = \frac{1}{b_1b_2} [A_o]^T \int_{S_e} x_i [\Psi_o]^T [\Psi_o] ds_1 ds_2 [A_o] \tag{22c}$$

$$R_i^{\pm} = \frac{1}{h_1 b_2} [A_o]^T \int_{x=x_i^{\pm}} [\psi_o]^T \, dx_{3-i} \tag{22d}$$

where x_i^* denotes the x_i-coordinate, x_i^+ or x_i^-, of the flow face, and where $\psi_{o,o} = \psi_o$. The integrals in Eqs. (22) are given in Table 1.

Once Eqs. (17) have been obtained for the typical element, the individual submatrices being defined by Eqs. (21), it is possible to generate a single master matrix equation for the total lake using assembly techniques developed for problems in solid mechanics (5). This has not been done in this instance because of the nonlinear nature of Eq. (17b), i.e., observe that D_{44} given by Eq. (21h) is a function of nodal velocities. Instead, the assemblage is postponed until the governing equations of the solution procedure have been developed.

SOLUTION TECHNIQUE

To generate numerical solutions to the governing nonlinear algebraic equations developed in the preceding section, an iterative solution technique is required. The technique described in this section is a well-known convergence acceleration scheme--the Newton-Raphson method.

In essence, the Newton-Raphson method (9) consists of obtaining the improvement ϵ^r to an initial guess u^r by means of the equation

$$\epsilon^r = -[J(u^r)] \, f(u^r) \tag{23a}$$

where $f(u) = 0$ are the governing nonlinear equations, and

$$J_{ij}(u^r) = \frac{\partial f_i(u^r)}{\partial u_j^r} \tag{23b}$$

$\int \Psi_o^T \Psi_o \, 1/b_1 b_2$			
1	$\bar{x}_1$	$\bar{\bar{x}}_1 x_2$	$\bar{x}_2$
$\bar{x}_1$	$\bar{\bar{x}}_1$	$\bar{\bar{x}}_1 x_2$	$\bar{x}_1 x_2$
$\bar{x}_1 x_2$	$\bar{\bar{x}}_1 x_2$	$\bar{\bar{x}}_1 \bar{\bar{x}}_2$	$\bar{x}_1 \bar{\bar{x}}_2$
$\bar{x}_2$	$\bar{x}_1 x_2$	$\bar{x}_1 \bar{\bar{x}}_2$	$\bar{\bar{x}}_2$

$\int \Psi_o^T \Psi_{o,1}/b_1 b_2$			
0	1	$\bar{x}_2$	0
0	$\bar{x}_1$	$\bar{x}_1 x_2$	0
0	$\bar{x}_1 x_2$	$\bar{x}_1 \bar{\bar{x}}_2$	0
0	$\bar{x}_2$	$\bar{\bar{x}}_2$	0

$\int \Psi_o^T \Psi_{o,2}/b_1 b_2$			
0	0	$\bar{x}_1$	1
0	0	$\bar{\bar{x}}_1$	$\bar{x}_1$
0	0	$\bar{\bar{x}}_1 x_2$	$\bar{x}_1 \bar{\bar{x}}_2$
0	0	$\bar{x}_1 x_2$	$\bar{\bar{x}}_2$

$\int \Psi_{o,1}^T \Psi_{o,1} \, b_1 b_2$			
0	0	0	0
0	1	$\bar{x}_2$	0
0	$\bar{x}_2$	$\bar{\bar{x}}_2$	0
0	0	0	0

$\int \Psi_{o,1}^T \Psi_{o,2}/b_1 b_2$			
0	0	0	0
0	0	$\bar{x}_1$	1
0	0	$\bar{x}_1 x_2$	$\bar{x}_2$
0	0	0	0

$\int \Psi_{,2}^T \Psi_{,2}/b_1 b_2$			
0	0	0	0
0	0	0	0
0	0	$\bar{\bar{x}}_1$	$\bar{\bar{x}}_1$
0	0	$\bar{x}_1$	1

$\int \Psi_o / b_1 b_2$			
1	$\bar{x}_1$	$\bar{x}_1 \bar{x}_2$	x_2

$\int x_1 \Psi_o^T \Psi_o / b_1 b_2$			
$\bar{x}_1$	$\bar{\bar{x}}_1$	$\bar{\bar{x}}_1 x_2$	$\bar{x}_1 x_2$
$\bar{\bar{x}}_1$	$\bar{\bar{x}}_1$	$\bar{\bar{x}}_1 x_2$	$\bar{\bar{x}}_1 x_2$
$\bar{\bar{x}}_1 x_2$	$\bar{\bar{x}}_1 x_2$	$\bar{\bar{x}}_1 x_2$	$\bar{\bar{x}}_1 x_2$
$\bar{x}_1 x_2$	$\bar{\bar{x}}_1 x_2$	$\bar{\bar{x}}_1 x_2$	$\bar{x}_1 \bar{\bar{x}}_2$

$\int x_2 \Psi_o^T \Psi_o / b_1 b_2$			
$\bar{\bar{x}}_2$	$\bar{x}_1 x_2$	$\bar{x}_1 \bar{\bar{x}}_2$	$\bar{\bar{x}}_2$
$\bar{x}_1 x_2$	$\bar{\bar{x}}_1 x_2$	$\bar{\bar{x}}_1 \bar{\bar{x}}_2$	$\bar{\bar{x}}_1 \bar{\bar{x}}_2$
$\bar{x}_1 \bar{\bar{x}}_2$	$\bar{\bar{x}}_1 \bar{\bar{x}}_2$	$\bar{\bar{x}}_1 \bar{\bar{x}}_2$	$\bar{\bar{x}}_1 \bar{\bar{x}}_2$
$\bar{\bar{x}}_2$	$\bar{x}_1 \bar{\bar{x}}_2$	$\bar{x}_1 \bar{\bar{x}}_2$	$\bar{\bar{x}}_2$

$\bar{x}_i = (x_i^+ + x_i^-)/2$

$\bar{\bar{x}}_i = x_i^+ x_i^- + b_i^2/3$

$\bar{\bar{\bar{x}}}_i = \bar{x}_i (x_i^{+2} \; x_i^{-2})/2$

Table 1. Integrals in Eqs. (22) ($[M_{ij}] = [M_{ji}^T]$)

The new approximation is then given by

$$u^{r+1} = u^r + \epsilon^r \tag{23c}$$

Therefore, given an initial guess, u^o, repeated application of Eqs. (23) will yield successively better approximations to the true solution.

The primary tasks in the application of the Newton-Raphson method to the current problem are to generate [J] and to obtain an initial guess in the neighborhood of the true solution. For the typical element the governing equations are given by Eqs. (17). Applying Eq. (23b), one obtains for each element

$$\sum_{e_v} \left(\begin{bmatrix} J_{11} & J_{12} & J_{13} & J_{14} \\ J_{21} & J_{22} & J_{23} & J_{24} \\ J_{31} & J_{32} & J_{33} & J_{34} \end{bmatrix} \begin{bmatrix} \epsilon_{u_1} \\ \epsilon_{u_2} \\ \epsilon_{u_3} \\ \epsilon_H \end{bmatrix} + \right.$$

$$\left. \begin{bmatrix} D_{11} & D_{12} & D_{13} & D_{14} \\ D_{21} & D_{22} & D_{23} & D_{24} \\ D_{31} & D_{32} & D_{33} & D_{34} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2^r \\ u_3^r \\ H^r \end{bmatrix} - \begin{bmatrix} F_1 \\ F_2 \\ Q \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{24a}$$

$$\sum_{e_S} \left(\begin{bmatrix} J_{41} & J_{42} & J_{43} & J_{44} \end{bmatrix} \begin{bmatrix} \epsilon_{u_{1S}} \\ \epsilon_{u_{2S}} \\ \epsilon_{u_{3S}} \\ \epsilon_H \end{bmatrix} + \right.$$

$$+ \quad [D_{41} \quad D_{42} \quad D_{43} \quad D_{44}] \begin{bmatrix} u^r_{1S} \\ u^r_{2S} \\ u^r_{3S} \\ H^r \end{bmatrix}) = 0 \qquad (24b)$$

where

$$J_{ij} = D_{ij} + \Delta J_{ij} \qquad (24c)$$

$$\Delta J_{ij} = 0 \qquad i = 1,2,3;\ j = 1,2,3,4 \qquad (24d)$$

$$\Delta J_{4j} = 0 \qquad j = 3,4 \qquad (24e)$$

$$\Delta J_{4j} = \frac{1}{b_1 b_2}\left[[M_{oo}]\cdot([A_{2\cdot j}]\ H^r) + [M_{3-j}]\cdot([A_3]\ H^r)\right] \qquad j = 1,2 \qquad (24f)$$

In Eq. (24f) A_i = ith row of matrix A given by Eq. (20b) and $[B]\cdot(f)$ denotes multiplication of the matrix B by the scalar f.

If it is assumed for the moment that an initial guess U_i^o, H^o at every node is readily available, then the assemblage and solution scheme could be as follows:

1) Calculate all D_{ij} for element e. (If this element adjoins the free surface calculate all D_{4j} except D_{44}).

2) For a surface element use appropriate subvectors of U_i^o and H^o to calculate D_{44}^o and the elemental vector

$$P^o = \left[\begin{array}{c|c} D_{ij} & D_{i4} \\ \hline D_{4i} & D_{44}^o \end{array}\right] \left[\begin{array}{c} U_i^o \\ \hline H^o \end{array}\right] - \begin{bmatrix} F_i \\ - \\ 0 \end{bmatrix} \qquad (25)$$

3) Add subvectors of P^o to appropriate locations of master matrix P_T^o arranged by nodal groupings of variables.

4) Generate all J^o_{ij} for element e using Eq. (24c) and add submatrices to appropriate locations of master matrix $[J^o_T]$ also arranged by nodal groupings of variables (note that only J^o_{41} and J^o_{42} differ from values stored in submatrices of D and that J_{i4} are associated with the corresponding surface nodes for element e).

5) Account for boundary conditions by setting the forcing function to zero at boundary nodes and by replacing the rows and columns of J^o_S associated with boundary nodes by rows and columns of zeros with an identity matrix in the diagonal position.

6) Solve for the correction vector ϵ_T.

$$\epsilon_T = [J^o_T]^{-1} \quad P^o_T \tag{26}$$

and update the most recent trial solution.

7) If ϵ_T is sufficiently small, output the final solution. If ϵ_T is not sufficiently small, repeat steps 2 through 6.

The problem remains of determining an initial guess, U^o_i and H^o. Since the governing equations are only slightly nonlinear, the equivalent linear solution to Eqs. (17) can be used as an initial guess. The matrix D_{i4} is suppressed and Eq. (17a) yields a system matrix of the form

$$\left[\begin{array}{c|c} D_{ij_T} & 0 \\ \hline D_{4i_T} & D_{44_T} \end{array}\right] \left[\begin{array}{c} U_{i_T} \\ \hline H_T \end{array}\right] = \left[\begin{array}{c} F_{i_T} \\ \hline 0 \end{array}\right] \tag{27}$$

The solution to Eq. (27) can be seen to be

$$U_{i_T} = [D_{ij_T}]^{-1} \quad F_{i_T}$$

$$H = [D_{44}]_T (U_i)^{-1} [D_{4i}]_T U_{i_T}$$

which are then taken as the initial guesses for the solution process outlined above.

The selection procedure for an initial guess can in fact be incorporated easily as a substep between Steps 1 and 2 above. For the first iteration: 1) set U^o and H equal to zero; 2) suppress D_{i4}, D_{4i}, and J_{4i}; 3) set J_{44} and D_{44} equal to identity matrices; 4) solve for ϵ_T and update; 5) insert the proper values of D_{i4}, D_{4i}, D_{i4}, J_{i4} and J_{4i}; 6) set $J_{4i} = D_{i4}$ and resolve for ϵ_T; and finally, 7) rejoin solution process given above at Step 2.

SUMMARY

A finite element model of the massive circulation in shallow lake basins has been formulated based on the method of weighted residuals. Unlike previous analytic and finite element models, the three-dimensional character of the basin was accounted for by using rectangular **parallelo**piped elements with associated surface nodes.

The governing field equations assumed in the derivation of the model are those due to Ekman (1) in which vertical momentum transfer is neglected compared to gravitational forces implying a hydrostatic pressure distribution. Also, inertial forces are neglected compared to the coriolis and surface wind forces.

The field equations and their corresponding finite element analogues are linear. However, the auxiliary equations expressing the free surface boundary conditions is nonlinear in nature. Therefore, an iterative solution technique, the Newton-Raphson method, is the

basis for solution to the algebraic matrix equation. An outline of this method as applied to the problem at hand is given along with a scheme for the automatic generation of an initial trial solution required for the Newton-Raphson method.

Numerical results to verify and demonstrate the formulations presented herein are as yet unavailable. A computer program to obtain such results has been written and is currently being debugged. One difficulty anticipated in verifying the model is that there are no exact, or even approximate, results available for three-dimensional flow in a lake basin. Only depth-averaged two-dimensional flows (with inherent inaccuracies) have been considered. With such models only relative rather than absolute values for the free surface elevation are obtainable. Also, experimental data suitable for comparison purposes are scarce and scattered.

It is feasible to consider various extensions of the finite element model presented herein. Steady-state flow in thermally stratified lakes can be incorporated by allowing different values for the pertinent fluid properties in different elements. Transient flow capabilities can be added to the linearized model by retaining the linear portion of the convective derivative in the Navier-Stokes equation. If the assumption of negligible inertial forces is not used, the fully nonlinear field equations can be modeled by a finite element analogue as described in Reference 10. Such a model would be useful to describe the complicated flow in the regions of inlet or outlet rivers where in fact the Coriolis and surface wind forces are negligible compared to the inertial forces rather than the converse as assumed herein.

REFERENCES

1. Ekman, V. W., "On the Influence of the Earth's Rotation on Ocean Currents," Arkiv. Met. Astr. Fysik., Vol. 2, No. 4, 1905.

2. Lai, R.Y.S., and Rumer, R. R., "Circulation in an Idealized Rectangular Basin," Proc., 13th Conf. Great Lakes Res., 1970, Buffalo, N.Y.

3. Cheng, R. T., "Numerical Investigations of Lake Circulation Around Islands by the Finite Element Method," to appear in the International Journal for Numerical Methods in Engineering.

4. Cheng, R. T., and Tung, C., "Wind Driven Lake Circulations by the Finite Element Method," Proc., 13th Conf. Great Lakes Res., 1970, Buffalo, New York.

5. Przemieniecki, J. S., Theory of Matrix Structural Analysis, McGraw-Hill, N.Y., 1968.

6. Leonard, J.W., and Bramlette, T. T., "Finite Element Solutions to Differential Equations," Journal of the Engng. Mechs. Div., ASCE, Vol. 96, No. EM6, p. 1277, December 1970.

7. Finlayson, B.A. and Scriven, L.E., "The Method of Weighted Residuals--A Review," Appl. Mech. Rev., Vol. 19, p. 735, 1966.

8. Csanady, G.T., "Large Scale Motion in the Great Lakes," J. Geophys. Res., Vol. 72, p. 4151, 1967.

9. Carnahan, B. et al, Applied Numerical Methods, John Wiley and Sons, N.Y., 1969.

10. Leonard, J.W., "Galerkin, Finite Element Formulations for Incompressible Flow," Report No. 9500-920181, Bell Aerospace Company, Buffalo, New York, 1970.

*U.S.Government Printing Office: 1974 — 657-013/53