INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS, VOL. 1, 17–43 (1981)

THE CAUSE AND CURE (?) OF THE SPURIOUS PRESSURES GENERATED BY CERTAIN FEM SOLUTIONS OF THE INCOMPRESSIBLE NAVIER–STOKES EQUATIONS: PART 1

R. L. SANI

CIRES/NOAA, University of Colorado, Boulder, Colorado, U.S.A.

P. M. GRESHO AND R. L. LEE

Lawrence Livermore Laboratory, Livermore, California, U.S.A.

D. F. GRIFFITHS

University of Dundee, Dundee, Scotland

SUMMARY

The spurious pressures and ostensibly acceptable velocities which sometimes result from certain FEM approximate solutions of the incompressible Navier–Stokes equations are explained in detail. The concept of pressure modes, physical and spurious, pure and impure, is introduced and their effects on discretized solutions is analysed, in the context of both mixed interpolation and penalty approaches. Pressure filtering schemes, which are capable of recovering useful pressures from otherwise polluted numerical results, are developed for two particular elements in two-dimensions and one element in three-dimensions. Implications regarding the effect of spurious pressure modes on accuracy and ultimate convergence with mesh refinement are discussed and a list of unanswered questions presented. Sufficient numerical examples are discussed to corroborate the theory presented herein.

KEY WORDS Finite Element Navier–Stokes Incompressible Flows Penalty Methods Pressure Filters

CONTENTS

0271–2091/81/010017–28$01.00
© U.S. Government
Received August 1980
Revised December 1980

I. INTRODUCTION

Discretized approximations to the incompressible Navier–Stokes equations, in the primitive variable (velocity–pressure) formulation, especially when generated via the Galerkin Finite Element Method (GFEM), have been plagued with confusion regarding the 'appropriate' workable combinations of velocity and pressure approximations. Early on it was discovered by Hood and Taylor[1,2] that equal-order interpolation on conforming quadrilateral elements, wherein the same basis functions are used for representing velocity and pressure, causes difficulty in the pressure solution. They obtained better results when using mixed interpolation (the basis functions for pressure were one order lower than those for velocity), and suggested an explanation in the form of 'balancing residuals' from momentum and continuity equations. Their explanation, although intuitively appealing, was judged inadequate by Olson and Tuann[3] who explained the results in terms of the eigenvalues of a single element (equal interpolation always generated one or more eigenvectors which contained only pressure and corresponded to zero eigenvalues; they claimed that these were spurious and were the cause of the failure). While a combination of the above two explanations appears to have satisfied most researchers (at least in practice, since essentially everyone (?) now effectively uses mixed interpolation), there are lingering doubts in the minds of some (e.g. Richards[4]) that the complete theoretical explanation is contained in these analyses. As shown herein, these doubts are well justified, even though the work of Olson and Tuann was more or less in the right direction.

Even when mixed interpolation is employed, however, there are cases where numerical difficulties are encountered. In particular, the simplest possible approximation employs piecewise linear approximation for velocity (bilinear on quadrilaterals) and piecewise constant approximation for pressure. This element has been found to work well in some cases and poorly in others; in solid mechanics, see Argyris *et al.*[5] and Nagtegaal *et al.*,[6] and in fluid mechanics, see Fabayo,[7] Huyakorn *et al.*,[8] Hughes *et al.*,[9] and Lee *et al.*[10] For certain combinations of boundary conditions and element distributions over a domain, the solutions display acceptable velocities but totally spurious pressures; the pressures are afflicted with the 'checkerboard syndrome,' wherein they display oscillations which are frequently of one sign on all 'black' squares and of the opposite sign on all 'red' squares. These pressure patterns have also been encountered using certain finite difference discretization techniques (Fortin,[11] Pracht and Brackbill[12]) so that the affliction is not intrinsic to FEM formulations; for example, Chorin[13] has encountered four spurious pressure patterns when solving a consistently centred difference approximation to the Navier–Stokes equations in two-dimensions (eight in 3-D). Finally, similar behaviour can also occur using higher-order elements; e.g. the quadrilateral element with biquadratic velocity (9-node) and bilinear pressure with nodes at the 2×2 Gauss points can display a checkerboard syndrome.

This spurious pressure problem is the principal item addressed in this paper, which is a detailed expansion of the summary version presented earlier;[14] in particular, we have extensively studied this 'problem' on 2-D grids composed of certain quadrilateral finite elements. The results and techniques we shall present also apply in 3-D (at least for the simplest elements, which exhibit many pressure modes) and for related 2-D and 3-D finite difference discretizations (Caldwell[15]), as well as for the related incompressible computations in solid mechanics. The techniques developed herein also shed additional light on the question of why equal interpolation fails.

In the remainder of this paper, we will define and characterize these 'zero energy pressure modes,' both theoretically and numerically, and, for two particular elements, present simple

and effective methods for extracting good physical pressures from the polluted numerical results. Additionally, we will point out some problems associated with particular element distributions and boundary conditions which may bear on the theoretical proof of convergence for certain elements.

II. THE CONTINUUM EQUATIONS AND THEIR FEM DISCRETIZATION

The setting for our discussion will be the steady Stokes equations (or equivalently, the equations of incompressible isotropic linear elasticity):

$$\nabla \cdot \boldsymbol{\tau} = \mathbf{0} \tag{1a}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{1b}$$

where $\tau_{ij} = -p\delta_{ij} + S_{ij}$ is the symmetric stress tensor,

$$S_{ij} = \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{1c}$$

$\mathbf{u} = (u, v)$ is the velocity, p is the pressure, and μ is the (constant) viscosity of the fluid. Since the pressure modes arise from the discretization of ∇p and $\nabla \cdot \mathbf{u}$, these simplest equations suffice and the results immediately generalize to unsteady, nonlinear (viscous or inviscid), and nonisothermal (Boussinesq) flows; and our numerical results substantiate this.

In the related penalty method,[9,16] the solenoidal constraint on the velocity field is slightly weakened, to $p = -\lambda \nabla \cdot \mathbf{u}$, where λ is the penalty parameter. The results from the penalty approximation will be close to those using (1b) if the following inequalities are satisfied: $e \ll \mu/\lambda \ll 1$, where e is the 'unit roundoff level' of the computer; for example, on our CDC–7600, $e \approx 10^{-14}$ and the penalty method 'works well' for $\sim 10^5 < \lambda/\mu < \sim 10^9$. With perfect arithmetic ($e = 0$), the penalty results would converge to those from (1) as $\lambda \to \infty$. For large but finite λ, the results should typically differ by $O(1/\lambda)$; actually, we usually observe 6 digits of agreement for $10^6 \leq \lambda \leq 10^8$. Since the pressures obtained from the penalty method can also be plagued with the checkerboard syndrome, it is appropriate to consider it also.

The GFEM discretized approximation is applied to the following weak form of (1), with some initial notational laxity, which we quickly rectify in (4):

$$\int_\Omega \boldsymbol{\tau} \cdot \nabla \phi_i = \int_{\partial\Omega} \phi_i \boldsymbol{\tau} \cdot \mathbf{n} \tag{2a}$$

$$-\int_\Omega \psi_j \nabla \cdot \mathbf{u} = 0 \tag{2b}$$

over the domain Ω with boundary $\partial\Omega$, where $\boldsymbol{\tau} \cdot \mathbf{n}$ is the surface traction, ϕ_i represents the 'appropriate' basis function for velocity, ψ_j is any basis function for pressure (the minus sign is required to retain symmetry), and $\mathbf{n}$ is the outward pointing unit normal vector. The following approximations are now applied, using the appropriate piecewise polynomial basis functions:

$$\mathbf{u}^h = \sum_{k=1}^{N_0} \mathbf{u}_k \phi_k(\mathbf{x}), \tag{3a}$$

$$p^h = \sum_{k=1}^{M} P_k \psi_k(\mathbf{x}), \tag{3b}$$

where there are N_0 velocity nodes and M pressure 'nodes' in the discretized approximation. Inserting (3) into (2) leads to the GFEM equations, written in compact form as:

$$\mathbf{KU}+\mathbf{CP}=\mathbf{f} \tag{4a}$$

$$\mathbf{C}^{\mathrm{T}}\mathbf{U}=\mathbf{g} \tag{4b}$$

where $\mathbf{U}=(u_1, u_2, \ldots, u_{N_1}, v_1, v_2, \ldots, v_{N_2})^{\mathrm{T}}$, $\mathbf{P}=(P_1, P_2, \ldots, P_M)^{\mathrm{T}}$.

This symmetric algebraic system represents a condensed system in which N_1 and N_2 are, respectively, the number of 'free' u and v degrees of freedom. Consequently, the vectors $\mathbf{f}$ and $\mathbf{g}$ reflect the effect of imposed boundary conditions; $\mathbf{f}$ could represent prescribed boundary forces (*not* pressures[17]) or prescribed velocities but $\mathbf{g}$ corresponds only to prescribed velocities. Also,

$$K_{ij}=\mu\int_\Omega \left[\begin{array}{c|c} \left(2\dfrac{\partial\phi_i}{\partial x}\dfrac{\partial\phi_j}{\partial x}+\dfrac{\partial\phi_i}{\partial y}\dfrac{\partial\phi_j}{\partial y}\right) & \dfrac{\partial\phi_i}{\partial y}\dfrac{\partial\phi_j}{\partial x} \\ \hline \dfrac{\partial\phi_i}{\partial x}\dfrac{\partial\phi_j}{\partial y} & \left(\dfrac{\partial\phi_i}{\partial x}\dfrac{\partial\phi_j}{\partial x}+2\dfrac{\partial\phi_i}{\partial y}\dfrac{\partial\phi_j}{\partial y}\right) \end{array}\right] \tag{5}$$

is a positive-definite, symmetric matrix,

$$C_{ij}=-\int_\Omega \left[\begin{array}{c} \dfrac{\partial\phi_i}{\partial x}\psi_j \\ \hline \dfrac{\partial\phi_i}{\partial y}\psi_j \end{array}\right], \tag{6}$$

and the contribution, if any, to $\mathbf{f}$ from prescribed surface forces is

$$f_i=\int_{\partial\Omega} \left[\begin{array}{c} \phi_i\left[2n_x\dfrac{\partial u^h}{\partial x}+n_y\left(\dfrac{\partial u^h}{\partial y}+\dfrac{\partial v^h}{\partial x}\right)-p^h n_x\right] \\ \hline \phi_i\left[n_x\left(\dfrac{\partial u^h}{\partial y}+\dfrac{\partial v^h}{\partial x}\right)+2n_y\dfrac{\partial v^h}{\partial y}-p^h n_y\right] \end{array}\right], \tag{7}$$

where n_x, n_y are the x- and y-components of $\mathbf{n}$.

In the penalty approximation, the discretized form of the continuity equation (4b) becomes

$$\lambda\mathbf{C}^{\mathrm{T}}\mathbf{U}-\mathbf{MP}=\lambda\mathbf{g}, \tag{4c}$$

where $M_{ij}=\int_\Omega \psi_i\psi_j$. Combining (4a) and (4c) leads to the computational form of the penalty approximation,

$$[\mathbf{K}+\lambda(\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^{\mathrm{T}})]\mathbf{U}=\mathbf{f}+\lambda\mathbf{M}^{-1}\mathbf{g}, \tag{4d}$$

which is a symmetric linear system in velocity only (pressure can be recovered via (4c) applied as a post-processor). The consistent penalty matrix, $\mathbf{B}=\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^{\mathrm{T}}$, (see Appendix I) as well as the λ-portion of the right hand side, is simple to form only when discontinuous (C^{-1}) pressure approximation is employed, in which case it can be formed element-wise. Under certain conditions, a further economy accrues when employing the Lagrange family of approximations for velocity and (C^{-1}) pressure in that $\mathbf{B}$ can be efficiently formed via reduced quadrature methods, using the 'continuum' approach ($p=-\lambda\nabla\cdot\mathbf{u}$ is inserted into (1a) and (2a) and the equivalence theorem of Malkus and Hughes,[18] in which the $\{\psi_i\}$ basis need never be explicitly introduced. The use of reduced quadrature, however, is generally not recommended in the case of higher-order isoparametric elements with *curved* sides (e.g. biquadratic velocity) since the accuracy of the penalty method (relative to mixed interpolation) is then only attained via actual construction of the $\mathbf{B}$ matrix.[19] A final comment on

reduced quadrature: if full quadrature is employed on the penalty term, the result is equivalent to using equal-order interpolation for velocity and pressure; this form of equal-order interpolation (using C^0 velocity and C^{-1} pressure) is particularly untenable since it generates more continuity constraints than there are velocities to satisfy them.

III. THE BILINEAR VELOCITY–PIECEWISE CONSTANT PRESSURE ELEMENT

Specializing now to the bilinear element, where ϕ_i is piecewise bilinear and ψ_i is piecewise constant (unity on element i and zero on all other elements), we write the element matrix C_{ij}^e for the general isoparametric element (K_{ij} needn't be written in detail, since the pressure modes are associated solely with the rectangular and indefinite **C** matrix) using the notation shown in Figure 1 and the bilinear isoparametric mapping, e.g.

$$\int_\Omega \psi_i \frac{\partial \phi_j}{\partial x}\,dx\,dy = \int_{A_i} \frac{\partial \phi_j}{\partial x}\,dx\,dy = \int_{-1}^{1}\int_{-1}^{1}\left[\frac{\partial \phi_j^e}{\partial \xi}\frac{\partial \xi}{\partial x}+\frac{\partial \phi_j^e}{\partial \eta}\frac{\partial \eta}{\partial x}\right]J\,d\xi\,d\eta$$
$$= \int_{-1}^{1}\int_{-1}^{1}\left[\frac{\partial \phi_j^e}{\partial \xi}\frac{\partial y^e}{\partial \eta}-\frac{\partial \phi_j^e}{\partial \eta}\frac{\partial y^e}{\partial \xi}\right]d\xi\,d\eta$$

where $y^e = \sum_{i=1}^4 y_i \phi_i^e$ etc., A_i is the area of element i, and ϕ_i^e is the local (element level) basis function, to give

$$C_{ji}^e = \tfrac{1}{2}[y_4 - y_2,\ y_1 - y_3,\ y_2 - y_4,\ y_3 - y_1 \quad x_2 - x_4,\ x_3 - x_1,\ x_4 - x_2,\ x_1 - x_3] \tag{8}$$

from which the global equations corresponding to **CP** and $\mathbf{C}^T\mathbf{U}$ can be formed in the standard way by adding the appropriate element level contributions. Note that for the bilinear element, each global continuity equation is identical to the continuity equation applied over a single element, in the form $\mathbf{C}_m^T\mathbf{U}^m = \mathbf{0}$ for element m, $m = 1, 2, \ldots, M$, where $\mathbf{U}^m = (u_1^m, u_2^m, u_3^m, u_4^m, v_1^m, v_2^m, v_3^m, v_4^m)^T$. Each continuity equation describes an element level mass balance and is a consequence of the discontinuous pressure approximation; by contrast, continuous (C^0) pressure approximation leads to an overall mass balance but can never ensure element level mass balances.

This element, which is also referred to as the 4-node element, is considered in detail because it leads to a simple, but intriguing and important exposition of the occurrence and ramifications of pressure modes. In succeeding sections, we will generalize these notions to other elements and point out the impact of pressure modes—both on numerical simulations and theory.

The discretized continuity equations (constraint equations among the velocity variables on each element), in conjunction with the imposed boundary conditions, can lead to what are termed 'pressure modes'. (Note that the pressure is implicitly defined by the continuity

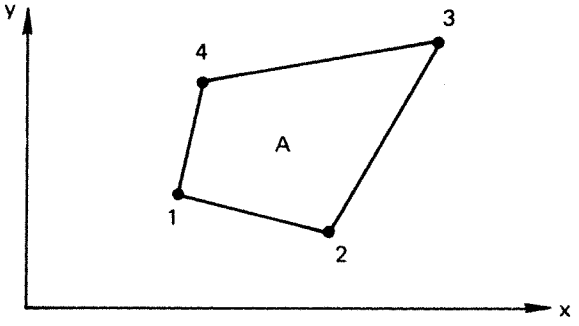

Figure 1. General bilinear element

constraints; there is one pressure per continuity equation.) These pressure modes (for *any* element) correspond to special solutions of (4a,b) with $\mathbf{f}=\mathbf{0}$ and $\mathbf{g}=\mathbf{0}$, for which $\mathbf{U}=\mathbf{0}$ and $\mathbf{CP}=\mathbf{0}$ for *nontrivial* $\mathbf{P}$. They may be spurious in that they do not correspond to actual (physical) pressures and, in a numerical calculation, they are imbedded in the numerical results in such a way that the physical pressures, which are still present, are often 'buried in the noise level.' It turns out that there are two pressure modes for this element, one of which is physical and the other spurious. From the viewpoint of linear algebra, these pressure modes correspond to solutions in the null space of the linear operator (matrix) in that the corresponding eigenvalues are zero.

In the mixed-interpolation (or Lagrange multiplier) method (4a,b) the concomitant matrix singularity causes these pressure modes to be present in arbitrary amounts. In the penalty method, however (4c,d), while (related) pressure modes can still occur, their magnitudes are not arbitrary; the amplitudes are automatically prescribed and are, in some sense, minimized (this will be explained in due course).

The discussion of the pressure modes will be aided by first considering the simple rectangular grid shown in Figure 2, in which a small number of elements is shown to simplify the presentation.

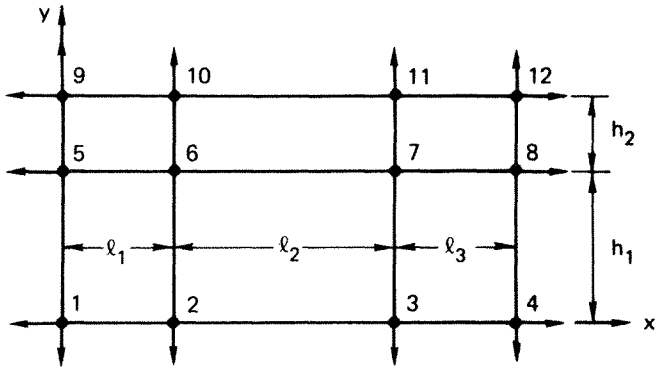

Figure 2. A simple mesh of quadrilateral elements

A. *The hydrostatic pressure mode*

As a prelude to the discussion of the more subtle checkerboard mode (the spurious mode; hereafter abbreviated as the CB mode), we first elucidate the simpler and much more common mode, the (physical) hydrostatic pressure mode. This pressure mode is more general in that it can occur for any type of velocity and pressure approximation and for any type of domain subdivision. Its existence depends solely on the form of the boundary conditions imposed on the domain. Thus, while particular portions of this discussion are limited to the 4-node element, the general results and conclusions apply to any type of element.

The individual discretized continuity equations are rewritten as

$$\int_{\Omega} \psi_i \nabla \cdot \mathbf{u}^h = 0, \qquad i = 1, 2, \ldots, M \tag{9}$$

where M is the total number of elements (and pressures) in the domain. Of course, in practice, Ω is replaced by A_i, since $\psi_i = 0$ on $\Omega \ominus A_i$. By simple addition of all M equations we arrive at the global mass balance,

$$\int_{\Omega} \nabla \cdot \mathbf{u}^h = 0 \tag{10a}$$

which, via the divergence theorem, is equivalent to

$$\int_{\partial\Omega} \mathbf{n} \cdot \mathbf{u}^h = 0. \tag{10b}$$

Note that since this global equation is a simple linear combination of the individual continuity equations, it is inherently contained in the M separate equations and must therefore be satisfied by the numerical solution. It is this fact which can lead to the hydrostatic pressure mode; (10b) represents a constraint equation among the normal velocity components at the boundary nodes.

For the mesh in Figure 2, the global constraint equation, which may also be formed from the appropriate element level matrices (8), takes the form

$$\begin{aligned}(u_4-u_1)h_1+(u_8-u_5)(h_1+h_2)+(u_{12}-u_9)h_2+\\ (v_9-v_1)l_1+(v_{10}-v_2)(l_1+l_2)+(v_{11}-v_3)(l_2+l_3)+(v_{12}-v_4)l_3=0,\end{aligned} \tag{11}$$

which is the constraint equation among normal velocity components on the entire boundary of the domain and is represented pictorially by the vectors in Figure 2. It is to be emphasized that the general result (10) is independent of grid size or shape and also applies to other representations of pressure and velocity.

Consideration of the applied boundary conditions then leads directly to the following conclusions (assuming that there are no other pressure modes present):

1. If the imposed boundary conditions are such that they do not themselves imply a constraint among the normal velocities on the entire boundary (i.e. normal traction conditions, applied over any portion of the boundary), then the constraint equation (10) is independent and indeed required, there are no redundant continuity equations, there is no zero eigenvalue in the assembled matrix, and there is no hydrostatic pressure mode (the pressure level is set by the normal traction boundary conditions).

2. If the imposed boundary conditions identically satisfy (i.e. *duplicate*) the same constraint equation (which can only occur if all normal velocity components on the boundary are specified) as is implied by the global mass balance ((10) in general, (11) for the simple example), then the system is consistent, but overspecified. There will then exist a hydrostatic pressure mode (constant pressure; to be shown later), there is one redundant continuity equation, and the assembled matrix will have a zero eigenvalue corresponding to this pressure mode. If this pressure mode is the only one present, it can be specified, with no adverse effects, by eliminating (omitting) any one of the discretized continuity equations and selecting an arbitrary constant for the corresponding pressure.

The simplest and most common example is that of a contained flow. Here the normal velocity components are zero at each node on the boundary which leads to the well known case wherein the pressure is determined only up to an arbitrary additive constant.

3. Finally, if the imposed boundary conditions violate the constraint equation among the normal velocities on the boundary (which again can only occur if all normal velocity components on $\partial\Omega$ are specified), the problem is ill-posed, the algebraic system will be inconsistent, and no solution is possible (owing to the lack of global mass balance implied by the erroneous boundary conditions).

B. *The checkerboard pressure mode*

We now return to the more subtle pressure mode, the spurious CB mode and ask the question: Are there, in addition to the simple hydrostatic case just discussed, any other ways in which the M continuity equations can be linearly combined such that the result is a

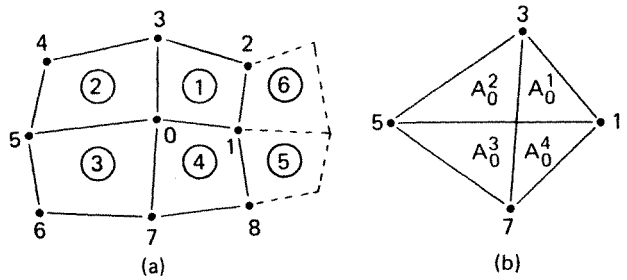

Figure 3. A general patch of elements

boundary node constraint equation (i.e. again, with all internal nodes eliminated)? To answer this question, consider first the patch of four elements shown in Figure 3(a). In order to show that it is possible to combine the discretized continuity equations on each element so as to eliminate u_0 and v_0 we may, without loss of generality, set $u_j = v_j = 0$, $j = 1, 2, \ldots, 8$.

Using the general element level matrix given by (8), the four continuity equations for this patch of elements may be written as

$$\mathbf{C}_0^{\mathrm{T}} \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix},$$

where

$$\mathbf{C}_0 = \begin{bmatrix} y_3 - y_1 & y_5 - y_3 & y_7 - y_5 & y_1 - y_7 \\ x_1 - x_3 & x_3 - x_5 & x_5 - x_7 & x_7 - x_1 \end{bmatrix}.$$

The rank of $\mathbf{C}_0$ is always two (except in the disallowed degenerate case when nodes 1, 3, 5 and 7 are collinear) and consequently there exist two linearly independent vectors $\mathbf{P}$ such that $\mathbf{P}^{\mathrm{T}}\mathbf{C}_0^{\mathrm{T}} = \mathbf{0}$. One such vector is given by $\mathbf{P} = \alpha(1, 1, 1, 1)^{\mathrm{T}}$ where α is an arbitrary constant, and corresponds to the hydrostatic mode discussed earlier. To describe the second such vector (the CB vector), we first construct the four triangles shown in Figure 3(b) (via straight lines interconnecting nodes 1, 3, 5 and 7 of Figure 3(a); the intersection of the two diagonals is generally *not* a node of the original grid). Then, omitting the algebraic details, it may be shown that $\mathbf{P}_0^{\mathrm{T}}\mathbf{C}_0^{\mathrm{T}} = \mathbf{0}$ for

$$\mathbf{P}_0 = \beta_0(1/A_0^1, -1/A_0^2, 1/A_0^3, -1/A_0^4)^{\mathrm{T}},$$

where β_0 is an arbitrary constant and A_j^k denotes the area of the triangle associated with node j which lies in element k. It is understood that the components of $\mathbf{P}_0$ are identified with the correspondingly numbered element. This multiplying vector necessarily causes u_0, v_0 to drop out when the linear combination $\mathbf{P}_0^{\mathrm{T}}\mathbf{C}_0^{\mathrm{T}}\mathbf{U}_0$ is formed. Since the internal velocities have been eliminated, a (spurious) 'boundary' constraint (and a CB mode) is always present for a patch of four elements taken in isolation; the nature of this constraint, and its potential 'propagation' through the mesh will now be described.

For a general mesh, the above construction must be performed at each internal node and then the local multiplying vectors $\{\mathbf{P}_j\}$ must be checked to determine whether or not they can be consistently assembled to form a *global* multiplying vector, which is required for the mesh to exhibit a CB mode. To illustrate this process, consider the enlarged patch of 6 elements shown in Figure 3(a) and let $\mathbf{P}_0, \mathbf{P}_1$, be the local multiplying vectors for nodes 0 and 1 respectively. The vector $\mathbf{P}_0$ is defined above and $\mathbf{P}_1$ is given by

$$\mathbf{P}_1 = \beta_1(1/A_1^6, -1/A_1^1, 1/A_1^4, -1/A_1^5)^{\mathrm{T}},$$

where the components of $\mathbf{P}_1$ are identified with elements 6, 1, 4 and 5 respectively. For consistency the first and fourth components of $\mathbf{P}_0$ should agree with the second and third

components of $\mathbf{P}_1$ respectively, since they are each identified with elements 1 and 4. This gives

$$\beta_0/A_0^1 = -\beta_1/A_1^1$$

and

$$-\beta_0/A_0^4 = \beta_1/A_1^4;$$

that is

$$A_0^1/A_0^4 = A_1^1/A_1^4 \tag{12}$$

(which also relates β_1 to β_0) as the condition on the *geometry* of the grid under which a boundary constraint could exist. If the grid satisfies (12), the appropriate linear combination of the six continuity equations which eliminates u_0, u_1, v_0 and v_1 is

$$\mathbf{P} = \beta_0(1/A_0^1, -1/A_0^2, 1/A_0^3, -1/A_0^4, A_1^4/(A_0^4A_1^5), -A_1^4/(A_0^4A_1^6)^{\mathrm{T}} \tag{12a}$$

which, for a uniform arrangement of rectangular elements, reduces to $P_j = \pm\beta_0/\text{Area of } \textit{element } j$. In both cases the checkerboard sign pattern is apparent. If the mesh in Figure 3(a) is extended in other directions, a condition similar to (12) must be satisfied at each of the nodes 3, 5, 7, . . . if all the internal variables are to be eliminated and a (potential—depending on boundary conditions) CB mode is to exist.

Regular assemblages of rectangular or parallelogram elements are simple examples of grids which identically satisfy all geometrical constraints of the form (12), but more irregular structures are also possible, an example of which will be presented later.

We shall now assume that the grids allow the complete elimination of internal velocities and try to elucidate the nature of the resulting boundary constraint. This constraint may conveniently be written in the form

$$\sum_{\substack{\text{boundary}\\ \text{nodes (j)}}} \sigma_j \begin{pmatrix} u_j \\ v_j \end{pmatrix}^{\mathrm{T}} \mathbf{a}_j = 0, \tag{13}$$

where $\sigma_j \pm 1$ (the sign is taken arbitrarily at any one boundary node and then alternates as the nodes are taken in sequence around the boundary) and the vectors $\mathbf{a}_j$ depend only on the geometry of the mesh in the neighbourhood of node j. For example, at a corner such as node 4 in Figure 3(a), $\mathbf{a}_4$ is perpendicular to the diagonal joining nodes 3 and 5. At a more general boundary point such as node 3 in Figure 4(a), the situation is a little more complicated. Define A_3^1 and A_3^2 to be the areas of the triangles $02P$ and $04P$ respectively, where $0P$ is perpendicular to the line joining nodes 4 and 2. It can then be shown that $\mathbf{a}_3$ is parallel to the vector

$$(A_3^1/A_0^1 - A_3^2/A_0^2)\mathbf{n} + h^2(1/A_0^1 + 1/A_0^2)\mathbf{t}, \tag{13a}$$

where h is the distance $0P$, A_0^1 and A_0^2 are defined as in Figure 3(b) and $\mathbf{n}$, $\mathbf{t}$ are, respectively, unit vectors orthogonal and parallel to the line 4–2 and serve to define normal and tangential directions at node 3. The vectors $\mathbf{a}_j$ are sketched in Figures 4(b) and 4(c) for two grids of square elements. For the special, but common (and important) case of a grid composed of rectangular elements, a simpler (degenerate) form of the boundary constraint equation is available. If the elements are considered to be alternately 'red' or 'black' as on a checkerboard, the special form of the boundary constraint equation is

$$\sum_{\text{red}} C_i/A_i - \sum_{\text{black}} C_i/A_i = 0, \tag{14}$$

where C_i represents the left-hand side of the discretized continuity equation ($[\mathbf{C}^{\mathrm{T}}\mathbf{U}]_i$) on the

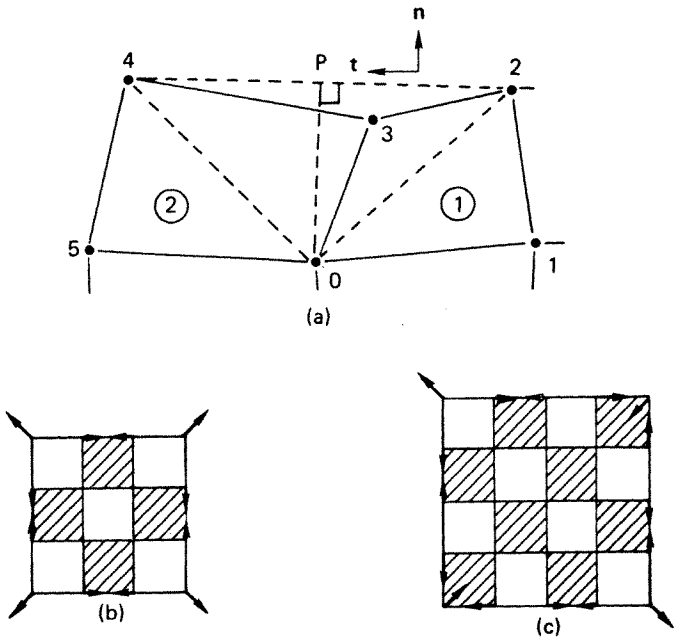

Figure 4. Element patch (a) and two CB grids (b), (c)

ith element which has area A_i. For example, applied to the grid of Figure 2, this yields

$$-\left(\frac{u_1}{l_1}+\frac{v_1}{h_1}\right)+u_2\left(\frac{1}{l_1}+\frac{1}{l_2}\right)-u_3\left(\frac{1}{l_2}+\frac{1}{l_3}\right)+\left(\frac{u_4}{l_3}-\frac{v_4}{h_1}\right)$$
$$+(v_5+v_8)\left(\frac{1}{h_1}+\frac{1}{h_2}\right)+\left(\frac{u_9}{l_1}-\frac{v_9}{h_2}\right)-u_{10}\left(\frac{1}{l_1}+\frac{1}{l_2}\right)+u_{11}\left(\frac{1}{l_2}+\frac{1}{l_3}\right)$$
$$-\left(\frac{u_{12}}{l_3}+\frac{v_{12}}{h_2}\right)=0, \tag{14a}$$

a constraint among the tangential boundary velocities. Equation (14) also describes the less common case of a grid composed of parallelograms.

When the grid is composed of square elements, another interpretation of the spurious boundary constraints, (10) and (14), is possible. Let $\partial\Omega_{\text{red}}$ ($\partial\Omega_{\text{black}}$) denote the intersection of red (black) elements with the boundary $\partial\Omega$. Since (10b) can be expressed as $\sum_{i=1}^{M} C_i = 0$ in the current notation, it is clear that (10) and (14) together imply that $\sum_{\text{red}} C_i = 0$ and $\sum_{\text{black}} C_i = 0$; i.e., the net flow across each of $\partial\Omega_{\text{red}}$ and $\partial\Omega_{\text{black}}$ must be zero separately.

Finally in terms of the notation developed in this section, the boundary constraint (10b), corresponding to the hydrostatic pressure mode, may be written

$$\sum_{\substack{\text{boundary}\\ \text{nodes (j)}}} \begin{pmatrix} u_j \\ v_j \end{pmatrix}^{\mathrm{T}} \mathbf{b}_j = 0, \tag{15}$$

where $\mathbf{b}_j$ is parallel to $\mathbf{a}_j$ when j is a corner node whilst, at a more general point such as node 3 of Figure 4(a), $\mathbf{b}_3$ is parallel to $(A_3^1+A_3^2)\mathbf{n}$, which should be compared with the expression (13a). It may be worth repeating that (15) is a valid boundary constraint because it has a

physical meaning (a global mass balance), but the boundary constraint represented by (13) (or (14) for rectangles) is a spurious artifact of the discretization.

Noting now the analogy between this boundary constraint equation (13) and that presented in the previous section ((10) or (11); or (15)), we can again draw the following conclusions by considering that (13) must be satisfied regardless of the applied boundary conditions (again ignoring, temporarily, the possibility of other pressure modes):

1. If the imposed boundary conditions (detailed in the next section) are such that they do not themselves imply a constraint among the velocities on the entire boundary, then the constraint equation (13) is independent and required, there are no redundant continuity equations, there is no zero eigenvalue in the assembled matrix, and there is no pressure mode. The resulting 'boundary equation' ((13) or (14)) generated by the continuity equations, is, however, extraneous (and spurious) and will be considered in more detail later.

2. If the imposed boundary conditions duplicate the same constraint equation, then the system is consistent, but over-specified. There will then exist a spurious CB pressure mode (related to the spurious boundary constraint), there is one redundant continuity equation, and the assembled matrix will have a zero eigenvalue corresponding to this pressure mode. If this pressure mode is the only one present, it can be 'specified' (but not eliminated) by omitting any one of the discretized continuity equations and selecting an arbitrary constant for the corresponding pressure. In addition, this pressure mode must be filtered if usable pressures are to be obtained; effective techniques will be presented in a later section.

3. Finally, if the imposed boundary conditions violate the (spurious) constraint of (13), the discretized problem is ill-posed, the algebraic system will be inconsistent, and no solution is possible. The conditions under which such a violation (which is *non*-physical) can occur are somewhat subtle, and fortunately do not appear to arise often; an example will be presented shortly.

To conclude this section, we consider the possibility of the simultaneous existence of both pressure modes since both constraint equations ((10) and (13)) are always applicable (there are but two pressure modes for this element):

1. If the applied boundary conditions imply no constraint equations among boundary velocities, then both implied constraint equations are required, there are no redundant continuity equations, no zero eigenvalues, and no pressure modes.

2. If the boundary conditions satisfy (duplicate) one of the constraint equations and do not violate the other, the system is consistent but over-specified, and there is one redundant continuity equation. Depending upon which of the constraint equations is satisfied by the boundary conditions, there will exist either a hydrostatic or a CB pressure mode. In either case, there is a zero eigenvalue in the matrix which can be removed by specifying one pressure and deleting the corresponding continuity equation. If the pressure mode is the spurious CB mode, further treatment will be required to filter this mode.

3. If the imposed boundary conditions duplicate both constraint equations, the system is consistent but over-specified; there are then two redundant continuity equations, one on a 'black' element and the other (in general) on a 'red'. The corresponding two zero eigenvalues in the matrix can be removed by specifying two pressures (at any values, the only restriction being that, in general, one must be black, the other red; two 'red pressures' may be specified only if the corresponding areas are not identical) and deleting the two corresponding continuity equations. Again in this case, post-processing will be required to filter the spurious CB mode. Note here that the specification of the pressure at two points in the flow field is physically absurd, but mathematically permissible; it is simply another manifestation of the spurious pressure mode.

4. Finally, if either of the constraints is violated by the imposed boundary conditions, the problem is ill-posed, the algebraic system will be inconsistent, and no solution is possible.

Further details on these matters will be presented below, and later, a sufficient number of numerical examples which substantiate these claims.

C. *Description of the pressure modes*

As mentioned earlier, the existence of a pressure mode is associated with a zero eigenvalue in the assembled system; i.e. the eigenvalue problem associated with (4a,b) is

$$\begin{aligned} \mathbf{K}\mathbf{w}_i + \mathbf{C}\mathbf{r}_i &= \lambda_i \mathbf{w}_i \\ \mathbf{C}^{\mathrm{T}}\mathbf{w}_i &= \lambda_i \mathbf{r}_i \end{aligned} \tag{16}$$

where $\{\lambda_i\}$ are the eigenvalues and $(\mathbf{w}_i, \mathbf{r}_i)^{\mathrm{T}}$ the eigenvectors, $i = 1, 2, \ldots, N+M$, and $N = N_1 + N_2$. When a pressure mode exists, there is a corresponding nonzero 'pressure' solution to (16) of the form

$$\begin{aligned} \lambda_i &= 0 \\ \mathbf{w}_i &= \mathbf{0} \end{aligned} \tag{17}$$

and

$$\mathbf{C}\mathbf{r}_i = \mathbf{0}$$

The eigenvector, $(\mathbf{0}, \mathbf{r}_i)^{\mathrm{T}}$, corresponding to (17), is a pressure mode and our goal here is to describe its form.

The clue to the characterization of these pressure modes is contained in the same linear combinations of continuity equations which were employed to display their existence. Let $\mathbf{z}^h(x)$ be the general velocity test function employed in the Galerkin approximation of (1a); then, as in (3a),

$$\mathbf{z}^h = \sum_{k=1}^{N_0} \mathbf{z}_k \phi_k(\mathbf{x}).$$

In accordance with the principles of Galerkin approximation, if the velocity $\mathbf{u}^h$ (or any component of $\mathbf{u}^h$) is prescribed at a boundary point, then $\mathbf{z}^h$ (or its corresponding component) must vanish at that point. For a pressure mode $\mathbf{r}_i$ to exist, we must have

$$\mathbf{Z}^{\mathrm{T}}\mathbf{C}\mathbf{r}_i = 0 \tag{18}$$

for nontrivial $\mathbf{r}_i$ and *all* global vectors $\mathbf{Z}$ which conform to the boundary conditions. We are now in a position to analyse the modes more closely.

1. *Hydrostatic pressure mode.* For this mode $\mathbf{r}_i = \mathbf{P}_{\mathrm{H}}$ where $\mathbf{P}_{\mathrm{H}} = \alpha(1, 1, \ldots, 1)^{\mathrm{T}}$ and α is an arbitrary constant. Now, from the earlier discussion of the boundary constraint, (15),

$$\mathbf{Z}^{\mathrm{T}}\mathbf{C}\mathbf{P}_{\mathrm{H}} = \mathbf{P}_{\mathrm{H}}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{Z} = \alpha \sum_{\substack{\text{boundary} \\ \text{nodes } (j)}} \mathbf{b}_j^{\mathrm{T}}\mathbf{z}^h \big|_j. \tag{19}$$

When the applied boundary conditions imply that $\mathbf{b}_j^{\mathrm{T}}\mathbf{u}^h \big|_j$ is specified at all boundary nodes. then these values must satisfy (15); furthermore, $\mathbf{b}_j^{\mathrm{T}}\mathbf{z}^h \big|_j$ must vanish for *all* functions $\mathbf{z}^h$ and boundary nodes j, which from (19) implies that $\mathbf{C}\mathbf{P}_{\mathrm{H}} = \mathbf{0}$. The hydrostatic pressure mode will therefore exist under the following boundary conditions:

a) u and v specified on $\partial\Omega$ (as in a contained flow for example),

b) u and v specified on part of $\partial\Omega$ and the normal velocity and tangential force (f_t) prescribed on the remainder. Here it is important to define the normal and tangential directions as indicated in Figure 4(a).

The hydrostatic pressure mode will not exist under any other type of boundary condition but the resulting solution $\mathbf{u}^h$ will satisfy (15).

2. *Checkerboard pressure mode.* We assume that the mesh satisfies the relation (12) at each internal node. Then, by choosing $\mathbf{r}_i = \mathbf{P}_c$ where the components of $\mathbf{P}_c$ define the linear combination of continuity equations which leads to (13), we have, from (13),

$$\mathbf{Z}^T\mathbf{C}\mathbf{P}_c = \mathbf{P}_c^T\mathbf{C}^T\mathbf{Z} = \sum_{\substack{\text{boundary} \\ \text{nodes (j)}}} \mathbf{a}_j^T\mathbf{z}^h \big|_j. \tag{20}$$

For a grid of simple rectangles, the components of $\mathbf{P}_c$ are $\pm 1/A_i$, where A_i is the area of the ith element and the plus sign corresponds to a 'black' element and the minus sign to a 'red.' $\mathbf{P}_c$ is more difficult to describe for a more general mesh which could support a CB mode (*cf.* (12a)). By analogy with the hydrostatic mode, when the applied boundary conditions imply that $\mathbf{a}_j^T\mathbf{u}^h \big|_j$ is specified at each boundary node, then these values must satisfy the constraint (13); furthermore such boundary conditions imply that $\mathbf{a}_j^T\mathbf{z}^h \big|_j = 0$ for *all* functions $\mathbf{z}^h$ and boundary nodes j. Consequently, $\mathbf{C}\mathbf{P}_c = \mathbf{0}$, indicating that a CB mode will exist. It follows therefore that a CB mode will be present if the velocity $\mathbf{u}^h$ is prescribed everywhere on $\partial\Omega$. In more general problems the normal force (f_n) and/or the tangential force (f_t) may be specified on part of the boundary. To simplify the discussion of these cases, let $\partial\Omega = \partial\Omega_1 \cup \partial\Omega_2$, where $\partial\Omega_1$ is *linear.* We shall suppose that both components of velocity are prescribed on $\partial\Omega_2$ whereas, on $\partial\Omega_1$

a) f_t and u_n are specified: all test functions $\mathbf{z}^h$ must satisfy $z_n^h = 0$, i.e. the test functions have nonzero component only in the tangential direction, $\mathbf{t}$. Using the representation (13a) for the vectors $\mathbf{a}_j$ on $\partial\Omega_1$, it follows that $\mathbf{a}_j^T\mathbf{z}^h \big|_j \neq 0$ and consequently $\mathbf{C}\mathbf{P}_c \neq \mathbf{0}$. This in turn implies that there cannot be a CB mode.

b) f_t and f_n are specified: no CB mode can exist,

c) f_n and u_t are specified: similar to case a), the test functions have nonzero component only in the normal direction $\mathbf{n}$ and therefore, from (20) and (13a), $\mathbf{a}_j^T\mathbf{z}^h \big|_j$ will vanish only for certain mesh configurations near $\partial\Omega_1$. For example, at a typical boundary node such as 3 in Figure 4(a) (the points 2, 3 and 4 lie on $\partial\Omega_1$ and must now be collinear) the geometric constraint deduced from (13a) is

$$A_0^2/A_0^1 = A_3^2/A_3^1. \tag{21}$$

This is satisfied for instance when all the elements which share an edge with $\partial\Omega_1$ are rectangular and identical, thus leading to a CB mode, an example of which is shown in Figure 5.

If the tangential boundary condition at the right side is changed from $v = 0$ to $f_t = 0$, the CB mode could no longer exist; we will return to such a case shortly, where we present another curious manifestation of the spurious constraint equation.

In this section we have presented the conditions under which pressure modes can exist and have shown the form of the associated eigenvectors. One of the reasons for describing the CB mode in such generality is to emphasize that the nature of the geometric constraints, (13) and (18), can make it extremely difficult to determine *a priori* whether or not a CB mode will be present. In general, the nonexistence of pressure modes is a sufficient condition for the existence and uniqueness of a solution to the algebraic system. Conversely, the existence of

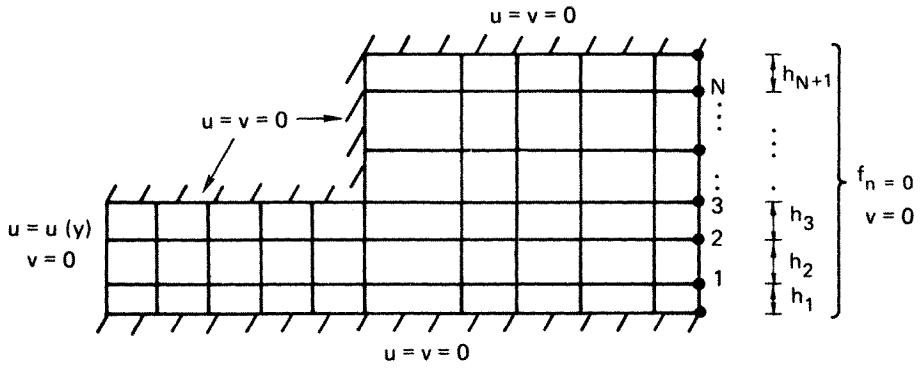

Figure 5. An example of a domain and boundary conditions which will cause a CB pressure mode

a pressure mode affects existence and uniqueness of a solution to the algebraic system. For example, there are common situations in which pressure modes exist and the algebraic solution exists, but is nonunique (in the pressure); furthermore, there are even situations in which a CB mode 'exists' and *precludes* a solution of the discretized Navier–Stokes equations! This ill-posedness is an artifact of the spurious CB constraint equation combined with boundary conditions which, while physically valid and meaningful, violate this constraint. An example of such a case is presented in the next section.

D. *Further implications of the spurious constraint*

The fact that the 'CB constraint equation,' ((13); (14) for rectangles), must *always* be satisfied by the discretized solution, carries other, rather serious implications over and above the possible existence of a CB mode. We will demonstrate two consequences of this extraneous constraint via simple examples. In the first example, the constraint on boundary velocities occurs in a flow which has no CB mode and, in the second example, it occurs in conjunction with the CB mode and can lead to an ill-posed problem.

Example 1: Tangential velocity constraints. We begin by returning to the grid in Figure 5, assumed to be composed of rectangular elements, and modify the outflow boundary condition from $v = 0$ to $f_t = 0$, so that no CB mode can exist. However, application of the CB constraint equation (14) to this grid gives, considering the imposed boundary conditions,

$$\left(\frac{1}{h_1}+\frac{1}{h_2}\right)v_1-\left(\frac{1}{h_2}+\frac{1}{h_3}\right)v_2+\left(\frac{1}{h_3}+\frac{1}{h_4}\right)v_3-\cdots\pm\left(\frac{1}{h_N}+\frac{1}{h_{N+1}}\right)v_N \doteq 0. \tag{22}$$

The existence of this spurious constraint equation, which will be satisfied by the numerical solution, is an artifact of the discretization with the bilinear element and it exists independently of, and in addition to, the proper constraints; viz, those imposed by the shear stress-free boundary condition and global mass conservation. If this 'element' converges to the solution of the Navier–Stokes equations, this constraint presumably is not too harmful (we have performed many numerical simulations which do, in fact, satisfy (22), but which still 'look reasonable') and in fact, must vanish as $h \to 0$.

It is also noteworthy that the same CB constraint equation applies *internally*; i.e. on the 'boundary' of *any* patch of elements (which could be completely internal or could include a portion of the domain boundary).

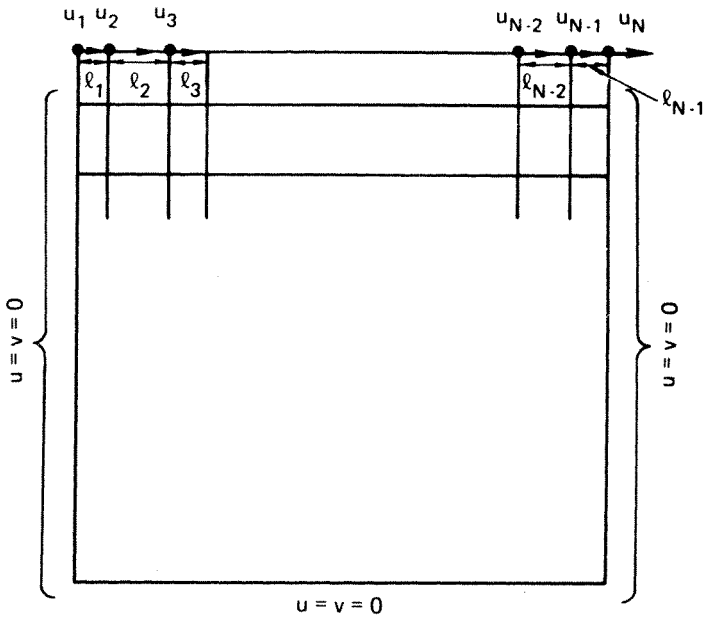

Figure 6. The driven cavity problem

We next consider a case where the CB constraint can be very harmful, after mentioning that if v is specified at the outflow boundary in this example in such a way that (22) is violated, the algebraic system will be inconsistent and no solution exists.

Example 2: The driven cavity. The popular lid-driven cavity problem is an appropriate example to demonstrate another important consequence of the CB constraint, and is depicted in Figure 6.

Application of (14) to this grid gives a different result, depending on whether $N-1$ (the number of elements across the top of the cavity) is even or odd:

For $N-1$ even,

$$\frac{u_1}{l_1}-\left(\frac{1}{l_1}+\frac{1}{l_2}\right)u_2+\left(\frac{1}{l_2}+\frac{1}{l_3}\right)u_3-\cdots-\left(\frac{1}{l_{N-2}}+\frac{1}{l_{N-1}}\right)u_{N-1}+u_N/l_{N-1}=0, \tag{23}$$

and for $N-1$ odd,

$$\frac{u_1}{l_1}-\left(\frac{1}{l_1}+\frac{1}{l_2}\right)u_2+\left(\frac{1}{l_2}+\frac{1}{l_3}\right)u_3-\cdots+\left(\frac{1}{l_{N-2}}+\frac{1}{l_{N-1}}\right)u_{N-1}+u_N/l_{N-1}=0. \tag{24}$$

Consider first the simpler case in which $u_i = u_0$, $i = 1, 2, \ldots, N$; i.e. the case of equal velocity at every node, including the first and last (a 'flow-through' cavity). In this case, both (23) and (24) are satisfied identically and a CB mode will exist (i.e. the CB constraint is automatically satisfied for either an even or odd mesh). If, however, we wish to compute the more difficult case of a contained flow, we must set $u_1 = u_N = 0$ and, for example, $u_i = u_0$, $i = 2, 3, \ldots, N-1$. In this case we obtain

$$-u_0\left(\frac{1}{l_1}+\frac{1}{l_{N-1}}\right)=0 \quad \text{for an even grid,}$$

and

$$u_0\left(-\frac{1}{l_1}+\frac{1}{l_{N-1}}\right)=0 \quad \text{for an odd grid.}$$

While the constraint equation on an odd grid can be satisfied, if and only if $l_1 = l_{N-1}$ (in which case the CB pressure mode, as well as the hydrostatic mode, exists), it can *never* be satisfied on an even grid (which also possesses the two pressure modes). Hence, the driven cavity problem, for these (mathematically permissible) boundary conditions, is *ill-posed* on any grid with an even number of elements across the top and on any odd-element grid which doesn't satisfy $l_1 = l_{N-1}$. These are clearly physically erroneous constraints and are forced upon the discretized system by the extraneous CB pressure mode. An even grid can, if desired, be employed for a contained flow ($u_1 = u_N = 0$) simulation, if the proper precautions are taken; e.g. for $u_i = u_0$; $i = 3, 4, \ldots, N-2$, (23) gives

$$u_0\left(\frac{1}{l_2}+\frac{1}{l_{N-2}}\right)-u_2\left(\frac{1}{l_1}+\frac{1}{l_2}\right)-u_{N-1}\left(\frac{1}{l_{N-2}}+\frac{1}{l_{N-1}}\right)=0$$

which is easily satisfied (e.g. for $l = \text{constant}$, $u_2 + u_{N-1} = u_0$ will suffice, and it is then reasonable to take $u_2 = u_{N-1} = 1/2u_0$; we have used this approach successfully and results were presented in Lee *et al.*[10] (If we attempt to solve the ill-posed algebraic system, we obtain pressures of $O(10^{14})$ and velocities which, while $O(1)$, are meaningless.)

These conclusions, which have been substantiated via numerical calculations, are also explainable from the viewpoint of linear algebra. When the coefficient matrix in (4a,b) is rank deficient, the algebraic system is inconsistent unless the right hand side vector has no projection into the null space of the matrix (it must lie in the range of the matrix). Since the null space (of dimension 2) comprises the two pressure modes, the solvability conditions for (4a,b) are

$$(\mathbf{g}, \mathbf{P}_H) = 0 \tag{25a}$$

$$(\mathbf{g}, \mathbf{P}_c) = 0 \tag{25b}$$

where (,) is the conventional inner product. Since $\mathbf{g}$ is composed from specified boundary velocities (it is obtained by transposing those terms in $\mathbf{C}^T\mathbf{U} = \mathbf{0}$ corresponding to prescribed velocities to the right hand side), it is perhaps not surprising that (25a) is precisely the global mass balance requirement ((10) or (11)) and (25b) is the CB constraint equation ((13) or (14) or, for the driven cavity example, (23) or (24)).

The net result, as demonstrated by these two examples, is that the CB constraint is rather insidious, far-reaching, and undoubtedly even has important implications regarding the ultimate proof of convergence of this FEM approximation (which proof, according to Fortin,[20] is 'still an open question'—perhaps it must remain that way).

E. *The impure checkerboard pressure mode*

We now address one of the most difficult (and ominous) of the pressure mode effects which we have encountered. One might optimistically expect, since the existence of the CB mode was proven under rather specialized conditions, that it would not occur under the more general conditions of a mesh composed of variously distorted isoparametric elements. Unfortunately, this is not the case and 'our troubles have just begun'; in fact, a 'residual' CB pattern appears to be present (under appropriate boundary conditions; viz, those which permit the existence of a pure CB mode) even in a mesh composed of severely distorted elements. However, since it does not display identical characteristics to what may be called

the pure CB mode, we have labelled this an 'impure' CB mode. In this section, we attempt to characterize this CB mode, the impurity of which hampers a rigorous analysis. The important fact is that it does persist in numerical calculations and thus, in a later section, we present effective means of dealing with both pure and impure CB modes.

Basically, the impure CB mode appears to exist in such a way that, while not pure (there is no corresponding zero eigenvalue with an associated pure pressure eigenvector—other than the hydrostatic pressure mode—and therefore, no associated redundant continuity equation), the pressure solution is still oscillatory and generally unacceptable without further post-processing. Our explanation of the impure mode (one which appears to explain essentially all of the results from a wide variety of numerical experiments) is one which considers it as a perturbation from the simpler pure CB mode and it is this approach which we shall present; i.e. any mesh which does not support a pure CB mode is to be interpreted as a perturbation (small in theory, but not necessarily in practice) from one which does.

To motivate this discussion, consider the following experimental results from the cavity problem (details of which are presented under Numerical Examples in Part 2):

1. On a mesh composed of simple rectangles, there exists a pure CB pressure mode for the driven or undriven cavity problem (by undriven we mean $\mathbf{f}=\mathbf{0}$, $\mathbf{g}=\mathbf{0}$ in (4a,b). Actually, to excite the CB mode in the undriven case, we must set the pressure on one element so that in fact $\mathbf{g}\neq\mathbf{0}$. In this case, the velocity is, of course, zero to machine roundoff).
2. If the mesh is perturbed by moving one or more internal nodes (a small but 'sufficient' distance), the undriven cavity displays no CB mode (except that caused by roundoff error for very small perturbations), but the driven cavity contains (often in very sizable amount) the impure CB pressure mode. Furthermore, in the perturbed mesh, our Gaussian elimination process (a frontal solver) produces one small, but nonzero pivot (again, for 'recognizable' perturbations), whereas when the pure CB mode exists, the solver produces one pivot of essentially zero (machine roundoff), which indicates a singular matrix. (For 'large' mesh perturbations, there are usually no clearly identifiable small pivots).
3. For the driven cavity, the perturbation in velocities is often inordinately large ($O(1)$) relative to the size of the mesh perturbation (e.g. 10^{-6} to 10^{-3})—a quite unsettling result.

It thus appears that the original CB theory is only partially applicable in that it would predict that the perturbed matrix would no longer have a zero eigenvalue (correct) and that a CB pressure mode would no longer exist (incorrect). It is toward the reconciliation of these matters, as well as that of the large velocity perturbation, that we present a perturbation analysis of the impure CB mode.

A crucial role is played in the analysis by the way in which a small perturbation of the nodes affects the continuity constraints. For a typical element (say element 1 of Figure 3(a)) the coordinates of the jth node (x_j, y_j) are perturbed to $(x_j+\varepsilon\alpha_j\cos\theta_j, y_j-\varepsilon\alpha_j\sin\theta_j)$ where $|\alpha_j|\leq 1$ $(j=0,1,2,3)$ and with the proviso that boundary nodes remain fixed. The continuity constraints on the original and perturbed elements are represented by $[\mathbf{C}^{\mathrm{T}}\mathbf{U}]_1$, and $[\mathbf{C}_\varepsilon^{\mathrm{T}}\mathbf{U}]_1$, respectively, where, at this point the global N-vector $\mathbf{U}$ is left unspecified. Defining $\boldsymbol{\nu}_j=(\alpha_j\sin\theta_j, \alpha_j\cos\theta_j)$, $j=0,1,2,3$, it can be shown by direct construction that

$$\begin{aligned}[\mathbf{C}_\varepsilon^{\mathrm{T}}\mathbf{U}]_1 &= [\mathbf{C}^{\mathrm{T}}\mathbf{U}]_1-\varepsilon\{(\mathbf{u}_2-\mathbf{u}_0)\cdot(\boldsymbol{\nu}_3-\boldsymbol{\nu}_1)+(\mathbf{u}_3-\mathbf{u}_1)\cdot(\boldsymbol{\nu}_0-\boldsymbol{\nu}_2)\}\\ &= [\mathbf{C}^{\mathrm{T}}\mathbf{U}]_1+\varepsilon\mathbf{c}_1^{\mathrm{T}}\mathbf{U},\end{aligned}$$

where the global N-vector $\mathbf{c}_1$ is independent of ε. We now suppose that the original element belongs to a grid of elements on which there is a pure CB mode $\mathbf{P}_c$ (normalized so that $\|\mathbf{P}_c\| = 1$). Then $\mathbf{P}_c^T \mathbf{C}^T \mathbf{U} = 0$ for *any* vector $\mathbf{U}$ and consequently the appropriate linear combination of all element equations of the above form leads to

$$\mathbf{P}_c^T \mathbf{C}_\varepsilon^T \mathbf{U} = \varepsilon \sum_{k=1}^{M} [\mathbf{P}_c]_k \mathbf{c}_k^T \mathbf{U} \equiv \varepsilon \mathbf{c}^T \mathbf{U}, \text{ say,} \tag{26}$$

where $\mathbf{c} = \sum_{k=1}^{M} [\mathbf{P}_c]_k \mathbf{c}_k$ is independent of ε and the construction of $\mathbf{c}_k$ on element k follows that typified above for $\mathbf{c}_1$ (an alternative way to construct $\mathbf{c}$ and to derive (27) is given at the end of Appendix II). Finally, if we take $\mathbf{U} = \mathbf{U}_\varepsilon$, where $\mathbf{U}_\varepsilon$ denotes the velocity solution on the perturbed grid, it must *also* satisfy $\mathbf{C}_\varepsilon^T \mathbf{U}_\varepsilon = \mathbf{0}$ which, from (26), leads to the spurious *internal* constraint

$$\mathbf{c}^T \mathbf{U}_\varepsilon = 0, \tag{27}$$

which can lead to an $O(1)$ perturbation in $\mathbf{U}$ *regardless* of the smallness of ε. By means of suitable perturbations of the nodes, the velocity $\mathbf{U}_\varepsilon$ can therefore be made to satisfy virtually any preassigned (nonphysical) constraint, which may preclude the possibility of convergence as $h \to 0$ (here, of course, ε must be proportional to h).

Further ramifications of this constraint can be deduced by analysing the eigen problems for the two grids. This analysis is somewhat lengthy and the details are therefore relegated to Appendix II; we present here only the final results. The eigen problem on the unperturbed grid (see (16)) possesses two zero eigenvalues ($\lambda_1 = \lambda_2 = 0$) and, corresponding to the nonzero eigenvalues $\lambda_3, \lambda_4, \ldots, \lambda_{N+M}$, a set of orthonormal eigenvectors $\begin{pmatrix} \mathbf{w}_j \\ \mathbf{r}_j \end{pmatrix}$, $j = 3,4, \ldots, N+M$. The solution of the system (4a,b) may be conveniently expressed in the form

$$\begin{pmatrix} \mathbf{U} \\ \mathbf{P} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{U}} \\ \tilde{\mathbf{P}} \end{pmatrix} + \gamma_H \begin{pmatrix} \mathbf{0} \\ \mathbf{P}_H \end{pmatrix} + \gamma_c \begin{pmatrix} \mathbf{0} \\ \mathbf{P}_c \end{pmatrix}, \tag{28}$$

where

$$\begin{pmatrix} \tilde{\mathbf{U}} \\ \tilde{\mathbf{P}} \end{pmatrix} = \sum_{j=3}^{N+M} \lambda_j^{-1} [(\mathbf{f}, \mathbf{w}_j) + (\mathbf{g}, \mathbf{r}_j)] \begin{pmatrix} \mathbf{w}_j \\ \mathbf{r}_j \end{pmatrix};$$

the coefficients γ_H, γ_c are arbitrary and the pressure modes are normalized so that $\|\mathbf{P}_H\| = \|\mathbf{P}_c\| = 1$. The corresponding solution on the perturbed grid is

$$\begin{pmatrix} \mathbf{U}_\varepsilon \\ \mathbf{P}_\varepsilon \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{U}} \\ \tilde{\mathbf{P}} \end{pmatrix} + \gamma_H \begin{pmatrix} \mathbf{0} \\ \mathbf{P}_H \end{pmatrix} + \gamma_c' \begin{pmatrix} \mathbf{0} \\ \varepsilon^{-1} \mathbf{P}_c' \end{pmatrix} - \gamma_c' \sum_{j=3}^{N+M} \lambda_j^{-1} (\mathbf{c}, \mathbf{w}_j) \begin{pmatrix} \mathbf{w}_j \\ \mathbf{r}_j \end{pmatrix}, \tag{29}$$

where γ_c' is a fixed constant of magnitude $O(1)$ and $\mathbf{P}_c' = \mathbf{P}_c - (\mathbf{P}_H, \mathbf{P}_c)\mathbf{P}_H$. Furthermore, the eigenvalues $\lambda_1^\varepsilon, \lambda_2^\varepsilon, \ldots, \lambda_{N+M}^\varepsilon$ of the perturbed system satisfy

$$\lambda_j^\varepsilon - \lambda_j = \begin{cases} 0, & j = 1, \text{ the hydrostatic mode} \\ O(\varepsilon^2), & j = 2, \text{ the CB mode} \\ O(\varepsilon), & j = 3, 4, \ldots, N+M. \end{cases} \tag{30}$$

Comparison of (28) and (29) reveals that there is in general an $O(1)$ perturbation of the velocity field and the amplitude of the CB pressure mode is $O(\varepsilon^{-1})$. Clearly some action must be taken to combat this behaviour if reliable results are to be obtained from this element. Before addressing ourselves to this question we point out that ε should not be

identified with round-off error effects since, for practical purposes, (29) accurately reflects the behaviour of the computed solution only when ε is at least large enough that an inequality of the form $1+\varepsilon^2 \neq 1$ is satisfied to machine precision; for smaller values of ε, the situation is better described by (28). This is the practical interpretation of the singular limiting process as $\varepsilon \to 0$ (the actual limit does not exist since the solution is not analytic at $\varepsilon = 0$). We may also think of ε as being a measure of the distance from a given (perturbed) grid to the nearest pure CB grid.

The significant difficulties associated with the singular limit can be more easily appreciated and the impure mode perhaps better understood by considering the following simple but powerful example with only two equations:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The matrix has eigenvalues $\lambda_1 = 1$, $\lambda_2 = 0$ and the linear system has the solution $\begin{pmatrix} 1 \\ a \end{pmatrix}$, where a is arbitrary. A small perturbation in the off-diagonal elements gives

$$\begin{pmatrix} 1 & \varepsilon \\ \varepsilon & 0 \end{pmatrix}\begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The matrix now has eigenvalues $\lambda_1^{\varepsilon} = [1+\sqrt{(1+4\varepsilon^2)}]/2 \cong 1+\varepsilon^2$ and $\lambda_2 = [1-\sqrt{(1+4\varepsilon^2)}]/2 \cong -\varepsilon^2$, and the algebraic system has the solution $\begin{pmatrix} 0 \\ \varepsilon^{-1} \end{pmatrix}$. The first component is changed by $O(1)$ and the second by $O(1/\varepsilon)$; furthermore, the unperturbed solution cannot be recovered by letting ε approach zero. (Although the eigenvalues are both perturbed to $O(\varepsilon^2)$, λ_1 would be perturbed to $O(\delta)$ if we had also perturbed the 1 in the matrix to $1+\delta$; but the perturbed solution would still be $(0, 1/\varepsilon)^{\mathrm{T}}$.)

These perturbation results can be used to explain the observed behaviour of the numerical solution to the perturbed cavity problem (and ostensibly any other situation in which an impure CB mode exists): For the undriven cavity, $\mathbf{f} = \mathbf{0}$ and $\mathbf{g} = \mathbf{0}$ and hence, (29) gives (since $\gamma_c' = 0$; see Appendix II)

$$\begin{pmatrix} \mathbf{U}_{\varepsilon} \\ \mathbf{P}_{\varepsilon} \end{pmatrix} = \gamma_{\mathrm{H}} \begin{pmatrix} \mathbf{0} \\ \mathbf{P}_{\mathrm{H}} \end{pmatrix}, \tag{31}$$

which agrees with our numerical result (there is no CB mode). For the driven cavity, $\mathbf{f}$ and $\mathbf{g}$ are of course nonzero, the solution is contaminated with a 'significant' amount of the CB pressure mode, and the velocity perturbations are $O(1)$, which also corroborates our numerical results. In both cases, the eigenvalue associated with the CB mode is nonzero (the smallest pivot varies like ε^2 when one pressure is specified).

Focusing now on the inordinately large velocity perturbation caused by the impure mode, we can suggest a practical and usually effective remedy. Recall first that in the presence of a hydrostatic mode and a *pure* CB mode, it is quite permissible to peg (specify) the pressures on two elements, one red and one black, and remove the continuity equations for these elements; the two deleted constraints will be satisfied implicitly. In the case of an impure CB mode, one can proceed similarly (at least for small ε): the pressure can be pegged on elements $M-1$ and M, say, and the momentum equations solved subject to the constraints

$$[\mathbf{C}_{\varepsilon}^{\mathrm{T}}\mathbf{U}_{\varepsilon}]_k = 0, \qquad k = 1, 2, \ldots, M-2.$$

From these constraints, (26), and the identity $\mathbf{P}_H^T\mathbf{C}_\varepsilon^T\mathbf{U}_\varepsilon = 0$, where $\mathbf{P}_H = \gamma_H(1, 1, \ldots, 1)^T$ is the hydrostatic mode, we can deduce that

$$-[\mathbf{C}_\varepsilon^T\mathbf{U}_\varepsilon]_{M-1} = [\mathbf{C}_\varepsilon^T\mathbf{U}_\varepsilon]_M = \varepsilon \mathbf{c}^T\mathbf{U}_\varepsilon/([\mathbf{P}_c]_M - [\mathbf{P}_c]_{M-1}). \tag{32}$$

Provided that the 'colour' of element $M-1$ is different than that of M, the corresponding components of the (pure) CB vector $\mathbf{P}_c$ will have opposite sign and the denominator on the right hand side of (32) cannot vanish. Thus, whilst there is not an *exact* mass balance on these two elements, the actual imbalance will be small in the troublesome cases when ε is small, and, importantly, the velocity perturbation is now small—$O(\varepsilon)$. Any attempt to achieve exact mass balance by pegging the pressure on only one element will of course cause (32) to degenerate to the constraint (27). Whether the loss of exact mass balances on two elements is 'worse than' the potentially $O(1)$ velocity error appears to depend on how far the perturbed mesh is from one which displays a pure CB, a determination that is generally impossible to make *a priori*. In the absence of a hydrostatic mode, the analogous procedure is to simply peg the pressure on any single element (one then trades an inexact element mass blance on one element for a more regular perturbed result—the velocity field will then be perturbed to $O(\varepsilon)$).

We have ignored boundary terms in the above analysis in order to present a clearer picture; the conclusions remain valid provided the boundary vector $\mathbf{g}$ satisfies the consistency conditions (25).

Finally, in addition to the impure CB mode described above, we have also encountered another sort of impure mode. This mode is seen under certain conditions on a mesh which would support a pure CB (e.g. a grid of rectangles) except that the boundary conditions, specified f_t along some portion of $\partial\Omega$, should preclude it. While we have not analysed this (typically small amplitude) impure mode in any detail, it seems that the pure CB eigenvector is perturbed by the change in boundary conditions in such a way that the projection of $\mathbf{g}$ (see (4b)) in this direction (e.g. at a specified inflow boundary) excites the impure mode and the resulting pressures are oscillatory near such a boundary.

F. *Filtering and smoothing techniques*

Since the CB pressure mode has been shown to be quite persistent, it must be filtered from the physical part of the pressure solution if usable pressure results are to be obtained. The techniques which we have developed to filter this pressure mode are directly related to a knowledge of the form of the CB pressure mode eigenvector. They apply rigorously to the appropriate pure CB modes; for impure modes, they are sometimes less effective, but usually still quite useful, especially when used in conjunction with grid smoothing (see below).

For a typical patch of four elements as shown in Figure 3(a), the CB vector $\mathbf{P}_c$ has components $(-1)^{j+1}\beta_0/A_0^j$ in element j ($j = 1, 2, 3, 4$). Recognizing that the computed pressure is the sum of a physical pressure and a spurious pressure mode, it is clear that the CB contribution to the sum $\sum_{j=1}^4 P_j A_0^j$ is zero and this forms the basis for the filters.

The proposed filtering techniques we shall describe also generate a smoothed (physical) pressure. The reason we refer to the pressure as smoothed is that the values will be available at the velocity nodes (or a closely related set of perturbed nodes, as described presently) rather than at the element centroids and can, if desired, be considered as continuous functions via representation by the velocity basis functions $\{\phi_i\}$.

Scheme 1

$$P_0 = \sum_{j=1}^4 P_j A_0^j / \sum_{j=1}^4 A_0^j, \tag{33}$$

where A_0^j ($j = 1, 2, 3, 4$) are the triangle areas defined in Figure 3(b), P_j is the computed (and polluted) pressure in element j, and P_0 is the smoothed presure at node 0 of Figure 3(a). Although this scheme always successfully filters the CB mode, it is cumbersome to implement. We therefore propose two alternatives, which often appear to work nearly as well and are simpler to implement.

Scheme 2

$$P_0 = \sum_{j=1}^{4} P_j \bar{A}_0^j \Big/ \sum_{j=1}^{4} \bar{A}_0^j, \tag{34}$$

where $\bar{A}_0^j$ ($j = 1, 2, 3, 4$) are the triangle areas shown in Figure 7. It is a straightforward though tedious exercise to show that the CB mode makes no contribution to the numerator of (34). If the mesh were such that node 0 were on the straight lines 1–5 and 3–7 (e.g. for a mesh of rectangles), it is clear that Scheme 2 would be identical to Scheme 1.

For a node lying on the boundary (for example node 3 in Figure 7) we define

$$P_3 = (P_1 A_0^1 + P_2 A_0^2)/(A_0^1 + A_0^2) \quad \text{for Scheme 1 or 2 (using } \bar{A}_0^i \text{ for the latter).}$$

The value provided by this expression should properly be associated with a point lying midway between node 0 and node 3; this can be corrected by taking instead the extrapolated value $2P_3 - P_0$. The pressure at a corner node (e.g. node 2 of Figure 7) can be deduced either by extrapolating linearly from the three other nodes of the element (e.g. for node 2, $P_2 = P_1 + P_3 - P_0$), or by a least squares process as described by Lee *et al.*,[10] or by the scheme employed by Hughes *et al.*[9]

Scheme 3. Another filter which is simpler yet and has generally been quite successful, is also defined by (33), but this time the quantities A_0^j ($j = 1, 2, 3, 4$) are interpreted more simply as the areas of the respective quadrilateral *elements*. This scheme does not always totally annihilate the CB mode (except for rectangular or parallelogram elements) and must therefore be used selectively. Also, boundary nodes still require additional treatment, as discussed above. Caldwell[15] has applied this technique (extended from 2-D in the obvious way—volume weighting) to successfully smooth CB modes obtained when using the 3-D BAAL finite difference code (Pracht and Brackbill[12]). We (PMG and RLL) also use volume weighting in our new 3-D FEM code using the 8-node brick element.

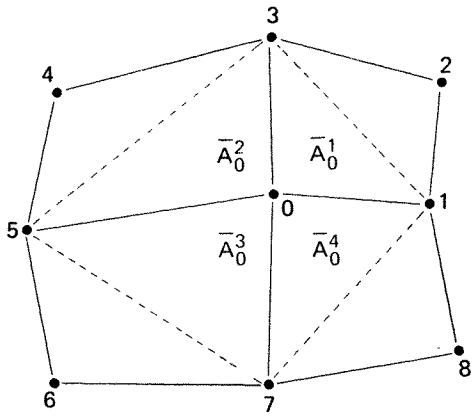

Figure 7. Triangular areas for Scheme 2

Finally, the basis function-weighted scheme described in Lee *et al.*,[10] currently has little to recommend it; it is theoretically less sound, is more work than Scheme 3, and usually gives results which differ little from those of Scheme 3. Also note that Schemes 1 and 2 are equivalent to Scheme 3 for grids composed of rectangles or parallelograms.

Numerical results based on these filters are described in a later section, as well as those in which the *nodes* are moved, as described next.

G. *Grid smoothing*

Improved results are obtained if the grid is smoothed (or relaxed) in the same manner as the pressures, a trick which is both useful and simple—it can be done simultaneously with the pressure smoothing (same DO LOOP). In fact, it is easy to show (via Taylor series) that if P_0 is given by (33), it is a second-order accurate approximation to a smoothed pressure *only* if its location (x_0, y_0) is given by

$$x_0 = \sum_{j=1}^{4} x_j A_0^j / \sum_{j=1}^{4} A_0^j, \tag{35}$$

where x_j is the x-coordinate of the centroid of element j, with an analogous equation for y_0; A_0^j is the same area used to smooth the pressure. (For a patch of four different-sized rectangles, this prescription moves the internal node to the geometric centre of the four-patch.) If the nodes are not re-located, the resulting smoothed pressure is only a first-order accurate result at the nodes. It seems probable that grid smoothing can be used to retain the optimal rate of convergence of the pressure on meshes which are quite distorted; we will present a rather convincing example of this later. After moving the internal nodes, the boundary nodes (save corners) can be moved in the *tangential* direction only and the previously discussed extrapolation methods employed.

H. *Extension to 3-D*

The extension of the CB analysis to 3-D, via the simplest element (8-node trilinear velocity, piecewise constant pressure on hexahedra) leads, somewhat surprisingly, to *multiple* CB pressure modes. On a grid composed of N_x by N_y by N_z brick-shaped elements, there exist (almost) as many CB modes as there are 'planes' of elements which contain specified tangential velocity components on the boundary. The 'worst case' is a fully contained flow, for which there are $N_x + N_y + N_z - 2$ pure CB pressure modes (and one hydrostatic mode)! For example, for the mesh shown in Figure 8, there are 10 pure CB modes and one hydrostatic mode for a contained flow, corresponding to the shaded elements (11 of 60 continuity equations are actually redundant). If the pressure is specified on the shaded elements, the 10 zero eigenvalues corresponding to the CB modes are eliminated (and of course the velocity solution is unaffected).

Other than the appearance of multiple modes, most of the 2-D results seem to carry over: e.g. the existence of impure modes on distorted meshes, the elimination of some of the modes when shear stress boundary conditions are employed, and the corresponding spurious constraints on tangential velocities. (A possibility which is unique to 3-D is the occurrence of *mixed* modes; some pure, and some impure.) Similarly, the obvious extension of our 2-D filtering and smoothing techniques is applicable (we have successfully filtered *all* of these (pure) CB modes using simple volume averaging) and, indeed, necessary; finally, the additional advantages of grid smoothing are also applicable.

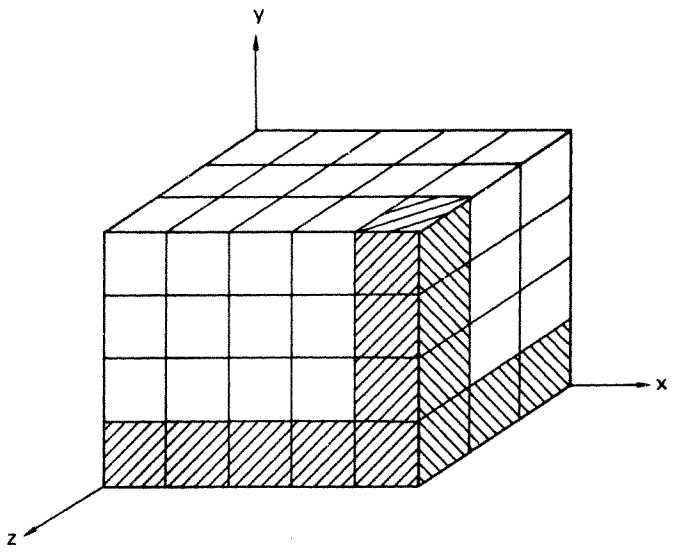

Figure 8. A simple 3-D grid

APPENDIX I: THE PENALTY MATRIX

Theorem

Let $\mathbf{W}_h$ and Φ_h denote, respectively, the velocity and pressure spaces and define the linear operator $\nabla_h: \mathbf{W}_h \rightarrow \Phi_h$ by $(\nabla_h \cdot \mathbf{v}, p) = (\nabla \cdot \mathbf{v}, p)$, for all $p \in \Phi_h$ and $\mathbf{v} \in \mathbf{W}_h$. then the consistent[19] penalty matrix $\mathbf{B}$ and the Lagrange multiplier matrix $\mathbf{C}$ are related by $\mathbf{B} = \mathbf{C}\mathbf{M}^{-1}\mathbf{C}^{\mathrm{T}}$ where $\mathbf{M}$ is the mass matrix on Φ_h. (Note: There does not seem to be any restriction on $\mathbf{W}_h$ and Φ_h.)

Proof. Let $\mathbf{V}$ and $\mathbf{P}$ denote the vectors of nodal values for $\mathbf{v} \in \mathbf{W}_h$ and $p \in \Phi_h$; then $\mathbf{C}$, $\mathbf{B}$ are given by

$$\mathbf{V}^{\mathrm{T}}\mathbf{C}\mathbf{P} = (\nabla \cdot \mathbf{v}, p) \text{ for all } \mathbf{v} \in \mathbf{W}_h,\ p \in \Phi_h$$

$$\mathbf{V}^{\mathrm{T}}\mathbf{B}\mathbf{U} = (\nabla_h \cdot \mathbf{v}, \nabla_h \cdot \mathbf{u}) \text{ for all } \mathbf{u}, \mathbf{v} \in \mathbf{W}_h.$$

Now $\nabla_h \cdot \mathbf{v} \in \Phi_h$ and therefore there exists a matrix $\mathbf{Q}$ such that $\nabla_h \cdot \mathbf{v} = \mathbf{V}^{\mathrm{T}}\mathbf{Q}\boldsymbol{\psi}(\mathbf{x})$, where $\boldsymbol{\psi}(\mathbf{x})$ denotes the vector of basis functions on Φ_h. Hence $(\nabla_h \cdot \mathbf{v}, \nabla_h \cdot \mathbf{u}) = \mathbf{V}^{\mathrm{T}}\mathbf{Q}\mathbf{M}\mathbf{Q}^{\mathrm{T}}\mathbf{U}$ with $\mathbf{P}^{\mathrm{T}}\mathbf{M}\mathbf{R} = (p, r)$ for all $p, r \in \Phi_h$. By definition of ∇_h,

$$(\nabla \cdot \mathbf{v}, p) = (\nabla_h \cdot \mathbf{v}, p) \text{ for all } \mathbf{v} \in \mathbf{W}_h,\ p \in \Phi_h$$

$$= \mathbf{V}^{\mathrm{T}}\mathbf{Q}\mathbf{M}\mathbf{P}.$$

Thus $\mathbf{C} = \mathbf{Q}\mathbf{M}$ and $\mathbf{B} = \mathbf{Q}\mathbf{M}\mathbf{Q}^{\mathrm{T}}$ giving $\mathbf{C}\mathbf{M}^{-1}\mathbf{C}^{\mathrm{T}} = \mathbf{Q}\mathbf{M}\mathbf{M}^{-1}\mathbf{M}\mathbf{Q}^{\mathrm{T}} = \mathbf{Q}\mathbf{M}\mathbf{Q}^{\mathrm{T}} = \mathbf{B}$ and the result is proved.

APPENDIX II: PERTURBATION ANALYSIS

1. *The unperturbed system*

We consider the solution of the $N+M$ square symmetric system

$$\begin{bmatrix} \mathbf{K} & \mathbf{C} \\ \mathbf{C}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{pmatrix} \mathbf{U} \\ \mathbf{P} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{g} \end{pmatrix} \tag{A1}$$

(see (4)), given that there exist two M-vectors $\mathbf{P}_H$, $\mathbf{P}_c$ (the pure hydrostatic and CB modes respectively) such that

$$\mathbf{C}\mathbf{P}_H = \mathbf{C}\mathbf{P}_c = \mathbf{0}.$$

The associated eigenvalue problem (see (16)) therefore has two zero eigenvalues ($\lambda_1 = \lambda_2 = 0$), the remaining $N+M-2$ eigenvalues being real and bounded strictly away from zero. Let $(\mathbf{w}_i, \mathbf{r}_i)^T$, $i = 1, 2, \ldots, N+M$, denote an orthogonal system of eigenvectors normalized so that $\|\mathbf{w}_i\|^2 + \|\mathbf{r}_i\|^2 = 1$. The eigenvectors corresponding to the zero eigenvalues are therefore

$$\begin{pmatrix}\mathbf{w}_1\\ \mathbf{r}_1\end{pmatrix} = \begin{pmatrix}\mathbf{0}\\ \mathbf{P}_H\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}\mathbf{w}_2\\ \mathbf{r}_2\end{pmatrix} = \alpha\begin{pmatrix}\mathbf{0}\\ \mathbf{P}'_c\end{pmatrix}; \qquad \alpha = \|\mathbf{P}'_c\|^{-1} \tag{A2}$$

where $\mathbf{P}'_c = \mathbf{P}_c - (\mathbf{P}_H, \mathbf{P}_c)\mathbf{P}_H$ denotes a CB mode with an adjusted hydrostatic level (to ensure orthonormality; note that the 'conventional' CB and hydrostatic eigenvectors are not orthogonal).

A necessary and sufficient condition that (A1) has at least one solution is that the right hand side vector has zero projection into the null space of the coefficient matrix. Since the null space is spanned by the vectors defined in (A2), this condition is equivalent to

$$(\mathbf{g}, \mathbf{P}_H) = (\mathbf{g}, \mathbf{P}_c) = 0 \tag{A3}$$

as previously mentioned in the text. Provided $\mathbf{g}$ satisfies (A3), the solution of (A1) may be written in the form

$$\begin{pmatrix}\mathbf{U}\\ \mathbf{P}\end{pmatrix} = \begin{pmatrix}\tilde{\mathbf{U}}\\ \tilde{\mathbf{P}}\end{pmatrix} + \gamma_H\begin{pmatrix}\mathbf{0}\\ \mathbf{P}_H\end{pmatrix} + \gamma_c\begin{pmatrix}\mathbf{0}\\ \mathbf{P}'_c\end{pmatrix} \tag{A4}$$

where

$$\begin{pmatrix}\tilde{\mathbf{U}}\\ \tilde{\mathbf{P}}\end{pmatrix} = \sum_{j=3}^{N+M} \lambda_j^{-1}[(\mathbf{f}, \mathbf{w}_j) + (\mathbf{g}, \mathbf{r}_j)]\begin{pmatrix}\mathbf{w}_j\\ \mathbf{r}_j\end{pmatrix}$$

and γ_H, γ_c are arbitrary constants.

2. *The perturbed system*

Let ε be a small parameter and perturb system (A1) to

$$\begin{bmatrix}\mathbf{K}_\varepsilon & \mathbf{C}_\varepsilon\\ \mathbf{C}_\varepsilon^T & \mathbf{0}\end{bmatrix}\begin{pmatrix}\mathbf{U}_\varepsilon\\ \mathbf{P}_\varepsilon\end{pmatrix} = \begin{pmatrix}\mathbf{f}_\varepsilon\\ \mathbf{g}_\varepsilon\end{pmatrix} \tag{A5}$$

such that

$$\|\mathbf{K}_\varepsilon - \mathbf{K}\| = O(\varepsilon), \|\mathbf{K}_\varepsilon^{-1} - \mathbf{K}^{-1}\| = O(\varepsilon), \mathbf{C}_\varepsilon = \mathbf{C} + \varepsilon\Gamma \tag{A6}$$

and the $N \times M$ matrix Γ (independent of ε) is such that $\Gamma\mathbf{P}_H = \mathbf{0}$, $\Gamma\mathbf{P}_c \neq \mathbf{0}$. We shall denote by λ_i^ε and $(\hat{\mathbf{w}}_i, \hat{\mathbf{r}}_i)^T$ the eigenvalues and eigenvectors of the perturbed coefficient matrix in (A5). In contrast to the unperturbed system, we now have only one zero eigenvalue ($\gamma_1 = 0$), the corresponding eigenvector being $(\mathbf{w}_1, \mathbf{r}_1)^T = (\mathbf{0}, \mathbf{P}_H)^T$.

Our goal is to construct a solution of (A5) comparable to (A4), but first we must analyse the perturbed eigen problem. Estimates of the perturbed eigenvalues can be deduced from the following result (the estimate of λ_2^ε is preliminary and will be refined later):

Theorem 1 [Isaacson and Keller (see Part 2, Reference 39, p. 141)]

Let $\mathbf{A}$ be an Hermitian matrix of order n and have eigenvalues $\{\mu_i\}$. If

$$\mathbf{A}\mathbf{x} - \mu\mathbf{x} \equiv \boldsymbol{\eta}, \qquad \mathbf{x} \neq \mathbf{0} \tag{A7}$$

then

$$\min_i |\mu - \mu_i| \leq \|\boldsymbol{\eta}\|/\|\mathbf{x}\|. \tag{A8}$$

To apply this result, we take $\mathbf{A}$ to be the coefficient matrix in (A5), $\mu = \lambda_j$, and $\mathbf{x} = (\mathbf{w}_j, \mathbf{r}_j)^T$. The residual vector $\boldsymbol{\eta}$ then satisfies,

$$\boldsymbol{\eta} = \begin{bmatrix} \mathbf{K}_\varepsilon - \mathbf{K} & \mathbf{C}_\varepsilon - \mathbf{C}^T \\ \mathbf{C}_\varepsilon^T - \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{pmatrix} \mathbf{w}_j \\ \mathbf{r}_j \end{pmatrix}$$

which, by means of the estimates (A6), gives $\|\boldsymbol{\eta}\| = O(\varepsilon)$. Hence,

$$\lambda_j - \lambda_j^\varepsilon = O(\varepsilon), \qquad j = 2, \ldots, N+M \tag{A9}$$

Corresponding estimates of the perturbed eigenvectors follow from

Theorem 2 (*Ibid*, p. 142)

For a Hermitian matrix $\mathbf{A}$ with eigenvalues $\{\mu_i\}$ and corresponding eigenvectors $\{\mathbf{v}_i\}$, let $V_{j,k}(k \geq j)$ denote the linear space spanned by $\mathbf{v}_j, \mathbf{v}_{j+1}, \ldots, \mathbf{v}_k$. If

$$|\mu - \mu_i| \begin{cases} \leq \|\boldsymbol{\eta}\|, & j \leq i \leq k \\ \geq d > 0 & \text{otherwise} \end{cases}$$

where μ, $\boldsymbol{\eta}$ and $\mathbf{x}$ are as in Theorem 1, then

$$\min_{\mathbf{v} \in \mathbf{V}_{j,k}} \|\mathbf{x} - \mathbf{v}\| \leq \|\boldsymbol{\eta}\|/d \tag{A10}$$

Consider an eigenvalue λ_j of the unperturbed system of multiplicity $(k-j+1)$ (so that $\lambda_j = \lambda_{j+1} = \cdots = \lambda_k$, $k \geq j$). The conditions of the theorem can be met by choosing $d = 1/2 \min_{\lambda_i \neq \lambda_j} |\lambda_i - \lambda_j|$ provided that ε is sufficiently small. For each $i, j \leq i \leq k$, take $\mathbf{x} = (\mathbf{w}_i, \mathbf{r}_i)^T$ so that $\|\boldsymbol{\eta}\| = O(\varepsilon)$. We can therefore construct a linearly independent set of vectors $(\hat{\mathbf{w}}_i, \hat{\mathbf{r}}_i)^T$, the points at which the minimum in (A10) is achieved, such that

$$\begin{pmatrix} \hat{\mathbf{w}}_i \\ \hat{\mathbf{r}}_i \end{pmatrix} = \begin{pmatrix} \mathbf{w}_i \\ \mathbf{r}_i \end{pmatrix} + O(\varepsilon). \tag{A11}$$

This result enables us to calculate the projection of the solution onto the linear space spanned by the eigenvectors $(\hat{\mathbf{w}}_i, \hat{\mathbf{r}}_i)^T$, $i = 3, 4, \ldots, N+M$:

$$\begin{aligned} \begin{pmatrix} \tilde{\mathbf{U}}_\varepsilon \\ \tilde{\mathbf{P}}_\varepsilon \end{pmatrix} &= \sum_{j=3}^{N+M} (\lambda_j^\varepsilon)^{-1} [(\mathbf{f}_\varepsilon, \hat{\mathbf{w}}_j) + (\mathbf{g}_\varepsilon, \hat{\mathbf{r}}_j)] \begin{pmatrix} \hat{\mathbf{w}}_j \\ \hat{\mathbf{r}}_j \end{pmatrix} \\ &= \begin{pmatrix} \tilde{\mathbf{U}} \\ \tilde{\mathbf{P}} \end{pmatrix} + O(\varepsilon) \end{aligned}$$

since, for $j \geq 3$, $|\lambda_j^\varepsilon| \geq \frac{1}{2} \min_{3 \leq i \leq N+M} |\lambda_i| > 0$ when ε is sufficiently small.

We now refine the estimate for λ_2^ε and the corresponding eigenvector. To this end we apply the theorems with the test vector $\mathbf{x} = (\varepsilon \mathbf{w}, \mathbf{P}_c' + \varepsilon \mathbf{r})^T$ and $\mu = \lambda_2 = 0$ where $(\mathbf{w}, \mathbf{r})^T$ constitutes any solution of the equations

$$\mathbf{K}\mathbf{w} + \mathbf{C}\mathbf{r} = -\mathbf{c}$$
$$\mathbf{C}^T\mathbf{w} = \mathbf{0},$$

where $\mathbf{c} \equiv \boldsymbol{\Gamma}\mathbf{P}_c'$ (see (26) and (A6)); note that $\mathbf{c}$ can also be expressed as $\mathbf{c} = \mathbf{C}_\varepsilon \mathbf{P}_c'$ since $\mathbf{C}\mathbf{P}_c' = \mathbf{0}$. Note that we may take

$$\begin{pmatrix} \mathbf{w} \\ \mathbf{r} \end{pmatrix} = -\sum_{j=3}^{N+M} \lambda_j^{-1}(\mathbf{c}, \mathbf{w}_j) \begin{pmatrix} \mathbf{w}_j \\ \mathbf{r}_j \end{pmatrix}$$

Now

$$\boldsymbol{\eta} = \varepsilon \begin{bmatrix} (\mathbf{K}_\varepsilon - \mathbf{K})\mathbf{w} + \varepsilon \boldsymbol{\Gamma}\mathbf{r} \\ \varepsilon \boldsymbol{\Gamma}^T \mathbf{w} \end{bmatrix}$$

so that $\|\boldsymbol{\eta}\| = O(\varepsilon^2)$, giving $\lambda_2^\varepsilon = O(\varepsilon^2)$ and

$$\begin{pmatrix} \hat{\mathbf{w}}_2 \\ \hat{\mathbf{r}}_2 \end{pmatrix} = \beta \begin{pmatrix} \varepsilon \mathbf{w} \\ \mathbf{P}_c' + \varepsilon \mathbf{r} \end{pmatrix} + O(\varepsilon^2)$$

where β ($\beta = 1 + O(\varepsilon)$) is a normalizing constant.

The solution of (A5) may now be expressed in the form

$$\begin{pmatrix} \mathbf{U}_\varepsilon \\ \mathbf{P}_\varepsilon \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{U}} \\ \tilde{\mathbf{P}} \end{pmatrix} + \gamma_H \begin{pmatrix} \mathbf{0} \\ \mathbf{P}_H \end{pmatrix} + \gamma_c' \begin{pmatrix} \mathbf{w} \\ \varepsilon^{-1}\mathbf{P}_c' + \mathbf{r} \end{pmatrix}$$

where γ_H is an arbitrary constant and

$$\gamma_c' = (\varepsilon^2/\lambda_2^\varepsilon)\{(\mathbf{f}_\varepsilon, \mathbf{w}) + (\mathbf{g}_\varepsilon, \mathbf{r})\} + O(\varepsilon)$$

is a fixed constant whose magnitude is, in general, $O(1)$.

Finally, we rederive (27) in a more formal way:

(1) Define $\boldsymbol{\Gamma}$ via $\mathbf{C}_\varepsilon = \mathbf{C} + \varepsilon\boldsymbol{\Gamma}$ (see (A6))

(2) Form $\mathbf{C}_\varepsilon^T \mathbf{U}_\varepsilon = \mathbf{0} = \mathbf{C}^T \mathbf{U}_\varepsilon + \varepsilon \boldsymbol{\Gamma}^T \mathbf{U}_\varepsilon$

(3) Use $\mathbf{C}\mathbf{P}_c = \mathbf{0}$ to form

$\mathbf{C}_\varepsilon \mathbf{P}_c = \varepsilon \boldsymbol{\Gamma}\mathbf{P}_c \equiv \varepsilon \mathbf{c}$ where $\mathbf{c}$ has elements $O(1)$ and

$\mathbf{P}_c^T \mathbf{C}_\varepsilon^T \mathbf{U}_\varepsilon = 0 = \varepsilon \mathbf{P}_c^T \boldsymbol{\Gamma}^T \mathbf{U}_\varepsilon = \varepsilon \mathbf{c}^T \mathbf{U}_\varepsilon$, which is (27) in the text.

REFERENCES

1. C. Taylor and P. Hood, 'A numerical solution of the Navier–Stokes equations using FEM technique, *Computers and Fluids*, **1**, 73–100 (1973).
2. P. Hood and C. Taylor, 'Navier–Stokes equations using mixed interpolation', in *Int. Symp. on Finite Element Method in Flow Problems*, Proceedings, Swansea, Wales (1974).
3. M. Olson and S. Tuann, 'Primitive variables versus stream function finite element solutions of the Navier–Stokes equations', in *Finite Elements in Fluids*, Volume 3, Wiley, Chichester, 1978, pp. 73–89.
4. G. D. Richards, 'Finite elements for incompressible flow', *M.Sc. Dissertation*, Dept. Math., University of Reading, UK (1978).
5. J. H. Argyris, P. C. Dunne, T. Angelopoulos and B. Bichat, 'Large natural strains and some special difficulties due to non-linearity and incompressibility in finite elements', *Comp. Methods Appl. Mech. and Eng.*, **4**, 219–278 (1974).
6. J. C. Nagtegaal, D. M. Parks and J. R. Rice, 'On numerically accurate finite element solutions in the fully plastic range,' *Comp. Methods Appl. Mech. and Eng.*, **4**, 153–177 (1974).
7. O. R. Fabayo, 'Bilinear finite elements for incompressible flow', *M.Sc. Dissertation*, Dept. Math., University of Dundee, Scotland (1977).
8. P. A. Huyakorn, C. Taylor, R. L. Lee and P. M. Gresho, 'A comparison of various mixed interpolation finite elements in the velocity-pressure formulation of the Navier–Stokes equations', *Computers and Fluids*, **6**, 25–35 (1978).

9. T. J. R. Hughes, W. K. Liu and A. Brooks, 'Finite elements analysis of incompressible viscous flows by the penalty formulation', *J. Comp. Phys.*, **30,** 1–15 (1979).
10. R. L. Lee, P. M. Gresho and R. L. Sani, 'Smoothing techniques for certain primitive variable solutions of the Navier–Stokes equations', *Int. J. num. Meth. Engng.* **14,** 1785–1804 (1979).
11. M. Fortin, 'Numerical solution of steady state Navier–Stokes equations', in *Numerical Methods in Fluid Dynamics* (Ed. J. J. Smolderen), Agard Lecture Series No. 48, AGARD–LS–48 (1972).
12. W. E. Pracht and J. V. Brackbill, 'BAAL: A code for calculating three-dimensional fluid flows at all speeds with an Eulerian–Lagrangian computing mesh', *Los Alamos Scientific Laboratory Report LA–6342* (1976).
13. A. Chorin, personal communication and, 'On the convergence of discrete approximations to the Navier–Stokes equations', *Math. of Comp,* **23,** 341–353 (1969).
14. R. Sani, P. Gresho and R. Lee, 'On the spurious pressures generated by certain GFEM solutions of the incompressible Navier–Stokes equations', *Third Int. Conf. on Finite Elements in Flow Problems,* Proceedings, Banff, Alberta, Canada (1980).
15. C. Caldwell, Westinghouse Electric Corporation, Pittsburgh, PA, private communication (1979).
16. M. Bercovier and M. Engelman, 'A finite element for the numerical solution of viscous incompressible flows', *J. Comp. Phys.*, **30,** 181–201 (1979).
17. P. M. Gresho, R. L. Lee and R. L. Sani, 'On the time-dependent solution of the incompressible Navier–Stokes equations in two and three dimensions', in *Recent Advances in Numerical Methods in Fluids,* Pineridge Press, Swansea, U.K., 1980.
18. D. S. Malkus and T. J. R. Hughes, 'Mixed finite element methods—reduced and selective integration techniques: A unification of concepts', *Comp. Methods Appl. Mech. and Eng.*, **15,** 63–81 (1978).
19. M. Bercovier, R. Sani, P. Gresho and M. Engelman, in preparation.
20. M. Fortin, 'An analysis of the convergence of mixed finite element methods', *RAIRO,* **11,** R–3, 341–354 (1977).

CONTENTS OF PART 2

The remainder of this paper (Part 2), to appear in the next issue of the journal, contains:

(1) Generalization to Other Elements
(2) Remarks on Earlier Theory and the Penalty Approach
(3) Numerical Examples
(4) Discussion
(5) Conclusions

ACKNOWLEDGEMENTS

This work was performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore National Laboratory under contract No. W–7405–Eng–48. RLS would like to acknowledge support from both Lawrence Livermore National Laboratory and the U.S. Army Research Office (Grant DAA629–79–6–0045).

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately-owned rights.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.