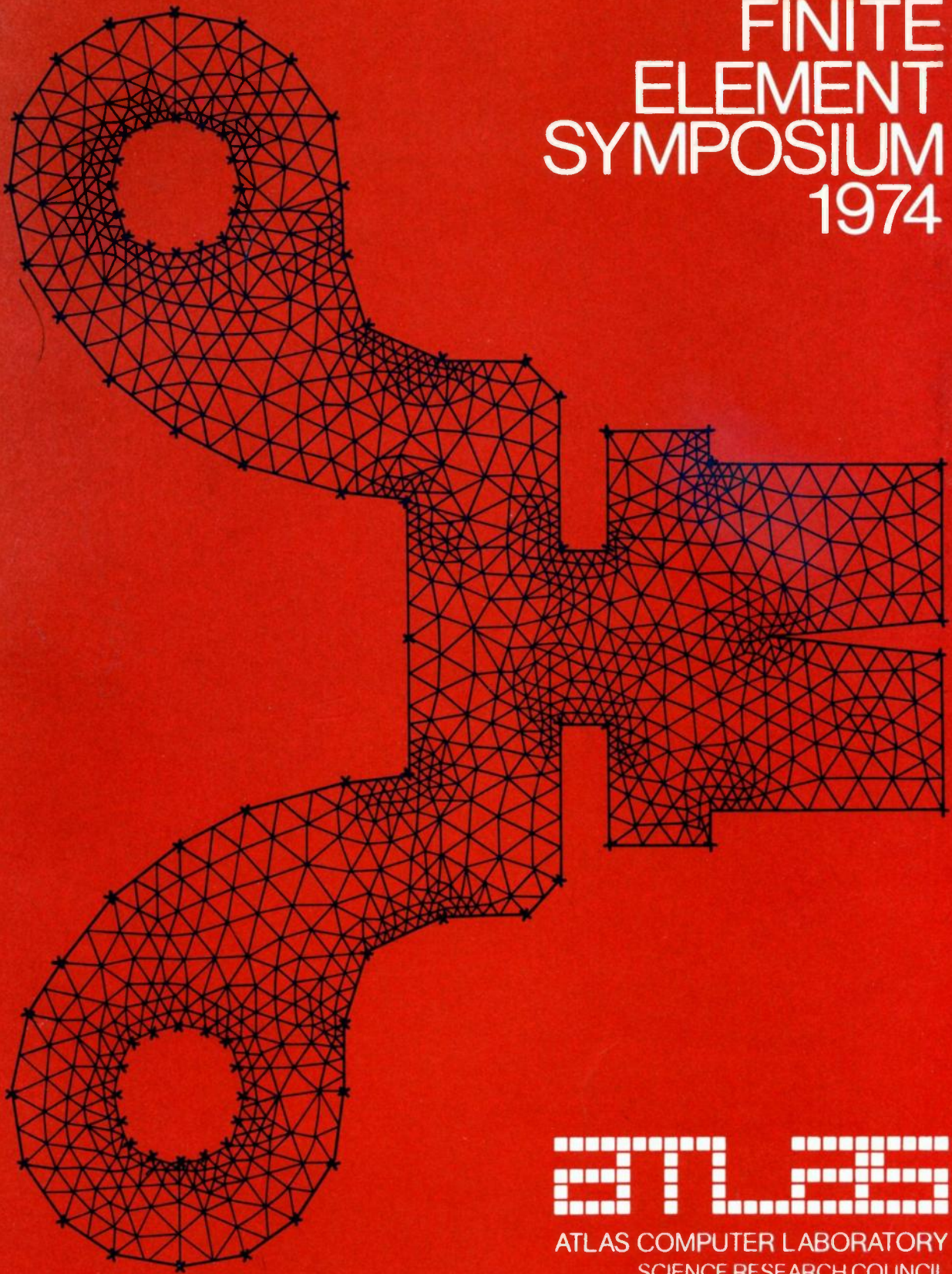# FINITE ELEMENT SYMPOSIUM 1974

**ATLAS**

ATLAS COMPUTER LABORATORY
SCIENCE RESEARCH COUNCIL

PROCEEDINGS
OF THE

# FINITE ELEMENT SYMPOSIUM

26-28 MARCH 1974

EDITED BY

# JEAN E CROW

# FOREWORD

The seeds of this Symposium were sown in the Autumn of 1973, when an increasing
number of requests from Atlas Laboratory's engineering users for a comprehensive
finite element package highlighted the necessity of ACL support for this type of
computation.   It became clear from informal discussions with existing users that
two courses of action would be desirable – firstly, an assessment of the needs of
engineers for finite element programs and numerical support, and secondly, the
bringing together of people from the engineering and numerical fields to exchange
ideas on the application of the Finite Element Method to problems of current and
future interest.

This Symposium was organised as part of an attempt to satisfy both the above criteria.
To the extent that the papers presented gave useful and up-to-date information about
the different aspects of the Method, and that the forty participants were able to
benefit from a cross-fertilisation of ideas between the different disciplines, the
symposium can be regarded as a success.

The Symposium took place on 26, 27 and 28 March 1974 at SRC's Cosener's House, a
pleasant country house overlooking the River Thames at Abingdon.   Fifteen papers
were presented, together with periods of discussion, and there was a most useful
session in which a Panel of eminent engineers, chaired by Professor Rosenbrock of
UMIST, answered questions on SRC policy in the field of engineering computing. The
meeting was also pleased to welcome Professor Edwards, chairman of the Science
Research Council, as guest speaker at the Symposium Dinner.

The papers which were presented have been edited and retyped, with figures redrawn,
and are now published in this volume for the benefit of symposium participants as
well as those who were unable to take part.   Anyone requiring extra copies should
contact the Librarian of the Atlas Laboratory who will be pleased to supply them
free-of-charge to applicants within UK universities.   Applicants requiring copies
to be sent outside the UK may be asked to contribute to the cost of postage.

It is appropriate here to express my gratitude to all those involved in the organisation of the symposium - the staff at Cosener's House for making us so comfortable, the administrative and secretarial staff at ACL for their hard work and cooperation, and particularly Di Byfield, who also acted as a most capable receptionist during the meeting and has since been a tower of strength in the preparation of these Proceedings.

It is my earnest hope that the interest which was expressed in our Symposium will continue to inspire cooperation between workers in the different fields of engineering, numerical mathematics, and computing. The Finite Element Method is a powerful computational tool in engineering and scientific research. It is undoubtedly the most versatile of numerical techniques for solving a wide variety of structural and fluid problems, and given continued inter-disciplinary cooperation and SRC support, can take its place as the acknowledged leader of modern engineering computing methods.

Jean E Crow
Editor and Symposium Organiser

# CONTENTS

# THE ROLE OF THE SCIENCE RESEARCH COUNCIL AND THE ATLAS COMPUTER LABORATORY

E B Fossey

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 0QY

## INTRODUCTION

This paper is intended to provide some background information against which the
Finite Element Symposium may be set.    This is important because unless it is given
the outcome of the Symposium may not be properly understood.    I do not intend
giving a complete historical background, but I hope that I have provided enough
detail to allow the present position and purposes of the Laboratory to be seen in
their full context.

## THE SCIENCE RESEARCH COUNCIL

The Council came into being in 1965 because the Government of that time adopted the
recommendations of the Trend report on the organisation of civil science within the
United Kingdom.    It took over much of the support for scientific and engineering
research previously provided by the then Department of Scientific and Industrial
Research, the responsibility for the Royal Observatories at Edinburgh and Greenwich
(actually Herstmonceux Castle), the support for the United Kingdom's part in Space
Research, and lastly but by no means least the whole function of the National
Institute for Research in Nuclear Science (NIRNS).    This last body was set up some
years earlier, chiefly to provide the Universities of the United Kingdom with
facilities for research in high energy physics.    The Rutherford Laboratory (then
the Rutherford High Energy Laboratory) came into being as the home of the NIMROD
accelerator, a project which was costly even by present day standards.    The Atlas
Computer Laboratory founded in December 1961 was also a part of NIRNS for reasons
which we shall see later.

From the beginning the Council's activities in the scientific and engineering fields
have taken place either directly within its own laboratories or indirectly through

the provision of grants to Universities to allow the purchase of capital equipment
and the support of research staff. These hallmarks of SRC remain unchanged
although there has been over the years a shift in emphasis towards greater selectivity
and, perhaps more importantly for us today, there is the need to judge the timeliness
and promise of particular pieces of proposed work. We are undoubtedly going to be
faced with a time of financial stringency and so this yardstick will become
increasingly more important.

I do not wish to spend a long time discussing the Council's organisation. A quick
glance at the SRC Annual Report will give much useful detail of this kind. I do
want now to turn to the other component in my title, the Atlas Computer Laboratory.

## THE ATLAS COMPUTER LABORATORY

I have already indicated that the Laboratory was founded in December 1961. Its
purpose was to house the large Atlas I computer and to provide facilities on it in
equal proportion to three groups of customer, namely, the Atomic Energy Research
Establishment (AERE), the Rutherford High Energy Laboratory (RHEL), and the British
Universities. Precisely because the Laboratory was to serve these groups of user
the management of the Laboratory was placed under NIRNS, and not under the United
Kingdom Atomic Energy Authority which, through AERE, had carried out the negotiations
with Ferranti Limited concerning the Atlas I computer. Atlas was, even at the time
of its installation in 1964 and of the introduction of the service on it in October
of that year, a computer with a power very considerably in excess of that available
in many British Universities. In time both AERE and RHEL stopped using the
facilities as the result of obtaining machines of their own. The fraction of time
used by the Universities rose, of course, but other Government Departments and the
Laboratory itself used some time for work of their own.

The main uses of Atlas by the Universities arose as one would expect in the physical
sciences, but engineering calculations always figured in the tables among the
largest sources of work. Some of that work was a precursor to the work reported in
this symposium. A most significant development during the formative years was the
way in which the adoption of the X-RAY system as a standard program for use by
crystallographers came about. A meeting of most of the crystallographers in the
UK was arranged and on their advice the Laboratory set about mounting on Atlas the
X-RAY 63 system developed at the University of Maryland. Once the transfer had
been completed, the system was given active support not as one might have expected

by trained crystallographers but by people experienced in the maintenance of large systems who could provide also active advice on its use. This proved to be a most valuable and highly appreciated Laboratory service.

During the late 1960's the impact of the Computer Board in improving the University computer equipment led to a lessening of the position of the Atlas Laboratory as a source of extreme computer power. It also led to the decision in 1971 by the Science Research Council to align the services of the Laboratory much more with scientific aims of Council. The present procedure for obtaining time at the Atlas Computer Laboratory developed out of that decision, and it was reinforced by the allocation to the Laboratory of a 20% share in the Rutherford Laboratory's large IBM 370/195 computer installation. This share, together with the ICL 1906A computer at the Atlas Laboratory, now constitute the facilities whose resources are in the main distributed to users on the approval of the SRC specialist committees.

We commonly talk of time awarded by the Committees as 'guaranteed" time. We mean that a user may expect to obtain each week an amount of time during the period of the award. It gives him the advantage of knowing that this work will not be dependent upon the vagaries of the demands of other users and that he may therefore plan not only the use of the time awarded but perhaps more importantly of his own time. We believe that this can contribute very significantly to the pace with which the investigation or research is carried out.

As a result of the Laboratory's greater identification with the general research aims of the SRC, the service aspect of the Laboratory has been preserved. Without it, the Laboratory could not hope to be an effective instrument in contributing to the development of computational science, that is, any scientific or engineering project in which the computer plays a central part. (Much work in Fluid Dynamics is of this character.) However, a proposal has been agreed that the Laboratory should play an active part in advancing the knowledge and techniques used in a few selected areas of computational science. It is to act as a focus for workers in fields selected to collaborate under the general scientific direction of a steering panel chosen by an SRC Board. No suitable name has been found for this kind of activity, but it has been referred to as a "Meeting House". One such pilot project is now under way and the Laboratory looks forward to learning a great deal about this exciting development, and to playing a full part in the programme of work.

– 4 –

SUMMARY

What we see is therefore that the Council supports selected scientific programmes of timeliness and promise; that the Laboratory plays a significant part in this through providing computing on a guaranteed footing to users whose applications have received approval from the SRC specialist committees; that the Laboratory can play an active part in the development of software considered to meet an overall need; and that the Laboratory may in the future be embroiled much more deeply through Meeting House activity.

# CIRCULATION PROBLEMS

C A Brebbia
R Adey


Department of Civil Engineering
The University
Southampton
SO9 4NH

## INTRODUCTION

The environmental consequences and increasing cost of engineering works in tidal
zones, lakes etc has made the prediction of the motions of water bodies increasingly
important.    The main objectives of tidal models are to predict the motion of the
tide at certain locations during or after construction of the engineering works, and
detailed prediction of velocities for water quality studies in areas where known
tidal data is remote from the region of interest.    This prediction can be attempted
using the shallow water equations deduced in this paper.


For the case of harbours and similar water bodies it may be important to predict the
harmonic response of the system, which will affect moorings, structures, etc.    The
equation governing this phenomenon can be obtained as a particular case of the more
general shallow water equations previously deduced.


Finally, by linearizing the shallow water equations the flow in lakes, cooling ponds
and similar bodies can be approximated to provide an initial estimate of the circu-
lation.    Once the preliminary studies are finished the approximation can then be
checked against the full equations.


## SHALLOW - WATER EQUATIONS

The present state of the art and the lack of suitable data does not seem to justify
more complex mathematical models for circulation in coastal regions, lakes, etc than
those based on the numerical solution of the shallow water equations.    Fully three
dimensional solutions are not warranted at this stage as they would require a large
amount of extra data and computer time.

Finite difference solutions for circulation problems have been implemented in the past. (Abbott et al 1973; Heaps 1973; Leenderstse 1970; Leenderstse and Gritton 1971; Reid and Boding 1968; Water Quality Office 1971). They suffer from lack of mesh flexibility and, in some cases, difficulty in the satisfaction of the boundary conditions. The method of finite elements allows here for a great flexibility in the analysis grid and the advantages of having to satisfy only the essential boundary conditions.

In what follows a consistent derivation of the vertically averaged equations for long wave propagation is presented.

The governing equations for the fluid, neglecting temperature effects, can be written

$$- \frac{\partial p}{\partial x_k} + \frac{\partial \tau_{ik}}{\partial x_i} + \rho b_k = \frac{D(\rho v_k)}{Dt} \tag{1}$$

$$i,k = 1,2,3$$

$$\frac{\partial(\rho v_i)}{\partial x_i} + \frac{\partial \rho}{\partial t} = 0 \tag{2}$$

These equations are difficult to apply for the solution of shallow water problems because of
   (a)   the presence of the free surface
   (b)   the variable nature of the boundary when the tide rises and falls
   (c)   the large number of variables present in the solution.

These difficulties can be solved by simplifying the equations into what are called the shallow water equations. The first simplification we will introduce is to reduce the third momentum relationship to

$$- \frac{\partial p}{\partial x_3} = - \rho g \tag{3}$$

where the negative sign for g is due to the direction of the axis relative to gravity (see Figure 1).

Equation (3) implies that we have neglected all acceleration terms and corresponding

stresses.   Integrating (3) we have

$$p = \int_{x_3}^{\eta} \rho g\, dx_3 = \rho g(\eta - x_3) + p_a \qquad (4)$$

where $p_a$ is the pressure acting on the surface of the water, usually atmospheric pressure.   $\eta$ is the elevation of the free surface.

The remaining two momentum equations – in $x_1$ $x_2$ directions – can now be written

$$\frac{D(\rho v_k)}{Dt} = \frac{\partial (\rho v_k)}{\partial t} + \frac{\partial (\rho v_k v_i)}{\partial x_i} = -\frac{\partial p}{\partial x_k} + \frac{\partial \tau_{ik}}{\partial x_i} + b_k \qquad (5)$$

The above formula represents two equations, one for $k = 1$, the other for $k = 2$.   The subindex i in this case indicates summation, for $i = 1,2,3$.   The $v_k$ are average velocities, $\rho$ is the variable mass density and $\tau$ is the sum of viscous plus Reynolds stresses ($\tau_{ik} = \tau_{ki}$).

Our aim is now to integrate (2) and (5) with respect to x .   This gives for the continuity equation

$$\int_{-h}^{\eta} \left( \frac{\partial (\rho v_i)}{\partial x_i} + \frac{\partial \rho}{\partial t} \right) dx_3 = 0 \qquad (6)$$

where h is the depth from a datum surface (not necessarily horizontal) and $\eta$ is the free surface variable (Figure 1).
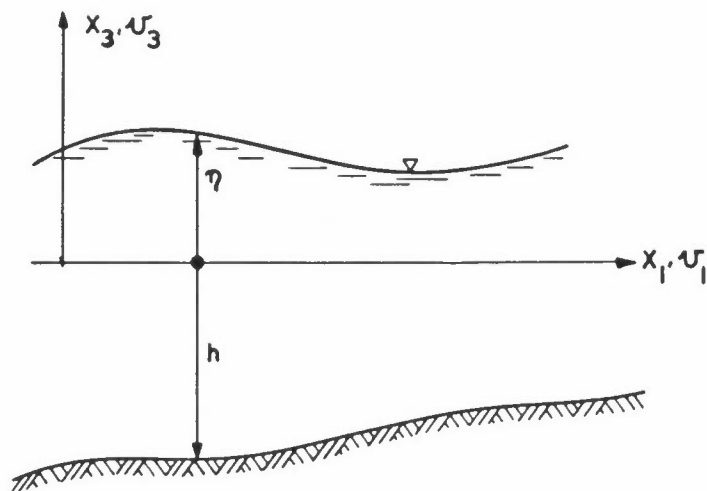


Figure 1:  Geometrical notation for shallow water equations

Let us now define a momentum flux variable $q_k$ (mass per unit length and time) such that

$$q_k = \int_{-h}^{\eta} \rho v_k \, dx_3 = \rho \int_{-h}^{\eta} v_k \, dx_3 \qquad (7)$$

Note that $\rho \, (x_1, x_2)$ is assumed not to be a function of $x_3$.

In order to integrate (6) we need to use the kinematic condition and Leibnitz rule for partial differentiation of an integral between variable limits. The latter gives for instance

$$\frac{\partial}{\partial x_1} \int_{h_1(x_1 x_2)}^{h_2(x_1 x_2)} f(x_1 x_2 x_3) dx_3 = \int_{h_1}^{h_2} \frac{\partial f}{\partial x_1} \, dx_3 + f\big|_{h_2} \frac{\partial h_2}{\partial x_1} + f\big|_{h_1} \frac{\partial h_1}{\partial x_1} \qquad (8)$$

and similarly for $x_2$.

The kinematic relationship for the free surface can be written

$$v_3\big|_{x_3 = \eta} = \frac{D\eta}{Dt} = \frac{\partial \eta}{\partial t} + v_1\big|_\eta \frac{\partial \eta}{\partial x_1} + v_2\big|_\eta \frac{\partial \eta}{\partial x_2} \qquad (9)$$

Applying (7) to (9) to equation (6) we obtain

$$\frac{\partial q_i}{\partial x_i} + \frac{\partial (\rho H)}{\partial t} = 0 \qquad (10)$$

where $H = h + \eta$.

To integrate now the two momentum equations (5) with respect to $x_3$, we define the following instantaneous velocities

$$v_1 = \bar{v}_1(x_1 \, x_2 \, t) + v_1'(x_1 \, x_2 \, x_3 \, t)$$

$$v_2 = \bar{v}_2(x_1 \, x_2 \, t) + v_2'(x_1 \, x_2 \, x_3 \, t) \qquad (11)$$

where $\bar{v}$ denotes the vertically averaged velocities and $v'$ the vertical deviations. Hence

$$< v_k > = \int_{-h}^{\eta} v_k \, dx_3 = \frac{1}{\rho} q_k \quad , \quad \bar{v}_k = \frac{1}{H} < v_k > \qquad (12)$$

since $< v_k' > = 0$.

We will assume that the body forces are only those due to Coriolis effects. Thus, for the northern hemisphere

$$b_1 = \rho f v_2 \quad , \qquad b_2 = \rho f v_1 \qquad (13)$$

One can also assume that the surface and bottom slopes are small with respect to unity. Hence we can approximate the internal stress components as follows (see Figure 2)

$$\tau_1|_s = \{ - \tau_{11} \frac{\partial \eta}{\partial x_1} - \tau_{12} \frac{\partial \eta}{\partial x_2} + \tau_{13} \}_{\text{surface}}$$

$$\tag{14}$$

$$\tau_1|_b = \{ \tau_{11} \frac{\partial h}{\partial x_1} + \tau_{12} \frac{\partial h}{\partial x_2} - \tau_{13} \}_{\text{bottom}}$$

and similarly for $\tau_2|_s$ and $\tau_2|_b$. Note that the $\tau$'s can be interpreted as external force components applied at the top and bottom.



$$\left(\frac{\partial \eta}{\partial x_1}\right)^2 << 1 \ , \ \left(\frac{\partial \eta}{\partial x_2}\right)^2 << 1 \qquad dA \simeq dx_1 \, dx_2$$
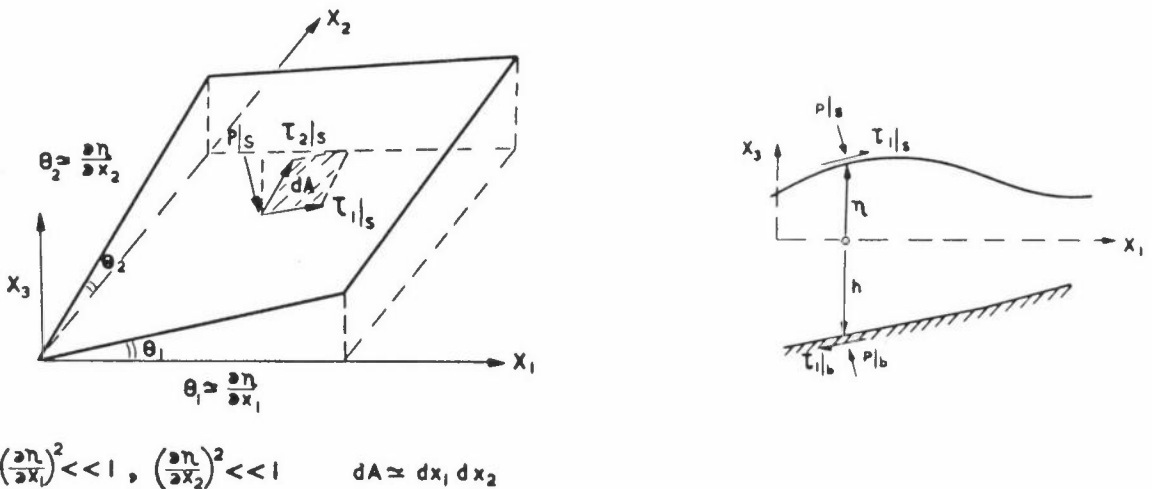
Figure 2: Surface notation

We will now substitute (11) to (13) into the momentum equations integrated with respect to $x_3$. One also needs to apply Leibnitz rule plus the kinematic condition to obtain the following result

$$\frac{\partial q_1}{\partial t} + \frac{\partial}{\partial x_1}\left(\frac{q_1^2}{H}\right) + \frac{\partial}{\partial x_2}\left(\frac{q_1 q_2}{H}\right) = -\frac{\partial N_p}{\partial x_1} + \frac{\partial N_{11}}{\partial x_1} + \frac{\partial N_{12}}{\partial x_2}$$

$$+ f_1 q_2 + p|_s \frac{\partial \eta}{\partial x_1} + \tau_1|_s + p|_b \frac{\partial h}{\partial x_1} - \tau_1|_b$$

(15)

$$\frac{\partial q_2}{\partial t} + \frac{\partial}{\partial x_1}\left(\frac{q_1 q_2}{H}\right) + \frac{\partial}{\partial x_2}\left(\frac{q_2^2}{H}\right) = -\frac{\partial N_p}{\partial x_2} + \frac{\partial N_{22}}{\partial x_2} + \frac{\partial N_{12}}{\partial x_1}$$

$$- f_2 q_1 + p|_s \frac{\partial \eta}{\partial x_2} + \tau_2|_s + p|_b \frac{\partial h}{\partial x_2} - \tau_2|_b$$

where

$$N_p = \langle N_p \rangle = \int_{-h}^{\eta} p \, dx_3 - \rho g \frac{H^2}{2} + H p_a$$

$$N_{11} = \langle \tau_{11} \rangle - \langle \rho v_1' v_1' \rangle$$

(16)

$$N_{22} = \langle \tau_{22} \rangle - \langle \rho v_2' v_2' \rangle$$

$$N_{12} = \langle \tau_{12} \rangle - \langle \rho v_1' v_2' \rangle$$

Furthermore, the $N_{ik}$ terms can be approximated by

$$N_{11} \simeq 2\varepsilon_{11} \frac{\partial q_1}{\partial x_1}$$

$$N_{22} \simeq 2\varepsilon_{22} \frac{\partial q_2}{\partial x_2}$$

(17)

$$N_{12} \simeq \varepsilon_{12}\left(\frac{\partial q_1}{\partial x_2} + \frac{\partial q_2}{\partial x_1}\right)$$

The $\varepsilon_{ik}$ are generalized viscosity coefficients (Connor and Wang 1974). For isotropic behaviour they become $\varepsilon_{11} = \varepsilon_{22} = \varepsilon_{12} = \varepsilon$.

The bottom shear stresses are usually given by the following relationships

$$\tau_1\big|_b = (\frac{g}{c^2}) \; \frac{1}{\rho} \; \frac{q_1(q_1^2 + q_2^2)^{\frac{1}{2}}}{H^2}$$

$$\tau_2\big|_b = (\frac{g}{c^2}) \; \frac{1}{\rho} \; \frac{q_2(q_1^2 + q_2^2)^{\frac{1}{2}}}{H^2}$$

(18)

where $\rho$ is the density of the water, g the acceleration due to gravity and c the friction factor or Chezy's coefficient. Note that $g/c^2$ is a dimensionless quantity. Sometimes a dimensionless friction , $c_m$, is used instead of Chezy's coefficient. The relationship between them is simply

$$c_m = \frac{g}{c^2}$$

(19)

The shear components on the surface of the water are generally due to wind and can be expressed as

$$\tau_1\big|_s = \gamma^2 \rho_a \; W^2 \; \cos \Theta$$

$$\tau_2\big|_s = \gamma^2 \rho_a \; W^2 \; \sin \Theta$$

(20)

where W is the velocity of the wind, $\rho_a$ the density of air and $\Theta$ the angle between the $x_1$ axis and the direction of the wind. $\gamma^2$ is a dimensionless coefficient called the wind stress coefficient; its value is given as approximately 0.0026 (Leenderstse and Gritton 1971).

Lastly we will assume that $\rho = \rho_o$ in all the terms except the pressure force terms.

Equation (15) can now be written

$$\frac{\partial q_1}{\partial t} + \frac{\partial}{\partial x_1}(\frac{q_1^2}{H}) + \frac{\partial}{\partial x_2}(\frac{q_1 q_2}{H}) = \frac{\partial}{\partial x_1}(N_{11} - N_p) + \frac{\partial N_{12}}{\partial x_2} + B_1$$

(21)

$$\frac{\partial q_2}{\partial t} + \frac{\partial}{\partial x_1}(\frac{q_1 q_2}{H}) + \frac{\partial}{\partial x_2}(\frac{q_2^2}{H}) = \frac{\partial}{\partial x_2}(N_{22} - N_p) + \frac{\partial N_{12}}{\partial x_1} + B_2$$

where the $B_1$ and $B_2$ terms are given by

$$B_1 = fq_2 + \gamma^2 \rho_a W^2 \cos \Theta - (\frac{g}{c^2}) \frac{1}{\rho} \frac{q_1(q_1^2 + q_2^2)^{\frac{1}{2}}}{H^2}$$

$$+ p_a \frac{\partial H}{\partial x_1} + \rho g H \frac{\partial h}{\partial x_1}$$

(22)

$$B_2 = - fq_1 + \gamma^2 \rho_a W^2 \sin \Theta - (\frac{g}{c^2}) \frac{1}{\rho} \frac{q_2(q_1^2 + q_2^2)^{\frac{1}{2}}}{H^2}$$

$$+ p_a \frac{\partial H}{\partial x_1} + \rho g H \frac{\partial h}{\partial x_2}$$

## Boundary Conditions

To solve the resulting system of equations (21) under condition (10) we need to establish the necessary boundary conditions. Consider that the S boundary (Figure 3) consists of two parts, land type boundaries $S_1$ and ocean type $S_2$.
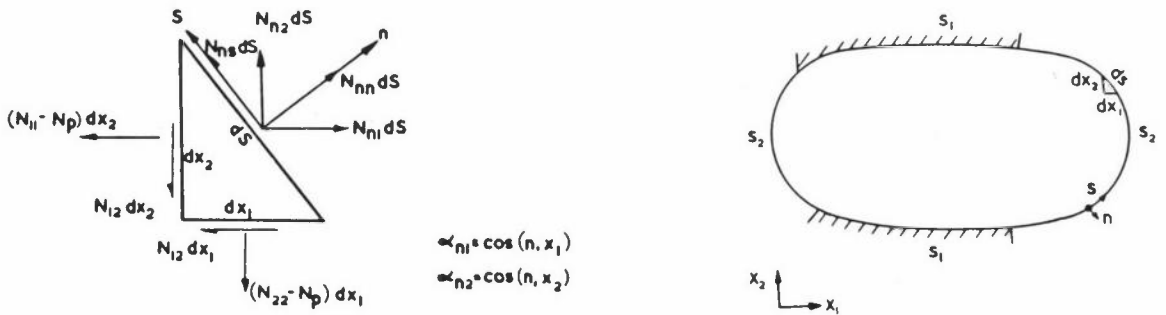


Figure 3: Boundary definitions

On the boundary reference system s-n, the momentum flux can be written as

$$q_n = \int_{-h}^{n} \rho v_n \, dx_3 = \alpha_{n1} q_1 + \alpha_{n2} q_2$$

(23)

$$q_s = \int_{-h}^{n} \rho v_s \, dx_3 = - \alpha_{n2} q_1 + \alpha_{n1} q_1$$

where $\alpha_{n1} = \cos(n, x_1)$, $\alpha_{n2} = \cos(n, x_2)$.

For the force resultants we have

$$N_{n1} = \alpha_{n1} (N_{11} - N_p) + \alpha_{n2} N_{12}$$

$$\tag{24}$$

$$N_{n2} = \alpha_{n1} N_{12} + \alpha_{n2} (N_{22} - N_p)$$

We can transform again to calculate $N_{nn}$ and $N_{ns}$

$$N_{nn} = \alpha_{n1} N_{n1} + \alpha_{n2} N_{n2}$$

$$\tag{25}$$

$$N_{ns} = - \alpha_{n2} N_{n1} + \alpha_{n1} N_{n2}$$

On land type boundaries we usually have

$$q_n = 0 \qquad\qquad \text{on } S_1 \tag{26}$$

However, if a river enters into the water we can specify its mass input as

$$q_n = \bar{q}_n = \{q\}_{\text{river inflow}}$$

$$q_s = 0 \qquad\qquad\qquad\qquad \substack{\text{on part of } S_1 \\ \text{boundary.}} \quad\tag{27}$$

On an ocean boundary, in principle, one has to prescribe the normal and tangential forces

$$N_{nn} = \bar{N}_{nn}$$

$$\qquad\qquad \text{on } S_2 \tag{28}$$

$$N_{ns} = \bar{N}_{ns}$$

but as the eddy viscosity terms in (21) are generally neglected the tangential flux or velocities can not be prescribed.   Hence our boundary conditions reduce to

$$q_n = 0 \quad \text{or } q_n = \bar{q}_n \qquad\qquad \text{on } S_1$$

plus

$$\tag{29}$$

$$N_{nn} = \bar{N}_{nn} = - N_p \qquad\qquad \text{on } S_2$$

## FINITE ELEMENT FORMULATION

In order to formulate our finite element model let us write equations (21) plus boundary conditions (29) and (10) plus condition (27) in the following weighted residual way

$$\iint \{ \frac{\partial q_1}{\partial t} + \frac{\partial}{\partial x_1} (\frac{q_1^2}{H}) + \frac{\partial}{\partial x_2} (\frac{q_2 q_1}{H}) + \frac{\partial N_p}{\partial x_1} - B_1 \} \, \delta q_1 \, dA$$

$$= \int_{S_2} \alpha_{n1} (N_p - \bar{N}_p) \, \delta q_1 \, dA \qquad (30)$$

$$\iint \{ \frac{\partial q_2}{\partial t} + \frac{\partial}{\partial x_1} (\frac{q_2 q_1}{H}) + \frac{\partial}{\partial x_2} (\frac{q_2^2}{H}) + \frac{\partial N_p}{\partial x_2} - B_2 \} \, \delta q_2 \, dA$$

$$= \int_{S_2} \alpha_{n2} (N_p - \bar{N}_p) \, \delta q_2 \, dS \qquad (31)$$

$$\iint \{ \frac{\partial q_1}{\partial x_1} + \frac{\partial q_2}{\partial x_2} + \frac{\partial (\rho H)}{\partial t} \} \, \delta H \, dA = \int_{S_1} (q_n - \bar{q}_n) \, \delta H \, dS \qquad (32)$$

where $\delta q_1$, $\delta q_2$, $\delta H$ are variations which satisfy the boundary conditions in mass flux and elevation. Note that if the elevation is known on $S_2$ the boundary conditions $N_p = \bar{N}_p$ will be identically satisfied (see equation (16)), hence the boundary terms on $S_2$ do not need to be included in (31) and (32).

The $B_i$ and $\partial N_p / \partial x_i$ terms can be combined to give

$$B_1^* = B_1 - \frac{\partial N_p}{\partial x_1}$$

$$= f q_2 + \gamma^2 \rho_a W^2 \cos \Theta - (\frac{g}{c^2}) \frac{1}{\rho} \frac{q_1 (q_1^2 + q_2^2)^{\frac{1}{2}}}{H^2} \quad \rho g H \frac{\partial (H-h)}{\partial x_1}$$

$$\qquad (33)$$

$$B_2^* = B_2 - \frac{\partial N_p}{\partial x_2}$$

$$= - f q_1 + \gamma^2 \rho_a W^2 \sin \Theta - (\frac{g}{c^2}) \frac{1}{\rho} \frac{q_2 (q_1^2 + q_2^2)^{\frac{1}{2}}}{H^2} - \rho g H \frac{\partial (H-h)}{\partial x_2}$$

The continuity equation (32) can be integrated by parts to render a simpler expression, ie

$$\iint \{ q_1 \frac{\partial \delta H}{\partial x_1} + q_2 \frac{\partial \delta H}{\partial x_2} - \frac{\partial (\rho H)}{\partial t} \delta H \} \, dA = \int_{S_1} \bar{q}_n \, \delta H \, dS \qquad (34)$$

which has then to be solved together with the following final momentum equations

$$\iint \{ \frac{\partial q_1}{\partial t} + \frac{\partial}{\partial x_1} (\frac{q_1^2}{H}) + \frac{\partial}{\partial x_2} (\frac{q_1 q_2}{H}) - B_1^* \} \cdot \delta q_1 \, dA = 0 \qquad (35)$$

$$\iint \{ \frac{\partial q_2}{\partial t} + \frac{\partial}{\partial x_2} (\frac{q_2 q_1}{H}) + \frac{\partial}{\partial x_2} (\frac{q_2^2}{H}) - B_2^* \} \, \delta q_2 \, dA = 0$$

The above variational statements are the starting point for the finite element models which we are now going to develop. Let us assume that the same interpolation functions apply for $q_i$ and $H$, ie

$$q_1 = \underset{\sim}{\emptyset} \, \underset{\sim}{q}_1^n \quad , \quad q_2 = \underset{\sim}{\emptyset} \, \underset{\sim}{q}_2^n \quad , \quad H = \underset{\sim}{\emptyset} \, \underset{\sim}{H}^n \qquad (36)$$

Substituting these values into the momentum and continuity equations (34)-(35) we can obtain

$$\delta \underset{\sim}{q}_1^{n,T} \, ( \underset{\sim}{M} \frac{\partial \underset{\sim}{q}_1^n}{\partial t} - \underset{\sim}{F}_1 ) = 0$$

$$\delta \underset{\sim}{q}_2^{n,T} \, ( \underset{\sim}{M} \frac{\partial \underset{\sim}{q}_2^n}{\partial t} - \underset{\sim}{F}_2 ) = 0 \qquad (37)$$

$$\delta \underset{\sim}{H}^{n,T} \, ( \underset{\sim}{M}^* \frac{\partial \underset{\sim}{H}^n}{\partial t} - \underset{\sim}{F}_H ) = 0$$

where

$$\underset{\sim}{M} = \iint \underset{\sim}{\emptyset}^T \underset{\sim}{\emptyset} \, dA \quad , \quad \underset{\sim}{M}^* = \iint \rho \, \underset{\sim}{\emptyset}^T \underset{\sim}{\emptyset} \, dA \qquad (38)$$

$$\underset{\sim}{F}_1 = \iint (\underset{\sim}{\emptyset}^T B_1^* - \underset{\sim}{\emptyset}^T A_1) \, dA$$

$$\underset{\sim}{F}_2 = \iint (\underset{\sim}{\phi}^T B_2^\star - \underset{\sim}{\phi}^T A_2) \, dA \tag{38}$$

$$\underset{\sim}{F}_H = \iint (\underset{\sim}{\phi}_{,1}^T q_1 + \underset{\sim}{\phi}_{,2}^T q_2) \, dA - \int_{S_1} \underset{\sim}{\phi}^T \bar{q}_n \, dS$$

The $A_1$ and $A_2$ stand for the advective part of the particle derivative, ie

$$A_1 = \frac{\partial}{\partial x_1} (\frac{q_1^2}{H}) + \frac{\partial}{\partial x_2} (\frac{q_1 q_2}{H})$$

$$\tag{39}$$

$$A_2 = \frac{\partial}{\partial x_1} (\frac{q_1 q_2}{H}) + \frac{\partial}{\partial x_2} (\frac{q_2^2}{H})$$

Equations (37) are a highly non-linear system of equations valid over one element. We can then assemble (37) for the whole continuum and write the resultant

$$\underset{\sim}{M} \underset{\sim}{\dot{Q}} = \underset{\sim}{F} (\underset{\sim}{Q}_o) \tag{40}$$

expression for which we assume the mass flux and elevation boundary conditions has been taken into account. $\underset{\sim}{M}$ is the global mass matrix, represents the derivatives with respect to time of the mass flux and elevation of all the nodes. All the other terms are included in $\underset{\sim}{F}$ and computed at $t = t_o$, or if an iterative method is applied at the end of the time step, they are the values obtained at the previous iteration. Integration of (40) may be attempted using an explicit method of solution such as Runge-Kutta or Euler technique or an implicit method, eg trapezoidal rule, Galerkin's, etc.

## EXAMPLE 1

A mathematical tidal model for the Solent was developed applying the theory described above (Figure 4). The model used 86 three node triangular elements with 58 nodes. The motion of shallow waters is strongly dependent on the bottom topography, hence the finite element mesh was chosen in such a way as to represent the complex channels and shallow flats (Figures 4(a) to 4(c)). The elevation, H, and two momentum flux parameters, $q_x$ and $q_y$, were taken as nodal variables. This gives a total of 174 unknowns.

Figure 4(a): The Solent
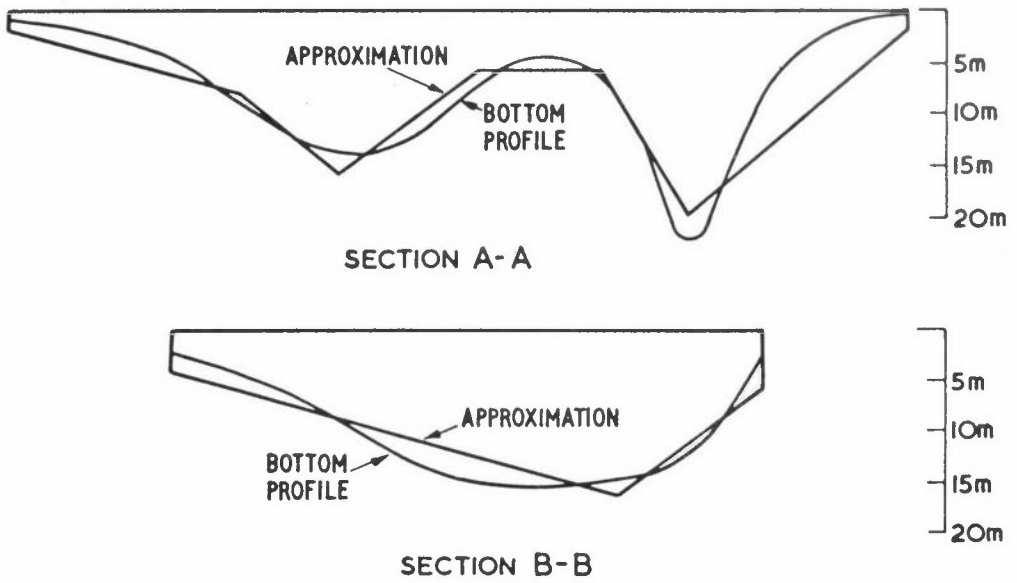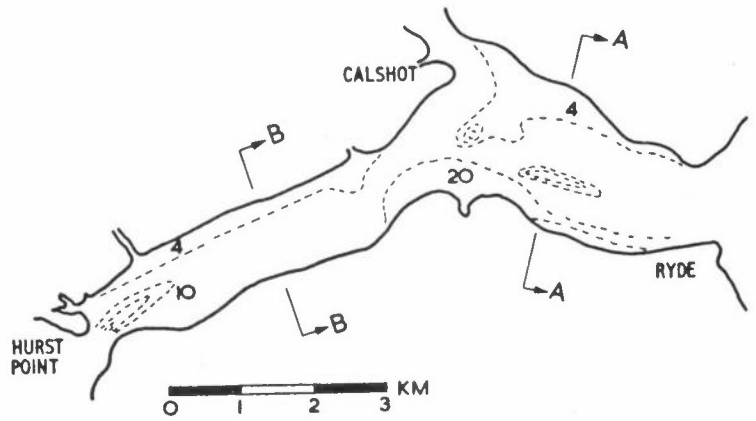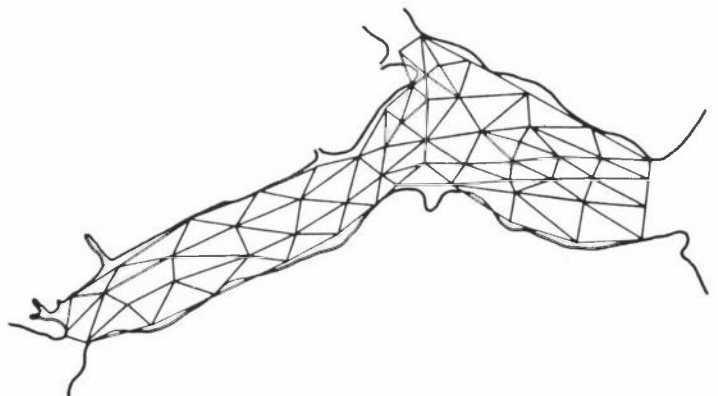
SECTION A-A

SECTION B-B

Figure 4(b): Cross sections and finite element approximation

Figure 4(c): Finite element mesh

The ocean type boundary conditions (elevation only) were obtained from the
Admiralty Tide tables for the Solent area.   The model can take into consideration
wind effects, bottom friction and Coriolis, although the latter was ignored for the
Solent.

·The time integration was carried out using fourth order Runge-Kutta and the model
was started with zero elevation and flux initial conditions.

The time step was selected using the Friedrichs-Courant-Lewy finite difference
criterion

$$\Delta t \; \leq \; \frac{\Delta x}{k \lambda} \tag{a}$$

where $\lambda$ is the celerity equal to $\sqrt{gh}$.   h was taken to be depth in the deepest part
of the Solent.

For a regular finite difference grid the value of k is $\sqrt{2}$.   This gives a $\Delta t$ of
approximately 20 seconds.   After several tests this value was found to be too small
and a $\Delta t = 120$ seconds was finally adopted.

The computer time for each $\Delta t$ was around one second on an ICL 1906A;   that is
approximately 6.5 minutes per tidal cycle.   Usually 3 to 4 tidal cycles are
required for any problem.

For this example, bottom friction (but not wind and Coriolis effects) was considered
with a value for Chezy's coefficient of

$$c^2 \; = \; 100 \; \; m/sec^2$$

throughout all the model.   Local variations together with a more realistic value
of the Chezy's coefficient are to be studied.

Typical results for mean velocities and tidal elevations are shown in Figures 4(e)
and 4(f).   The velocities are in agreement with those of the Admiralty charts.
The values of tidal elevations at Cowes depicted in Figure 4(f) closely follow
those of the chart.

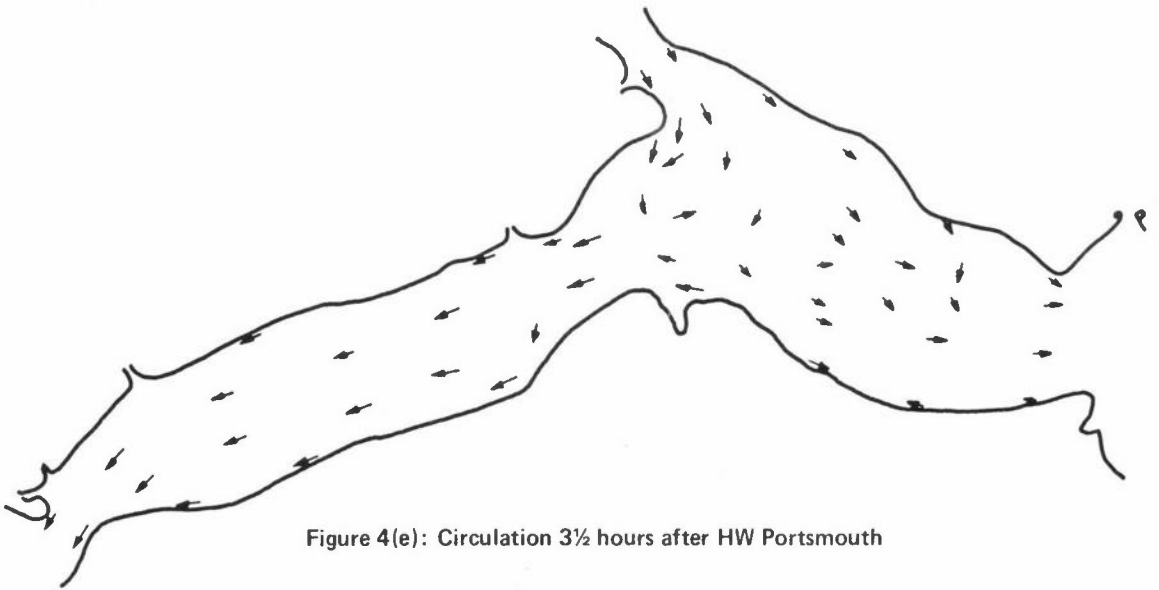Figure 4(d): Typical Spring tidal curve for
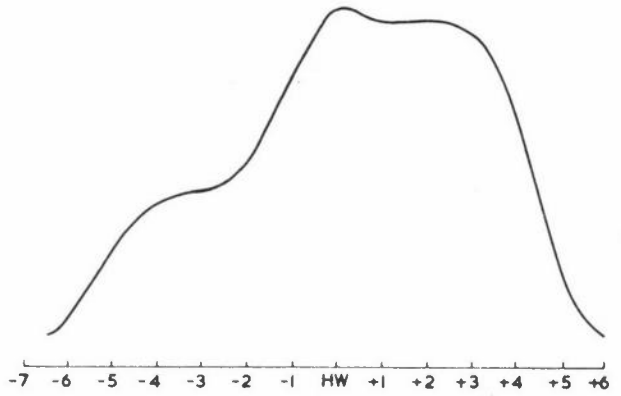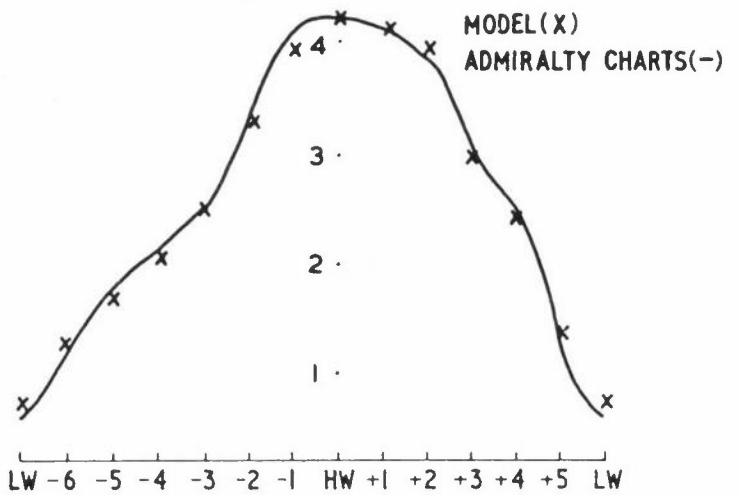Southampton Water

-7  -6  -5  -4  -3  -2  -1  HW  +1  +2  +3  +4  +5  +6

Figure 4(e): Circulation 3½ hours after HW Portsmouth

MODEL(X)
ADMIRALTY CHARTS(–)

Figure 4(f): Tidal height at Cowes

LW –6  -5  -4  -3  -2  -1  HW  +1  +2  +3  +4  +5  LW

## HARMONIC RESPONSE OF WATER BODIES

We will study how to predict the harmonic response of water bodies. We neglect the convective, friction, Coriolos, eddy viscosities and wind terms in equation (21). This gives

$$\frac{\partial q_1}{\partial t} = \rho g H \frac{\partial \eta}{\partial x_1}$$

$$\frac{\partial q_2}{\partial t} = \rho g H \frac{\partial \eta}{\partial x_2}$$

(41)

We also assume that the amplitude $\eta$ is small by comparison with the initial depth, h. Hence (41) becomes

$$\frac{\partial q_1}{\partial t} = \rho g h \frac{\partial \eta}{\partial x_1}$$

$$\frac{\partial q_2}{\partial t} = \rho g h \frac{\partial \eta}{\partial x_2}$$

(42)

For the constant density case, the continuity equation (10) can be written

$$\frac{\partial q_1}{\partial x_1} + \frac{\partial q_2}{\partial x_2} + \rho \frac{\partial \eta}{\partial t} = 0$$

(43)

On a <u>land</u> boundary, we prescribe the normal integrated velocity

$$q_n = \alpha_{n1} q_1 + \alpha_{n2} q_2 = \bar{q}_n \qquad\qquad \text{on } S_2$$

(44)

On an <u>ocean</u> boundary, the elevation is specified

$$\eta = \bar{\eta} (x_1 \ x_2 \ t) \qquad\qquad \text{on } S_1$$

(45)

The differential equations (42) and (43) can be transformed to a single differential equation of second order by differentiating (42) by $x_1$ and $x_2$ respectively and substituting them into the continuity equation (43) differentiated with respect to t. This gives

$$\frac{\partial}{\partial x_1} (h \frac{\partial \eta}{\partial x_1}) + \frac{\partial}{\partial x_2} (h \frac{\partial \eta}{\partial x_1}) - \frac{1}{g} \frac{\partial^2 \eta}{\partial t^2} = 0$$

(46)

where

$$\eta = \bar{\eta} \quad \text{on } S_1$$

$$h \frac{\partial \eta}{\partial x_n} = -\frac{\partial}{\partial t} \left(\frac{1}{g} \bar{q}_n\right) \quad \text{on } S_2$$

The normal velocity boundary condition has been replaced with the equilibrium equation for the normal direction which is consistent as the order of the differential equation governing the problem has increased.

Resonance and harmonic response due to tidal excitation can be investigated by expressing $\eta$ as

$$\eta(x_1 \ x_2 \ t) = \xi (x_1 \ x_2) \ e^{i\omega t} \tag{47}$$

where $\omega$ is the circular frequency. Then (43) becomes

$$\frac{\partial}{\partial x_1} \left(h \frac{\partial \xi}{\partial x_1}\right) + \frac{\partial}{\partial x_2} \left(h \frac{\partial \xi}{\partial x_2}\right) + \frac{\omega^2}{g} \xi = 0 \tag{48}$$

For tidal excitation

$$\xi = \bar{\xi} \quad \quad \text{on } S_1$$

$$\frac{\partial \xi}{\partial n} = f \quad \quad \text{on } S_2$$

For harmonic resonance, one considers a different set of homogeneous boundary conditions

$$\xi = 0 \ \text{on } S_1 \quad , \quad \frac{\partial \xi}{\partial n} = 0 \ \text{on } S_2$$

and determines the frequencies and modal shapes.

## Finite Element Formulation

From (46) and the boundary conditions we can write the following weighted residual expression

$$\iint \left[ \frac{\partial}{\partial x_1} (h \frac{\partial \xi}{\partial x_1}) + \frac{\partial}{\partial x_2} (h \frac{\partial \xi}{\partial x_2}) + \frac{\omega^2}{g} \xi \right] \delta\xi \ dx_1 \ dx_2$$

$$= \int_{S_2} (h \frac{\partial \xi}{\partial n} - f ) \delta\xi \ dS \qquad (49)$$

Integrating by parts we have

$$\iint (h \frac{\partial \xi}{\partial x_1} \frac{\partial \delta\xi}{\partial x_1} + h \frac{\partial \xi}{\partial x_2} \frac{\partial \delta\xi}{\partial x_2} - \frac{\omega^2}{g} \xi \delta\xi) dx_1 \ dx_2 = \int_{S_2} f \ \delta\xi \ dS \qquad (50)$$

If $f = 0$ the RHS of equation (50) disappears. Assume that the $\xi$ variable can be approximated on each element by

$$\xi = \underset{\sim}{\phi} \ \underset{\sim}{\xi}^n \qquad (51)$$

where $\underset{\sim}{\phi}$ is an interpolation function. $\underset{\sim}{\xi}^n$ are the nodal unknowns.

For an element we have

$$\underset{\sim}{\delta\xi}^{n,T} \iint \left[ h(\underset{\sim}{\phi}^T_{,1} \underset{\sim}{\phi}_{,1} + \underset{\sim}{\phi}^T_{,2} \underset{\sim}{\phi}_{,2}) - \frac{\omega^2}{g} \underset{\sim}{\phi}^T \underset{\sim}{\phi} \right] dx_1 \ dx_2 \ \underset{\sim}{\xi}^n = \underset{\sim}{\delta\xi}^{n,T} \int \underset{\sim}{\phi}^T f \ dS \qquad (52)$$

We can write (52) as

$$\underset{\sim}{K} \ \underset{\sim}{\xi}^n - \omega^2 \ \underset{\sim}{M} \ \underset{\sim}{\xi}^n = \underset{\sim}{P} \qquad (53)$$

where

$$\underset{\sim}{K} = \iint h (\underset{\sim}{\phi}^T_{,1} \underset{\sim}{\phi}_{,1} + \underset{\sim}{\phi}^T_{,2} \underset{\sim}{\phi}_{,2}) \ dx_1 \ dx_2$$

$$\underset{\sim}{M} = \frac{1}{g} \iint \underset{\sim}{\phi}^T \underset{\sim}{\phi} \ dx_1 \ dx_2 \ , \qquad (54)$$

$$\underset{\sim}{P} = \int \underset{\sim}{\phi}^T f \ dS$$

For the whole continuum we have,

$$\underset{\sim}{K} \ \underset{\sim}{\xi} - \omega^2 \underset{\sim}{M} \underset{\sim}{\xi} = \underset{\sim}{P} \qquad (55)$$

where $\underset{\sim}{K}$ and $\underset{\sim}{M}$ are both symmetric.

## EXAMPLE 2

As an illustration the case of Duncan Basin, built during the Second World War in Table Bay Harbour, South Africa, was analysed. This basin has been extensively studied, (Wilson 1954) as the features of the Bay are such that some of the seiche frequencies are greatly amplified (Figure 5(a)). This fact has been demonstrated by model experiment, harmonic analysis of seiches programmes and simple theoretical analysis, which can give reasonable results here as the shape of the basin is rectangular.
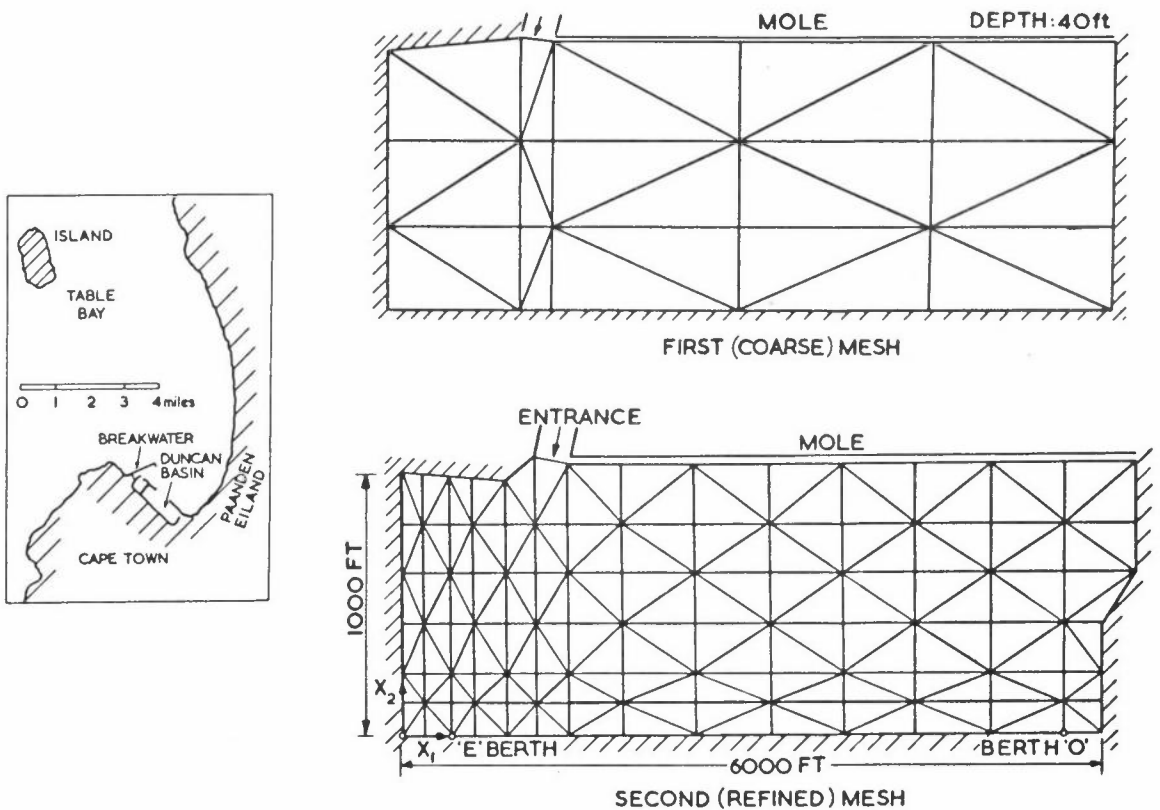


Figure 5(a): Finite element idealisation of Duncan Basin

It was decided to carry out an eigenvalue-eigenvector finite element analysis using a coarse mesh with 30 six node elements and 77 nodes. In this way the first natural frequencies were estimated (broken lines in Figure 5(b), $T = \frac{2\pi}{\omega}$).

Afterwards a more refined finite element analysis was carried out by dividing the basin into 168 elements, which gives 377 nodal unknowns. Periods T from 1 to 15 minutes were chosen and a unit elevation $\bar{\xi}$ input at the entrance of the harbour.

If we know the frequency ω of the waves and its elevation $\bar{\xi}$ at a given (ocean type) boundary, equation (55) can be written

$$( \underset{\sim}{K} - \omega^2 \underset{\sim}{M} ) \; \underset{\sim}{\xi} = \underset{\sim}{\bar{P}}$$

where the elements of the $\underset{\sim}{\bar{P}}$ vector are due to multiplying the known $\bar{\xi}$ elevations by the elements of $\underset{\sim}{K}$ and $\omega^2 \underset{\sim}{M}$ and passing the result to the right hand side.

The results for berths E and O are plotted in Figure 5(b) for periods T ranging from 1 to 15 minutes (full lines).
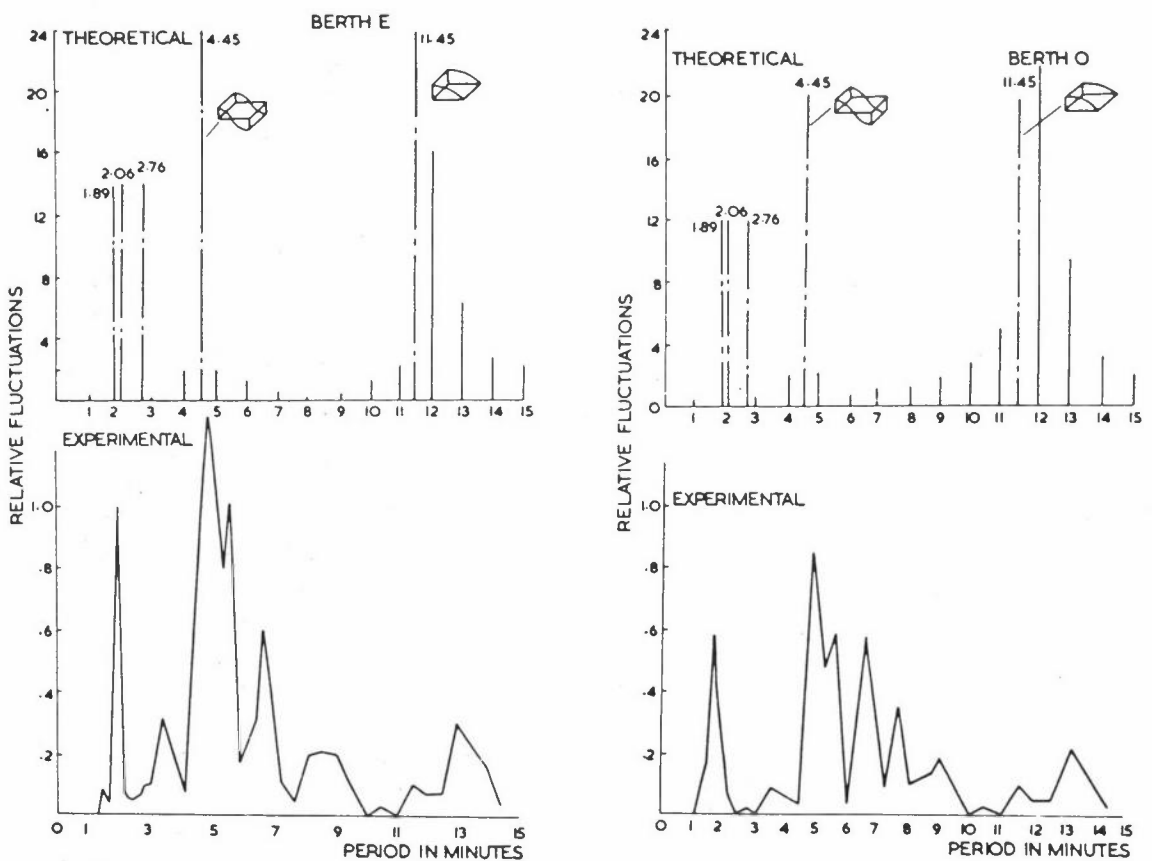


Figure 5(b): Harmonic response analysis of Duncan Basin

The finite element frequencies are in agreement with the experimental frequencies found on the physical model. The error in amplitude values is due to the lack of damping in the harmonic study.

## CIRCULATION IN LAKES

Flow in lakes, cooling ponds and other water bodies can be approximated to provide an initial estimate of the circulation, which can then be checked against the full shallow water equations. This flow is governed by linearized equations which are obtained in first instance by neglecting the inertia terms in the momentum equations, ie

$$-fq_2 + \rho gH \frac{\partial \eta}{\partial x_1} + (\tau_1|_s - \tau_1|_b) = 0$$

$$fq_1 + \rho gH \frac{\partial \eta}{\partial x_2} + (\tau_2|_s - \tau_2|_b) = 0$$

(56)

and the time dependent terms in the continuity formula

$$\frac{\partial q_1}{\partial x_1} + \frac{\partial q_2}{\partial x_2} = 0$$

(57)

If the $\eta$ values are much smaller than the h we can write $H \simeq h$ in (56), hence

$$-fq_2 + \rho gh \frac{\partial \eta}{\partial x_1} + (\tau_1|_s - \tau_1|_b) = 0$$

$$fq_1 + \rho gh \frac{\partial \eta}{\partial x_2} + (\tau_2|_s - \tau_2|_b) = 0$$

(58)

The momentum components can now be defined as

$$\frac{q_i}{H} \simeq \frac{q_i}{h} = \frac{\rho}{h} \int_{-h}^{\eta} v_i \, dx_3$$

(59)

The $\tau|_s$ terms are due to wind stresses and $\tau|_b$ are bottom friction stress components. The latter are assumed here to be linearly proportional to the mean momentum components

$$\tau_1|_b = \gamma q_1 \, , \qquad \tau_2|_b = \gamma q_2$$

(60)

We can now cross-differentiate equations (58) assuming that the derivatives of h are negligible (ie the bottom slope is small), and afterwards subtract both equations. This gives, taking continuity into consideration, the following equation

$$(\frac{\partial \tau_1|_b}{\partial x_2} - \frac{\partial \tau_2|_s}{\partial x_1}) = \gamma(\frac{\partial q_1}{\partial x_2} - \frac{\partial q_2}{\partial x_1})$$

(61)

One can propose a stream function such as

$$q_1 = \frac{\partial \psi}{\partial x_2} \quad , \quad q_2 = - \frac{\partial \psi}{\partial x_1} \tag{62}$$

Whence formula (61) becomes

$$W = \gamma \nabla^2 \psi \tag{63}$$

where

$$W = (\frac{\partial \tau_1|_s}{\partial x_2} - \frac{\partial \tau_2|_s}{\partial x_1})$$

The boundary conditions associated with this equation are

$$\frac{\partial \psi}{\partial n} = 0 \qquad \qquad \text{on 'land' boundaries}$$

$$\psi = \bar{\psi} \qquad \qquad \text{on 'inlet' boundaries} \tag{64}$$

Equation (63) plus conditions (64) can be written in variational form and the finite element technique can then be applied.   Any of the programs already developed for inviscid flow can be used.


Note that in the above formulation we have included the Coriolis parameter but assumed that it is constant for all the lake, ie the lake is small enough to neglect local variations in the Coriolis forces.

## REFERENCES

ABBOTT M B, DANESGAARD A and RODENHUIS G S, System 21, Jupiter, Journal of Hydraulic Research, Vol 11, No 1 (1973)

CONNOR J J and WANG J, Finite Element Modelling of Hydrodynamic Circulation, in Numerical Methods in Fluid Dynamics, edited by C A Brebbia and J J Connor, Pentech Press (1974)

HEAPS N S, Three Dimensional Model of the Irish Sea, Geophys J R Astr Soc, 35, 99-120 (1973)

LEENDERSTSE J J, A Water-Quality Simulation Model for Well-Mixed Estuaries and Coastal Seas, Vol 1, Principles of Computation. Memorandum RM 6230-RC, The Rand Corporation, Santa Monica, California (1970)

LEENDERSTSE J J and GRITTON E C, A Water-Quality Simulation Model for Well Mixed Estuaries and Coastal Seas: Vol II, Computational Procedures. Report R-708-NYC, The Rand Corporation, New York (1971)

REID R O and BODINE B R, Numerical Models for Storm Surge in Galveston Bay, Journal of the Waterways and Harbors Division, ASCE, Vol 94, No WW1 (1968)

WATER QUALITY OFFICE, Environmental Protection Agency, Estuarine Modelling: An Assessment, Water Pollution Control Research Series (1971)

WILSON B W, Surge Problems in Table Bay Harbour, South Africa, 5th Conference on Coastal Engineering, Grenoble (1954)

# STRUCTURAL ENGINEERING

R D Henshell

Department of Mechanical Engineering
The University
Nottingham
NG7 2RD

## DEFINITION OF STRUCTURAL ENGINEERING

Some while ago the term 'structural engineering' was used primarily in Civil
Engineering fields.   The use of the term 'structure' has increased with the advent
of the finite element method and it is my first intention to provide some definition.
In this paper I shall be concerned with the use of the finite element method for
obtaining some knowledge about stresses, strains, displacements or natural frequencies
of vibration of any body which is made of solid material.   This encompasses an
enormously wide range of engineering applications such as bridges, aircraft, machine
tools, engine components and drilling rigs.   This rather wide field of application
arises because the mathematical techniques and in many cases the computer programs
themselves are identical for all these cases.   The stiffened plate elements used to
solve a bridge problem may be identical to those used to model an aircraft fuselage.
The same solid elements may be used for modelling dams, rock strata and diesel engine
crankshafts.   All problems of practical interest happen to be highly irregular.

## THEORETICAL BASIS FOR THE FINITE ELEMENT METHOD APPLIED
## TO STRUCTURAL PROBLEMS

In a finite element analysis we proceed by dividing the irregular structure into a
number of simple elements of finite size.   Whether we are interested in the displace-
ments or stresses or the natural frequencies of our structure we next make some
assumption about how the displacement within this element varies.   There are alter-
native but less common formulations in which we make assumptions about the stresses
or strains.   The displacement within the element can be characterised by a list $\{u\}$
of displacements at the element nodes.   It happens that nature behaves in such a way
that energy is minimised in any structural situation.   We therefore write down an
expression for energy storage which involves a product of stresses with strains.   We
can obtain the strains by differentiating $\{u\}$ and in most situations the stresses will
be linear in the strains.   Thus our energy becomes a quadratic $\{u\}$

Strain energy, $\quad T = \frac{1}{2}\{u_e\}^T [S_e] \{u_e\}$ (1)

The use of Castigliano's theorem tells us that a differential of strain energy with respect to displacement is a force and therefore we have a linear relationship between the forces on an element and the element displacements

$$\{F_e\} \;=\; \frac{\partial(T)}{\partial\{u_e\}} \;=\; [S_e]\{u_e\}$$ (2)

The $[S_e]$ matrix, which is referred to as the element stiffness matrix, can be found for each of the elements in the structure. In order to form a mathematical model of the complete structure we observe that the displacements between adjacent elements must satisfy certain continuity requirements and additionally that the forces acting at a node must be in equilibrium. The process of joining together the finite elements is often termed merging and in static problems involving known applied loads the resulting set of equations is very similar to equation (2) except that all the matrices are very much larger. The overall set of equations will be written as

$$[S]\{u\} \;=\; \{F\}$$ (3)

In equation (3) the matrix $[S]$ is square, symmetric and very large and can be arranged in a highly banded form. Fortunately $[S]$ appears to be quite well conditioned even if there are thousands of rows and columns. The set of equations (3) is normally solved using Gaussian elimination or Choleski factorisation. The former method appears to be in most common usage now especially if a frontal technique is adopted to minimise core requirements.

## TYPES OF FINITE ELEMENT

There are a number of quite different types of structural finite element (see Figure 1). The first group, beam elements, are essentially one dimensional although they may be curved and they may be used to form three dimensional structures.

Membrane type elements may be used for two dimensional situations in which the displacements are all in the plane of the element. Figure 2 shows a two dimensional mesh for a turbine disc.
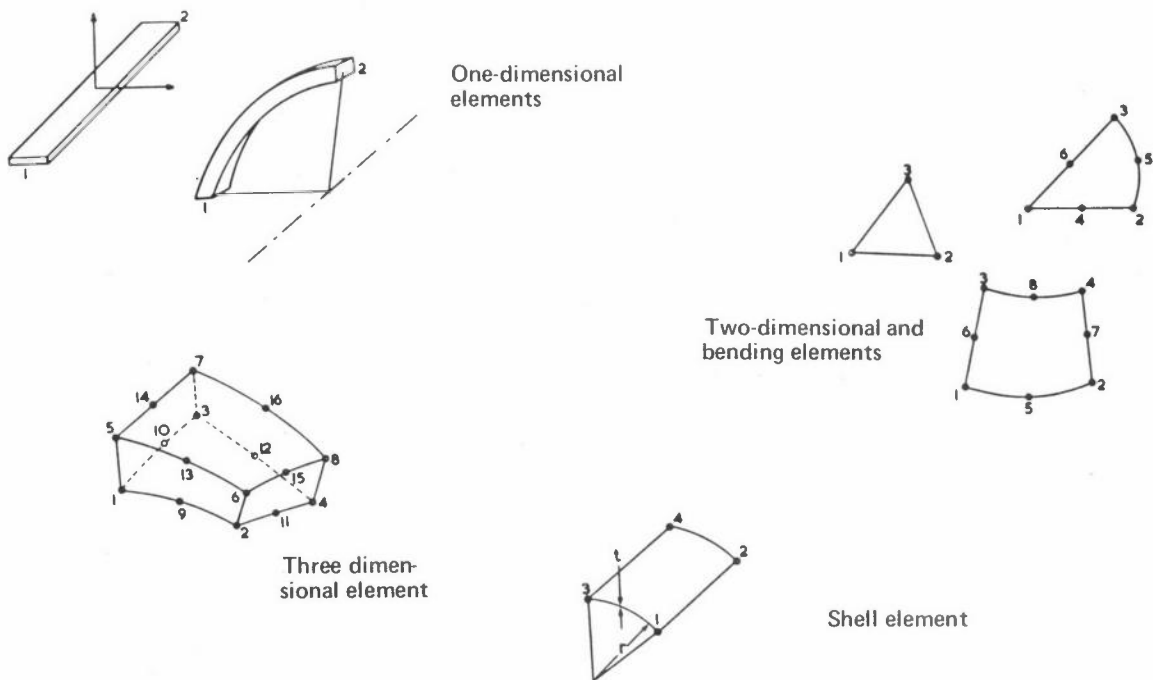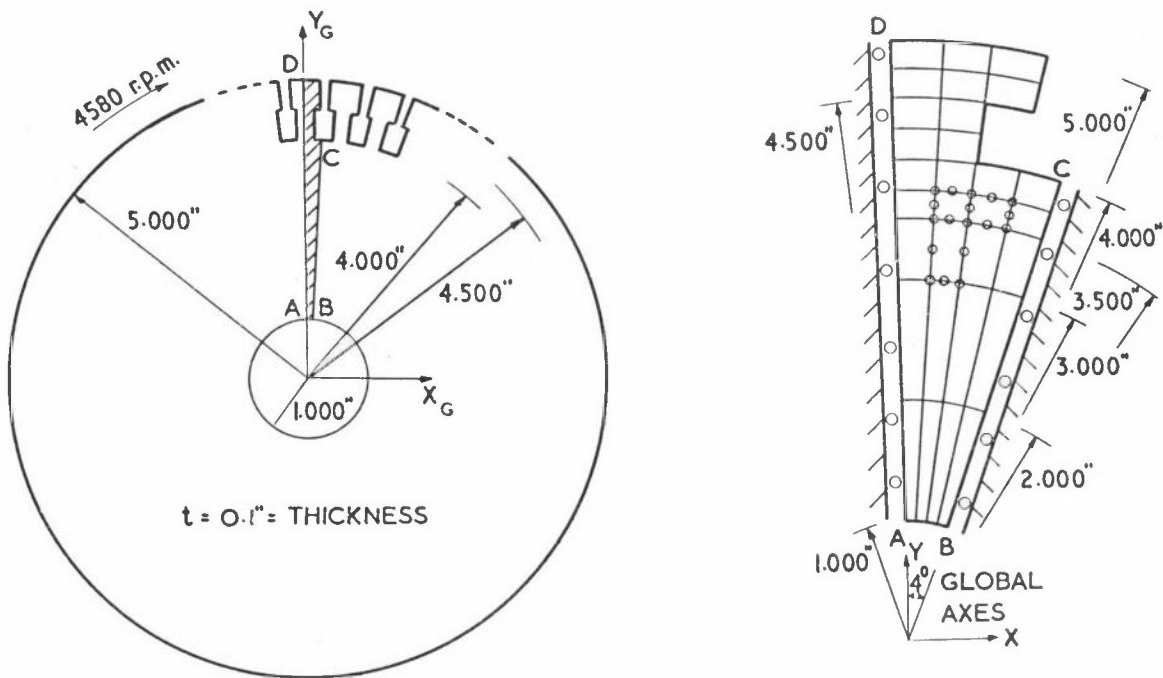
Figure 1: Types of finite element



Figure 2: A two-dimensional mesh used for analysing a turbine disc

Plate bending elements have the same geometry as the membrane elements except that the energy storage mechanisms are associated with displacement perpendicular to the plane of the element.

The membrane and plate bending families may be combined to give shell elements in which transverse and membrane displacements are considered. Figure 3 is an example of a civil engineering structure that involves shell type elements. Shell type elements can usually be extended to include curvature of the element plane.
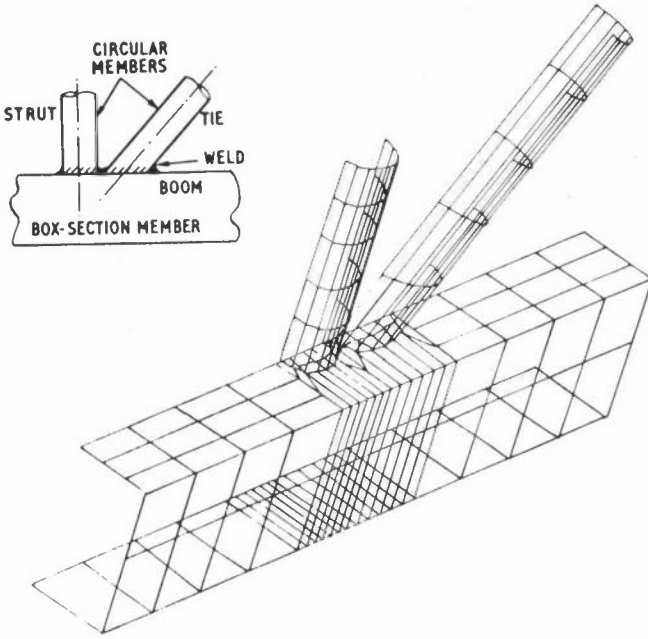


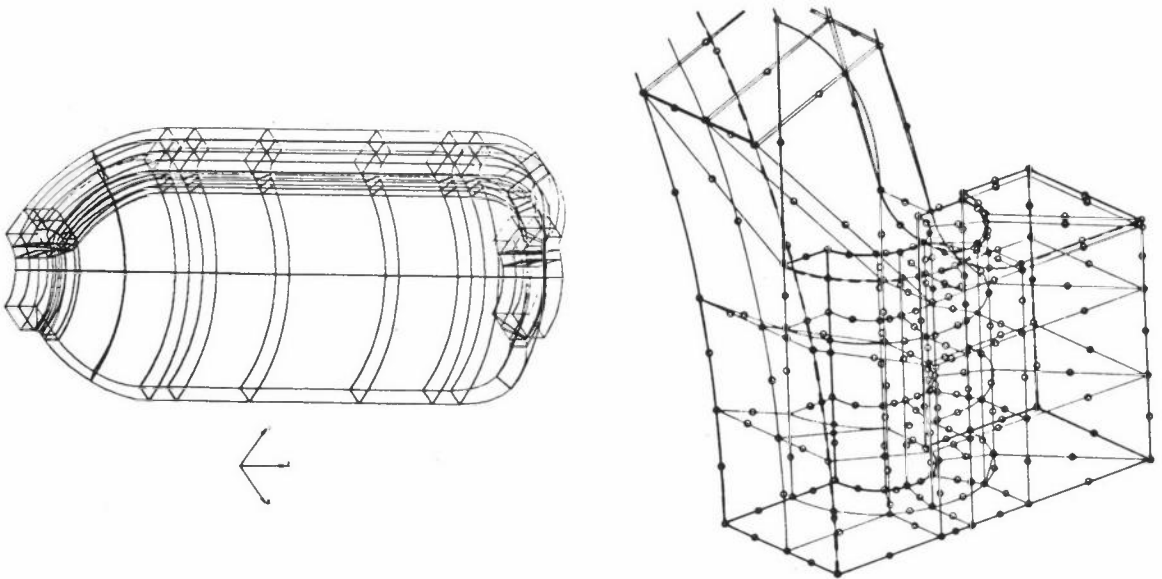Figure 3: A truss problem involving shell-type elements



Figure 4: Parts of a turbine casing idealised using three-dimensional elements

Full three dimensional elements are only used if none of the simpler structural elements is capable of achieving the required complexity of displacement distribution. Figure 4 shows portions of a turbine casing which comes into this category.

In general it is possible to mix all types of finite element together in one idealisation provided that the appropriate continuity and equilibrium constraints are applied between the elements. Figure 5 shows an idealisation of part of an aircraft structure in which curved shell, flat shell and curved beam elements are all used together in the same idealisation.



**CURVED BEAM ELEMENTS**    **PLATE ELEMENTS WITH MEMBRANE AND BENDING ACTION SUPERIMPOSED**    **CURVED SHELL ELEMENTS WITH VARYING THICKNESS**
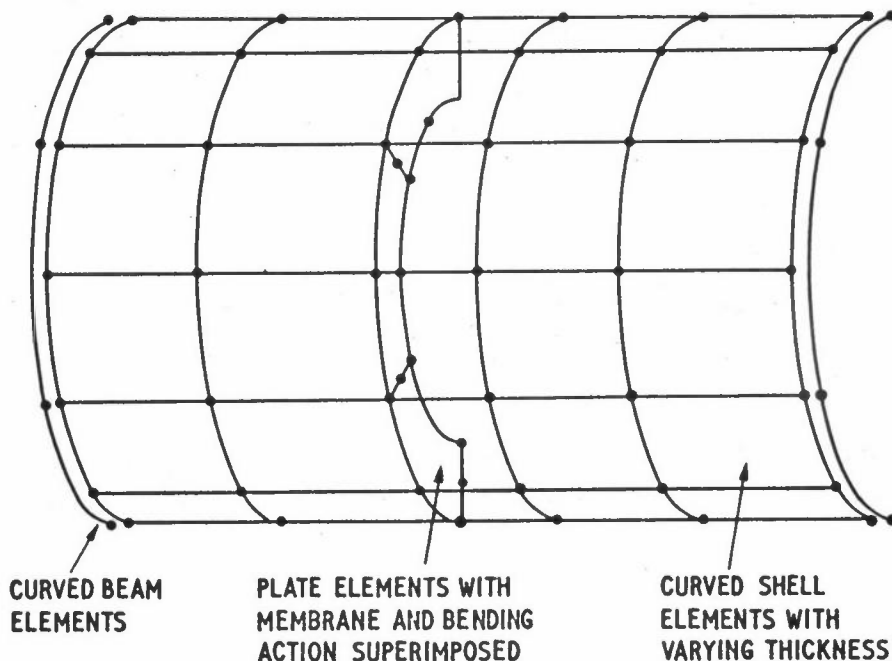
Figure 5: An idealisation involving a mixture of shell, plate and beam elements. (Only one half of the idealisation is shown since there is symmetry about the centre plane. Symmetric or antisymmetric boundary conditions are used to obtain all the modes of vibration.)

When the finite element method first became popular very simple finite elements were used. It became clear that more complicated elements gave better solutions for a given amount of data preparation and computing effort. Elements may be made more complicated in one of two ways

    (i)   by increasing the number of nodes per element;

   (ii)   by characterising the displacement within the element in terms of nodal strains as well as nodal displacements.

The first of these two methods has proved to be of enormous value and the isoparametric elements are generally regarded as being the best compromise from considerations of accuracy and ease of practical use.

## TYPES OF STRUCTURE ANALYSES

The most commonly occuring problem is that defined by equation (3) in which we wish to examine the behaviour of a structure under the influence of some known loads. Solution of the equations for {u} gives us the displacement of a system which may be differentiated to give strains. Stresses may be found from strains using linear material laws. Often the list of forces { F } will be derived from a distributed loading such as a pressure or gravity. A computer program can easily calculate {F}; thus a distributed loading problem is virtually identical to a problem in which known point loads are applied.

We often wish to calculate the natural frequencies of vibration of a structure. In such cases we subtract the kinetic energy from the strain energy to give a Langrangian

$$\text{Langrangian,} \quad L = T - V \tag{4}$$

where V is the kinetic energy. The Langrangian is minimised with respect to the displacements again and under sinusoidal conditions we obtain the eigenvalue problem

$$[S] \{u\} - \omega^2 [M] \{u\} = \{0\} \tag{5}$$

where $[M]$ is the system mass matrix and $\omega$ is the radian frequency. The solution of equation (5) for the eigenvalues $\omega^2$ and the eigenvectors {u} appears to be a formidable task for systems of practical size. However, it happens that we can use static condensation or eigenvalue economisation in order to reduce the thousands of degrees of freedom to at most 100 without any appreciable loss of accuracy. The solution to equation (5) gives a steady state vibration. Transient effects may be studied by writing

$$[S] \{u\} - [M] \{\ddot{u}\} = \{F(t)\} \tag{6}$$

where {F(t)} represents some arbitrary forcing function. Damping may be included into equation (6) in various ways. It should be noted however that wave propagation problems cannot be solved by the finite element method in its usual form. Present methods only give standing wave solutions and are therefore inapplicable in shock situations. For many structures there is a particular load which will cause the body to buckle. As usual the strain energy of the structure is important. We also need to consider the work done by the load on second order displacements. By following through the mathematics we derive an eigenvalue problem which is similar to that of equation (5)

$$[S] \{\underline{u}\} - \alpha [S'] \{u\} = \{0\} \tag{7}$$

where $[S']$ is the second order or geometric stiffness matrix calculated for a particular applied load and $\alpha$ is the number by which that load needs to be multiplied in order to cause $\{u\}$ to take up any required value; that is, to cause the structure to buckle.

All the situations considered so far have been within the realms of small displacements, first order theories and linear stress-strain relation. Wherever possible structural analysts will use simple theories because they tend to be at least an order of magnitude cheaper to use than the corresponding second order theories. However, there are a wide range of structural problems in which non-linearity is important.

The simplest class of non-linear problems is that in which the displacements which the structure undergoes are sufficiently large to change the geometry to such an extent that the stiffness matrix is not constant as the load is applied. As a rule of thumb these geometric non-linearities will occur when the displacement of a shell surface approaches the thickness of that surface. In the large displacement problem we know that $[S]$ is a function of $\{u\}$ but we have no motivation for finding a relation between them since equation (3) would become a very large non-linear set of equations in $\{u\}$. Instead we proceed by observing that the structure is still reasonably linear for small increments of load and we can recalculate or adjust $[S]$ as the total load is applied in steps. Figure 6 shows how the non-linear effects in a food storage container affect the final deflection.
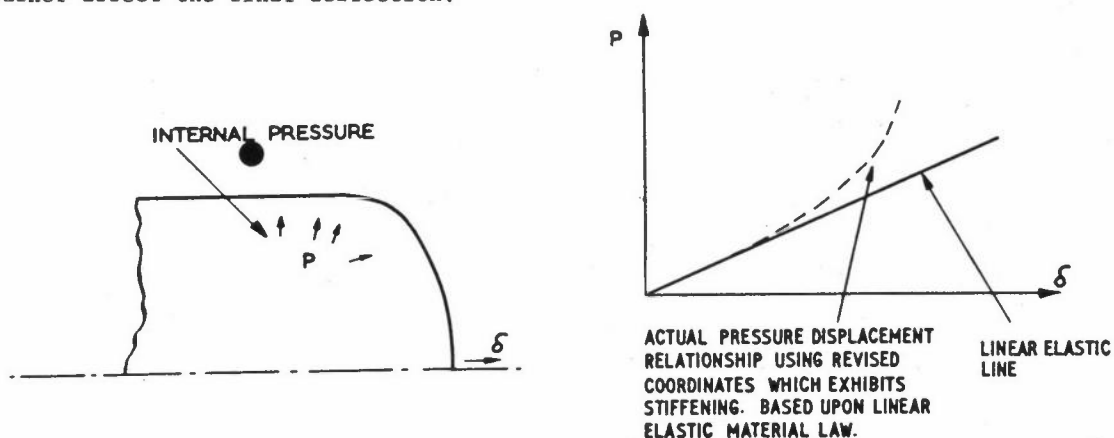


Figure 6: Non-linear behaviour of a thin pressure vessel

The materials of which structures are made often exhibit non-linearities in the stress strain relationships. At a particular stress level a material will yield and flow plastically (see Figure 7). The most common way of solving a problem involving plastic flow is to calculate first of all the stresses and strains which occur when a small increment of the load is applied. This calculation is based upon the linear elastic load and therefore some of the stresses and strains calculated do not lie on the true material curves. The amount of discrepancy between a typical point and the true curve is equated to an initial strain and a set of loads which would be necessary
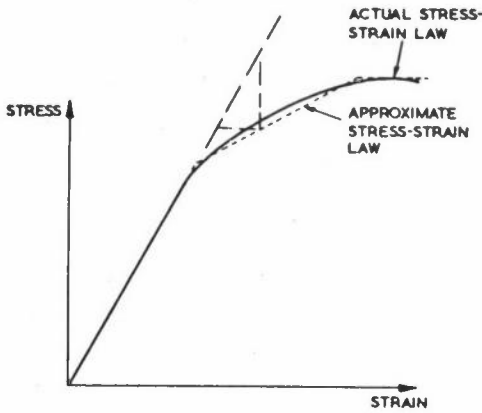
Figure 7: Stress-strain relationship for an
elastic-plastic material

to move these points back to the material curve are calculated. The process is not
exact and an iterative technique is needed. When all points lie sufficiently close
to the true material curve a further increment of load is applied. Figure 8 shows the
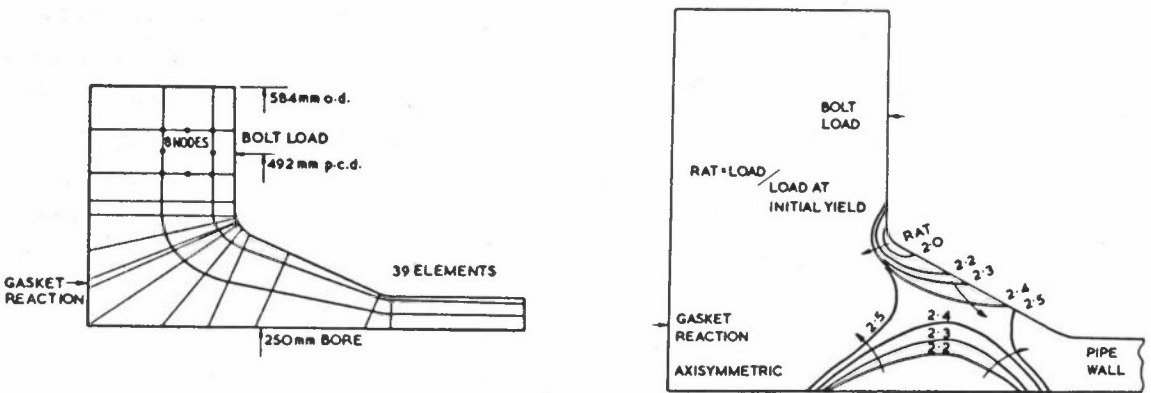manner in which the yield surface varies with load for a bolted flange joint. In



Figure 8: The growth of the plastic region in a pipe flange

linear elastic analysis it is assumed that if a set of forces are constant in time
then the stresses and displacements are also constants. In practice if we take a
tension bar and apply a load then there will be a rate of straining which is a highly
non-linear function of temperature and stress. The solution for displacements and
stresses in a structure is therefore a function of time and we proceed by considering
a particular time at which we know all the temperatures, displacements and stresses.
With a knowledge of material properties we can calculate the rate of straining. After
a suitably small time increment we can determine the loads which are implied by the
change of strain and carry out a linear elastic solution for a set of incremental
displacements. Figure 9 gives an example of a creep solution.


The creep problem involves a time variation of displacements and this leads naturally
on to a concept which has been recently gaining popularity in finite element
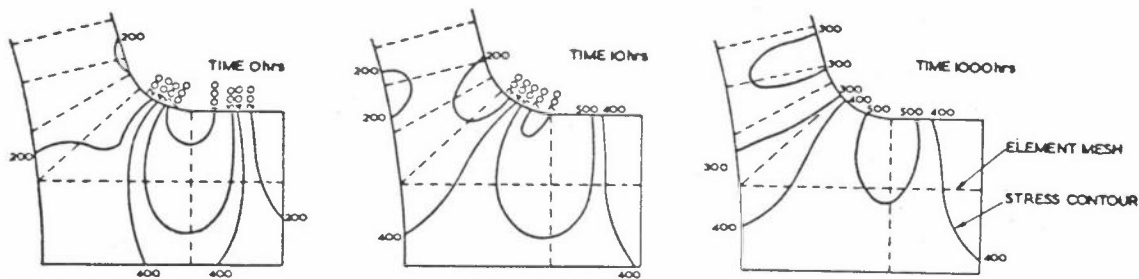structural analysis.

Figure 9: Creep results for part of a turbine (Stress contours showing the manner in which a stress distribution is redistributed with time. The idealisation represents the intersection between a toroid and a cylinder in a lead model of a turbine casing.)

Until a few years ago a time dependent finite element solution was obtained by using finite elements in space and the finite difference method in time. Recently a number of workers have investigated the possibility of using finite elements in the time coordinate. This technique may well become quite standard within a few years.



Figure 10: Research effort in finite element methods
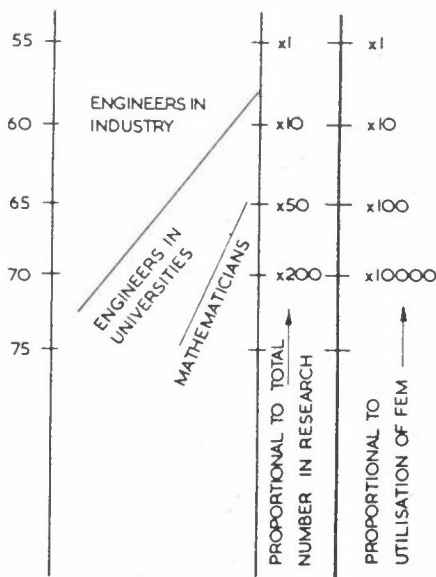
## THE FUTURE OF FINITE ELEMENTS

It is inevitable that this discussion of future trends is highly personal.

Let us consider first the theoretical methods associated with finite elements. Finite elements were originally developed by engineers in order to solve their problems. It is only quite recently in the history of finite elements that considerable mathematical investigations of their properties have been carried out. Figure 10 shows approximately how research trends have progressed. This process will certainly continue, it

is probable however that the feed-back to engineers using the method will be in terms of improved understanding of the method and its properties rather than any startling new techniques. The days in which everyone is developing a new and better element seem to be nearly over, at least as regards structural applications of the finite element method.

There are probably less than 10,000 people in the UK who have a reasonable working knowledge of the finite element method. This is a very small number considering the enormous increase in the usage which will take place in the next 10 years. There is a very pressing need to provide widespread education. All Universities and Polytechnics ought to include a course on the subject for all students and not for just a selected few who happen to choose a particular option. Since finite element methods are applicable to an enormous range of problems a finite element program ought to be very general. This implies large programs. Large finite element programs will become more readily available. There is a danger that these large programs might be treated as black boxes and in order to prevent this they must be transparent. Users must not be in awe of the code. In particular large programs should :

   (i)     be written in FORTRAN

   (ii)    have both internal (comment) and external (manuals) documentation.

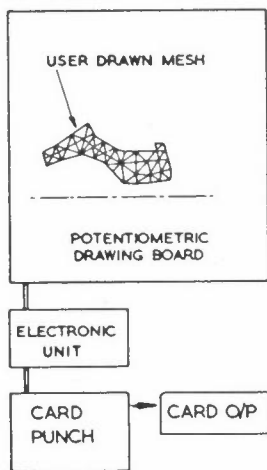   (iii)   be written in such a way that users can modify or 'play' with them.



Figure 11: Finite element data produced with the aid of a digitiser
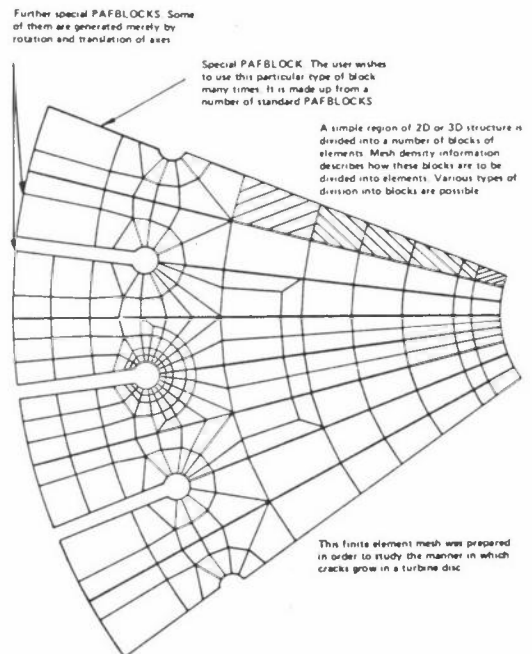


Figure 12: Finite element mesh generation with the aid of PAFBLOCKS

Even with the latest mesh generation aids it is still a rather laborious process to generate a finite element mesh for a complicated structure. The following mesh generation aids which are the subject of much development will probably be augmented by new techniques in coming year.

(1)     Use of special purpose programs which are limited to narrow classes of structure.

(2)     The use of a digitiser (see Figure 11).

(3)     The use of building block systems (see Figure 12).

(4)     Interactive graphics.

(5)     The use of mesh generation languages.

The problem of interpreting output is similar to that of mesh generation and similar developments can be expected in this field. It is imperative that complex finite element idealisations be checked thoroughly. This is possible with the aid of various views and exploded plots, with numbered nodes and/or elements. It is already common for the results of a two dimensional stress analysis to be plotted in various ways to show the principal stresses as vectors (Figure 13) or to show stress contours (Figure 14). There is no really satisfactory method at the moment available for
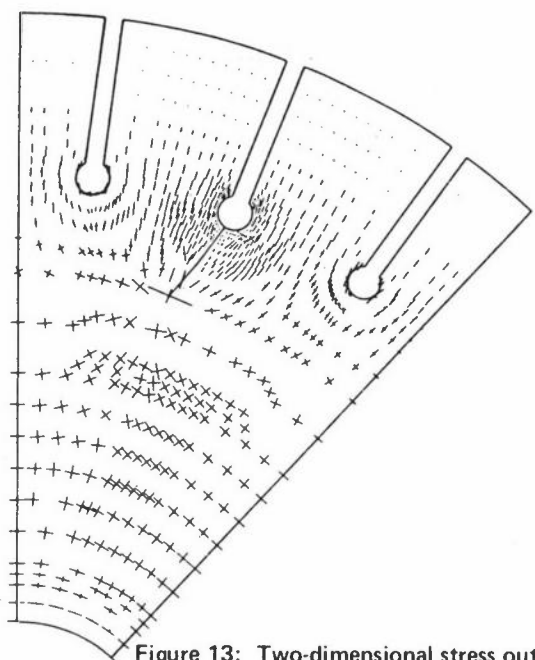


Figure 13:  Two-dimensional stress output — vectors representing principal stresses in a portion of a turbine disc. (The arms of the crosses represent, to scale, the magnitudes and directions of the principal stresses)
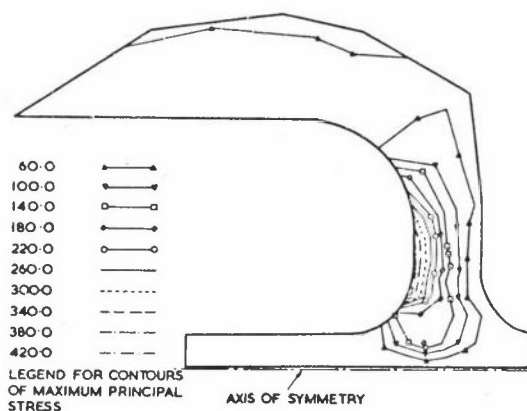
Figure 14:  Two-dimensional stress output — contours of stress in a gyroscope flywheel.

giving three dimensional stresses in a convenient form.    This situation ought to be remedied.

In the case of really complicated structures it is inevitable that a finite element analysis will involve considerable simplification.    Research is necessary in order to provide guide lines so that the simplified mesh provides a reasonable accuracy.

Since we expect the finite element method to become more widely available and to be used on more complicated structures the need for building safeguards into finite element systems will become even more essential than it is today.    Graphical checks on idealisations, such as that shown in Figure 15 are now universal and will become more sophisticated.
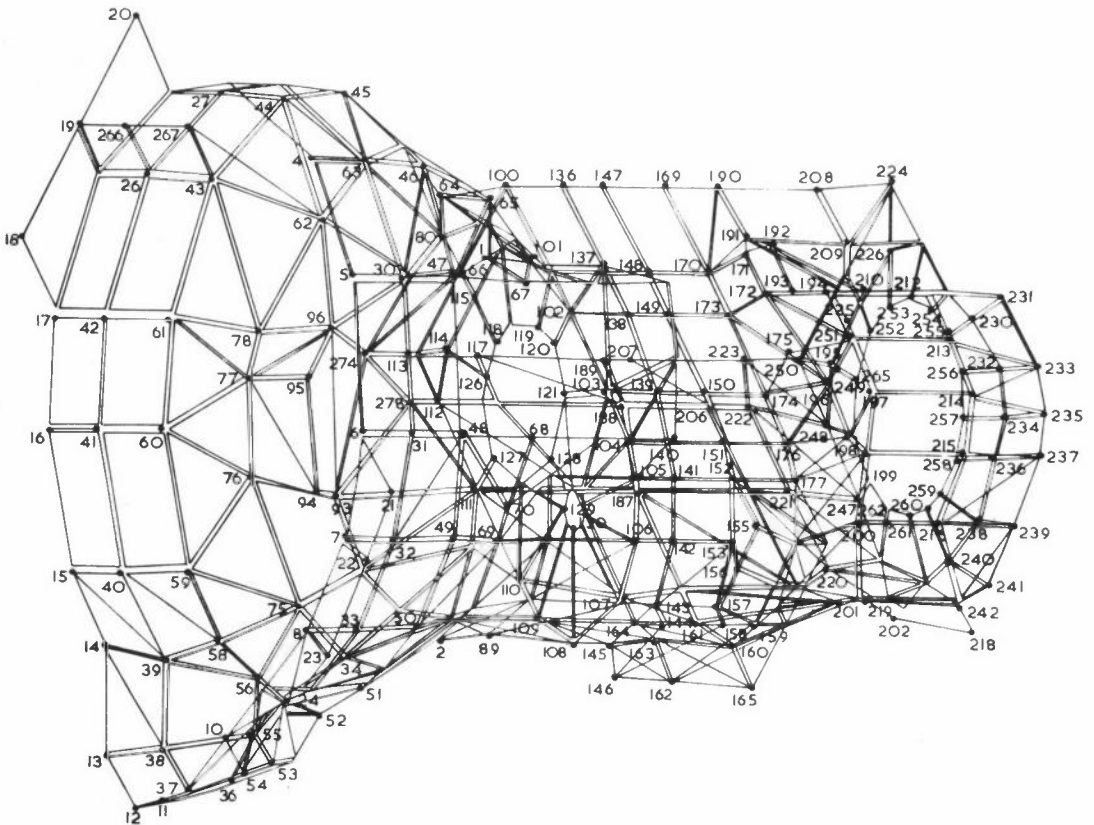


**Figure 15:  Data checking with the aid of a graph plotter**

In conclusion I consider whether or not the finite element method has a long term future.    For the forseeable future engineers will be interested in the analysis of very complicated shapes.    Closed form solutions are always likely to be impossible for practical problems and we shall have to resort to dissecting the structure into components of manageable simplicity.    The finite difference method (FDM) was

developed before the finite element method (FEM). The FDM turns out to be less convenient and less accurate than a high order FEM. Because of this there has been a tendency to improve the FDM as shown in Figure 16. A recent newcomer to the field has been the boundary integral method (BIM). These elements may not always have the same properties as those that we know of today, but is very likely that large parts of the existing technology will continue to be directly useful and indeed the method of the future may well be called the xxxxx finite element method.



Figure 16: The relationship of various methods of solving structural engineering problems

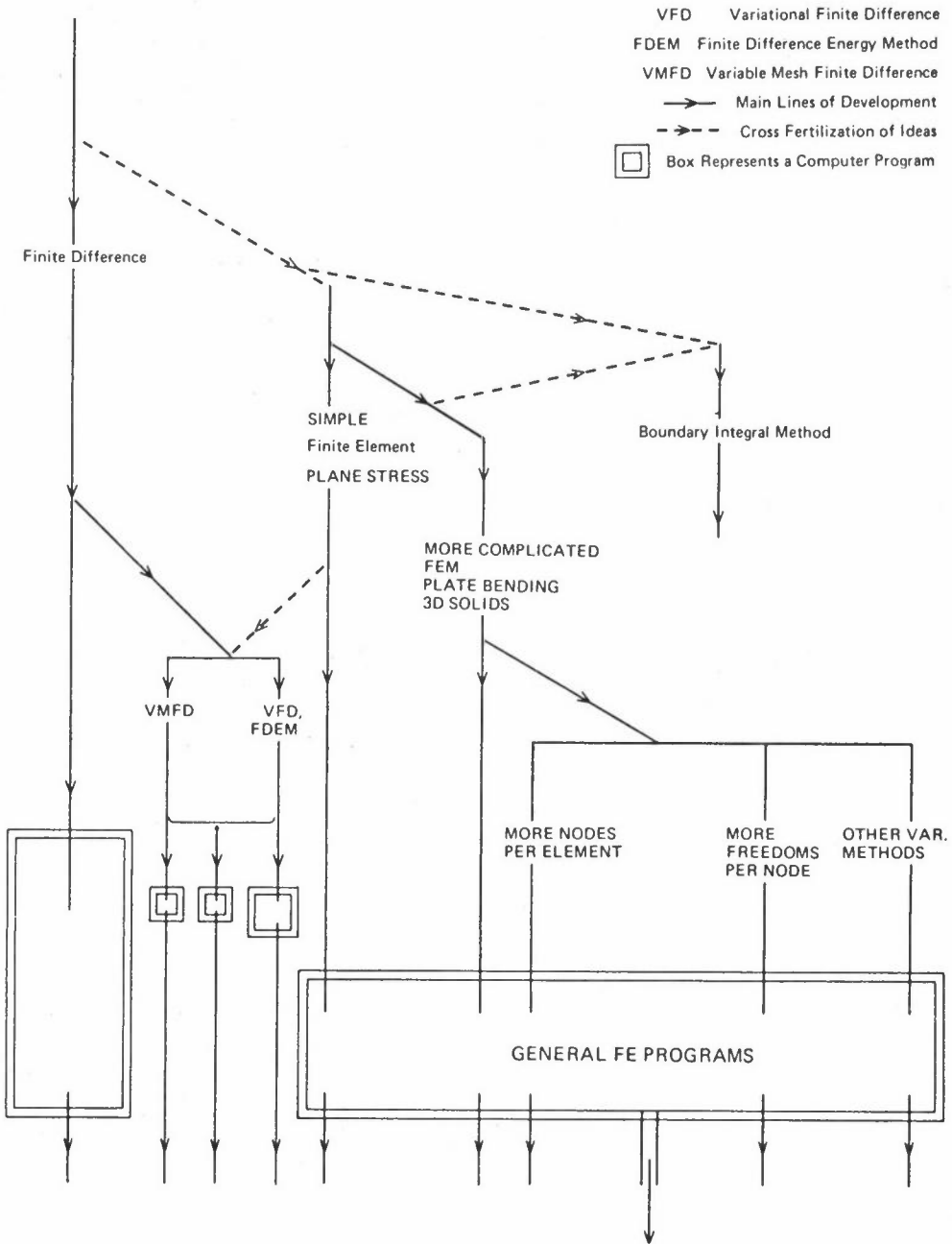# A FINITE ELEMENT FORMULATION OF HEAT CONDUCTION AND HEAT AND MASS TRANSFER PROBLEMS

R W Lewis

Department of Civil Engineering
University of Wales
Singleton Park
Swansea
SA2 8PP

## ABSTRACT

The paper presents techniques for dealing with transient heat conduction problems
with non-linear physical properties and boundary conditions.   Also investigated is
the coupled phenomenon of heat and mass transfer in porous bodies.   The numerical
results obtained for both these types of problems are compared with known analytical
results and indicate good agreement.

## INTRODUCTION

The finite element method has been applied successfully on many occasions to linear
situations involving heat conduction problems (Wilson and Nickell 1966;  Zienkiewicz
1971;  Zienkiewicz and Parekh 1970) but relatively little work has been reported for
temperature dependent thermophysical properties and/or non-linear radiation-convection
boundary conditions.   In this paper a method is presented of solving simultaneously
for temperature dependence of thermal conductivity, heat capacity, rate of internal
heat generation and surface heat transfer coefficients.   A three time level
difference scheme is used for the time marching which obviates the necessity for
iterations within a time step.

In the general formulation, latent heat effects are approximated by rapid variations
of heat capacity within a narrow temperature range as shown in Figure 1.   If the
phase change occurs with almost no variation it is convenient to define a new variable
H (enthalpy) and then estimate average heat capacity values based on well known
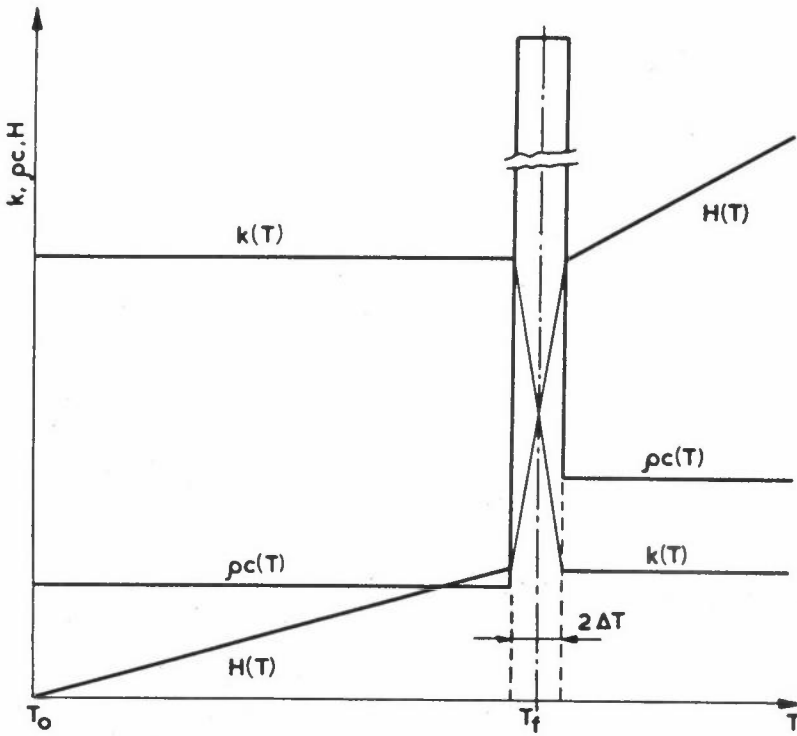enthalpy properties.

**Figure 1: Variation of thermal conductivity, heat capacity and enthalpy with temperature**

The phenomenon of heat and mass transfer in porous bodies is in many cases not accompanied by filtration flows of liquid and/or vapour. Extensive reviews of typical examples are quoted in the literature (Bonacina and Comini 1971; Little 1961; Luikov 1966) and a widely accepted mathematical model has been established through the use of the thermodynamics of irreversible processes (Luikov 1966). The resulting system of coupled partial differential equations are extremely difficult to solve even for conditions of constant total pressure which is the case considered here. In consequence, very few analytical solutions are available for comparison with the finite element model. However, the cases considered in this paper indicate very favourable agreement between numerical and analytical results.

## MATHEMATICAL MODELS

### Non-linear heat conduction with phase change

The problem considered in this paper is governed, in a region $\Omega$, by a quasilinear parabolic equation of the following form:-

$$\frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right) + \frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right) + \frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right) + Q = \rho c\frac{\partial T}{\partial t} \tag{1}$$

with the following boundary conditions:

$$T = T_w \quad \text{on } \Gamma_1 \tag{1a}$$

and

$$k\left(\frac{\partial T}{\partial x}\ell_x + \frac{\partial T}{\partial y}\ell_y + \frac{\partial T}{\partial z}\ell_z\right) + q + q_c + q_r = 0 \quad \text{on } \Gamma_2 \tag{1b}$$

where k is the thermal conductivity, $\rho c$ is the heat capacity, Q is the rate of internal heat generation and all are dependent on temperature. The terms $\ell_x$, $\ell_y$ and $\ell_z$ are the direction cosines of the outward normal to the boundary surface while q, $q_c$ and $q_r$ represent the imposed heat flux and the rates of heat flow per unit area due to convection and radiation. The following relationships are usually written:-

$$q_c = \alpha_c(T - T_{ac}) \tag{2a}$$

and

$$q_r = \varepsilon\sigma(T^4 - T_{ar}^4) = \alpha_r(T - T_{ar}) \tag{2b}$$

where $\alpha_c$ is the convective heat transfer coefficient, $\varepsilon$ is the emissivity, $\sigma$ is the Stefan constant, $\alpha_r$ is a parameter related to the effects of radiation by the expression

$$\alpha_r = \varepsilon\sigma(T^2 + T_{ar}^2)(T + T_{ar}) \tag{2c}$$

and $T_{ac}$ and $T_{ar}$ are the equilibrium temperatures for which no convection or radiation occurs.

In general k, $\rho c$, Q and $\alpha_r$ are thus temperature dependent functions with k, $\rho c$ and Q being known numerically and $\alpha_r$ being expressed explicitly by equation (2c).

## Coupled heat and mass transfer in porous bodies

If the assumption of constant pressure throughout the moist body is correct then the exchange of heat and mass is described by the following equations (Luikov and Mikhailov 1965).

$$\rho c_q \frac{\partial T}{\partial t} = \lambda_q \nabla^2 T + E r \rho c_m \frac{\partial u}{\partial t} \tag{3a}$$

$$\rho c_m \frac{\partial u}{\partial t} = \lambda_q \delta \nabla^2 T + \lambda_m \nabla^2 u \tag{3b}$$

where T and u are the heat and mass transfer potentials, $\rho c$ the heat capacity, $\lambda$ the conductivity, E the ratio of the vapour diffusion coefficient to the coefficient of the total diffusion of vapour and liquid and $\delta$ is the temperature gradient coefficient.

The problem as defined by equation (3) can be re-written in a generalised form as a system of interconnected parabolic equations:

$$C_q \frac{\partial T}{\partial \theta} = L_q \left(\frac{\partial^2 T}{\partial X^2} + \frac{\partial^2 T}{\partial Y^2} + \frac{\partial^2 T}{\partial Z^2} + L_\varepsilon \left(\frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2} + \frac{\partial^2 U}{\partial Z^2}\right)\right) \tag{4a}$$

$$C_m \frac{\partial U}{\partial \theta} = L_\delta \left(\frac{\partial^2 T}{\partial X^2} + \frac{\partial^2 T}{\partial Y^2} + \frac{\partial^2 T}{\partial Z^2}\right) + L_m \left(\frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2} + \frac{\partial^2 U}{\partial Z^2}\right) \tag{4b}$$

with the following boundary conditions

$$T = T_w \quad \text{on } \Gamma_1 \tag{4c}$$

$$L_q \left(\frac{\partial T}{\partial X}\gamma_x + \frac{\partial T}{\partial Y}\gamma_y + \frac{\partial T}{\partial Z}\gamma_z\right) + J_q + A_q(T - T_a) + A_\varepsilon(U - U_a) = 0 \quad \text{on } \Gamma_1 \tag{4d}$$

$$U = U_w \quad \text{on } \Gamma_3 \tag{4e}$$

$$L_m\left(\frac{\partial U}{\partial X}\gamma_x + \frac{\partial U}{\partial Y}\gamma_y + \frac{\partial U}{\partial Z}\gamma_z\right) + J_m + A_\delta(T - T_a) + A_m(U - U_a) = 0 \quad \text{on } \Gamma_4 \tag{4f}$$

In the above expressions $T = T/T_a$, $U = u/u_a$, $\theta = t/t_o$, $X = x/e$, $Y = y/e$, $Z = z/e$ are dimensionless variables, the C's are generalised capacities, whilst the L's, A's and J's can be thought of respectively as generalised transfer coefficient and specific fluxes.

Lewis et al (1974) have shown that equations (4) can be made symmetrical by a suitable definition of the coefficients, and it is also convenient to write equation (4d) and (4f) in the following form:

$$L_q\left(\frac{\partial T}{\partial X}\gamma_x + \frac{\partial T}{\partial Y}\gamma_y + \frac{\partial T}{\partial Z}\gamma_z\right) + J_q^* = 0 \quad \text{on } \Gamma_2 \tag{5a}$$

$$L_m\left(\frac{\partial U}{\partial X}\gamma_x + \frac{\partial U}{\partial Y}\gamma_y + \frac{\partial U}{\partial Z}\gamma_z\right) + J_m^* = 0 \quad \text{on } \Gamma_4 \tag{5b}$$

where $J_q^*$ and $J_m^*$ are defined by the following expressions:

$$J_q^* = J_q + A_q(T - T_a) + A_\varepsilon(U - U_a)$$

$$J_m^* = J_m + A_\delta(T - T_a) + A_m(U - U_a)$$

and

$$L_\varepsilon = L_\delta = \varepsilon r \lambda_m \, \delta / \lambda_q$$

## FINITE ELEMENT FORMULATION

The unknown functions in both the problems are approximated throughout the solution domain by the relationships:

$$T \cong \bar{T} = \sum_{r=1}^{K} N_r(X,Y,Z) T_r(\theta) = \underset{\sim}{N} \bullet \underset{\sim}{T} \tag{6a}$$

and

$$U \cong \bar{U} = \sum_{r=1}^{K} N_r(X,Y,Z) U_r(\theta) = \underset{\sim}{N} \bullet \underset{\sim}{U} \tag{6b}$$

where $N_r$ are the usual shape functions defined piecewise element by element, $T_r$ (or $\underset{\sim}{T}$) and $U_r$ (or $\underset{\sim}{U}$) being the nodal parameters.

The simultaneous equations, allowing the solution for K values of $T_r$ in the case of the heat conduction problem and for 2K values of $T_r$ and $U_r$ in the coupled heat and mass transfer problem are obtained using Galerkin's method as shown by Lewis and Garner (1972) and Zienkiewicz and Parekh (1970).

In the case of the non-linear heat conduction problem the discretized equations can be written in matrix form as:

$$K \underset{\sim}{T} + C \underset{\sim}{\dot{T}} + \underset{\sim}{F} = 0 \tag{7}$$

where typical matrix elements are

$$K_{ij} = \int_{\Omega^e} k \left( \frac{\partial N_i}{\partial X} \frac{\partial N_j}{\partial X} + \frac{\partial N_i}{\partial Y} \frac{\partial N_j}{\partial Y} + \frac{\partial N_i}{\partial Z} \frac{\partial N_j}{\partial Z} \right) d\Omega \;\; + \;\; \int_{\Gamma_2^e} (\alpha_c + \alpha_r) N_i N_j d\Gamma \tag{8}$$

$$C_{ij} = \int_{\Omega^e} \rho c \, N_i N_j d\Omega \tag{9}$$

and

$$F_i = -\int_{\Omega^e} N_i Q d\Omega + \int_{\Gamma_2^e} N_i (q - \alpha_c T_{ac} - \alpha_r T_{ar}) d\Gamma \tag{10}$$

The equations for the coupled heat and mass transfer problem are more complex but can be transformed into a pseudo 'one-degree of freedom' problem. The resulting equations written in matrix form are:

$$K \underset{\sim}{\emptyset} + C \underset{\sim}{\dot{\emptyset}} + \underset{\sim}{J} = 0 \tag{11a}$$

where K and C are 2K x 2K symmetric matrices

$$K = \begin{bmatrix} K_q & K_\varepsilon \\ K_\delta & K_m \end{bmatrix} \tag{11b}$$

$$C = \begin{bmatrix} C_q & 0 \\ 0 & C_m \end{bmatrix} \tag{11c}$$

$\underset{\sim}{\emptyset}$ and $\underset{\sim}{J}$ are 2K vectors

$$\underset{\sim}{\emptyset} = [\underset{\sim}{T}, \underset{\sim}{U}]^T \tag{11d}$$

$$\underset{\sim}{J} = [\underset{\sim}{J}_q, \underset{\sim}{J}_m]^T \tag{11e}$$

The matrices $K_q$, $K_\varepsilon$, $K_\delta$ and $K_m$ are of the following form

$$K_{ij} = \int_{\Omega^e} L(\frac{\partial N_i}{\partial X} \frac{\partial N_j}{\partial X} + \frac{\partial N_i}{\partial Y} \frac{\partial N_j}{\partial Y} + \frac{\partial N_i}{\partial Z} \frac{\partial N_j}{\partial Z}) d\Omega \tag{12a}$$

The $C_q$ and $C_m$ matrices are of the form

$$C_{ij} = \int_{\Omega^e} C N_i N_j d\Omega \tag{12b}$$

and the 'load' vectors $\underset{\sim}{J}_q$ and $\underset{\sim}{J}_m$ are as follows:

$$(J_i)_q = \int_{\Gamma^e} (J_q^* + J_m^* L_\varepsilon/L_m) N_i d\Gamma \tag{12c}$$

$$(J_i)_m = \int_{\Gamma^e} (J_m^* + J_q^* L_\delta/L_q) N_i d\Gamma \tag{12d}$$

The equations (11a) are non-linear as the generalised fluxes $J_q^*$ and $J_m^*$ depend on the values of the potentials at the external nodes, but the non-linearity is confined to the vector $\underset{\sim}{J}$.

The time integration scheme for both these types of problems utilizes a three-level method first proposed by Lees (1966). The governing equations are approximated as follows:-

$$K_t(\underset{\sim}{T}_{t+\Delta t} + \underset{\sim}{T}_t + \underset{\sim}{T}_{t-\Delta t})/3 + C_t(\underset{\sim}{T}_{t+\Delta t} - \underset{\sim}{T}_{t-\Delta t})/(2\Delta t) + \underset{\sim}{F}_t = 0 \qquad (13)$$

This, after some algebra, results in the following recurrence formula for final integration:

$$\underset{\sim}{T}_{t+\Delta t} = -\left[K_t + \frac{3}{2\Delta t} C_t\right]^{-1}(K_t\underset{\sim}{T}_t + K_t\underset{\sim}{T}_{t-\Delta t} - \frac{3}{2\Delta t} C_t\underset{\sim}{T}_{t-\Delta t} + 3\underset{\sim}{F}_t) \qquad (14)$$

This method has the advantage of not requiring iteration as the matrices K, C and vector $\underset{\sim}{F}$ are evaluated at the mid-time interval and known values of $\underset{\sim}{T}$ can be obtained for any two consecutive time steps. This algorithm has also been shown to be unconditionally stable and convergent in the context of both finite difference and finite element methods.

The heat conduction problems solved in the content of this paper are non-linear and also include a phase change as the melting or freezing points are reached. Therefore, attention has to be paid to the sharp peak present in the heat capacity versus temperature curve. This problem is circumvented by interpretation of enthalpy, which is a smooth function of temperature even in the phase change zone (see Figure 1), and is given by:

$$H = \int_{To}^{T} \rho c \, dT \qquad (15)$$

Thus, the heat capacity values at the integrating points can be approximated by determining the gradient of enthalpy with respect to temperature, ie

$$<\rho c>_{x,y,z} \cong \frac{1}{3} \left(\frac{\partial H}{\partial x} \frac{\partial x}{\partial T} + \frac{\partial H}{\partial y} \frac{\partial y}{\partial T} + \frac{\partial H}{\partial z} \frac{\partial z}{\partial T}\right) \qquad (16)$$

## RESULTS

The first example to be investigated was the solidification of an infinite slab of liquid which has been solved exactly by several authors (Luikov 1968).    The problem is two dimensional and has thermal dependent heat capacity and thermal conductivity.

The heat capacity is assumed to vary in a discontinuous manner (as a pulse) between the liquid and solid phases while the thermal conductivity varies linearly as shown in Figure 1.    The progress of the freezing front and the spatial temperature distributions are compared with analytical results in Figures 2 and 3.
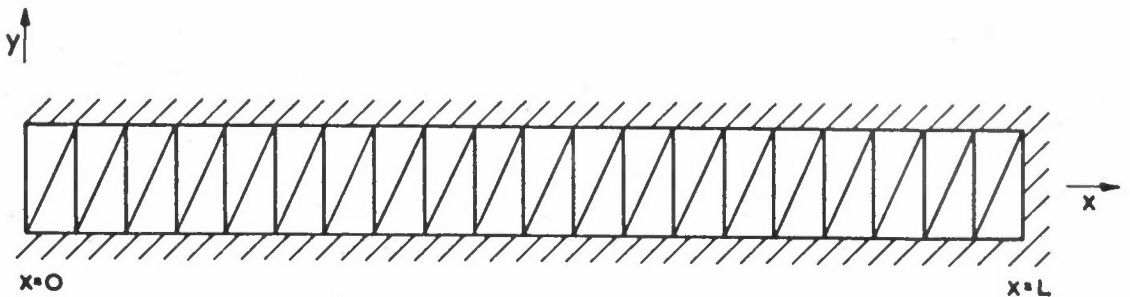
Figure 2(a):  Solidification of a slab in liquid — finite element mesh. Non conductive boundaries are assumed throughout except at the face x = 0. Slab thickness L = 1m
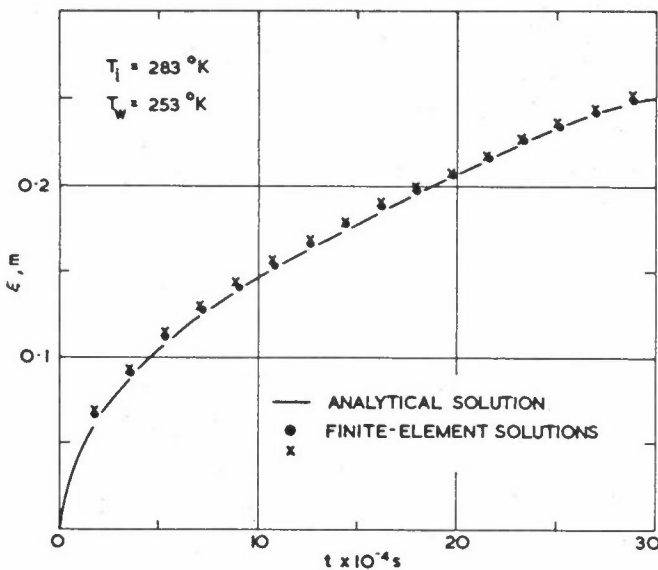
Figure 2(b): Solidification of a slab in liquid — progress of the freezing front

The position of the freezing front is given by the isotherm $T_f - \Delta T$.    The initial temperature was taken to be above freezing as previous investigations have shown that such conditions are more critical numerically than those with water initially at the freezing point.    The phase interval $2\Delta T$ was assumed as 0.5K but the results proved to be relatively insensitive to this value.

Figure 3: Temperature distributions at different time values during the freezing process referred to in Figure 2.
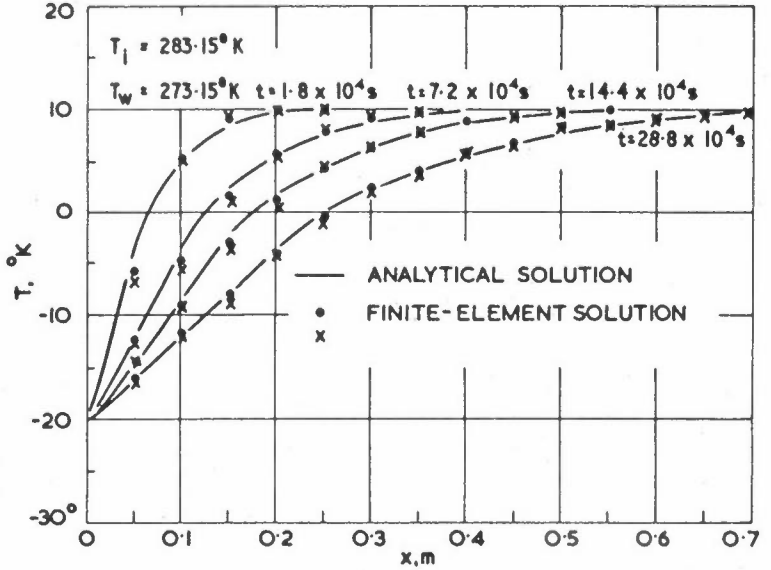
Since the slab is of finite dimensions the comparison with the theory was terminated when temperature changed appreciably at X = L. It was found that the results agreed very well with the theory with maximum errors being less than two per cent.



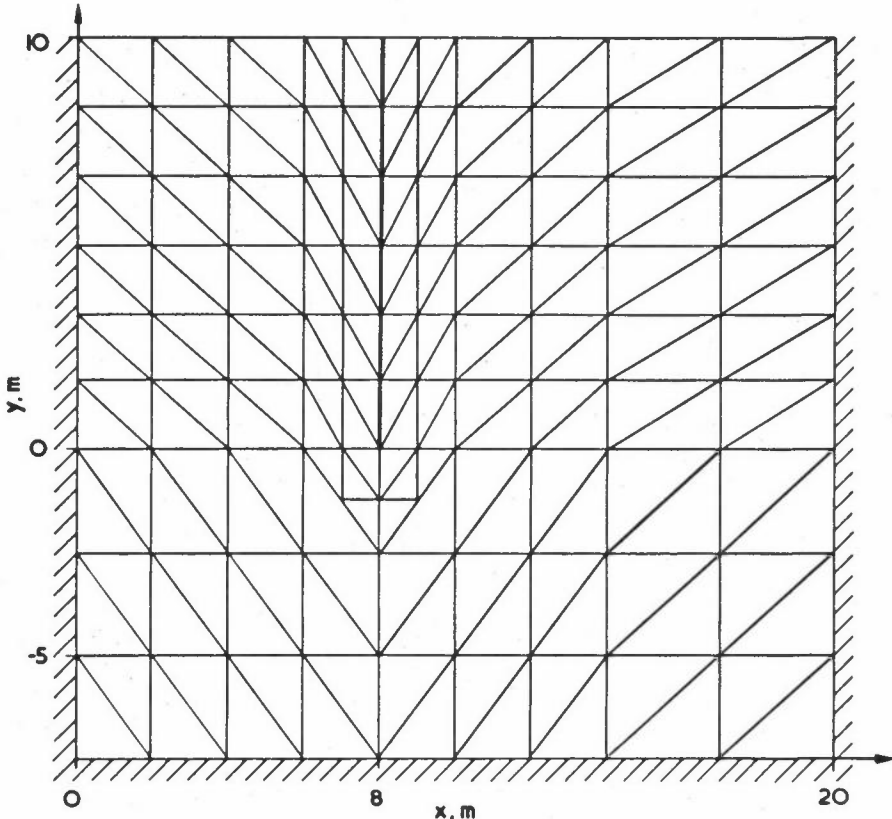Figure 4(a): Freezing of a moist soil (sand). Transverse section illustrating the finite element mesh used. The thick line corresponds to the freezing sheath. The infinite boundary is replaced by a finite one which is assumed to be non-conductive. Convection heat transfer takes place at the surface.

The next example concerns a practical problem of interest in the civil engineering industry for which no exact solutions exist. The ground region inside two lengths of cooling pipes is frozen in order to stabilise the soil. The time progression of the freezing front along with the relevant aspects of the problem are presented in Figure 4. The same material properties are used as given by Aguirre-Peunte et al (1967).

In the case of the heat and mass transfer problem two examples were run in order to assess the accuracy of the model. The first example concerns a slab, insulated at the face X = 0, with convective boundary conditions at the face X = x/e = 1. The moisture and temperature values in the slab are assumed to be zero initially with constant values of the equilibrium transfer potentials in the surrounding medium.
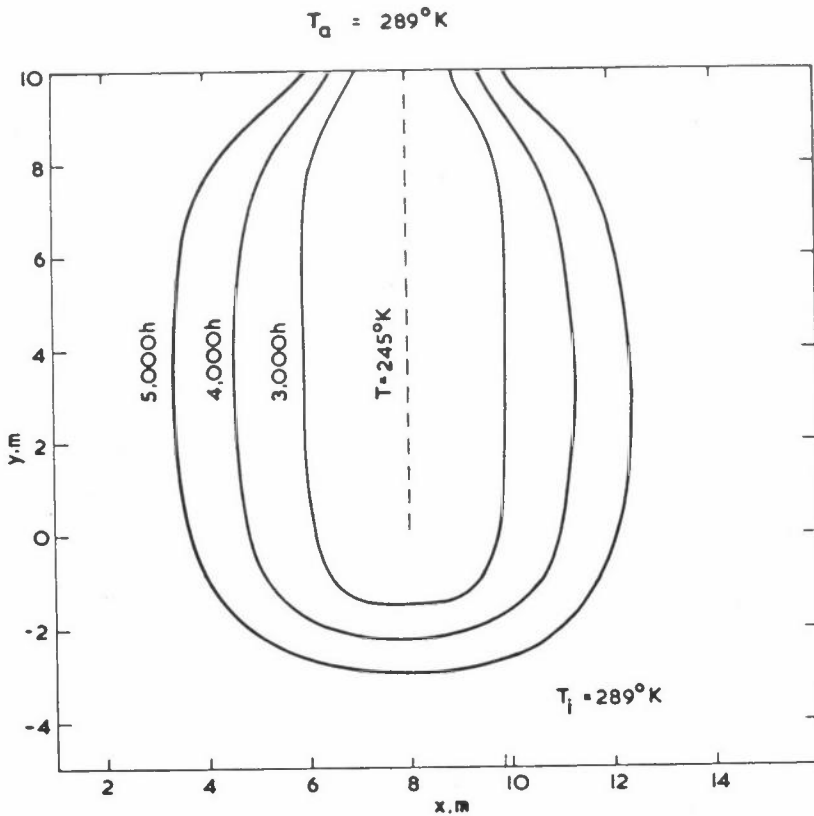


Figure 4(b): Freezing front positions as computed by the finite element method. The freezing sheath is maintained at 245°K. External temperature is 289°K and the convective heat transfer coefficient is $\alpha$ = 20W/m² K. Material properties (Aguirre-Peunte et al, 1967): k = 2.32 W/mK; $\rho c$ = 2.04 x 10⁶ J/m³ K (frozen) k = 1.65 W/mK; $\rho c$ = 2.815 x 10⁶ J/m³ K (unfrozen) Latent heat effect $\lambda$ = 120 x 10⁶. Assumed width of the phase change interval $2\Delta T$ = 1K

The analytical solution as given by Luikov and Mikhailov (1965) is discussed in detail elsewhere (R W Lewis et al 1974) and will not be repeated here. The finite

element mesh used is shown in Figure 5 and essentially represents a two-dimensional domain as the upper and lower boundaries are non conductive. The comparison shown in Figure 6 demonstrates good agreement between the analytical and finite element solutions with errors never greater than one per cent.
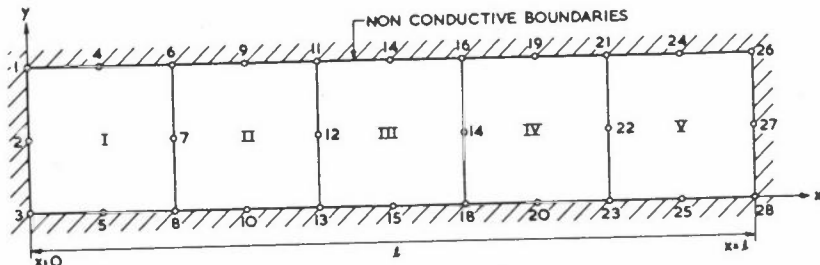


Figure 5: "One-dimensional" finite element mesh used in the numerical examples. Five parabolic isoparametric elements are employed.
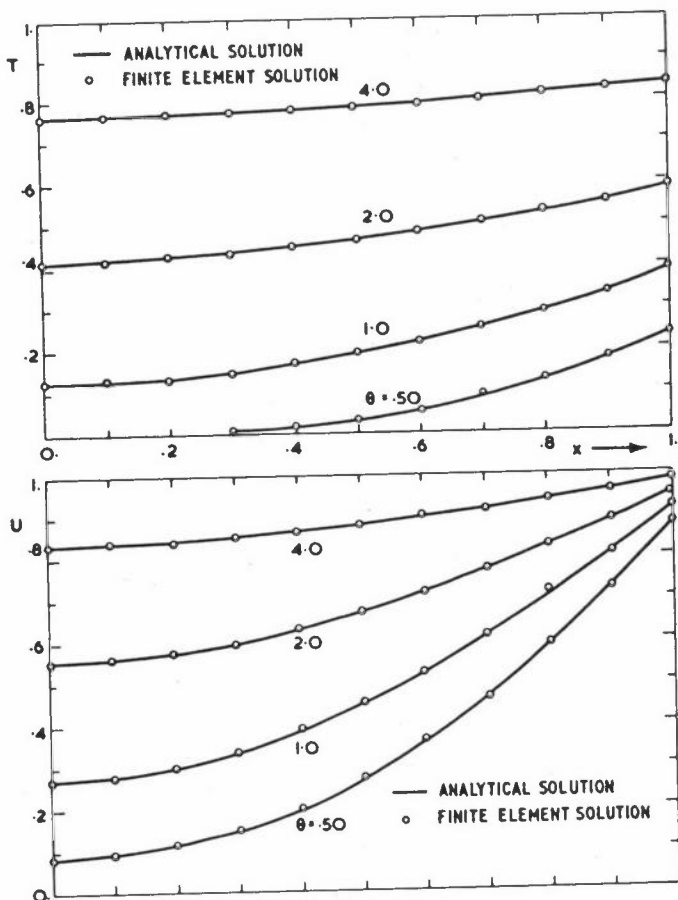


Figure 6: Comparison between analytical and finite element solutions for a one-dimensional heat and mass transfer problem in the case: Lu = 0.3, $\epsilon$ = 0.5, Ko = 1.2, Psn = 0.5, $Bi_q$ = 1.0, $Bi_m$ = 10.0

The second example utilised the same mesh and initial conditions as for the first problem but boundary conditions of the first kind are assumed at the face X = x/e = 1. Again the finite element results are shown to agree well with the analytical solution as demonstrated in Figure 7.
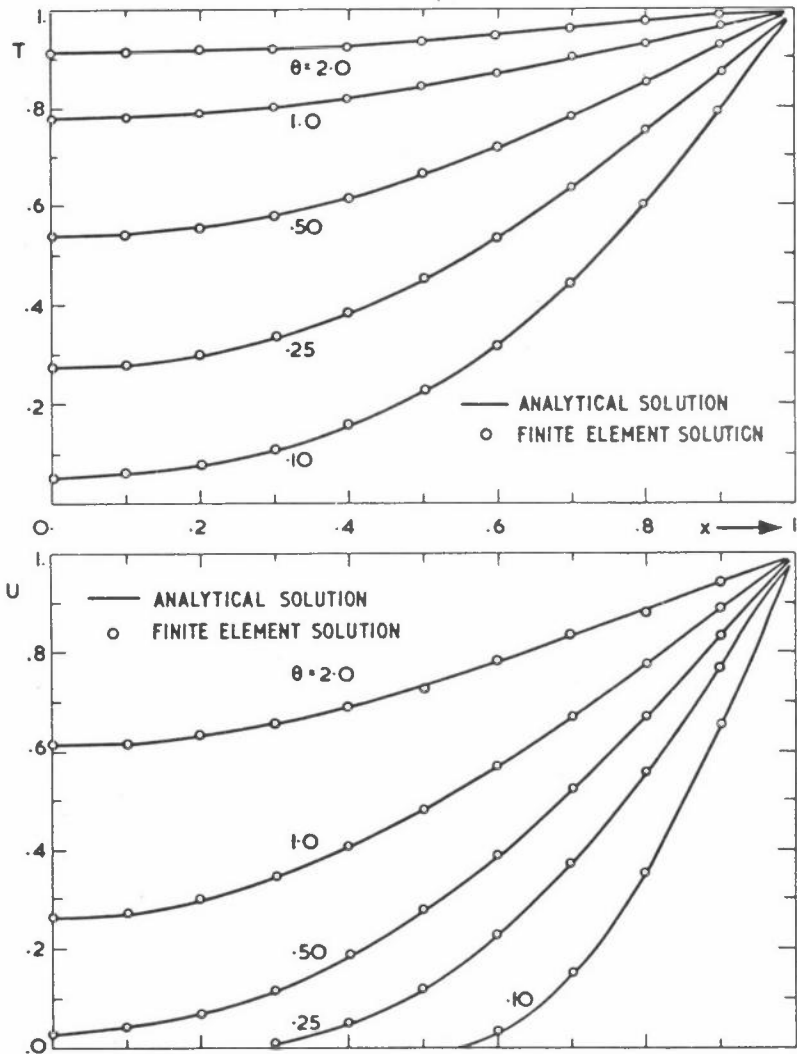


Figure 7: Comparison between analytical and finite element solutions for a one-dimensional heat and mass transfer problem in the case: Lu = 0.3, $\epsilon$ = 0.5, Ko = 1.2, Psn = 0.5, $Bi_q = Bi_m \simeq \infty$

- 55 -

## CONCLUSIONS

The approach to non-linear heat conduction with phase change and heat and mass transfer problems have been shown to be amenable to solution by the finite element method.

The versatility of the method in dealing with complicated boundaries and/or physical property variations along with the phase change relationship opens the door to many practical applications.   In particular, the problems of heat and mass exchanges in buildings and determination of thermal stresses in castings etc, can now be tackled effectively.

## ACKNOWLEDGEMENTS

The author wises to acknowledge the valuable contributions of Professor O C Zienkiewicz,  Dr G Comini and Mr C Humpheson towards the research presented in this paper.

## REFERENCES

AGUIRRE-PEUNTE J, Le FUR B and SZANTO I, "Congelation unidimensionele d'un echantillon de sable sature d'eau" in Phenomenes de Transport avec Changement de Phase Dan les Milieux Poreux au Colloidaux, 149-165, Ed CNRS Paris (1967)

BONACINA C and COMINI G, "Mass and heat transfer during cooling, freezing and cold storage of foodstuffs", Bull IIF 51, 1539-1558 (1971)

LEES M, "A linear three-level difference scheme for quasilinear parabolic equations", Maths Comp 20, 516-622 (1966)

LEWIS R W, COMINI G and HUMPHESON C, "Finite element formulation of some heat and mass transfer problems", Submitted to Int Journal for Heat and Mass Transfer, (February 1974)

LEWIS R W and GARNER R W, "Finite element solution of coupled electrokinetic and hydrodynamic flow in porous media", Int J Num Meth Eng 5, 41-55 (1972)

LITTLE A L, Foundations, Edward Arnold, London (1961)

LUIKOV A V, Heat and Mass Transfer in Capillary-porous Bodies, Pergamon, Oxford (1966)

LUIKOV A V, Analytical Heat Diffusion Theory, Academic Press, New York (1968)

LUIKOV A V, and MIKHAILOV Y A, Theory of Energy and Mass Transfer, Pergamon, Oxford (1965)

WILSON E L and NICKELL R E, "Application of the finite element method to heat conduction analysis", Nuclear Eng Design 4, 276-286 (1966)

ZIENKIEWICZ O C, The Finite Element Method in Engineering Science, 2nd edition, McGraw-Hill, London (1971)

ZIENKIEWICZ O C, and PAREKH C J, "Transient field problems: two dimensional and three dimensional analysis by isoparametric finite elements", Int J Num Meth Eng 2, 61-71 (1970)

# A STUDY OF THE MHD STABILITY OF VARIOUS PINCH CONFIGURATIONS USING THE FINITE ELEMENT METHOD

L R T Gardner
G A Gardner

School of Mathematics and Computer Science
University College of North Wales
Bangor
LL57 2UW

## INTRODUCTION

The aim of the research effort in controlled thermonuclear fusion is the containment and heating of a plasma for a time sufficiently long to permit fusion of the deuterium nuclei. Various plasma containment schemes have been studied and their equilibria examined. Equilibrium, although a necessary condition, is insufficient since for controlled fusion to be attained we must also have stability.

Early analytic studies (Shafranov 1958; Tayler 1957a; Tayler 1957b) of plasma stability concentrated on the linear pinch configuration in which a plasma column of circular cross section with an axially flowing current is compressed and contained by the $J \times B$ force acting radially inwards. This analytic work has recently been generalised and extended by Shafranov (1970) and our numerical calculation for the linear pinch will be compared, where possible, with his results.

The linear pinch is an example of an open ended system from which plasma can escape through its ends. This loss can be prevented by bending the two ends together thus forming a torus. A configuration of considerable current interest is an axially symmetric toroidal system (TOKAMAK) (Artsimovich 1972) in which hot plasma is contained by the magnetic field of a current flowing within it. A very strong axial magnetic field parallel to the current serves to suppress the principal hydromagnetic instabilities. Experiments (Artsimovich 1972; Artsimovich and Shafranov 1972) indicate that a further improvement in stability can be achieved by making the cross-section non circular. A toroidal belt pinch of rectangular cross-section ISAR IV is now operating at IPP Garching, and a larger similar device ISAR TX is under construction (Nuclear Fusion 1973). As an example of a pinch with non circular cross-section we will consider a cylindrical pinch of rectangular section. This is a

first step towards the numerical prediction of the stability ranges for a belt pinch of similar cross-section.

In the work reported here the stability of cylindrical pinches is studied using the (numerical) finite element method (Zienkiewicz 1971) to implement the hydromagnetic energy principle of ideal MHD theory (Rose and Clark 1965).

When a plasma containing a current $\underset{\sim}{J}$, magnetic field $\underset{\sim}{B}$ and of pressure p undergoes small pertubations $\underset{\sim}{\xi}$ away from a state of static equilibrium the change in the hydromagnetic energy of the system may be written (Bernstein et al 1958)

$$W(\underset{\sim}{\xi}) = \tfrac{1}{2}\int\left[Q^2 + \underset{\sim}{J}.\underset{\sim}{\xi} \times \underset{\sim}{Q} + \nabla.\underset{\sim}{\xi}(\underset{\sim}{\xi}.\nabla p) + \gamma p(\nabla.\underset{\sim}{\xi})^2\right]d\tau \qquad (1)$$

where the integral ranges over the volume of the device, $\underset{\sim}{Q}$ is the perturbed magnetic field and $\gamma$ the ratio of specific heats.

In the finite element method the volume of the pinch is divided into elements with chosen nodes. Within each element the perturbations are found by interpolating between the nodal values which are taken as parameters of a Lagrangian L

$$L = T - W(\underset{\sim}{\xi})$$

where T is the kinetic energy.

When the stationary values of the Lagrangian with respect to these variables is sought an eigenvalue problem is obtained. The stability of the configuration is determined by the eigenvalues.

## LINEAR PINCH MODEL

The linear pinch is modelled as a right circular cylindrical column of plasma, radius a, surrounded by an annular vacuum which is bounded on its outer surface, radius b, by a perfectly conducting wall. Let us introduce cylindrical polar coordinates $r, \theta, z$ about the axis of symmetry. The plasma is assumed to be incompressible, to have infinite conductivity and to contain a magnetic field $\underset{\sim}{B}$ which has no radial

component (Newcombe 1960). A current with density $\underset{\sim}{J} = \underset{\sim}{\nabla} \times \underset{\sim}{B}$ flows within the plasma. The vacuum magnetic field is $\underset{\sim}{B}_v$.

The equilibrium state of the plasma is governed by

$$\underset{\sim}{J} = \underset{\sim}{\nabla} \times \underset{\sim}{B}$$

$$\nabla p = \underset{\sim}{J} \times \underset{\sim}{B} \tag{2}$$

$$\nabla \cdot \underset{\sim}{B} = 0$$

The configuration has cylindrical symmetry hence all quantities are functions only of r. Equation (2) thus reduces to

$$J_r = 0 \qquad J_\theta = -\partial_r B_z, \qquad J_z = \frac{1}{r}\partial_r(rB_\theta)$$

$$\tag{3}$$

$$\partial_r p + B_z \partial_r B_z + \frac{B_\theta}{r} \partial_r(rB_\theta) = 0$$

Equation (2.3) is automatically satisfied.

We will assume further that $B_z$ is a constant so that there is no azimuthal current. In the vacuum p is zero hence

$$B_\theta = \frac{a}{r} B_{\theta a} \tag{4}$$

where $B_{\theta a}$ is the azimuthal field at the plasma surface. Any azimuthal field $B_\theta$ which is a function only of r is a possible equilibrium configuration for the plasma subject only to the boundary condition $B_\theta = B_{\theta a}$ at r = a.

When the plasma undergoes small perturbations $\underset{\sim}{\xi}$ the hydromagnetic energy W of the system is

$$W(\underset{\sim}{\xi}) = \frac{1}{2} \int_{\text{PLASMA}} [Q^2 + \underset{\sim}{J} \cdot \underset{\sim}{\xi} \times \underset{\sim}{Q}] d\tau + \frac{1}{2} \int_{\text{VACUUM}} Q_v^2 \, d\tau \tag{5}$$

where the quantities $\underset{\sim}{Q}$, $\underset{\sim}{Q}_v$ are the perturbed magnetic fields within the plasma and vacuum respectively·

For the plasma

$$\underset{\sim}{Q} = \underset{\sim}{\nabla} \times (\underset{\sim}{\xi} \times \underset{\sim}{B}) \qquad (6)$$

and for the vacuum

$$\underset{\sim}{Q}_v = \underset{\sim}{\nabla} \times \underset{\sim}{A} \qquad (7)$$

where A is the vector potential which must satisfy

$$\underset{\sim}{n} \times \underset{\sim}{A} = -(\underset{\sim}{n}.\underset{\sim}{\xi})\underset{\sim}{B} \qquad (8)$$

at the plasma vacuum interface r = a, and

$$\underset{\sim}{n} \times \underset{\sim}{A} = 0 \qquad (9)$$

at the bounding wall r = b.

It is convenient to extend the definition of $\underset{\sim}{\xi}$ into the vacuum region and to write the vector potential (Newcombe 1960)

$$\underset{\sim}{A} = \underset{\sim}{\xi} \times \underset{\sim}{B}_v$$

so that the perturbed magnetic field

$$\underset{\sim}{Q}_v = \underset{\sim}{\nabla} \times (\underset{\sim}{\xi} \times \underset{\sim}{B}_v)$$

has the same form as that in the plasma. The boundary condition (8) is automatically satisfied if $\underset{\sim}{B}(a) = \underset{\sim}{B}_v(a)$, and (9) is satisfied by requiring $\xi_r = 0$ at r = b.

To enable direct comparison with earlier work (Shafranov 1970; Boyd et al 1973; Takeda et al 1973) we have expressed the results of our calculations in terms of a variable nq defined by

$$nq = \frac{krB_z}{B_\theta}$$

or

$$nq_a = \frac{kaB_z}{B_{\theta a}}$$

a key parameter in discussions of pinch stability.

The displacements $\xi$ can be expressed in terms of normal modes in cylindrical polar coordinates (Newcombe 1960) by

$$\underset{\sim}{\xi} = \left[\xi_r(r), \; \xi_\theta(r), \; \xi_z(r)\right] \exp\left[im\theta + ikz + iwt\right] \tag{10}$$

where $\xi_r$ is real and $\xi_\theta$, $\xi_z$ imaginary.

The equation of continuity now becomes

$$\frac{1}{r} \, \partial_r(r\xi_r) + \frac{im\xi_\theta}{r} + ik\xi_z = 0 \tag{11}$$

For the perturbed motion the Lagrangian is

$$L = \omega^2 T - (W_p + W_v) \tag{12}$$

where

$$T = \frac{\pi}{2} \int \rho \, (\xi_r^2 + \xi_\theta^2 + \xi_z^2) r dr \tag{13}$$

The plasma hydromagnetic energy $W_p$, given by the first integral in equation (5), becomes now (Newcombe 1960)

$$W_p = \frac{\pi}{2} \int_0^a \left[\Lambda + \frac{k^2 r^2 + m^2}{r}(\zeta - \zeta_o)^2\right] r dr \tag{14}$$

where

$$\Lambda = \frac{1}{k^2 r^2 + m^2} \left[(krB_z + mB_\theta) \, \partial_r \xi_r + (krB_z - mB_\theta) \frac{\xi_r}{r}\right]^2$$

$$+ \left[(krB_z + mB_\theta)^2 - 2B_\theta \partial_r(rB_\theta)\right] \frac{\xi_r^2}{r^2} \tag{15}$$

$$\zeta_o = \frac{r}{k^2 r^2 + m^2} \left[(krB_\theta - mB_z) \, \partial_r \xi_r - (krB_\theta + mB_z) \frac{\xi_r}{r}\right] \tag{16}$$

and

$$\zeta = i(\xi_\theta B_z - \xi_z B_\theta) \tag{17}$$

The number of independent variables can easily be reduced for this problem by minimising the potential energy with respect to $\zeta$ simply by setting (Newcombe 1960)

$$\zeta = \zeta_o \tag{18}$$

in equation (14).   The displacements $\xi_\theta$, $\xi_z$ are related to $\xi_r$ by equations (11) and (18) and may be elminated from the Lagrangian which then depends only upon $\xi_r$.

To implement the finite element method the plasma vacuum radius is divided into N line elements in such a way that a node occurs at the plasma-vacuum interface, eg by points $r_o = 0$, $r_1$, $r_2$,...,$r_k = a$,..,$r_N = b$ and nodal parameters $\xi_o, \xi_1 \ldots, \xi_N$ assigned.   Within an element the displacement $\xi_r$ is obtained by linearly inter-polating between nodal values.

It is found, by considering the radial component of the perturbed field $Q_r$, that $\xi_r$ given by

$$\xi_r = \frac{-rQr}{iB_\theta(m+nq)} \tag{19}$$

can become infinite when $m + nq = 0$.   This event may occur since $nq = \frac{r^2}{a^2} nq(a)$ varies across the vacuum.   In situations of interest this does occur so an alternative variable $\eta$ defined by (Shafranov 1970)

$$\eta = -B_\theta(m + nq)\xi_r \tag{20}$$

will be used within the vacuum.   The boundary conditions to be satisfied by $\eta$ are

$$\eta(a) = -B_{\theta a}(m + nq(a))\xi_r(a)$$

and

$$\eta(b) = 0$$

The hydromagnetic energy for a typical plasma element i can, after integrating along the element, be written

$$W_p^i = (\xi_i, \xi_{i-1})K^i(\xi_i, \xi_{i-1})^T$$

and for a typical vacuum element

$$W_v^i = (\eta_i, \eta_{i-1})K^i(\eta_i, \eta_{i-1})^T$$

$K^i$ is a 2 × 2 matrix corresponding to the element stiffness matrix for solid structures (Zienkiewicz 1971).   Similarly for the kinetic energy of plasma element  i  we have

$$T^i = (\xi_i, \xi_{i-1}) M^i (\xi_i, \xi_{i-1})^T$$

where $M^i$ corresponds to the element mass matrix (Zienkiewicz 1971).

The total kinetic and potential energies are obtained by summing contributions from all elements. Assembling the element matrices together the Lagrangian is given by

$$L = \omega^2 \underset{\sim}{\delta}^T M \underset{\sim}{\delta} - \underset{\sim}{\delta}^T K \underset{\sim}{\delta} \tag{21}$$

where

$$\underset{\sim}{\delta}^T = (\xi_0, \xi_1 \cdots \xi_k, \eta_{k+1} \cdots \eta_n) \tag{22}$$

is an $N + 1$ vector. Node k occurs at the plasma-vacuum interface. The quantities M and K are $n + 1$ square matrices.

A matrix equation .

$$(\omega^2 M - K) \underset{\sim}{\delta} = 0 \tag{23}$$

is obtained when we seek a stationary value of the Lagrangian. The problem of MHD stability has become an eigenvalue problem. The positive eigenvalues $\omega^2$ give the frequency of a mode and the associated eigenvector $\underset{\sim}{\delta}$ the amplitude of nodal perturbations. A negative eigenvalue implies that the associated mode is unstable, with growth rate $\sqrt{|\omega^2|}$.

## LINEAR PINCH RESULTS

### CASE 1   Constant Current Density

When the current density $J_z$ is constant the magnetic field in the plasma is

$$\underset{\sim}{B} = (0, \frac{B_{\theta a} r}{a}, B_z) \tag{24}$$

and in the vacuum

$$\underset{\sim}{B}_v = (0, \frac{B_{\theta a} a}{r}, B_z) \tag{25}$$

where $B_{\theta a}$, $B_z$ are constant, so that nq is constant over the plasma. Through a normal mode analysis Shafranov (1970) has shown that when the axial wave number k is small

$$\xi_r \propto r^{-m-1} \tag{26}$$

and

$$\omega^2 = - \frac{B_{\theta a}^2}{4\pi\rho a^2} \left[ 2(m + nq) + \frac{2(m + nq)^2}{1 - (\frac{a}{b})^{-2m}} \right] \tag{27}$$

When the plasma extends out to the bounding wall so that a = b these expressions are replaced by (Shafranov 1970)

$$\xi_r \propto \frac{1}{r} J_m(z_m \frac{r}{a}) \tag{28}$$

$$\omega^2 = - \frac{B_{\theta a}^2}{4\pi\rho a^2} \left[ 2(m + nq) \frac{ka}{z_m} + (m + nq)^2 \right] \tag{29}$$

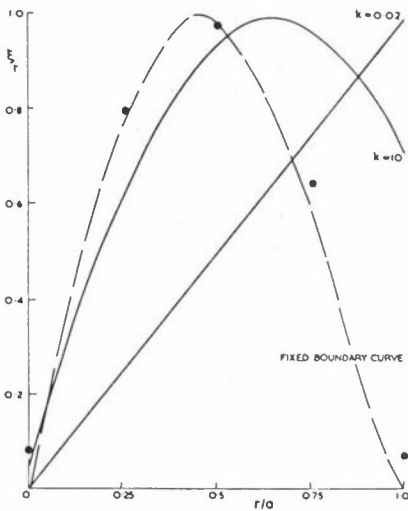where $z_m$ is the first zero of the mth order Bessel function $J_m(z)$.



Figure 1: Radial variation of the perturbation $\xi_r$ for m = -2 mode in a linear pinch with free boundary for various axial wave numbers k and a/b = 0.5. Points ● are for k = 50
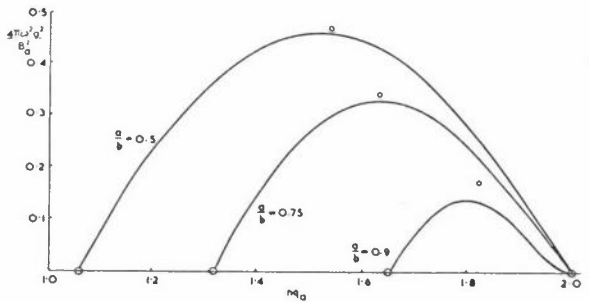
Figure 2: Square of the normalised growth rates for linear pinch with free boundary plotted against $nq_a = kaB_z/B_a$ for m = -2 mode, k = 0.2 and various values of a/b. The points o are Shafranov's an analytic results

We have determined the nodal amplitudes and growth rate for the mode m = -2, k = 0.2 (Boyd et al 1974) and have compared these with Shafranov's analytic expressions in Figures 1 and 2. Excellent agreement is found for various a/b ratios. Good

agreement is also obtained for the no vacuum case (Boyd et al 1974). We have also
determined nodal amplitudes for modes m = -1, -3, k = 0.2, a/b = 0.5. In the former
case (Boyd et al 1974) $\xi_r$ is constant over the plasma, and in the latter case $\xi_r$
varies as $r^2$, in complete agreement with equation (26). It was found that 16
elements along the plasma's vacuum radius were sufficient for convergence of the
solution.

## CASE 2    Exponential Current Density

In the latter stages of tokamak discharges it has been observed that the current
density becomes peaked on the cylinder axis. We have therefore studied current
densities of the form (Grossmann and Ortolani 1973)

$$J = J_0 \exp\left(-\frac{\alpha r}{a}\right)$$

with a plasma density variation $\rho = \rho_0 (1 - \frac{r^4}{a^4})$, typical of tokamaks

In Figure 3 the growth rate of mode m = -1, with $\alpha = 3$, is plotted as a function of
$nq_a$.



Figure 3:  The growth rate of the m = -1 mode
plotted as a function of $nq_a$ for the
current distribution $J = J_0 \exp(-3r/a)$
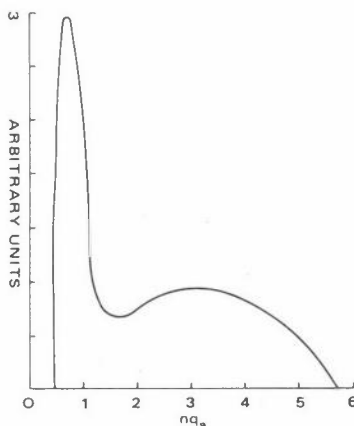
The position of maximum growth rate and the instabilities ranges agree with those
reported by Grossmann and Ortolani (1973). We also confirmed that the radial
distribution is peaked on the cylinder axis;  more highly peaked for larger $nq_a$
values.

For this case about 40 elements of equal size were needed to ensure convergence of
the solution. When a more judicious arrangement of elements was made the number
required for convergence was considerably reduced.

**CASE 3**  Bennett Pinch

It has been suggested (Morse and Friedberg 1970;  Harris and Kaneko 1971) that the pressure and field configuration of the Bennett pinch (Bennett 1934) are especially stable.

The density variation over the plasma is $\rho = \rho_0 \, (1 + \frac{r^2}{a^2})^{-2}$ and $\underset{\sim}{B} = (0, \frac{2B_{\theta a} ar}{a^2 + r^2}, B_z)$ where $B_{\theta a}$, $B_z$ are constants.  $\underset{\sim v}{B}$ is again given by equation (25), and the axial current by

$$J_z = \frac{4B_{\theta a} a^3}{(a^2 + r^2)^2}$$

We have determined the growth rate of the mode m = -1, k = 0.2 for various a/b ratios.   These calculations are given in Figure 4.   Analytically it has been shown that the Bennett pinch is stable if $nq \geq - \frac{10}{3}m$ (Harris and Kaneko 1971);  our results are in agreement with this conclusion.   The m = -1 mode of the Bennett pinch has the same ranges of instability as the constant current case (Shafranov 1970) and seemingly higher growth rates than those predicted by equation (27).   In these computations 16 radial finite elements were used.
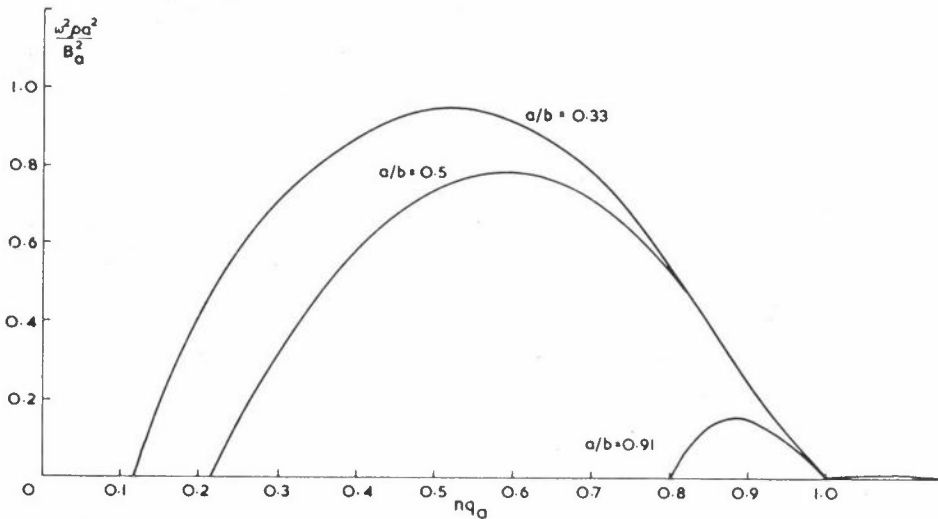


Figure 4:  Square of the normalised growth rate for the kink mode m = -1 of a Bennett pinch with free boundary plotted as a function of $nq_a$ for various values of a/b

## RECTANGULAR PINCH MODEL

The pinch is modelled as a right cylinder of rectangular section, sides 2a, 2b, bounded by a perfectly conducting wall; – there is no vacuum in this case. Rectangular cartesian coordinates are set up with the z-axis along the cylinder axis and the x, y axes parallel to cross-section sides. The plasma is assumed to have infinite conductivity and to contain an axially flowing current, and a magnetic field $\underset{\sim}{B} = (B_x, B_y, B_z)$ where $B_z$ is constant and the other two components are independent of z. The equilibrium distribution of magnetic fields etc was provided to us by Dr Glenn Bateman of IPP Garching in numerical form.

When the plasma undergoes small perturbations the hydromagnetic energy of the system is given by equation (1). The displacements may be expanded in normal modes by

$$\underset{\sim}{\xi}(x,y,z,t) = \left[ \xi_x(x,y), \xi_y(x,y), \xi_z(x,y) \right] \exp(ikz + i\omega t)$$

The vector's components are in general complex and will be written

$$\xi_x = \xi_{x1} + i\xi_{x2} \; ; \qquad \xi_y = \xi_{y1} + i\xi_{y2}; \qquad \xi_z = \xi_{z1} + i\xi_{z2}$$

In this general case we have six nodal parameters. This number is not reduced if we consider only the incompressible modes so this assumption is not made. For convenience we shall study only the kink modes, for which $\xi_z = 0$, so that the nodal parameters are reduced to 4.

The hydromagnetic energy is given by

$$W(\underset{\sim}{\xi}) = \frac{1}{2} \int_{-a}^{a} \int_{-b}^{b} \Lambda \; dx \; dy$$

where

$$\Lambda = (\xi_{x1}^2 + \xi_{x2}^2 + \xi_{y1}^2 + \xi_{y2}^2)\{k^2 B_z^2 + (\partial_x B_x)^2 + \partial_x B_y \partial_y B_x\}$$

$$+ (\xi_{x2}\partial_y\xi_{x1} - \xi_{x1}\partial_y\xi_{x2})(-2kB_y B_z) + (\xi_{x1}\partial_y\xi_{x1} + \xi_{x2}\partial_y\xi_{x2})(2B_y\partial_y B_y)$$

$$+ (\xi_{x2}\partial_y\xi_{y1} - \xi_{x1}\partial_y\xi_{y2})(-2kB_x B_z) + (\xi_{x1}\partial_y\xi_{y1} + \xi_{x2}\partial_y\xi_{y2})(-2B_x\partial_y B_y - JB_y)$$

$$+ (\xi_{y2}\partial_x\xi_{y1} - \xi_{y1}\partial_x\xi_{y2})(2kB_y B_z) + (\xi_{y1}\partial_x\xi_{x1} + \xi_{y2}\partial_x\xi_{x2})(-2B_y\partial_x B_x + JB_x)$$

$$+ (\xi_{y2}\partial_x\xi_{x1} - \xi_{y1}\partial_x\xi_{x2})(-2kB_x B_z) + (\xi_{y1}\partial_x\xi_{y1} + \xi_{y2}\partial_x\xi_{y2})(2B_x\partial_x B_x)$$

$$+ (\xi_{y1}\partial_y\xi_{x1} + \xi_{y2}\partial_y\xi_{x2})(-B_y\{\partial_x B_y + \partial_y B_x\}) + (\xi_{y1}\partial_y\xi_{y1} + \xi_{y2}\partial_y\xi_{y2})(2B_x\partial_x B_y)$$

$$+ (\xi_{x1}\partial_x\xi_{y1} + \xi_{x2}\partial_x\xi_{y2})(-B_x\{\partial_x B_y + \partial_y B_x\}) + (\xi_{x1}\partial_x\xi_{x1} + \xi_{x2}\partial_x\xi_{x2})(2B_y\partial_y B_x)$$

$$+ \{(\partial_y\xi_{x1})^2 + (\partial_y\xi_{x2})^2\}B_y^2 + \{(\partial_x\xi_{y1})^2 + (\partial_x\xi_{y2})^2\}B_x^2$$

$$+ (\partial_y\xi_{x1}\partial_y\xi_{y1} + \partial_y\xi_{x2}\partial_y\xi_{y2} + \partial_x\xi_{y1}\partial_x\xi_{x1} + \partial_x\xi_{y2}\partial_x\xi_{x2})(-B_x B_y)$$

$$+ \{(\partial_x\xi_{x1})^2 + (\partial_x\xi_{x2})^2\}(B_y^2 + B_z^2 + \gamma p) + \{(\partial_y\xi_{y1})^2 + (\partial_y\xi_{y2})^2\}(B_x^2 + B_z^2 + \gamma p)$$

$$+ (\partial_x\xi_{x1}\partial_y\xi_{y1} + \partial_x\xi_{x2}\partial_y\xi_{y2})(B_z^2 + \gamma p)$$

The kinetic energy T possessed by the plasma, density $\rho$, during the perturbations is calculated from

$$T = \frac{\omega^2}{4} \int_{-a}^{a} \int_{-b}^{b} \rho(\xi_{x1}^2 + \xi_{x2}^2 + \xi_{y1}^2 + \xi_{y2}^2) \, dy dx$$

and the Lagrangian of the perturbations formed.

The cross-section of the pinch was divided into rectangular finite elements with a node at each corner, there are thus 16 nodal parameters per element. In this implementation all x,y modes are considered at the same time, but obviously only the lowest modes will be determined accurately. Accuracy again depends on the number of finite elements used.

For the present configuration the relevant parameter $nq_o$ is given by

$$nq_o = \frac{2kB_z}{\pi} \left[ \int_0^b \frac{dy}{B_y(a,y)} + \int_a^Q \frac{dx}{B_x(x,b)} \right]$$

and all growth rates will be presented in terms of this variable.

## RECTANGULAR PINCH RESULTS

Here we report calculations for the rectangular pinch containing a uniform current density. Two cases have been considered which will have direct comparison with analytic studies being undertaken at IPP Garching.

(1)    At the rectangular boundary $\xi_n = 0$.

(2)    At the rectangular boundary $\xi_n$ is free.

For case (1) we have been unable to detect any instabilities so far.    Instability results for case (2) for a square pinch and a rectangular pinch, sides in ratio 1:3, are shown in Figure 5.    The results for $nq_a$ less than 1 are for a first order mode;   other results are for higher modes.    In all cases the growth rate of the most unstable mode is plotted.    The computations indicate that growth rates for the rectangular pinch are higher than those for the square pinch.    In all the calculations designated 0,+ 4×4 finite elements were used.    For the square pinch convergence of the solution was tested for $nq_o = 0.5$ by running cases with 2×2 and 3×3 elements. Growth rates for these cases are shown by × in Figure 5.
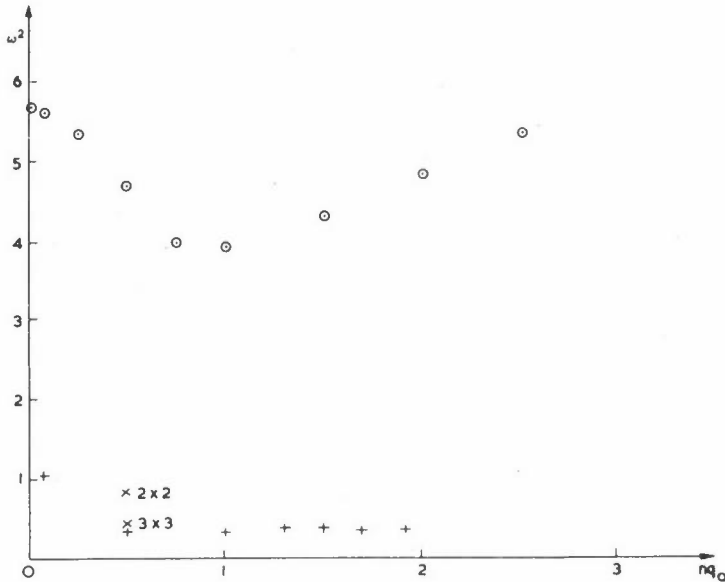


Figure 5:  Square of the growth rate for kink modes of a square pinch +, and rectangular pinch (1:3) ⊙, plotted as functions of $nq_o$. m = -1, k = 0.2, $\gamma$ = 5/3.  4 x 4 finite elements. Convergence of calculation for $nq_o$ = .5 is demonstrated for the square pinch

Neither case is really physical;   no hot plasma can be allowed to touch the container walls, but is constrained by a magnetic field, which extends into the vacuum and tends to improve stability.    In this situation the equilibrium form of the plasma is not strictly rectangular.    However the true situation probably lies somewhere between the two cases considered above.    The next step in this program is to study the realistic equilibrium configuration.

## REFERENCES

ARTSIMOVICH L A, Nuclear Fusion 12, 215 (1972)

ARTSIMOVICH L A and SHAFRANOV V D, JETP Letters, 15, 51 (1972)

BENNETT W J,  Phys Rev 45, 890 (1934)

BERNSTEIN I B,  FRIEMANN E A,  KRUSKAL M D,  KULSRUD R M,  Proc Roy Soc.
    A244, 17 (1958)

BOYD T J M,  GARDNER G A  and GARDNER L R T,  Finite Element Methods
    in Fluid Mechanics, Chapter 25, J Wiley, London (1974)
    Finite Element Methods in Flow problems, 309, UAH Press (1974)

BOYD T J M,  GARDNER G A  and  GARDNER L R T,  Nuclear Fusion 13, 764 (1973)

GROSSMANN W  and  ORTALANI S,  European Conference on Controlled Fusion
    and Plasma Physics, Moscow, Vol 1, 83 (1973)

HARRIS F G  and  KANEKO S,  Phys Fluids, 14, 987 (1971)

MORSE R L  and FRIEDBERG J P,  Phys Fluids 13, 533 (1970)

NEWCOMBE W A,  Ann Phys 10, 232 (1960)

NUCLEAR FUSION:  World Survey of Major Facilities in Controlled Fusion Research (1973)

ROSE D J  and CLARK M,  Plasma and Controlled Fusion,   MIT Press (1965)

SHAFRANOV V D,  Plasma Physics and the Problem of Controlled Thermonuclear
    Reactions, Vol 4, Pergamon, Oxford (1958)

SHAFRANOV V D,  Sov Phys Techn Phys 15, 175 (1970)

TAKEDA T,  SHIMOMURA Y,  OHTA M,  and YOSHIKAWA M, Phys Fluids 15, 2193 (1973)

TAYLER R J,  Proc Phys Soc B70, 31 (1957)a

TAYLER R J,  Proc Phys Tech Phys 15, 175 (1957)b

ZIENKIEWICZ O C,  The Finite Element Method in Engineering Science,
    McGraw-Hill  London (1971)

# APPLICATION OF THE FINITE ELEMENT METHOD TO METEOROLOGICAL PROBLEMS

M J P Cullen

Meteorological Office
London Road
Bracknell
Berkshire
RG12 2SZ

## EQUATIONS ARISING IN METEOROLOGICAL PROBLEMS

The type of equations that have to be solved in the meteorological forecasting problem can be typified by the shallow water equations

$$\frac{\partial u}{\partial t} + \underset{\sim}{u} \cdot \nabla \underset{\sim}{u} + \nabla \phi + 2 \underset{\sim}{\Omega}_\nu \times \underset{\sim}{u} = 0$$

$$\frac{\partial \phi}{\partial t} + \nabla (\phi \underset{\sim}{u}) = 0$$

Where the vector operations are two-dimensional. When the full three-dimensional problem is solved the scales of motion in the horizontal direction are completely different from those in the vertical and so a two dimensional problem can represent very well a large number of the motions present in the three dimensional case. This set of equations has the following important properties for numerical solution:

(1) An initial value problem has to be solved to evolve a given initial field into a future forecast field. In practical cases there are also forcing terms so that the evolution is controlled both by the initial data and by the forcing. It is debatable for how long the initial data has influence in the atmospheric case.

(2) The equations are nonlinear and can describe interactions between different scales of motion. In the three dimensional case the equations can produce discontinuities which are controlled in the atmosphere by the development of small scale turbulence.

(3) The equations describe wave type motions, both gravity waves where the pressure gradient balances the acceleration and slow moving waves due to the variation of the vertical component of $\Omega$ over the earth's surface.

(4)   The equations satisfy certain conservation laws which are important for
      long term integrations.

## RELATION OF THE FINITE ELEMENT METHOD TO OTHER METHODS OF SOLUTION

Historically, most solutions of these equations have been obtained by the finite
difference method.   Recently the Galerkin approach has been used with a basis of
wave functions $e^{ikx}$, or spherical harmonics for global integrations.   We consider
the advantages of these methods and the finite element method for this type of
problem:

(1)   If we consider the equation

$$\frac{\partial u}{\partial t} = Lu$$

and write the finite element projection as $Q$ , so that $Qu$ is some
approximation to u, then the error can be written as

$$Lu - QLQu = (I-Q)Lu + QL(I-Q)u$$

The first term on the right hand side is the error in making a discrete
approximation to Lu and can be studied by the usual methods to give an
order of convergence.   This error is unavoidable.   The second term
represents the divergence of the discrete solution actually obtained from
the discrete approximation to the true solution (see Strang & Fix (1973)).
It accumulates through time, and must be analysed as well.

(2)   In approaching a nonlinear problem interactions may lead to fields which
cannot be resolved by the discrete grid.   If only values at a discrete
set of points are used a field which has three wavelengths for every four
gridpoints will look identical to one with only one wavelength for every
four gridpoints.   If information propagates from point to point this gives
rise to errors known as 'aliasing'.   This error can be avoided by using
the Galerkin technique which uses information from the whole field.
Finite element Galerkin methods should do this while the usual finite
difference approach does not.

(3)   Most discrete schemes do not reproduce the phase speeds of waves correctly
and if the errors vary with wavelength wrong phase relationships develop in
the flow.   A scheme is thus required which reduces phase speed errors as

far as possible.

(4)     Certain conservation laws are obeyed automatically when the Galerkin method is used.     However, the full set is only obtained when the basis functions are closed under such operations as multiplication and differentiation or the extra manipulations needed to get the laws cannot be done.     Finite difference methods have to be specially designed to obey conservation laws and to satisfy a large number of them is very difficult.

(5)     The nonlinearity also involves a lot of computation so that in a Galerkin method as few inner products as possible must be non-zero.

These considerations suggest that a finite element Galerkin procedure should have several useful properties for this type of equation, but a simpler collocation procedure will probably not gain anything over the usual finite difference schemes.

## FOURIER ANALYSIS OF THE FINITE ELEMENT METHOD

Using piecewise linear interpolations in one dimension the finite element Galerkin approximation to

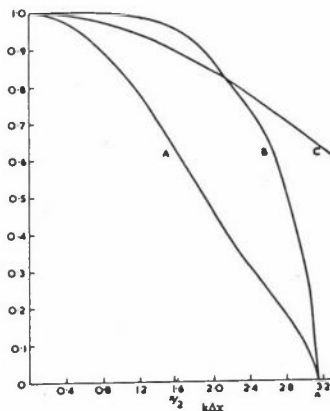$$\frac{\partial u}{\partial t} = c \frac{\partial u}{\partial x}$$

is

$$\frac{1}{6} \dot{u}_{n-1} + \frac{2}{3} \dot{u}_n + \frac{1}{6} \dot{u}_{n+1} = c \left(\frac{u_{n+1} - u_{n-1}}{2\delta x}\right) \tag{1}$$

The accuracy of this formula as opposed to the usual

$$\dot{u}_n = c \left(\frac{u_{n+1} - u_{n-1}}{2\delta x}\right)$$

is shown in Figure 1.



Figure 1: Accuracy of solution of $\partial u / \partial t = \partial u / \partial x$ with $u = e^{ikx}$
and exact line integration.
A:     Centred finite difference
B:     Linear finite element
C:     Finite difference with double resolution

In many cases greater accuracy can be obtained from (1) than by using the simpler formula with twice as many points.

Using the same interpolation scheme the finite element approximation to

$$\dot{u} = uv$$

gives

$$\frac{1}{6}\dot{u}_{n-1} + \frac{2}{3}\dot{u}_n + \frac{1}{6}\dot{u}_{n+1} = \frac{1}{12}(u_{n-1}v_{n-1} + u_{n-1}v_n + u_n v_{n-1}) +$$

$$\frac{1}{2}u_n v_n + \frac{1}{12}(u_n v_{n+1} + u_{n+1}v_n + u_{n+1}v_{n+1})$$

Figure 2 shows the result if $u = e^{ipx}$ and $v = e^{ip'x}$ in the cases $p=p'$ and $p'=0$. Thus interactions leading to wave-lengths too small to be resolved are strongly damped. A point collocation scheme would not do this and large aliasing errors would result.
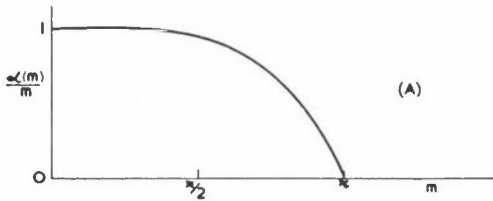


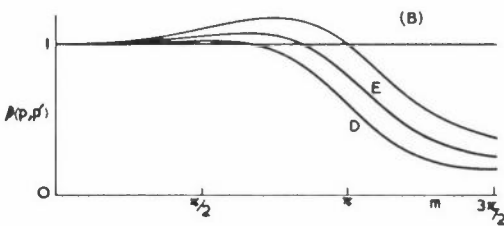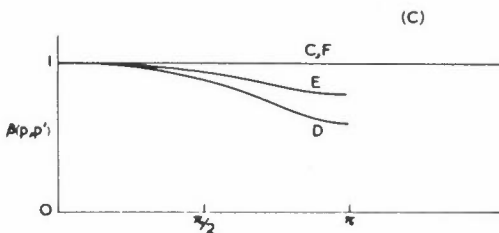Figure 2: (A) Accuracy of solution of $\partial u/\partial t = \partial u/\partial x$ with $u = e^{cmx}$ and $\delta x = 1$

(B) Accuracy of solution of $\partial u/\partial t = u^2$ with $u = e^{\ell px}$. The top curve is the usual finite element solution, curves E and D contain extra damping.

(C) Accuracy of solution of $\partial u/\partial t = uu_0$ ($u_0$ constant) showing effect of extra damping on this effectively linear term

In the case $u = u\dfrac{\partial v}{\partial x}$

with $u = e^{ip'x}$, $v = e^{ipx}$ we obtain

$$\dot{u} = \beta(p,p')\, e^{i(p+p')x}$$

where

$$\beta(p,p') = \frac{2(-\sin p' + 2\sin p + \sin(p+p'))}{4 + 2\cos(p+p')} \tag{2}$$

if the interpolation functions are substituted directly into the equation. However, if we use an approximation theory approach and seek, given the model values, best approximations to $u$ and $\dfrac{\partial v}{\partial x}$ and then perform the multiplications we can choose $\dfrac{\partial v}{\partial x}$ to be piecewise linear also, instead of being a piecewise constant. We then obtain a much more satisfactory representation of $u\dfrac{\partial v}{\partial x}$ and

$$\beta(p,p') = \frac{6\sin p}{4 + 2\cos p} \cdot \frac{3 + \cos p + \cos p' + \cos(p+p')}{4 + 2\cos(p+p')} \tag{3}$$

In some cases this gives the same result as (2), in the case $p = p' = \pi/2$ we obtain a value $0.96p$ from (3) and $0.6p$ from (2), the true value being $p$. This is found to be important when the nonlinear terms are an important part of the evolution.

We now use the Fourier method to analyse the error component $QL(I-Q)u$ discussed in the last section and consider the two operators $L = u_x$ and $L = uv$. Substitute $u = e^{ikx}$, then $Qu = \Sigma\, \alpha(k)\, e^{ikx_n}\chi_n$.

where

$$\alpha(k) = \frac{12}{4 + 2\cos k}\,\frac{1}{k^2}\,(1 - \cos k),$$

and we have again used linear elements. Then if $L = u_x$

$$QLu = \Sigma\, ik\, \alpha(k)\, e^{ikx_n}\chi_n$$

$$QLQu = \Sigma\, \frac{6\sin k}{4 + 2\cos k}\, \alpha(k)\, e^{ikx_n}\chi_n = \Sigma\, ik\,(1 + O(k^4))\alpha(k)\, e^{ikx_n}\chi_n$$

so that $QL(I-Q)u$ is of order $k^4$.

A similar calculation shows that with $L(u,v) = uv$ and $u = e^{ikx}$, $v = e^{i\ell x}$, $QL(u,v) - QL(Qu,Qv)$ is of order $k^4 + k^2\ell^2 + \ell^4$. Thus we have achieved fourth order accuracy with linear elements. Strang and Fix quote the result that $QL(I-Q)u$ with $L = u_x$ is only $O(k^3)$ using Hermite cubics. Thus linear elements have especially good convergence properties for this component of the error. It appears that using $m^{th}$ order spline basis functions on a regular mesh gives $QL(I-Q)u = O(k^{2(m+1)})$ for $L = u_x$.

Thus for evolutionary problems with this type of operator $L$ the optimal choice of higher order elements may not be the same as the conventional quadratic and cubic schemes. If the mesh is very irregular the convergence probably reverts to $O(k^{m-1})$ for $k^{th}$ order approximations. (See Thomée (1973) for the general result.)

## RESULTS

Figures 3 and 4 show the solution of $\frac{\partial C}{\partial t} = \underset{\sim}{u}.\nabla C$ where $C$ is initially $C_o + C_1 x$ and $u = u_o(-y,x)$. Figure 3 is obtained by a fourth order finite difference scheme on a 32 x 32 mesh and Figure 4 by a bilinear finite element scheme on 16 x 16 rectangles.
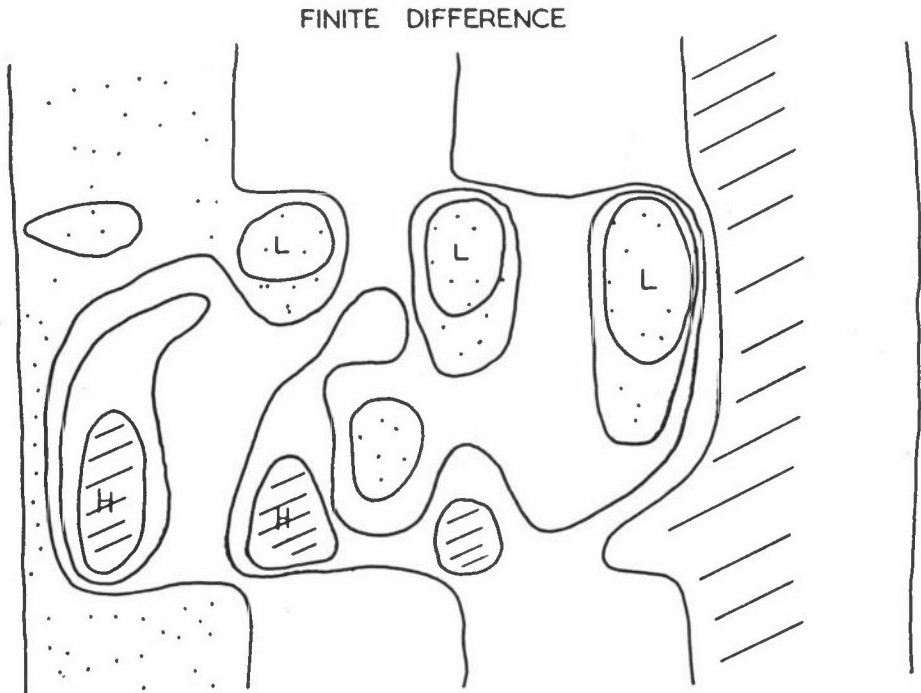
### FINITE DIFFERENCE



Figure 3: Solution of $\partial c/\partial t = u.\nabla c$ with u, c as given in text and fourth order difference scheme
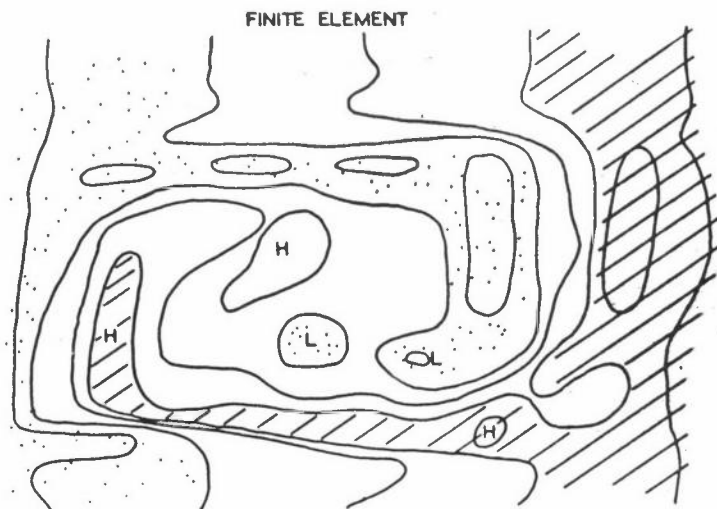
FINITE ELEMENT



Figure 4: As Figure 3 after some time with bilinear element scheme

Figures 5 and 6 show the solution of the shallow water equations in a channel obtained by a finite element Galerkin procedure on a regular triangular mesh, and by a point collocation procedure which allows aliasing errors. The two fields have very large differences for this reason.



Figure 5: Integration of shallow water equations after 2 days using linear elements and Galerkin
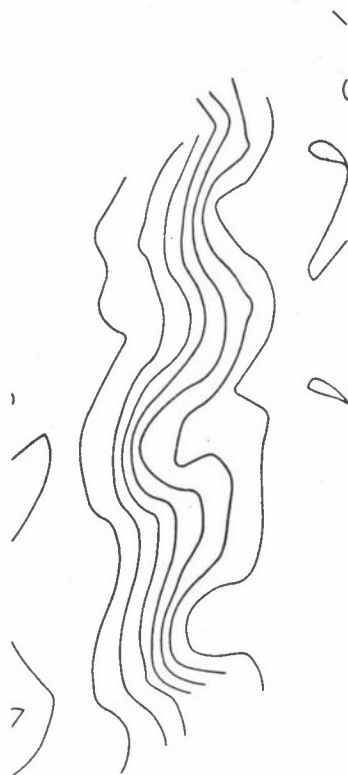
Figure 6: As Figure 5 but using a collocation formula $(uv)_n = u_n v_n$ for non-linear terms

Figures 7 and 8 show the solution of the shallow water equations on the sphere by
finite element, finite difference and spectral methods for initial data containing
a zonal wavenumber 8.    All integrations show this wave breaking down after 5 days,
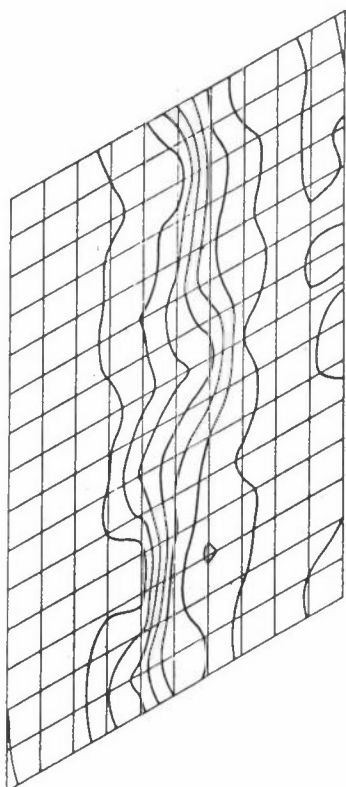except a finite difference integration with 4000 points on the sphere.    The finite
element model correctly predicting the breakdown has only 1000 points.    The computer
time required for the two models is similar.



(a)   3 days

(b)   4 days

(c)   5 days
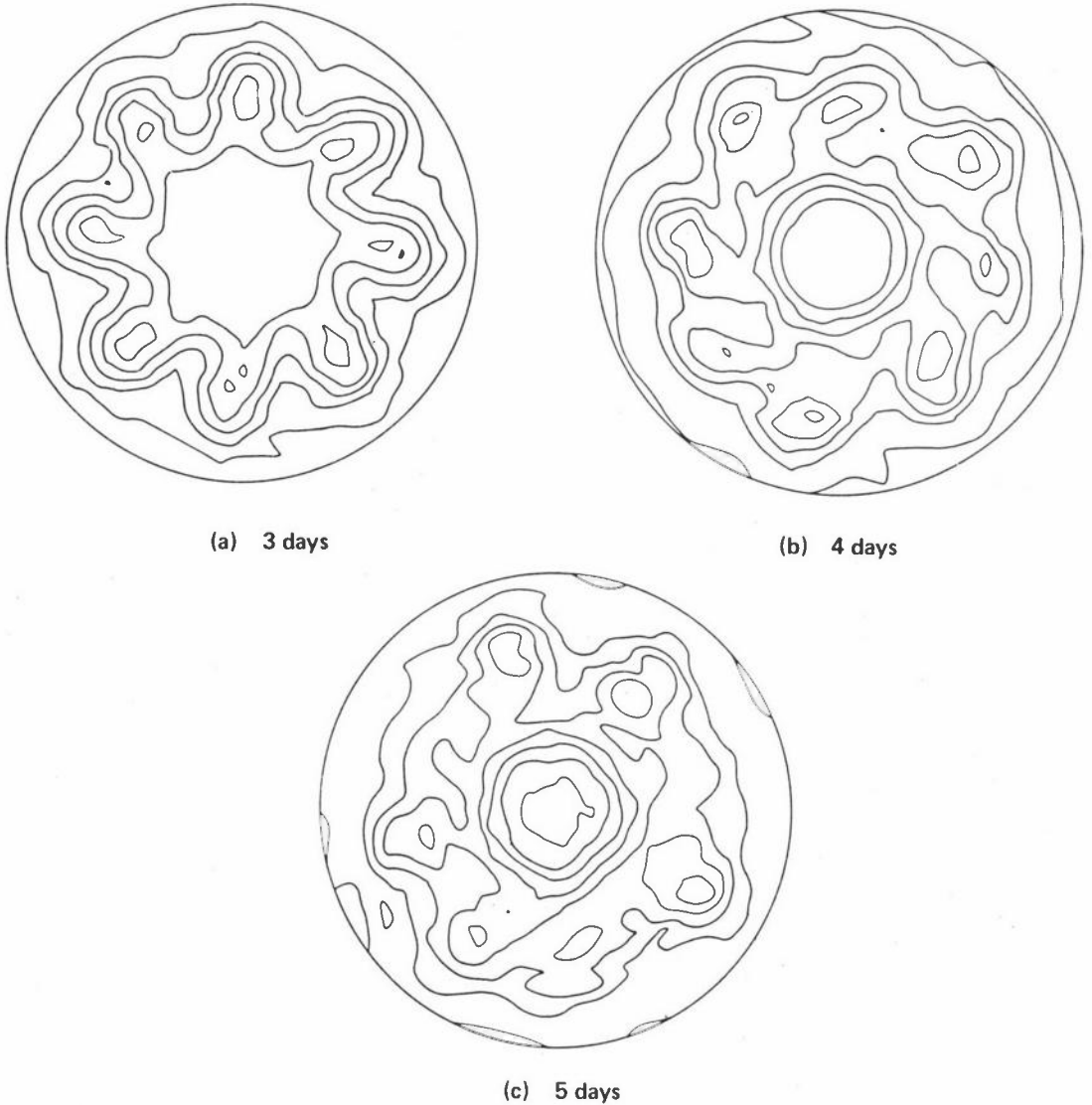
**Figure 7:  Integration of shallow water equations on sphere using linear elements**

Figure 9 shows an integration of the shallow water equations on an irregular mesh
and Figure 10 an integration with a higher order finite element scheme.    Both have
large errors due to the $QL(I-Q)u$ term;    methods of dealing with these errors are
under discussion.

Figure 8: As Figure 7 using finite difference and spectral methods
(a) low order spectral;
(b) low order difference;
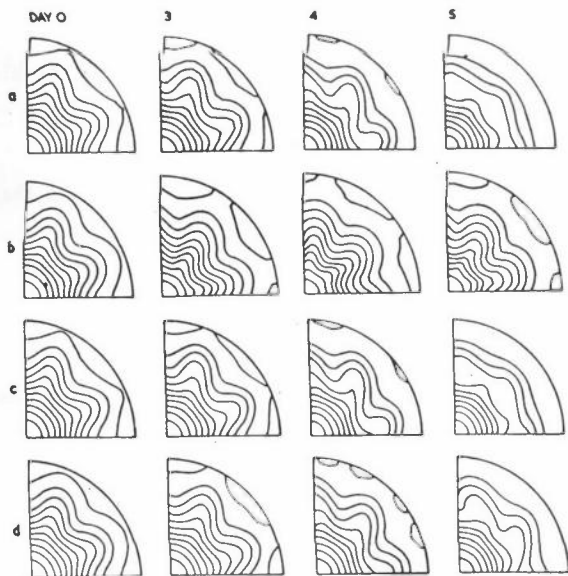(c) high order spectral;
(d) high order difference (Doron et al 1974)



Figure 9: Integration of shallow water equations on an irregular mesh (after one day)
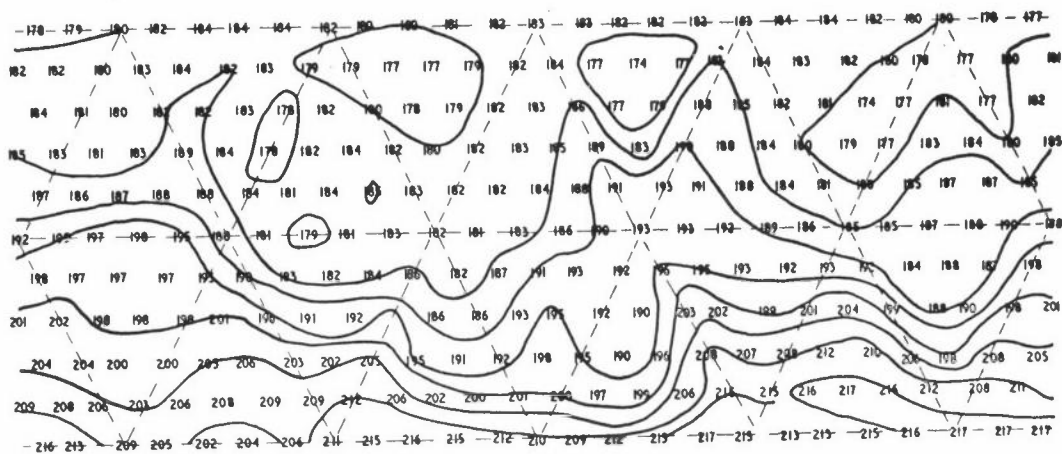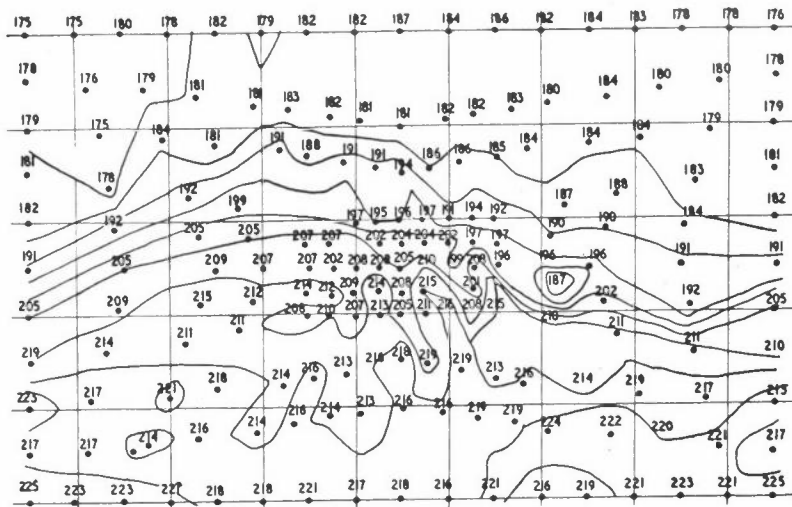


Figure 10: Integration of shallow water equations using a cubic scheme (after one day)

Special types of extra "artificial viscosity" terms of the same order as the errors
in the solution of the equation are found to allow successful integration of the problem
in Figure 9 without having a very large effect on a regular problem.    It is
reasonable to have to use such a term because if the true solution contains
unresolvable features and the approximate and true solutions conserve energy, then the
unresolved features must be replaced by erroneous ones of the same energy.

The solution of linear meteorological type problems is described in Cullen (1973),
and the solution of non-linear problems is discussed in Cullen (1974).    The solutions
on a sphere will appear in the Quarterly Journal of the Royal Meteorological
Society.

Finite element methods are also used for the shallow water equations in one dimension
by Wong (1972).

## CONCLUSION

The linear finite element approach seems to be competitive with finite difference
schemes since the computing time required for 16x16 elements is similar to that for
32x32 differences.    However, this probably only holds for schemes with special
convergence properties (apparently splines on a regular mesh and the other finite element
schemes may be too expensive to use competitively in this type of problem).

- 81 -

## REFERENCES

CULLEN M J P, J Inst Math Applics <u>11</u> 15 (1973)

CULLEN M J P, J Inst Math Applics to appear (1974)

DORAN E, HOLLINGSWORTH A, HOSKINS B, SIMMONS A, A Comparison of Grid-Point and
    Spectral Methods in a Meteorological Problem, Q J Roy Met Soc to appear (1974)

STRANG G and FIX G, An Analysis of the Finite Element Method, Chapter 7,
    Prentice-Hall (1973)

THOMEE V, Spline Galerkin Methods for Initial Value Problems with Constant
    Coefficients, in 'Numerical Solution of Differential Equations', Lecture Notes
    in Mathematics, Springer-Verlag (1973)

WONG H H, HALPEN P, DOUGLAS J, DUPONT T, Numerical Solutions of the One-Dimensional
    Primitive Equations Using Galerkin Approximations with Localised Basis Functions,
    Mon Weath Rev <u>100</u> 10 738 (1972)

# PAFEC 70+
# PROGRAM FOR AUTOMATIC FINITE ELEMENT CALCULATIONS

D A C Parkes

Department of Mechanical Engineering
The University
Nottingham
NG7 2RD

## HISTORY

The idea of developing a general finite element program, initially for the use of
research workers within the Department of Mechanical Engineering, was proposed approxi-
mately 8 years ago, 1965-1966.   Research workers who were individually pursuing different
activities associated with finite element programs realised that the large amount of
duplication of programs could be overcome by using common subroutines in each program.
This had the immediate effect of reducing the time spent by each research worker to get
his program working as he did not, for example, have to write and test merge and solu-
tion subroutines.   These early researchers prepared their subroutines on a very broad
basis and by the late 1960's the original PAFEC system contained a large number of
shell and beam type elements together with a very general merge subroutine and Choleski
solution subroutine.   At this time a number of SRC and Ministry contracts as well as
SRC studentships with a finite element basis were starting.   The research workers
associated with these contracts were persuaded to use those areas of PAFEC which would
be useful to them and hopefully the end product of their work would be available to
subsequent workers.

One of the major problems at this time, 1969, was in selecting a computer language to
use.   The initial programs had been written in Atlas Autocode for use on the Manchester
Atlas and also for the Chilton Atlas and Egdon Algol for use on the Nottingham
University KDF-9.   Because of updating problems inherent with two copies of a com-
puter program together with paper tape difficulties, a common language using punched
cards was chosen, Fortran II.   It was then possible to use both the Hartran compiler
on the Chilton Atlas and the Egdon Fortran compiler on the KDF-9 with the same card
deck.   The program immediately became available to other universities as well as being
commercially available.

By 1971 the form of the program was becoming unmanageable as the original ideas had
been modified and added to as new facilities within PAFEC became available.  The pro-
gram was radically modified to a file based system leaving the user interface as
simple as possible.   All of the subroutines had to be altered and retested at the
same time as new data handling facilities were being generated.   These activities
were timed to coincide with the change to ICL 1906A computers at both Chilton and
Nottingham.   This present system, PAFEC 70+, has now been in use for two years and is
infinitely more attractive to both university and commercial users than the original
system.   PAFEC 70+ is currently being added to as a by-product of many research
activities and is well documented in a 350 page user's manual.   A list of the major
contributors and their research topics is given in Appendix 1.

## PAFEC 70+ :- WHAT IS IT?

### Problem Types

PAFEC 70+ is capable of solving many types of structural problems.

(1)    Vibration Analysis - Natural frequency calculations for eigenvectors defined
       by all the degrees of freedom for simple structures or master degrees of
       freedom for large structures.

(2)    External Loading - Static loads applied at any degree of freedom in structure.
       Solution for all degrees of freedom.

(3)    Damped Vibration Analysis - Dynamic response to forcing functions with
       hysteritic damping either constant or variable over the structure.

(4)    Gravity Loading - Self weight of structure in particularly defined co-ordinate
       directions depending on type of structure.

(5)    Inertia Loading - Structure rotates about the global z-axis for non-
       axisymmetric cases.   Axisymmetric cases also available.

(6)    Pressure Loading - Internal and external loading dependent on the element
       topology definition.

(7)    Thermal Deformation - Temperatures defined at all structural nodes.

(8)    Steady State Temperature Calculation - Solution of Laplace equations from
       some known temperatures and heat flows.

(9)     Temperature Transient Calculation - Solution of Laplace equations from an
        initial defined temperature profile, (also used in combination with (8)
        above), for any number of time increments for a step or ramp temperature
        change.

(10)    Prescribed Displacement - Solution for whole structure from a number of
        known displacements.

In addition to the static loading type of problems, types (2), (4)-(7) above,
additional subroutines for the solution of fracture mechanics, creep and plasticity
problems are available.

## Element Types

A very large number of finite elements have been developed at Nottingham and are
available with PAFEC 70+.   Other elements developed elsewhere have been programmed
and added to the system.   The present system is now a reduced set of the best and
most useful of all these different element types.

## Classification of Elements in the PAFEC 70+ Scheme

1.  Beam Elements

    1.1    Simple beam element with a capability for flexure in two orthogonal planes,
           torsion and extension.
    1.2    As simple beam element (1.1) but with inclusion of shear deformation
           and rotary inertia.
    1.3    Offset simple beam element for use in building up sections or stiffened
           plates.
    1.4    Initially curved beam element with shear deformation and rotary inertia.

2.  2-Dimensional Elements - for the analysis of plane stress, plane strain or
axisymmetric structures.

    2.1    Simple constant strain 3-node triangle.
    2.2    3-node isoparametric triangle.
    2.3    6-node isoparametric triangle.
    2.4    9-node isoparametric triangle.
    2.5    4-node isoparametric quadrilateral.
    2.6    8-node isoparametric quadrilateral.
    2.7    12-node isoparametric quadrilateral.

3. <u>3-Dimensional Elements</u> - for the analysis of structures with variations of strain in 3 directions.

    3.1   6-node isoparametric wedge.

    3.2   15-node isoparametric wedge.

    3.3   24-node isoparametric wedge.

    3.4   8-node isoparametric brick.

    3.5   20-node isoparametric brick.

    3.6   32-node isoparametric brick.

    3.7   16-node superparametric quadrilateral shell (Ahmad type).

    3.8   20-node isoparametric transition brick (Joins 3.7 to 3.2 or 3.5).

4. <u>Plate Bending Elements</u>

    4.1   3-node curvilinear triangle (various shape functions).

    4.2   6-node curvilinear triangle.

    4.3   4-node curvilinear quadrilateral (various shape functions).

    4.4   8-node curvilinear quadrilateral.

5. <u>Shell Elements</u>

    5.1   3-node Hybrid stretching and bending triangle.

    5.2   6-node combined stretching and bending triangle (2.3 and 4.2).

    5.3   4-node combined stretching and bending quadrilateral (2.5 and 4.3).

    5.4   8-node combined stretching and bending quadrilateral (2.5 and 4.4).

    5.5   4-node cylindrically curved hybrid triangle.

    5.6   4-node cylindrically curved hybrid quadrilateral.

    5.7   2, 3 or 4-node axisymmetric thin shell for use with axisymmetric or harmonically varying loadings.

6. <u>Anisotropic Elements</u>

    6.1   4-node isoparametric membrane quadrilateral.

    6.2   8-node isoparametric membrane quadrilateral.

    6.3   4-node isoparametric bending quadrilateral.

    6.4   8-node isoparametric bending quadrilateral.

    6.5   20-node isoparametric solid brick element.

7. <u>Laplace Elements</u>

    7.1   3-node isoparametric plane or axisymmetric triangle.

    7.2   6-node isoparametric plane or axisymmetric triangle.

7.3     4-node isoparametric plane or axisymmetric quadrilateral.

7.4     8-node isoparametric plane or axisymmetric quadrilateral.

7.5     15-node isoparametric solid wedge.

7.6     20-node isoparametric solid brick.

The coding for many other elements is available and it is a simple task to add these subroutines into the PAFEC 70+ scheme.   It is a fairly easy matter to generate the extra elements in the isoparametric families due to the present simplicity of the coding.

Solution Methods

For any static problem there is available a choice of solution depending on the size of the problem.

(1)     Full or Banded Stiffness Matrix where the whole matrix can be held completely in core.   The Choleski banded solution procedure is used to find the displacements.

(2)     Semi-Automatic Solution procedure uses the Choleski banded solution to solve for a set of master degrees of freedom, the reduced stiffness matrix being held in core.   The user defines which degrees of freedom are inactive and can be reduced out using Gaussian reduction as the solution proceeds.   The disadvantages of this method are that the reduction data can be rather lengthy and tedious and as no backing store is used all the reduced-out displacements are lost.

(3)     An Automatic Frontal Solution Technique uses both Gaussian and Guyan reduction to maintain a small in-core stiffness and mass matrix if required.   The use of fast access backing store enables all displacements to be calculated.   A separate program is available to preprocess the data so that the solution is completely automatic.

For vibration problems the 3 solution methods mentioned above are used to form a final square stiffness and mass matrix.   The eigenvalues and eigenvectors are found for this stiffness and mass matrix using a Householder Tridiagonalisation followed by either a Sturms Sequence algorithm or QL algorithm.

Transient analyses such as creep and plasticity of large problems are solved using the resolution technique of the frontal solution with time or load incrementing techniques.

Steady state and transient temperature or prescribed displacement calculations
are treated as

(1)    Partitioned solutions where the matrices can be held in core.

(2)    Frontal solutions with the known values treated as master degrees of freedom.

A Crank-Nicholson time marching technique is used for transient temperature calculations.


Data Inputs

The data for any PAFEC 70+ job can take a number of forms.    The original PAFEC data
was rather long and required not only a description of every node in the structure
but also every degree of freedom.    Recent modifications have made it possible to
reduce the data by defining constraints rather than every degree of freedom.    Mesh
generation techniques have also reduced the data input by reducing both the number
of nodes necessary to define the idealisation as well as the topology of all the
elements.    Now only the chassis nodes and topology together with constraint and
loading information is necessary.


An example of a typical simple structure is shown in Figure 1(a).    The data required
by the program is shown in Figure 1(b) and may be loaded in this form.    Alternatively
using a mesh and data generation facility the data can be drastically reduced,
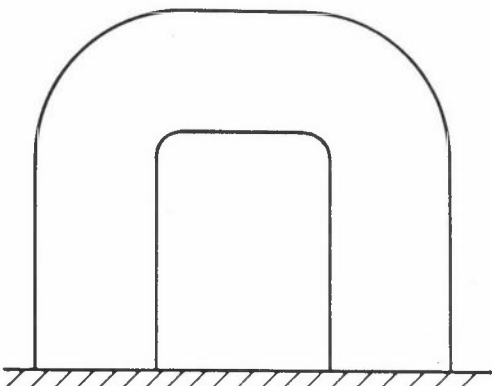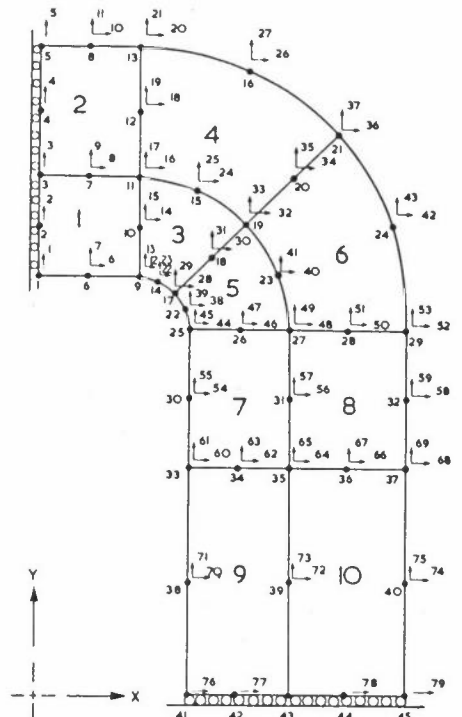Figure 1(c).

## IDEALISATION OF HOOK STRUCTURE



Figure 1(a):  Idealisation of hook structure

```
*** PAFEC 70+JOB ***   10 ELEMENTS - BANDED SOLUTION      TITLE
    2   45   79     79     79     5    22     1     4    10    ⎤  Job Control Card
          0.000        1.000        0.000     0     1    ⎤
          0.000        1.120        0.000     0     2    │
          0.000        1.240        0.000     0     3    │
          0.000        1.395        0.000     0     4    │
          0.000        1.550        0.000     0     5    │
          0.120        1.000        0.000     6     7    │
          0.120        1.240        0.000     8     9    │
          0.120        1.550        o.000    10    11    │
          0.240        1.000        0.000    12    13    │
          0.240        1.120        0.000    14    15    │
          0.240        1.240        0.000    16    17    │
          0.240        1.395        0.000    18    19    │
          0.240        1.550        0.000    20    21    │
          0.286        0.991        0.000    22    23    │
          0.378        1.213        0.000    24    25    │
          0.497        1.499        0.000    26    27    │
          0.325        0.965        0.000    28    29    │
          0.410        1.050        0.000    30    31    │
          0.495        1.135        0.000    32    33    │
          0.603        1.242        0.000    34    35    │
          0.711        1.349        0.000    36    37    ⎬  CPDDC
          0.351        0.926        0.000    38    39    │
          0.573        1.018        0.000    40    41    │
          0.841        1.131        0.000    42    43    │
          0.360        0.880        0.000    44    45    │
          0.480        0.880        0.000    46    47    │
          0.600        0.880        0.000    48    49    │
          0.738        0.880        0.000    50    51    │
          0.875        0.880        0.000    52    53    │
          0.360        0.715        0.000    54    55    │
          0.600        0.715        0.000    56    57    │
          0.875        0.715        0.000    58    59    │
          0.360        0.550        0.000    60    61    │
          0.480        0.550        0.000    62    63    │
          0.600        0.550        0.000    64    65    │
          0.738        0.550        0.000    66    67    │
          0.875        0.550        0.000    68    69    │
          0.360        0.275        0.000    70    71    │
          0.600        0.275        0.000    72    73    │
          0.875        0.275        0.000    74    75    │
          0.360        0.000        0.000    76     0    │
          0.480        0.000        0.000    77     0    │
          0.600        0.000        0.000     0     0    │
          0.738        0.000        0.000    78     0    │
          0.875        0.000        0.000    79     0    ⎦
36210   10    0                                          ⎤  Element Control Card
    1    9    3    11     6     2    10     7    0.300    ⎤
       3.0E7       0.250       0.300                      │
    3   11    5    13     7     4    12     8    0.000    │
    9   17   11    19    14    10    18    15    0.000    │
   11   19   13    21    15    12    20    16    0.000    │
   17   25   19    27    22    18    26    23    0.000    ⎬  Element Topology
   19   27   21    29    23    20    28    24    0.000    │
   25   33   27    35    30    26    34    31    0.000    │
   27   35   29    37    31    28    36    32    0.000    │
   33   41   35    43    38    34    42    39    0.000    │
   35   43   37    45    39    36    44    40    0.000    ⎦
     9999.0                                               ⎤
    1    1        1.0                                     │
    7    2        1.0                                     │
   13    3        1.0                                     │
    1    4        1.0                                     │
    7    4        2.0                                     ⎬  Loading Information
   13    4        1.0                                     │
    0    0        0.0                                     │
    1    9    0                                           │
    0    0    0                                           ⎦
```

Figure 1(b): Standard PAFEC data for hook structure

```
    10 ELEMENT HOOK - PAFBLOCKS SOLUTION - BANDED          Title
      2    10          8         5    1    1    4    3 _   Control Card
        0.000      1.000
        0.000      1.550
        0.240      1.000
        0.240      1.550
        0.325      0.965                                    Chassis CPDDC
        0.711      1.349
        0.360      0.880
        0.875      0.880
        0.360      0.000
        0.875      0.000                                _
  36210
        1     1    2              1
        1     3    2    4                        0.300
          3.0E7      0.250      0.300
        1
       24    55
  36210
        1     2    2              1         1
        3     7    4    8    5              6
        1     1                                            Pafblocks topology
       24    55
        1     5    2
        0     0    0
  36210
        1     2    2              1
        7     9    8   10
        3     5
       24    55                                        _
        7777.0
        0.000      1.000      0.000    2    1    1.000
        0.120      1.000      0.000    2    2    1.000
        0.240      1.000      0.000    2    3    1.000
        0.000      1.000      0.000    2    4    1.000     Loading Information
        0.120      1.000      0.000    2    4    2.000
        0.240      1.000      0.000    2    4    1.000
        0.000      0.000      0.000    0    0    0.000 _
        1    10    0                                    _
        0     0    0                                       Stressing Data
    -9000     1    1                                    _
        0.000      1.000      0.000                     _
    -9000     2    2                                       Constraint Data
        0.360      0.000      0.000
        0     0    0                                    _
  ZZZZ                                                  _
```

Figure 1(c):  PAFBLOCKS data for hook structure

Drawing Facilities

A very important requirement of all finite element programs is the ability to present information about the structure in a pictorial form.   The PAFEC 70+ program uses graph plotters, generally running off line, to produce a wide variety of plots.

   (1)    Drawing of all or part of the structure in "blown-up" form.
          (The elements are slightly separated from one another and the
          structure appears to have disintegrated.)

(2)     As (1) but with nodal circles drawn in.

(3)     As (2) but with node numbers added.

(4)     As (1) but with element numbers added.

(5)     All or part of structure not "blown up".

(6)     As (5) dotted with degree of freedom arrows and numbers drawn.

(7)     As (5) dotted with constraint arrows drawn.

(8)     Outline of structure.

(9)     Stress vectors on (8) or (5)

(10)    Displaced shape plot.


The drawing facility is used initially to check the data as errors are usually
immediately apparent.    Presentation of graphical output for reports saves time and
by suitably scaling the plot the output can be used directly.    An example of a plot
used for data checking is shown in Figure 2.
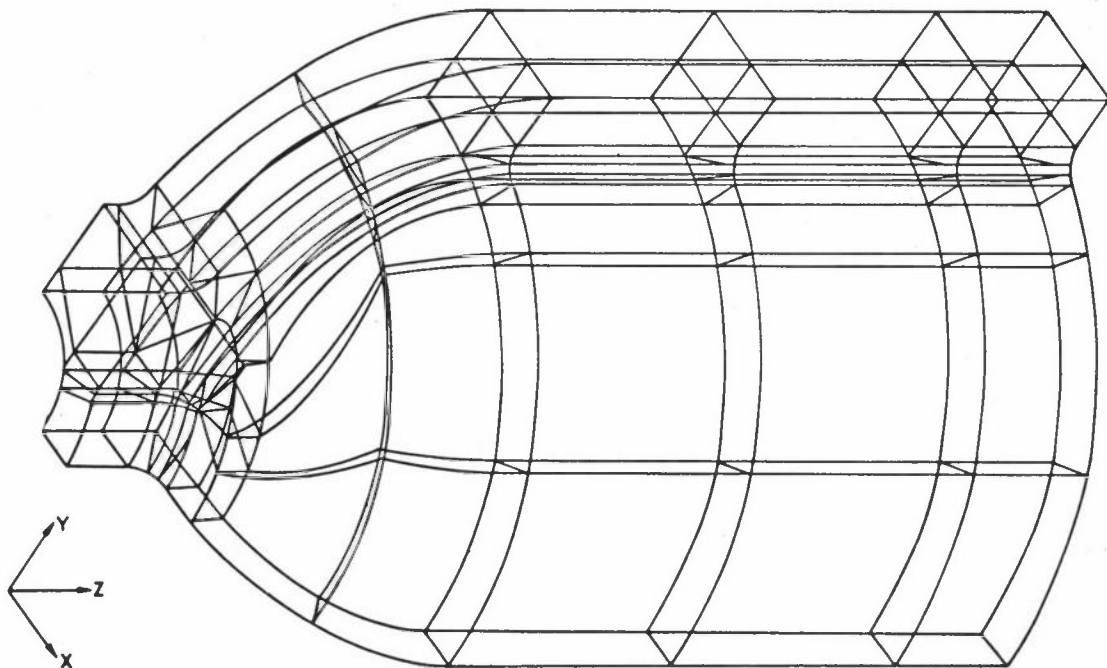


Figure 2:  Idealisation of flanged turbine casing

## SOFTWARE

The PAFEC 70+ program is a file based system.    Each program is expected to run in 2
or more parts

    (1)    Card data read

              Mesh and/or Data generation

              Plotting

              Write Input data for program to magnetic tape

    (2)    Read Input data for program from magnetic tape

              Form element matrices and merge

              Solve for displacements, eigenvalues, etc.

              Calculate stresses

              Plot stresses

There are many variations to the basic concept but the main intention is that the user
checks his output and graph plots from the first program before commencing the second
more expensive part of the program.

The PAFEC 70+ program is broken down into 12 separate libraries.    It would be possible
to run with as few as 4 libraries but owing to the ICL restriction on filestore size
together with the restriction of the single pass consolidator (link-editor) the 12
libraries have been chosen.    Their titles and contents are as follows

    PAFECLIBCI - All data handling, polynomial, matrix pack and some solution
                  subroutines

    PAFECMRGCI - Merge subroutine for full or banded matrices

    PAFECSEMCI - Merge subroutine for semi-automatic solution

    PAFECAUTCI - Merge and solution subroutines for frontal solution

    PAFECBSHCI - All Beam and Shell type element subroutines

    PAFECEL2CI - All 2-dimensional Membrane and Bending element subroutines

    PAFECEL3CI - All 3-dimensional element subroutines

    PAFECBLKCI - Mesh Generation subroutines

    PAFECDGNCI - Data Generation constraint subroutines

PAFECDRWCI - Plotting subroutines

PAFECTMPCI - Laplacian element and solution subroutines

PAFECMATCI - Matrix pack written in assembler code.


For any program only a few of these libraries will be required depending on the problem but they have to be called in the order depicted by Figure 3.    It is possible from the data to generate not only the main program but also the information required to determine the libraries required by the program.
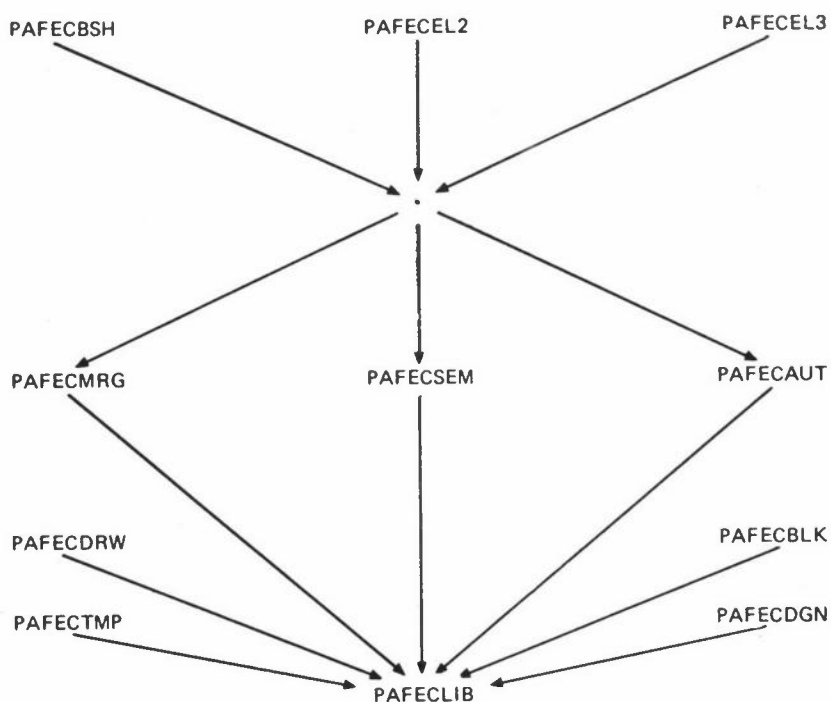


Figure 3:  PAFEC library layout
(It is important that the library files are called in the order indicated by the arrows so that no subroutines are omitted by a single pass consolidator (eg a job using R36210 with automatic front solution required "semi-compiled" PAFECL2, PAFECAUT, PAFECLIB libraries to be scanned. The library PAFECMATCI contains matrix handling subroutines written in PLAN. This library of subroutines is available for ICL 1900 series computers.)

Two macros are available at present for ICL GEORGE operating systems which only require a data file for each job.    The macros generate 2 programs as described previously, the default conditions being a limited jobtime, a plot of the structure, a mesh generation if the control parameters are correct, data generation and frontal solution.    Other options are available by altering the default options by insertion in the datafile.

The macros enable the user to be guaranteed that the job control and main program including the dimension statements will be correct for every program.    Up to 50% of research programs fail on their first run due to either main program or job control errors.

## USE OF PAFEC 70+

PAFEC 70+ was originally written as a research tool but has since become available as a program for production analyses.    The program is highly segmented which enables the research activities with PAFEC 70+ to carry on unimpeded.    Modifications and new facilities are easily incorporated into the program due to this segmentation.

The program is freely available to academic establishments and is at present being used by a number of universities and polytechnics.    Interaction with other finite element research groups should hopefully cut down on repetitive development work and yield a larger and generally more useful program.    Increase in academic usage would require support from a research organisation as the present servicing is supported from commercial activities.

## FUTURE OF PAFEC 70+

Present developments are aimed at improving the input and output facilities of the program.    The mesh generation program PAFBLOCKS is being extended and it is hoped that the recently developed data language PAFTRAN can be suitably modified to incorporate the PAFBLOCKS subroutines.    The mesh generation is being further extended to incorporate a number of regularly shaped structures.    A recent addition to the ease of data preparation is a digitising table to be used in conjunction with the mesh generation facility.

Extra output facilities being developed include contour plotting to complement the present stress vector plotting facility.    Other modifications anticipated in the near future are a frontal blocking solution technique and possibly a banded partitioned solution.    A standard program is being developed for incorporating the individual creep and plasticity subroutines more generally into the main program.    Development of elements includes a doubly curved shell and a general non-symmetric loading axisymmetric isoparametric element.

## APPENDIX I    CONTRIBUTORS TO PAFEC 70+

BATCHELOR A, The Correlation of Roadway Displacements with the Stress Redistribution and Strata Movements caused by Longwall Mining, PhD Thesis (1972)

BICKLEY A, Elastic Distortion of Crankshafts, PhD Thesis (1968)

BISHOP P T, Failure of Reinforced Plastics caused by Stress Concentrations, PhD Thesis (1973)

BOND T J, Some Considerations of the Finite Element Method in Stress Analysis, PhD Thesis (1972)

COHEN E, Transformer Noise, PhD Thesis (1968)

DAVIS R, Advances in Beam Finite Elements and Applications to Stiffened Plates, PhD Thesis (1972)

DASGUPTA A, The Behaviour of Joints in Tubular Trusses, PhD Thesis (1970)

DICKENSON S M, Flexural Vibration of Rectangular Plate Systems, PhD Thesis (1966)

EDWARDS G, Cylindrical Shell Hybrid Finite Elements, PhD Thesis to be submitted (1974)

HENSHELL R, Transmission of Vibration in Damped Elastic Structures, PhD Thesis (1967)

JOHNSON K, The Mechanical Behaviour of Laminated Carbon Fibre Reinforced Plastics, PhD Thesis (1973)

JONES C, Stress Analysis of Aircraft Components and Advanced Aspects of Finite Element Software, PhD Thesis to be submitted (1975)

MAKOJU J O, The Finite Element Analysis of Thick Shells, PhD Thesis to be submitted (1976)

NEALE B K, Vibration of Shell Structures, PhD Thesis (1971)

ONG J H, Finite Element Analysis of Transmission Systems, PhD Thesis to be submitted (1976)

PARKES D A C, Finite Element Analysis of Stresses and Creep in Turbine Casings, PhD Thesis (1973)

RIGBY N, Hybrid Shell Finite Elements, PhD Thesis to be submitted (1976)

ROPER C, An Analytical Study of the Strength and Flexibility of Tubular Joints, MPhil Thesis to be submitted (1974)

SAFAVI F, Automatic Data Checking for Structural Analysis using Graphical Techniques, MSc Thesis (1971)

SHAW K, Finite Element Analysis of Crack Propagation in a non-Uniform bi-axial Stress Field, PhD Thesis to be submitted (1974)

SULLIVAN C, Stresses in Hooks, MPhil Thesis (1972)
        Finite Element Analysis of Aircraft Structures,
        PhD Thesis to be submitted (1975)

SWANNELL J, Plasticity and Creep Using Finite Element Techniques with Special Reference to Taper Hub Flanges, PhD Thesis to be submitted (1974)

WALTERS D, Paftran; A Finite Element Mesh Generation Language, PhD Thesis (1973)

WALTERS J, The Stress Analysis of Joint Links using Anisotropic Materials, PhD Thesis to be submitted (1975)

WEBSTER J J, Free Vibration of Shells of Revolution, PhD Thesis (1968)

WINDSOR A A, Vibration of Shells using Finite Elements, PhD Thesis to be submitted (1975)

# ASAS – A GENERAL PURPOSE FINITE ELEMENT SYSTEM

J B Spooner

Atkins Research and Development
Ashley Road
Epsom
Surrey

## DESIGN OBJECTIVES

ASAS is a large scale, general purpose finite element program based on the matrix
displacement method.   Although originally based on a special purpose package
delivered to AWRE  Aldermaston, its designers set themselves the objective that the
program should be commercially attractive and thus generate sufficient revenue
through sales, rentals, leases and bureau royalties to cover the not inconsiderable
development costs together with a reasonable profit margin.

In considering how this requirement should be met, further objectives were isolated
by the design team:  many were rejected as impracticable and the following represent
its assessment of the essential combination.

It was decided right from the very first that the program should be general purpose,
that is, not restricted to any one class of engineering problem.   This was made
possible by the fact that the underlying mathematics and modelling techniques are
common to all structural problems.   Above all, however, it should be easy to use
by the design engineer with little or no experience of writing computer programs.
It was thought that this would go a long way to help break down the aura of
mystique built round the finite element method – probably as a result of the
technical literature doing relatively little to encourage discussion on the practical
application of the method or, more likely, as a result of unfortunate experiences
with some early program.

In spite of trying to achieve generality and ease of use, the program was also
expected to be cheap to use, again in an attempt to attract back to the method
those many engineers frightened off by the high cost of some early analyses.
It should again be noted that the term cheap to use embraces all activities, from
setting up the initial data to interpretation of the final results, and does not,

therefore, refer only to the running costs on the computer. Early experience of the finite element method indicated that computing contributed about sixty percent of the total bill for any job - this nowadays has sunk to about forty percent of the total cost. Note, however, the total cost may be only one-third of that for the corresponding analysis done five years ago.

Another essential requirement was that the program should be flexible, that is, easy to extend, modify and adapt. This was primarily in recognition of the fact that, although a great effort was to be expended in trying to make the program genuinely general purpose, it would have been foolish to pretend that all requirements of all branches of engineering could be met and that addition would always be expected.

A further and not inconsiderable requirement was that the program should be easily transportable from one computer make or configuration to another: many finite element programs have fallen at this hurdle. In spite of great care in the design of a system, this is always a point at which some difficulty can be expected.

The final requirement was that the program should be well documented, serviced and supported at all levels. Furthermore, it should be shown to be a living system - that is, development does not stop after the first few successful runs, but that it continues so that new facilities are offered, new applications found and that user experience is circulated from one installation to another.

## THE OBJECTIVES IN GREATER DETAIL

It was only to be anticipated that the satisfaction of the above concurrent requirements placed severe demands on the program design. The next section amplifies each main requirement and following sections describe design features implemented in ASAS to meet these objectives, giving some indication, where possible, of the effectiveness of the approach. It should be recognised that there is a strong cross-coupling between several of the declared objectives, and this is clarified in the following sections.

Generality

It can be assumed that if the description "general purpose" is applied to a finite element program it implies suitability for use by engineers in the many branches of

the profession ranging, for example, from university research workers in finite element theory to high-pressure vessel designers. However, "suitability for use" implies in turn that the program provides adequate modelling facilities for both structure and loading, performs reasonably efficiently in all different applications and is therefore not noticeably more effective in one field at the expense of another.

A good indication of the generality of a finite element program is given by its library of finite elements. Obviously, the more elements of different type and facility, the greater the structural modelling capability available to the user.

The elements listed below form the basic library of displacement elements: a new range of force continuity and special elements has recently been introduced and used in the analysis of several aircraft structures.

| | |
|---|---|
| Flange (bar) elements | 2 and 3 node axial elements with constant and linearly varying cross-section. |
| Beam and grillage elements | 2 node beam in 2-D and 3-D with and without shear area. Curved (circular arc) beam element. |
| Membrane elements | 3 and 6 node triangular elements with constant and linearly varying strain field, 4 and 8 node quadrilateral elements. |
| Solid elements | 4 and 10 node tetrahedra. 8, 20 and 32 node isoparametric solids. 6 and 15 node isoparametric wedges. 6 node triangular axisymmetric element for axisymmetric loading. 9, 18 and 30 node axisymmetric sector elements for non-axisymmetric loading. |
| Plate bending elements | 3 node (18 degrees of freedom) triangular element. 8 node curved, quadrilateral isoparametric thin/thick element including transverse shear, 24 degrees of freedom. |

| | |
|---|---|
| Shell elements | 3 node axisymmetric thin shell (line) element. |
| | 3 node thin shell triangular element, 27 degrees of freedom (cubic interpolation of geometry and displacements). |
| | 3 node thin shell triangular element, 54 degrees of freedom (quintic inter- polation of geometry and displacements). |
| | 3 node thin shell triangular element, 18 degrees of freedom (flat plate bending element plus membrane element). |
| Special purpose elements | Rectangular shear wall element. |
| | Warped quadrilateral shear panel with force continuity. Axial load carrying element with force continuity for use with warped quadrilateral. |
| | Warped semi-monocoque element (stringer stiffened membrane element). |

Many of the above elements allow the consideration of orthotropic as well as isotropic material behaviour. This library of elements is not static and will be added to as demand necessitates.

To complement an extensive range of elements for modelling many different structures, it is necessary to include facilities for handling various loading types. ASAS accepts the following loadings:

| | |
|---|---|
| Nodal forces | (point loads, moments, torques) |
| Line and surface forces | (line loads on beams, edge loads on membrane structures, pressure loads on bending, shell and solid elements) |
| Body forces | (self weight, centrifugal and inertia forces) |
| Initial strains | (usually arising from a temperature distribution) |
| Prescribed displacements. | |

These loading types may be combined in any way in one or many different load cases. There is virtually no restriction to the number of load cases that can be handled in one run (the maximum to date is 108).

Should a consideration of other types of loading be necessary, the program should be sufficiently adaptable to enable the user to readily incorporate his requirements. This is covered later under "flexibility".

Easy to use

There seems little point in developing a general purpose finite element program that is not easy to use in all its phases - data preparation and correction, execution of the analysis on the computer, and interpretation of the results. Experience shows that given the easy to use program, straightforward guidelines and possibly a little assistance in the early stages, design engineers are well able to make a success first of small analyses and later, as experience grows, of more complex analyses.

Experience also shows that ease of use of a finite element program is virtually synonymous with ease of use and completeness of its user manual. This must contain a clear description of how all data is to be prepared and what is necessary to execute a job on the particular computer installation, using a minimum of systems analyst's or mathematical jargon. It has also been found that pre-printed, self descriptive data preparation sheets not only make a good impression but, together with the user manual, make data preparation less of an error prone process than is usually the case.

Having briefly mentioned 'external to program' aids to ease of use, mention should be made of internal features that have been programmed in ASAS. First the program accepts the description of the structure's geometry, element topology, loading and support conditions punched on the data cards in very compact form. This data is then interpreted and expanded by the program and a comprehensive series of checks made for consistency. Should errors be found, they are reported by the program using straightforward English language descriptions rather than complicated codes. The program cannot proceed with an analysis if errors have been found and not corrected.

Warnings of data giving rise to suspect idealisation are also given, but the user is allowed the option of overriding them if he has deliberately set up his problem

that way. At this point, provision is also made for a graphical display of the idealisation. Should the data be consistent and any warning messages overridden, the program proceeds automatically to any stage selected by the user. Usually, after the data checking phase, he will select to carry a calculation through to displacements but could, of course, carry it through to the output of stresses. All these options are under full control of the user and are declared using mnemonic descriptions at the beginning of the data input. Automatic restart facilities are also available so that the program can be stopped and restarted at various points through job execution.

Finally, the results of the analyses are clearly formatted with title and headings for each set and column of numbers, and facilities also exist whereby the output is automatically stored for post-processing - for example, factored load combinations.

## Cheap to Use

'Cheap to use' and 'easy to use' are in many ways closely interwoven. For example, compact data input, good data checking and plotting facilities not only make the program easy to use, but simultaneously reduce the overall costs of data preparation. It has been found advantageous to include a substantial volume of code in the data checking phase at the expense of slightly extended run times in this area of the program - this cost disadvantage is heavily overweighed by the cost savings in quick location of errors.

Once the data is as correct as can be checked by the computer, it is used to establish the idealised structure and loading. At the solution stage, two alternative paths are available. For small problems the frontal solution technique is employed, whilst for large problems a partitioned sparse matrix scheme is used. The appropriate solution routine is, normally, automatically selected by the main program. Both obviously, recognise the banded nature of the structural stiffness matrix, though in the frontal solution this matrix is not formally assembled.

## Flexibility

It has already been stated that at the design stage it is impossible to anticipate all the uses the program may be put to, so it is imperative that it is easy to modify or extend to meet specialist requirements. This is achieved by a combination of modular program design plus full system documentation.

## Machine Independence

It has been stated that a program cannot be called machine independent until it is completely operational on at least two distinct ranges of hardware. For this reason, ASAS is written in ASA Standard FORTRAN and the first version relies on sequential access of backing store files through standard FORTRAN Input/Output instructions. Although this introduces a measure of inefficiency at the assembly and solution stages, it has not proved as penalising as many would suspect. The benefit of machine dependent facilities, notably in the two areas mentioned above, has been recognised, and a version of ASAS using direct access facilities to backing files has been produced. The program is currently mounted on RXDS SIGMA 5 and SIGMA 7, UNIVAC 1108 (Exec 2 and Exec 8), UNIVAC 1106 (Exec 8), IBM 360/75 and IBM 370/150, and conversions are being prepared for the ICL 1900 series and CDC 6600 machines.

## Service at All Levels

The success of a general purpose finite element program depends to a very large extent on the back-up provided. Back-up ranges from hand holding and first-aid treatment during initial use of the program to access to the originators of the program or equally qualified personnel in case of unexpected or unusual difficulties, or should modifications or new developments be required. Furthermore, the program should be seen to be a living thing with guaranteed support and adequate development to come.

An elasto-plastic version of ASAS and a thermal analyser ASASHEAT have been produced and the design team is currently extending sub-structuring facilities and the calculation of dynamic response. Figure 1 summarises current ASAS analytical capabilities.

## THE DESIGN OF ASAS

This section describes how the framework of ASAS was planned so as to meet the objectives declared in earlier sections.

ASAS is subdivided into a number of modules, each performing a distinct operation, linked together by an overall steering or header program. The modules are designed mainly for use with other types of problems. The interface between each module is
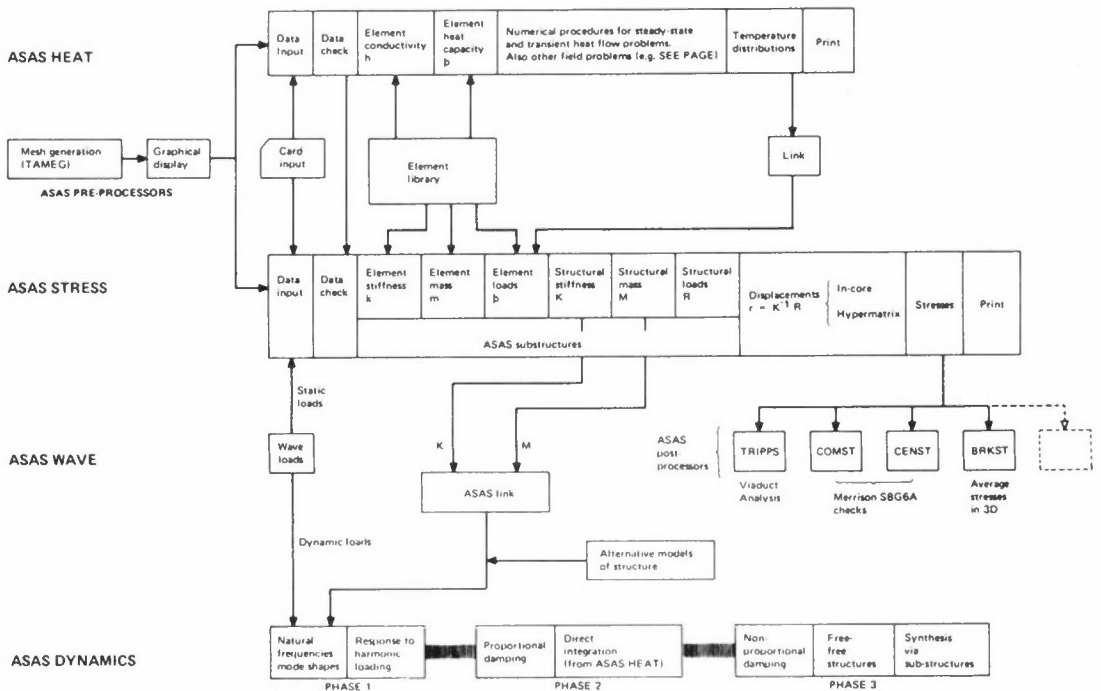
**Figure 1: ASAS analytical capabilities**

clearly defined by the use of data blocks, which simplify the work involved in the
addition of new facilities, and allow internal information to be used outside the
main program for both pre- and post-processing work.    Mesh generation and stress
plotting are examples.    Another advantage of this modularity is the ease with
which improved versions of existing modules may be added to the system.    For example,
a complete double precision version of the structural stiffness assembly routine can
be incorporated without disturbing other parts of the system.

Envisaging the continuous development and addition of new finite elements, there
are clearly defined points where new element dependent routines may be added.    These
points appear at the element checking stage, the calculation of the element stiff-
ness, element loads and element stresses.    The element dependent routines can be
written with no knowledge of ASAS, as their introduction to the main system requires
only a simple 'capping' routine that accepts the element data in the format provided
and returns the element characteristic matrices in a similar manner.

Key information, giving certain details about the elements, is kept permanently in
core in FORTRAN arrays termed 'libraries'.    Typically, they contain the number and
types of freedoms and the number and types of stresses on an element.    The element
dependent stages are guided by references to these libraries and hence there are

no imposed restrictions on the type of element employed, and the number of nodal parameters permitted at a node. Any number of elements of any type can meet at a node providing they have some common freedoms, although the number of parameters on one element may differ from that on another and may vary from node to node.

An essential feature of ASAS is that it is designed to handle both small and large problems by the use of dynamic storage allocation and automatic core overflow techniques. The method used within ASAS recognises the simple flow of data inherent in the finite element method, while giving the primary advantage of core allocation at job execution time. This is achieved by regarding the available data area as a single, one-dimensional FORTRAN array termed 'freestore'. From this, any amount of space can be called for and used by each module of the program. The information created can then be written to a backing file and the space returned to freestore for use in some other part of the program. This simple approach has the advantage that no program space is required for complicated control software and that no organisational overheads are incurred during job execution.

ASAS is capable of performing an efficient analysis of most practical engineering and research problems. This flexibility is achieved by controlling the storage of internal information, word packing and a choice of solution routines. The user, however, does not have to make a choice as to which modules are necessary for an efficient solution to his problem; this is decided in the header program.

Ideally, there should be no restrictions on the size of problem that a structural analysis system is capable of handling. Practical limitations, such as the amount of core and backing store available on a particular configuration place certain restrictions on the size of problems that ASAS can tackle. Typically, a problem with 3,000 unknowns and a large number of load cases will run successfully on a 32 bit, 32K word store machine, and a 64K word machine will solve problems very much larger than this without difficulty. Substructuring facilities are currently being implemented to remove some of the difficulties associated with the very large problem, particularly in the description of the idealisation and with the solution of the equilibrium equations.

A restart facility exists in ASAS which allows the user a degree of control in the execution of large jobs and in salvaging jobs in the event of inadequate limits specification (time, number of output pages), or if a machine failure occurs. In all, there are twelve restart stages in the program.

# NEWPAC – A PROGRAM FOR LINEAR ELASTIC STRUCTURAL ANALYSIS BY FINITE ELEMENTS

A Scholes

British Railways Board
London Road
Derby

## INTRODUCTION

The need for a general purpose structural analysis program was formulated in British
Rail's Research and Development Division in Derby in 1966.  A wide variety of
railway structures, eg coaches, wheels, bridges, required detailed stress calculations
and the traditional methods were proving inadequate.  The Division had developed
expertise in the finite element method using some small and special purpose programs
and came to the conclusion, along with some other organisations at that time, that
there was no suitable general purpose program available from external sources and hence
that a program, written in house and thus more closely allied to railway engineers'
needs, would be most appropriate.

This latter point was most important since it was considered that the successful intro-
duction of more exact computational techniques into the design offices of British
Rail would be eased by the local availability of any programs and of qualified support
staff.

The major aims in producing a program were formulated as:-

   (i)    Capability for 3-D static and dynamic analysis using finite elements of all
          structures likely to be of interest to railway engineers.

   (ii)   No arbitrary limits (as far as possible) in regard to the complexity of
          structure it can tackle.

   (iii)  Simple form of data input which requires little learning.  The data can
          now be checked using interactive graphics.

   (iv)   Modularity in the program so allowing ease of amendment and additions.

(v)     Use of standard high level language (FORTRAN IV) with minimum of assembler
        level language.

The aim of this paper is to show how we attempted to meet these objectives and to allow
our attempts to be compared and discussed in the light of other solutions.


## History

The program NEWPAC was begun in 1966 and was in productive use at the end of 1969.  In
the meantime, smaller programs, particularly one called OLDPAC, were developed to allow
the wider introduction of finite element methods into the design offices and to gain
experience of users reactions to, for example, different forms of data input (West 1969).
All those programs were developed on an ICL 1909 computer with five magnetic tape
decks.   NEWPAC was subsequently transferred to an IBM System 360 (and later a 370)
computer, working under the OS/MFT2 operating system.   The subsequent description refers
to the latest version of the program.   All facilities, particularly those involving
graphical output, may not be available with all versions of the program (see user guide,
(Patel et al 1974)).


## ANALYSIS PERFORMED

The various types of analysis which may be performed are:-

(i)     Static analysis to calculate the displacements and stresses due to prescribed
        loads, displacements and temperature fields.

(ii)    Static analysis as (i) but including terms due to geometric stiffness, eg
        where tensile or compressive loads in end-load members are taken to have
        stiffening or de-stabilising effects respectively.

(iii)   Calculation of natural frequencies and corresponding modal displacements and
        stresses.

(iv)    Calculation of buckling load factors and the corresponding buckled shapes.

(v)     Steady-state analysis of temperature or other potential problems.

(vi)    Transient analysis for mechanical systems with mass, damping and stiffness
        to calculate displacements and stresses under time-varying loads.

(vii)   Thermal transient analysis to calculate temperature-histories under varying heat load inputs.


## Static Analysis

This is the basic finite element procedure which involves the following steps:-

(i)   The idealisation of the structure under consideration as a number of elements (members), connected at nodes (joints).   The movements of each node are defined in terms of displacements (both linear and rotational).

(ii)   the specification of this idealisation (as nodal co-ordinates and element properties) and also of the loading and constraints imposed on the structure in a form suitable to the computer.   This is done on specially designed "input" data sheets.

(iii)   the computer analysis itself.   The input information is used to calculate the relationships between the forces on each element and the displacements of its connecting nodes.   These relationships are called "element stiffness matrices" and are used for setting up simultaneous equations which, in matrix notation, can be represented as

$$[P] = [K] \bullet [d] \tag{1}$$

where $[d]$ is the vector (set) of the unknown nodal displacements and $[P]$ is the vector of the loads applied at the nodes.   The matrix $[K]$ is referred to as the "assembled stiffness matrix" since all the forces on all the elements as assembled into the complete structure are taken into account.

The vector $[P]$ may, in fact, also be a matrix, if more than one load case is applied, and may be assembled from element load matrices if loads are applied to an element rather than only as nodal forces.

The stress/displacement characteristics of each element are also generated in the computer.   This may be expressed in matrix form as

$$[S] = [F] \bullet [d] \tag{2}$$

where $[S]$ is the vector of stresses and $[F]$ the "stress-matrix".   The loads $[P]$ are given and the computer determines the displacements and then calculates the stresses for each element.   It is also possible by a modification of the above procedure to specify some of the displacements $[d]$ .

(iv)    The basic stresses calculated may be further modified, eg by combination, or by calculation of principal stresses.

## Static Analysis with Geometric Stiffness

Further element stiffness matrices are derived for elements for which stiffness modification due to end-load effects is to be considered.    These are assembled with the other element matrices giving

$$[P] = [K + K_G] \bullet [d] \tag{3}$$

The rest of the static analysis is unchanged.

## Natural Frequencies

Additional element matrices are derived for the nodal forces arising from mass effects. These matrices are assembled into a structure mass matrix $[M]$ where $[M] \bullet w^2 \bullet [d]$ are the set of nodal forces arising from inertia effects from displacements $[d]$ at frequency w. The combined equation including inertia forces and stiffness forces is thus

$$(-w^2 [M] + [K]) \bullet [d] = 0 \tag{4}$$

Since the structure is assumed in free vibration, the applied forces are zero.    This is an eigenvalue problem which is solved to give natural frequencies and mode shapes $[d]$ . During this process, the number of displacements in $[d]$ may be reduced by a condensation procedure (Zienkiewicz and Cheung 1967).  Stresses may be calculated thereafter.

## Buckling Analysis

The element geometric stiffness matrices mentioned above are proportional in magnitude to the end load or direct stresses in the respective elements.    If this is assumed to be true for the whole structure, ie if all such effects are proportional with proportionality factor $\lambda$, then equation (3) can be replaced by

$$([K] + \lambda [K_G]) \bullet [d] = 0 \tag{5}$$

Since the structure is buckling, the applied loads other than those producing geometric stiffnesses are zero.    The condensation procedure can also be applied in buckling analysis.

Steady State Temperatures

This analysis is similar to static stress analysis with the variable at each node the temperature rather than displacements.   Element matrices are calculated and assembled and the equation (1) solved where $[d]$ are the nodal temperatures and $[P]$ the heat inputs and outputs.   This analysis may also be used for other potential problems, such as electric fields and fluid flows.

Mechanical Transients

For each element, mass, damping and stiffness matrices are calculated and assembled to give the following equation for the force balance at the nodes.

$$[M] \bullet [\ddot{d}] + [C] \bullet [\dot{d}] + [K] \bullet [d] = [P] \tag{6}$$

The applied loads $[P]$ may also vary with time.   The equation is solved time-step by time-step throughout the time interval specified.   The mass, stiffness and damping matrices may be condensed as above.   Stresses may be calculated as required.

Thermal Transients

This analysis results in a similar equation to the time-history analysis, except that there is no mass matrix.   The $[C]$ matrix is called the specific heat matrix and the $[K]$ matrix, the conductivity matrix.

## THEORY

No theoretical details are given in this note.   The elements details are covered in Prempeh and Patel (1971) and Sunley (1971).   The solution procedure used in Choleski's factorization;   the eigenvalue algorithm is by Jacobi (Rolston and Wilf 1962) and the time history iteration is Newmark's beta method (Chan et al 1962).

## NEWPAC ORGANISATION

The analysis of a structure using a complex finite element program involves several steps including:-   idealisation, data input, data checking and preferably display, the

analysis itself and output printing and display.    The way in which these steps are re
lated and the other auxiliary programs which may be used with NEWPAC are shown in the
block diagram (Figure 1) and described below.
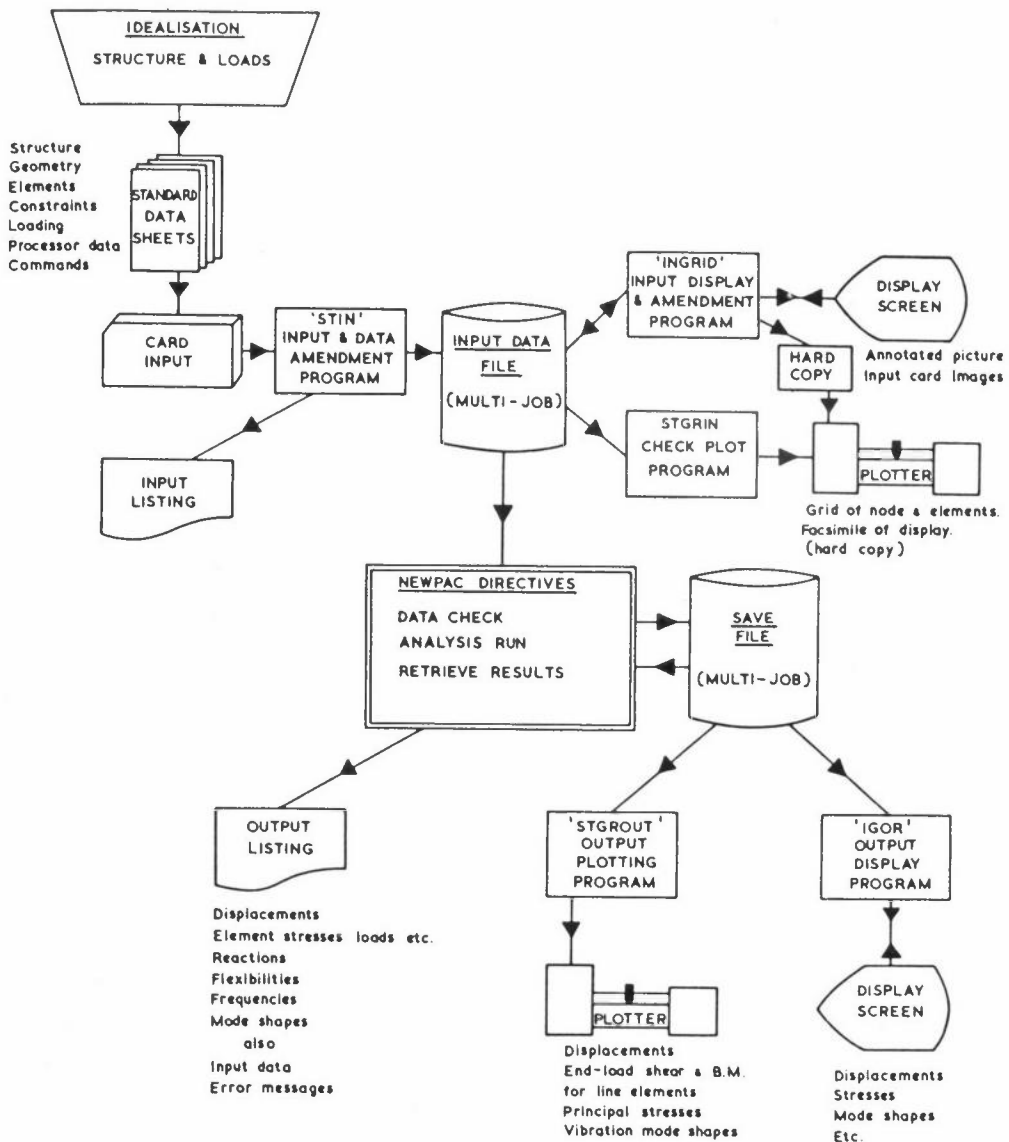


Figure 1:   Block diagram — NEWPAC system

Input

Each analysis will start from an IDEALISATION of the structure to be analysed.    A
complex structure will generally be made up of different types of elements and the user
must have sufficient knowledge of the behaviour of the various element types so that the
idealisation produces the accuracy and detail of results required.    The idealised
structure has then to be specified on STANDARD DATA SHEETS given in the format shown on

Figure 2; this includes the degrees of freedom required at each node, the node numbering, the nodal coordinates, element names and parameters, structure supports and applied loads. This data is punched on CARDS and may be entered in an INPUT DATA FILE. For some specific configurations of structure, data may be produced in the required format by special purpose programs.



Figure 2 (part 1): NEWPAC input data formats

Figure 2 (part 2): NEWPAC input data formats

## Data Checking

Complex structures require a considerable amount of data to be input and mistakes are likely to occur. The checking facilities offered within the NEWPAC system are programs STGRIN which can draw on an incremental plotter scaled pictures of the idealisation (Figure 3); program INGRID which displays views of the idealisation on

an interactive display and allows data amendment (West and Scholes 1972); and NEWPAC itself running under the DATACHECK directive. This directive terminates NEWPAC after format and consistency checks are made on geometry, elements, loads, constraints etc , and each error found produces an explanatory note on the OUTPUT LISTING. Other useful information, eg degree of freedom numbers, element band-widths, load summations are produced.
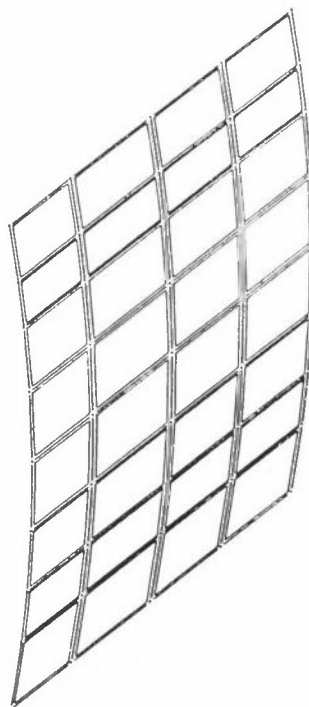
**Figure 3: Input plot of idealised structure**

```
NEWPAC INPUT PLOT                                            Y
                                                            ↑  ↗ X
C.T.LOADING VEHICLE CURVED DOOR                             └──→ Z
GRID SCALE =  0.0350         UNITS = MMS
TRI VIEW  FROM -1.00  0.60   0.60          VERTICAL AXIS = Y
```
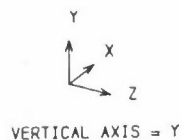
### Analysis

The second mode of operation of NEWPAC is for a main or ANALYSIS RUN; all the commands used and the amount of core store required are now checked. If there are no errors, the main analysis required is completed and results are printed and also retained on the SAVE FILE.

### Output Processing

A third mode of operation is to RETRIEVE RESULTS from the SAVE FILE, when certain further analysis or processing is possible. Additionally results held on SAVE FILE may be assessed by program STGROUT which plots the calculated deflections and

stresses superimposed on a structure grid (Figure 4) and program IGOR which draws chosen results on an interactive display (West and Scholes 1972).
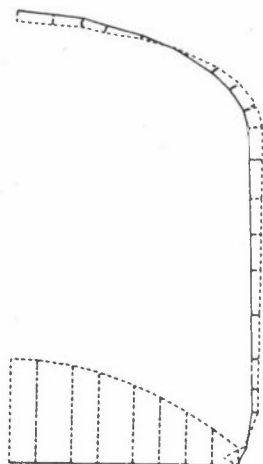


Figure 4: Output plot of bending moments plotted normal to element axes

```
NEWPAC OUTPUT PLOT
MK 3 COACH FRAME                    RUN NO.   TWO
GRID SCALE = 0.0500        UNITS = MM
Y-Z VIEW        VERTICAL AXIS = Y

LOAD CASE NO.  1 ( 0.001   N/MM2   )
MOMT 1 PLOTTED         RESULTS SCALE =.750E 05 NEWTON-MM/MM
MAXIMUM RESULTS VALUE FOR THIS LOAD CASE =   0.2199E 07
RESULTS FOR ALL RELEVANT ELEMENT TYPES PLOTTED
```

Commands

NEWPAC is guided as to the user's requirements by DIRECTIVES and COMMANDS. The former specify the modes of operation as mentioned above, DATACHECK, ANALYSIS RUN or RETRIEVE RESULTS. The latter specify the analysis and output required. Some examples are HARMONICS, LOAD COMBINATION, PRINCIPAL STRESS, REACTIONS, TRANSIENTS.

Some sequences of COMMANDS are illogical or impossible and NEWPAC determines on input whether the sequence specified may be executed in the core store available.

An error message is produced if the job cannot be run and the user is often saved an embarrassing and expensive failure.

Program Limitations

The structure of NEWPAC is such that it can run on various partition sizes on an IBM 360 or 370 computer; thus different sized structures can be analysed using the same program. However, there are many parameters which determine whether a given structure

will fit into the amount of core and disk space available.    The following limitations
apply generally but are not exhaustive.

No of nodes:                4000

No of displacements:        from nil to six per node

No of elements:             limited only by backing storage, normally more than 3000:

No of constraints:          sum of dependent and independent constraint entries within
                            any bandwidth to be less than 250

The maximum displacement number difference B and number of applied load cases, NL, for
static runs and the number of master displacements for vibration and buckling analysis,
NPD, are limited (approx) as given below.

| Core size | B | NL | NPD |
|-----------|------|----|-----|
| 180 k | 130* | 40 | 50 |
| 240 k | 180 | 50 | 60 |
| 300 k | 220 | 50 | 70 |

*100 only for vibration and buckling analysis.

Exact limits for any partition size are calculated and printed by NEWPAC.

## ELEMENTS

The NEWPAC program contains a large number of different types of finite elements.
These may be grouped as:- (a) Line elements, (b) Planar elements, (c) Axi-symmetric
elements, (d) Field elements and (e) Point elements.    Each element type has a code
name, a mnemonic formed by two initial letters of the element type name, followed by
the number of nodes to which it is incident.

For each element in the program a stiffness matrix is generated.    Other matrices,
such as the stress, mass, damping and different types of loading matrices, are
calculated depending upon the type of analysis being performed and the parameters
provided for the particular element.

The input for each element consists of an element name (user selected), the element type, the nodes to which it is connected and parameters describing its properties. For some elements a reference plane is also required. Details of the format of this input are shown in Figure 2 and details of each element type available are shown in Table 1.

## TABLE 1
## ELEMENTS AVAILABLE IN NEWPAC

| Code | Description | Displacement Form |
|------|-------------|-------------------|
| | Beams | |
| EL2 | End Load Element | linear |
| EL3 | " " " (with mid-side node) | quadratic |
| BM2 | General BeaM (with shear, torsion) | cubic bending, linear torsion etc. |
| GB2 | Grillage Beam | reduced form of BM2 |
| PB2 | Plane Frame Beam | " " " |
| BG2 | Beam with Geometric Stiffness | BM2 plus |
| BM4 | BeaM (with four corner nodes) | generalised form of PB2 |
| RB4 | Referred Beam | BM2 with rigid offsets |
| | Triangular elements | |
| MT3 | Membrane Triangle | linear |
| MT6 | " " (with mid-side nodes) | quadratic |
| BT3 | Bending Triangle (Shieh,Lee,Parmalee) | quadratic in sub-triangles |
| BT6 | " " (with mid-side nodes) | cubic (Razzaque) |
| | Quadrilateral elements | |
| MR4 | Membrane Rectangle | linear with xy terms |
| WR4 | Web Rectangle | " (but constant shear) |
| SR4 | Shear Rectangle | shear only |
| BR4 | Bending Rectangle (Zienkiewicz) | cubic with two quartic terms |
| MQ4 | Membrane Quadrilateral | derived from 4 MT3 |
| MQ8 | " " (with mid-side nodes) | " " 4 MT6 |
| BQ4 | Bending Quadrilateral | " " 4 BT3 |
| BQ8 | " " (with mid-side nodes) | cubic in sub-triangles (de Veubeke) |
| | Axisymmetric elements | |
| AT3 | Axisymmetric Triangle | linear (constant circum) |
| AA3 | Axisymmetric Anisotropic Triangle | " " |
| HT3 | Harmonic Triangle | " (harmonic circum) |
| HA3 | Harmonic Anisotropic Triangle | " " |
| AT6 | Axisymmetric Triangle (with mid-side nodes) | quadratic (constant circum) |
| HT6 | Harmonic Triangle (with mid-side nodes) | " (harmonic circum) |
| | Field elements | |
| FT3 | Field Triangle (for heat flow problems) | linear |
| FA3 | Field Axisymmetric Triangle | linear |
| FA6 | " " " (with mid-side nodes) | quadratic |
| AB2 | Axisymmetric Boundary element | linear |
| AB3 | " " " (with mid-side nodes) | quadratic |
| | Point elements | |
| PS1 | Point Stiffness | - |
| PD1 | Point Damping | - |
| PM1 | Point Mass | - |

Element Matrices

A range of element matrices may be calculated, although not all are provided for each element type. A stiffness matrix is, of course, calculated for all elements (the conductivity matrix for field elements used in heat flow problems). A "stress-matrix" is also available which allows the calculation of various stresses and stress resultants.

For beam elements, these stresses are the direct and bending stress resultants at each end. Where such an element abuts along its length to a shear carrying panel, the transfer of load between the two elements is assumed to be as a constant shear flow and the end load in the beam adjusted accordingly. Further, fibre stress can be calculated at any specified position on the beam cross-section.

The stress matrices provided for the planar elements give stress resultants (such as moments/unit length) and the stresses on both surfaces of the element (where different). Depending on the element, stresses are evaluated at the corner nodes or the element centroid. Principal stresses may also be calculated and plotted (Figure 5).

Consistent mass matrices are provided for all beam elements. However, since the mass condensation procedure is almost invariably used, it is not thought worthwhile to extend this to other elements and all other element matrices are of lumped mass type. Additional lumped masses are often necessary and these may be provided by point mass elements.

Only a very few element damping matrices are provided. However the facility exists to specify the structure damping matrix to be used for mechanical transient analysis as

$$[C] = [C_e] + a[M] + b[K]$$

where $[C_e]$ is the damping matrix assembled from the element matrices, $[M]$ and $[K]$ are the structure mass and stiffness matrices and a and b are constants.

A geometric stiffness matrix is only available with the general beam element at present, although all that is required to extend this facility to other elements is the insertion of the appropriate element routine.
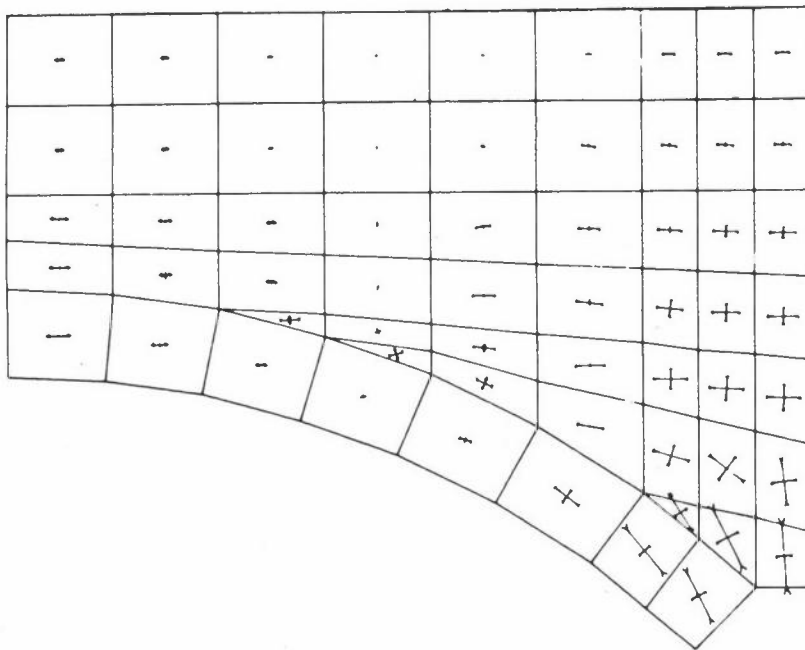
NEWPAC OUTPUT PLOT

  MASONRY  ARCH  BRIDGE                 RUN NO.    ONE

GRID SCALE =  0.0300        UNITS  =  M.
X-Y VIEW           VERTICAL AXIS = Y


LOAD CASE NO.   5 ( HRO1+SF WT+OM+LL )

PSTT    PLOTTED        RESULTS SCALE =.875E 08 N/SQ.M/M
(RESULTS SCALE TRUE ONLY IF ELEMENT PLANE PERP TO VIEW)
MAXIMUM RESULTS VALUE FOR THIS LOAD CASE =   0.1107E 07
RESULTS FOR ALL RELEVANT ELEMENT TYPES PLOTTED
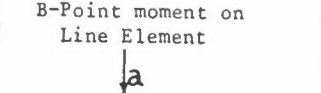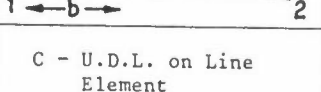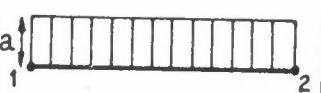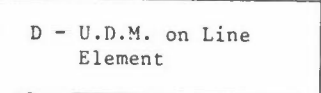
**Figure 5:  Output plot of principal stresses**

Element loads may be specified in addition to nodal loads, the program automatically generating the corresponding nodal load matrices which are subsequently assembled.  Beam elements may be subject to six types of point and line load (Figure 6).   Planar elements may be subject to uniform pressure loads normal to the panel or in one of the axis directions.

Element Axes

A right handed Cartesian set is used for nodal geometry definition.   Each element has a local set, generally with one axis parallel to one side, which is used for stress definition.

| ELEMENT TYPE | ELEMENT LOADS | LOAD TYPE | LOAD DIRN | NO.OF INPUT LINES | LOAD VALUES REQUIRED |
|---|---|---|---|---|---|
| BEAM ELEMENTS | A-Point force on Line Element  | AX AY AZ | LX LY LZ | 2 | a=magnitude of force b=distance of force from end 1. |
| | B-Point moment on Line Element  | BX BY BZ | RX RY RZ | 2 | a=magnitude of moment b=Distance of moment from end 1 |
| | C - U.D.L. on Line Element  | CX CY CZ | LX LY LZ | 1 | a-U.D.L. in force/unit length of beam |
| | D - U.D.M. on Line Element  | DX DY DZ | RX RY RZ | 1 | a=U.D.M. in moment/unit length of beam |
| | E-Linearly varying Load on Line Element  | EX EY EX | LX LY LZ | 4 | a=force - End 1 b=force - End 2 c=distance of a from End 1 d=distance of b from End 2 |
| | F-Linearly varying moment on Line Element  | FX FY FZ | RX RY RZ | 4 | a=moment - End 1 b=moment - End 2 c=distance of a from End 1 d=distance of b from end 2 |
| PLANAR ELEMENTS | U-Normal pressure load or pressure loads along one of the structure axes | UN<br><br>UX UY UZ | Normal to Panel<br><br>LX LY LZ | <br><br><br>1 | a=pressure in force per unit area |
| AXISYMMETRIC ELEMENTS | TH-Thermal Load | TH | | 1 | a=Temperature |
| | RT-Rotational Load | RT | | 1 | a=Circular frequency (Radians/Unit Time) |

a, b, c, d are the load values to be entered on the first, second, third and fourth lines of input data.

**Figure 6: Element loads**

## OTHER FACILITIES

### Axisymmetric Harmonic Analysis

For axisymmetric structures under non-axisymmetric loads, an analysis in terms of
Fourier series circumferentially may be made. NEWPAC automatically generates the
equations for all sets of harmonics, solves and recombines the separate harmonics
to give the results required at specified angular intervals. Some of these
elements have anisotropic material properties.

### Matrix input and output

A precomputed stiffness matrix may be input and a specified flexibility matrix
calculated and output.

### Load Combination

Linear combinations of deflection and stress results from different load cases may
be made. These load cases may be from different analysis runs, having different
constraints. This facility thus allows the combination of symmetric and anti-
symmetric loading on a symmetric structure, where advantage has been taken of
symmetry to analyse only part (generally half) of the structure.

### Constraints

Any displacement may be specified as having zero value, as being in a linear
relationship with other displacements, as having a pre-specified value or as being
the (summed) combination of the last two.

### Save File

All results are usually saved on a magnetic tape file, from whence they may be
reprocessed (for instance by further load combinations or to find further fibre
stresses) or plotted.

## SOME PROGRAMMING DETAILS

NEWPAC is a heavily overlaid program, having over 130 subroutines of which 50
are for generating element stiffness, stress, mass and load matrices. All but
5 of these routines are written in FORTRAN; the ASSEMBLER routines are generally

small and perform such functions as packing and unpacking words, finding the extent of core available, allowing input to be buffered and re-read in different formats.

Intermediate data is transferred between main routines via disc and a number of temporary data sets are used for this purpose.

One, labelled the OUTWELL is organised in a special way and is used for data to be retained throughout the program. The others are used as scratch areas and will generally hold only one or two "blocks" of data. A block is here defined as an adjacent set of FORTRAN logical records limited by a trailer record. Each block of data is also preceded by a header record, which gives a code specifying the type of block following. This allows this data set to be searched for a particular block. Flexibility regarding which data set any particular block of data is allocated to is maintained by holding channel identifiers as parameters in the calling sequence of subroutines.

Common

Two main COMMON areas are used, one, CONTROL, holding the main problem parameters, both structural (degrees of freedom, number of load cases etc) as well as computational (allocation of input/output channels). The other COMMON area is always arranged to be at the bottom of each branch of the overlay tree and expands to the limit of the area available. In this extendible COMMON area, generally only one variable is declared eg, $A(1)$. Assembler routines are used to determine the limit of core available to the COMMON area, leaving an appropriate amount for input/output buffers and other OS requirements, and for determining the relative position of any variable, such as $A(1)$ above, in a COMMON area. The subroutine writer then has the maximum possible amount of core available to him, though this will vary according to the partition size in which the program is run. However he must now work within one single dimensional array and distinguish different parts of the data he is working on by relative shifts, ie what would normally be $B(I)$ becomes $A(IB + I)$ and $D(I,J)$ becomes $A(ID + I \times N + J)$, where the size of D is $M \times N$. These additional book-keeping activities are essential if core is to be allocated at run-time, an absolute necessity if the program is not to be considerably constrained by the array dimensioning requirements of FORTRAN.

During the assembly, constraining and solution phases of the program, a triangular part only of the lower band of the structure matrix is held in core. This triangle is of semi-band width and effectively moves down the lower band of the structure as

columns of finished data are written out and new rows read in. The mapping of the triangle onto the single dimensional array mentioned above is such as not to require any core rearrangement at any time. A marker is also held of the first non-zero element in each row and this together with the mapping arrangment minimises the computation while maximising the bandwidth which can be held in a particular core size. A flow chart of the program is given as Figure 7.
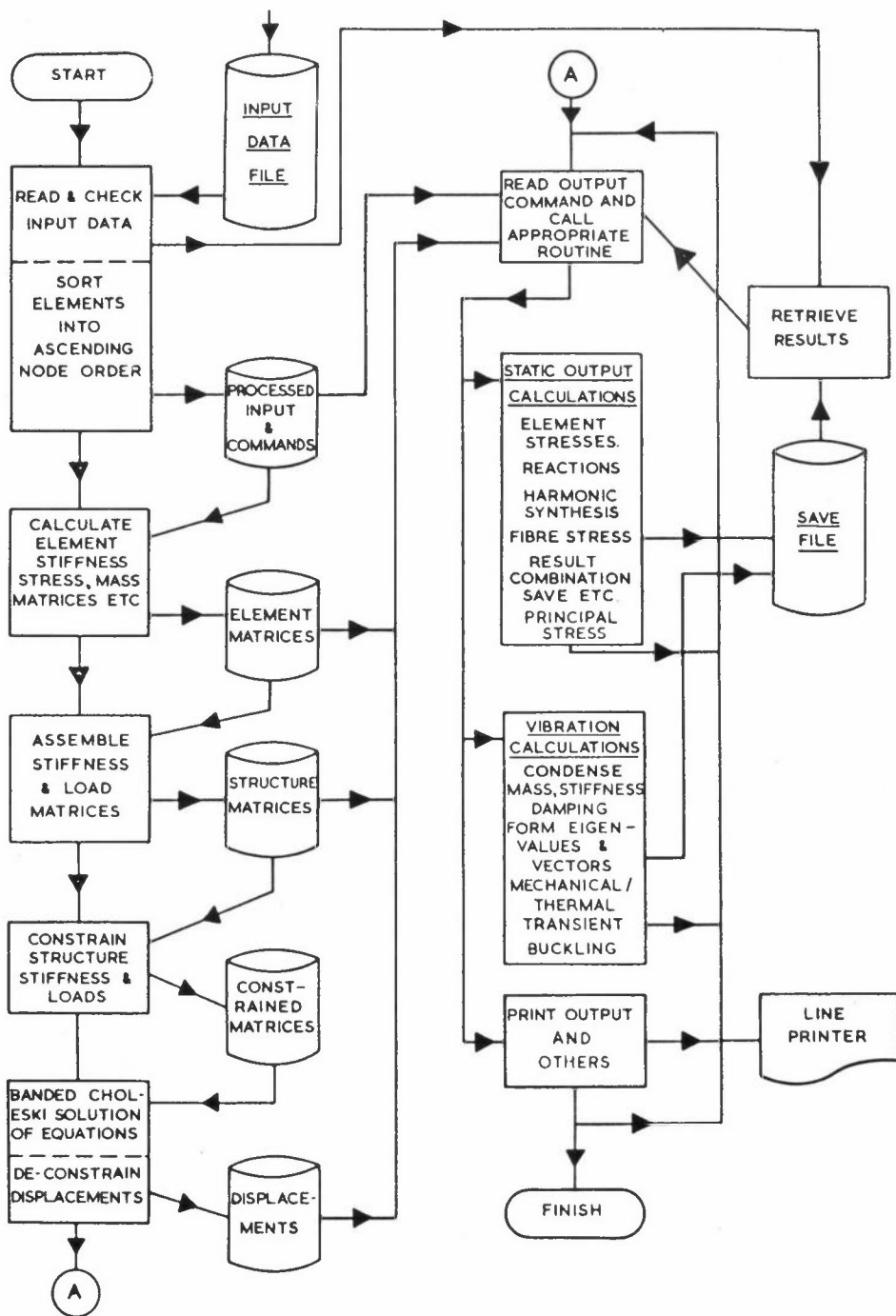


Figure 7: NEWPAC flow chart

## APPLICATIONS

Details of some of the mechanical and civil applications of NEWPAC in British Rail
are given in previous papers (Dodd 1972 and Scholes 1974). The structures analysed
at Derby as well as including railway vehicle superstructures and components, such as
bogies and wheels, have encompassed a wide variety of bridges, in steel, concrete or
brick and frame structures. Studies have also been made on details of rail joints,
pantographs on electric trains and of the overhead electrified wire itself.

The program has also been used outside BR for such diverse structures as blast
furnace hoods and bus frames and by civil engineering consultants for bridges and
dock gates.

Computational details of some of the larger jobs run on BR's IBM 370/145 are
included below (Table 2).

## TABLE 2
## SOME LARGER JOBS

| DESCRIPTION | TYPE OF RUN | DEGREES OF FREEDOM | SEMI-BAND-WIDTH | CPU (mins) |
|---|---|---|---|---|
| Coach (half) | Static | 2440 | 137 | 112 |
| Coach (quarter) | Nat frequency | 966 | 75 | 40 |
| Bus (whole) | Static | 1374 | 209 | 130 |
| Wheel | Harmonic | 6558 | 33 | 420* |
| Bridge | Static | 1920 | 77 | 361* |

* ICL 1909

## REMARKS

It is hoped that an impression has been given of a robust system, simple to use,
where the emphasis has been on providing the user with as general a program as
possible with a large number of time saving facilities on output. On input, it
has generally been left to the user to provide his own data generation program,
mainly because of the wide variety of structural types analysed.

The program has been designed to be modular in concept and operation and this has eased the enhancements made although requiring additional effort in holding source code, in overlay design and in documentation.

As the program became more complex, the process of logical checking of commands and other data to guard the unskilled user assumed a greater importance, although it undoubtedly saves much of the user's computing time.

NEWPAC is currently a mature program, well used but with plenty of potential for future enhancement to meet the needs of engineers and designers. For instance, procedures for optimisation and for checking against allowable stresses are being planned.

## ACKNOWLEDGEMENTS

## REFERENCES

CHAN S P, COX H L, BENFIELD W A, Transient Analysis of Forced Vibration of Complex Structural Mechanical Systems, Int Royal Aero Soc (1962)

DODD R J M, Application of the Finite Element Program NEWPAC to Railway Vehicle Structures, I Mech E Conference on Finite Elements, London October (1972)

PATEL N L, PREMPEH E O A, SCHOLES A, NEWPAC - A User's Guide, EMMA 15, R and D Division, British Railways Board, Issue 3 (1974)

PREMPEH E O A and PATEL N L, A Detailed Description of Finite Elements Available in NEWPAC, EMMA 19, R and D Division, British Railways Board (1971)

ROLSTON A and WILF H S, Mathematical Methods for Digital Computers, Wiley and Sons, New York (as cited by IBM Scientific Subroutines Package - Programmer's Manual) (1962)

SCHOLES A, Some Uses of Computers for the Stress Analysis of Bridges on British Rail, to be published in Civil Engineering (1974)

SUNLEY V K, The Derivation of Stiffness and Stress Matrices for the Axisymmetric Finite Elements of NEWPAC, EMMA 20, R and D Division, British Railways Board (1971)

WEST P E, Computers in the Design of Railway Structures, Computer Aided Design (1969)

WEST P E and SCHOLES A, The Introduction of Graphics into Railway Engineering, presented at On-Line 72, Brunel University (1972)

ZIENKIEWICZ O C and CHEUNG V K, The Finite Element Method in Structural and Continuum Mechanics, McGraw-Hill Publishing Company

# MESH GENERATION

J K Reid

Theoretical Physics Division
Atomic Energy Research Establishment
Harwell
Oxfordshire
0X11 ORA

## SUMMARY

A number of techniques for automatic mesh generation will be reviewed and some current
work of the author will be described.

## INTRODUCTION

When setting up a finite-element model it has been usual to position the elements and
number the variables manually.   This procedure is very tedious and prone to error and
its cost can exceed by a substantial factor the cost of running the final program that
produces the solution.   These disadvantages can be mitigated in a number of ways, for
example by the use of a good input data format allowing repeating elements and groups
of elements to be specified in a condensed form and by the extensive use of graphical
facilities to check the grid and even to specify it with the help of light-pen.   Here
we discuss the alternative of making the computer generate the whole grid.   As well
as avoiding the disadvantages we have mentioned of the simple manual approach it has at
least the potential of permitting the computer to choose for itself mesh sizes that
are everywhere appropriate both for the adequate representation of the region and for
the adequate representation of the solution.   This would be particularly useful, for
example, if a sequence of problems of widely varying smoothness were to be solved over
the same geometrical region.

## REVIEW OF TECHNIQUES

A number of authors have begun from a sequence of points on the boundary, fitting an
internal triangulation to them.   George (1971) successively removes one triangle at a
time, as illustrated in Figure 1, his algorithm trying to ensure that each is nearly
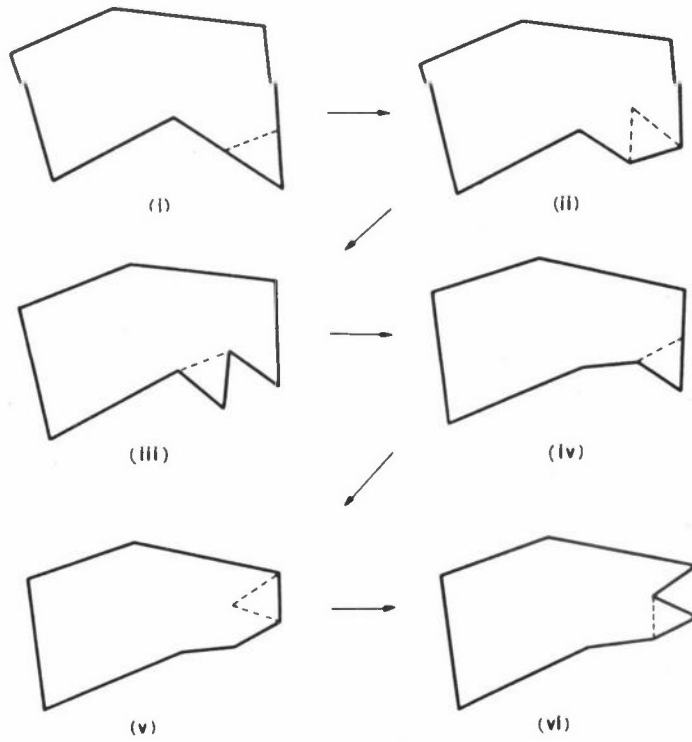
Figure 1: Illustration of George's triangulation algorithm

equilateral. A similar algorithm has been proposed by Bykat (1972) but its details
are different. If he cannot remove a triangle of reasonable shape (Figure 2 (a)) the
he introduces a cut (Figure 2(b)) which gives his removal code scope to proceed
(Figure 2(c)). George finishes by smoothing with general sweeps of the grid moving
points successively so that their coordinates are the means of the coordinates of
their immediate neighbours in the grid. An example of a final George triangulation
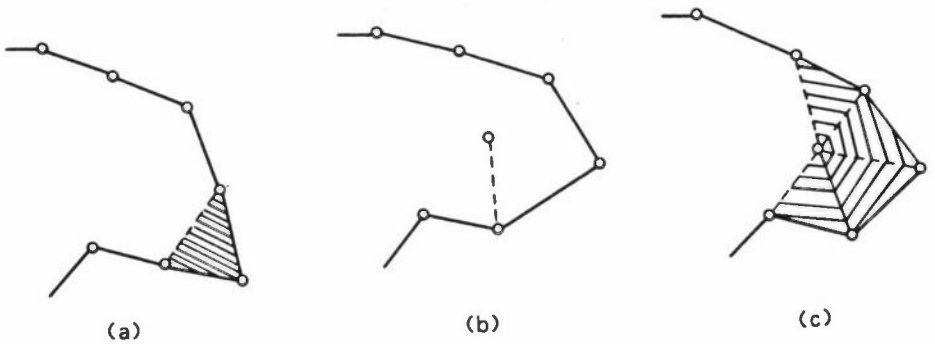


Figure 2: Illustration of Bykat's triangulation algorithm

shown in Figure 3, Bykat's triangulation of the same region is shown in Figure 4 and his triangulation of a more complicated region is shown in Figure 5.    Collier, Ellington and Rees (1971) have developed an algorithm based on the same idea and this is incorporated in TRESS, a finite element code for stress analysis problems.    A
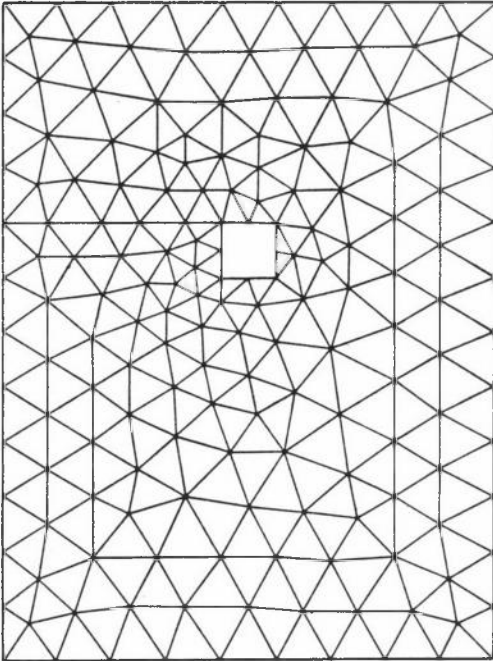


Figure 3:  A final George triangulation
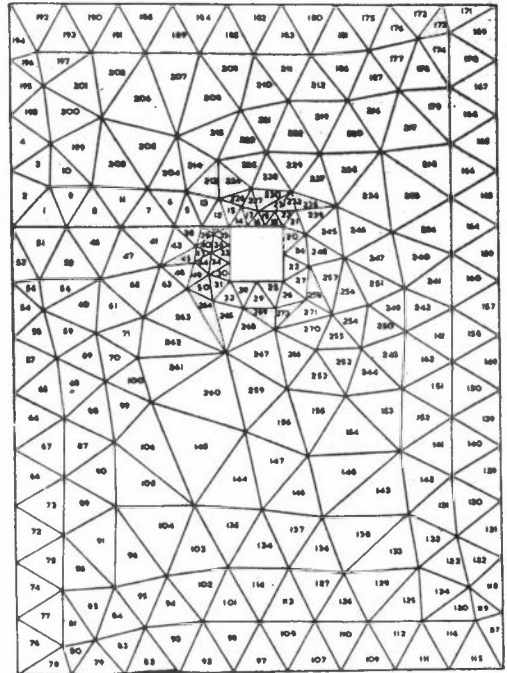


Figure 4:  Bykat's triangulation of the region of Figure 3
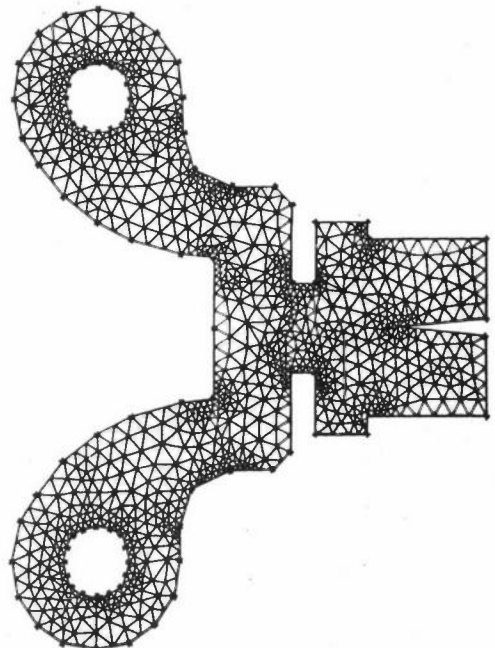
Figure 5:  Another Bykat triangulation

similar form of input has been used by Winslow (1964).   A fictitious potential is
associated with each grid point and Laplace's equation is solved so that equipotentials
provide one set of grid lines and orthogonal trajectories another.   This generates a
smooth quadrilateral grid which can be subdivided into triangles.   The solution of
Laplace's equation seems rather a "sledgehammer for cracking a nut", and one wonders
whether really complicated regions can be handled but it has the virtue of producing
smooth variations of triangle size, generally agreed to be desirable.

Another approach involves the user in providing an original gross triangulation which
is subdivided by an automatic procedure.   George (1971) begins with triangles, which
may vary in size (see Figure 6, for example), and divides each uniformly into the same
number of parts.   Zienkiewicz and Phillips (1971) begin with quadrilaterals, allow the
user to specify where on each quadrilateral side grid points are wanted, and fill in
the interior with lines parallel to the quadrilateral sides.   The idea is given
greater power by superimposing isoparametric mapping and by identifying topologically
different quadrilateral sides to allow complicated structures such as box girders with
internal diaphragms to be handled.   The results are impressive but the technique
really falls into the category of a mechanical aid to a hand method for all the
essential decisions about the shapes and positions of elements are made manually.



**Figure 6:  Another George triangulation**
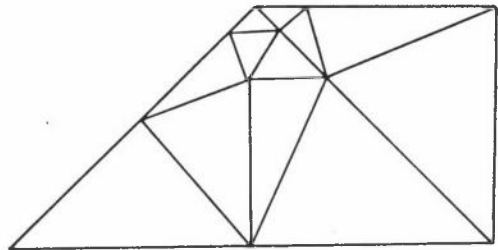
## THE AUTHOR'S WORK

### The Early Stages

My own work (Reid (1970), Reid and Turner (1970)) involved taking an original uniform
equilateral grid and distorting it by moving grid points near the boundary onto the
boundary.   In Figure 7(a) the points $1,2,\ldots16$ are points of the original uniform grid
and the points $1',2',3',4',5'$ are points of intersection of the boundary with grid
lines;   the distorted grid is shown in Figure 7(a) and is formed by moving point 1 to
position $1'$, 2 to $2'$, etc.   I was able to show that, with the help of some rules
exploiting the freedom that sometimes exists over to which boundary point a grid point
should be moved, no triangle could have an angle outside the range $\{\cot^{-1} 3\sqrt{3}, 5\pi/6\}$
$\approx \{11°, 150°\}$ and that with a smooth boundary and a sufficiently fine grid these bounds

could be improved to $\{\cot^{-1} (5/\sqrt{3}), \ 2\pi/3\} \approx \{19^{\circ}, \ 120^{\circ}\}$. The procedure produced use-
ful triangulations, particularly for the practical problem of A B Turner that first
aroused my interest in the subject, and one of the triangulations produced for him is
shown in Figure 7 (b). It has, however, been justly criticized for the fact that the
grid on which it is based is uniform so that an unreasonably large number of elements
may be produced just because the geometry of the region or bad behaviour of the
solution may require this in one small area. A further disadvantage is that I allowed
for holes but not for internal interfaces.
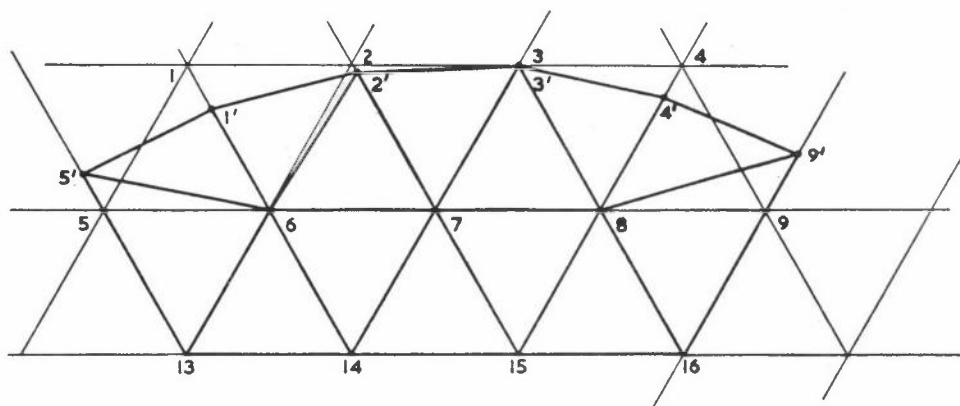


Figure 7(a): Grid distortion at a boundary
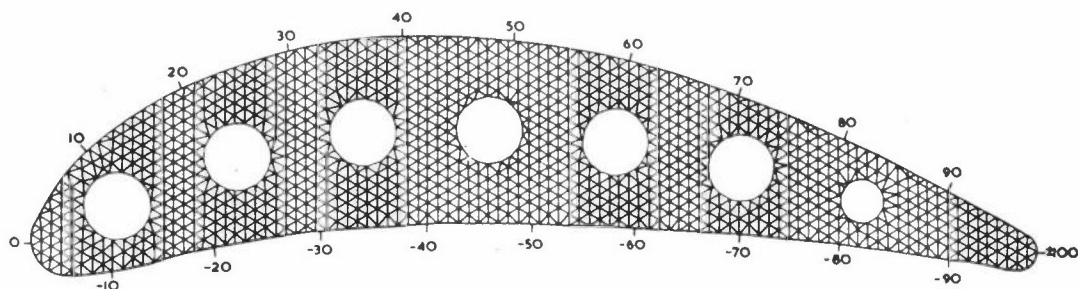


Figure 7(b): A final triangulation incorporating boundary grid distortion

Recent Developments

My current work is aimed at overcoming these deficiencies and also at allowing the
automatic choice of suitable grid sizes everywhere in the region. The work is not yet
complete but I propose to tell you of my progress so far. I have replaced the earlier
uniform grid by a grid consisting of equilateral triangles and bisected equilateral
triangles, an example of which is shown in Figure 8. I take an original equilateral
triangle and subdivide it successively to any depth (within reason) but see no need to
be able to reverse the process. Using only equilateral triangles and bisected
equilateral triangles means that dividing one triangle may necessitate dividing several

neighbouring ones too.   For example dividing the triangle 6-7-9 of Figure 8 will
necessitate triangles 3-5-6 and 7-8-9 having their bisectors removed and then being
divided into four.   The topology of this structure is very simple and can conveniently
be stored using three pointers in association with each node giving the neighbouring
nodes in the directions 0, $2\pi/3$, $4\pi/3$ respectively (see Figure 8).   In each case the
dummy value zero is used when there is no neighbour in the relevant direction either



Figure 8:  A Reid triangulation
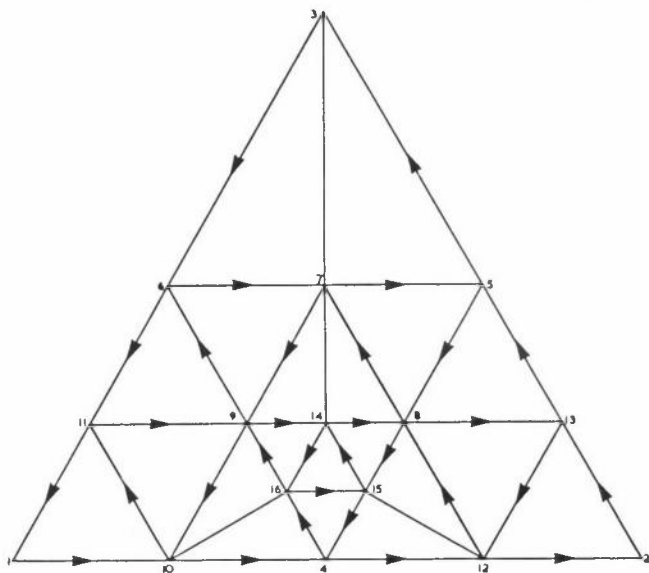
because we are at the edge of the big triangle or because of a local change of mesh-
size.   It is actually unnecessary to hold the coordinates of the nodes and I have
written a subroutine which holds a pointer consisting of a node number, its x and y
coordinates and the local mesh size.   The subroutine is able to move this pointer
around to find which triangle contains any given point.   However I have decided
against this scheme in favour of holding the coordinates explicitly because this gives
code that will execute faster and I do not anticipate shortage of storage on our
machine;   in any case the extra storage can be used later to hold the finite-element
matrix.


Following my earlier work I fit the real boundary by distorting the grid.   Each smooth
section of boundary or internal interface is specified by the user in parametric form
(eg (t,t), t = 0,1, representing the line joining (0,0) to (1,1)) and is represented
by a sequence of adjacent grid points each of which is moved to a nearby point of the
boundary.   An obvious logical problem exists in the choice of this sequence.   It must
be such that the topology of the given region is correct; for example separate internal
holes must not be connected to each other or to the outside of the region.   Also we
want to be able to handle efficiently the case where a solution on a finer grid is
wanted because the present grid has not produced a solution that is sufficiently
accurate everywhere.   In fact, I am encouraged by the work of Rice (1973) to hope that

many singularities could be treated automatically without the user even knowing of
their existence.)    I therefore use separate double-linked lists for the boundaries
so that each can be traversed easily in either direction and extra points introduced
conveniently.    Each element of one of these lists contains, in addition to forward
and backward pointers to other elements in the same list, the coordinates of a point
on the boundary and a pointer to an associated node in the undistorted grid.    It is
also necessary to give each node a pointer to its associated boundary point (if any).
I choose the node to associate with a boundary point by the simple device of taking
the nearest.    This means that each node has a "catchment area" of possible positions
of boundary points associated with it.    The catchment area of point 8 of Figure 8 is
shown in Figure 9.    We aim to find a sequence of boundary points in distinct adjacent

Figure 9:  Catchment area of point 8 of
           Figure 8
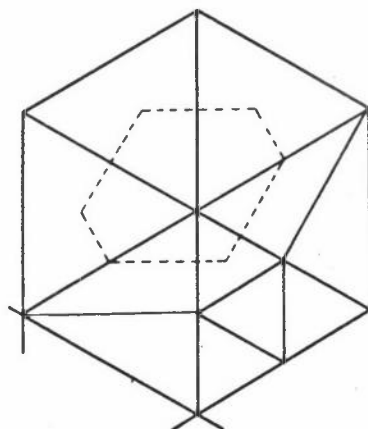
catchment areas.    If adjacent boundary points are not in adjacent catchment areas then
we search between them for a boundary point in another catchment area and insert it
into the boundary linked list.    Two distinct boundary points must not be associated
with the same node, for we cannot move the node to both.    Occasionally this can
happen because two adjacent boundary points are in the same catchment area and in this
case we simply remove one of them from its linked list.    More usually it is because
the local step length is too large so that two different boundaries or remote parts of
the same boundary have entered the same catchment area;    in this case it is clear that
a local refinement of the mesh is needed if the geometry is to be represented properly
so we divide into four parts the equilateral triangle containing the midpoint of the
line joining the two boundary points which were associated with the same node.
Dividing any triangle alters the catchment area of its vertices so we reset to the
dummy value zero their pointers to boundary points in case the old assignment is now
wrong.    Following a division we backtrack along the boundary on which we are currently
working until an unaffected point is reached and recommence from there.    Of course other
boundaries may be affected too so we continue until all of them are swept successively
without any divisions being caused.    Another reason for requiring subdivision is when
a distorted triangle is found to be pathologically thin for it can be caused by the

situation illustrated in Figure 10 where the equilateral triangle is distorted to the nearly flat triangle 1'2'3'.   The situation does not occur with the divided region because the boundary is then near a grid line.   Currently I subdivide if any angle is less than $0.3^c \approx 17.2^o$ and I hope to show that such a limit will not cause any difficulties.



Figure 10:  A situation that produces a pathologically thin triangle
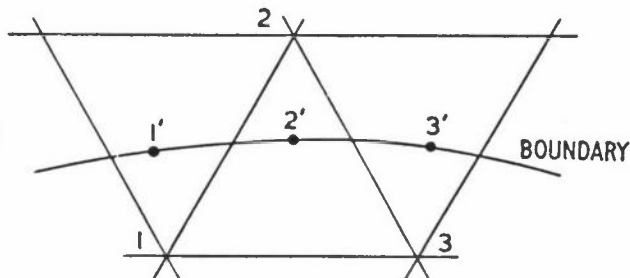
For data input I require the user to write a subroutine which gives each smooth section of the boundary in parametric form, and to specify region numbers on both sides of each boundary.   He may also specify a desired local step-size for each point on any boundary and (separately) for each point of the whole region.   An original undistorted triangulation satisfying the whole region requirement is set up and as each boundary point is found a check is made on the local step-size.   An example of the use of the "whole region" step-size control is shown in Figure 11 where the function $h(x,y) = 0.2(x^2+y^2)^{1/3}$ was used to allow for a potential singularity at the origin
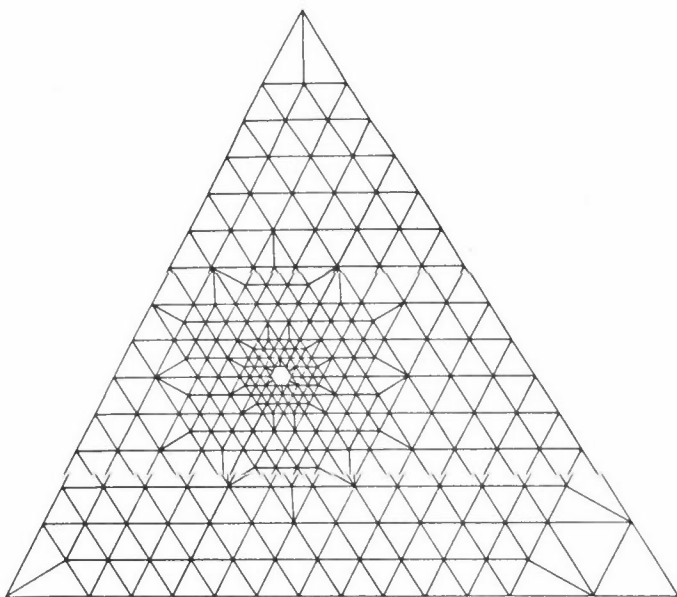


Figure 11:  An example of "whole region" step size control
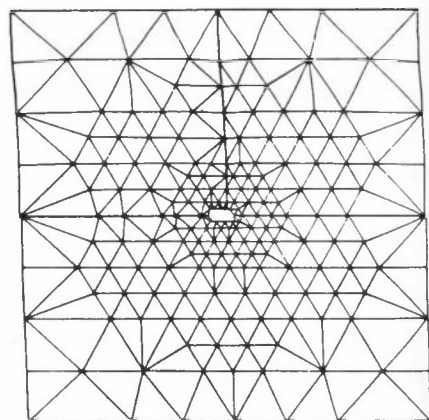


Figure 12:  Distorted grid for the example of Figure 11

caused by an internal interface.   The distorted grid for this case is shown in Figure 12.   A triangulation of a simpler region is shown in Figure 13 and the same region is shown retriangulated in Figure 14 after a refinement in the top left-hand corner, as might be caused by the poor behaviour of the solution of a differential equation there.   A more complicated region with two circular holes and six circular
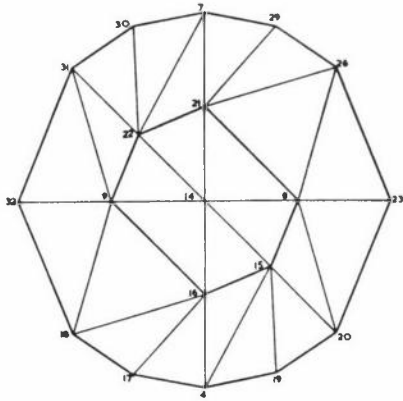
**Figure 13: A triangulation of a simple region**
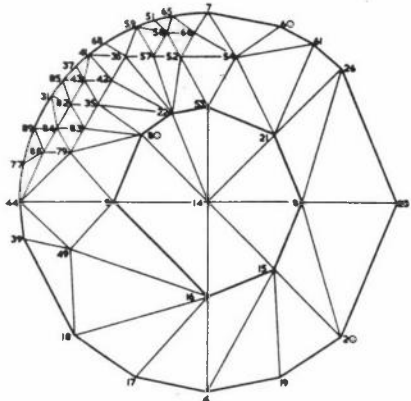


**Figure 14: Retriangulation of the region of Figure 13 after refinement**

internal interfaces is shown triangulated in Figures 15 to 17, the three figures being produced by using successively smaller values for the required mesh size as the boundaries are traversed.
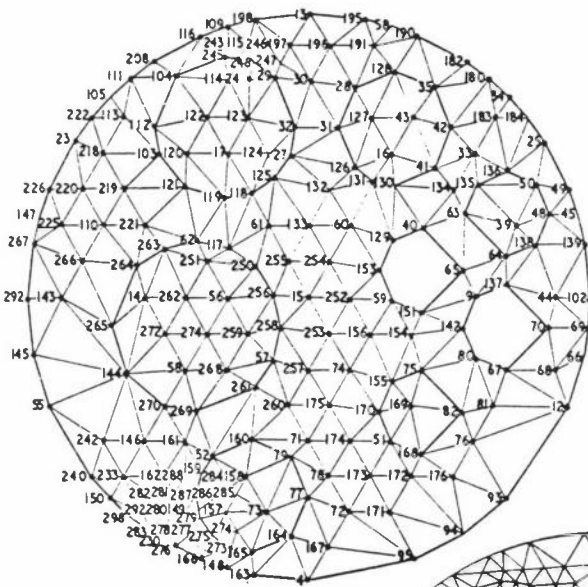


**Figure 15**
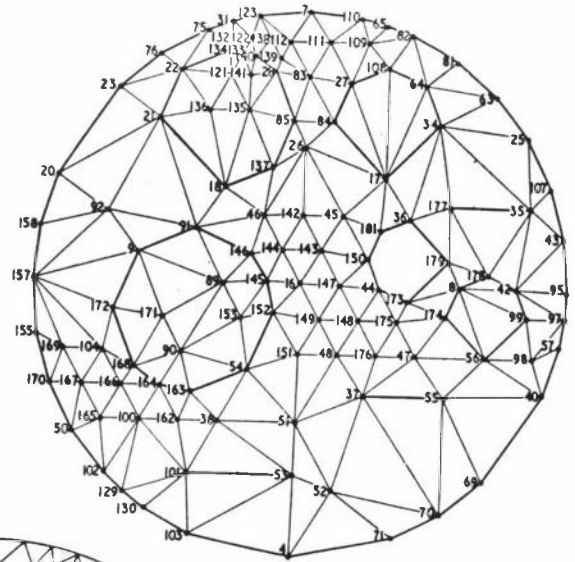


**Figure 16**
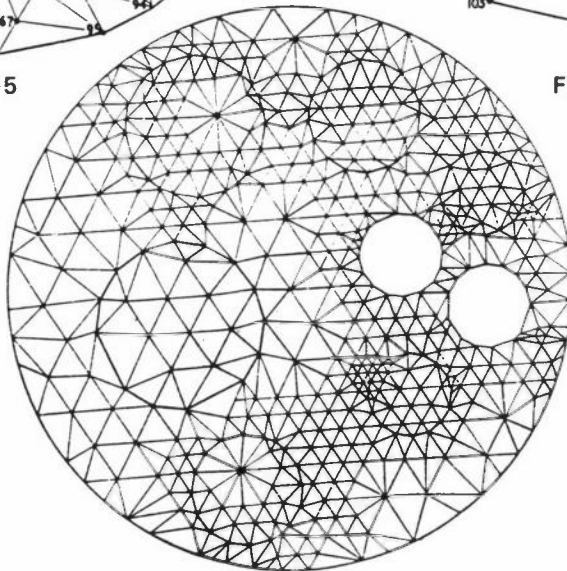


**Figure 17**

## REFERENCES

BYKAT A, Automatic Triangulation of Two-Dimensional Regions,
Report ICSI 420, Institute of Computer Sciences, London (1972)

COLLIER W D, ELLINGTON J P and REES P M, TRESS, A Finite Element Code for
Stress Analysis Problems, UKAEA Risley report, TRG 1813(R) HMSO (1971)

GEORGE J A, Computer Implementation of the Finite Element Method,
Report STAN-CS-71-208, Stanford University (1971)

REID J K, On the Construction and Convergence of a Finite Element Solution
of Laplace's equation.   J Inst Maths Applics, $\underline{9}$, 1-13 (1970)

REID J K and TURNER A B, Fortran Subroutines for the Solution of Laplace's
equation over a General Region in Two Dimensions, Report TP 422, AERE,
Harwell (1970)

RICE J R, An Educational Adaptive Quadrature Algorithm.   SIGNUM Newsletter
$\underline{8}$, 2, 27-41 (1973)

WINSLOW A M, Equipotential Zoning for Two-Dimensional Meshes, Report VCRL-7312,
Lawrence Radiation Laboratory, University of California (1964)

ZIENKIEWICZ O C and PHILLIPS D V, An Automatic Mesh Generation Scheme for Plane
and Curved Surfaces by "Isoparametric" Coordinates, Int J Num Meth Engng,
$\underline{3}$, 519-528 (1971)

# MARGINAL STABILITY: EXPERIMENTS WITH THE DIFFUSION EQUATION

Winifred L Wood

Department of Mathematics
University of Reading
Whiteknights
Reading
Berks

## SUMMARY

Should Crank-Nicholson noise effects be dealt with as and when they arise or should they be avoided?   This paper looks at both strategies in a series of experiments with the linear diffusion equation with finite element and finite difference methods and reaches a tentative conclusion that a simple averaging process is the best.

## INTRODUCTION

This paper contains an account of a number of experiments made by Roland Lewis of Swansea and the author.   These were inspired by troublesome noise effects experienced when using the Crank-Nicholson method in time together with finite element discretisation in space.   We decided to take a simple linear heat conduction problem to which we knew the exact answer and to conduct experiments to test various ways of dealing with the noise and to compare the accuracies of these and other time-stepping schemes.

So that standard element subroutines can be used the problem is taken to be that of heat conduction in a bar 4 units long and 1 unit wide;  ie made up of four square elements.   The elements used are "serendipity" type (Zienkiewicz 1971):  linear, quadratic and cubic with 4, 8, 12 nodes respectively.   The problem becomes essentially one-dimensional by supposing the bar is insulated along its length as well as at one end.   It is initially at zero temperature and a temperature of unity is applied at the other end for $t > 0$.

In following sections the exact solution is obtained and compared with the solution of the space-discretised problem and the solution of the generalised two-level time difference scheme.   The origin of the noise effects is discussed and two possible

ways of dealing with them are analysed.    Two methods which do not produce noise
effects are also analysed;    these are the backward difference method and a method
suggested by Richtmyer and Morton (1967) which has the same accuracy as Crank-
Nicholson.    Results for the corresponding finite difference method are also included
for comparison.    All these methods have been assessed for their average percentage
relative errors and the results are presented in Tables I, II and III.

## THE PROBLEM

The problem considered is represented by the non-dimensionalised linear heat
conduction equation

$$\frac{\partial c}{\partial t} = \frac{\partial^2 c}{\partial x^2}, \qquad 0 < x < 4 \tag{1}$$

with $c = 0$ at $x = 0$, $\frac{\partial c}{\partial t} = 0$ at $x = 4$; $t > 0$

$$\left.\begin{array}{l} \\ \\ \\ \end{array}\right\} \tag{2}$$

and $c(x, 0)$ given, $0 < x \leq 4$

This represents heat conduction in a bar of length 4 units;    c   is the temperature.

For time   $t > 0$   the end   $x = 0$   of the bar is maintained at unit temperature.    The
other end of the bar is insulated so the steady state solution is unit temperature
for the whole bar.
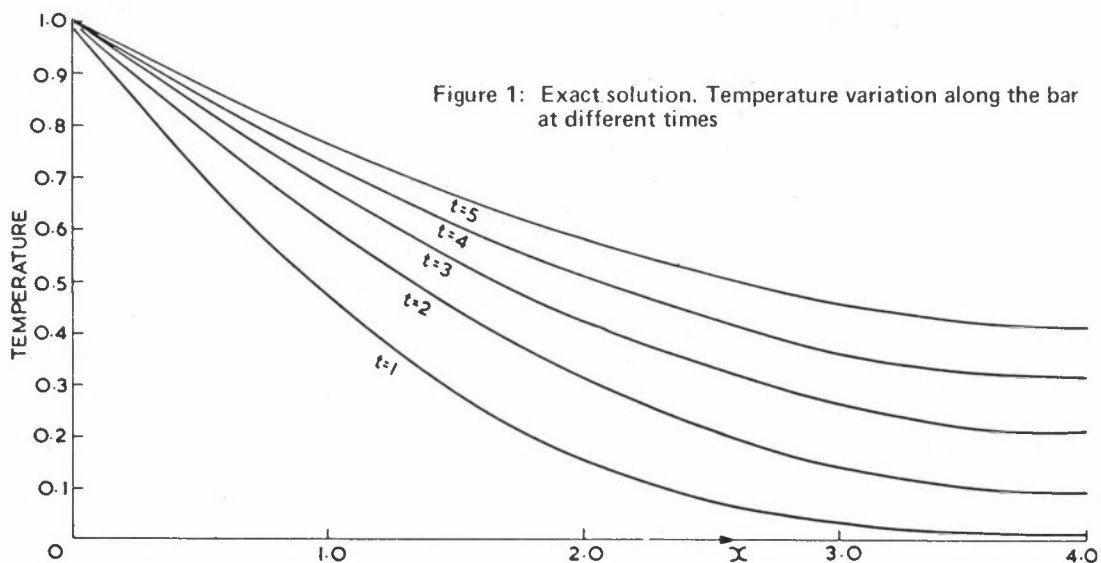


Figure 1: Exact solution. Temperature variation along the bar at different times

If the bar is initially at temperature $c = 0$ the exact solution is

$$c = 1 - \sum_{k=1}^{\infty} \frac{4}{\pi(2k - 1)} \exp\{- \left[\frac{(2k - 1)\pi}{8}\right]^2 t\} \sin\left[\frac{(2k - 1)\pi x}{8}\right] \qquad (3)$$

As $k$ increases the exponential factor decreases so rapidly that for practical purposes only the first few terms have any influence. This point is important when comparisons are made with the approximate solution. To fix ideas the exact solution is shown in Figure 1.

Suppose the spatial finite element discretisation of equations (1) and (2) gives the system of ordinary differential equations

$$N\underline{\dot{c}} = -K(\underline{c} - \underline{c}^{\infty}) \qquad (4)$$

where $\underline{c}$ is the vector of unknowns at the nodes at time $t$ and
$\underline{c}^{\infty}$ is the steady state value of $\underline{c}$,
$N$ is the mass matrix $\{(\phi_i, \phi_j)\}$ and
$K$ is the stiffness matrix $\left\{\left(\frac{d\phi_i}{dx}, \frac{d\phi_j}{dx}\right)\right\}$

where $\{\phi_i\}$ is the set of basis functions.

Assuming $A = N^{-1}K$ has a complete set of eigenvectors $\underline{u}_i$ with corresponding eigenvalues $\mu_i$ we can write

$$\underline{c}(0) = \Sigma \alpha_i \underline{u}_i, \qquad \underline{c}^{\infty} = \Sigma \gamma_i \underline{u}_i \qquad (5)$$

The solution of the system of equations (4) is then

$$\underline{c} = \Sigma \left[\gamma_i + (\alpha_i - \gamma_i)e^{-\mu_i t}\right]\underline{u}_i \qquad (6)$$

However if equation (4) is replaced by the time-difference scheme

$$N \frac{(\underline{c}^{n+1} - \underline{c}^n)}{\Delta t} = -K\left[\theta \underline{c}^{n+1} + (1 - \theta)\underline{c}^n\right] + K\underline{c}^{\infty}, \quad 0 \le \theta \le 1 \qquad (7)$$

this is the recurrence relation

$$(I + \theta \Delta t\, A)\underline{c}^{n+1} = \left[I - (1 - \theta)\Delta t\, A\right]\underline{c}^n + \Delta t\, A\underline{c}^{\infty} \qquad (8)$$

with general solution

$$\underline{c}^n = \Sigma\left\{\gamma_i + (\alpha_i - \gamma_i)\left[\frac{1 - \Delta t(1 - \theta)\mu_i}{1 + \Delta t\theta\mu_i}\right]^n\right\}\underline{u}_i \qquad (9)$$

Comparing equations (6) and (9) we see that $e^{-\mu_i t}$ is being approximated by

$$\left[\frac{1 - \Delta t(1 - \theta)\mu_i}{1 + \Delta t\theta\mu_i}\right]^n \tag{10}$$

The scheme is stable for any value of $\Delta t$ if $\frac{1}{2} \le \theta \le 1$.

The expression (10) gives values of alternating sign if

$$\Delta t > \frac{1}{(1- \theta)\mu_i} \tag{11}$$

Thus there is no oscillation effect with the backward difference scheme ($\theta = 1$).

The Crank-Nicholson (or trapezoidal rule) scheme with $\theta = \frac{1}{2}$ is equivalent to approximating $e^{-\mu_i \Delta t}$ by the (1,1) Padé approximation

$$\frac{1 - \frac{1}{2}\Delta t\mu_i}{1 + \frac{1}{2}\Delta t\mu_i}$$

This scheme is very popular because it is unconditionally stable and has accuracy $O((\Delta t)^2)$. The formula

$$\left[\frac{1 - \frac{1}{2}\Delta t\mu_i}{1 + \frac{1}{2}\Delta t\mu_i}\right]^n \tag{12}$$

where $n\Delta t = t$, gives a very good representation of $e^{-\mu_i t}$ for smaller values of $\mu_i\Delta t$ but if $\mu_i\Delta t$ becomes large the expression (12) $\to (-1)^n$ whereas $e^{-\mu_i t} \to 0$. This is the origin of the Crank-Nicholson noise effect and the reason why the method may be described as "marginally" stable.

There is no oscillation from any mode if

$$\Delta t < \Delta t_{crit} = \frac{2}{\max \mu_i} \tag{13}$$

The greater the number of degrees of freedom the smaller is the critical value of $\Delta t$ because the $\mu_i$ are approximations to the eigenvalues

$$\frac{(2i - 1)^2 \pi^2}{64} \tag{14}$$

of the original problem which appear in equation (3).

The $\mu_i$ obtained with the finite element method are overestimates of the corresponding eigenvalues (14). Hence with n degrees of freedom

$$\max \mu_i > \frac{(2n - 1)^2 \pi^2}{64} \tag{15}$$

and $\Delta t_{crit} < \dfrac{128}{(2n - 1)^2 \pi^2}$ (16)

For each of the three types of element used, linear, quadratic and cubic the matrices N, K have been recovered and their eigenvalues found. The corresponding values of $\Delta t_{crit}$ are

linear element      0.187
quadratic element  0.049
cubic element      0.012

As the number of degrees of freedom is increased either by increasing the number of elements or by increasing the degree of the basis functions, the critical value of $\Delta t$ is decreased. It is clearly out of the question to avoid noise effects by keeping $\Delta t$ below the critical value.

Figures 2 and 3 show typical results from solving equation (8) with zero initial starting vector ($\alpha_i = 0$ for all i), for $\Delta t = 0.5$ and 4.0. The result for $\Delta t = 4.0$ shows an effect which is also noticeable for $\Delta t = 8.0$, namely a dying away
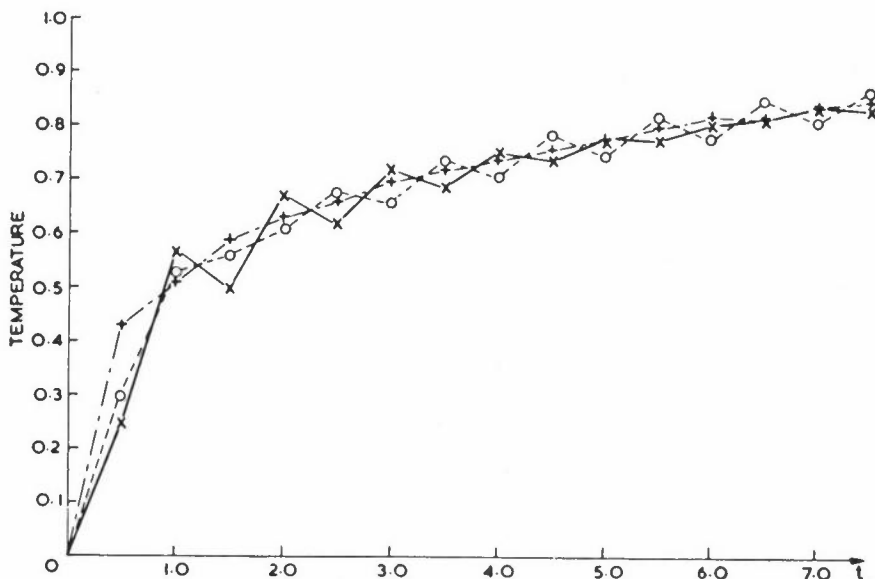


Figure 2: Temperature variation at x = 1.0 for $\Delta t$ = 0.5 (Crank-Nicholson)
+ Linear  x Quadratic  0 Cubic
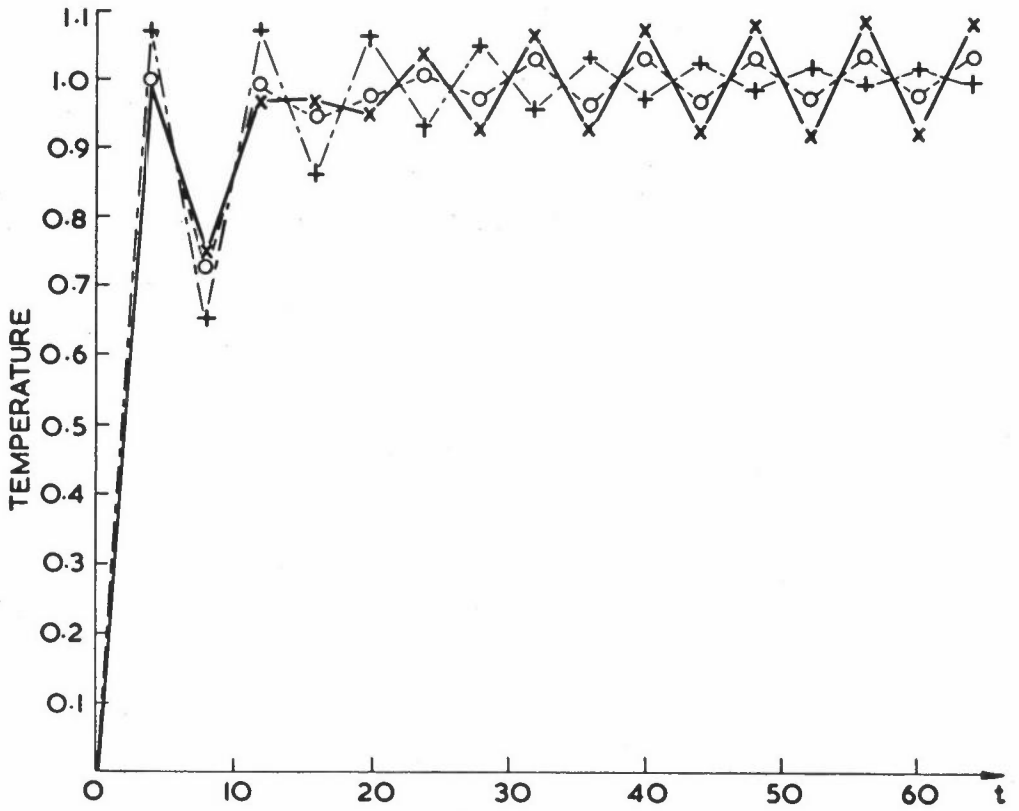
Figure 3:  Temperature variation at x = 1.0 for $\Delta t$ = 4.0 (Crank-Nicholson)
+ Linear   x Quadratic   0 Cubic

and subsequent increase in the amplitude of the oscillations for the quadratic
and cubic elements.   The $\Delta t$ = 8.0 case when run for 100 time steps to see how this
develops merely continues with the hunting about the true solution.   The factors are
-0.956, -0.988, -0.996 for the linear, quadratic and cubic elements respectively.
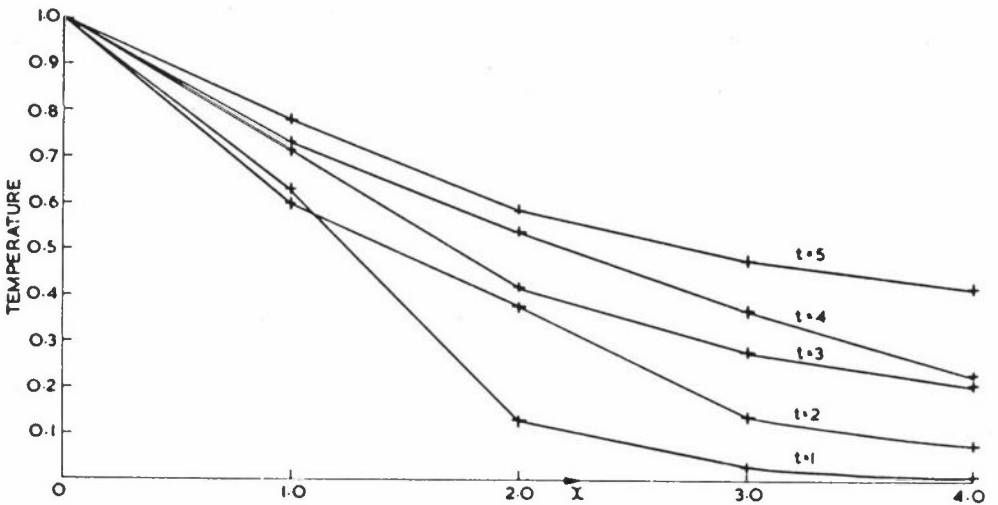


Figure 4:  Linear element. (Crank-Nicholson) $\Delta t$ = 1.0
Variation of temperature along bar at different times

Figures 4, 5, 6 show the variation of temperature along the bar for the first few time steps for the three different elements.
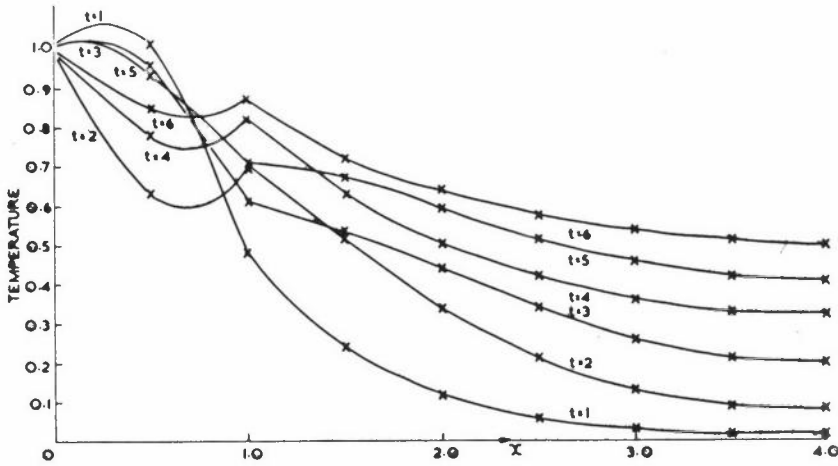


**Figure 5:** Quadratic element. (Crank-Nicholson) Δt = 1.0
Variation of temperature along bar at different times
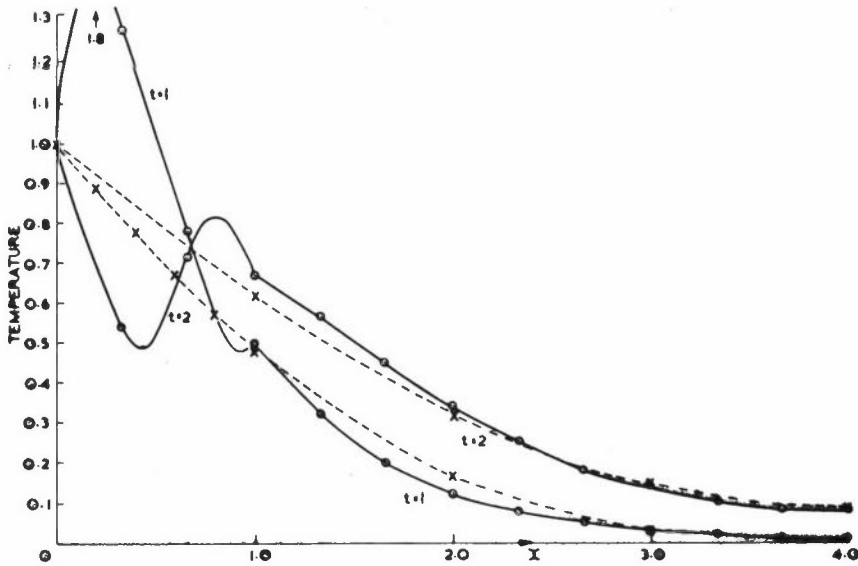


**Figure 6:** Cubic element. (Crank-Nicholson) Δt = 1.0
Variation of temperature along bar for t = 1,2
Computed results 0   Exact results x

## METHODS OF DEALING WITH THE NOISE EFFECT

(1)  One way of dealing with the noise effect from Crank-Nicholson is to take the mean of the initial values and the values at the end of one time step (with Crank-Nicholson) and use these values as new starting values thus obtaining solutions at $\frac{\Delta t}{2}$, $3\frac{\Delta t}{2}$, $5\frac{\Delta t}{2}$ etc.   The result with this linear problem is just the same from continually taking the mean of succssive values.

If $\quad \underline{v}^n = \dfrac{\underline{c}^n + \underline{c}^{n+1}}{2}$

$\hfill(17)$

then from equation (8) with $\theta = \frac{1}{2}$,

$$(I + \tfrac{1}{2}\Delta t\ A)\underline{v}^{n+1} = (I - \tfrac{1}{2}\Delta t\ A)\underline{v}^n + \Delta t\ A\underline{c}^{\infty} \qquad(18)$$

Hence $\underline{v}^n$ satisfies the same recurrence relation as $\underline{c}^n$ but with starting value

$$\underline{v}^o = \frac{\underline{c}^o + \underline{c}^1}{2}$$

$$= \Sigma\ \frac{(2\alpha_i + \Delta t\mu_i\gamma_i)\underline{u}_i}{2 + \Delta t\mu_i} \qquad (\text{from (9) with } \theta = \tfrac{1}{2})$$

Hence the solution of equation (18) is

$$\underline{v}^n = \Sigma\left\{\gamma_i + \frac{(1 - \tfrac{1}{2}\Delta t\mu_i)^n}{(1 + \tfrac{1}{2}\Delta t\mu_i)^{n+1}}\ (\alpha_i - \gamma_i)\right\}\underline{u}_i \qquad(19)$$

This now has $\exp\left[-\mu_i(t + \tfrac{1}{2}\Delta t)\right]$ approximated by

$$\frac{(1 - \tfrac{1}{2}\Delta t\mu_i)^n}{(1 + \tfrac{1}{2}\Delta t\mu_i)^{n+1}} \qquad(20)$$

There is still oscillation if $\frac{1}{2}\Delta t\mu_i < 1$ but for large $\Delta t\mu_i$ the expression (20) does $\to 0$ so that taking the mean of successive values obtained from Crank-Nicholson does damp out the noise effect.

Figures 7, 8 illustrate the improvement made by averaging.
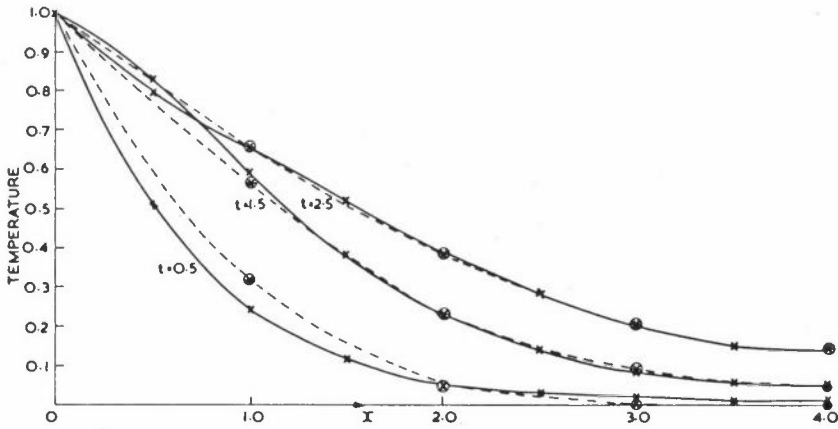
Figure 7: Quadratic element. (Averaged Crank-Nicholson). $\Delta t = 1.0$
Variation of temperature along bar at different times.
Computed results x   Exact results ⊛
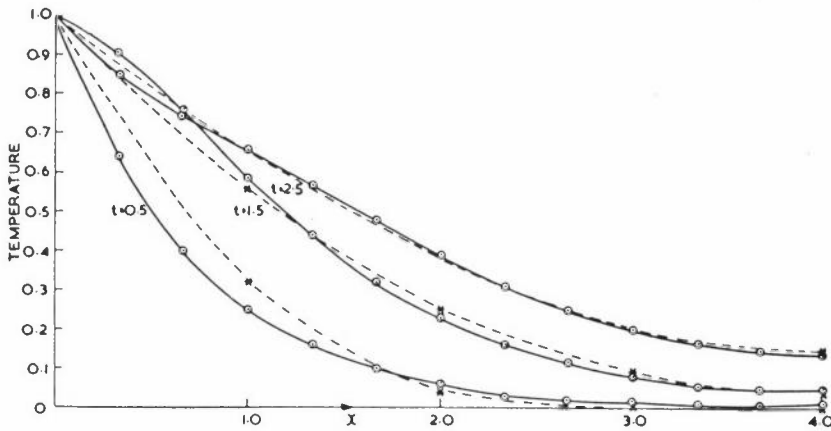


Figure 8: Cubic element. (Averaged Crank-Nicholson) $\Delta t = 1.0$
Variation of temperature along bar at different times.
Computed results ϴ   Exact results ∗

(2)    Another method for dealing with the noise effects is equivalent in the
present problem to raising the temperature at the end $x = 0$ by $\frac{1}{m}$ in each
of  m  time steps instead of making it jump to unity in one time step.

We write the recurrence relation (8) with $\theta = \frac{1}{2}$ in the form

$$\underline{c}^{n+1} = B\underline{c}^n + \underline{h} \tag{21}$$

where

$$\mathbf{B} = (I + \tfrac{1}{2}\Delta t \ A)^{-1}(I - \tfrac{1}{2}\Delta t \ A)$$

$$\underline{h} = (I + \tfrac{1}{2}\Delta t \ A)^{-1}\Delta t \ A\underline{c}^{\infty}$$

(22)

Then the m-step method takes

$$\underline{\tilde{c}}^k = B\underline{\tilde{c}}^{k-1} + \frac{k}{m} \ \underline{h} \ , \qquad 1 \le k \le m \tag{23}$$

Put $\underline{\omega}^k = \underline{\tilde{c}}^k - \underline{\tilde{c}}^{k-1}, \qquad 1 \le k \le m$

then $\underline{\omega}^k = B\underline{\omega}^{k-1} + \frac{1}{m}h, \quad 2 \le k \le m$

The matrices A and B have a common basis of eigenvectors
and the eigenvalues of B are

$$\lambda_i = \frac{2 - \Delta t\mu_i}{2 + \Delta t\mu_i} \tag{24}$$

With $\underline{\tilde{c}}^0 = \underline{0}$ and $\underline{h} = \Sigma\beta_i\underline{u}_i$ we have

$$\underline{\omega}^1 = \sum_i \frac{\beta_i}{m} \ \underline{u}_i$$

and $\underline{\omega}^k = B^{k-1}\underline{\omega}^1 + \frac{1}{m}(I - B)^{-1}(I - B^{k-1})\underline{h}$

(25)

Hence

$$\underline{\tilde{c}}^k - \underline{\tilde{c}}^{k-1} = \Sigma \ \frac{\beta_i}{m(1 - \lambda_i)}(1 - \lambda_i^k)\underline{u}_i$$

$$\underline{\tilde{c}}^{k-1} - \underline{\tilde{c}}^{k-2} = \Sigma \frac{\beta_i}{m(1 - \lambda_i)}(1 - \lambda_i^{k-1})\underline{u}_i$$

$$- - - - - - - - - -$$

$$\underline{\tilde{c}}^1 - \underline{\tilde{c}}^0 = \Sigma \frac{\beta_i}{m(1 - \lambda_i)}(1 - \lambda_i)\underline{u}_i$$

and $\underline{\tilde{c}}^0 = \underline{0}$

Hence addition gives

$$\underset{\sim}{\underline{c}}^k = \Sigma \left[ \frac{\beta_i k}{m(1 - \lambda_i)} - \frac{\beta_i \lambda_i (1 - \lambda_i^k)}{m(1 - \lambda_i)^2} \right] \underline{u}_i$$

$$= \Sigma \frac{\gamma_i}{m} \left[ k - \frac{\lambda_i (1 - \lambda_i^k)}{1 - \lambda_i} \right] \underline{u}_i \tag{26}$$

$$\text{since } \beta_i = (1 - \lambda_i)\gamma_i$$

Hence at the end of the m steps we take

$$\underline{c}^1 = \underset{\sim}{\underline{c}}^m = \Sigma \gamma_i \left[ 1 - \frac{(\lambda_i + \lambda_i^2 + \ldots + \lambda_i^m)}{m} \right] \underline{u}_i \tag{27}$$

instead of

$$\Sigma \gamma_i (1 - \lambda_i)\underline{u}_i$$

ie $\lambda_i$ is replaced by

$$\frac{\lambda_i + \lambda_i^2 + \ldots + \lambda_i^m}{m} \tag{28}$$

The usual Crank-Nicholson method then starts with the new value of $\underline{c}_s^1 = \underset{\sim}{\underline{c}}^m$. This gives

$$\underset{\sim}{\underline{c}}_s^n = \Sigma \left[ \gamma_i + (\theta_i - \gamma_i)\lambda_i^{n-1} \right]\underline{u}_i$$

$$\text{where } \theta_i = \gamma_i \left[ 1 - \frac{\lambda_i(1 - \lambda_i^m)}{m(1 - \lambda_i)} \right] \tag{29}$$

If m is even and $\lambda_i \approx -1$ then, from equation (28), the corresponding component vector will be almost eliminated. With the linear, quadratic and cubic elements and $\Delta t = 2.0$ the highest frequencies correspond to values of $\lambda_i$ approximately $-0.83$, $-0.95$, $-0.98$ respectively. They are of course nearer to $-1$ with higher values of $\Delta t$.

Results are presented with three ways of using the m-step start:

(a) as above, finishing at $t = \Delta t$;
(b) starting at $t = \Delta t$. This gives better accuracy.

With this method we take

$$\underline{c}^m = \underline{\tilde{c}}^m$$

and proceed with Crank-Nicholson from there

ie $\lambda_i^m$ is replaced by $\dfrac{\lambda_i + \lambda_i^2 + \ldots + \lambda_i^m}{m}$ ;

(c) starting at $\dfrac{m-1}{2}$ steps back

ie $\underline{c}^{\frac{m+1}{2}} = \underline{\tilde{c}}^m$ .

This is a kind of average of (a) and (b) and gives the best accuracy of the step methods.

Figure 9 illustrates (a) and (b) for m = 2, $\Delta t$ = 2.0.
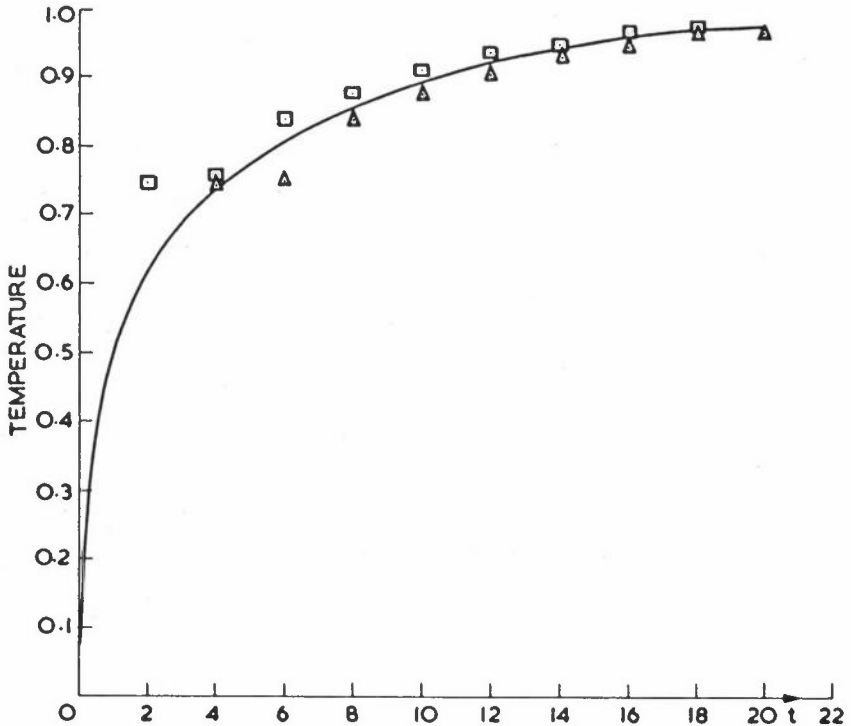


Figure 9: Quadratic element    2 step start
(a) ▫    (b) △    Exact solution –

(3)    Richtmyer and Morton (1967) recommend a scheme (p 190) of the same order of accuracy as Crank-Nicholson and designed to damp out high frequency

oscillation. The equivalent with the finite element discretisation is

$$\frac{3}{2\Delta t} N(\underline{c}^{n+1} - \underline{c}^n) - \frac{1}{2\Delta t} N(\underline{c}^n - \underline{c}^{n-1}) = -K\underline{c}^{n+1} + K\underline{c}^{\infty} \tag{30}$$

Results have been obtained from this, using an averaged Crank-Nicholson value for the $\underline{c}^1$ necessary at the start. Also, for comparison, some results have been obtained using a zero-zero start.

The amplification factor $\xi_i$ corresponding to the eigenvalue $\mu_i$ is given by

$$(3 + 2\Delta t\mu_i)\xi_i^2 - 4\xi_i + 1 = 0 \tag{31}$$

Put $\mu_i \Delta t = \alpha$. Then if $\alpha \leq \frac{1}{2}$, one root approximates $e^{-\alpha}$, the other is the spurious root

$$\frac{2 - \sqrt{1 - 2\alpha}}{3 + 2\alpha} < \frac{2}{3}$$

When $\alpha > \frac{1}{2}$ the roots of (31) are complex and of modulus $\rho = \left[\dfrac{1}{\sqrt{3 + 2\alpha}}\right]$

which $\to 0$ for large $\alpha$.

## COMPARISON WITH THE FINITE DIFFERENCE METHOD

For comparison results are also obtained using the standard finite difference method for discretising in space. This corresponds to using a linear element with a lumped mass matrix. If all the terms of the mass matrix are positive (as they are with the linear element) then lumping by making the diagonal terms equal to the row sums cannot decrease the eigenvalues. The spectral radius, in particular, must be increased (because, with an irreducible matrix it could not previously have been a boundary point of the union of the Gerschgorin circles).

If we assume the matrices N, K have a common set of eigenvectors (as is true with the simple problem currently being studied) with corresponding eigenvalues $\mu_N$, $\mu_K$ respectively, it is evident that

$$\Delta t_{crit} = 2 \min \frac{\mu_N}{\mu_K} \tag{32}$$

is increased by lumping.

The matrix A now has the general form

$$\widetilde{A} = \begin{bmatrix} 2 & -1 & & & \bigcirc \\ -1 & & & & \\ & & \ddots & & \\ & & & -1 & -1 \\ \bigcirc & & & -2 & 2 \end{bmatrix}$$

with n mesh points (using a fictitious node to bring in the right hand boundary condition).

The eigenvalues of $\widetilde{A}$ are $4 \sin^2\left[\dfrac{(2p - 1)\pi}{4n}\right]$ , $p = 1,2, \ldots, n$.

Hence the spectral radius $\rho(\widetilde{A}) = 4 \sin^2\left[\dfrac{(2n - 1)\pi}{4n}\right] < 4$ and

$$\Delta t_{crit} > 0.5 \tag{33}$$

Finite difference results are included in the Tables for comparison for $\Delta t = 0.25$, 0.5 and 2.0. Only the $\Delta t = 2.0$ results were averaged as the others are below $\Delta t_{crit}$. The errors for $\Delta t = 2.0$ are alternately positive and negative as expected.

## SUMMARY OF RESULTS

Tables I, II and III present the comparison of results from the above methods based on average percentage relative error. They are all compared on the solutions given up to a time of about 10 units except one result in Table I which shows how the average improves with more time steps. It is noticeable also how in some cases the average is improved when the first result is omitted.

The conclusion seems to be that the method of averaging with Crank-Nicholson and omitting the first result is a simple and reasonably effective way of dealing with the noise. The results with this are better than with more elaborate methods. The quadratic and cubic elements give a better performance than the linear element especially with the smaller time steps. The finite difference method shows up extremely well.

AVERAGE PERCENTAGE RELATIVE ERROR TABLES

**Table I:** $x = 1.0$, $\Delta t = 2.0$
(values in brackets omit the first result)

| | Linear | Quadratic | Cubic | No. steps |
|---|---|---|---|---|
| Backward difference | 2.88 | 3.79 | 3.71 | 12 |
| Crank-Nicholson | 13.08 | 9.08 | 6.72 | 5 |
| Crank-Nicholson | 6.02 (3.23) | 8.29 (7.32) | 4.27 (2.70) | 12 (11) |
| Averaged C-N | 2.00 (1.17) | 2.85 (1.13) | 2.62 (1.05) | 13 (12) |
| 2-step start (a) | 3.77 | 3.49 | 3.48 | 10 |
| (b) | 1.30 | 1.58 | 1.49 | 10 |
| (c) | 1.38 | 1.33 | 1.22 | 10 |
| 3-step start (a) | 4.90 | 5.12 | 4.47 | 10 |
| (b) | 2.94 | 4.14 | 3.39 | 10 |
| (c) | 1.59 | 2.26 | 1.17 | 10 |
| Finite difference and Crank-Nicholson | 2.7 (1.0) | | | 13 (12) |
| Averaged F.D. C-N | 2.26 (0.77) | | | 13 (12) |
| Richtmyer–Morton averaged C-N start | 2.12 | 2.56 | 2.53 | 12 |

**Table II**: x = 1.0, $\Delta$t = 0.5
(values in brackets omit the first result)

|  | Linear | Quadratic | Cubic | No. steps |
|---|---|---|---|---|
| Backward difference | 0.71 | 3.45 | 3.11 | 20 |
| Crank-Nicholson | 3.14 | 4.10 | 3.38 | 20 |
| Averaged C–N | 3.59 (1.84) | 1.29 (0.36) | 0.71 (0.41) | 20 (19) |
| Finite difference with C–N | 0.54 (0.14) |  |  | 20 (19) |

**Table III**: x = 1.0, $\Delta$t = 0.25

|  | Linear | Quadratic | Cubic | No. steps |
|---|---|---|---|---|
| Backward difference | 2.98 | 3.76 | 2.78 | 20 |
| Crank-Nicholson | 3.89 | 1.92 | 1.56 | 40 |
| Averaged C–N | 7.32 (2.49) | 0.35 (0.31) | 1.87 (0.36) | 41 (40) |
| Finite difference with C–N | 0.90 (0.19) |  |  | 40 (39) |
| Richtmyer–Morton averaged C–N start | 0.28 | 2.20 |  | 39 |

## REFERENCES

RICHTMYER R D  and  MORTON K W,  Difference Methods for Initial Value
    Problems, Interscience (1967)
ZIENKIEWICZ O C,  The Finite Element Method, McGraw Hill (1971)

# A TECHNIQUE FOR DEALING WITH SLOWLY DECAYING NOISE

B M Irons


Department of Engineering
Faculty of Engineering
The University of Calgary
Calgary
Alberta
Canada

## INTRODUCTION

Slowly decaying noise is a hazard all too familiar to those researchers who must integrate large problems in time. It is not enough that a numerical scheme be "stable" in the strict mathematical sense: if a problem is nearly unstable, and nonlinear, roundoff alone can eventually drive it mad. Thus the problem we are discussing is neither trivial nor academic. Indeed, according to the researches of J Tinsley Oden, the behaviour of the noise means success or failure in, for example, a nonlinear problem that encounters shock waves.


Desperate problems engender desperate measures. It is our intention here to initiate research into algorithms which modify or "doctor" the given forcing function F, very slightly, to avoid at source those features which generate the noise. The changes are not always such as to make the function "smoother" in any ordinary sense. The sub-routine that makes the adjustments is appropriately called DOCTOR.


Before proceeding further, we must define the sort of noise we are concerned with. We see it frequently, as a perturbation (on an otherwise fairly accurate solution) which almost exactly changes sign in successive timesteps. In particular, the Crank-Nicholson scheme generates zigzag noise. It was this observation that led certain early workers, notably Chandra Parekh and Ken Fullard, to invent a simple technique – now widely used – which virtually eliminates the most persistent noise of the type we are considering. They took the response vectors q in successive timesteps, and averaged them. With this vector, $q_{i+\frac{1}{2}} = (q_i + q_{i+1})$ they then re-started at time $t_{i+\frac{1}{2}}$.


We seek here to combine the two techniques. Accordingly, we shall describe another subroutine HUSH, which decides when to re-start, and when the doctoring technique can cope alone.

## GENERAL OBSERVATIONS

It was a remark of Argyris that prompted the present work. He observed that a step in the forcing function introduced as in Figure 1(a) generated considerable noise; whereas Figure 1(b) did not. (He was not using Crank-Nicholson.) According to our view, this was because he had two steps instead of one. The first alone produces a zigzagging after-effect, but the second produces the opposite after-effect, which cancels it.

Figures 1(c) to (i) were pure conjecture, but at the time it was a useful exercise to discover how a step reinforces or cancels a later step, depending (a) on whether it is of opposite sign, and (b) whether it occurs an odd or an even number of timesteps later: this we shall see is crucial. In the general case, what determines the residual noise is (sum of jumps at odd timesteps) minus (sum of jumps at even timesteps).
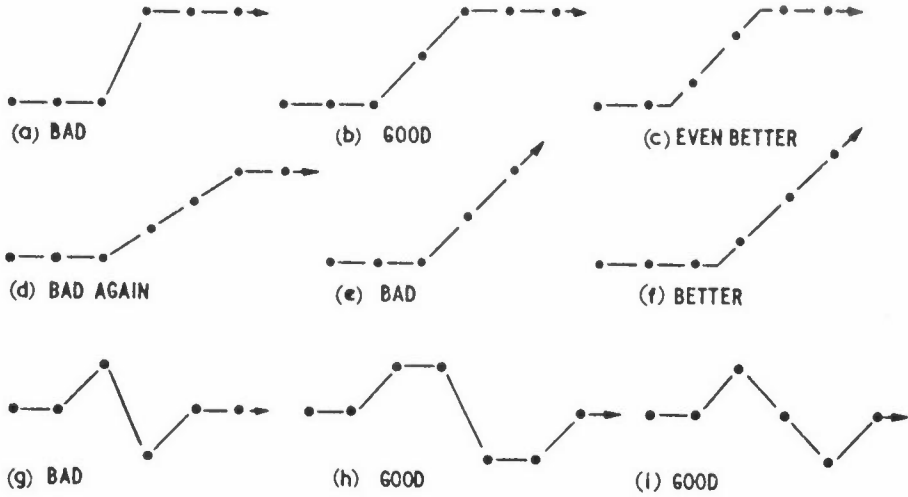


Figure 1: Showing good and bad representations of forcing functions

We can express this idea in finite difference form, writing the noise at step i caused by the doctored forcing function G:

$$(\text{noise})_i = \Delta G_{i-1} - \Delta G_{i-2} + \Delta G_{i-3} \cdots$$

$$= (G_i - G_{i-1}) - (G_{i-1} - G_{i-2}) + (G_{i-2} - G_{i-3}) \cdots$$

$$= \begin{bmatrix} G_1 & G_2 & G_3 & G_4 & \cdots \end{bmatrix} \{\ldots\ 2,\ -2,\ 2,\ -2,\ 1\}^T$$

If the decay ratio R is not exactly $-1$, as assumed here,

$$(\text{noise})_i = \Delta G_{i-1} + R\Delta G_{i-2} + R^2\Delta G_{i-3}\cdots$$

$$= \begin{bmatrix} G_1 & G_2 & G_3 & \cdots & G_i \end{bmatrix} \{ \cdots R^4 - R^3, \ R^3 - R^2, \ R^2 - R, \ R - 1, \ 1 \}$$

$$= G^T s_i \ (\text{say})$$

This is a rather poor first approximation to the vector $s_i$. To get a second approximation we shall compute a transient response according to the Crank-Nicholson scheme. This part of the work should be generally useful.

## NOISE ANALYSIS IN THE CRANK-NICHOLSON PROCESS

We shall take as our trial equation $q = \alpha(G - q)$, so that q tends to follow G after some lapse of time. The recurrence formula according to Crank-Nicholson is, with unit steplength:

$$q_{i+1}^c - q_i^c = \tfrac{1}{2}\alpha\{G_i + G_{i+1} - q_i^c - q_{i+1}^c\}$$

We assume

$$q_o = G_o = 0, \text{ and } G_1 = G_2 \ldots = 1, \text{ giving}$$

$$q_i^c = \frac{\alpha}{1 + \tfrac{1}{2}\alpha} \ (1 + R + R^2 + \ldots + R^{i-1} + \tfrac{1}{2}R^i)$$

where

$$R = (1 - \tfrac{1}{2}\alpha)/(1 + \tfrac{1}{2}\alpha) = \text{decay ratio}$$

$q_i^c$ converges to 1, but in a zigzag fashion if $\alpha$ is large. We shall compare this with the "correct" answer, and we shall regard the error as noise. Taking the view that G represents a step function from 0 to 1, at $t = \tfrac{1}{2}$,

$$q_i^s = 1 - e^{-\alpha(t-\tfrac{1}{2})} \text{ for } t \geqslant 1$$

Alternatively, we may take the more generous view that the data implies a continuous function, a ramp that goes linearly from $G_o = 0$ to $G_1 = 1$, and that takes a constant value, zero or one, elsewhere. The exact solution becomes

$$q_i^R = t - \alpha^{-1}(1-e^{-\alpha t}) \text{ for } 0 < t \leq 1$$

$$q_i^R = 1 - \alpha^{-1}(e^{\alpha}-1)e^{-\alpha t} \text{ for } t \geq 1$$

We now tabulate the consequences of these two views.

| R | $\alpha$ | Method | Values at steps | | | | |
|---|---|---|---|---|---|---|---|
| -.95 | 78 | Step | 0., | 1.0000, | 1,0000, | ... | |
| | | Ramp | 0., | .9872, | 1.0000, | ... | |
| | | Crank-Nicholson | 0., | .975, | 1.0237, | .9774, | 1.0214 |
| -.9 | 38 | Step | 0., | 1.000, | 1.0000, | ... | |
| | | Ramp | 0., | .9737, | 1.0000, | ... | |
| | | Crank-Nicholson | 0., | .95, | 1.045, | .9595, | 1.0364 |
| -.8 | 18 | Step | 0., | .9999, | 1,0000, | 1.0000, | ... |
| | | Ramp | 0., | .9444, | 1.0000, | 1.0000, | ... |
| | | Crank-Nicholson | 0., | .9, | 1.08, | .936, | 1.0512 |
| -.5 | 6 | Step | 0., | .9502, | .9999, | 1.0000, | ... |
| | | Ramp | 0., | .8337, | .9996, | 1.0000, | ... |
| | | Crank-Nicholson | 0., | .75, | 1.125, | .9375, | 1.0312 |
| 0 | 2 | Step | 0., | .6321, | .9502, | .9932, | .9991 |
| | | Ramp | 0., | .5677, | .9415, | .9921, | .9989 |
| | | Crank-Nicholson | 0., | .5, | 1.0, | 1.0, | 1.0 .. |

## A REFINEMENT TO DEAL WITH RAMP FUNCTIONS

The table shows two analytic comparisons with Crank-Nicholson. Both are interesting, but practical considerations predispose us irresistibly towards the second interpretation. We frequently meet examples in which G climbs, fairly steadily, over the whole time range. We should expect the initial noise to decay to zero, yet according to equation (1) it will not. We remark that a ramp function $G = kt$ over 200 timesteps is identical

to the sum of 200 short ramps, each extending over a single timestep: the implication is that if we synthesize the response from ramps we shall cover this important case.

To this end we shall modify $s_i$, following the clue that the error in the first time-step in the table is only half what we should expect, and at the end of the unit ramp we find only half of the zigzag value, on the evidence from later timesteps:

(a)    We seek a formula which gives evanescent noise in an extended ramp.

(b)    The noise must decay after a single step also.

A formula which satisfies these requirements, and which is easily computed, is

$$(\text{noise})_i = \left[ \ldots \ R^4 - R^3, \ R^3 - R^2, \ R^2 - R, \ R - q, \ q \ \right] \{G\}$$

where

$$q = - \frac{R}{1-R} \approx \frac{1}{2} \tag{1}$$

## A PRACTICAL OVERALL MEASURE OF NOISE

It is now an easy matter to state our intentions in general terms. We shall aim to minimise the sum of $(\text{noise})_i^2$ over all the timesteps:

$$(\text{noise})^2 = G^T s_1 s_1^T G + G^T s_2 s_2^T G + \ldots$$

$$= G^T S G \text{ say} \tag{2}$$

(We note in passing that S is positive definite, but not very well-behaved.) This minimization is subject to the doctoring terms d being reasonably small, again in a least-squares sense:

$$d_1^2 + d_2^2 + d_3^2 + \ldots \quad = d^T d = (G-F)^T (G-F) \tag{3}$$

where F = undoctored forcing function.

Now that the problem is defined, we hope it will be possible to do even better than the "good" examples of Figure 1 and, furthermore, to deal with general input data.

Computational technique: Subroutine DOCTOR will compute SP many times. We are fortunate that $[S]$ - a large matrix - need never be stored. Instead, it computes SP thus:

$$SP = (s_1^T P)s_1 + (s_2^T P)s_2 + \ldots \tag{4}$$

We use a single vector for the $s_i$ and update it as required. This is more difficult, but it takes a little less computing time, as coded in DOCTOR, than the matrix version.

Note: Any approach based on S admits that the direction of marching is significant. Noise is the residue of history, and there is a difference between past and future.

## STRATEGIES FOR OPTIMISING THE NOISE REDUCTION

Let us now examine critically three strategies which we could adopt to minimise the noise:

(1) We could fix the amount of doctoring, $d^T d$, and seek the greatest noise reduction.

(2) Perhaps we may find the amount of doctoring more than adequate; we could aim instead for some acceptable noise level with minimum $d^T d$.

(3) The last strategy, actually used in DOCTOR, emerges from the algebra of versions 1 and 2 above. Suppose we minimise $G^T SG$, subject to the constraint $d^T d = \delta = $ constant, imposed by the Lagrange multiplier $\lambda$:

$$G^T SG + \lambda(d^T d - \delta) = \text{stationary} \tag{5}$$

or, substituting d = G-F and differentiating:

$$SG = -\lambda (G-F) = -\lambda d \tag{6}$$

$$d^T d = \delta \tag{7}$$

But here strategy (3) is different: we pre-assign to $\lambda$ some value which gives reasonable smoothing, and we let $\delta$ take what value it will. Thus, fixing $\rho = \lambda^{-1}$, we merely solve the equations for G:

$$\left[ I + \rho S \right] G = F \tag{8}$$

This philosophy gives a simple robust program. Let us enquire how it affects the user.

## CONSEQUENCES OF THE STRATEGY CHOSEN

(a)    We can demonstrate that provided $\lambda > 0$ in (6), then $G^T SG$ is a global minimum for $d^T d$ = constant.

For, perturbating d to d + $\varepsilon$,

$$d^T d + 2 \, d^T \varepsilon + \varepsilon^T \varepsilon = d^T d$$

$$\therefore \quad d^T \varepsilon = - \tfrac{1}{2} \, \varepsilon^T \varepsilon \tag{9}$$

Hence $G^T SG$ becomes $(G + \varepsilon)^T S (G + \varepsilon)$

$$= G^T SG + 2 \, \varepsilon^T SG + \varepsilon^T S \varepsilon$$

But from (6) we have $SG = - \lambda d$, giving

$$(G + \varepsilon)^T S (G + \varepsilon) = G^T SG - 2\lambda d^T \varepsilon + \varepsilon^T S \varepsilon$$

$$= G^T SG + \lambda \varepsilon^T \varepsilon + \varepsilon^T S \varepsilon$$

$$> G^T SG \tag{10}$$

for any nonzero $\varepsilon$ because S is positive definite. Thus whatever positive value we give to $\lambda$, we have a global optimum, in this special sense.

Conversely we can show, by a similar argument, that $d^T d$ is a global minimum for $G^T SG$ = constant.

(b)    This result will not convince a sceptic of the wisdom of choosing $\rho$ rather than a level of doctoring or of noise reduction. It is therefore reassuring to discover that $\rho$ is related monotonically and continuously to both of these. Let us introduce the eigenvectors and eigenvalues of S:

$$Sv_i = \lambda_i v_i \qquad\qquad \text{where } v_i^T v_i = 1$$

and let us write

$$F = \Sigma c_i v_i$$

Then

$$G^T SG = \Sigma \lambda_i (\frac{c_i}{1+\rho\lambda_i})^2 \qquad\qquad (11)$$

and

$$d^T d = \Sigma (\frac{c_i \rho\lambda_i}{1+\rho\lambda_i})^2 \qquad\qquad (12)$$

Clearly with increasing $\rho$, the noise $G^T SG$ decreases continuously from $F^T SF$ to zero. Meanwhile, the doctoring $d^T d$ increases continuously from zero to $F^T F$. There is a (1,1,1) correspondence relating $\rho$, $d^T d$ and $G^T SG$. All the relations are monotonic and continuous within the allowed ranges.

(c)  Writing:

$$F = \Sigma c_i v_i$$

$$G = \Sigma (\frac{c_i}{1+\rho\lambda_i}) v_i$$

We observe that if $\lambda_i$ is very small, $v_i$ carries over from F into G almost unchanged in amplitude. On the other hand, if $\lambda_i$ is large, ie $v_i$ is a zigzag function, it is greatly reduced in amplitude. The amount of doctoring $d_i^T d_i$ will evidently depend on the data submitted. If there is a consistent zigzag error pattern, it will be virtually obliterated, at the cost of a good deal of doctoring. If there is a single sharp jump, we may expect compensatory zigzags before and after. But if the noise is random, ie there is no consistent zigzag pattern, we should aim for moderate doctoring, and moderate noise reduction. This will dictate our choice of $\rho$.

Note: Hand calculations lead us to expect $[S]$ to have a conditioning number $\lambda_{max}/\lambda_{min} \approx N^5$, for R = -1. For the sum of the roots is $(2N^2 - 3\frac{1}{2}N + 1\frac{3}{4})$ - the largest root being of the same order - and the product of the roots is only $2^{-N}$. Evidently $[S]$ could be a brutal tool with N in the hundreds. But we observe that $[I + \rho S]$ has a conditioning number of order $2\rho N^2$ which we can handle confidently in solving the equations.

Since it is easy and economical to multiply by $[S]$, we use conjugate gradients to compute $[I + \rho S]^{-1} F$. It is a bad technique for solving finite element equations. But we believe this is because the initial direction is too wildly different from the final answer. In the present case the zigzag responses emerge very quickly and it is a good algorithm.

## AN ANALYSIS OF THE AVERAGING PROCEDURE

Before discussing HUSH, we must define exactly what happens to the noise amplitude when we average; we seek some formula:

$$(\text{noise})_{i+\frac{1}{2}} = A(\text{noise})_i + B(\text{noise})_{i+1}$$

To this end, we supplement the table, and this helps us to deduce B. From

$$q_{\frac{1}{2}}^{c} - q_{\frac{1}{2}}^{r} = \left[\tfrac{1}{4}\alpha/(1 + \tfrac{1}{2}\alpha)\right] - \left[\tfrac{1}{2} - \alpha^{-1}(1 - e^{-\frac{1}{2}\alpha})\right]$$

$$\approx \frac{1}{\alpha(\tfrac{1}{2}\alpha + 1)}$$

$$q_{1}^{c} - q_{1}^{r} = \left[\tfrac{1}{2}\alpha/(1 + \tfrac{1}{2}\alpha)\right] - \left[1 - \alpha^{-1}(1 - e^{-\alpha})\right]$$

$$\approx \frac{1 - \tfrac{1}{2}\alpha}{\alpha(\tfrac{1}{2}\alpha + 1)}$$

Because $q_{o}^{c} - q_{o}^{r} = 0$, this gives B. To get A we invoke the case of a decay during the dwell after a single step:

$$(\text{noise})_i = 1, \quad (\text{noise})_{i+1} = R, \quad (\text{noise})_{i+\frac{1}{2}} = \tfrac{1}{2}(1 + R)$$

The formula for the residue of noise after averaging thus has A = O:

$$(\text{noise})_{i+\frac{1}{2}} = (\text{zero}) \bullet (\text{noise})_i - \frac{1}{\tfrac{1}{2}\alpha - 1}(\text{noise})_{i+1}$$

$$= \frac{1 + R}{2R}(\text{noise})_{i+1} \tag{13}$$

on substituting $\tfrac{1}{2}\alpha = \frac{1 - R}{1 + R}$.

| R | −.95 | −.9 | −.8 | −.5 | 0 |
|---|---|---|---|---|---|
| $\alpha$ | 78 | 38 | 18 | 6 | 2 |
| Value at first halfstep from averaging Crank-Nicholson | .4875 | .475 | .45 | .375 | .25 |
| Exact value for ramp | .4872 | .4737 | .4445 | .3416 | .1839 |

We now enquire how to introduce this initial noise into the next segment to be smoothed. Let us write the noise vector thus:

$$\{\text{noise amplitude}\} = \begin{bmatrix} q & & & \\ R - q & q & & \\ R^2 - R & R - q & q & \\ R^3 - R^2 & R^2 - R & R - q & q \end{bmatrix} \begin{Bmatrix} G_1 \\ G_2 \\ G_3 \\ G_4 \end{Bmatrix}$$

Because this implies that $G_o = 0$, if we put $G_1 = G_2 = G_3 = G_4 = 1$, the noise amplitudes are q, R, $R^2$. $R^3$. Apart from the first, it is exactly as if unit noise amplitude had existed at time $t_1$. Therefore, we contrive to add $(\text{noise})_{\frac{1}{2}}$ to all the values of F to be smoothed, and, finally, we subtract it from the G values.

## THE MIXED STRATEGY — SUBROUTINE "HUSH"

In a large marching problem, DOCTOR will account for only a small part of the computing cost. Therefore we advocate the mixed approach of subroutine HUSH which, before the marching starts, both doctors the forcing functions and decides at which timesteps we should apply the averaging technique to eliminate accumulated noise. The organisation of HUSH is roughly as follows:

(1)    We doctor F for steps 1 to $N_Z$, the full range.

(2)    We compute the noise amplitudes, $s_i^T G$, and we observe that the noise first becomes intolerable at step $N_{STOP}$.

(3)    We doctor F again for steps $N_A = 1$ to $N_Z = N_{STOP}$.

(4)    We record in G our intention to suppress the noise at $N_B$ and to re-start at timestep $N_{STOP} - \frac{1}{2}$.

(5)    We doctor F again for steps $N_A = N_{STOP} - \frac{1}{2}$ to $N_Z$.


Therefore the first task in subroutine HUSH is to find the interpolated values of F at the half timesteps.   We use cubic interpolation, except in the end intervals, where we use linear interpolation.   (These are usually unimportant.)


In order to satisfy the requirements of HUSH, subroutine DOCTOR must do slightly more. It is not enough to assume a zero starting value: we must include the effect of the value at which we averaged, and also the residual noise.


A vexed question now arises – if the noise becomes intolerable at $N_{STOP}$, should we smooth up to and including $N_{STOP}$, or should we go back one timestep?   In an extreme case where intense noise suddenly strikes, we shall probably have to inch forward over the crisis, in a series of averaged half-steps.   Our view is that the undoctored values should be used here.   Accordingly, HUSH regards a noise amplitude of DINLIM $\star \sqrt{FF}$ as merely objectionable, and smooths as far as $N_{STOP}$.   However, if the noise amplitude reaches the higher level of DINBAD $\star \sqrt{FF}$ it smooths only as far as $N_{STOP} - 1$.


## PRELIMINARY EXPERIMENTS

The random forcing function F was generated automatically.   A vector $V_R$ of numbers randomly distributed between $-\frac{1}{2}$ and $\frac{1}{2}$ was created.   Then F was smoothed from these:

$$F_i = \sum_{-\infty}^{\infty} V_{i+j} e^{-\lambda_j^2}$$

The problem chosen was a spring of uniform section, with uniformly distributed viscous drag. There were 40 elements, of lengths $1.05^{i-1}$. The longest element was earthed, and the force was applied at the opposite end. This problem should give a broad band of eigenvalues, yet it is not too expensive to compute.

It was run in three forms:

    (a)    with the original F.

    (b)    with the doctored F.

    (c)    with F interpolated to quarter-intervals.

The results of these experiments, so far, is somewhat inconclusive. The noise can certainly be reduced, but the doctoring inevitably alters the values produced. As far as accuracy is concerned, as measured by comparing (a) - (c) with (b) - (c), there is no improvement, the doctoring introduces errors in the response of the same order as the noise removed!

## CONCLUSIONS

In seeking a technique for doctoring the input forcing function, so as to reduce the noise, a simple and adequate representation of the noise generation in the Crank-Nicholson process has been evolved. In particular, it was found necessary to consider a ramp rather than a step function. Other workers are recommended to do likewise.

However, the results so far from the doctored forcing function are disappointing. Perhaps there will be more profitable applications. Or perhaps variants will give a better performance.

In any case equations (1) and (13), which give the noise with and without averaging, should be useful as a tool of analysis even if further attempts at synthesis should fail.

## ADDENDUM
## A FOURTH STRATEGY FOR OPTIMISING THE NOISE REDUCTION

In addition to the techniques outlined in the section entitled "Strategies for Optimising the Noise Reduction" there is another radically different approach which leads to a much simpler algorithm. ⌐he philosophy is as follows:

(1)    We decide on a value of $\alpha\Delta t$, and hence of $R = (1-\tfrac{1}{2}\alpha)/(1+\tfrac{1}{2}\alpha)$ at which we want the greatest suppression of noise. Presumably the $\alpha$ chosen would be near the highest eigenvalue, as estimated from the smallest element.

(2)    We consider, as the exact solution, the analytic solution based on the values of the original forcing function F, connected by straight lines.

(3)    We define G as the values which would give _exactly_ these analytic values, via the Crank-Nicholson process, assuming the chosen decay ratio R.

This is potentially an exact method, but for simplicity we shall ignore the exponential terms as in deriving equation (1). The following formulae are very accurate. Indeed with $\alpha\Delta t = 18$, $R = -0.8$, they are accurate to four decimals.

$q_i$ = Crank-Nicholson value based on G

$$= \frac{1-R}{2}(G_i + G_{i-1}) + R q_{i-1}$$

$q_i^*$ = analytic value based on unit ramps

$$= \frac{1-3R}{2(1-R)} F_i + \frac{1+R}{2(1-R)} F_{i-1}$$

Putting $q_i^* = q_i = q_{i-1}^* = q_{i-1}$ we find

$$\frac{1-R}{2}(G_i + G_{i-1}) + \frac{R(1-3R)}{2(1-R)} F_{i-1} + \frac{R(1+R)}{2(1-R)} F_{i-2}$$

$$= \frac{1-3R}{2(1-R)} R_i + \frac{1+R}{2(1-R)} F_{i-1}$$

$$G_i = -G_{i-1} + \frac{1-3R}{(1-R)^2} F_i + \frac{1+3R^2}{(1-R)^2} F_{i-1} - \frac{R+R^2}{(1-R)^2} F_{i-2}$$

To apply this very simple recurrence relation we must assume that $F_{-1} = F_0 = G_0$. For example, with R = -0.8 we have the following step function:

F = 0      0      1         1         1         1

G = 0      0      1.0493    0.9012    1.0988    0.9012    etc.

This is a remarkable result, in that G is now zigzagging with constant amplitude. Yet before the jump there is no zigzag.

We have not yet tried this promising variant.  If it is successful, it will be necessary to re-formulate the averaging process, so as to detect the noise buildup at another, presumably lower, value of $\alpha$, and then to discover how to re-start the process.

# THE PATCH TEST

B M Irons

Department of Engineering
Faculty of Engineering
The University of Calgary
Calgary
Alberta
Canada

## SUMMARY

The patch test is presented in a form accessible to engineers, emphasizing its usefulness as a debugging tool, and as a practical guide in element research.   The applications are straightforward.   A final section enquires into the implications of the patch test.   In particular we show that it is not possible with a given nodal configuration to beat the patch performance of an element which conforms according to the classical criteria.

## INTENTIONS

Irons and Razzaque (1972) gave a historical introduction to the patch test for convergence, emphasizing its usefulness to engineers.   It seems to have failed.   Admittedly people now talk knowledgeably about the test:   but few workers would automatically seek its advice whenever a new set of interpolating functions or a new technique for integration came into use.   In my opinion it should be regarded as unprofessional to publish any such development without giving the patch test performance.   I hope in this paper to persuade others, beyond the handful of students I have supervised, to use it routinely as a debugging aid and research guide, and to record the results for posterity.

What then is the patch test?   Often it has been merely a computer run, using a perfectly standard program, on a compact assemblage of perhaps six elements.   We give the external nodes of this patch the displacements, slopes etc corresponding to some state of constant stress.   For example, if the plate bending elements of Figure 1 are to be subjected to a quadratic patch test, we put say $w = x^2 + 3xy$, giving $\partial w/\partial x = 2x + 3y$ and $\partial w/\partial y = 3x$ and hence also the prescribed slopes at certain boundary nodes.   The nodes inside the patch are unloaded.   The final answers from the computer

will include twenty or more stresses, and if we are lucky they are all equal to the values we intended:  if not, then we must discover what has gone wrong.
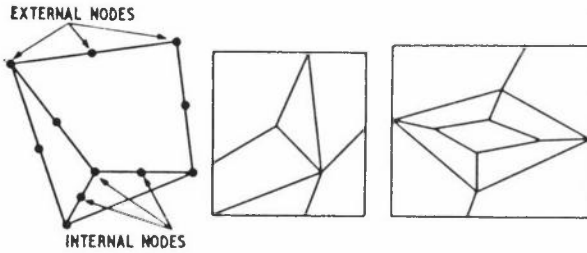


Figure 1:  These patches in 2 dimensions have actually been used.

So anybody can do a classical patch test.   Moreover, it performs several important roles in engineering:

(1)    With certain provisos, it establishes fine-mesh convergence.
       <u>Note</u>:    Engineers are seldom if ever interested in high accuracy.
       In assessing a new element we should let the patch test take care of the fine-mesh performance, and concentrate our attentions on those meshes which give the permitted errors (about ± 5% in most applications), and also on even coarser meshes to see how sensitive the element is in this critical region. A good element is foolproof in every sense.

(2)    It provides one more criterion for accepting or discarding a new element.
       <u>Note</u>:    Of course, there should be no question of rejecting an element <u>because</u> it is non-conforming, <u>provided</u> it passes the patch test and therefore converges. But there are elements which have been heavily sold, eg our nonconforming triangle (Bazeley et al 1965), and those Ahmad layered shells and plates (Ahmad 1969) which fail the test for quadrilateral geometry. They should perhaps be discarded, as a matter of principle;   we now have many elements to choose from, for virtually every job.

       (In practice the patch test has usually acted as nurse rather than as executioner.   The one simple, permissive criterion has encouraged me more often to modify a formulation than to abandon it.   And, being simple, it does not distract one's attention away from all the other important criteria - ease of implementation, ruggedness etc.)

(3)    To some extent, it helps us to debug our programs.
       <u>Note</u>: Debugging tends to be a progressive affair for most of us. Initially we might work on a single element, until it will accept an arbitrary constant strain, given the appropriate nodal values.   The second stage might be to check that a patch of elements is in equilibrium - this test sometimes gives a more revealing diagnosis than actually solving the

equations. Again, let us remind ourselves that the patch test (like any other) can fail spuriously owing to an error in the data! This is why I suggest that the correct values and slopes should at some stage be prescribed at <u>all</u> the nodes. The printout should then give the correct stresses, even for most of the elements known to fail the patch test. Elements that pass, however, will give zero reactions at the internal nodes of the patch. If this approach gives correct stresses and <u>nonzero</u> reactions, the evidence is very decisive. For what can then be wrong? - only such details as faulty node numbering and nonuniform material properties, details which are relatively easy to check.

(Of course there is a missing ingredient. This technique overlooks the possibility of a <u>singular assembled matrix</u>. The third test, then, might be to fix just enough deflections to prevent rigid body motions, and to apply to the patch boundaries the forces already calculated in the second stage. If there is a danger that the element might give a singular assembled matrix, this procedure will encourage it to misbehave under the controlled research conditions of the patch test rather than later in a commercial job. But should we regard this as an appropriate task for the patch test? - a rhetorical question, maybe, but the subject of <u>mechanisms</u>, and especially of those that can propagate and infect the whole problem, is more difficult in practice than that of convergence, and is at least as deserving of <u>separate</u> research effort.)

(4)  Even more important, it helps us to debug our thoughts.
      <u>Note:</u>  I should hate to admit how often my predictions have been wrong, and therefore how much the computer has taught me. A patch test is an experiment, whose results often acquire a much greater significance if they contribute towards some general theory or working hypothesis. It pays to think <u>and</u> to do experiments. The frenzied thought which follows an unexpected result is especially productive and we learn far more when some fancied element <u>fails</u> the patch test.

Of course, to do any good in patch testing we must be decisive about what we do, clear about why we are doing it, and explicit in our reporting of the results. For example, in the isoparametric elements a lower integration rule will often suffice if the patch contains no curved edges. For me, a 3-element test in two dimensions as in Figure 1 sometimes carries sufficient conviction; I never remember going above 9 elements and that was when the patch boundary was square. The geometry must be arbitrary, with no pattern at all, and the stress must usually be arbitrary. For we shall conclude from this one test, or a second for the nervous operator, that <u>any</u> combination of elements will give patch equilibrium.

This conclusion carries less risk than might be supposed. For one geometry that passes the patch test fortuitously, we expect to find an infinity that do not. (We have the opposite problem in detecting a propagating mechanism.)

Evidently the test is useful, and easy to do. Perhaps it is not always very easy to think about, so we shall adopt here an "educational" rather than an "academic" tone. However I cannot resist the temptation to include a section at the end about an interesting new area revealed by the patch test.

## INTRODUCING THE PATCH TEST

In teaching any new concept, one seeks to strike a balance between, on the one hand, presenting some application which is so trivial that it hardly registers, and, on the other hand, presenting an application so obscure that a class will lose the principle in the details. For example, in teaching isoparametric elements the best choice seems to be Ian Taig's bilinear quadrilateral of 1957, not because it was the first, but - more important - because it avoids both triviality and obscurity. Good teaching is not <u>always</u> good history.

I have only recently recognised a uniquely entertaining introduction to the patch test. Without some variational crime, the test will seem trivial, and to a beginner non-conformity will seem the most audacious crime yet perpetrated. A continuous beam in bending is wilfully treated as discontinuous in the example chosen: as we shall show him, the crime goes unpunished. At least he will make a mental note of our defence counsel.

A shape function for this extraordinary beam is illustrated in Figure 2. The numbers 1, 2... represent the nodes, while A, B... are midway between these nodes. The slopes are not represented as nodal variables. Instead, we have $w_1$, $w_2$ ... as unknowns, the deflections at the nodes.
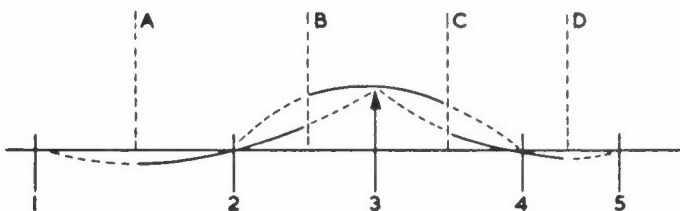


Figure 2: Showing the response to unit deflection w at node 3

The construction of Figure 2 is to draw parabolas through the values at 1, 2, 3, at 2, 3, 4, and at 3, 4, 5. But each parabola extends only over half the range: that through 2, 3 and 4 extends only from B to C. The deflections are therefore zero left of A and right of D. Furthermore, it follows that at C, for example, there is a discontinuity of deflection, but the slope is continuous. For the slope at C, the mid-point of segment 34, is the mean slope $(w_4 - w_3)/(x_4 - x_3)$ for both parabolas. We have the curious phenomenon of continuous first derivatives with discontinuous values - as if we had $C^{(1)}$ without $C^{(0)}$.

This completes the geometric arguments, and we now appeal to structural theory. We postulate, for the patch test, that the points $w_1$, $w_2$ ...$w_8$, $w_9$, $w_{10}$ all lie on some parabola, as in Figure 3. (This is not altogether magic. We could achieve such a
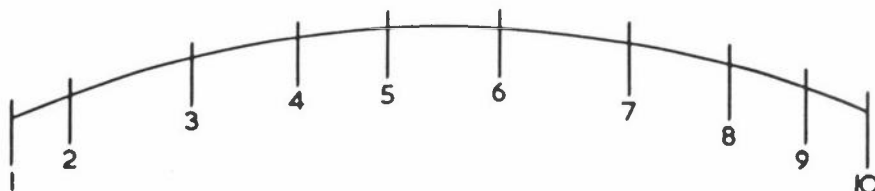


Figure 3: An assemblage of beam elements of Figure 2

state by fixing $w_1$, $w_2$, $w_9$ and $w_{10}$: crudely, this is like fixing the values and the slopes at the two ends.) If we now let nodes 3...8 go unloaded, so that they can choose their preferred $w_i$ on the basis of minimum strain energy, we have a recognizable patch test.

We now argue that nodes 3...8 will prefer their positions on the given parabola of Figure 3, rather than seek some perturbated positions, involving the discontinuities of Figure 2. In other words, in the context of this patch test we do not have conformity in the usual, obligatory sense, but we do have continuity because the model prefers not to misbehave. To reach this conclusion, we consider what happens if we give, say, node 3 a virtual deflection, thus perturbating the smooth parabola of Figure 3 by a function scaled from that of Figure 2. If this perturbation causes nonzero virtual work, then node 3 is not in equilibrium. But the virtual work is zero. For the bending moment M is constant in Figure 3. The work caused by a virtual deflection $\delta$ is

$$\int M(\frac{d^2\delta}{dx^2})dx = M \left[\frac{d\delta}{dx}\right]_{x_{1\frac{1}{2}}}^{x_{9\frac{1}{2}}} = 0$$

integrated over all the elements; this is zero because $d\delta/dx$ is continuous, and because both end pairs of nodes are fixed. (Of course, it is possible to integrate this in the

ordinary way, with considerable effort, but we do not recommend it: the more direct technique is too useful in the sequel.)

What is more, when we try to reproduce this on the computer, it works, just as predicted. It is safe practice always to check one's conclusions in this way. The error rate of computed patch tests has been considerably less than the error-rate of my own thought processes.

Note: Those who already know these particular finite difference assumptions may well be astonished: the appeal of finite elements to the imagination is strong, so that physical absurdities are emphasized.

## IMPLICATIONS OF THE PATCH TEST

The patch test is merely the translation into finite element terminology of the test for consistency in finite differences. We have been trying to solve the beam equation, $d^4w/dx^4 = q(x)$, whose variational formulation in terms of strain energy involves w and its derivatives up to $d^2w/dx^2$, the curvature. The case to watch is $d^2w/dx^2 = $ constant, giving uniform bending stresses.

According to the classical approach;

(1)    The element must be able to reproduce any chosen case, $w = a + bx + cx^2$, giving $d^2w/dx^2 = $ constant.

(2)    The shape functions must guarantee continuity of w and dw/dx between elements, regardless of the nodal values.

According to the patch test, we have convergence only if a group, or "patch", of elements do not misbehave when we attempt to reproduce $w = a + bx + cx$ by prescribing the nodal values at the boundaries of the patch.

One's first reaction might well be that the patch test is a less permissive criterion than the classical requirements, because we can choose arbitrarily the geometry of every element in the patch, to some extent. Thanks to the discontinuous beam, we can see already that it is more permissive. Indeed, I have had to use it at various times to justify the following crimes:

(1)    Elements without the continuity of the classical requirements.    (These have
       always been common, but not all have converged.)

(2)    Elements in which the energy is approximately integrated.    (This is the most
       important case.)

(3)    Elements whose shape functions are rational functions with singularities.
       (They might have a little historical interest:  they are relatively poor
       performers.)

(4)    Elements with no clear physical basis at the time of their discovery (Irons
       and Razzague 1972;  Irons and Razzaque 1962).    Exploration is often motivated
       by fake reasoning.

(5)    Bending elements in which the Kirchhoff assumption that normals remain
       normal is imposed discretely, eg, at certain points in the element
       (Stricklin et al 1969).    (These are usually good performers.)


A worker might, at this moment, be planning an even more wicked crime.    Again, the
patch test defence will be that if he refines his mesh sufficiently then the patches,
as well as the elements, will be small.    Assuming that his answers are good, the
variation of stress over any patch will be small.    Therefore he has applied a test
very similar to what will actually be required of the element, in some real job, but
with fine mesh.


Yes, very similar conditions - but not identical.    I am always frightened that a patch
of elements will be abnormally sensitive to some particular perturbation from constant
stress conditions.    Alas, this really happens, and it can be calamitous.    A pertur-
bation which carries very little strain energy is called a "near-mechanism": one which
carries no strain energy at all, a "mechanism".    According to the computation, it
resembles a rigid body motion, as if the structure were unsupported, or were supported
on a flexible spring.    The phenomenon has several ways of making a nuisance of itself:

(a)    The answers are large random numbers, of order $10^{N-2}$ too large if the
       computer gives N-decimal accuracy.

(b)    The answers are random numbers, a little larger than expected in general
       magnitude.    We see this if the combination of forces that excites the
       mechanism happens to be zero.

(c)     The pilot examples are successful, and perhaps so are the first few
        commercial jobs.    Then a large, expensive, and important commercial job
        fails with the symptoms of (b) above.    We have a near-mechanism.

These phenomena are examples of instability, in which a spurious perturbation or
parasitic response takes control.    The problem normally arises only if each element
has too many zero-energy modes, ie more than the rigid body modes.    It is possible,
if unlikely, that an assemblage of elements can find a way of combining their mechanisms
without contradiction.    Occasionally, we find a collective mechanism where the
individual elements have none.

In general, we can teach our students to estimate roughly the side-effects in a solu-
tion by considering a "matched solution" corresponding to some chosen exact solution:
here we substitute the exact nodal values into the finite element model.    The
matched solution gives a very good bound on the potential energy for a conforming
model.    But if instabilities are present, then a crazy set of nodal deflections can
give slightly lower potential energy, so the exercise is worthless.    Perhaps the
answer will be to avoid using elements with spurious mechanisms, together with those
that fail the patch test.    Yet this too might be unacceptable, because so many of
the best performers have mechanisms.

In practice, such instabilities are not always revealed by the patch test.    Again,
the patch test says nothing about the manner or order of convergence.    It does not
progress towards the limiting case:    it is the limiting case - exactly.

## DEFINITIONS RELEVANT TO THE PATCH TEST

It remains to define the patch test in a way that we can interpret in an unfamiliar
case, chosen for convenience only.    The elements cover the whole region without
overlapping.

An element is a region of integration.

A patch is a collection of elements containing internal nodes of every type
envisaged.    (We consider only those nodes which control the elements of the patch.)

An _internal_ node is a node completely surrounded by elements, a node whose _zone of_ _influence_ does not trespass outside the patch:  otherwise we have an _external node._

The _target-state_ for the patch test of lowest order is $\varepsilon$ = constant, where $\varepsilon$ contains the highest derivatives present in the functional.   In a higher patch test the target-state may be any simple analytic solution.

The _patch test_ may be a computer run which seeks to reproduce the target-state _exactly_ (ie to 6 or more decimals) by any means short of prescribing the values at the internal nodes.   (For example, in a higher patch test we may have to load the internal nodes, in exactly the same way as in a real problem.)   Or the patch test may be a computer run which prescribes _all_ the nodal values and merely calculates the residuals.   Or it may be a _theoretical argument_ which concludes that the residuals vanish at the internal nodes.

## APPLICATIONS OF THE PATCH TEST

The _Isoparametric Elements_ (Irons 1966) are economic only because _low-order integration_ suffices.   If we had to integrate the coefficients accurately, these elements would have been too costly;  but we know that inaccurate coefficients can sometimes yield quite accurate final answers.   This was expected:  a simple argument leads to the conclusion that the patch test will be satisfied if the rule chosen can integrate element volumes accurately.   The argument went in three steps (Irons 1966):

(1)  An accurately integrated iso-P element conforms and accepts constant stresses, so it passes the test.

(2)  Under constant stress, a less accurately integrated element may give identi-cal nodal loads with the nodal deflections of the patch test, but not otherwise.

(3)  Therefore if 2 replaces 1, the patch remains in equilibrium.   We say that 1 and 2 are _mixable_, in the sense that we could mix them indiscriminately in a patch.

If we can postulate that in the limit all the edges are relatively straight, and the midside nodes are truly central − the sides are _linear_ − the position is even better.

In this case, all we need integrate accurately is a determinant (Irons and Razzaque 1972), typified by the force in the x-direction at node i due to a given state of stress throughout a 20-node brick:

$$- \int \int \int \begin{vmatrix} \partial x/\partial \xi & \partial x/\partial \eta & \partial x/\partial \zeta & \sigma_{xx} \\ \partial y/\partial \xi & \partial y/\partial \eta & \partial y/\partial \zeta & \sigma_{xy} \\ \partial z/\partial \xi & \partial z/\partial \eta & \partial z/\partial \zeta & \sigma_{xz} \\ \partial N_i/\partial \xi & \partial N_i/\partial \eta & \partial N_i/\partial \zeta & 0 \end{vmatrix} \; d\xi \quad d\eta \quad d\zeta$$

But this is for a classical patch test, with prescribed deflections at the external nodes. However we should be able simply to apply the appropriate tractions over all the patch boundaries. Success depends also on integrating expressions like that for the x-force at node i due to tractions over the face $\xi = 1$:

$$- \int \int N_i \begin{vmatrix} \partial x/\partial \eta & \partial x/\partial \zeta & \sigma_{xx} \\ \partial y/\partial \eta & \partial y/\partial \zeta & \sigma_{xy} \\ \partial z/\partial \eta & \partial z/\partial \zeta & \sigma_{xz} \end{vmatrix} \; d\eta \quad d\zeta$$

The requirements for product-Gauss rules are identical.

If the edges are linear, x for example has no terms like $\xi^2 \eta \xi$, although these still occur in $N_i$. Then for the 20-node brick we can show that the 2 x 2 x 2 Gauss rule suffices: a reassuring conclusion in view of the popularity of 2 x 2 x 2 integration. (It does however give mechanisms.)

Note: The bi-quadratic quadrilateral, with its 9 nodes (the last being at the centre) passes the quadratic patch test. When the internal freedoms have been eliminated, it looks like a variant of the 8-node element. The norm would be 3 x 3 integration.

The 8-node Ahmad layered plate has a less happy ending. I had always assumed that a quadrilateral with linear sides would allow constant bending stress, because one can treat the membrane layer $\zeta$ = constant as a case of plane stress, where the patch test is satisfied even with 2 x 2 integration. Razzaque had his doubts, and some two years ago he quietly did the test. It is easy to see the fallacy in my argument.

To achieve constant membrane strain in the layer $\zeta$ = constant with normal remaining normal, it is kinematically necessary to have a lateral deflection w varying quadratically with x and y. With quadrilateral geometry $x = a + b\xi + c\eta + d\xi\eta$, so that $x^2$ contains the term $\xi^2\eta^2$. This is not available, which is why Razzaque's patch test failed. Thanks to a timely experiment, I did not repeat this particular blunder in Semiloof (Irons 1973a; Irons 1973b).

It is doubly unfortunate that the first patch test argument depended on symmetry. Such arguments have never been used since. Moreover, the conclusions on the Ari Adini Parallelogram (Adini & Clough 1961) were weaker than they need have been, and it was only when R H Gallagher suggested that a patch test on parallelograms of varying sizes would succeed that we tried such a mesh. A restrictive argument actually delayed progress.

With a little imagination, we can discuss this more general case with less difficulty than the original demonstration. Considering a unit square, the basis for the shape functions can be expressed in terms of the cubics of Figure 4. We have the four products $L_i(x) L_j(y)$, i and j = 0 or 1. Then we have the eight products, $L_i(x) S_j(y)$ and $S_i(x) L_j(y)$. These twelve functions are activated by the value and two slopes at each of the four corners.
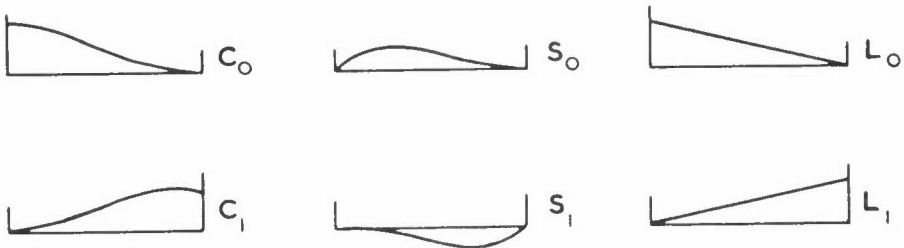


Figure 4: Constituent cubics for the Ari Adini rectangle

We imagine that a patch of such elements is in a target-state of constant bending moments. Now comes the trick: we consider that the patch is perturbed by the products $C_i(x) C_j(y)$, $C_i(x) S_j(y)$ and $S_i(x) C_j(y)$ in which C consistently replaces L. With these functions we can preserve slope continuity between elements. It follows that if we actually used L instead of C, the change in slope would be the same at opposite edges of the elements, so that as much virtual work would be lost at one as would be gained at the other: therefore the total would be zero.

An even simpler argument of the same type applies to the Veubeke constant stress equilibrium triangle which connects to its neighbours at three midside nodes. Any
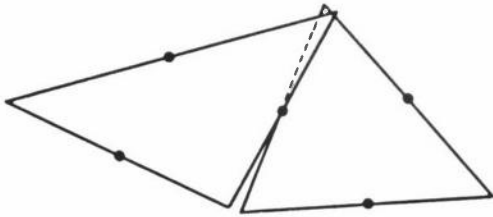
Figure 5: Nonconformity between two
Veubeke elements
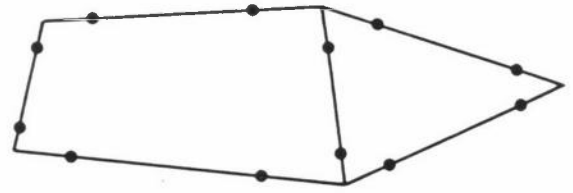


Figure 6: Diagram of elements with Loof nodes

nonconformity must be like that depicted in Figure 5 which does zero work in a field
of constant stress.

Elements with <u>Loof nodes</u> at the two Gauss points along each side, as in Figure 6, can
be treated in the same way.  The shape functions are cubic along the edges, for both
the triangle and the quadrilateral.   The assumed stress field is still constant, so
we can use the 2-point Gauss rule to compute exactly the work due to nonconformity,
ie zero.

The argument for the Veubeke triangle must be extended slightly for the <u>Morley</u>
<u>equilibrium plate bending triangle</u>.   This has quadratic lateral deflection w, giving
uniform bending stress controlled by w at the corners and normal slopes at the midsides.
The top membrane, for example, is under constant strain, and the in-plane deflection
normal to the edge is common to two neighbouring elements, being controlled by the
normal slope at the midside.    Finally the slope along the side is given by
$(W_B - W_A)/AB$ in Figure 7, so that the deflection along the edge is also common between
neighbouring elements.    Therefore nonconformity in that membrane can do no work in a
patch test.
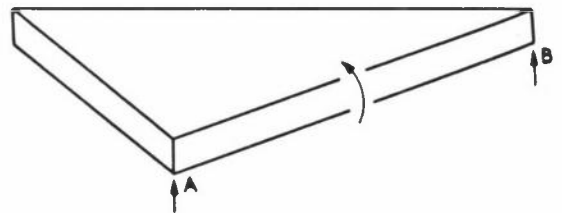
Figure 7: The Morley bending triangle



We have now encountered all the ingredients in the synthesis which enabled the
"Semiloof" shell element to pass the test - the first Ahmad layered model to do so -
for general quadrilaterals and triangles with linear sides and again when the 2 x 2
Gauss integration was changed to 5-point integration to remove the mechanisms.   This
promising element owes its very existence to the patch test (Irons 1973a;   Irons 1973b).

## ALCHEMY AND THE PATCH TEST

We have seen that the patch test transcends the displacement or equilibrium formula-
tions. Let us now enquire what is possible with "nondescript" elements, perhaps with
no physical basis whatsoever. What other properties must a formulation have, given
only that it passes the patch test? We must try to discover what is intrinsically
necessary, and also what is intrinsically possible. Can we for example find a four-
node element, with four degrees of freedom, that passes a patch test with quadratic
variation of the variable? The quick answer is obviously, no. For we could choose
as target-state a quadratic variation giving zero values at the four nodes - six are
needed to define it - and the four zero nodal values would presumably give zero
responses everywhere. But

(a)  If the stresses are to be calculated at certain prescribed points, it is
     conceivable that zero nodal values always give zero stress at those points,
     whatever the quadratic variation.

(b)  In say an elastic problem we could, perhaps, invoke for example the equations
     of equilibrium, which operate in the patch test, and hence reduce the number
     of independent coefficients to less than six per variable.

These are wild thoughts. But our next is slightly more sober. Suppose we use
hexagonal elements, with six simple nodes at the corners? We should have just enough
degrees of freedom to define a quadratic variation, and provided the six vertices do
not lie on a conic section the shape functions certainly exist. If it were possible
to create a mesh of decagons, could it pass a cubic patch test? These are still
flippant suggestions, but they lead to the serious general question; is it possible
for an element to pass a higher patch test by virtue of its number of degrees of
freedom, regardless of the dispostion of nodes along a typical side? In particular,
can we ever do better than a conforming model with the same nodal configuration?
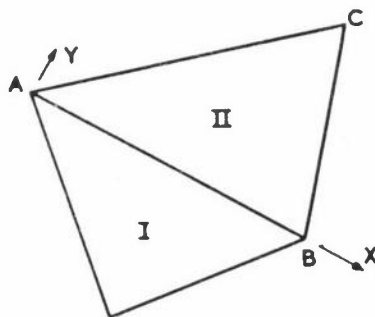
Figure 8: Two elements considered together,
          to develop conformity requirements

To be more precise, let us consider what are the unavoidable constraints in designing
a nodal configuration for a conforming element to do a given job. In Figure 8, elements
I and II have $C^{(1)}$ continuity across AB, regardless of the nodal values. In order to

make deductions we assume that all the nodal values of element I are zero, so that its response must be zero everywhere. In particular w and $\partial w/\partial y$ are zero at all points along AB. We assume that II can give all the cubic terms:

$$w = a + (bx + cy) + (dx^2 + exy + fy^2) + (gx^3 + hx^2y + ixy^2 + jy^3)$$

The minimum connectivity requirement for $C^{(1)}$ is $w = 0$ at $y = 0$ giving $a = b = d = g = 0$, and $\partial w/\partial y = 0$ at $y = 0$ giving $c = e = h = 0$. This implies four simple nodes along AB, with w only, and three nodes with $\partial w/\partial y$, as nodal variable. We say $n_s$ = number of nodal tractions = 7. We may replace some of the w by $\partial w/\partial x$ or $\partial^2 w/\partial x^2$, and some of the $\partial w/\partial y$ by $\partial^2 w/\partial x \partial y$ etc. Our assumption initially is that there are not more than $n_s$ in all.

Note: The importance in later arguments is that we shall have just enough nodal tractions to give the correct work done at the element boundaries due to bending moment and shear force, in going from one target-state to another.

We add the corollary that if, for example, we had $\partial^2 w/\partial y^2$ at a midside node on AB in Figure 8, a variable superfluous to the $n_s$ required, these arguments would still hold. For such a variable must control a shape function with zero w and normal slope on all the element boundaries. It would give patch equilibrium in each of the elements concerned, considered separately. As far as the patch test is concerned, it makes no difference if we enforce equality between neighbours.

Thus $n_s$ may exceed the minimum number of tractions without necessarily excluding an element from the theorems we shall develop. We must certainly exclude an element which could, in principle, pass a higher patch test than it in fact does, for example the "quadratic" isoparametric elements which pass only a linear patch test. To clarify the subtle distinction, we have three just-successful configurations in Figure 9.
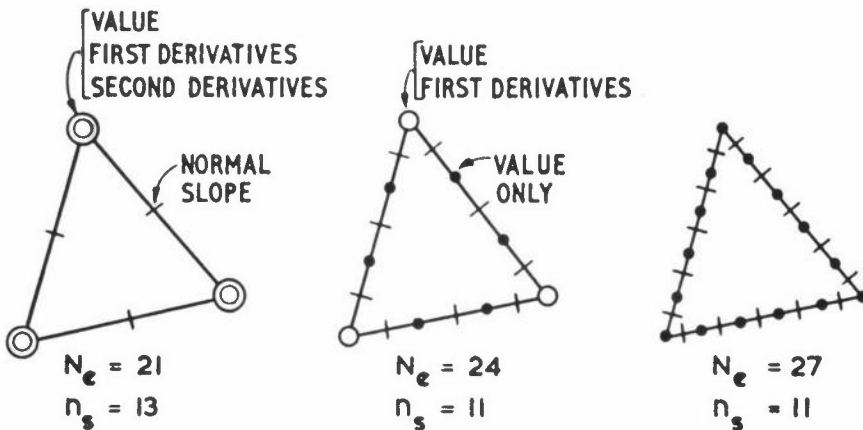


Figure 9: Three quintic triangles for plate bending, with $N_T = 21$

It happens that the first is the only version giving analytic shape functions – but this is not relevant to the count of degrees of freedom. The second is just feasible in a displacement formulation, because the functions with singularities give finite (but high) strain energy. The third gives indeterminate <u>slopes</u> at the corners; but Pian's approach remains open even in a fourth version, without any corner nodes, so that the corner <u>values</u> are indeterminate, as with Loof nodes.

The first gives us <u>excess continuity</u>, but here – as usually – it is not of the simple variety envisaged in the corollary above. As a variable, $\partial^2 w/\partial y^2$ at A in Figure 8 would not now be wasted, but would be needed for the continuity of side AC with another neighbour. Indeed, the three second derivatives at a corner node are called upon to do <u>four</u> jobs. Notwithstanding $n_s$ being 13 instead of 11, by arguments of nodal valency we should have <u>fewer</u> assembled degrees of freedom (Irons 1973). (See also appendix.)

Our alchemy is concerned with <u>super-patch behaviour</u> which would imply <u>super-convergence.</u> That is, we seek a "nondescript" element which passes a higher patch test than would be possible with a conforming element. Later we shall prove that we cannot succeed. But is it worth the effort? Indeed yes: this is not fool's gold. For example, Figure 10(a) shows an attractive 14-node brick element which could perhaps give excellent performance. But could it pass a quadratic patch test, in view of the fact that each face has only 5 nodes, whereas 6 are needed to define a quadratic variation over a plane? Again, Figure 10(b) shows a plate bending element, of familiar nodal configuration – the value and the two slopes at each corner. It has 12 degrees of freedom, and a complete cubic in x, y requires 10. We doubt however whether such an element can pass a cubic patch test, because the nodes along any one side contain only enough information to define a <u>linear</u> variation of normal slope. We can say decisively that a conforming displacement model cannot pass. Could a hybrid succeed?



(a)          (b)
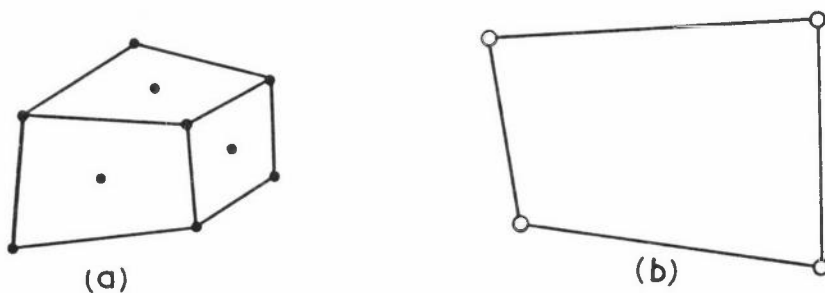
Figure 10: A plausible replacement for the 20-node isoparametric brick, and a nondescript plate bending element

make deductions we assume that <u>all</u> the nodal values of element I are zero, so that its response must be zero everywhere.    In particular w and $\partial w/\partial y$ are zero at all points along AB.    We assume that II can give all the cubic terms:

$$w = a + (bx + cy) + (dx^2 + exy + fy^2) + (gx^3 + hx^2y + ixy^2 + jy^3)$$

The minimum connectivity requirement for $C^{(1)}$ is w = 0 at y = 0 giving a = b = d = g = 0, and $\partial w/\partial y = 0$ at y = 0 giving c = e = h = 0.    This implies four simple nodes along AB, with w only, and three nodes with $\partial w/\partial y$, as nodal variable.    We say $n_s$ = number of <u>nodal tractions</u> = 7.    We may replace some of the w by $\partial w/\partial x$ or $\partial^2 w/\partial x^2$, and some of the $\partial w/\partial y$ by $\partial^2 w/\partial x \partial y$ etc.  Our assumption initially is that there are not more than $n_s$ in all.

<u>Note</u>:  The importance in later arguments is that we shall have just enough nodal tractions to give the <u>correct work done</u> at the element boundaries due to bending moment and shear force, in going from one target-state to another.

We add the <u>corollary</u> that if, for example, we had $\partial^2 w/\partial y^2$ at a midside node on AB in Figure 8, a variable <u>superfluous</u> to the $n_s$ required, these arguments would still hold. For such a variable must control a shape function with zero w and normal slope on all the element boundaries.    It would give patch equilibrium in each of the elements concerned, considered separately.    As far as the patch test is concerned, it makes no difference if we enforce equality between neighbours.

Thus $n_s$ may exceed the minimum number of tractions without necessarily excluding an element from the theorems we shall develop.    We must certainly exclude an element which could, in principle, pass a higher patch test than it in fact does, for example the "quadratic" isoparametric elements which pass only a <u>linear</u> patch test.    To clarify the subtle distinction, we have three just-successful configurations in Figure 9.



Figure 9: Three quintic triangles for plate bending, with $N_T = 21$

This brings us to our conclusion. For in Figure 12 we have introduced a very small
element at each vertex. According to the lemma for small elements we can argue that
the forces they exert in a field of constant stress are insignificant. Ordinary
dimensional analysis provides this result for the moment, and a better argument will
soon appear. Therefore the boundary forces shown in Figure 12 are entirely due to
the elongated boundary elements. We conclude that for a given formulation each side
makes its individual, calculable contributions to the boundary forces at the nodes on
that side. And this conclusion is inescapable - it follows directly from the nature
of the patch test, for the most criminal elements. Perhaps the patch test is less
permissive than we thought.

Note: We now return to the lemma for small elements. In Figure 12 the slender
elements generate certain nodal forces. If we now discard the small elements, these
forces will change only slightly, on account of the change in boundary geometry. We
interpret this change in terms of the small elements taking their small share of the
traction, and without proving the lemma strictly this is enough. We have separated
the slender elements.
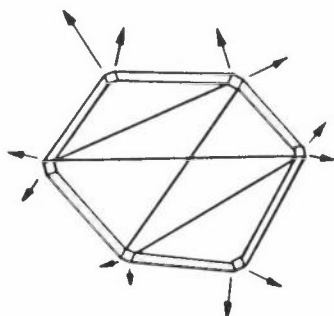


Figure 12: A variant on the inner
patch
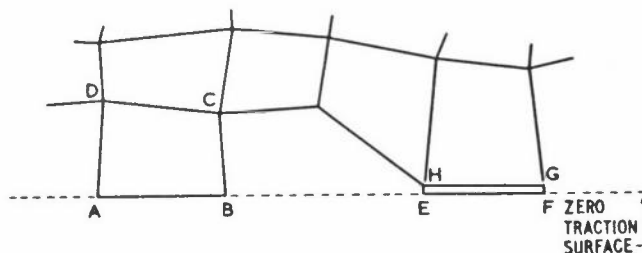


Figure 13: A patch with zero traction

Let us now consider the implications of the patch in Figure 13, with a flat surface,
across which there is zero traction; thus the available target-states are limited.
We shall prove the zero traction theorem, that the nodal forces on AB are zero if the
tractions are zero. For we could make ABCD a needle-shaped element like EFGH without
changing these forces, and we can make the forces on EF arbitrarily small by reducing
the depth of EFGH. (The forces in BC, including the corner node at B, need not be
small.)

For the next theorem we shall need the lemma of element strain energy, more obvious in
its statement than in its implications. Of course, if the stress everywhere repro-
duces the target-state exactly, then the strain energy integrated over the element
must be correct. Therefore, it makes no difference if the formulation studiously
avoids energy principles - the patch test introduces them inescapably in this sense,
provided that we consider only the target-states.

Our next theorem is strongly suggested by the zero-traction theorem, but at this point we must introduce the assumption that the nodal configuration along each side would be just acceptable in a conforming model, as discussed above.  Thus we suppose that element ABCD in Figure 14(a) has $N_e$ nodal variables and that there are $N_T \leqslant N_e$ target-states.  To prove the mixability theorem we shall try to compute the patch forces at the nodes on AB by considering only $n_s$ of the target-states, fixing to zero $(N_T - n_s)$ of the nodal variables not on AB.  (It may be necessary to choose carefully which to fix.)  Let us imagine that the remaining $(N_e - N_T)$ non-target responses are suppressed, by constraining the remaining variables.  We now perturb the patch, to another target-state only slightly different;  hence we do work directly on the element through the nodes on AB, but we also allow energy to be transmitted from neighbouring elements through any consequent perturbations of the remaining $(N_e - N_T)$ constrained variables.
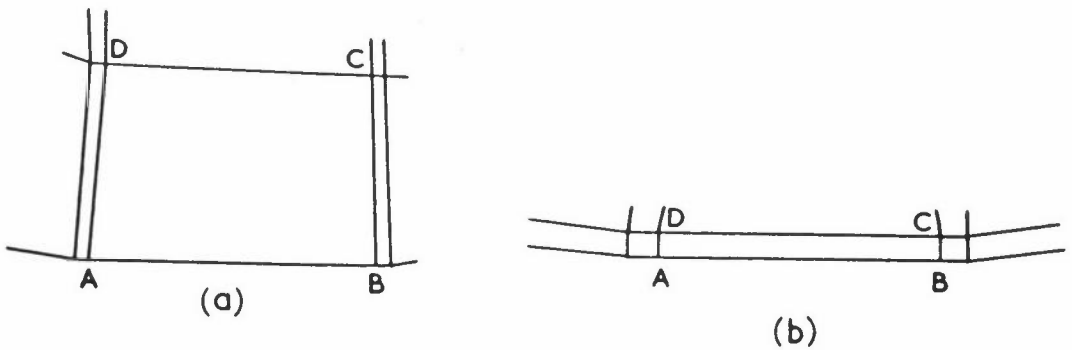


**Figure 14:** Patches for arguing the mixability theorem

To complete the argument, we go back and repeat the perturbation, but this time we do not allow the $(N_e - N_T)$ variables, previously constrained, to depart from their initial positions.  Thus we prevent the nodes not on AB from receiving energy from neighbouring elements, and in so doing we commit only a second order error in the total potential energy, beacuse of patch equilibrium at the internal nodes.  The only energy now entering the element is that transmitted via the nodal tractions on AB, and it must therefore equal the virtual work computed along AB, and along AD and BC for the corner nodes, as a continuum, considering only the target-states.  We now move on to Figure 14(b) in which the virtual work along AD and BC will be relatively small.  The argument still holds, for any perturbation of AB.

But from the point of view of virtual work, the nodal values along AB just define the values and, if necessary, the slopes on AB for any target-state.  The nodal values are correct, we assume.  Thus whatever the inventor thinks is happening along AB, it is as if the target-state values were there in any patch test.  Therefore we can compute the tractions for the patch test, if only we assume that the virtual work done at the nodes is correct.  It would be difficult to sustain a model, for example, with an unsymmetrical stiffness matrix:  the work done would be patch-dependent.  Thus the

argument brings us especially close to Pian's viewpoint, although by a totally un-familiar route.

Note:    It is fortuitous that all adequately integrated iso-P elements are mixable, including those whose patch behaviour is not as good as would be suggested by the nodal configuration.


The theorem just proved has a close affinity with the super-convergence theorem, our last.    Indeed, it is possible to use the patch of Figure 14(b), or even that of Figure 12 with the original lemma for small elements, but we choose here to vary the argument.    The rudimentary patch of Figure 15 contains only a slender element, flanked by two small square elements.    (Indeed it is not strictly a patch at all, without any internal corner nodes.)    There is little change if we remove the external patch tractions from the small elements.    If we wish to discard all energy principles, including virtual work, then we can proceed taking this statement as a premise. Other-wise we can argue it from structural principles, ie if element CDFE has fixed displacements along DF and CE, then the change in the strain energy of CDFE cannot exceed a relatively small quantity.

A    C                                                    E    G

B    D                                                    F    H

Figure 15:  Patch for the super-convergence theorem

Lemma    We prove the statement just made by dividing the inter-element boundaries of Figure 15 into three regions:

Z = DF + CE, whose displacements remain fixed according to the target-state.

Q = CD + EF, the inter-element region which remains free of external loads.

L = (AC + AB + BD) + (FH + HG + EG), whose displacements are initially those of the target-state;   but subsequently L becomes free, ie unloaded.

When the patch is initially set up, Z and L are given the target-state displacements, and an easily calculable amount of work becomes strain energy.    Then L is slowly un-loaded, and as the boundary moves from displacements $1$ to $1^*$, the patch does work on the decreasing forces on L, reducing the strain energy by the same amount.    (The forces on Z also change, but because Z remains fixed no work is done.)    We note that $(1^* - 1)$ is the displacement caused by applying load L to the unloaded patch, earthed along Z, giving the required change in strain energy exactly.    However, we prefer something larger - a bound - and something more easily calculated.    In withstanding

the end-loads L, the slender element must stiffen the small elements, so we now remove it, thus isolating the small elements. Therefore we set Q to zero, and we apply loads L to the small elements, giving the strain energy e as our bound on the change in strain energy. e is normally small.

Note: e can be infinite if the square elements have mechanisms - for we are no longer dealing with target-states. This proviso was implied more subtly in Figure 14, where the removal of the constraints could have caused infinite deflections elsewhere.

Having removed the tractions from the small elements, we cannot expect to reproduce any target-state exactly, but the modified patch test must evidently give nearly the correct stress field. The conclusion follows. For the number of distinct stress fields available (which are nearly target-states), plus the rigid body motions, must be 2n, just as for a conforming model. The 2n items of information - the nodal tractions - entering the element along sides EC and DF must define the displacement field adequately for any target-state when EC comes close to DF, which is the same requirement as for a conforming element. Otherwise we can argue that with less than 2n tractions, or with inappropriate tractions, the virtual work in going from one target-state to another might be wrong.

We conclude, sadly, that <u>super-patch behaviour is intrinsically impossible</u>, and that workers should look elsewhere for profitable research. If the patch test has led us into an alchemist's dream, it has also taken us a step nearer to modern chemistry.

## CONCLUSIONS

Most of the material reported here was unknown to me at the time of our 1972 paper (Irons and Razzaque 1972). I did not expect this outcome, which reflects the indispensibility of the patch test in my research. One's attitude towards those tools that one uses only occasionally does not change from year to year, whereas there are other tools for which one constantly finds new uses and new adaptations. For years I imagined that all my fellow workers used the patch test in their secret thoughts, and I was only gradually persuaded otherwise. I needed it, and I assumed others needed it too.

The usefulness of the patch test requires no pedagogic thumping - the examples speak for themselves. Therefore I should like, finally, to emphasize the subtle and characteristic philosophy that was struggling to emerge in the final section above. It gave me

great pleasure to discover that the patch test itself is leading us towards the concept of the hybrid elements, and not the other way around. For the patch test is more fundamental and universal than any single formulation. It begins to appear that Pian's technical note of 1964 (Pian 1964) held the whole future, not just a part of it.

## APPENDIX

I once thought that the frontwidth was independent of how the necessary nodal tractions were distributed between corner, midside and mid-face nodes. This is far from true. Figure 16 considers possible nodal configurations for a quintic plate bending element, and we imagine a rectangular problem $n$ elements in width, treated with a horizontal front (indicated by the full lines) progressing vertically as the equations are reduced.

(a)     frontwidth     $=$     $7n + 21$

(b)     frontwidth     $=$     $8n + 24$

(c)     frontwidth     $=$     $6n + 26$

(The last is probably of academic interest only.)
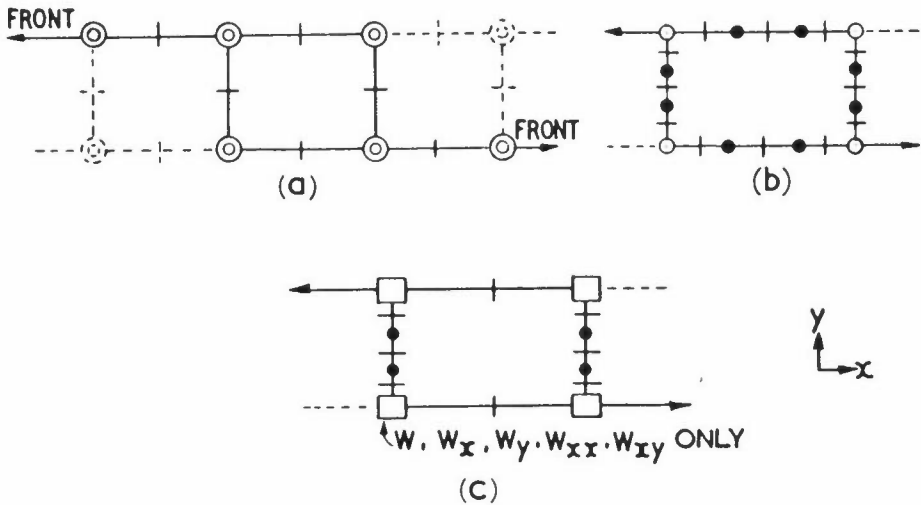


Figure 16: Nodal configurations for discussing frontwidth

## REFERENCES

ADINI A and CLOUGH R W, "Analysis of Plate Bending by the Finite Element Method", Report to National Science Foundation, USA, G.7337 (1961)

AHMAD S, "Curved Finite Elements in the Analysis of Solid Shell and Plate Structures", PhD Thesis, University of Wales (1969)

BAZELEY G P, CHEUNG Y K, IRONS B M and ZIENKIEWICZ O C, "Triangular Elements in Bending - Conforming and Nonconforming Solutions", Proc Conf Matrix Methods in Struc Mechs, AFIT, Wright-Patterson Air Force Base, Ohio (1965)

FRAEIJS DE VEUBEKE B, "Displacement and Equilibrium Models in the Finite Element Method", Stress Analysis, ed Zienkiewicz and Holister, Wiley (1965)

GARFIELD WILLIAMS J, "Studies in Pseudo-Hybrid Plate Bending Elements and the Patch Test", BSc Thesis, Civil Engineering, University of Wales, Swansea (1974)

IRONS B, "Engineering Applications of Numerical Integration in Stiffness Methods", J AIAA, $\underline{4}$, 2035-2037 (1966)

IRONS B, Lecture Notes, International Research Seminar on the Theory and Application of Finite Elements, a NATO Advanced Study Institute, Calgary University (1973)

IRONS B, "Un Nouvel Element de Coques Generales - Semiloof". Methods de Calcul Scientifique et Technique, Colloques IRIA, Rocquencourt, Paris (1973)

IRONS B and RAZZAQUE A, "Shape Function Formulations for Elements Other than Displacement Models", Conference on Variational Methods in Engineering, Southampton University (1962)

MORLEY L S D, "On the Constant Moment Plate Bending Element", Journal of Stress Analysis $\underline{6}$, 20-24 (1971)

PIAN T H H, "Derivation of Element Stiffness Matrices by assumed Stress Distribution", JAIAA $\underline{2}$, 1332-4, (1964)

STRANG G and FIX G J, "An Analysis of the Finite Element Method", Prentice-Hall (1973)

# PREPARATION OF FINITE ELEMENT DATA

D A C Parkes

Department of Mechanical Engineering
The University
Nottingham
NG7 2RD

## PREPARATION OF FINITE ELEMENT DATA

Data prepared for a finite element analysis of a structure is set out in a particular manner. Having selected a suitable idealisation of the structure and chosen the method of solution to be used the data can be prepared. In general it is necessary to give each node in the structure a label or number and calculate the coordinates of these nodes in some suitable coordinate system.

The figure shows the particular method of defining the data necessary to run a finite element analysis using PAFEC 70+. Detailed descriptions of each of the options are given in the sections referred to by the flow chart (Nottingham University 1974). The basis of the PAFEC 70+ data is:

(1)    Control Integers
(2)    Nodal Coordinate Data
(3)    Element Topology
(4)    Loading Information
(5)    Constraint Data

All finite element programs will require the above 5 sets of information but these may be supplied in a different manner to those described for PAFEC.

## REFERENCES

NOTTINGHAM UNIVERSITY, PAFEC 70+ Data Preparation Guide (1974)

## DECISION CHART FOR DATA PREPARATION

Note: All variables are fixed field formats.
Integers are those beginning I-N.
Real data is fieldwidth of 10.
Integer data is fieldwidth of 5 right
justified except where stated.

START

Is data only system available

yes →

*Pafec Macro Option Cards*

PAFEC JOB

Option Cards

ZZZZ          Ref. Section 1.0

no

Job Description Card
TITLE OF JOB — PINCHED CYLINDER —
Ref Section 1.1

Program Control Card
IM  IP  IDT  ID  IDF  IX  IW  IK  ILO  IE  ICD  JDC  ISS  ICA  ICB  ICC
Ref Section 1.2

ICC  > 1 →

Drawing Data Card(s)
ALPHA        BETA        XLTH        BRTH    NP   NE1  NE2  IN
Ref Section 1.3

0

Cpddc Array Data — Co-ordinates, Degrees of Freedom, Local Direction Cosines

| X | Y | Z | IU | IV | IW | IOX | IOY | IOZ | CONT | If IW ≥ 1 |
| X | Y | Z | U | V | W | | OX | OY | | If IW = 0 |
| OZ | CONT | | | | | | | | | |

or

DCA12  DCA13  DCA21  DCA22  DCA23   If IX=18,CONT≠9999.0
DCA11  DCA12  DCA13  DCA21  DCA22  DCA23  If CONT=INODE  Ref Section 1.4

Standard Element Data          or          Pafblocks Data
IERN   IL   ISN   IB   MW   NW   NA   N   NEI   ISSN
ITP   NN   NM   NL   NK  MESH   INM  ICIRC  INEST
Element Topology and Material Properties          Pafblock Topology and Material Properties
Ref Section 1.5.1          IR1  IR2  IR3  IR4  (Mesh Spacing)
N1   N2   N3   (Circular Arc)
INB   IER   (Nesting)   Ref Section 1.5.2

IW  1 →   IM  2,4,5   IM  6   IM  12   IM  7,8   IM  10,11

0

Back Substitution Data
for Vibration Problems
IC    N (format I5,I10)
Ref Section 1.6.1

Centrifugal Speed
OMEGA rad/s
Ref Section 1.6.6

Pressure or Thermal
Loading Data
X1
Ref Section 1.6.7

Transient
Temperature Data
DELTA  IKN  INS  JP  TMAX  IQ
N1   N2   SPRING (INS cards)
N    T   TDOT (IKN cards)
NQ   Q   QDOT (IQ cards)
Ref Section 1.6.10

X1  9999.0 → NODE XNODE      8888.0 → XNODE (IP-1 cards)

IM  11 → TEMP (IP cards)
Ref Section 1.6.11

PRESSURE

Steady State
Temperature Data
IKN  INS  IQ
N1   N2   SPRING(INS cards)
N    T   (IKN cards)
N    Q   (IQ cards)
Ref Section 1.6.9

External Load Data
F1
Ref Section 1.7

F1  7777.0 →   F1  5555.0  6666.0  9999.0
8888.0

FNODE(IDT-1 cards)
Ref Section 1.7.1

X  Y  Z  IDISP  ILO  F
Ref Section 1.7.2

NODE  IDISP  ILO  F
Ref Section 1.7.3

IDEF  ILO  F
Ref Section 1.7.4

IK  0 →

Constraint Data (If applicable)
INODE  IDISP  IREPT
Ref Section 1.9

Stressing Data
N1   N2   N3
Ref Section 1.8

# COMPUTER GRAPHICS IN SCIENCE AND ENGINEERING

J W E Lewis

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 0QY

## INTRODUCTION

Computer graphics is basically very simple, consisting of combinations of three operations.

(a)    Draw a line

(b)    Not draw a line (remove the pencil from the paper)

(c)    Finish the picture (ie replace drawing with a fresh sheet of paper)

These operations are used not only by computers but by human draughtsmen, designers and artists.

Engineering and Science are based on exact laws with a mathematical superstructure. This means that unlike "pictorial" art, scientific and engineering graphics are in the main numerically definable making them easy to produce by digital methods.

What is required to produce computer graphics, besides access to a digital computer? These requirements can be subdivided into hardware and software requirements.

## GRAPHICS HARDWARE

We shall first consider the main types of hardware that are presently available.

(a)    Flat bed and drum plotters
       The flat bed plotter is simply a large table with a pen gantry such that the pen can move anywhere on a sheet of paper placed on the table.

Drum plotters consist of a roll of paper, a drum on which the plotting is done and a take up roller. The pen traverses the width of the paper and the paper moves back and forth to give the required drawing capacity. Such plotters are usually up to 36 inches wide and allow plots up to 140 feet long, an advantage to wallpaper designers.

Flat bed plotters have the advantage of producing good quality plots of often very large size. However, they do tend to be rather slow.

Drum plotters on the other hand are less accurate, usually a little faster than flat bed plotters, but generally do not produce pictures greater than 36 inches in width.

(b)  Visual Display Units

These devices are based on cathode ray tubes. Two types are generally available, first the refreshed type where the picture is continually regenerated from some digital representation held in the controlling computer's memory. Second, the storage tube devices. These are marketed by Tektronix and use the phosphor and CRT technology developed for their memory oscilloscopes. The picture is stored on a long lived tube phosphor and after transmitting the picture to the VDU the computer plays no part in keeping the picture on the screen.

(c)  Microfilm plotters

These are a cross breed between (a) and (b). In some ways they offer greater flexibility over the (a) type plotters, but they possess certain disadvantages. They consist of a special CRT on which the picture is generated. This picture is photographed either onto photographic paper which is processed to produce a paper picture, or on to photographic film to produce a microfilm image. This latter facility can be used to compact documents and also gives the possibility of making computer generated films.

## GRAPHICS SOFTWARE

Let us now consider computer graphics software, that is, how the various types of plotter can be driven.

Pictures consist of a large number of lines, or vectors defined by the coordinates of their beginning and end points (or in some cases just the end point, the starting

point being defined by the previous vector's end point).   Thus any but the simplest drawing graphics is fairly demanding on the output of vectors to the plotter; especially when it is remembered that most plotters can only draw straight lines, and curves must be approximated by a number of straight lines.   Plotters are generally slow devices compared to computer CPU and other peripherals, and it is better to run such devices off line.   Essentially the user's program, which might be very large, generates a magnetic tape.   This tape is either taken to the plotter which is physically separate from the computer or it is played back to a connected plotter using a small efficient program.   The SD4020 at ACL works in the former mode whilst the drum plotters at many universities work in the latter mode.   VDUs are rather different;  storage tubes can be run at relatively high speed and are usually run online to the computer but off line to the users program. the interface being via a disc file.   In smaller installations such devices can be run online to the users program with a limited amount of interaction.

The refreshed devices require continual servicing by the driving computer.   They are generally used for interactive applications and enable the use of light pens etc.

Let us now briefly consider interaction especially as far as it concerns graphics. Drum plotters and microfilm plotters are obviously one way devices.   Storage tubes offer a limited degree of interaction in that lines can be added to a picture but not removed without replotting the entire picture.

Refresh displays where the picture is effectively being replotted continually do not suffer from this disadvantage and are particularly suitable for heavy interactive work.

Interactive graphics where graphical information flows both ways is an expensive activity and should be resorted to only when other graphical methods cannot be used or there is some tremendous advantage in using interactive techniques.   Generally interactive methods do not help when a problem can be defined algorithmically, and this would cover most problems encountered in science and rather less in engineering.

In dealing with heuristic problems, eg design work, they are almost essential.   Such methods usually require specialised hardware.   A particularly good example where interactive graphics has proved valuable in an algorithmic case is the National Economy.   The Central Statistical Office has an interactive graphics facility which enables economists with no computing training to manipulate statistics, correlate,

predict etc both rapidly and simply.    The economic gains in such a case are obvious,
although in many cases it is not so clear cut.


Between fully interactive graphics and graph plotter type graphics lies online
graphics where a storage tube device is used to view generated pictures after the
completion of a job.    Here the flow is one way.    Such methods enable rapid
development of graphical programs.    Although the operations of graphics are
fundamentally simple various problems can arise concerning scaling of data,
generating characters, drawing arches and graphics.    This is the area of software.
Much graphics software is very crude, the user having to manipulate the plotter as if
he were drawing the graph himself.    A lot of effort has gone into packages that
allow automatic graph production, scaling and character generation.


## COMPUTER MOVIES

Earlier I mentioned the possibility of making computer movies.    A movie consists of
a number of frames, each differing slightly from the previous frame, such that when
the film is shown an apparently smooth action is portrayed.    Silent films are shown
at 18 frames per second (fps) and sound films at 24 fps.    Thus a few minutes of
film requires quite a lot of frames.    A lot of software effort has gone into pro-
ducing programs to allow users to use some of the tricks of cinematography easily
eg zooms, pans, titling etc.    They may not be of professional standard but scientists
and engineers are primarily interested in research films ie subject content and
things like "nice" titling are just frills.


I will now show a short film entitled "The Flexpede".    It is not about anything
scientific, and is really very frivolous.    It is very simple both in terms of what
it tries to do and in the graphical shapes that are manipulated.    There is basically
one shape plus a few others that are linked together to produce what I feel is a
very effective little film.


In both science and engineering computer movies can be useful for studying the time
evolution of a system in a qualitative fashion.    Often one might wish to expand or
compress the time scale.    In the case of molecules moving in a liquid at speeds of up
to several metres a second some expansion of the time scale is required.    In the case
of galaxies however where there is little change over a million years the time scale
must be compressed.

Of course, it is possible to map some other variable onto the real time axis of the
film eg stress on a component as a linear function of time.

## CONCLUSION

I would like to sum up by making the following points:

(a)   Computer Graphics are essential for the automatic production of graphs and
charts, which can lead to a more rapid and thorough understanding of the
numerical information emerging from the computer.

(b)   It is important to choose the right graphical technique for a given type
of problem.

(c)   Computer generated films can now give a time dependent display facility,
not cheaply but more easily and cheaply than electronic displays.

# INTERACTIVE GRAPHICS FOR FINITE ELEMENTS

G A Butlin

Department of Engineering
The University
Leicester
LE1 7RH

## INTRODUCTION

In 1968 the Engineering Department at Leicester University took delivery of an ICL
(then Elliott) "4280 advanced display" (purchased with funds provided by the Science
Research Council) to be attached to the University's 4130 computer for "investigating
engineering uses of digital computer graphical display equipment" under the leader-
ship of Professor G D S MacLellan.   Further funds were subsequently provided for a
disc and a teletypewriter.   The author spent four months in the early part of 1968
in the computer-aided design group led by Dr D T Ross in Project MAC (Butlin 1968).
The fundamental graphics and data handling techniques and software that have subse-
quently formed the basis of all the work at Leicester have close links with the AED
approach to computer applications (which is described by Ross (1967)).   Almost all
the programming has been done in Fortran, the primary aim being to ease transfer of
programs to other computing systems.

The hardware configuration used for all the work described in this paper is shown
in diagrammatic form in Figure 1 and the display terminal is shown in Figure 2.
Software appropriate for operating the display in a timeshared mode was never forth-
coming from ICL and consequently whenever the display has been used the 4130 has
always had to be dedicated to it.

Among the many investigations of the use of the display in engineering that have been
made in Leicester such as hydraulic control system design, mechanism design, traffic
control simulation, fluid flow visualization etc the main interest has been the
improvement of the finite element method in continuum mechanics.   The long term aim
here has been to provide the means for a designer to improve his understanding of
structural behaviour through the combined use of interactive graphics and the finite
element method.

Two early systems that explored the potential of this combination were

    (1)    LUISA1 (Leicester University Interactive Structural Analysis - 1) and

    (2)    a slab bridge design program .

While limited to certain relatively small structures these exploratory systems led to the identification of critical problem areas such as:

    (i)    digital representation of a complex three-dimensional object in a form immediately suitable for engineering analysis;

    (ii)    techniques for rapid re-analysis and assessment of performance in · connection with engineering design modification;

    (iii)    programming aids for large, changing interactive graphics systems in Fortran;

    (iv)    operating system, data paging and program overlay requirements for severe real time applications;

    (v)    ergonomics and operational strategy of intense man-machine interaction;

    (vi)    the introduction of new design/analysis procedures into existing practices.

Research in these areas is in progress and forms the basis of a four year programme for which funds are being sought.

Recent work has concentrated on producing systems for the rapid generation of data input to and rapid assessment of output from finite element structural analysis. Such systems are seen as an attempt to meet the present need that is illustrated in Figure 4 (Tocher 1970), and also as a basis for future design/analysis systems on the lines of our early exploratory design systems but capable of handling complex structures.

This paper presents the facilities and example applications of:

    (1)    LUISA1 and the slab bridge design programs;

    (2)    an input data generation system;

(3)   an output assessment system;

followed by notes on system design.


(1) <u>Design Systems</u>

   (i)   LUISA1 - was designed as an aid to a structural engineer in investigating
         the effects of changes in element configuration and properties in simple
         two-dimensional structures.   All data defining topology, geometry, displace-
         ment constraints and applied forces are generated, checked and, if desired,
         modified at the display using the lightpen, function buttons and an adjacent
         teletype keyboard.   Figures 5(a)-(r) illustrate some of these features.
         Because of the nature of the 4100 Fortran software the system is limited to
         about 70 elements and it has been possible only to provide for triangular
         elements and for a selected range of types of displacement constraint, applied
         loading and display of stresses.   The system has been fully described in
         earlier publications (Butlin and Hubbold 1969;   Butlin 1969;   Butlin and
         Leckie 1970).


   (ii)   a slab bridge design program.


As a result of a joint project by E W H Gifford and Partners and the Leicester CAD
group a program was written (by Miss Rachel Britton under the direction of M J Platts)
to make possible rapid interaction with computer analysis in the early stages of a
structural design using finite element methods.   It provides the means for an engineer
to study alternative solutions to a design problem at the conceptual stage by making
full use of interaction and display with both input data and computed results.   The
design program for rectangular bridges produced in this project enables an engineer
to obtain an overall picture of the behaviour of a rectangular pre-stressed concrete
bridge deck under pre-stressing, loading and support conditions which can be varied on
command.   (Typical display screen layouts are shown in Figures 6 (a) and 6(b).)   It was
used to good effect in the design of the Newport Harbour 3-span bridge in the Isle of
Wight (shown in Figure 7 (a)), resulting in an improved pre-stressing scheme (shown in
Figure 7(b)).   Programming aspects are described in Britton & Platts (1971) and
experience of the system in use is described in Platts and Britton (1972).


(2) <u>Input Data Generation System</u>

A system of Fortran programs known as the LUISA2 family has been developed to permit
interactive computer graphics to aid the rapid generation of finite element data.
It consists of seven subsystems:

    (i)     LUISA-OUTLIN for structural shape definition

    (ii)    LUISA-PAMGEN for parametric generation of finite element meshes

  (iii)    LUISA-FEMMOD for finite element mesh generation and modification

   (iv)    LUISA-COMPAS for assembly of component meshes into large structures

    (v)     LUISA-ELTYPE for element type and material property assignment

   (vi)    LUISA-NODNUM for nodal numbering

  (vii)    LUISA-LOADBC for specification of applied loads and displacement boundary
conditions.

Each subsystem has access to a library of structural components which forms a data base on disc and provides the means for passing data between subsystems.   Each subsystem is operated on an interactive refresh display terminal.   All input is from the lightpen, function buttons and keyboard of the display terminal;   approximate values usually being input through the tracking cross (which follows the lightpen) with visual assessment of magnitude, and accurate numerical values usually being input through the keyboard;   except in subsystems OUTLIN and PAMGEN where some initial data may come from cards.   Facilities for dimensioning, zooming and windowing, hardcopy plotting and dumping and restoring are made available throughout the systems.

In each subsystem provision is made for re-entry, retrieval of old data, modification and re-filing.   A sample of sequences of operation through the subsystems that the user may follow is shown in Figure 8.

Figures 9 to 14 present a sample of currently available facilities.

    (i)    Figures 9(a)-(f) show the way a simplified outline may be input from cards, checked visually, adjusted interactively with subsystem OUTLIN and passed to the interactive element mesh generation subsystem PAMGEN which provides parametric control of mesh density and grading through a set of patches representing the area enclosed within the outline.

    (ii)   Figures 10(a)-(f) show two further examples with an indication of the total time taken, from designing and punching the initial card input to filing the element data on disc and requesting hard copy output.

(iii)    FEMMOD provides an alternative method of mesh generation that includes more detailed control of nodal coordinates and element assemblies. Regular regions of a mesh can be generated by a simple command and subsequently 'edited', joined and copied with additions and deletions of single elements where required. Figures 5(a)-(c) show this method in use within LUISA1. Figure 11(b) shows some meshes that have been generated this way and filed as substructures. Figures 12(a)-(f) show how FEMMOD may be used to edit an existing mesh.

(iv)    Figures 11(a)-(i) show the use of COMPAS in assembling a set of component meshes, or substructures to form a representation of a structure that is too large to be held in core at one time. Examples are shown of a dam/ foundation (11(a)-(d)), a ship bulkhead (11(e)-(f)), a bridge diaphragm (11(g)-(i)) and a centre-line girder of a ship tank (17(a)-(b)). All the component meshes have first been generated with FEMMOD and then reduced to the form shown eg in Figure 11(e) to make them available to COMPAS for interactive assembly. Coding is in hand to permit meshes generated by PAMGEN to be passed to COMPAS.

(v)    LUISA-ELTYPE is at present being coded and will provide the means to change the fundamental geometric types of triangle and quadrilateral to element types such as linear strain triangle, axisymmetric quadrilateral with mid-side nodes etc and in addition will provide the means to assign material properties to appropriate regions of a mesh. The geometric and topological data structure is of a fundamental form that is not limited to structural problems. ELTYPE forms the interface with the physical system that is being represented. Another physical system would need an alternative ELTYPE (and LOADBC).

(vi)    Figures 13(a)-(d) show how one may control the "shape" of the stiffness matrix of a substructure by numbering the nodes with the lightpen and thereby reduce the required storage space and computation time. However such a technique on its own is of limited value and there is a need to combine it with those of Gill's SHELL system (Gill 1972) and some automatic band width optimization program.

(vii)    Applied loads and displacement boundary conditions may be specified with LOADBC which uses standard interactive identification procedures in common with other subsystems. Rapid selection of element nodes is achieved with a combined use of lightpen and function buttons, and numerical values may be keyed in where required with an immediate visual check always available. Figures 14(a)-(b) show the display of loads and boundary conditions for the

dam/foundation and the centre-line girder.   Loads and boundary conditions
may be specified at the reduced structure level or the element mesh (sub-
structure) level.


Coding is in hand to permit 2-D components like those used in the examples
of Figures 11(e), 11(a) and 17(a) to be assembled into a 3-D configuration
like that shown in Figure 15(a) which represents a structure like that
shown in Figure 15(b).   (These two figures (15(a) and (b)) are taken from
publications of Det Norske Veritas.)


Progress with fundamental techniques for generating and modifying a finite
element representation of a 3-D solid object is shown in Figures 16(a)-(b)
which show a short section of a spine-beam bridge that has been inter-
actively generated on the display screen composed of 3-D "bricks" from
which an appropriate finite element mesh could be generated.   This work
is the subject of a PhD research project, and is aimed at handling such
structures as those shown in Figure 18.


Transfer of data from the LUISA2 family to input files for ASAS has been
organised through simple interface routines.   To code interface routines
for other such analysis systems would be a relatively simple task.


## (3) An Output Assessment System

While many finite element analysis systems include some facilities for presenting
results in a graphical form on passive devices such as digital and micro-film plotters
relatively little use is made of the more active devices such as storage or refresh
crt display (Butlin 1972b;   International Ship Structures Congress 1973).   An
exception is the very imaginative work being done at Utah (Christiansen 1971) where
impressively realistic displays of structural displacements and stress patterns are
being produced with a special purpose display capable of rendering a full range of
tones.


The current work in Leicester is designed to especially exploit the interactive
facility in conjunction with a convential crt display.   Figures 17(a)-(ℓ) illustrate
some of the display options that may be interactively requested by the user to aid
him in assessing the behaviour of a structure and to prepare a record in hard copy
form of only the most relevant analysis results.

Incorporation of these facilities into the previously mentioned 3-D work is the subject of an MSc project.

(4) System Design

Features of the design of the LUISA1 and bridge systems of section (1) are described in Butlin and Hubbold (1969); Britton and Platts (1971); and are now only of historical interest.

Fortran is not the most appropriate language for coding these interactive graphics applications. Algol would be preferable but by no means as effective as AED or Algol-68. However the ogre of transferability has forced us to compromise and to establish a methodology for the production of large interactive graphics systems in Fortran. There follows some brief notes under the headings of graphics, interaction and data handling:

## Graphics

A package of routines called FRED (International Computers Ltd 1969) written in assembly code provides all the basic facilities in Fortran for picture organisation and display. An additional package (Butlin 1973b) written in Fortran aids the organisation of menus of options, warnings, error messages etc.

## Interaction

Control of interaction from Fortran is provided by other FRED routines. An "interaction processor" package of Fortran programs (Butlin 1973a; Butlin 1972a) provides a framework within which each subsystem is designed and its operation assessed prior to implementation and ensures compatibility in the display screen control language between the subsystems which are designed by different people. It includes means for:

(a) defining commands with data, using a simple numerical association between command names and their corresponding subroutines;

(b) storing a set of options so as to minimise the interaction necessary to communicate an instruction to the computer;

(c) introducing optional diagnostics;

(d)     modifying a large system to optimise its operation in the light of
        experience or to produce different versions suited to a novice, for
        example, instead of a skilled user.


Data Handling

While some of the LUISA2 family of subsystems have the same type of data structure
others have quite distinctive structures.   Each data structure has been designed to
suit the set of operations that need to be performed on it.   As data is read from
the disc data base into core or out to disc it is transformed to or from the
appropriate in-core data structure.   All subsystems use a free storage package which
manages the allocation and return of core storage.   Rather than garbage collection
its strategy includes an attempt to prevent the fragmentation of unused space.
Basic data structuring techniques are described in Butlin (1973c).


In OUTLIN, COMPAS and the 3-D system a paging package (Hubbold 1971;   Hubbold 1972)
is used to manage core/disc data transfers and has been designed to satisfy the
special requirements of on-line demands.


# DISCUSSION

Although applications of interactive graphics are now in regular production use in
some aeronautical and aerospace organisations the current benefits to be found are
still a long way short of what had been expected in the mid-sixties.


The most widespread productive applications to date in structural engineering have
employed the rather more passive devices (eg digital and micro-filmplotters) for
checking input data and for displaying results.   The extension to the more active
lightpen and refresh display has been restrained by the special programming problems,
lack of appropriate general purpose software, lack of uniformity of hardware and the
difficulty in assessing benefits.


However as a result of research in Leicester some software has been produced for
production structural analysis and in a form that makes possible its transfer to
other hardware.   The graphics packages mentioned in section (4) have
been converted to run on a DEC system 10 at Time Sharing Ltd and as a first trial the

subsystem LUISA-PAMGEN has been transferred to this computer to be operated from a Tektronix 4002A storage tube display shown in Figure 3. Support is now required for the commercial implementation of this and the other LUISA subsystems.

The problems of achieving the longer term aims of providing design engineers with aids to explore alternative solutions have been brought into focus. Some of these problems have already been solved and a large research programme is in hand to tackle the others over the next few years.

## ACKNOWLEDGEMENTS

## REFERENCES

BRITTON R A and PLATTS M J, A Bridge Design Program, Leicester
University Engineering Department Report 71-8 (1971)

BUTLIN G A, Report on a visit to the Massachusetts Institute of
Technology, Leicester University Engineering Department 68-14 (1968)

BUTLIN G A, Man-Machine Interactive Structural Analysis as a Preliminary Design
Aid, 29th Meeting of AGARD Structures and Materials Panel, Istanbul Symposium
on Optimisation Paper 37.2 (1969)

BUTLIN G A, An Interaction Processor Package, BCS Seminar on Language Processing
(to be published by the NCC) (1972a)

BUTLIN G A, Computer Graphics for Structural Analysis - A Brief Survey,
Leicester University Engineering Department (1972b)

BUTLIN G A, An Interaction Processor Package - User Manual, Leicester University
Engineering Department Report 73-3 (1973a)

BUTLIN G A, A Message Package - User Manual, Leicester University Engineering
Department Report 73-2 (1973b)

BUTLIN G A, 1st Tutorial Data Structures Conference, Cambridge (to be
published by the NCC) (1973c)

BUTLIN G A and HUBBOLD R J, A Scheme for Man-Machine Interactive Structural
Analysis, Inst Electrical Engineers Conference Publication, no 51, 615-625 (1969)

BUTLIN G A and LECKIE F A, Response Times in the Application of Interactive Graphics
to Structural Analysis, IUTAM Symposium on High Speed Computing of Elastic
Structures, Liege (1970)

CHRISTIANSEN H N, Displays of Kinematic and Elastic Systems, Proc Third Conf
Matrix Methods in Structural Mechanics, Wright-Patterson Airforce Base (1971)

GILL J I, Computer Aided Design of Shell Structures using the Finite Element Method,
PhD Thesis Cambridge University (1972)

HUBBOLD R J, A Scheme for Software Paging of List Data Structures, Symposium on
Computer Graphics, Gelleschaft fur Informatik, Technical University of Berlin (1971)

HUBBOLD R J, Software Paging of List Data Structures for Interactive Engineering Design,
IEE Conference on Computer-aided Design, Southampton (1972)

INTERNATIONAL COMPUTERS LTD, FRED (Fortran Routines for the Elliott Display)
4100 Manual (1969)

INTERNATIONAL SHIP STRUCTURES CONGRESS, Committee report no 9 (1973)

PLATTS M J and BRITTON R A, Computer Graphics in Creative Design Work, Int Symposium
on Computer-aided Structural Design, University of Warwick (1972)

ROSS D T , The AED Approach to Generalised Computer-aided Design,
    ACM National Meeting (1967)

TOCHER J L, FELIPPA, Computer Graphics Applied to Production Structural Analysis,
    IUTAM Conference on High Speed Computing of Elastic Structures, Liege (1970)
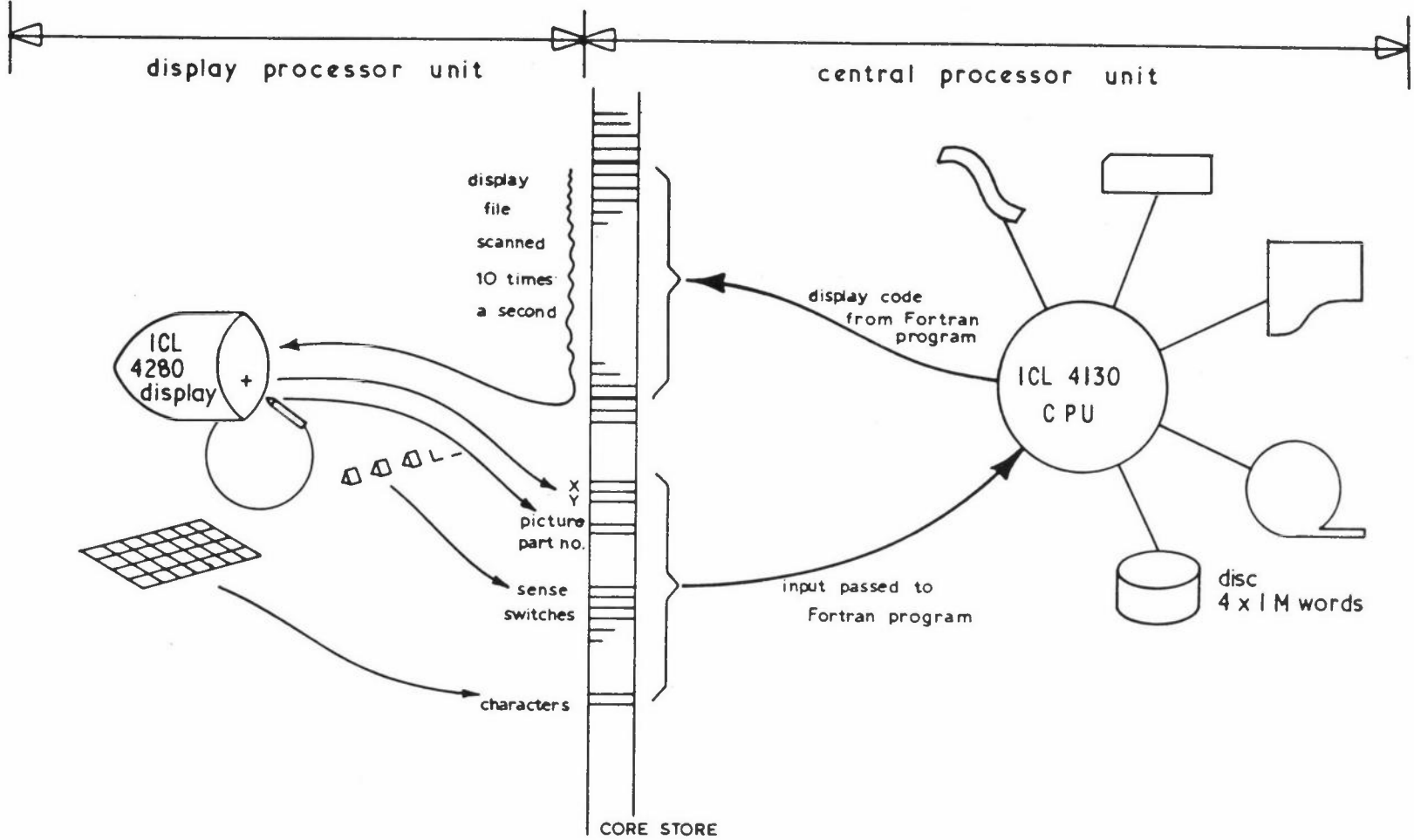
**Figure 1:** Diagrammatic representation of ICL 4130/4280 hardware and fundamental software

Figure 2: ICL 4280 refresh display terminal with
lightpen (and foot pedal), function buttons
and keyboard



Figure 3: Tektronix 4002A storage tube display with
joystick, keyboard and hard copy unit

| PRODUCTION USAGE OF TYPICAL PROGRAMS | PRODUCTION EFFORT | | TYPICAL DEVELOPMENT ACTIVITIES | DEVELOPMENT EFFORT | |
|---|---|---|---|---|---|
| | MANPOWER % | MACHINE TIME % | | MANPOWER % | MACHINE TIME % |
| ●DATA PREPARATION<br>●DATA CHECKING | 50 | 40 | ●DATA GENERATORS<br>●GRAPHIC DISPLAYS | 10 | 10 |
| ●SOLUTION PHASE | 10 | 50 | ●NEW FINITE ELEMENTS<br>●BETTER EQUATION SOLVERS<br>●ADDITIONAL FEATURES<br>●NONLINEAR CAPABILITIES | 80 | 80 |
| ●INTERPRETATION OF RESULTS | 40 | 10 | ●GRAPHIC DISPLAYS | 10 | 10 |

Figure 4: (From Tocher, 1970) Typical work breakdown comparison between production use of finite element programs and development effort on new analysis programs
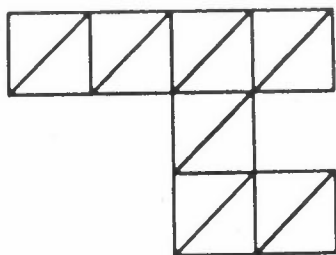
(a) Basic topology generated by copying and joining
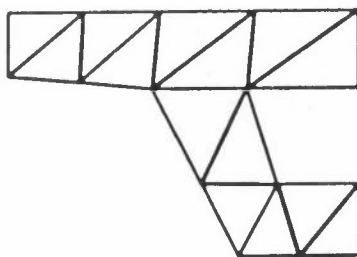
(b) Distorted to fit required dimensions
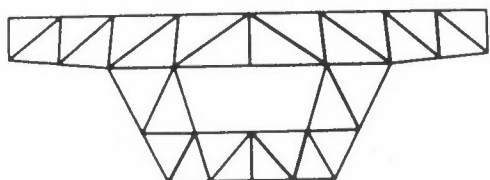
(c) Copied and joined again to produce required mesh
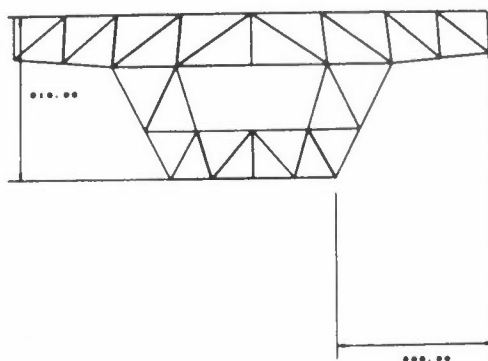
(d) Dimensions checked
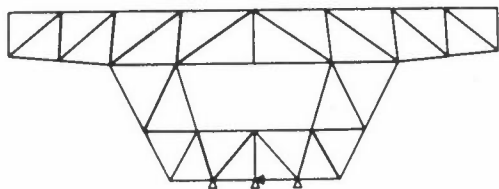
(e) Displacement constraints specified
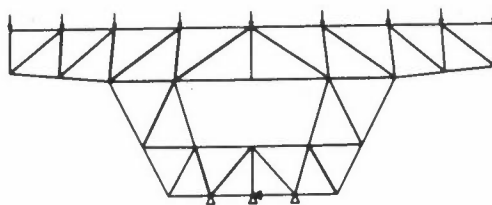
(f) Loads applied

Figure 5 (part 1): LUISA1
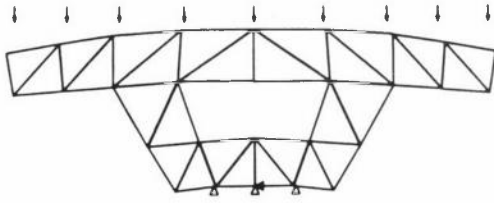
(g) Displaced shape
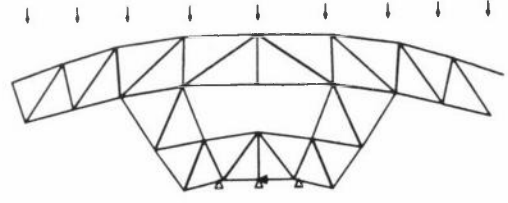
(h) With increased scale on display of displacements
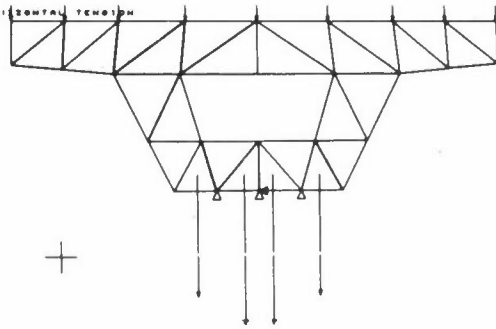
(i) Sigma x displayed for four elements from datum at tracking cross
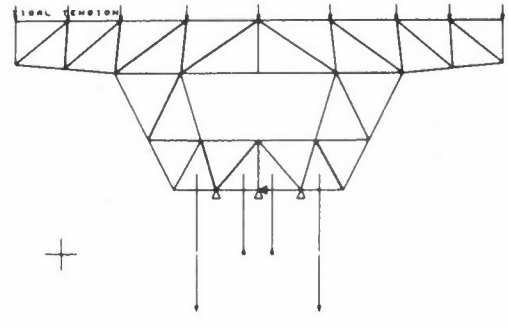
(j) Sigma y
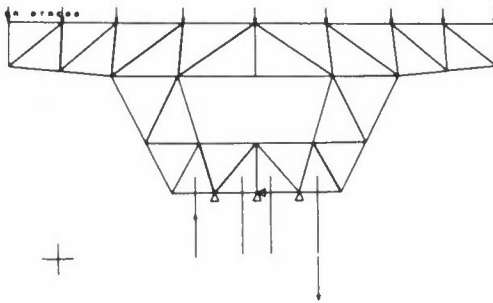
(k) Sigma xy
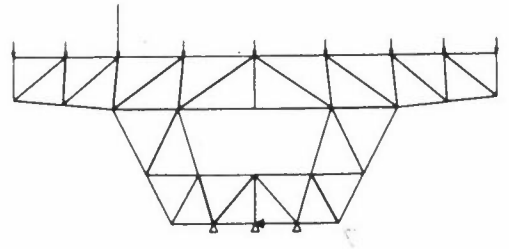
(l) Second load case

Figure 5 (part 2): LUISA1
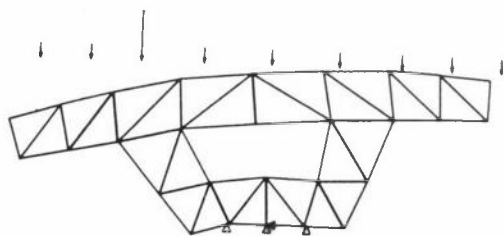
(m) Displacements
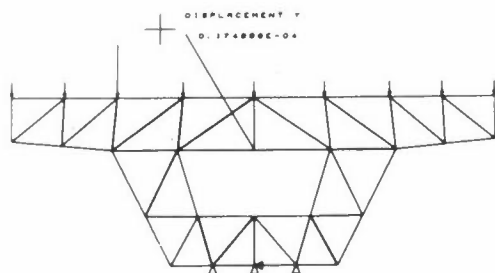
(n) Numerical value of displacement for node chosen with lightpen
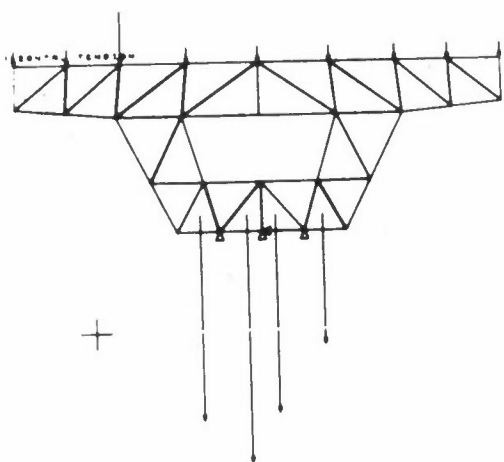
(o) Sigma x
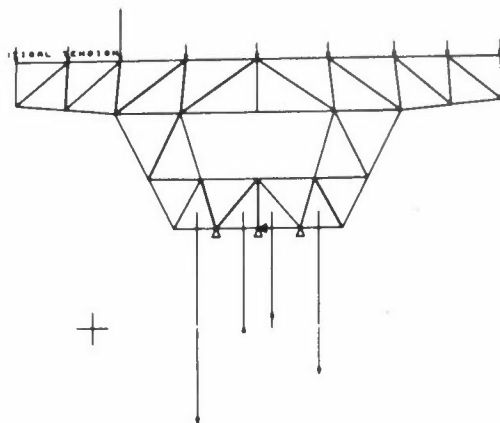
(p) Sigma y
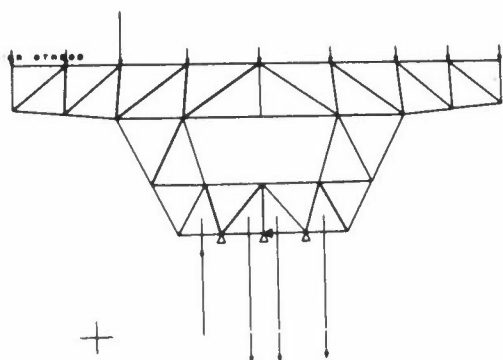
(q) Sigma xy
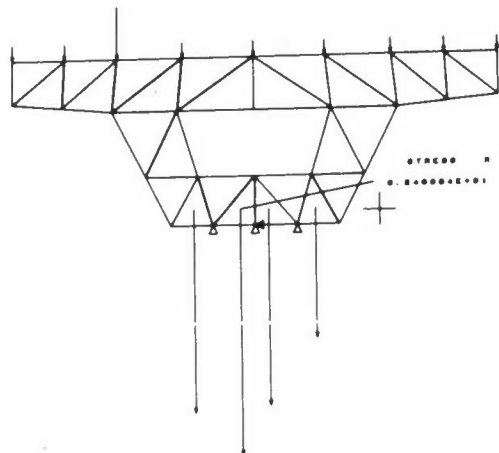
(r) Numerical value of sigma x for an element chosen with lightpen

Figure 5 (part 3): LUISA1
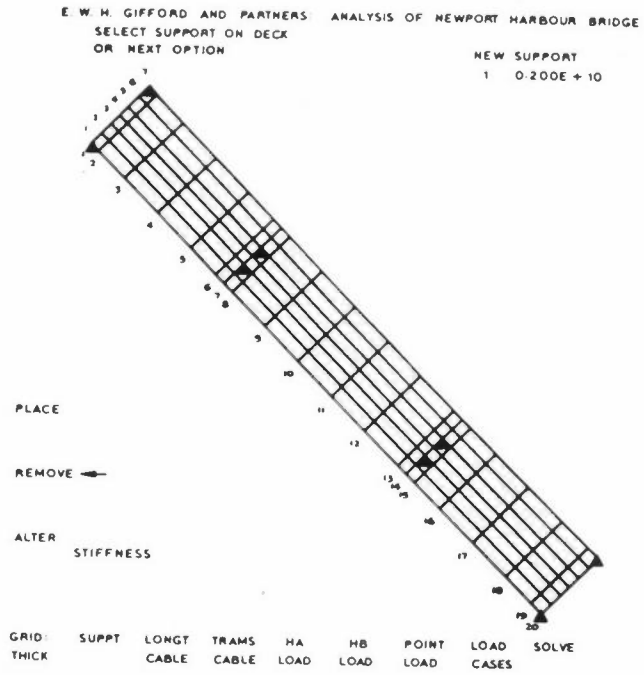
Figure 6(a): Typical display screen layout — showing setting
of supports
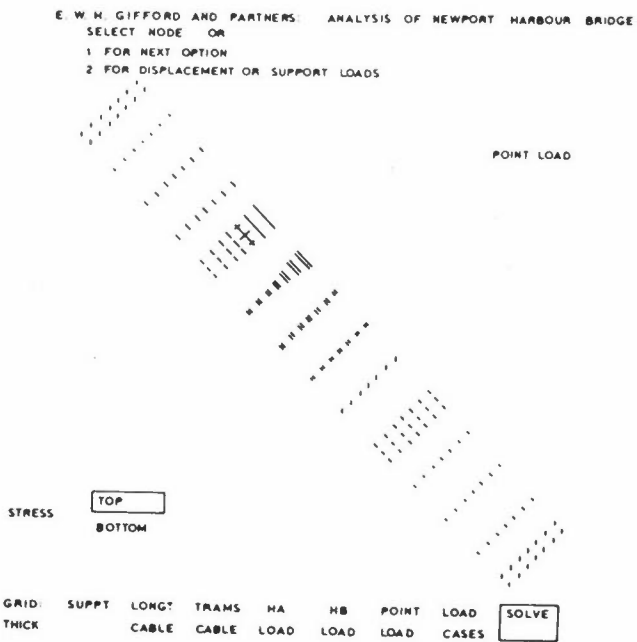


Figure 6(b): Typical display screen layout — showing display
of stress for point load
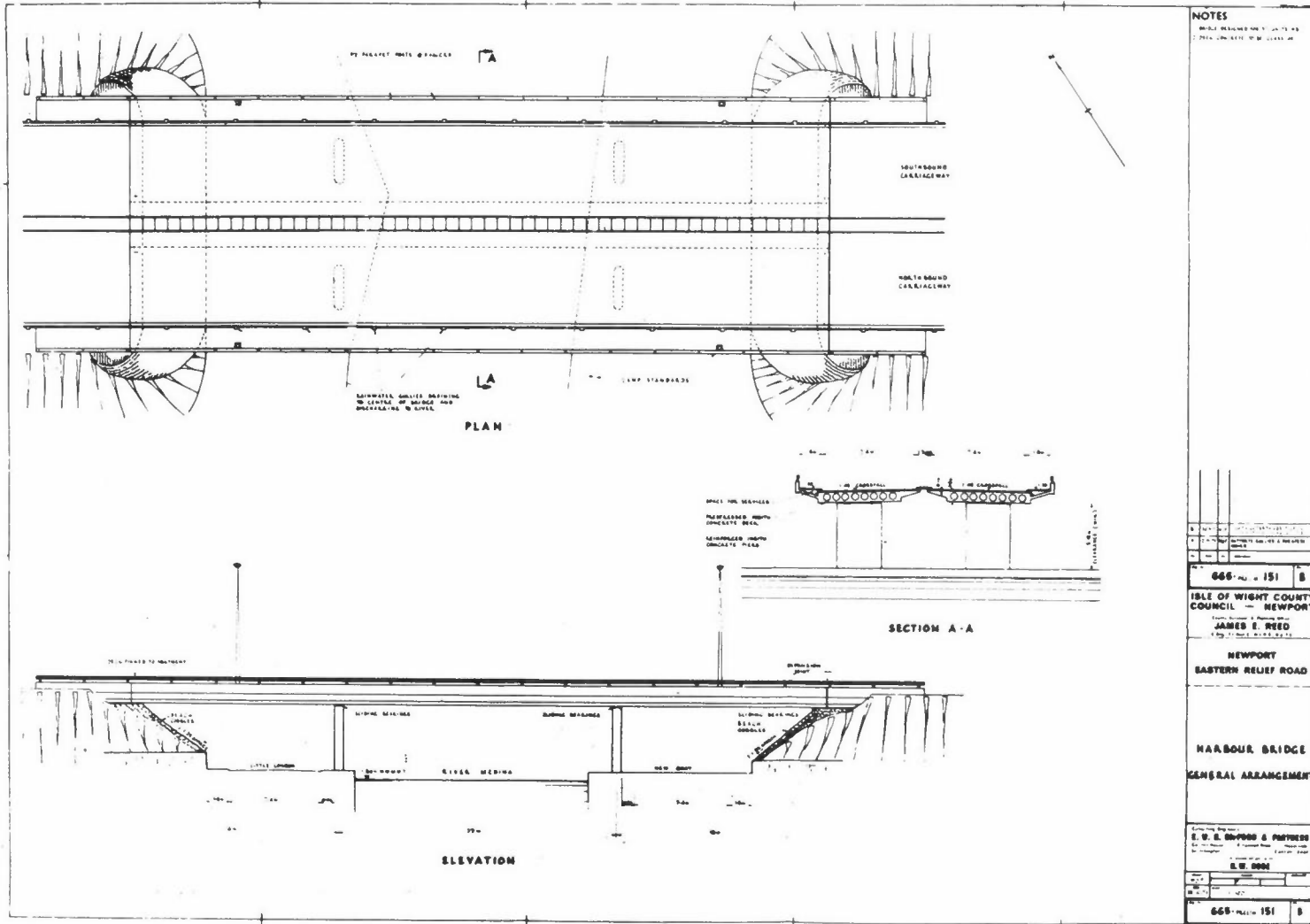
Figure 7(a): Newport Harbour Bridge – general arrangement
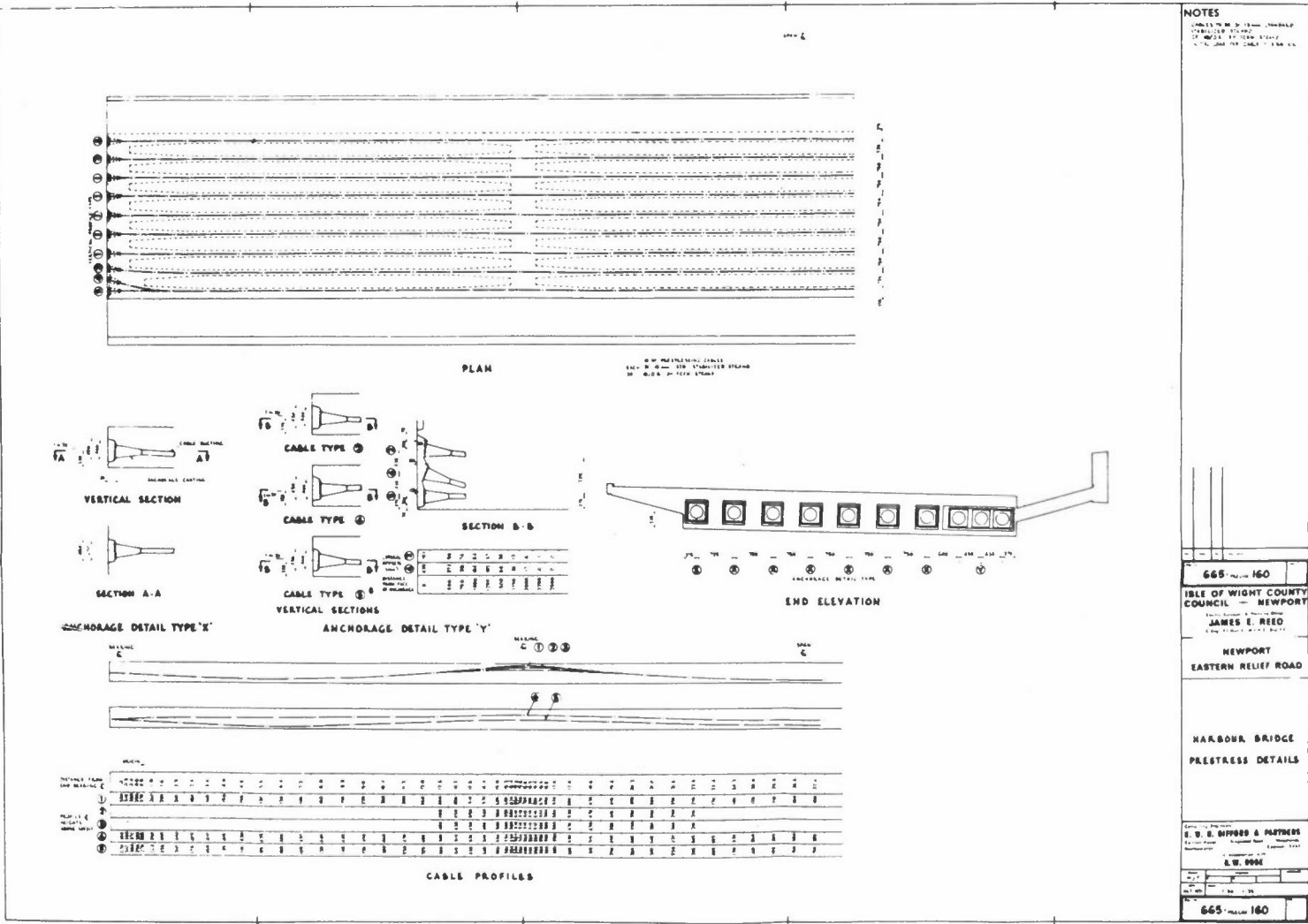
Figure 7(b): Newport Harbour Bridge – prestress details
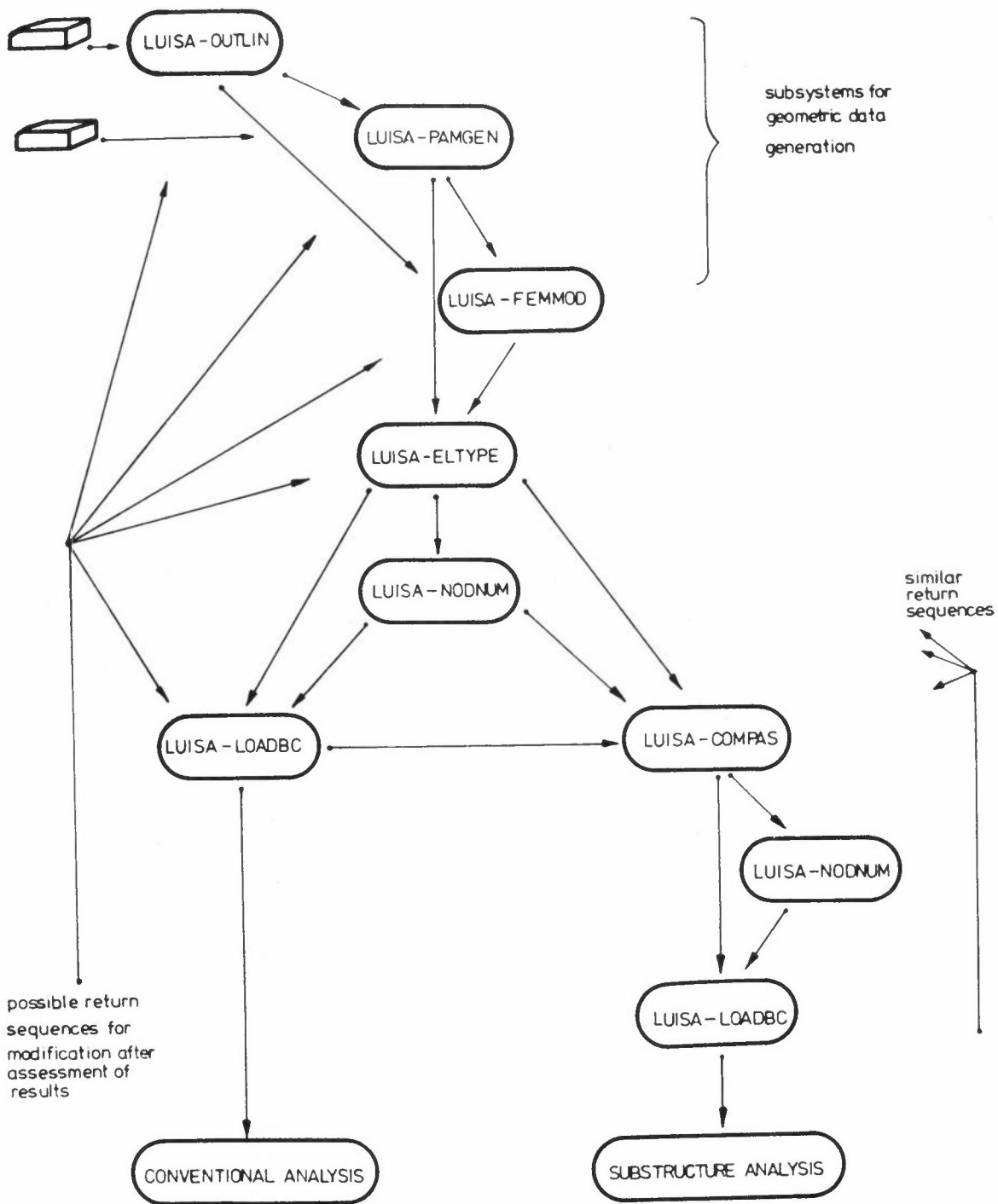
Figure 8: User sequences for LUISA subsystems
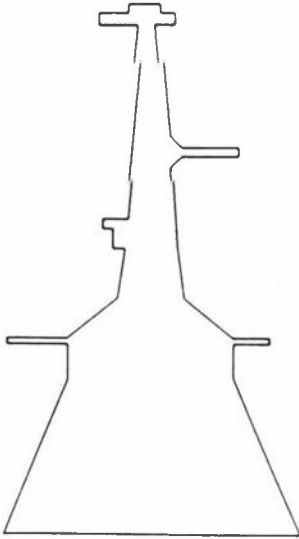
(a) Simplified outline shape of turbine disc input from cards
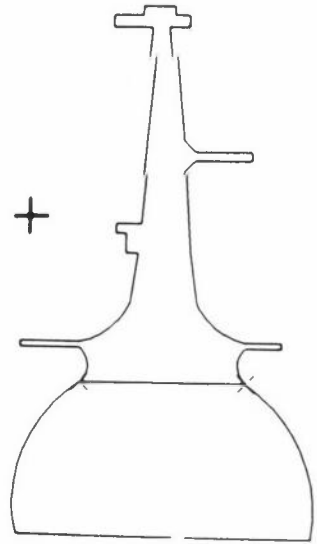
(b) Outline shape adjusted and first line drawn to divide into plane patches
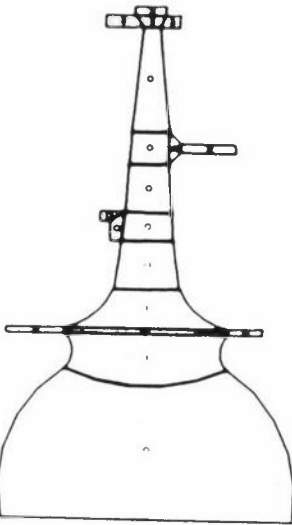
(c) Division into patches complete and mesh types chosen (0, 1 etc)
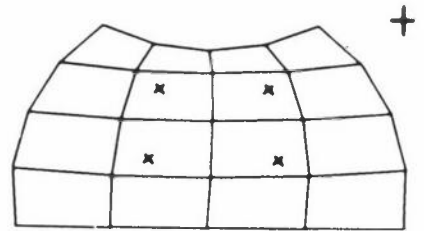
(d) Single patch selected, scaled up and internal parametric control points (the four crosses) being used to adjust mesh
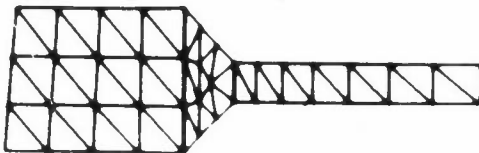
(e) Small group of patches selected for adjustment of grading
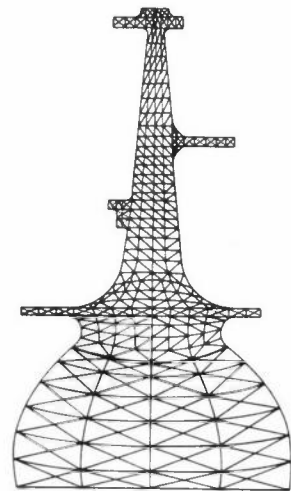
(f) Full mesh showing different mesh types and consequent grading

Figure 9: Parametric mesh generation
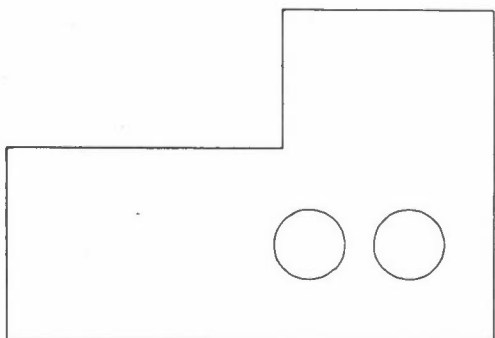
(a) Outline of underground railway tunnels input from cards
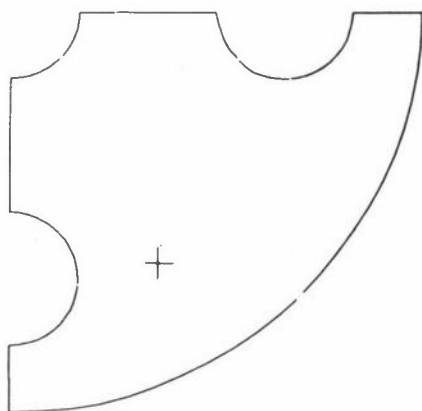
(d) Outline of quarter of disc input from cards
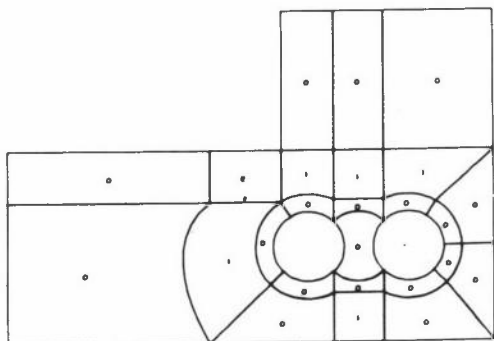
(b) Subdivided interactively into patches and mesh types selected
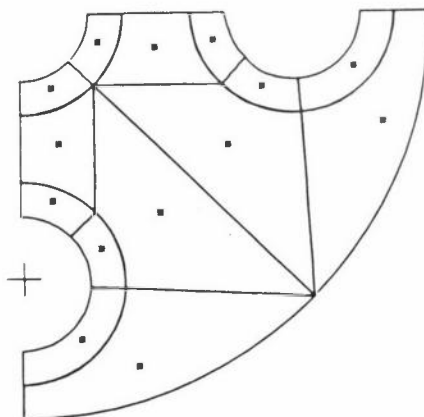
(e) Subdivided interactively into patches
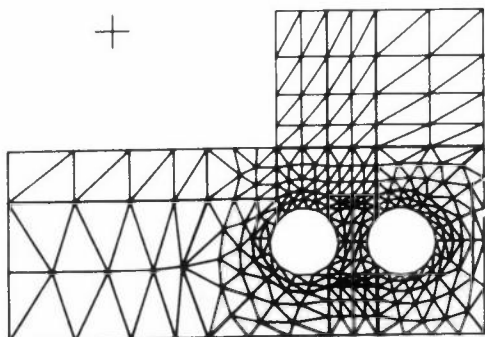
(c) Resulting mesh; total elapsed time for stages 1-3 = c 1 hour (computer time is a fraction of this)
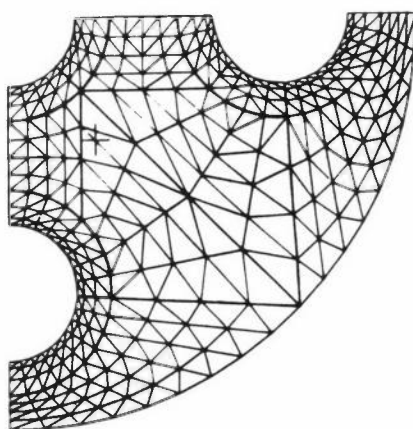
(f) Resulting mesh; total elapsed time = c 45 mins (computer time is a fraction of this)

Figure 10: Example meshes generated with OUTLIN and PAMGEN subsystems of LUISA

(a) Library of substructures
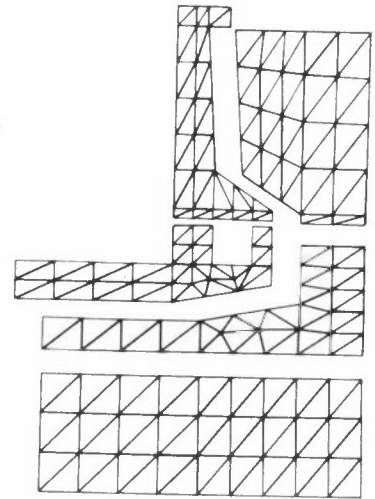


(b) Meshes of some substructures in library
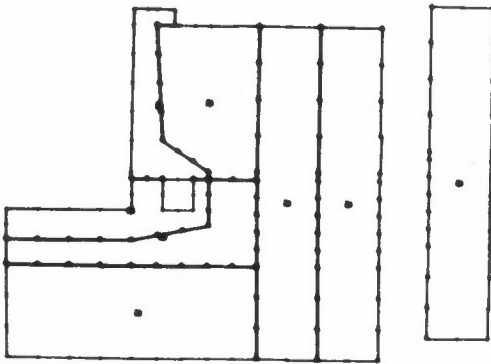


(c) Substructures being assembled (in reduced form) into dam/foundation example
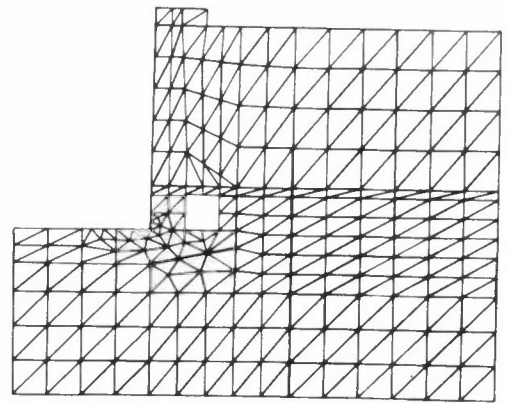


(d) Full mesh of dam/foundation; total elapsed time taken = c 2 hours (computer time is a fraction of this)
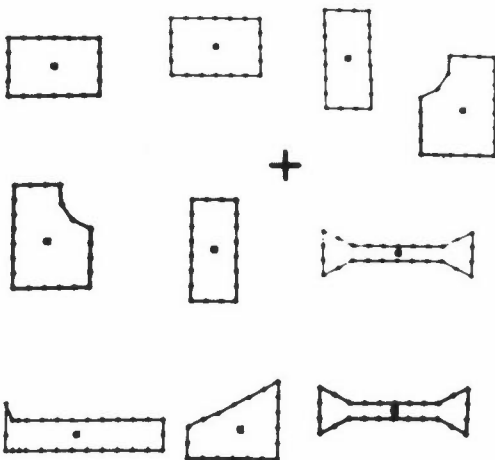


(e) A selection of substructures shown in reduced form in another library
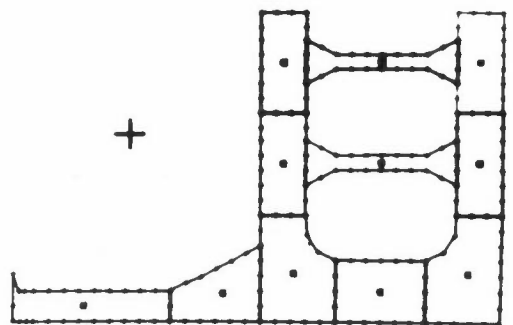


(f) assembled to represent part of a ship bulkhead

Figure 11 (part 1):  Substructure assembly
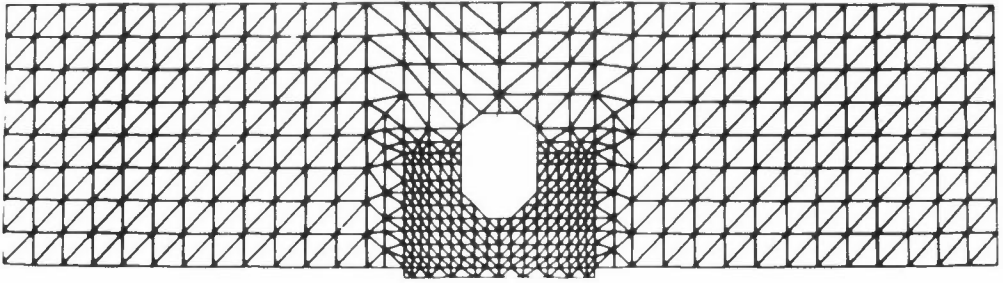
(g) Meshes of substructures for a bridge diaphragm
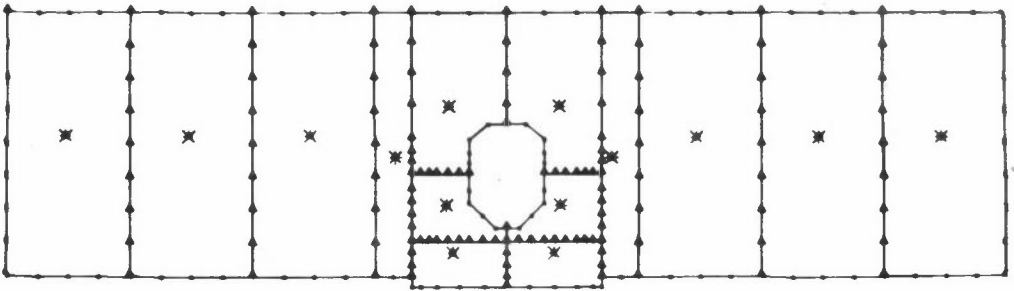


(h) Assembled substructures in reduced form
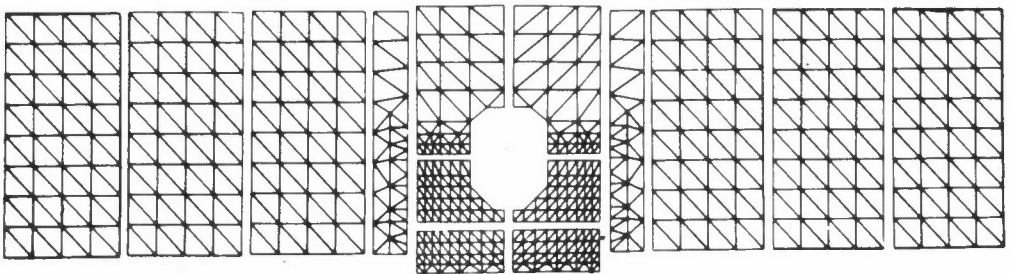


(i) Resulting mesh

Figure 11 (part 2): Substructure assembly
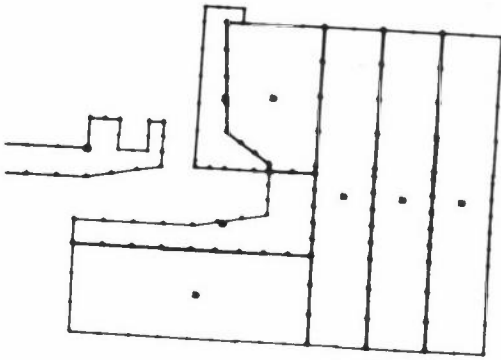
(a) Substructure extracted for modification
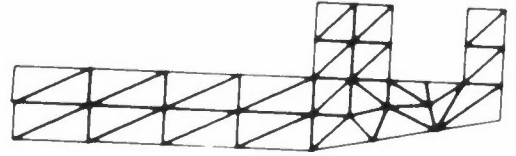
(b) Mesh retrieved from library file
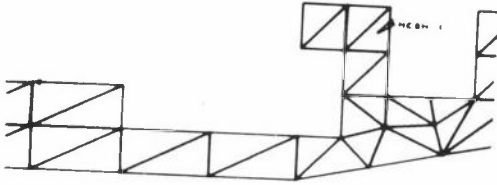
(c) Expanded view of area to be modified showing elements deleted
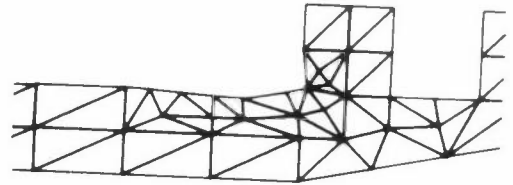
(d) New, finer mesh sketched in
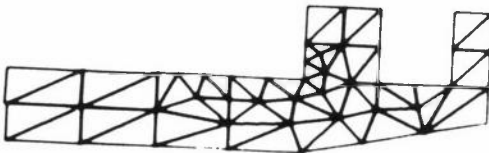
(e) Boundary nodes lined up and internal nodes adjusted
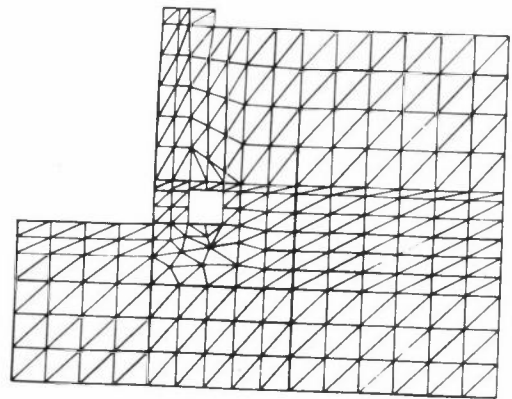
(f) Full modified mesh after rejoining modified substructure

Figure 12: Mesh modification

B1 NUMBER ALL NODES; B2 SOME; B3 ZOOM;       PLOT

B4 CENTRE ZOOM; B5 VIEW MATRIX; B6 EXIT,
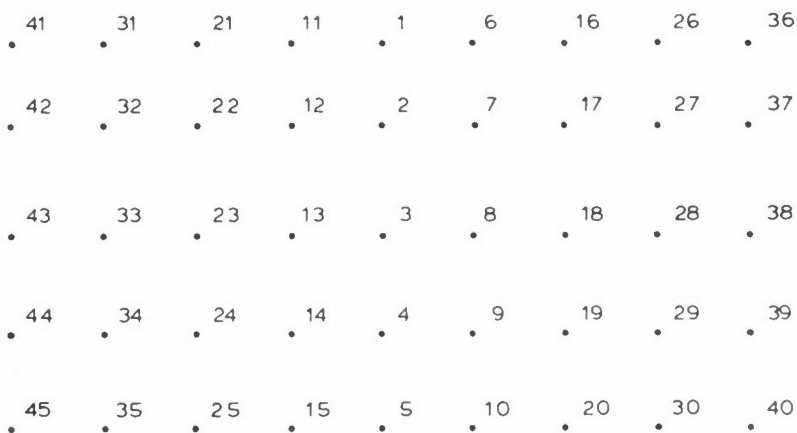


Figure 13(a): Internal nodes (from 1 to 35) numbered sequentially starting
with the left-most column of internal nodes

OUT 1 TO SAVE NUMBERING REPRESENTED IN MATRIX       PLOT
2 TO LOOK AT PREVIOUSLY SAVED NUMBERING
3 TO LOOK AT CURRENT NUMBERING
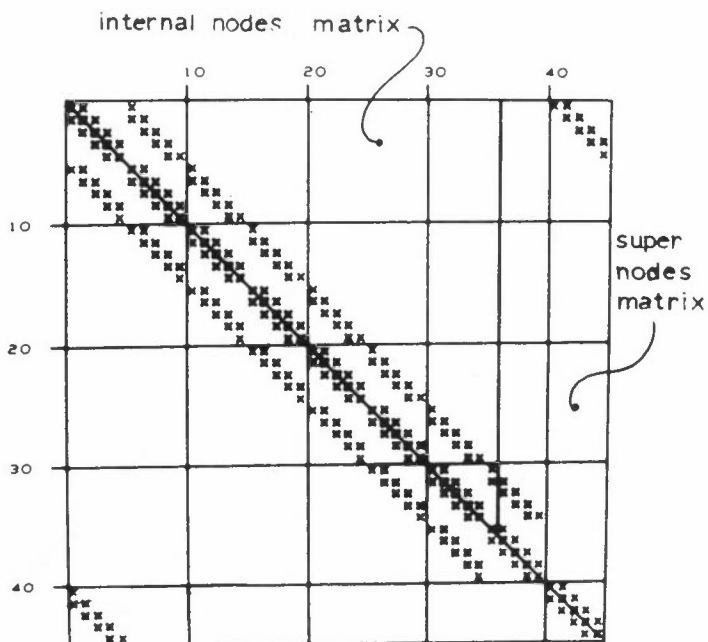


Figure 13(b): The corresponding matrix pattern showing a well banded
internal matrix (Band width = 12 nodes) but an ill shaped
super nodes matrix with non zeros at top right

B1  NUMBER  ALL  NODES;   B2  SOME;   B3  ZOOM;          PLOT

B4  CENTRE  ZOOM;   B5  VIEW  MATRIX;   B6  EXIT;

```
 41      1       6      11      16      21      26      31      36
 •       •       •       •       •       •       •       •       •

 42      2       7      12      17      22      27      32      37
 •       •       •       •       •       •       •       •       •

 43      3       8      13      18      23      28      33      38
 •       •       •       •       •       •       •       •       •

 44      4       9      14      19      24      29      34      39
 •       •       •       •       •       •       •       •       •

 45      5      10      15      20      25      30      35      40
 •       •       •       •       •       •       •       •       •
```
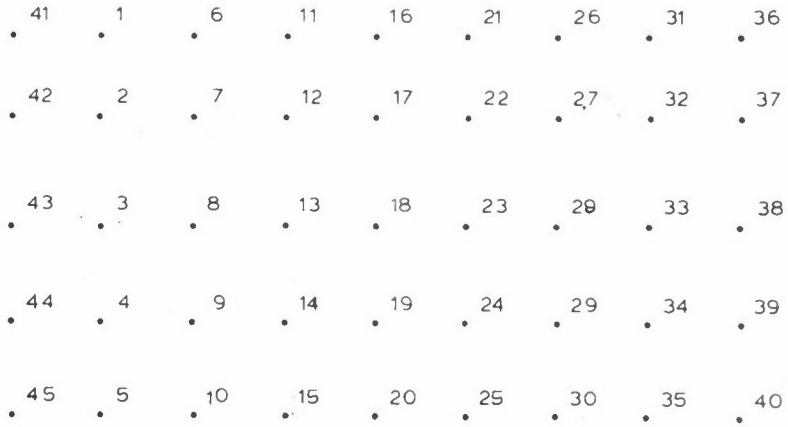
Figure 13(c):  Internal nodes numbered alternately starting with the middle
               column of nodes

OUT 1 TO SAVE NUMBERING REPRESENTED IN MATRIX
    2 TO LOOK AT PREVIOUSLY SAVED NUMBERING                          PLOT
    3 TO LOOK AT CURRENT NUMBERING



Figure 13(d):  The corresponding matrix pattern with a larger band width
               for the internal nodes matrix while the super nodes matrix
               has no non-zeros at top right
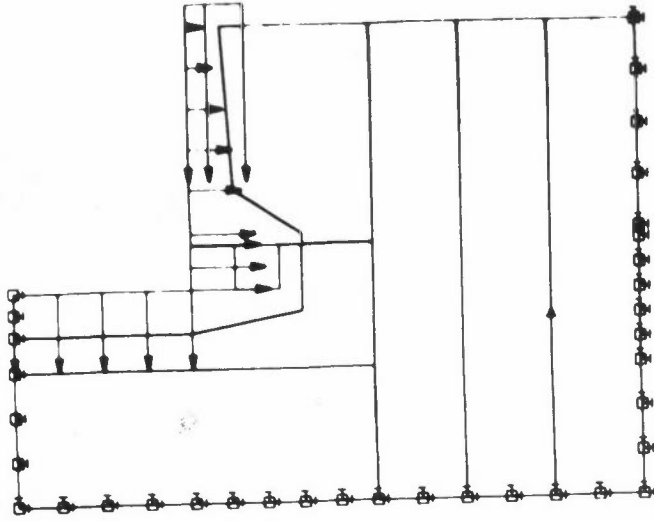
(a) Loads and boundary conditions of dam/foundation
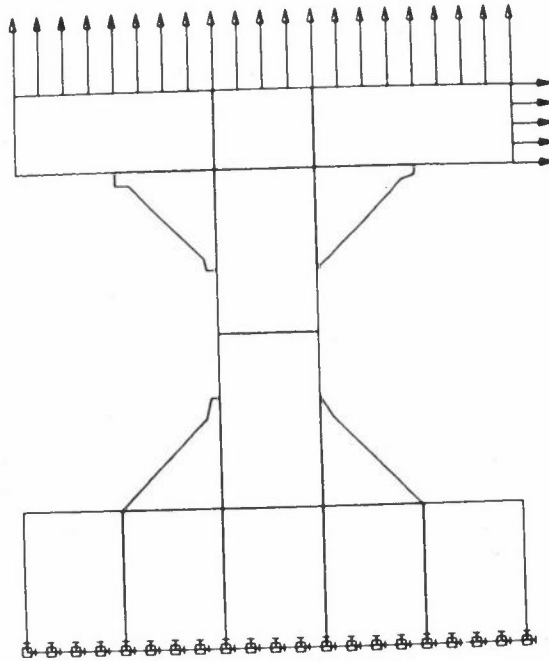


(b) Loads and boundary conditions of ship girder

Figure 14: Loads and boundary conditions
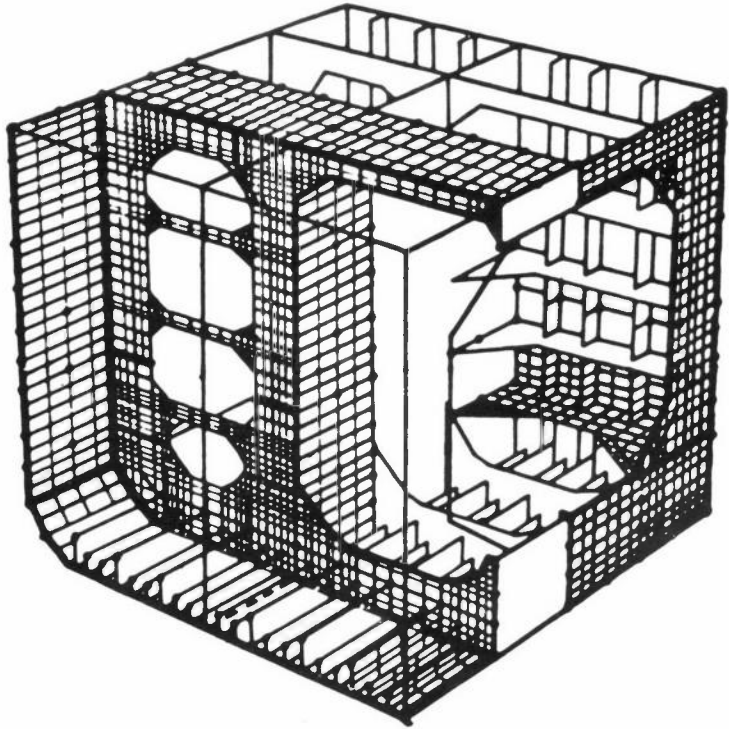
Figure 15(a): (reproduced from a SESAM paper) Diagrammatic view of finite element idealisation of part of the structure of a tanker
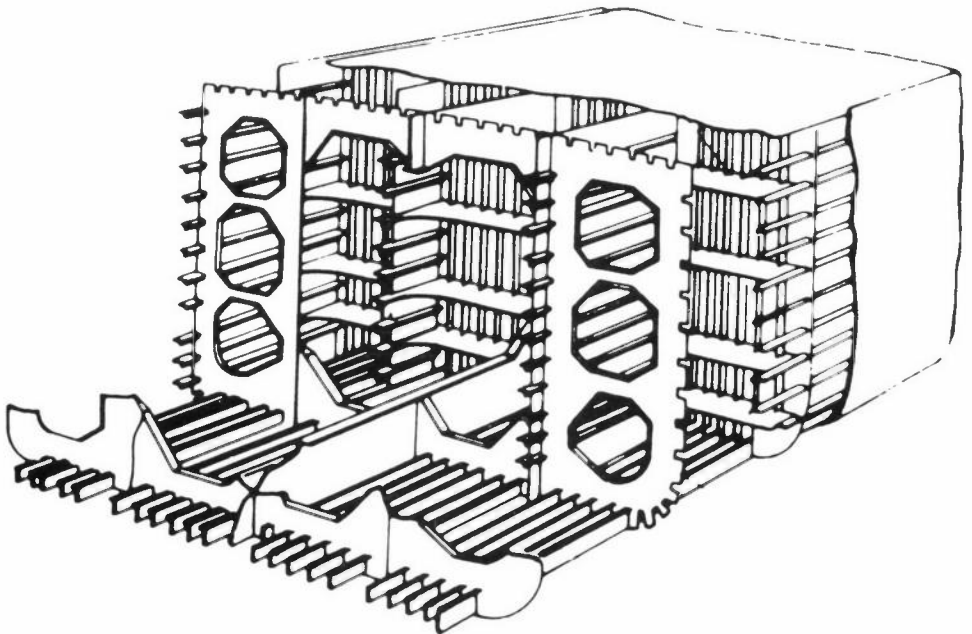


Figure 15(b): (reproduced from "Optimisation and Automated Design of Structures", edited by Johannes Moe and Kaare M Gisvold, Trondheim 1971) Typical internal structure of a tanker
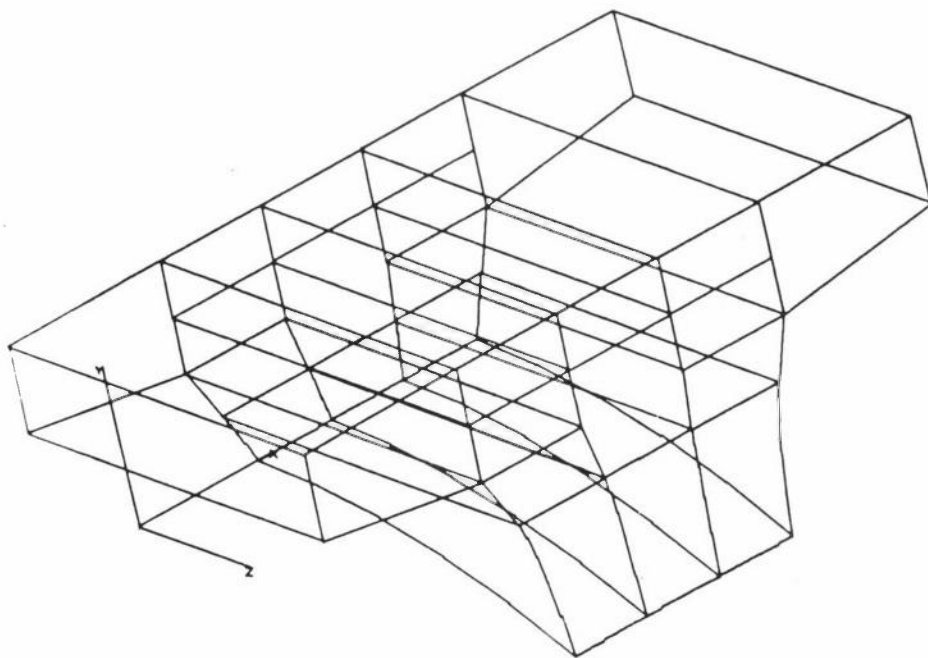
Figure 16(a): Section of spine beam bridge composed of 3-D "bricks"
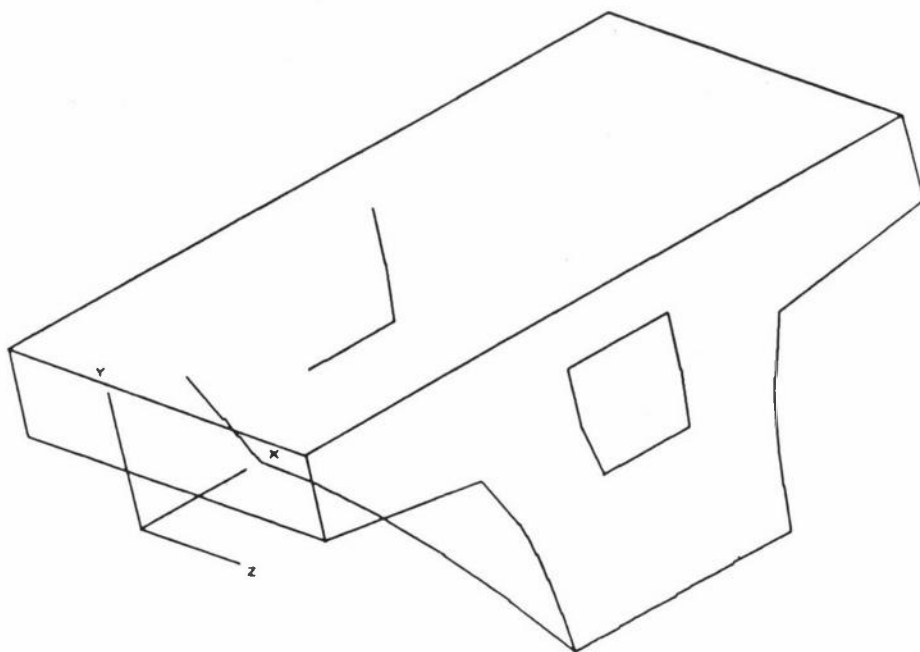


Figure 16(b): Same set of "bricks" but with some lines removed to improve visual
appreciation
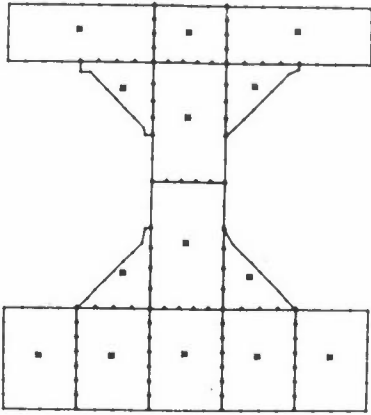
(a) Assembled substructures of a centre line girder of a ship
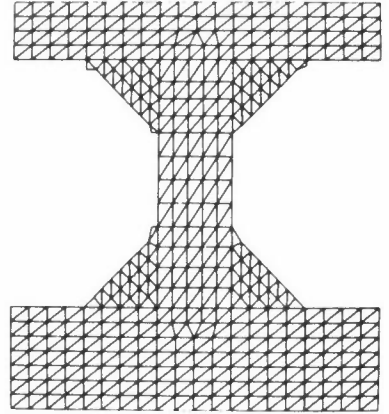
(b) Corresponding full finite element mesh
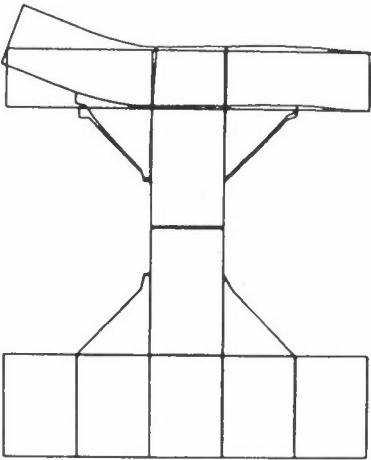
(c) Overall displaced shape under loads and boundary conditions shown in figure 14(b)
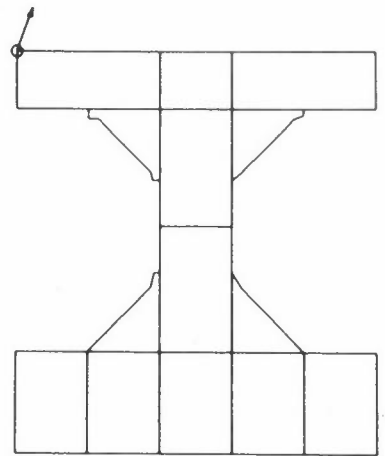
(d) Display of numerical values of a displacement following identification of a node
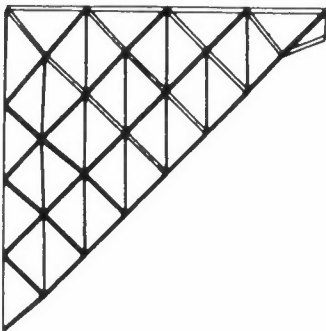
(e) Displaced shape of a single component (that has been selected with the lightpen)
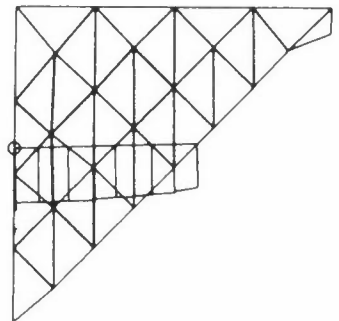
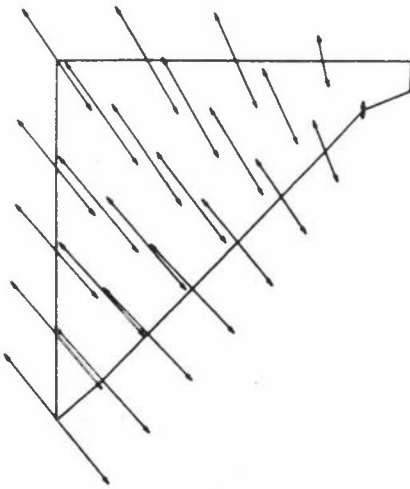(f) Stress profile along a section chosen with the lightpen
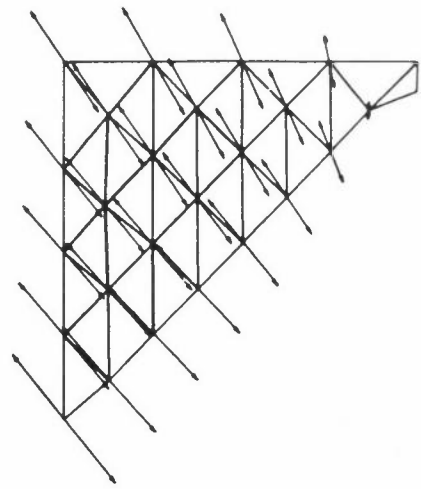
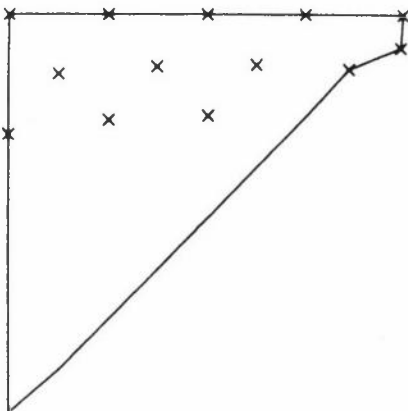Figure 17 (part 1): Output assessment
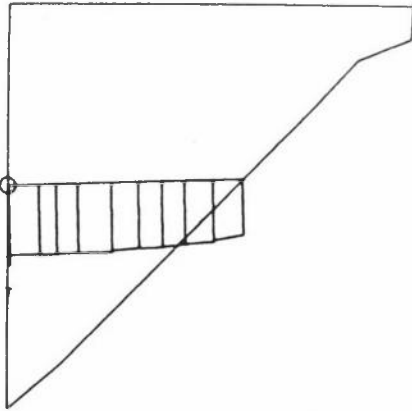
(g) As (f) but without mesh
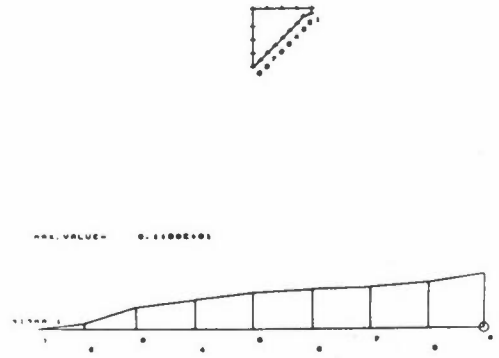
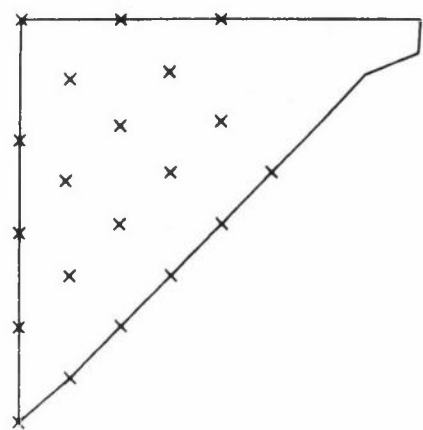(h) Stress profile along boundary with correlation numbers

(i) Principal stresses at each node

(j) As (i) but with mesh

(k) Nodes marked at which compression stress exceeds a specified critical value

(l) As (k) but for tension

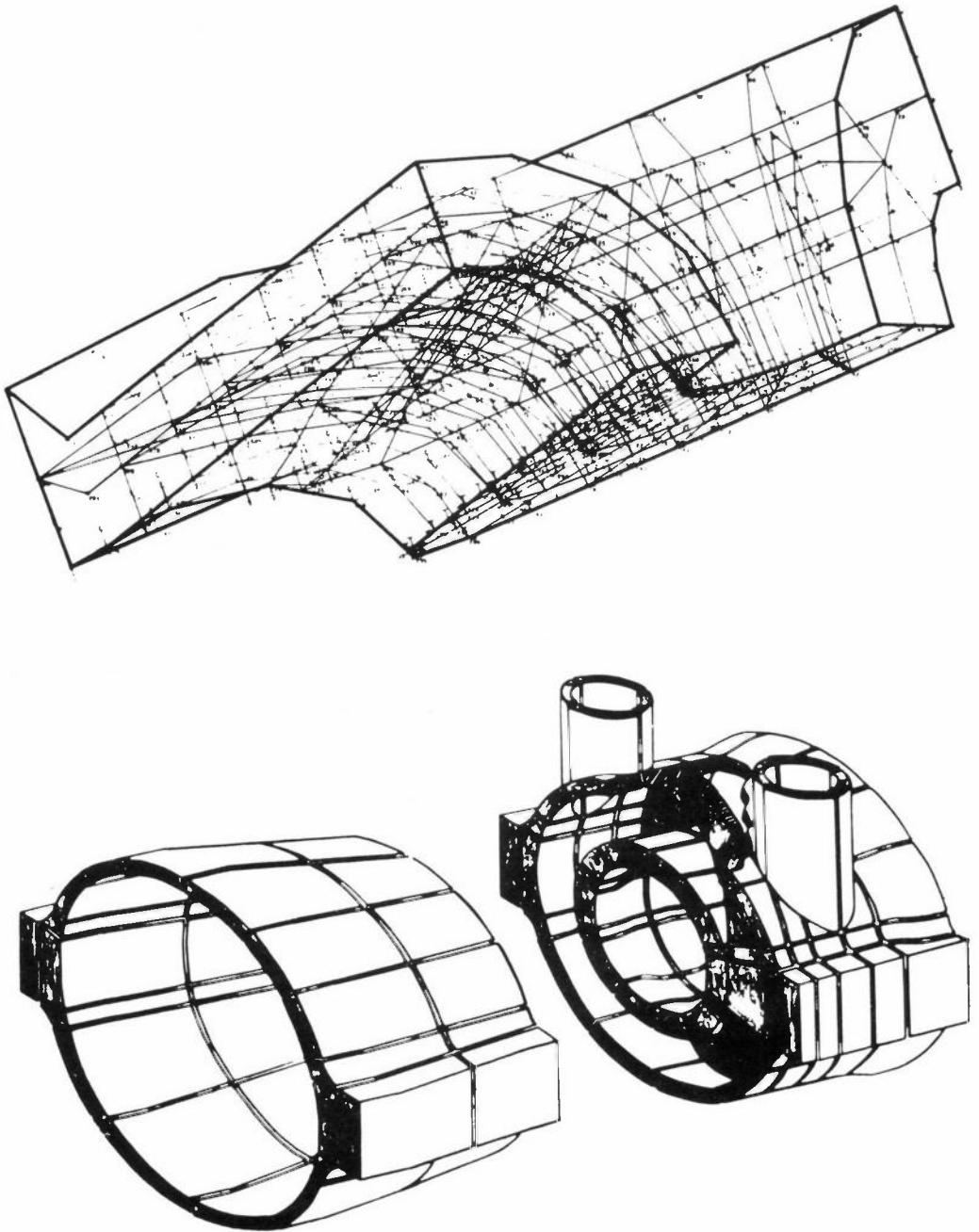Figure 17 (part 2): Output assessment

**Figure 18:** Typical finite element assemblies that will be more rapidly generated and checked with the aid of interactive computer graphics

Top: Basic element mesh in highly stressed region of 45° — pipe intersection (reproduced from "Application of computerised methods in analysis and design of ship structures, marine structures and machinery," Det Norske Veritas, Oslo 1972)

Bottom: Finite element representation of a marine steam turbine (by courtesy of GEC Power Engineering Ltd)

# LIST OF PARTICIPANTS

MR R A ADEY

Department of Civil Engineering
The University
Southampton
SO9 4NH

DR C A BREBBIA

Department of Civil Engineering
The University
Southampton
SO9 4NH

DR G A BUTLIN

Department of Engineering
The University
Leicester
LE1 7RH

MR S CHANDA

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 OQY

MS J E CROW

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 OQY

MR M J P CULLEN

Meteorological Office
London Road
Bracknell
Berkshire

MR P DIMMER

Department of Engineering
The University
Leicester
LE1 7RH

MR N J DISERENS

Rutherford Laboratory
Chilton
Didcot
Oxfordshire

DR D M DUGGAN

Department of Engineering Science
The University
Oxford

PROFESSOR S F EDWARDS

Science Research Council
State House
High Holborn
London WC1

MR A J EGGINTON

Engineering Division
Science Research Council
State House
High Holborn
London WC1

MR F B ELLERBY

Department of Mathematics
The University
Claverton Road
Bath
BA2 2A7

MR J P ELLINGTON

Central Technical Services
UKAEA
Risley
Warrington
Lancashire

DR J ENDERBY

Central Technical Services
UKAEA
Risley
Warrington
Lancashire

MR E B FOSSEY

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 0QY

DR L R T GARDNER

School of Mathematics and Computer
 Science
University College of North Wales
Bangor
LL57 2UW

MR J E HAILSTONE

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 0QY

DR D R HAYHURST

Department of Engineering
The University
Leicester
LE1 7RH

DR R D HENSHELL

Department of Mechanical Engineering
The University
Nottingham
NG7 2RD

PROFESSOR F R A HOPGOOD

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 0QY

DR J HOWLETT

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 0QY

DR R HURST

Wallsend Research Station
British Ship Research Association
Wallsend
Northumberland
NE28 6UY

MR B M IRONS

Department of Engineering
Faculty of Engineering
The University of Calgary
Calgary
Alberta
Canada

MR J L KIRTON

Department of Civil and Structural
 Engineering
UMIST
PO Box 88
Sackville Street
Manchester
M60 1QD

DR J W E LEWIS

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 OQY

DR R W LEWIS

Department of Civil Engineering
University of Wales
Singleton Park
Swansea
SA2 8PP

DR E LIGHTFOOT

Department of Engineering Science
The University
Oxford

PROFESSOR G D S MACLELLAN

Department of Engineering
The University
Leicester
LE1 7RH

PROFESSOR K W MORTON

Department of Mathematics
The University
Whiteknights Park
Reading
Berkshire

MR M NEWMAN

Rutherford Laboratory
Chilton
Didcot
Oxfordshire

DR D A C PARKES

Department of Mechanical Engineering
The University
Nottingham
NG7 2RD

DR J K REID

Theoretical Physics Division
Atomic Energy Research Establishment
Harwell
Oxfordshire

MR C L ROBERTS

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 OQY

PROFESSOR H H ROSENBROCK

Control Systems Centre
UMIST
PO Box 88
Manchester
M60 1QD

DR V R SAUNDERS

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 OQY

MR A SCHOLES

Research and Development Division
British Railways Board
London Road
Derby
DE2 8UP

MR W D SHAW

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 OQY

MR A SPENCE

School of Mathematics
University of Bath
Claverton Down
Bath
BA2 7AY

MR J B SPOONER

Atkins Research and Development
Ashley Road
Epsom
Surrey

MISS B STOKOE

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 OQY

DR A WEXLER

Department of Electrical Engineering
University of Manitoba
Winnipeg
Canada
R3T 2N2

MR R C G WILLIAMS

Atlas Computer Laboratory
Chilton
Didcot
Oxfordshire
OX11 OQY

DR W L WOOD

Department of Mathematics
The University
Whiteknights
Reading
Berkshire