NASA Contractor Report 195293

Analysis and Control of Hourglass Instabilities in Underintegrated Linear and Nonlinear Elasticity

Olivier P. Jacquotte and J. Tinsley Oden
The University of Texas at Austin
Austin, Texas

March 1994

Prepared for
Lewis Research Center
Under Grant NAG3–329

National Aeronautics and
Space Administration

ANALYSIS AND CONTROL OF HOURGLASS INSTABILITIES IN UNDERINTEGRATED LINEAR AND NONLINEAR ELASTICITY*

Olivier P. Jacquotte and J. Tinsley Oden
Texas Institute for Computational Mechanics
The University of Texas at Austin
Austin, Texas, U.S.A.

July 1985

TICOM Report 85-10

*This work was supported by Grant NAG3-329 from the Lewis Research Center, NASA.

PART I: INTRODUCTION

1.1 General

Among various choices involved in applying the finite element method for the resolution of problems of fluid or solid mechanics, the choice of the element type associated with a given mesh and that of the integration rule used on these elements often decide the properties of the approximation of the operator to be discretized. A bad approximation generally results in a kernel larger than the kernel of the continuous operator, and these spurious elements of this kernel may appear in the discrete solution. This solution then can exhibit undesirable oscillations and instabilities.

In particular, for the general class of problem:

$$Au - B^*p = f$$

$$Bu = 0$$

oscillations and instabilities may arise from a bad approximation of the governing operator A or of the constraint operator B.

As far as the constraint is conerned, numerous theoretical and numerical studies of pressure instabilities have been done over the last decade [7, 10, 33, 36, 45]. In these studies, it has been proven that a key stability condition, the L.B.B. condition, must be satisfied in order to have existence and uniqueness of a solution [31, 9, 2], and that the constant appearing in this condition must be independent of the mesh size in order to have stability of the element considered [39, 40, 41]. Several stable and unstable elements have been studied with respect to the L.B.B. condition [19, 20, 42,43, 44].

Our purpose in this report is to concentrate on a bad finite element approximation of the governing operator obtained when underintegration is used in numerical code for several model problems: the Poisson problem, the linear elasticity problem, and for problems in the nonlinear theory of elasticity. For each of these problems, the reasons for the occurrence of instabilities will be given, way to control or eliminate them will be presented, and theorems of existence, uniqueness and convergence for the given methods will be established. Finally, numerical results are given which illustrate the theory.

1.2 Major Results

Historical background as well as precise definitions and notations will be given in the introductions of Parts II and III. However, we list here the major results we have obtained:

1) In underintegrated finite element methods, a rank-deficiency counting for each element cannot work to predict the exact number of spurious modes of the gobal stiffness matrix.

2) When the spurious modes are known exactly, it is posssible to eliminate them a-posteriori and obtain a unique, stable and accurate solution.

3) Even when the spurious zero-energy modes cannot be predicted from rank-deficiency observations, oscillations can be predicted from the eigenvalue analysis of the stiffness matrix.

4) Concentrated forces are the most effective source of oscillations; a way to handle such loads without exciting oscillations is theoretically obtained and its numerical implementation turns out to be efficient.

5) In linear elasticity, the spurious modes and their construction is described.

6) Stabilization methods proposed by several authors may not work.

7) For a general mesh, an element-by-element spurious mode control is presented, as well as its numerical results: this control is efficient and leads to an accurate solution.

8) This control is cheap, easy to implement, and can be used independently of the material.

9) In nonlinear incompressible elasticity, several numerical results are presented and confirm that the results previously listed can be extended to nonlinear problems.

1.3 Report Organization

Part II is devoted to the analysis and control of hourglass modes in underintegrated finite element methods. The framework of the theoretical study is the simple two-dimensional Poisson problem solved using bilinear elements; the comparison involves the 4- and 1-integration point rules. We prove that for some boundary conditions, the stiffness matrix is rank-deficient and spurious modes appear, but we also prove that the solution obtained from the underintegrated problem can be processed a-posteriori in order to obtain a stable and convergent solution in which the spurious modes have been eliminated. Moreover, the method of proof allows us to demonstrate precisely why spurious modes are excited, even though they cannot be predicted by a rank-deficiency of the stiffness matrix. Numerical experiments are discussed which not only illustrate the theory, but also generalize the results to various elements (8- and 9-node elements) and operators (linear elasticity operator).

In Part III, we numerically investigate the behavior of underintegrated 8- and 9-node elements associated with linear discontinuous pressures for the analysis of problems in finite elasticity. We observe that whereas the 9-node element is stable, the 8-node element exhibits pressure oscillations. We study the performance of the control presented in Part II to the control of oscillations appearing in the finite element solution of rubber elasticity when underintegration is practiced.

PART II: ANALYSIS OF INSTABILITIES IN UNDERINTEGRATED FINITE ELEMENT METHODS

2.1 Introduction

For many years, a special type of numerical instability has been observed in finite difference approximations of flow fields, which has been referred to as "hourglassing", "keystoning", or "chickenwiring". These graphic terms refer to geometrical patterns which appear in computed flow fields (e.g. velocities) and which emerge as spurious oscillations superimposed on an otherwise smooth field, the spurious oscillations often taking a zig-zag form which resembles an hourglass or a chickenwire mesh. These spurious modes can be amplified upon refining the mesh, and to control such numerical instabilities, various schemes for incorporating "hourglass viscosity" or "hourglass damping" have been proposed by some authors.

It is now known that hourglass modes can arise from an incomplete (or poor) approximation of the kernel of the operators in the momentum equations in flow or solid mechanics problems (or, more generally, of the principal part of the operator in the governing differential equation of a given boundary-value problem). For example, in addition to the rigid body motions residing in the kernel of the standard operators appearing in the equilibrium (momentum) equations of solid and fluid mechanics, one finds hourglass modes in various crude discrete models of these operators.

In recent years, the occurrence of hourglass instabilities in

underintegrated finite element approximations has been observed. In the implementation of most finite element methods, integrals defining stiffness matrices are evaluated using numerical quadrature schemes. To improve computational efficiency, the practice of *underintegration* is often employed, by which is meant the use of a quadrature rule of an order lower than that required to integrate polynomial integrands exactly. This can produce rank-deficient stiffness matrices or, equivalently, an expanded kernel of the equilibrium operation which contains spurious hourglass modes, and the result is again a numerically unstable scheme.

In order to overcome this difficulty, artificial stiffness or viscosity methods, or other stabilization methods have been proposed by several authors (e.g. [3-6, 18, 30]). These methods involve computing an underintegrated matrix, and then adding a stabilization matrix which effectively eliminates the hourglass modes. They turn out to be fairly general and have been used for a long time in numerous codes. Whereas all of these methods based on intuitive feeling give good numerical results, their mathematical study remains often non-existent.

The most interesting challenge is to solve the problem using only the crude rank-deficient underintegrated stiffness matrix, the solution is obtained up to within an arbitrary spurious mode, and then to eliminate these modes from the solution in a post-processing operation.

Unfortunately, even when the stiffness matrix is rank-sufficient, similar oscillations are observed when underintegration is used. In that case, the process of the excitations of modes similar to the hourglass modes is not completely understood and these modes have never been

mathematically studied.

In this report, we give precise mathematical justifications and answers to the questions previously mentioned. The next Section (Section 2.2) is devoted to the proof that the Stabilization method is mathematically justified. Then, in Section 2.3, we present a method which involves solving an underintegrated and not well-posed problem, then in a-posteriori eliminating the unknown degree of freedom. The proof of the accuracy of the method is given in Section 2.4, and its numerical aspects and results are described in Section 2.5. In Section 2.6, we examine the case in which spurious oscillations cannot be predicted from the rank-deficiency of the stiffness matrix and we analyze why these modes may be excited. Finally, we apply the previous considerations to the linear elasticity problem.

It should be noted that the method and its results cannot be embedded in a classical elliptic theory: Strang's ellipticity condition [49] is here violated and this non-elliptic method cannot be studied by the classical theory of finite element methods and numerical integration [11, 12, 51, 52]. Note in these references that, both Ciarlet and Wahlbin crucially suppose the exactitude of the numerical scheme for the polynomials considered. This polynomial invariance plays a decisive role in their error estimations. We also refer to Girault [21, 22] for his approach to the same kind of problem, non-elliptic because of the use of partially underintegrated stiffness matrix, but where hourglass modes did not appear.

2.2 A-Priori Hourglass Control

2.2.1 Introduction. This section is devoted to mathematical preliminaries to several methods consisting of adding a stabilization matrix to the underintegrated matrix. For clarity, we shall confine our attention to a simple model problem. Let Ω be a regular domain in $\mathbb{R}^2$ with boundary $\partial\Omega$ and consider the model Neumann problem,

(P_0) Find $u = u(x,y)$ such that

$$\left.\begin{aligned} -\Delta u &= f && \text{on } \Omega \\ \frac{\partial u}{\partial n} &= 0 && \text{on } \partial\Omega \end{aligned}\right\} \tag{2.1}$$

where f is an $L^2(\Omega)$-function satisfying

$$\int_\Omega f \, dxdy = 0 \tag{2.2}$$

The questions of the existence and uniqueness of solutions to (2.1) (which are well-known) are taken up in Part II.

We shall first consider a finite element approximation of (2.1) constructed using Q_1-elements, i.e., four-node quadrilateral elements over which bilinear shape functions are used. Most of our notations and results are reproductions of those of Flanagan [18] and Belytschko [3, 5, 6]. Then we will attempt to extend our results to the Q_2-elements (nine-node, biquadratic elements) and will indicate in which ways they differ from Belytschko's [4].

The construction of finite element approximations of (2.1) involves the calculation of the stiffness matrix $\underset{\sim}{K}_e$ for a typical finite ele-

ment Ω_e , which is given by the formula,

$$\underset{\sim}{K}_e = \int_\Omega \nabla \underset{\sim}{N}^t \cdot \nabla \underset{\sim}{N} \, dxdy \tag{2.3}$$

where $\underset{\sim}{N}$ is a vector representing the bilinear or biquadratic shape functions in each element Ω_e , $1 \leq e \leq E$.

When Q_1-(respectively Q_2-) elements are used to discretize the domain Ω , $\underset{\sim}{K}_e$ is a 4x4 matrix (resp. 9x9) and the $\underset{\sim}{N}$'s contain four bilinear (resp. nine biquadratic) shape functions. We will distinguish exact-, full-, and under-integrations. The full integration is obtained using the number of Gauss integration points necessary to obtain the exact integration on regular square elements: 4 (resp. 9) points in our study. The underintegration will involve the Gauss rule of lower order: 1 (resp. 4) points. The stiffness matrix associated with a rule involving k points will be denoted $\underset{\sim}{K}_e^{(k)}$, $k = 1,4,9$.

Several authors [3, 5, 6, 18, 30] proposed to add to the underintegrated stiffness matrix a stabilization matrix which exhibits several special properties. In this section, we will prove that these properties are indeed satisfied and that the exact stiffness matrix $\underset{\sim}{K}_e$ can be computed by this method. This will be accomplished by first carrying out the integration (2.3) exactly.

We first introduce some notations. Suppose that element Ω_e is defined by the coordinates of its nodes (x^I, y^I), $1 \leq I \leq p$, $p = 4$ or 9 . We introduce the isoparametric mapping from a master element

$$\hat{\Omega} = \left[-\frac{1}{2}, +\frac{1}{2}\right] \times \left[-\frac{1}{2}, +\frac{1}{2}\right]$$

to Ω_e such that

$$\left.\begin{aligned} x &= \sum_{I=1}^{P} x^I N_I(\xi,\eta) \\ y &= \sum_{I=1}^{P} y^I N_I(\xi,\eta) \end{aligned}\right\} \tag{2.4}$$

where N_I , $1 \leq I \leq p$, are the shape functions for the quadrilateral element on the master element. The node numbering convention is shown in Figure 1.

The stiffness matrix $\underset{\sim}{K}_e$ is evaluated in (2.3) using the mapping (2.4) from $\hat{\Omega}$ to Ω_e :

$$\left.\begin{aligned} \underset{\sim}{K}_e &= \int_{\Omega} \underset{\sim}{\nabla N}^T \underset{\sim}{\nabla N}\, dxdy \\ &= \int_{\hat{\Omega}} |J|\, \underset{\sim}{\hat{\nabla} N}^T \left[\frac{d\xi}{dx}\right]^T \left[\frac{d\xi}{dx}\right] \underset{\sim}{\hat{\nabla} N}\, d\xi d\eta \end{aligned}\right\} \tag{2.5}$$

where $\left[\frac{d\xi}{dx}\right]$ is the Jacobian matrix of the mapping from $\hat{\Omega}$ to Ω_e , J is the inverse of its determinant, and where the gradients of the shape functions are derived with respect to the master element coordinates (ξ,η) . These matrices can be computed using (2.4) :

$$\left.\begin{aligned} x &= \underset{\sim}{x}^T \cdot \underset{\sim}{N} \\ y &= \underset{\sim}{y}^T \cdot \underset{\sim}{N} \end{aligned}\right\} \tag{2.6}$$

$$\left[\frac{dx}{d\xi}\right] = \begin{bmatrix} \underset{\sim}{x}^T \frac{dN}{d\xi} & \underset{\sim}{y}^T \frac{dN}{d\xi} \\ \underset{\sim}{x}^T \frac{dN}{d\eta} & \underset{\sim}{y}^T \frac{dN}{d\eta} \end{bmatrix} \tag{2.7}$$

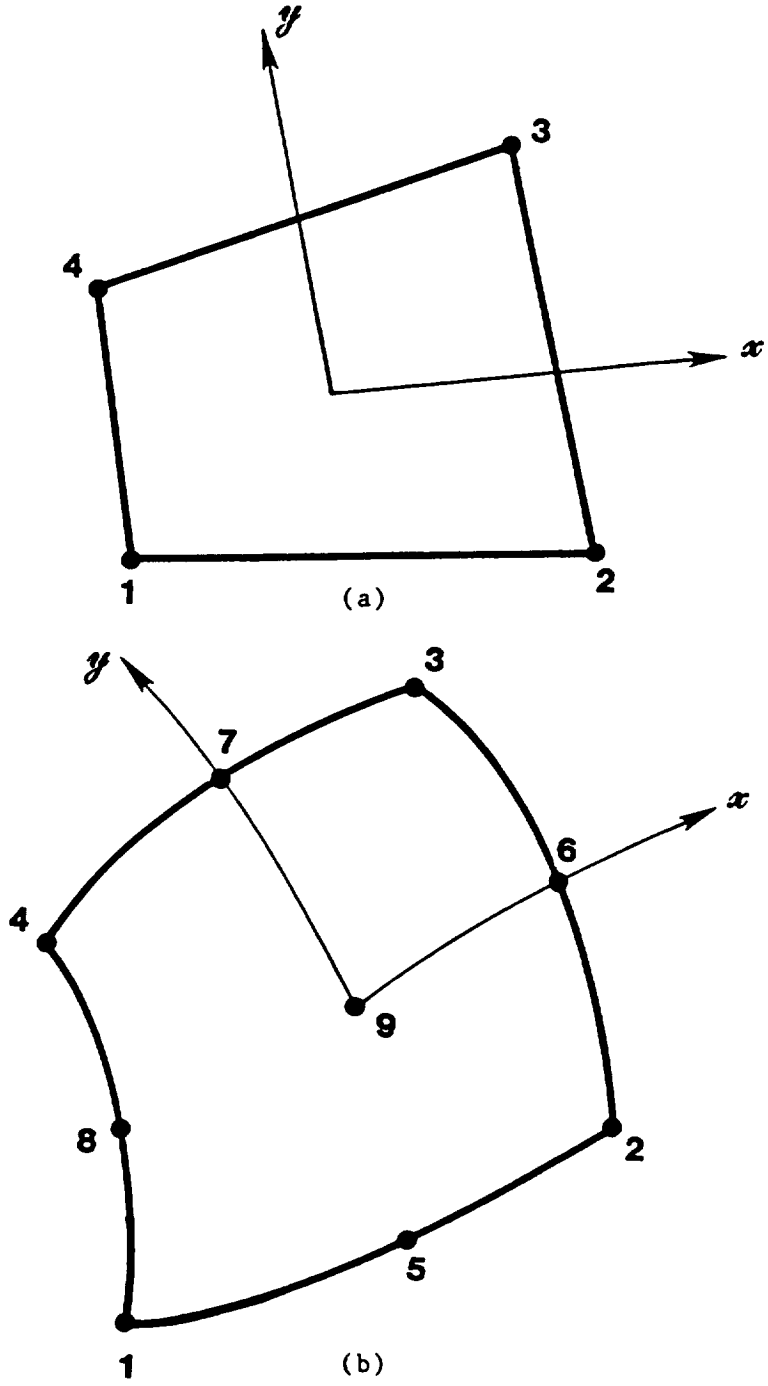


Figure 1. Node Numbering for a) 4-Node Element; b) 9-Node Element.

$$\left[\frac{d\xi}{dx}\right] = \left[\frac{dx}{d\xi}\right]^{-1} = \frac{1}{J}\begin{bmatrix} y^T \frac{dN}{d\eta} & -y^T \frac{dN}{d\xi} \\ -x^T \frac{dN}{d\eta} & x^T \frac{dN}{d\xi} \end{bmatrix} \tag{2.8}$$

$$\hat{\nabla}N = \begin{bmatrix} \frac{dN}{d\xi}^T \\ \frac{dN}{d\eta}^T \end{bmatrix} \tag{2.9}$$

where J is the Jacobian of the mapping

$$J = \det \frac{dx}{d\xi} \tag{2.10}$$

Finally, we obtain the expression,

$$\underset{\sim}{K}_e = \int_{\hat{\Omega}} \left(\frac{\underset{\sim}{A}\underset{\sim}{x}\underset{\sim}{x}^T\underset{\sim}{A}^T}{\underset{\sim}{y}^T\underset{\sim}{A}\underset{\sim}{x}} + \frac{\underset{\sim}{A}\underset{\sim}{y}\underset{\sim}{y}^T\underset{\sim}{A}^T}{\underset{\sim}{y}^T\underset{\sim}{A}\underset{\sim}{x}} \right) d\xi d\eta \tag{2.11}$$

where $\underset{\sim}{A}$ is the antisymmetric matrix

$$\underset{\sim}{A} = \frac{dN}{d\eta}\frac{dN}{d\xi}^T - \frac{dN}{d\xi}\frac{dN}{d\eta}^T \tag{2.12}$$

the Jacobian J can be expressed as $\underset{\sim}{y}^T\underset{\sim}{A}\underset{\sim}{x}$.

A study of $\underset{\sim}{K}_e$ expressed as in (2.11) and the properties of the matrix $\underset{\sim}{A}$ will then enable us to study the effect of the underintegration of the stiffness matrix. We will first concentrate on the 4 node-element and derive the exact expression of the stabilization matrix. Then we will discuss what form this matrix may take for the 9-node element.

2.2.2. The stabilization matrix for the bilinear element. For the bilinear element, the shape function vector can be written as

$$\underset{\sim}{N} = \frac{1}{4}\underset{\sim}{t} - \frac{\xi}{2}\underset{\sim}{s} + \frac{\eta}{2}\underset{\sim}{s'} + \xi\eta\,\underset{\sim}{h} \tag{2.13}$$

where

$$\begin{aligned} \underset{\sim}{s}^T &= (1, -1, -1, 1) \\ \underset{\sim}{s'}^T &= (1, 1, -1, -1) \\ \underset{\sim}{t}^T &= (1, 1, 1, 1) \\ \underset{\sim}{h}^T &= (1, -1, 1, -1) \end{aligned} \tag{2.14}$$

then the explicit form of the (4x4) matrix A is

$$\underset{\sim}{A} = \frac{1}{4}(\underset{\sim}{s'}\underset{\sim}{s}^T - \underset{\sim}{s}\underset{\sim}{s'}^T) + \frac{\xi}{2}(\underset{\sim}{s}\underset{\sim}{h}^T - \underset{\sim}{h}\underset{\sim}{s}^T) + \frac{\eta}{2}(\underset{\sim}{h}\underset{\sim}{s'}^T - \underset{\sim}{s'}\underset{\sim}{h}^T) \tag{2.15}$$

for $(\xi,\eta) = (0, 0)$, we obtain A_0 which satisfies

$$\underset{\sim}{y}^T\underset{\sim}{A}_0\underset{\sim}{x} = \underset{\sim}{y}^T\underset{\sim}{A}\Big|_{\xi=\eta=0}\underset{\sim}{x} = \frac{1}{2}(y_{24}x_{13} + y_{31}x_{24}) \tag{2.16}$$

which is merely the area of the element Ω_e, noted $|\Omega_e|$.

At this point, we can define precisely the matrix resulting from a 1-point rule; this underintegrated matrix, denoted by $\underset{\sim}{K}_e^{(1)}$, is given by

$$\underset{\sim}{K}_e^{(1)} = \frac{\underset{\sim}{A}_0\underset{\sim}{x}\underset{\sim}{x}^T\underset{\sim}{A}_0^T}{|\Omega_e|} + \frac{\underset{\sim}{A}_0\underset{\sim}{y}\underset{\sim}{y}^T\underset{\sim}{A}_0^T}{|\Omega_e|} \tag{2.17}$$

Also, if we denote $B = (\underset{\sim}{b}_1, \underset{\sim}{b}_2)^T$, the discrete approximations of the gradient $\nabla\underset{\sim}{N}$ evaluated at the integration point is given by

$$\begin{cases} \underset{\sim}{b}_1 = -\underset{\sim}{A}_0\underset{\sim}{y} \\ \underset{\sim}{b}_2 = \underset{\sim}{A}_0\underset{\sim}{x} \end{cases} \tag{2.18}$$

and,therefore,(2.17) takes the usual form

$$\underset{\sim e}{K}^{(1)} = \frac{1}{|\Omega_e|} (\underset{\sim 1}{b}_1 \underset{\sim 1}{b}_1^T + \underset{\sim 2}{b}_2 \underset{\sim 2}{b}_2^T) \tag{2.19}$$

The rank deficiency of $K_e^{(1)}$ can now be verified.

Indeed, from (2.14) and (2.15) we note that

$$\left.\begin{aligned} \underset{\sim 0}{A}\underset{\sim}{h} &= 0 \\ \underset{\sim 0}{A}\underset{\sim}{t} &= 0 \end{aligned}\right\} \tag{2.20}$$

and then

$$\left.\begin{aligned} \underset{\sim e}{K}^1 \underset{\sim}{h} &= 0 \\ \underset{\sim e}{K}^1 \underset{\sim}{t} &= 0 \end{aligned}\right\} \tag{2.21}$$

Therefore, if we consider H and T the global hourglass and translation, and $\underset{\sim}{K}^{(1)}$ the assembled underintegrated stiffness matrix, we have

$$\left.\begin{aligned} \underset{\sim}{K}^{(1)} \cdot H &= \sum_e \underset{\sim e}{K}^{(1)} \cdot H = \sum_e \underset{\sim e}{K}^{(1)} \cdot h = 0 \\ \text{also} \quad \underset{\sim}{K}^{(1)} \cdot T &= 0 \end{aligned}\right\} \tag{2.22}$$

and this proves the rank deficiency of $\underset{\sim}{K}^{(1)}$. Note that this "±1" pattern is independent of the regularity of the mesh and that h will take alternating values +1 and -1 at neighbor nodes as shown in Fig. 2.

Our goal will now be to calculate a matrix K_e^{stab} such that, if added to $\underset{\sim e}{K}^{(1)}$, we obtain the exact stiffness matrix $\underset{\sim e}{K}$ given by (2.11). This expression does not seem easy to integrate, but the image of certain vectors mapped by this matrix can be easily computed using

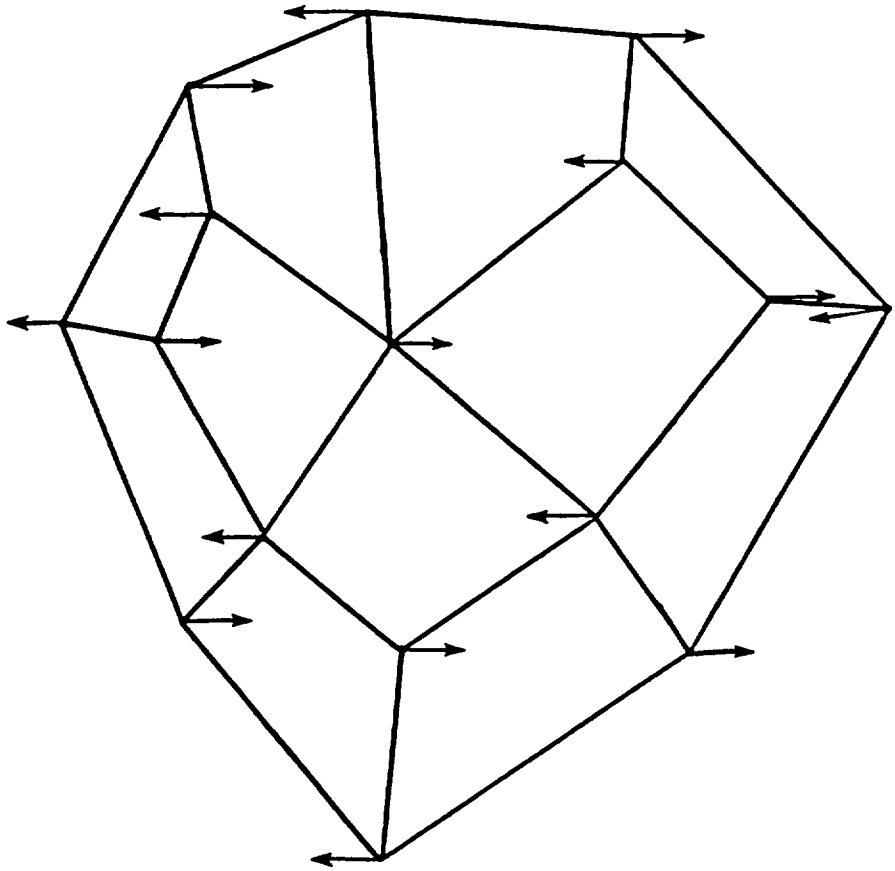


Figure 2. "±1" Pattern of the Hourglass Mode in an Arbitrary Mesh.

orthogonality relations previously obtained (2.20) and the fact that

$$\begin{aligned} \underset{\sim}{b}_1^T \underset{\sim}{x} &= \underset{\sim}{b}_2^T \underset{\sim}{y} = |\Omega_e| \\ \underset{\sim}{b}_1^T \underset{\sim}{y} &= \underset{\sim}{b}_2^T \underset{\sim}{x} = 0 \end{aligned} \tag{2.23}$$

we obtain:

$$\left.\begin{aligned} \underset{\sim e}{K}\, \underset{\sim}{t} &= \underset{\sim}{0} \\ \underset{\sim e}{K}\, \underset{\sim}{x} &= \underset{\sim}{b}_1 \\ \underset{\sim e}{K}\, \underset{\sim}{y} &= \underset{\sim}{b}_2 \end{aligned}\right\} \tag{2.24}$$

Equation (2.24) is not sufficient to compute $\underset{\sim e}{K}$ because it gives only 9 out of the 10 coefficients of $\underset{\sim e}{K}$ (4 x 4, symmetric). It is enough to know $\underset{\sim}{X}^t \underset{\sim e}{K}\, \underset{\sim}{X}$, where $\underset{\sim}{t}$, $\underset{\sim}{x}$, $\underset{\sim}{y}$, and $\underset{\sim}{X}$ form a set of independent vectors. That is the case for $\underset{\sim}{X} = \underset{\sim}{h}$ because

$$\det(\underset{\sim}{x}, \underset{\sim}{y}, \underset{\sim}{t}, \underset{\sim}{h}) = 4A \neq 0$$

provided the element is not singular. Then the knowledge of $\underset{\sim}{h}^T \underset{\sim e}{K}\, \underset{\sim}{h}$ and the relations define uniquely $\underset{\sim e}{K}$. If we set

$$\underset{\sim}{h}^T \underset{\sim e}{K}\, \underset{\sim}{h} = 16\, \bar{\varepsilon} \tag{2.25}$$

then $\underset{\sim e}{K}$ is given by

$$\underset{\sim e}{K} = \underset{\sim e}{K}^{(1)} + \bar{\varepsilon}\, \underset{\sim}{\gamma}\underset{\sim}{\gamma}^T \tag{2.26}$$

whereas, again, given by (2.25), $\bar{\varepsilon}$ is a scalar, and

$$\underset{\sim}{\gamma} = \underset{\sim}{h} - \frac{\underset{\sim}{h}^T \underset{\sim}{x}}{|\Omega_e|}\, \underset{\sim}{b}_1 - \frac{\underset{\sim}{h}^t \underset{\sim}{y}}{|\Omega_e|}\, \underset{\sim}{b}_2 \tag{2.27}$$

While it is difficult to express $\bar{\varepsilon}$ nicely as function of $\underset{\sim}{x}$ and $\underset{\sim}{y}$, its exact value can be written

$$\bar{\varepsilon} = \frac{1}{4}\int_{\hat{\Omega}} \frac{\left[\xi(\underset{\sim}{s}^T\underset{\sim}{x}) - \eta(\underset{\sim}{s}'^T\underset{\sim}{x})\right]^2 + \left[\xi(\underset{\sim}{s}^T\underset{\sim}{y}) - \eta(\underset{\sim}{s}'^T\underset{\sim}{y})\right]^2}{|\Omega_e| + \xi(y_{43}x_{12} + y_{12}x_{34}) + \eta(y_{32}x_{14}+y_{14}x_{23})} \, d\xi d\eta \tag{2.28}$$

We observe that for parallelogram elements, the denominator is constant and its value is the area of the domain $|\Omega_e|$. In this case

$$\bar{\varepsilon} = \frac{1}{24|\Omega_e|}(x_{13}^2 + x_{24}^2 + y_{13}^2 + y_{24}^2) \tag{2.29}$$

or for rectangular elements

$$\bar{\varepsilon} = \frac{\ell_x^2 + \ell_y^2}{12\,\ell_x\ell_y} \tag{2.30}$$

where ℓ_x and ℓ_y are the lengths of the sides of the rectangular element. Also note that for such parallelogram elements $\underset{\sim}{\gamma}$ reduces to $\underset{\sim}{h}$.

The expression (2.26) is often used to eliminate a-priori spurious modes for the kernel of $\underset{\sim}{K}$, but the determination of $\bar{\varepsilon}$ remains a problem. The choice $\bar{\varepsilon} = 0$ leads to the underintegrated matrix and to the method to be studied in the next section. On the other hand, a cheaper way than the full integration of the whole matrix would be to fully integrate $\bar{\varepsilon}$ given by (2.28). This method would lead again to the full integration and is cheaper because it needs only one 4 x 4 integration per element instead of 10. A more common practice is to take for $\bar{\varepsilon}$ a simple value independent of the geometry of the element, which

is often the value obtained for a square 1/6 or sometimes any arbitrary constant, as used in [5, 6].

2.2.3. The stabilization matrix for the biquadratic element. In this section, we will study the effect of the underintegration of the 9 x 9 stiffness matrix obtained in (2.11) with nine-node elements when a 4- Gauss integration point rule is used. Whereas Belytschko et al. [4] intuitively obtain another "$\underset{\sim}{\gamma}.\underset{\sim}{\gamma}^T$" stabilization, we prove that this decomposition is not even valid for regular meshes. We then propose a decomposition derived on regular mesh.

But first we exhibit the spurious modes out of $\underset{\sim}{K}_e^{(4)}$. For the biquadratic element, the shape function vector can be written as

$$\underset{\sim}{N} = \underset{\sim}{S}\ \underset{\sim}{\xi} \tag{2.31}$$

where $\underset{\sim}{S}$ is the 9 x 9 matrix

$$\underset{\sim}{S} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & -2 & -2 & 4 \\ 0 & 0 & 0 & -1 & 0 & 0 & 2 & -2 & 4 \\ 0 & 0 & 0 & 1 & 0 & 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & -1 & 0 & 0 & -2 & 2 & 4 \\ 0 & 0 & -1 & 0 & 0 & 2 & 0 & 4 & -8 \\ 0 & 1 & 0 & 0 & 2 & 0 & -4 & 0 & -8 \\ 0 & 0 & 1 & 0 & 0 & 2 & 0 & -4 & -8 \\ 0 & -1 & 0 & 0 & 2 & 0 & 4 & 0 & -8 \\ 1 & 0 & 0 & 0 & -4 & -4 & 0 & 0 & 16 \end{bmatrix} \tag{2.32}$$

and

$$\underset{\sim}{\xi} = \left[1,\ \xi,\ \eta,\ \xi\eta,\ \xi^2,\ \eta^2,\ \xi\eta^2,\ \xi^2\eta,\ \xi^2\eta^2\right]^T \tag{2.33}$$

The integration rule we are interested in involves four integration points $(\xi^\alpha,\ \eta^\alpha)$, $\alpha = 1,\ 4$. Associated with each of them, we denote

by $\underset{\sim}{A}_\alpha$ and J_α the corresponding matrix $\underset{\sim}{A}$ and Jacobian. The under-integrated matrix is then

$$\underset{\sim}{K}_e^{(4)} = \sum_{\alpha=1}^{4} \frac{\underset{\sim}{A}_\alpha \underset{\sim}{x}\underset{\sim}{x}^T \underset{\sim}{A}_\alpha + \underset{\sim}{A}_\alpha \underset{\sim}{y}\underset{\sim}{y}^T \underset{\sim}{A}_\alpha}{J_\alpha} \tag{2.34}$$

and can also be written

$$\underset{\sim}{K}_e^{(4)} = \sum_{\alpha=1}^{4} \frac{1}{J_\alpha} \left(\underset{\sim}{b}_1^\alpha \underset{\sim}{b}_1^{\alpha T} + \underset{\sim}{b}_2^\alpha \underset{\sim}{b}_2^{\alpha T}\right) \tag{2.35}$$

where

$$\underset{\sim}{B}^\alpha = \begin{bmatrix} \underset{\sim}{b}_1^{\alpha T} \\ \underset{\sim}{b}_2^{\alpha T} \end{bmatrix} = \begin{bmatrix} (-A_\alpha y)^T \\ (A_\alpha x)^T \end{bmatrix} \tag{2.36}$$

generalizes (2.18) to a 4- point rule.

The rank deficiency of $\underset{\sim}{K}_e^{(4)}$ can now be verified. Indeed if we call $\underset{\sim}{t}$ and $\underset{\sim}{h}$ the vectors defined by

$$\begin{aligned} \underset{\sim}{t}^T &= \left[1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1,\ 1\right] \\ \underset{\sim}{h}^T &= \left[1,\ 1,\ 1,\ 1,\ -1,\ -1,\ -1,\ -1,\ 0\right] \end{aligned} \tag{2.37}$$

we easily obtain:

$$\underset{\sim}{t}^T \cdot \underset{\sim}{N} = \underset{\sim}{t}^T \cdot \underset{\sim}{S} \cdot \underset{\sim}{\xi} = 1$$

and

$$\underset{\sim}{h}^T \cdot \underset{\sim}{N} = \underset{\sim}{h}^T \cdot \underset{\sim}{S} \cdot \underset{\sim}{\xi} = -4(\xi^2 + \eta^2 - 12\ \xi^2\eta^2)$$

and then differentiating these expressions and using (2.12) we get:

$$\underset{\sim}{A} \cdot \underset{\sim}{t} = 0$$

$$\underset{\sim}{A} \cdot \underset{\sim}{h} = -8 \left[\eta(1-12\xi^2) \frac{\underset{\sim}{dN}}{d\xi} + \xi(1-12\eta^2) \frac{\underset{\sim}{dN}}{d\eta}\right]$$

the second expression vanishes when the point (ξ,η) is one of the four

Gauss integration points of $\hat{\Omega}$

$$(\xi^{\alpha}, \eta^{\alpha}) = (\pm \frac{1}{2\sqrt{3}}, \pm \frac{1}{2\sqrt{3}}) \; ; \; \alpha = 1, 4 \tag{2.38}$$

Therefore we have

$$\underset{\sim}{A}_{\alpha} \cdot \underset{\sim}{t} = \underset{\sim}{A}_{\alpha} \cdot \underset{\sim}{h} = 0 \qquad \alpha = 1, 4 \tag{2.39}$$

and, consequently,

$$\underset{\sim}{K}_{e}^{(4)} \cdot t = \underset{\sim}{K}_{e}^{(4)} \cdot \underset{\sim}{h} = 0 \tag{2.40}$$

which proves the rank deficiency of $\underset{\sim}{K}_{e}^{(4)}$. Once again, we note that the pattern of $\underset{\sim}{h}$ defined in (2.37) is independent of the geometry of the element and is therefore valid for a rectangular mesh as well as for irregular element meshes.

The analysis of the decomposition of underintegrated and stabilizing matrices for this element cannot be completed as completely as that for the 4-node element. However, Belytschko and co-workers [3] have intuitively arrived at a decomposition similar to (2.26) where $\underset{\sim}{\gamma}$ and ϵ are

$$\underset{\sim}{\gamma} = \underset{\sim}{h} - \frac{1}{4} \underset{\sim}{h}^{T} \cdot \underset{\sim}{x} \sum_{\alpha=1}^{4} \frac{\underset{\sim}{b}_{1}^{\alpha}}{J_{\alpha}} - \frac{1}{4} \underset{\sim}{h}^{T} \cdot \underset{\sim}{y} \sum_{\alpha=1}^{4} \frac{\underset{\sim}{b}_{2}^{\alpha}}{J_{\alpha}} \tag{2.41}$$

$$\epsilon = \frac{1}{100} \sum_{\alpha=1}^{4} \frac{1}{J_{\alpha}} \left[\underset{\sim}{b}_{1}^{\alpha T} \cdot \underset{\sim}{b}_{1}^{\alpha} + \underset{\sim}{b}_{2}^{\alpha T} \cdot \underset{\sim}{b}_{2}^{\alpha} \right] \tag{2.42}$$

This decomposition does in fact satisfy several properties also satisfied by the exact matrix,

$$\underset{\sim}{K}_{e} \cdot \underset{\sim}{t} = 0$$

$$\underset{\sim}{K}_{e} \cdot \underset{\sim}{x} = \sum_{\alpha=1}^{4} \underset{\sim}{b}_{1}^{\alpha}$$

$$\underset{\sim}{K}_e \cdot \underset{\sim}{y} = \sum_{\alpha=1}^{4} \underset{\sim}{b}_2^{\alpha} .$$

but for a simple square element*, $\underset{\sim}{K}_e$ and its decomposition (2.26) do not coincide. Indeed, for this simple geometry, the calculation of (2.11) can be carried out explicitly and the polynomial in (ξ,η) obtained can be split into one part exactly integrated with 4 Gauss points, and another part of higher order that requires 9 points. This calculation leads to the decomposition:

$$\left.\begin{aligned} K_{xx}^{(9)} &= K_{xx}^{(4)} + \Omega_e \left(\frac{1}{45}\, \underset{\sim}{s}_7 \cdot \underset{\sim}{s}_7^T + \frac{4}{135}\, \underset{\sim}{s}_9 \cdot \underset{\sim}{s}_9^T\right) \\ K_{yy}^{(9)} &= K_{xx}^{(4)} + \Omega_e \left(\frac{1}{45}\, \underset{\sim}{s}_8 \cdot \underset{\sim}{s}_8^T + \frac{4}{135}\, \underset{\sim}{s}_9 \cdot \underset{\sim}{s}_9^T\right) \end{aligned}\right\} \qquad (2.43)$$

where

$$\left.\begin{aligned} K_{xx} &= \int_{\hat{\Omega}} \frac{\underset{\sim}{A} \cdot \underset{\sim}{y} \cdot \underset{\sim}{y}^T \cdot \underset{\sim}{A}^T}{\underset{\sim}{y}^T \underset{\sim}{A} \underset{\sim}{x}}\, d\xi d\eta \\ K_{yy} &= \int_{\hat{\Omega}} \frac{\underset{\sim}{A} \cdot \underset{\sim}{x} \cdot \underset{\sim}{x}^T \cdot \underset{\sim}{A}^T}{\underset{\sim}{y}^T \cdot \underset{\sim}{A} \cdot \underset{\sim}{x}}\, d\xi d\eta \end{aligned}\right\} \qquad (2.44)$$

and $\underset{\sim}{s}_7$, $\underset{\sim}{s}_8$, $\underset{\sim}{s}_9$ are the 7th, 8th and 9th column vectors of $\underset{\sim}{S}$ (2.32). These vectors correspond to the higher order of $\underset{\sim}{\xi}$ (2.33) that cannot be exactly integrated by a 4-point rule. The form taken by the stabilization matrix involves now three matrices $(\underset{\sim}{s}_i \cdot \underset{\sim}{s}_i^T$, i = 7,8,9), is exact for a square element and cannot coincide with the decomposition found in [4]. Finally, we note that both decompositions were used in our a-posteriori control described in Section 2.7 on a regular mesh,

---

* or also for a geometry for which the Jacobian is constant.

and optimal rates of convergence were only obtained with the decomposition (2.43).

2.2.4. The stabilization matrix for a general heat transfer equation. In this paragraph we give the stabilization matrix for a slightly more complicated operator. The case of the linear elasticity operator is discussed later.

Let us consider the case in which the operator is defined by

$$A = \underset{\sim}{\beta}^T \underset{\sim}{C} \underset{\sim}{\beta} \tag{2.45}$$

where

$$\underset{\sim}{\beta} = \left(\frac{\partial}{\partial x} \quad \frac{\partial}{\partial y}\right)^T$$

and

$$\underset{\sim}{C} = \begin{pmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{pmatrix}$$

Then the stiffness matrix associated with this operator is given by

$$\underset{\sim e}{K} = \int_\Omega \nabla \underset{\sim}{N}^T \cdot \underset{\sim}{C} \cdot \nabla \underset{\sim}{N} \, dxdy \tag{2.46}$$

The generalization of the stabilization decomposition when Q_1 elements are used can then be written

$$\underset{\sim e}{K} = \underset{\sim e}{K}^{(1)} + \bar{\varepsilon} \, \underset{\sim}{\gamma} \cdot \underset{\sim}{\gamma}^T \tag{2.47}$$

where

$$\underset{\sim e}{K}^{(1)} = \frac{1}{|\Omega_e|} \, \underset{\sim}{B}^T \underset{\sim}{C} \underset{\sim}{B} \tag{2.48}$$

$$\bar{\varepsilon} = C_{11} \, \bar{\varepsilon}_{xx} + (C_{12} + C_{21}) \bar{\varepsilon}_{xy} + C_{22} \, \bar{\varepsilon}_{yy} \tag{2.49}$$

and

$$\left.\begin{aligned}
\bar{\varepsilon}_{xx} &= \frac{1}{4}\int_{\hat{\Omega}} \frac{1}{J}(\xi \underset{\sim}{s}^{T}\cdot \underset{\sim}{y} - \eta \underset{\sim}{s}'^{T}\cdot \underset{\sim}{y})^2 \, d\xi d\eta \\
\bar{\varepsilon}_{yy} &= \frac{1}{4}\int_{\hat{\Omega}} \frac{1}{J}(\xi \underset{\sim}{s}^{T}\cdot \underset{\sim}{x} - \eta \underset{\sim}{s}'^{T}\cdot \underset{\sim}{y})^2 \, d\xi d\eta \\
\bar{\varepsilon}_{xy} &= \frac{1}{4}\int_{\hat{\Omega}} \frac{1}{J}(\xi \underset{\sim}{s}^{T}\cdot \underset{\sim}{x} - \eta \underset{\sim}{s}^{T}\cdot \underset{\sim}{x})(\xi \underset{\sim}{s}^{T}\cdot \underset{\sim}{y} - \eta \underset{\sim}{s}'^{T}\cdot \underset{\sim}{y}) \, d\xi d\eta
\end{aligned}\right\} \quad (2.50)$$

The quantities $\underset{\sim}{\gamma}$, $\underset{\sim}{B}$ and J are the ones defined previously. Expressions similar to those given in (2.29) and (2.30) can be used to simplify $\bar{\varepsilon}$.

For a regular geometry, and corresponding to (2.29) and (2.30), we have

$$\left.\begin{aligned}
\bar{\varepsilon}_{xx} &= \frac{1}{24(\Omega_e)} \; (y_{13}^2 + y_{24}^2) \\
\bar{\varepsilon}_{yy} &= \frac{1}{24(\Omega_e)} \; (x_{13}^2 + x_{24}^2) \\
\bar{\varepsilon}_{xy} &= \frac{1}{24(\Omega_e)} \; (x_{13}\, y_{13} + x_{24}\, y_{24})
\end{aligned}\right\} \quad (2.51)$$

As far as the 9-node element is concerned, the decomposition can be obtained only for regular elements. First we note that

$$\underset{\sim}{K}_{xy}^{(9)} = \underset{\sim}{K}_{xy}^{(4)} \quad (2.52)$$

where the notations are similar to (2.43) and (2.44). Therefore, the decomposition can be written:

$$\begin{aligned}
\underset{\sim}{K}_{e}^{(9)} = \underset{\sim}{K}_{e}^{(4)} &+ \Omega_e \, C_{11}\left(\frac{1}{45}\, \underset{\sim}{s}_7 \, \underset{\sim}{s}_7^{T} + \frac{4}{135}\, \underset{\sim}{s}_9 \, \underset{\sim}{s}_9^{T}\right) \\
&+ \Omega_e \, C_{22}\left(\frac{1}{45}\, \underset{\sim}{s}_8 \, \underset{\sim}{s}_8^{T} + \frac{4}{135}\, \underset{\sim}{s}_9 \, \underset{\sim}{s}_9^{T}\right)
\end{aligned} \quad (2.53)$$

2.3 A-Posteriori Hourglass Control

2.3.1. Introduction and preliminaries. The basic ideas are more easily understood when demonstrated for the same simple model problem. We still focus on the model Neumann problem P_0 or its variational equivalent P.

Let Ω be a regular (e.g. Lipschitz) domain in $\mathbb{R}^2$ with boundary $\partial\Omega$ and let f be a given L^2-function. Problem P_0 is then,

(P_0) Find u such that

$$\left.\begin{aligned} -\Delta u &= f \text{ in } \Omega \\ \frac{\partial u}{\partial n} &= 0 \text{ on } \partial\Omega \end{aligned}\right\} \tag{3.1}$$

where the data f satisfies the compatibility condition,

$$\int_\Omega f dx = 0 \tag{3.2}$$

Later we shall put further restrictions on Ω and on f (e.g. we will need $f \in H(\Omega)$). The kernel of the governing operator $A = (-\Delta, \frac{\partial}{\partial n})$ in (3.1) is, of course, the space of constants. Thus, whenever (3.2) holds, there exists a solution to (3.1) which is unique up to an arbitrary constant.

To formulate a variational statement of problem P_0, we introduce the spaces and inner products*,

---

* The elements of V (and $L^2(\Omega)/\mathbb{R}$) are cosets $[v]$ such that $u \in [v]$ implies that $u, v \in H^1(\Omega)$ (or $L^2(\Omega)$) and $v - u \in \mathbb{R}$. Throughout this paper we frequently refer to functions v in V, meaning, of course, that v is a representative function in the coset $[v]$.

$$V = H^1(\Omega)/\mathbb{R}$$

$$(u,v)_1 = \text{an inner product on } V$$

$$= \int_\Omega \nabla u \cdot \nabla v dx \; ; \; u, v \in V$$

$$(f,g)_0 = \text{an inner product on } L^2(\Omega)/\mathbb{R}$$

$$= \int_\Omega fgdx - \frac{1}{\text{meas } \Omega} \int_\Omega fdx \int_\Omega gdx \tag{3.3}$$

Three remarks are in order:

i) The norm $\|\cdot\|_0$ associated with the inner product $(\cdot,\cdot)_0$ is the canonical norm on the quotient space $L^2(\Omega)/\mathbb{R}$,

$$\|f\|_0 = \inf_{\lambda \in \mathbb{R}} \|f+\lambda\|_{L^2(\Omega)} \tag{3.4}$$

ii) According to Temam [50], there exists a constant C_0 , depending only on Ω , such that

$$\|v\|_1 \geq C_0 \|v\|_0 \qquad \forall v \in V \tag{3.5}$$

iii) For all f satisfying the compatibility condition (3.2) and any $v \in V$, we have

$$\begin{aligned} (f,v)_0 = \int_\Omega fvdx &\leq \|f\|_0 \; \|v\|_0 \\ &\leq \frac{1}{C_0} \|f\|_0 \; \|v\|_1 \end{aligned} \tag{3.6}$$

With these relations now established, we consider the variational statement of P_0 as problem P :

(P) Find $u \in V$ such that

$$(u,v)_1 = (f,v)_0 \qquad \forall v \in V \tag{3.7}$$

We can easily verify that any solution of P_0 is a solution of P and, conversely, the solution of P satisfies the condition of P_0 in at least, a distributional sense. Moreover, since the bilinear form $(\cdot,\cdot)_1$ is continuous and coercive on V and since the linear form $(f,\cdot)_0$ is continuous on V if (and only if) f satisfies (3.2), the following result is an immediate consequence of the Lax-Milgram Theorem:

THEOREM I. Let f *satisfy* (3.2).

Then there exists one and only one solution $u \in V$ *to problem* P *and this solution depends continuously on the data* f . □

We now consider a finite element approximation of the problem P. Let us now construct a finite element approximation of problem P. We begin by introducing a partition Q of Ω into E finite elements so that

$$\Omega = \bigcup_{e=1}^{E} \Omega_e$$

We shall assume that Ω is such that it can be partitioned in this fashion into four-node quadrilateral elements over which bilinear shape functions are defined. Thus, if

$Q_1(\Omega_e)$ = space of bilinear functions defined on Ω_e

we can introduce the finite-dimensional space

$$V^h = \left\{ v^h \in C^0(\Omega) \text{ such that } v^h|_{\Omega_e} \in Q_1(\Omega_e),\ 1 \le e \le E \ / \ \mathbf{R} \subset V \right\} \tag{3.8}$$

wherein, as usual, the label h is the mesh parameter (e.g. , $h = \max_{1\le e\le E} \mathrm{dia}(\Omega_e)$). The functions in V^h are continuous and are still defined up to an arbitrary constant.

Our finite-element approximation of problem P is embodied in the discrete problem,

$$(P_h) \quad \text{Find } u^h \in V^h \text{ such that}$$
$$(u^h, v^h)_1 = (f, v^h)_0 \qquad \forall v^h \in V^h \tag{3.9}$$

where, again, f satisfies condition (3.2).

In analogy with Theorem V, we have:

THEOREM II. Let f satisfy (3.2). Then there is one and only one solution u^h to problem P_h in V^h and this solution depends continuously on the data f. □

In examining the convergence of such finite element approximations, we shall confine our attention throughout this section to regular mesh refinements. In such cases, we have the a priori asymptotic error estimates,

$$\| u-u^h \|_1 = O(h) \ , \ \| u-u^h \|_0 = O(h^2) \tag{3.10}$$

2.3.2 The underintegrated problem. We now focus our attention on finite element approximations of problem P in which incomplete quadratures are used to evaluate the bilinear form $(\cdot,\cdot)_1$. To simplify this study, we shall now introduce some additional assumptions:

i) Ω is the unit square,

$$\Omega = (0,1) \times (0,1)$$

ii) The finite elements are the squares,

$$\Omega_{ij} = \left(\frac{i-1}{N}, \frac{i}{N}\right) \times \left(\frac{j-1}{N}, \frac{j}{N}\right)$$

$$1 \leq i, j \leq N$$

$$\overline{\Omega} = \bigcup_{1 \leq i,j \leq N} \overline{\Omega}_{ij}$$

iii) The data f is L^2-integrable; e.g.

$$f \in L^2(\Omega) \tag{3.11}$$

In this case, we take

$$h = \frac{1}{N}, \quad \dim V^h = (N+1)^2 - 1 = O(h^{-2}) \tag{3.12}$$

In P_h we can replace f by f^h, its L^2-projection on V^h is defined by

$$(f^h, v^h)_0 = (f, v^h)_0 \qquad \forall v^h \in V^h \tag{3.13}$$

For further use, we note that the projection satisfies

$$\| f^h \|_0 \leq \| f \|_0 \tag{3.14}$$

and can be chosen such that

$$\int_\Omega f^h \, dx = 0 \tag{3.15}$$

Now we turn to the issue of numerical integration of the stiffnesses. Let $I(\cdot,\cdot)$ denote a discrete inner product on $C^0(\Omega)$ defined by a numerical quadrature rule as follows:

$$I_G(f,g) = \sum_{e=1}^{E} I_e^G(f,g)$$
$$I_e^G(f,g) = \sum_{j=1}^{G} W_j^e f(\underset{\sim}{\xi}_j^e) g(\underset{\sim}{\xi}_j^e) \tag{3.16}$$

Here W_j^e are the quadrature weights and $\underset{\sim}{\xi}_j^e$ are the quadrature points for element e and G is the number of quadrature points used. Assuming that Gaussian quadrature is used, the choice $G=4$ (2x2 - Gauss rule) leads to an exact integration of the stiffnesses for each element:

$$(u^h, v^h)_1 = I_4(u^h, v^h) = \underset{\sim}{u}^T \underset{\sim}{K} \underset{\sim}{v} \tag{3.17}$$

for any $u^h, v^h \in V^h$. Here $\underset{\sim}{K}$ is the fully-integrated stiffness matrix and $\underset{\sim}{u}$ and $\underset{\sim}{v}$ are vectors of nodal degrees of freedom of u^h and v^h, respectively.

Instead of the correct bilinear form in (3.18), we wish to consider an underintegrated approximation to $(\cdot,\cdot)_1$ in which only one integration point per element is used:*

$$(u^h, v^h)_{1,h} = I_1(u^h, v^h) = u^T K^{(1)} v$$
$$\forall\, u^h, v^h \in V^h \tag{3.18}$$

Here $\underset{\sim}{K}^{(1)}$ is the underintegrated stiffness matrix. The difference between $(\cdot,\cdot)_1$ and $(\cdot,\cdot)_{1,h}$ (on V^h) is denoted $a'(\cdot,\cdot)$ and the corresponding stiffness matrix is $\underset{\sim}{K}^{stab}$:**

---

* Recall Section (2.2).

** Recall that $\underset{\sim}{K}^{stab} = \bar{\varepsilon} \underset{\sim}{\gamma} \underset{\sim}{\gamma}^T$ where $\bar{\varepsilon} = 1/6$ for a rectangular mesh and $\underset{\sim}{\gamma}$ is given by (2.27).

$$a'(u^h , v^h) = (u^h , v^h)_1 - (u^h , v^h)_{1,h}$$

$$= \underset{\sim}{u}^T \underset{\sim}{K}^{stab} \underset{\sim}{v} \qquad \forall u^h, v^h \in V^h \tag{3.19}$$

The "underintegrated problem",

(P_h^*) Find $u^h \in V^h$ such that

$$(u^h , v^h)_{1,h} = (f^h , v^h)_0 \quad \forall v^h \in V^h \tag{3.20}$$

is, in general, meaningless. This problem, in general, has no solution except for the special case in which f^h is orthogonal to the one-dimensional space of hourglass modes,

$$H = \{H \in V^h | (H, v^h)_{1,h} = 0 \qquad \forall v^h \in V^h\} \tag{3.21}$$

A way to overcome this difficulty is to note that the underintegration of the righthand side also leads to a rank-deficient linear form $(\cdot,\cdot)_{0,h}$:

$$(f^h , H)_{0,h} = 0 \quad , \quad \begin{matrix} \forall f^h \in V^h \\ \forall H \in H \end{matrix}$$

Note that if f^h satisfies (3.15) we also have

$$(f^h , 1)_{0,h} = 0 \tag{3.23}$$

Therefore we now consider the underintegrated problem $\bar{P}_h$:

$(\bar{P}_h)$ Find $\bar{u}^h \in \bar{V}^h$ such that

$$(\bar{u}^h , v^h)_{1,h} = (f^h , v^h)_{0,h} \quad \forall v^h \in \bar{V}^h \tag{3.24}$$

where

$$\overline{V}^h : V^h/H \tag{3.25}$$

We can now state and prove

THEOREM III. *There exists one and only one solution* $\overline{u}^h$ *to* $\overline{P}_h$.

Proof: This is an immediate consequence of the Lax-Milgram theorem. Since

$$(u^h , u^h)_{1,h} = 0 \Leftrightarrow u^h = \gamma_1 H + \gamma_2$$

we can consider $(\cdot,\cdot)_{1,h}^{\frac{1}{2}}$ as a norm on $\overline{V}^h$. It is therefore coercive and continuous on $\overline{V}^h$. As far as the continuity of the righthand side is concerned, a simple calculation shows that for any v^h in V^h we have

$$|(f^h, v^h)_{0,h}| \leq \| f^h\|_0 \ \| v^h\|_0$$

Also for any constants γ_1 and γ_2

$$|(f^h, v^h + \gamma_1 + \gamma_2 H)_{0,h}| = |(f^h, v^h)_{0,h}|$$

therefore

$$\begin{aligned} |(f^h, v^h)_{0,h}| &\leq \| f^h\|_0 \ \| v^h + \gamma_1 + \gamma_2 H\|_0 \quad \forall \gamma_1, \gamma_2 \\ &\leq \| f^h\|_0 \ \| v^h + \gamma_2 H\|_1 \quad \forall \gamma_2 \\ &\leq \alpha_h \| f^h\|_0 \ \| v^h\|_{1,h} \end{aligned}$$

Here we successively used (2.23), (2.22), (3,6) and the equivalence between the canonical norm of $\overline{V}^h$ and the norm $\| \cdot\|_{1,h}$ □

We have obtained a solution to the underintegrated problem $\overline{P}_h$.

This solution is unique in $\bar{V}^h$, from a computational point of view it is defined up to within an arbitrary hourglass mode. We now need a projection to obtain a reasonable solution from any representative $\bar{u}^h$ chosen.

2.3.3. Projection of the underintegrated solution. In order to construct this projection, we remark that, since u^h is a solution of P_h and since $H \in V^h$, u^h satisfies

$$(u^h, H)_1 = (f^h, H)_0 \tag{3.26}$$

We wish to extend $\bar{u}^h$ to all of V^h so that a new function $\tilde{u}^h \in V^h$ is obtained which contains an hourglass mode and which also satisfies (4.8). Thus, if π is an operator from $\bar{V}^h$ into V^h, we define

$$\tilde{u}^h = \pi\bar{u}^h = \bar{u}^h + \lambda_0 H, \quad \lambda_0 \in \mathbb{R} \tag{3.27}$$

$$(\tilde{u}^h, H)_1 = (f^h, H)_0$$

This latter requirement determines λ_0 uniquely as

$$\lambda_0 = \frac{1}{\|H\|_1^2}\left[(f^h, H)_0 - (\bar{u}^h, H)_1\right] \tag{3.29}$$

so that $\tilde{u}^h$ is uniquely determined as the function

$$\tilde{u}^h = \bar{u}^h + \frac{(f^h,H)_0}{\|H\|_1^2} H - \frac{(u^h,H)_1}{\|H\|_1^2} H$$

It is instructive to consider a geometrical interpretation of our projection defined in (4.9). Note that the "component" of the fully integrated solution u^h orthogonal (in V^h) to $H^{\perp}$ is $(u^h,H)_1 = (f^h,H)_0$,

as indicated in Fig. 3. The solutions $\bar{u}_h$ of $\bar{P}_h$ constitute the vectors generating a line "parallel to" the space H in the figure. The projection $\tilde{u}^h$ is then the vector defined by the orthogonal projection of u^h onto this line. Indeed, by construction,

$$(\tilde{u}^h - u^h, H)_1 = 0$$

At this point, we have established the following procedure for processing an underintegrated finite element approximation of problem P.

i) Compute the underintegrated bilinear and linear forms $(\cdot,\cdot)_{1,h}$ *and* $(f^h,\cdot)_{0,h}$

ii) Solve problem $\bar{P}_h$ *for* $\bar{u}^h$

iii) Compute $(\bar{u}^h, H)_1$

iv) Construct the enhanced solution $\tilde{u}^h$ *using* (3.30).

Thus, this procedure involves the computation of an underintegrated solution $\bar{u}_h$ to a reduced problem $\bar{P}_h$ and its enrichment via a post-processing operation to obtain a new approximation $\tilde{u}^h$. We shall now show that these post-processed solutions $\tilde{u}^h$ converge to the exact solution u of problem P as the mesh is refined, and, remarkably, these approximations converge at precisely the same rate as the fully-integrated solution!

Indeed we have:

THEOREM IV: Let u, u^h *and* $\bar{u}^h$ *be the solutions of* P, P_h *and* $\bar{P}_h$, *let* f *be in* $L^2(\Omega)$ *and satisfy* (3.2). *Let* $\tilde{u}^h$ *be ob-*

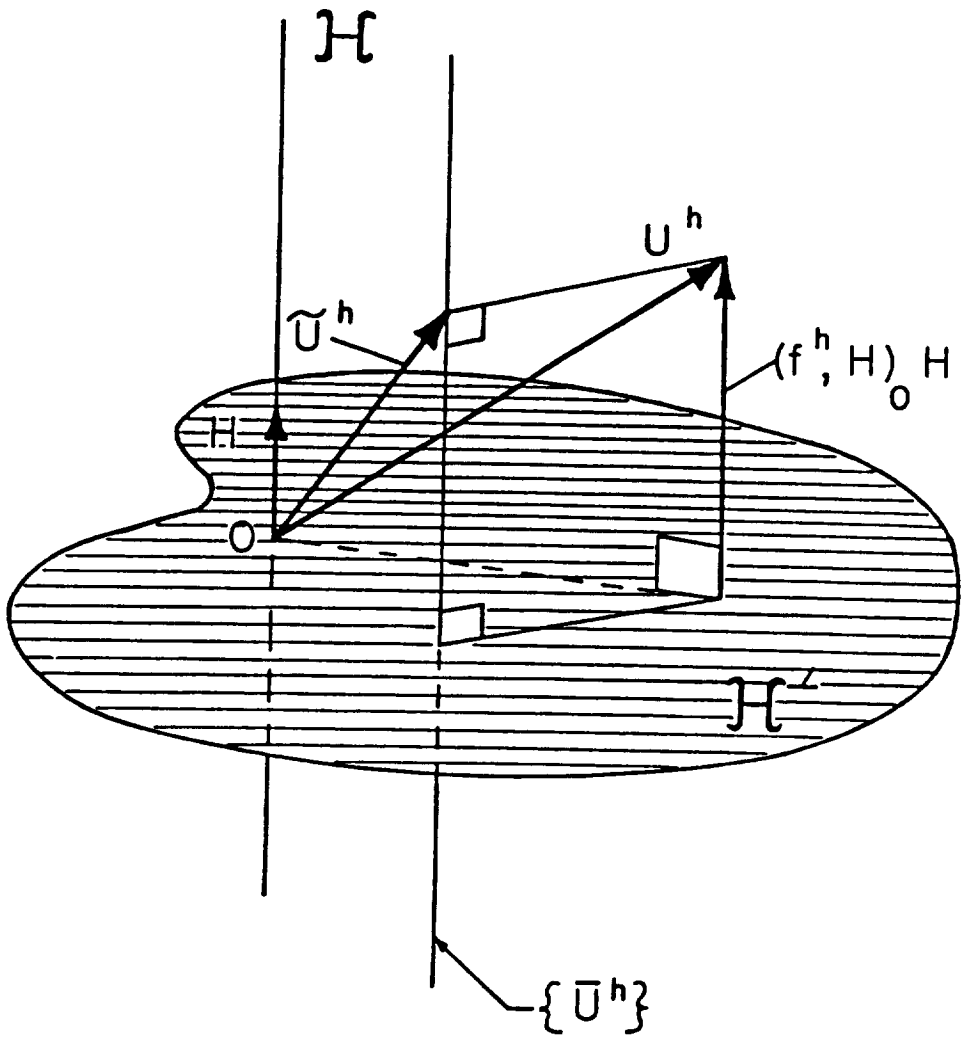


Figure 3. Geometrical Interpretation of the Projection $\tilde{u}^h = \pi \bar{u}^h$.

tained by the projection of $\bar{u}^h$ *defined in* (3.30). *Then we have the following error estimates for* $s = 0$ *and* 1

$$\| u^h - \tilde{u}^h \|_s \leq C_1 \, h^{2-s} \, \| f \|_0 \tag{3.31}$$

and

$$\| u - \tilde{u}^h \|_s \leq C_1 \, h^{2-s} \, \| f \|_0 \tag{3.32}$$

The next section will prove this theorem.

2.4. Convergence of the A-Posteriori Control

2.4.1 Introduction. This section is devoted to the proof of Theorem VIII. The method of proof relies on the tensor properties of the bilinear element and of the Gauss integration rules. The problems P_h and $\bar{P}_h$ will be explicitly solved using an orthonormal basis of eigenvectors of $(\cdot,\cdot)_1$, $(\cdot,\cdot)_{1,h}$ and $(\cdot,\cdot)_{0,h}$. Then we note that for a regular domain and mesh, $f \in L^2(\Omega)$ implies $u \in H^2(\Omega)$ and that

$$\| u - u^h \|_1 < Ch \| f \|_0 \tag{4.1}$$

Likewise, the Aubin-Nitsche method provides also

$$\| u - u^h \|_0 \leq C' h^2 \| f \|_0 \tag{4.2}$$

By the triangle inequality,

$$\| u - \tilde{u}^h \|_1 \leq Ch \, \| f \|_0 + \| u^h - \tilde{u}^h \|_1 \tag{4.3}$$

with a similar estimate in the $\| \cdot \|_0$-norm.

Thus, it suffices to estimate the relative error

$$e^h = \tilde{u}^h - u^h \tag{4.4}$$

The L^2- and H^1-norms of this relative error will be explicitly calculated and estimated.

2.4.2. Some one-dimensional results. For reasons to be made clear in the next subsection, it is convenient to review briefly some results on one-dimensional piecewise-linear approximations on a uniform mesh for $\Omega = (0,1)$. Our aim here is to establish concrete relationships between various bilinear forms $(\cdot,\cdot)_{0,h}$, $(\cdot,\cdot)_1$, and $(\cdot,\cdot)_{1,h}$ on spaces of piecewise-linear functions.

Let $\underset{\sim}{D}(k, \alpha)$ and $\underset{\sim}{I}$ denote the N+1-order matrices

$$\underset{\sim}{D}(k,\alpha)=\begin{bmatrix} k & \alpha & \cdot & \cdot & \cdot & 0 & 0 \\ \alpha & 2k & \cdot & \cdot & \cdot & 0 & 0 \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \\ 0 & 0 & \cdot & \cdot & \cdot & 2k & \alpha \\ 0 & 0 & \cdot & \cdot & \cdot & \alpha & k \end{bmatrix}, \underset{\sim}{I}'=\begin{bmatrix} 1 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 2 & \cdot & \cdot & \cdot & 0 & 0 \\ & \cdot & \cdot & \cdot & \cdot & \cdot & \\ 0 & 0 & \cdot & \cdot & \cdot & 2 & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 1 \end{bmatrix} \tag{4.5}$$

(i.e. $\underset{\sim}{I}' = (\underset{\sim}{D}(1,0)$. Then, for $\alpha \neq 0$, one can show that

$$\det \underset{\sim}{D}(k, \alpha) = (-\alpha)^{N+1} \det \underset{\sim}{D}(-\frac{k}{\alpha}, -1)$$

$$= (-\alpha)^{N+1} \det \underset{\sim}{D}(-\frac{k}{\alpha}) \tag{4.6}$$

where

$$\underset{\sim}{D}(k) \overset{\text{def}}{=} \underset{\sim}{D}(k, -1) \tag{4.7}$$

The values of k for which $\det \underset{\sim}{D}(k)$ vanishes are

$$k_i = \cos \frac{i\pi}{N} , \quad 0 \leq i \leq N \tag{4.8}$$

and the corresponding vectors $(\underset{\sim}{D}(k_i)\underset{\sim}{v}_i = \underset{\sim}{0})$ are

$$\underset{\sim}{v}_i = \{\cos \frac{ij\pi}{N}\} \ , \ 0 \leq j \leq N \tag{4.9}$$

The significance of the above matrices is that in one dimension, the discrete $H^1(0,1)$- , $L^2(0,1)$- and underintegrated $L^2(0,1)$-norms, on the space V_1^h of piecewise linear C_N^0-functions on a uniform mesh of N elements on $(0,1)$,

$$V_1^h = \{v^h \in C^0(0,1) \mid v^h \text{ is linear on } [eh, (e+1)h], \ e=0,\ldots,N-1\} \tag{4.10}$$

are associated with the matrices

$$\underset{\sim}{A}_0 = \frac{h}{3} D(1,\tfrac{1}{2}) \ , \ \underset{\sim}{A}_1 = \frac{1}{h} \underset{\sim}{D}(1, -1) \text{ and } \underset{\sim}{A}_{0,h} = \frac{h}{4} \underset{\sim}{D}(1,1) \tag{4.11}$$

respectively. In other words,

$$\| v^h \|_s^2 = \underset{\sim}{v} \underset{\sim}{A}_s \underset{\sim}{v} \qquad s = 0, \ 1 \ , \ (0,h) \tag{4.12}$$

where $\underset{\sim}{v}$ is the vector of nodal values of v^h .

By using (4.6) through (4.8), one can verify that the numbers α_i and β_i which render $\underset{\sim}{A}_{0,h} - \alpha_i \underset{\sim}{A}_0$ and $\underset{\sim}{A}_1 - \beta_i \underset{\sim}{A}_0$ singular are

$$\alpha_i = \frac{3(1 + \cos \frac{i\pi}{N})}{2(2 + \cos \frac{i\pi}{N})} \tag{4.13}$$

$$\beta_i = \frac{6}{h^2} \frac{1 - \cos \frac{i\pi}{N}}{2 + \cos \frac{i\pi}{N}} \tag{4.14}$$

In particular, let $\phi^i = \phi^i(x)$, $x \in [0,1]$ denote the piecewise

linear functions associated with the vectors v_i :

$$\left.\begin{aligned} &\phi^i(jh) = \cos \frac{ij\pi}{N} \ , \ 0 \leq i,j \leq N \\ &\text{span } \{\phi^i\}_{0\leq i\leq N} = V_1^h \end{aligned}\right\} \tag{4.15}$$

Then,

$$(v^h, \phi^i)_{0,h} = \alpha_i \, (v^h \ , \phi^i)_0 \qquad v^h \in V_1^h \tag{4.16}$$

$$(v^h, \phi^i)_1 = \beta_i \, (v^h, \phi^i)_0 \tag{4.17}$$

Notice that the base functions ϕ^i are orthogonal for each of the scalar products under consideration.

The following remarks are in order:

i) The denominators in (4.13) and (4.14) are non-zero.

ii) For $i = N$, $\alpha_i = 0$ and the corresponding eigenfunction is the one-dimensional hourglass mode:

$$(1, -1, 1, -1, \ldots)$$

iii) For $i = 0$, $\beta_i = 0$ and the corresponding eigenfunction is constant. Then we have the condition $(v^h,1)_1 = 0$ as expected.

2.4.3. Discrete norms for two dimensional meshes. The extension of the above results to two-dimensional rectangular meshes is straightforward. Since the bilinear basis functions for V^h are tensor products of piecewise linear functions of one variable, we can define

$$\phi^{ij}(x,y) = \phi^i(x)\phi^j(y)$$

$$0 \leq i,j \leq N \tag{4.18}$$

Further, let us normalize these basis functions so that

$$\| \phi^{ij} \|_0 = 1$$

We can then establish the following:

Lemma 1.1. For $v^h \in V^h$, *we have*

$$(v^h, \phi^{ij})_{0,h} = \alpha_i \alpha_j \ (v^h, \phi^{ij})_0 \tag{4.19}$$

$$(v^h, \phi^{ij})_1 = (\beta_i + \beta_j) \ (v^h, \phi^{ij})_0 \tag{4.20}$$

$$(v^h, \phi^{ij})_{1,h} = (\alpha_j \beta_i + \alpha_i \beta_j) \ (v^h, \phi^{ij})_0 \tag{4.21}$$

Moreover, if arbitrary $v^h \in V^h$ *is expressed in the form,*

$$\left.\begin{aligned} v^h &= \sum_{0 \le i,j \le N} v_{ij} \phi^{ij} \\ v_{ij} &= (v^h, \phi^{ij})_0 \end{aligned}\right\} \tag{4.22}$$

then

$$\| v^h \|_0^2 = \sum_{0 \le i,j \le N} v_{ij}^2 \tag{4.23}$$

$$\| v^h \|_1^2 = \sum_{0 \le i,j \le N} (\beta_i + \beta_j) v_{ij}^2 \tag{4.24}$$

Proof: First note that

$$\begin{aligned} (v^h, \phi^{ij})_{0,h} &= I(v^h \phi^i(x) \phi^j(y)) \\ &= \alpha_i \alpha_j \int_\Omega v^h \phi^i(x) \phi^j(y) \, dxdy \\ &= \alpha_i \alpha_j \ (v^h, \phi^{ij})_0 \end{aligned}$$

We also have

$$(v^h, \phi^{ij})_1 = \int_\Omega \frac{\partial v^h}{\partial x} \phi^{i'}(x)\, \phi^j(y) + \frac{\partial v^h}{\partial y} \phi^{j'}(y)\, \phi^i(x)\ dxdy$$

$$= \beta_i \int_\Omega v^h \phi^i(x) \phi^j(y)\ dxdy + \beta_j \int_\Omega v^h \phi^i(x) \phi^j(y)\ dxdy$$

$$= (\beta_i + \beta_j)\ (v^h, \phi^{ij})_0$$

Finally

$$(v^h, \phi^{ij})_{1,h} = I_1\left(\frac{\partial v^h}{\partial x} \phi^{i'}\phi^j + \frac{\partial v^h}{\partial y} \phi^i \phi^{j'}\right)$$

$$= \beta_i \alpha_j (v^h, \phi^{ij})_0 + \beta_j \alpha_i (v^h, \phi^{ij})_0$$

The norms (4.23) and (4.24) are then directly obtained □

In analogy with our remarks on the one-dimensional case, we observe that for $i = j = N$, $\phi^{ij} = H$, the two dimensional hourglass mode. Then

$$(v^h, H)_1 = \frac{24}{h^2} (v^h, H)_0 \tag{4.25}$$

and

$$(v^h, H)_{0,h} = 0 \tag{4.26}$$

Also, for $i = j = 0$, $\phi^{ij} = 1$ and the equilibrium condition (3.2) can be written

$$f_{00} = 0 \tag{4.27}$$

with

$$f_{ij} = (f^h, \phi^{ij})_0 \tag{4.28}$$

2.4.4. Explicit resolution of P_h and $(\bar{P}_h + \pi)$. With the above results in hand, let us now return to the fully-integrated finite-element approximate problem P_h given in (3.9). The solution u^h to that problem can be written

$$\left.\begin{aligned} u^h &= \sum_{0\le i,j\le N} u_{ij}\phi^{ij} \\ u_{ij} &= (u^h, \phi^{ij})_0 \end{aligned}\right\} \tag{4.29}$$

and since for $\left[v^h = \phi^{ij} \text{ in (3.9)}\right]$,

$$\begin{aligned} (u^h, \phi^{ij})_1 &= (\beta_i + \beta_j)\,(u^h, \phi^{ij})_0 \\ &= (f^h, \phi^{ij})_0 = f_{ij} \end{aligned}$$

we have

$$u_{ij} = \frac{1}{\beta_i + \beta_j}\, f_{ij} \;;\; (i,j) \neq (0,0) \tag{4.30}$$

Using constructions similar to those in (4.29) for the fully-integrated problem, we easily verify that the solution $\bar{u}^h$ to the underintegrated problem $\bar{P}_h$ is representable in the form,

$$\bar{u}^h = \sum_{\substack{(i,j)\neq(N,N)\\(i,j)\neq(0,0)}} \bar{u}_{ij}\phi^{ij} \tag{4.31}$$

with

$$\bar{u}_{ij} = \frac{\alpha_i\,\alpha_j}{\alpha_i\beta_j + \alpha_j\beta_i}\, f_{ij} \;;\; (i,j) \neq (0,0) \text{ and } (N,N) \tag{4.32}$$

The cases $(i,j) = (N,N)$ and $(i,j) = (0,0)$ correspond to the arbitrary hourglass mode and arbitrary constant, respectively.

The projected approximation $\tilde{u}^h$ defined by $\tilde{u}^h = \pi\bar{u}^h$ is constructed so that projections of $\tilde{u}^h$ and u^h coincide; i.e.

$$\left.\begin{aligned} \tilde{u}_{ij} &= \bar{u}_{ij} \qquad (i,j) \neq (0,0) \text{ and } (N,N) \\ \tilde{u}_{N,N} &= \bar{u}_{N,N} \end{aligned}\right\} \tag{4.33}$$

2.4.5. Proof of theorem IV. Since the error function $e^h = u^h - u^h$ is in V^h, we use (4.29) and (4.31) to obtain

$$e^h = \sum_{\substack{(i,j)\neq(N,N) \\ (i,j)\neq(0,0)}} e_{ij}\phi^{ij} \tag{4.34}$$

where

$$\begin{aligned} e_{ij} &= (e^h, \phi^{ij})_0 \\ &= (\tilde{u}^h, \phi^{ij})_0 - (u^h, \phi^{ij})_0 \\ &= \tilde{u}_{ij} - u_{ij} \end{aligned}$$

Thus, from (4.30) and (4.33),

$$e_{ij} = \left(\frac{\alpha_i\alpha_j}{\alpha_i\beta_j + \alpha_j\beta_i} - \frac{1}{\beta_i + \beta_j}\right) f_{ij} \tag{4.35}$$

Then, using (4.13) and (4.14), e_{ij} can be written as

$$e_{ij} = h^2 K_{ij} f_{ij} \tag{4.36}$$

where

$$K_{ij} = K(\cos \frac{i\pi}{N} , \cos \frac{j\pi}{N}) \tag{4.37}$$

and

$$K(x,y) = \frac{1}{4} \frac{(1+x)(1+y)}{(1+x)(1-y)+(1+y)(1-x)} - \frac{1}{6} \frac{(2+x)(2+y)}{(2+x)(1-y)+(2+y)(1-x)} \tag{4.38}$$

On the square $S = \left[-1,+1\right] \times \left[-1,+1\right] / \{(-1,-1), (1,1)\}$, $K(\cdot,\cdot)$ is bounded and there exists a positive constant K such that

$$|K(x,y)| \leq K \qquad \forall (x,y) \in S \tag{4.39}$$

Therefore we have

$$|K_{ij}| \leq K \qquad \forall (i,j) \neq (0,0) \text{ and } (N,N) \tag{4.40}$$

and we can obtain using (3.13) and (4.23)

$$\| e^h \|_0^2 = h^4 \sum_{\substack{(i,j)\neq(0,0) \\ (i,j)\neq(N,N)}} K_{ij}^2 f_{ij}^2$$

$$\leq h^4 K^2 \| f^h \|_0^2 \leq h^4 K^2 \| f \|_0^2$$

Also, after calculation and use of (4.24), (4.14) and (3.14), we have

$$\| e^h \|_1^2 = h^4 \sum_{\substack{(i,j)\neq(0,0) \\ (i,j)\neq(N,N)}} (\beta_i + \beta_j) K_{ij}^2 f_{ij}^2$$

$$\leq 12 h^2 K^2 \| f \|_0^2 \qquad \blacksquare$$

2.5. Implementation and Numerical Results of the A-Posteriori Control For the Laplace Equation.

In this section we first would like to indicate how the a-posteriori

control method is implemented, and how its time efficiency compares to the a-priori method. Then several numerical results will be given, illustrating the accuracy of the method and confirming the results obtained in the previous sections.

2.5.1. Implementation of the a-posteriori method. First let us indicate that from a mathematical point of view the problem $\bar{P}^h$ is well-posed but computationally, the matrix obtained from this formulation is singular and the dimension of its kernel is 2. Consequently, we must pick two nodes, fix them a value, and solve. The first value fixes the constant mode, and the second one fixes the hourglass mode to be eliminated later. Let us fix $\bar{u}^h$ equal to zero at the origin and at the next point on the boundary (coordinates : h,0) (Figure 4.a). According to the error estimates (3.30) and (3.31), we may write

$$u^h = \bar{u}^h + \lambda H + O(h^{2-\varepsilon}) \tag{5.1}$$

and therefore, if we normalize H such that its nodal values are 0 or 1, λ measures precisely the value of u^h at (h,0) (Figure 4.b), and approaches $u(h,0)$

$$\lambda = u(h,0) + O(h^{\sigma}) \tag{5.2}$$

But $u(h,0)$ is $O(h^2)$ for a smooth enough solution ($u(0,0) = 0$, $\partial u/\partial \eta(0,0) = 0$) and using L^{∞}-estimates [11], σ can be evaluated to $2-\varepsilon$, ε arbitrary. Finally, we have the estimate

$$\lambda = O(h^{2-\varepsilon}), \quad \varepsilon \text{ arbitrary} \tag{5.3}$$

Also, the choice of H leads to

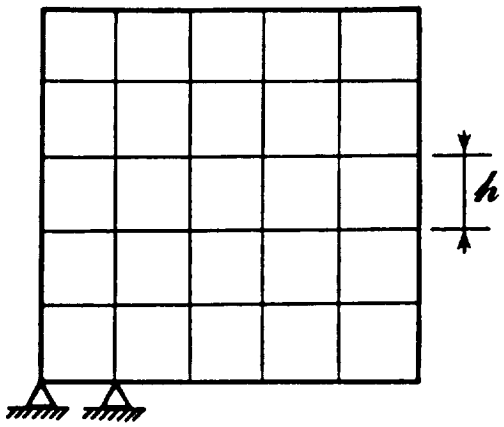


Figure 4.a. Two Fixed Nodes

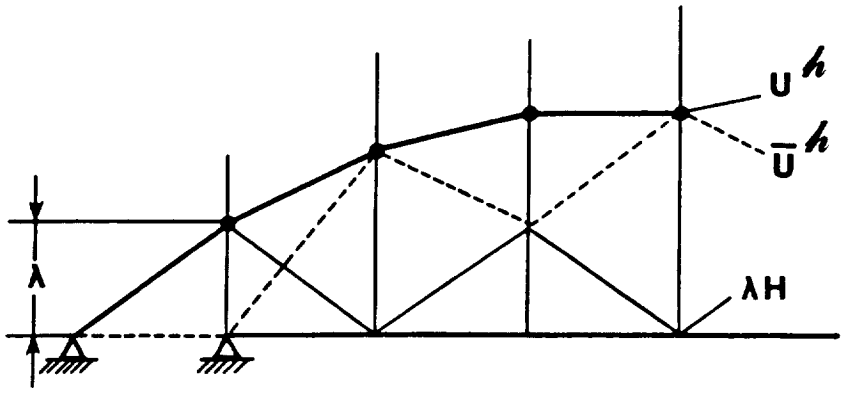


Figure 4.b. $\underset{\sim}{u}^h = \underset{\sim}{\bar{u}}^h + \lambda \underset{\sim}{H} + \underset{\sim}{0}(h^{2-s})$

Figure 4. Justification of the Omission of the Projection.

$$\|H\|_0 = 1/3$$

$$\|H\|_1 = \frac{2\sqrt{2}}{\sqrt{3}\,h} \tag{5.4}$$

and therefore we obtain

$$\|\lambda H\|_s = O(h^{2-s-\varepsilon}) \qquad s = 0,1 \tag{5.5}$$

and that proves that the post processor contribution λH can be neglected if the fixed nodes are chosen as indicated for this type of boundary condition. The error estimates of Theorem IV still hold up to within $h^{-\varepsilon}$.

Unfortunately this remark has two major drawbacks: it supposes that u is smooth ($u \in H^2(\Omega)$) and it is not valid to 9-node elements that will later be discussed.

Before discussing the implementation of (3.30), we indicate that this projection can be simplified. Indeed, taking $v^h = H$ in (4.25) we obtain

$$\left\| \frac{(f_1^h, H)_0}{\|H\|_1^2} H \right\| \leq \|f^h\|_0 \frac{\|H\|_0^2}{\|H\|_1^2} = \frac{h^2}{24} \|f^h\|_0 \tag{5.6}$$

$$\frac{(f_1^h, H)_0}{\|H\|_1^2} H \Big\|_1 \leq \|f^h\|_0 \frac{\|H\|_0}{\|H\|_1} = \frac{h}{2\sqrt{6}} \|f^h\|_0$$

Therefore we have

$$\left\| \frac{(f_1^h, H)_0}{\|H\|_1^2} H \right\|_s \leq C \|f^h\|_0 \; h^{2-s}, \quad s = 0,1 \tag{5.8}$$

and this term can be neglected without affecting the estimate of Theorem

VIII. The formula used in the post processor is then

$$\tilde{u}^h = \bar{u}^h - \lambda H \tag{5.9}$$

$$\lambda = (\bar{u}^h, H)_1 \, \| H \|_1^{-2} \tag{5.10}$$

In order to preserve the efficiency of the method provided by underintegration, one must find an efficient way to compute the parameter λ in (5.9). One way that suggests itself is to calculate the H^1 inner products of (5.10) using numerical integration. The use of a one point rule would be absurd and would lead to a ratio 0/0. The use of a 4 Gauss point rule has been numerically implemented and gives good results (similar to those to be presented next) but cost of this integration is expensive, as shown in Table 3. This method is therefore rejected.

We shall now describe a more efficient method with related numerical results shown in the next subsection. This method relies on the fact that, for the bilinear element, the stiffness matrix can be decomposed into two parts, one of which contains H in its kernel. The other part is such that the image of H is cheap to calculate.

This decomposition proved in Section 2.2.2 can be written as

$$\underset{\approx}{K}_e^{\text{exact}} = \underset{\approx}{K}_e^{\text{under}} + \bar{\varepsilon}_e \, \underset{\sim}{\gamma}_e \cdot \underset{\sim}{\gamma}_e^T \tag{5.11}$$

where $\underset{\approx}{K}_e^{\text{exact}}$ and $\underset{\approx}{K}_e^{\text{under}}$ respectively are the exact element stiffness matrix and its under-integrated form, ε_e and $\underset{\sim}{\gamma}_e$ are obtained from (2.27) and (2.28). In particular

$$\underset{\sim}{\gamma}_e = \underset{\sim}{h} - \frac{h \cdot x}{|\Omega_e|} \, \underset{\sim}{b}_1 - \frac{h \cdot y}{|\Omega_e|} \, \underset{\sim}{b}_2 \tag{5.12}$$

Even though ε_e given by (2.27) is still difficult to calculate, note that, during the element calculations, the vectors $\underset{\sim}{b}_1$ and $\underset{\sim}{b}_2$ and the Jacobian $|\Omega_e|$ are necessarily computed. Therefore γ_e is very easy to calculate.

The inner product $(\bar{u}^h, H)_1$ may, therefore, be calculated by using the decomposition (2.1) of $\underset{\approx}{K}^{exact}$. Introducing the nodal vectors $\underset{\sim}{\bar{U}}$ and $\underset{\sim}{H}$ associated with the function $\bar{u}^h$ and H, we have

$$(\bar{u}^h, H)_1 = \underset{\sim}{U}^T \underset{\approx}{K}^{exact} \underset{\sim}{H} = \underset{\sim}{U}^T \cdot \sum_e \bar{\varepsilon}_e \underset{\sim}{\gamma}_e \cdot \underset{\sim}{\gamma}_e^T \cdot \underset{\sim}{H}$$

$$= \sum \bar{\varepsilon}_e (\underset{\sim}{\gamma}_e^T \cdot \underset{\sim}{\bar{U}}_e)(\underset{\sim}{\gamma}_e^T \cdot \underset{\sim}{H}_e)$$

where $\underset{\sim}{\bar{U}}_e$ and $\underset{\sim}{H}_e$ are the values of $\underset{\sim}{\bar{U}}$ and $\underset{\sim}{H}$ at the nodes of the element e. We note that if the values of $\underset{\sim}{H}$ are $+1$ or -1, the scalar vector product $\underset{\sim}{\gamma}_e^T \cdot H_e$ is always ± 4. Therefore

$$(\bar{u}^h, H)_1 = 4 \sum_e \pm \bar{\varepsilon}_e \sum_{i=1} \gamma_e^i u_e^i \tag{5.13}$$

and

$$(H, H)_1 = 16 \sum_e \bar{\varepsilon}_e \tag{5.14}$$

These expressions are still exact since no approximation has been made on $\bar{\varepsilon}_e$. If we suppose that the Jacobian of the element is approximately constant (true for parallelogram element), $\bar{\varepsilon}_e$ is simply expressed as

$$\bar{\varepsilon}_e = \frac{1}{12|\Omega_e|} (\underset{\sim}{b}_1^T \underset{\sim}{b}_1 + \underset{\sim}{b}_2^T \underset{\sim}{b}_2) \tag{5.15}$$

The calculation of the approximate projection can be summarized in

the following algorithm:

- Loop on Elements
  - Calculate $\underset{\sim}{\gamma}_e$, ε_e using (2.2) and (2.6)
  - Calculate $\pm\varepsilon_e(\gamma_e^T \cdot \bar{U}_e)$
  - Add
    - $\lambda_2 = \lambda_1 \pm \varepsilon_e(\gamma_e^T \cdot \bar{U}_e)$
    - $\lambda_2 = \lambda_2 + \varepsilon_e$
- $\lambda = \lambda_1/4\lambda_2$
- Loop on Nodes
  - $\tilde{u}^h|_{Node} = \bar{u}^h|_{Node} \pm \lambda$

Remark: The notations previously used are essentially those found in the work of Belytschko and co-workers [5, 6] on stabilization methods. These methods rely on the decomposition (5.11) but the stabilization term $\varepsilon\underset{\sim}{\gamma}\cdot\underset{\sim}{\gamma}^T$ is a-priori added to the under-integrated matrix to prevent the spurious modes from the kernel of the stiffness matrix; whereas our control method uses the very same term a-posteriori, after solving with the underintegrated matrix. Therefore, our method seems to be cheaper than the stabilization methods as summarized in Table 1.

2.5.2. Numerical results.

2.5.2.a. Regular mesh of 4-node elements. In order to illustrate what has been stated, we have considered the Laplacian problem solved on a square domain partitioned into $N^2(=h^{-2})$ subdomains, for various values of N and we have studied the norms of the difference between the solution obtained with a full integration u^h (4 point rule

Table 1. Operations Cost per Element for Both Stabilization and A-posteriori Methods

Stabilization Method	Operations Cost		A-posteriori Method
Computations of $\bar{\varepsilon}_e$, $\underset{\sim}{\gamma}_e$	20x ; 21+	20x ; 21+	Computations of $\bar{\varepsilon}_e$, $\underset{\sim}{\gamma}_e$
Multiply $\gamma_e \cdot \gamma_e^T$	16x	4x	Multiply $U_e^T \cdot \gamma_e$
Multiply $\varepsilon_e \cdot \gamma_e \gamma_e^T$	16x	1x	Multiply $\varepsilon_e \cdot U_e^T \gamma_e$
Add $K_e + \varepsilon_e \gamma_e \gamma_e^T$	16+	2+	Add $\gamma_1 \pm \varepsilon_e U_e^T \gamma_e$ $\gamma_2 + \varepsilon_e$
		4+	Then: (4 nodes/element) Add $\bar{u}^h \pm \lambda$
TOTAL	52x ; 37+	29x ; 27+	TOTAL

and with underintegration (1 point rule) $\bar{u}^h$ and $\tilde{u}^h$ (before and after post processing). The results are shown as plot of $\text{Log}\| u^h - \bar{u}^h \|$ or $\text{Log}\| u - \tilde{u}^h \|_s$ in function of $|\text{Log } h|$, for $s = 0,1$. Data of various regularities have been used:

i) f_1 is a C^0-function, but not C^1 :

$$\begin{cases} f_1(x,y) = \frac{3}{2}(1-x) - \frac{14}{9}\, y & \text{if } Y_1(x,y) \geq 0 \\ f_1(x,y) = y(\frac{3}{2}\,(1-x) - y - \frac{5}{9}) & \text{if } Y_1(x,y) \leq 0 \end{cases}$$

where the C^1-discontinuity line is

$$Y_1(x,y) = \frac{3}{2}\,(1-x) - y$$

ii) f_2 is a non-continuous function

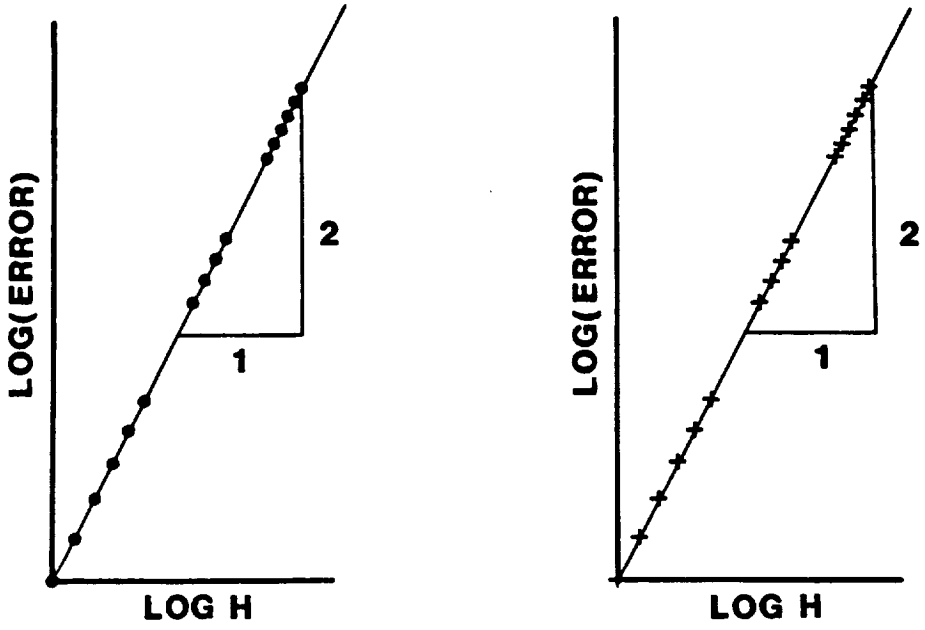
$$\begin{cases} f_2(x,y) = 1 & \text{if } Y_1(x,y) > 0 \\ f_2(x,y) = -2 & \text{if } Y_1(x,y) < 0 \end{cases}$$

where Y_1 is the same as in i).

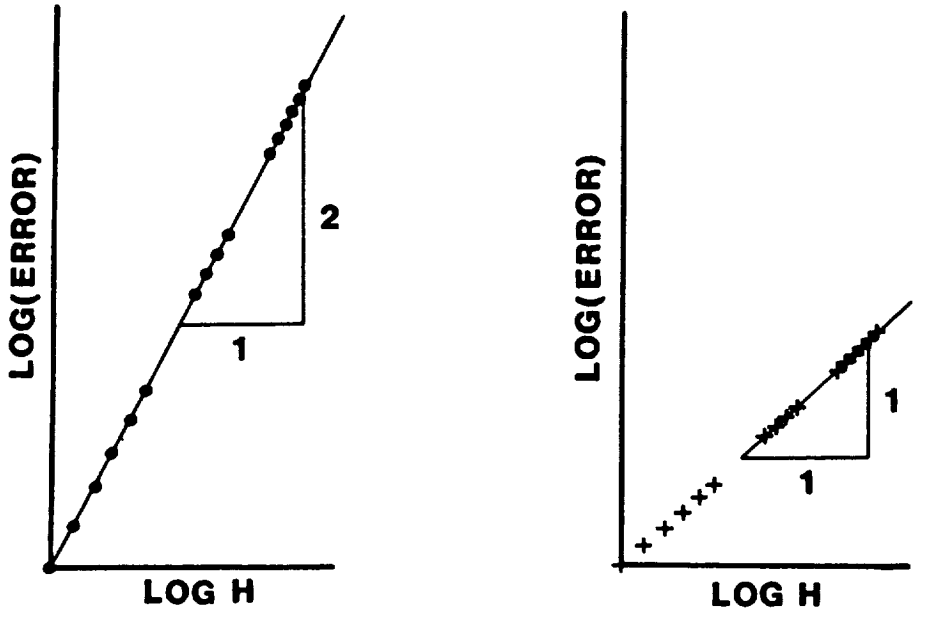
Remark: Both of these functions satisfy the compatibility condition (3.2).

Results obtained with the continuous function f_1 are shown in Figure 5. When the solution has been treated by the post processor (Fig. 5a.), bor both L^2 and H^1 norms, the representing points lie on lines of slope 2. This proves that whereas the estimate (1-13) is optimal for the L^2 norm ($s = 0$), it is not in the H^1 norm and seems to be in fact better than what was expected in our study. This does not affect in any case the comparison with the exact solution (1.14).

Figure 5.b shows the comparison with the crude solution $\bar{u}^h$, obtained with two fixed nodes, and not treated by the post-processor.



(a)



(b)

Figure 5. Results Obtained with a Continuous Data Function: Comparison between : a) u^h and $\tilde{u}^h$; b) u^h and $\bar{u}^h$.

Slopes 2 and 1 are observed and the loss (1. instead of 2. for the H^1 norm) corroborates the final remark of Section 3.32.

When the function f_2 is used (Figure 6), the points show oscillations around two lines of slope 2. (for the L^2-norm) and 1.65 (for the H^1-norm) proving that the estimate (3.31) still holds (Fig. 6.a). When the solution has not been treated by the post-processor (Fig. 6.b), the slope 1.65 becomes 1. as expected.

The next series of examples was intended to study the influence of a singularity (at the origin) for a unit square domain regularly partitioned. The data functions are of the form

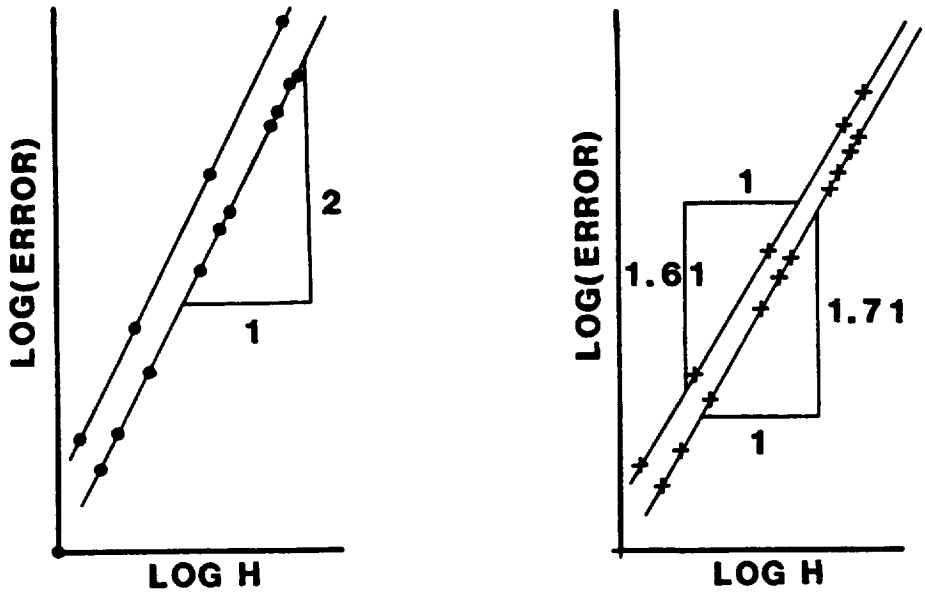
$$f_\alpha(x,y) = r^\alpha - C \quad , \quad \alpha > -2 \tag{5.16}$$

where C is a real number chosen such that the equilibrium condition (2.2) is satisfied. The family $\{f_\alpha\}$ represents various regularities of data:

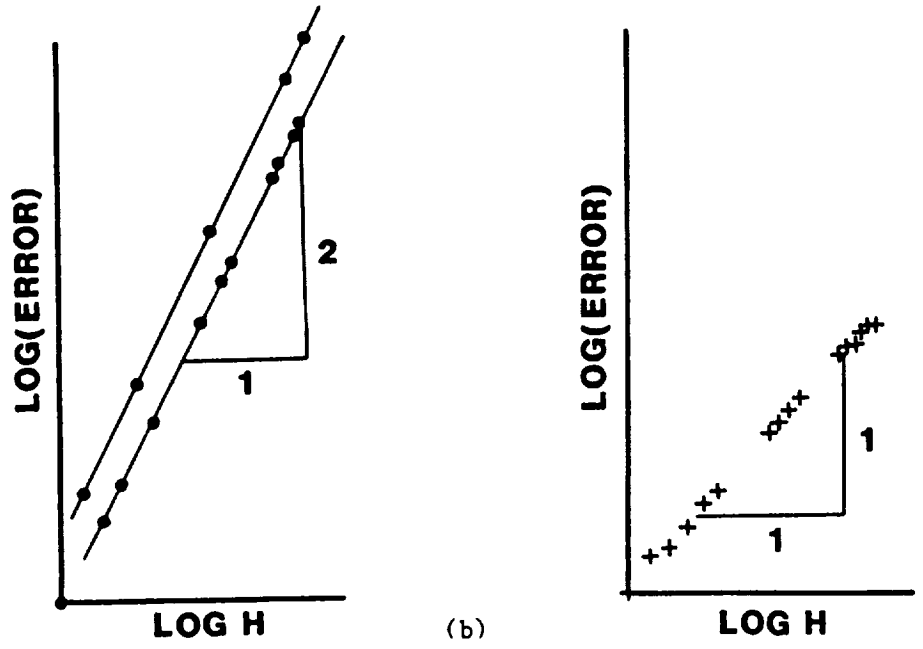
$$f_\alpha \in H^s(\Omega) \qquad \alpha > s - 1 \tag{5.17}$$

The result shown in Fig. 7.a is a plot of α (regularity) versus σ (rate of convergence of $\| \tilde{u}^h - u^h \|_{s=0,1}$) . The pattern of the (α, σ) plot seems to show a linear increase of slope 1 towards the maximum value 2 reached for $f \in L^2 (\alpha = -1)$ for the L^2-norm ($s=0$) . As far as the H^1-norm ($s=1$) of the error is concerned, the linear increase of slope 1 reached 1 for $f \in L^2$ but keeps increasing towards 2 . This shows that the expected error estimate

$$\| u^h - \tilde{u}^h \|_s \leq C\, h^k \, \| f \|_m \quad , \; s = 0,1 \tag{5.18}$$



(a)



(b)

Figure 6. Results Obtained with the Continuous Data f_2 : Comparison between : a) u^h and $\tilde{u}^h$; b) u^h and $\bar{u}^h$.

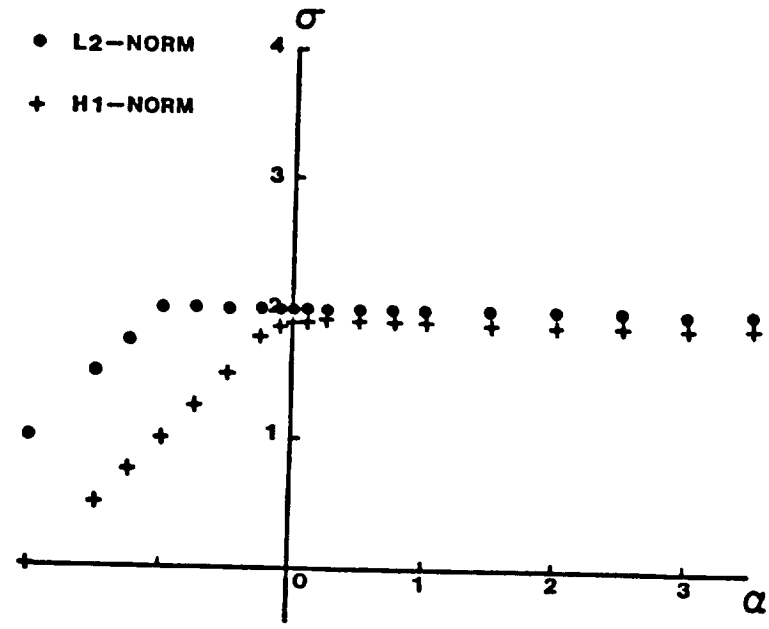


Figure 7.a. Bilinear Elements

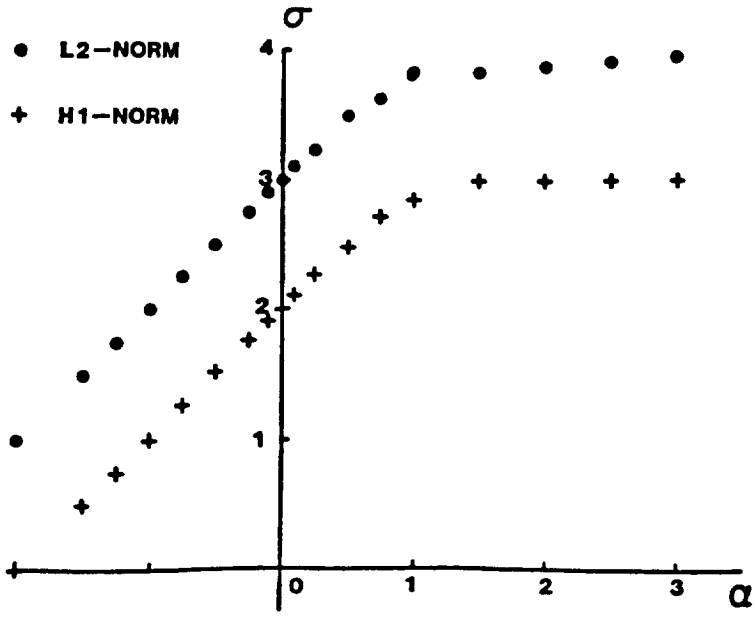


Figure 7.b. Biquadratic Elements

Figure 7. (α, σ) Plot for a Square Domain

where

$$k = 1 + \min(1, m) - s \tag{5.19}$$

is not optimal for $s = 1$ and $m > 0$. The estimate (3.32) remains optimal however.

In conclusion, these numerical results suggest that the method is accurate for regular meshes with a convergence rate equal to the fully integrated case.

2.5.2.b. Regular mesh of 8- and 9-node elements. Since the beginning of Part II we have not discussed the underintegration of the stiffness matrix of the 8-node elements. It is well known that this matrix is not rank-deficient, and the practice of the underintegration has been widely used with good results when the mesh is regular. Since there is not any spurious mode, the a-posteriori control previously described is not needed.

Unfortunately, the method of proof presented in Section 2.4 cannot be used because this element does not possess the nice tensor product properties on which the method relies. The only hope for a proof of convergence would be to obtain the result as a by-product of a result for the 9-node elements.

As far as this element is concerned (9-node element), we have proved (Section 2.2.3) that the underintegration of this element leads to a rank-deficient matrix; in fact, the procedure described in Section 2.3, for the resolution of the underintegrated problem and the projection of its solution is completely applicable to a mesh of 9-node elements. Thus, Theorem VII is valid and the projection defined in (3.3)

a convergence theorem is concerned, one can establish generalizations of (4.16) and (4.17) to 3-node, one-dimensional elements: there exist α_i , β_i , ϕ^i and ψ^i such that

$$(v^h, \phi^i)_{0,h} = \alpha_i (v^h, \phi^i)_0 \qquad \forall_{v^h}$$

$$(v^h, \psi^i)_1 = \beta_i (v^h, \psi^i)_0$$

Unfortunately, the basis functions ϕ^i and ψ^i are different for the L^2-underintegrated and H^1-norms and a lemma as Lemma 1.1 cannot be obtained.

However, in this subsection we will show numerical results obtained by use of the projection (3.30) for regular meshes of 9-node elements. Note that two types of control have been tested with similar results: the control only involving the term in $\gamma.\gamma^T$ predicted by Belytschko [4] and the complete control calculated with $s_i s_i^T$, i = 7, 8,9. (See Section 2.2.3). The results obtained with either of them are similar for this operator $(-\Delta)$.

For 8 and 9-node elements, the optimal rates of convergence are given by

$$\| u-u^h \|_s \leq C\, h^k \, \| f \|_m \qquad s = 0,1 \tag{5.20}$$

where

$$k = 2 + \min(1,m) - s \tag{5.21}$$

and the best rates of convergence $O(h^{3-s})$ are obtained when $f \in H^1(\Omega)$. The results obtained with functions presenting a singularity line (such as the functions f_1 and f_2 previously defined and others) are pre-

sented in Table 2 (first and second lines). We obtained 1.99 and 1.74 for a discontinuous function ($f \in L^2$), then 2.43 and 1.97 for a continuous, not C^1, function ($f \in H^1$), 2.95 and 1.95 for a C^1, not C^2 function ($f \in H^2$) and finally 4 and 3 for a C^∞ data. Therefore, the rates 3 and 2 are reached when f is at least H^2 or equivalent when the solution u is in H^4. In this case, the convergence rate (5.1) does not seem to be reached.

The second series of data involving the singularity at the origin (5.16) has been tested and results are shown in Fig. 6.b. The pattern of the (α,σ) plot shows linear increases of slope 1, the predicted values 3 and 2 are reached for $f \in H^1(r)$ according to (5.21), but the maximum values 4 and 3 are reached for $f \in H^2(\Omega)$.

2.5.2.c. Irregular mesh of 4- and 9- node elements. Finally, the method has been tested on the quarter unit disk shown in Fig. 8 with

$$f = r^\alpha - \frac{2}{\alpha+2} \qquad \alpha > -2$$

The plot (α,σ) is shown in Fig. 9 and we can point out:

- The general pattern is respected (linear increase towards a maximum value)
- The maximum values 4 and 3 (9-node elements) are reduced to values slightly lower than 3 and 2.

2.6 Excitation of Spurious Modes

The previous sections were devoted to the study of the Laplace equations with Neumann boundary conditions. The choice of these boundary conditions is convenient for the analysis of the hourglass instabil-

Table 2. Rate of Convergence $|\text{Log}\|e\|_s|$ v.s $|\text{Log } h|$ (0=0,1) for 8- and 9-node Elements

REGULARITY / BOUNDARY CONDITIONS AND ELEMENT	$f \in L^2$	$f \in H^1$	$f \in H^2$	$f \in C^\infty$
	$\in H^1$	$\in H^2$	$\in H^3$	
NEUMANN,9-NODE	1.99	2.43	2.95	4.00
+SPECIAL PROJECTION	1.74	1.97	1.94	3.00
NEUMANN,8-NODE EL.	1.99	2.00	2.97	4.00
	1.79	1.93	1.95	3.00
DIRICHLET,9-NODE EL.	2.35	2.85	2.99	3.00
	1.47	1.99	1.99	2.00
DIRICHLET,8-NODE EL.	2.30	2.71	2.99	3.00
	1.46	2.12	1.99	2.00
MIXED,9-NODE EL.	2.00	2.00	3.83	4.00
	1.67	2.28	2.84	3.00
MIXED,8-NODE EL.	2.00	2.00	3.84	4.00
	1.74	2.27	2.84	3.00

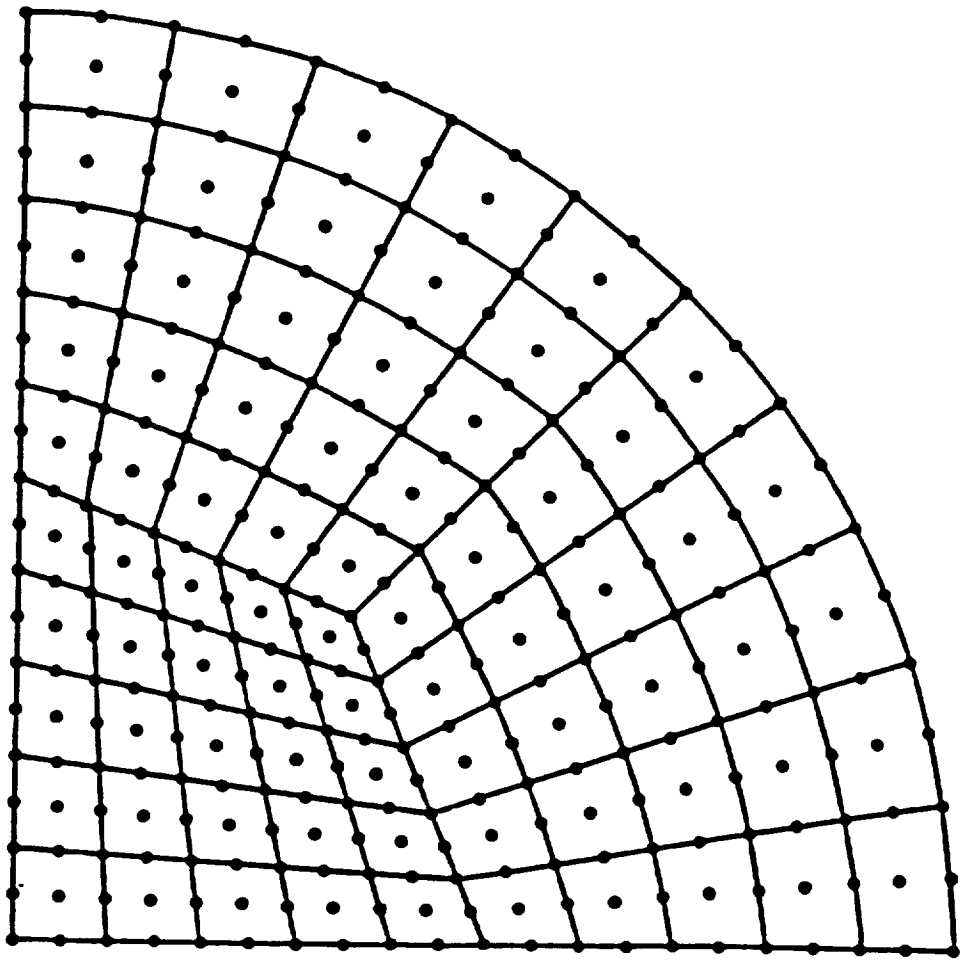


Figure 8. Typical Mesh on a Quarter Circle Domain.

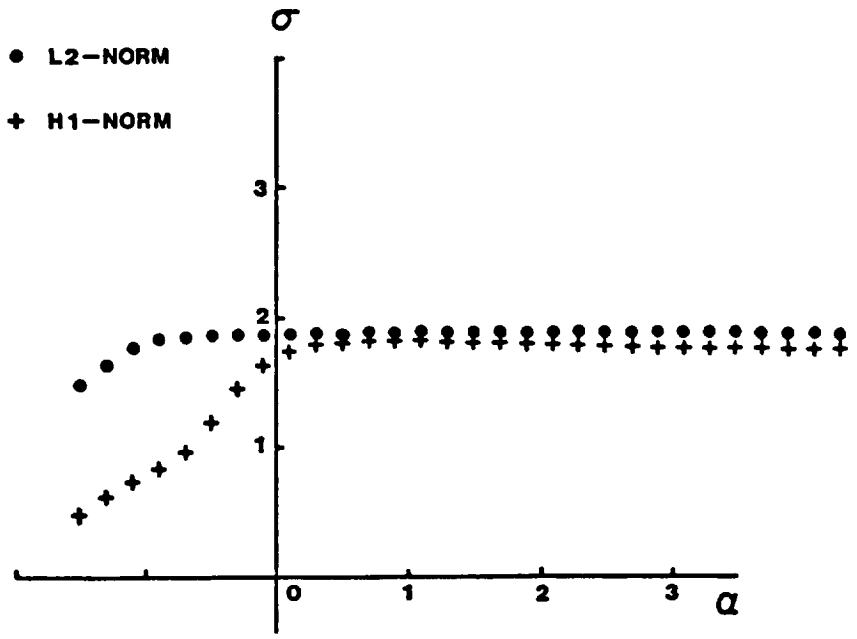


Figure 9.a. Bilinear Elements.

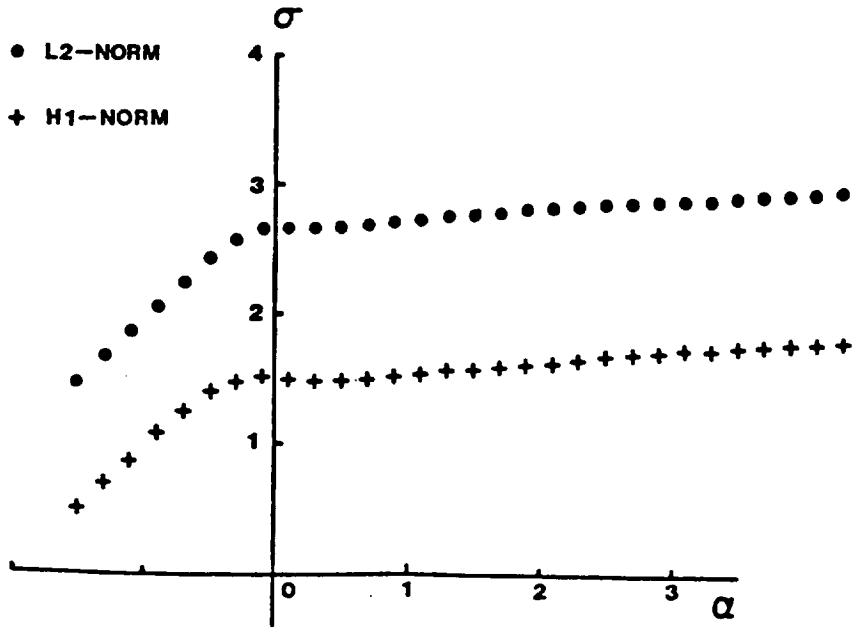


Figure 9.b. Biquadratic Elements.

Figure 9. (α,σ) Plot for a Quarter Circle Domain.

ities because these modes appear explicitly in the kernel of the underintegrated discrete operator. When Dirichlet conditions are applied on a part of the boundary, even though the kernel of the underintegrated stiffness matrix is not rank-deficient, instabilities may appear.

In this section we would like to study the influence the boundary conditions have on the solution of the underintegrated problem, and obtain results analogous to Theorem VIII. Also we would like to explain how the oscillations may be excited in certain problems. The method of proof is similar to that presented in Section 2.4. For various boundary conditions, we are able to exhibit the exact eigenvalues and eigenfunctions of the various linear and bilinear form involved. The explanation of the excitation of oscillations will result from the comparison of these eigenvalues. The procedure also allows us to study the underintegration of the operator $-\Delta+1$ and the control of resulting spurious modes. Numerical results will illustrate the theory.

2.6.1. The underintegrated problem with Dirichlet or mixed boundary conditions. This section is devoted to a generalization of the results obtained in Section 2.4 to the Laplacian equation with Dirichlet or mixed Dirichlet-Neumann boundary conditions. Only proofs for the Dirichlet case will be given in this section, but their equivalent for the mixed case can be found in Appendix A.

The Dirichlet case is simpler than the Neumann case because the hourglass mode does not belong to the new approximation space defined to handle the boundary condition. Therefore the stiffness matrix is no longer singular and can be inverted directly. In the variational formulation, similar to (3.7), the projection of the data function is not

necessary and the problem $\bar{P}^h$ is written:

$(\bar{P}^h)$ Find $\bar{u}^h \in V_0^h$ such that

$$(\bar{u}^h , v^h)_{1,h} = (f^h , v^h)_{0,h} \qquad v^h \in V_0^h \qquad (6.1)$$

where

$$V_0^h = \{v^h / v^h \in C^0(\bar{\Omega}) , v^h|_{\Omega ij} \in Q_1(\Omega ij) , 1 \leq i,j \leq N ,$$

$$v^h|_{\partial\Omega} = 0\}$$

Remarks

i) The fact that $(\cdot,\cdot)_{1,h}$ is not singular on V_0^h does not make the problem classically elliptic in the sense that the constant in the Lax Milgram Theorem is h-dependent.

ii) When Dirichlet or mixed boundary conditions are applied,

$$\text{Ker } A_1 = \text{Ker } A_{1,h} = \{0\}$$

Thus the post processor is not justified anymore and we may compare directly $\bar{u}^h$ and u^h.

This comparison is carried the same way as in Section 2.4 and a basis of the approximation-space V_0^h can be obtained. One useful basis is the common eigen-basis of the matrices of the H^1-, L^2-, and underintegrated H^1- or L^2- inner product. Let us consider the N-1 x N-1 matrix

$$D(k) = \begin{bmatrix} 2k & -1 & & 0 \\ -1 & & & \\ & & & -1 \\ 0 & & -1 & 2k \end{bmatrix}$$

The values for which det D(k) vanishes are

$$k_i = \cos \frac{i\pi}{N} \qquad 1 \leq i \leq N-1 \tag{6.2}$$

and the corresponding vectors $(D(k_i)v_i = 0)$ are

$$v_i = \{\sin \frac{ij\pi}{N}\}_{j=1,\ N-1} \tag{6.3}$$

Let $\phi^i = \phi^i(x)$, $x \in [0,1]$, denote the piecewise linear function associated with the vector v_i :

$$\phi^i(jh) = \sin \frac{ij\pi}{N} \qquad 1 \leq i,j \leq N-1 \tag{6.4}$$

$$\text{span}\{\phi^i\}_{1 \leq i \leq N-1} = V^h_{1,0} \tag{6.5}$$

where

$$V^h_{1,0} = \{v^h \in C^0(0,1),\ v^h(0) = v^h(1) = 0$$

$$v^h \text{ is linear on } eh,\ (e+1)h\ ,\ 0 \leq e \leq N-1\}$$

From this point, the remainder of the proof goes as in Section 2.4 and the variational problem and its underintegrated formulation can be explicitly solved and the decomposition (4.29), (4.30) and (4.31), (4.32) are obtained for $1 \leq i,j \leq N-1$, and we finally obtain the result for Dirichlet boundary conditions:

THEOREM V: *Let* f *be a function in* $L^2(\Omega)$. *Let* u *be the solution of* P :

$$P: \text{Find } u \in H^1_0 \ /\ (u,v)_1 = (f,v)_0 \qquad \forall v \in H^1_0 \tag{6.6}$$

Let f^h *be the* L^2*-projection of* f *onto* V^h *and let* $\bar{u}^h$ *be the solution of* $\bar{P}^h$:

$$\bar{P}^h\text{: Find } \bar{u}^h \in V_0^h \;/\; (\bar{u}^h, v^h)_{1,h} = (f^h, v^h)_{0,h} \quad \forall v^h \in V_0^h \qquad (6.7)$$

Then we have the following error estimate:

$$\| u - \bar{u}^h \|_s \leq C\, h^{2-s}\, \| f \|_0 \qquad s = 0,1 \qquad \square \qquad (6.8)$$

This theorem proves that the use of the underintegrated matrix does not affect the rate of convergence of the solution. The method is therefore accurate and efficient.

Data with various regularity have been tested for meshes of 4-, 8-, and 9-node elements, with various boundary conditions. Results are summarized in Tables 2 and 3. They indicate that the optimal rates of convergence are obtained for $f \in L^2(\Omega)$ for the 4-node case and $f \in H^2(\Omega)$ for the 8- and 9-node case.

2.6.2. The underintegration of the operator $-\Delta+1$. In this subsection we consider the underintegration of the operator associated with the problem

P_0 : Find $u \in H^1(\Omega)$ such that

$$\begin{cases} -\Delta u + u = f & \text{in } \Omega \\ \dfrac{\partial u}{\partial n} = 0 & \text{on } \partial\Omega \end{cases} \qquad (6.9)$$

The usual variational formulation of P_0 is

P : Find $u \in H^1(\Omega)$ such that

$$(u, v)_1 + (u, v)_0 = (f, v)_0\,, \qquad v \in H^1(\Omega) \qquad (6.10)$$

The results of existence, uniqueness of solutions of P are well-known

Table 3: Rate of Convergence $|\text{Log } \|e\|_s|$ v.s. $|\text{Log } h|$ for 4-node Elements (s=0,1)

OPERATOR AND BOUNDARY CONDITIONS	REGULARITY	$f \in L^2$ $\notin H^1$		$f \in H^1$ $\notin H^2$		$f \in C^\infty$	
$-\Delta$	NEUMANN +PROJECTION	1.99	1.61	2.00	2.00	2.00	2.00
	DIRICHLET	1.99	1.50	2.00	1.85	2.00	2.00
	MIXED	2.00	1.50	2.00	1.99	2.00	1.99
$-\Delta + 1$	NEUMANN +PROJECTION	2.00	1.50	2.00	2.00	2.00	2.00
	DIRICHLET	2.00	1.50	2.00	1.85	2.00	2.00
	MIXED	2.00	1.50	2.00	1.85	2.00	2.00

as are those for its discrete formulation:

P_h : Find $u^h \in V^h$ such that

$$(u^h, v^h)_1 + (u^h, v^h)_0 = (f, v^h)_0 , \quad \forall v^h \in V^h \tag{6.11}$$

where V^h is an approximation of $H^1(\Omega)$ using bilinear elements. The underintegration of $(\cdot,\cdot)_1 + (\cdot,\cdot)_0$ leads to the following underintegrated problem:

$\bar{P}_h$: Find $\bar{u}^h \in \bar{V}^h$ such that

$$(\bar{u}^h, v^h)_{1,h} + (\bar{u}^h, v^h)_{0,h} = (f, v^h)_{0,h} , \quad \forall v^h \in \bar{V}^h \tag{6.12}$$

where the choice of approximation space

$$\bar{V}^h = V^h/H \tag{6.13}$$

is justified by

$$(v^h, H)_{1,h} + (v^h, H)_{0,h} = 0 , \quad \forall v^h \in V^h \tag{6.14}$$

Then, the method of proof used in Section 2.3 allows us to obtain the existence and uniqueness of $\bar{u}^h$. A projection similar to (3.30) can be obtained by analogy: we have

$$(u^h, H)_1 + (u^h, H)_0 = (f, H)_0 \tag{6.15}$$

We therefore construct the projection as:

$$\tilde{u}^h = \pi\bar{u}^h = \bar{u} + \lambda_0 H \tag{6.16}$$

$$(\tilde{u}^h, H)_1 + (\tilde{u}^h, H)_0 = (f, H)_0 \tag{6.17}$$

This defines uniquely λ_0 as

$$\lambda_0 = \frac{(f,H)_0 - (\bar{u},H)_1 - (\bar{u},H)_0}{\|H\|_1^2 + \|H\|_0^2} \tag{6.18}$$

Similarly to what was done in Section 2.5.1, we can use (4.25), (5.6) through (5.8), simplify λ_0 without any loss of accuracy and still use (5.9), (5.10) for the projection

$$\tilde{u}^h = \bar{u}^h - \lambda H \qquad \text{(5.9)(repeated)}$$

$$\lambda = (\bar{u}^h, H)_1 \, \|H\|_1^{-2} \qquad \text{(5.10)(repeated)}$$

The proof of the convergence of $\tilde{u}^h$ towards u is again done by direct calculation of u^h and $\tilde{u}^h$: the explicit resolution of P_h and $(\bar{P}_h + \pi)$ leads to :

$$u_{i,j} = \frac{1}{1 + \beta_i + \beta_j} f_{ij} \qquad 0 \leq i,j \leq N \tag{6.19}$$

and

$$\begin{cases} \tilde{u}_{i,j} = \dfrac{\alpha_i \alpha_j}{\alpha_i\alpha_j + \alpha_i\beta_j + \alpha_j\beta_i} f_{ij} & \begin{matrix} 0 \leq i,j \leq N \\ (i,j) \neq (N,N) \end{matrix} \\ \\ \tilde{u}_{NN} = u_{NN} \end{cases} \tag{6.20}$$

These decompositions allow us to obtain $u^h - \tilde{u}^h$ as done in Section (2.4). Provided that $f \in L^2(\Omega)$, we can obtain

$$\|u^h - \tilde{u}^h\|_s \leq C \, h^{2-s} \|f\|_0 \qquad s = 0,1 \tag{6.21}$$

Once again, the underintegration does not seem to affect the rate of convergence. The result can also be obtained with various boundary

conditions. Numerical results are summarized in Tables 5 and 6 and confirm the theory.

2.6.3. Excitation of oscillations. The existence of spurious oscillations when underintegration is used is not only encountered when Neumann boundary conditions are applied on the whole boundary. In this subsection, we analyze precisely how modes that oscillate with wavelength of order h are excited when underintegration is used, whereas they are damped when the integration is exact.

For this discussion we consider the unit square

$$\Omega =]0,1[\times]0,1[\tag{6.22}$$

discretized into $N \times N$ elements. We consider the Laplace equation on Ω

$$\left.\begin{aligned} -\Delta u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \cap \{x = 0\} \\ \frac{\partial u}{\partial n} &= g \quad \text{on } \partial\Omega / \{x = 0\} \end{aligned}\right\} \tag{6.23}$$

For the first time we include two kinds of load: body forces and surface loads, and we will observe separately the effects of each of them.

The eigenfunctions associated with these particular mixed boundary conditions are constructed as in Section 2.4.

$$\chi^{ij}(x,y) = \psi^i(x)\phi^j(y), \qquad \begin{matrix} 1 \le i \le N \\ 0 \le j \le N \end{matrix} \tag{6.24}$$

where ϕ^j is defined in (4.15) (associated with Neumann boundary conditions at both ends) and ψ^i is similarly defined (see Appendix A). These functions are defined through sine and cosine functions and therefore oscillate. Among them we will distinguish "smooth" modes with

longer wavelengths (O(1)) from "irregular" modes with shorter wavelengths (O(h)). Smooth (respectively irregular) modes correspond to smaller (respectively larger) values of i or j. Examples of each extreme are shown in Fig. 10 for N=10.

The resolution of the fully integrated problem leads to the search for coefficients u_{ij} such that

$$u^h = \sum_{\substack{1\leq i\leq N \\ 0\leq j\leq N}} u_{ij}\, \chi^{ij} \tag{6.25}$$

The basis $\{\chi^{ij}\}$ is an eigenbasis for $(\cdot,\cdot)_1$ and therefore we have

$$u_{ij} = A_{ij}\left[(f,\ \chi^{ij})_0 + (g,\ \chi^{ij})_{0,\partial\Omega}\right] \tag{6.26}$$

where

$$A_{ij} = \frac{1}{\beta'_i + \beta_j} \tag{6.27}$$

with

$$\left.\begin{aligned} \beta'_i &= \frac{6}{h^2}\,\frac{1-\cos(\frac{i\pi}{N}-\frac{\pi}{2N})}{2+\cos(\frac{i\pi}{N}-\frac{\pi}{2N})} \\ \beta_j &= \frac{6}{h^2}\,\frac{1-\cos(\frac{j\pi}{N})}{2+\cos(\frac{j\pi}{N})} \end{aligned}\right\} \tag{6.28}$$

The values A_{ij} have been calculated exactly with these formulae and their values are reported in Table 4.a for N=10. The 20 highest values are in the shaded zone. We clearly can observe that

i) these values range from the highest value to 1% of this value,

ii) these values are associated with smooth modes (tensor products of smooth modes).

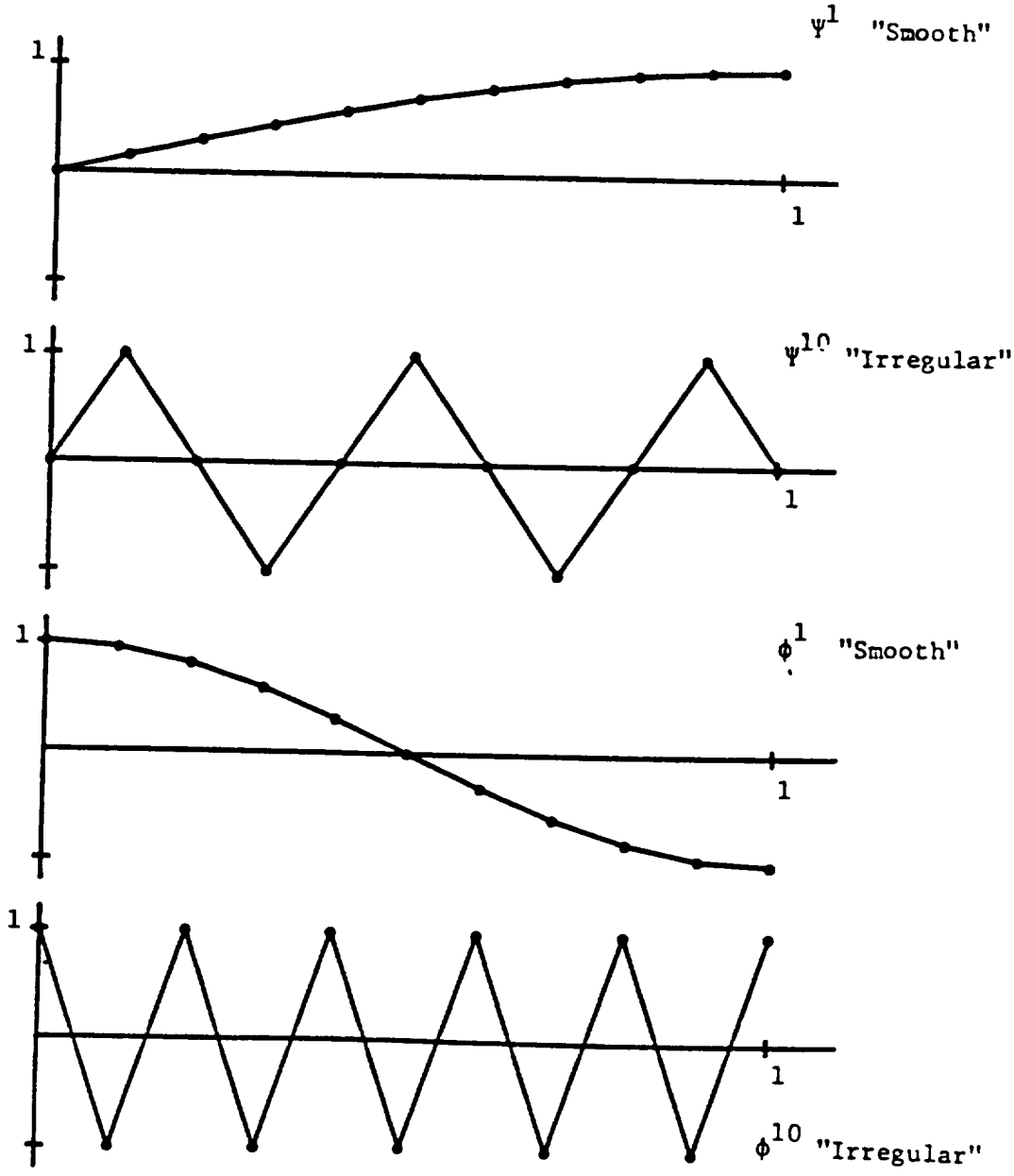


Figure 10. Examples of "Smooth" or "Irregular" Eigenfunctions.

Table 4: Arrays of Eigen Values A_{ij}, $\tilde{A}_{ij}$ and $\bar{A}_{ij}$

Table 4a. Array A_{ij}

j \ i	1	2	3	4	5	6	7	8	9	10
0	40.45	4.42	1.54	0.75	0.43	0.27	0.18	0.13	0.10	0.08
1	8.05	3.07	1.34	0.70	0.41	0.26	0.17	0.12	0.10	0.08
2	2.31	1.58	0.95	0.57	0.36	0.24	0.17	0.12	0.09	0.08
3	1.02	0.85	0.62	0.44	0.30	0.21	0.15	0.11	0.09	0.08
4	0.55	0.49	0.41	0.32	0.24	0.18	0.13	0.10	0.08	0.07
5	0.33	0.31	0.27	0.23	0.19	0.15	0.12	0.09	0.08	0.07
6	0.21	0.21	0.19	0.17	0.14	0.12	0.10	0.08	0.07	0.06
7	0.15	0.14	0.14	0.12	0.11	0.10	0.08	0.07	0.06	0.05
8	0.11	0.11	0.10	0.10	0.09	0.08	0.07	0.06	0.05	0.05
9	0.09	0.09	0.08	0.08	0.07	0.07	0.06	0.05	0.05	0.04
10	0.08	0.08	0.08	0.07	0.07	0.06	0.06	0.05	0.04	0.04

Table 4b. Array $\tilde{A}_{ij}$

j \ i	1	2	3	4	5	6	7	8	9	10
0	40.36	4.34	1.46	0.67	0.34	0.18	0.09	0.04	0.01	0.00
1	7.99	3.02	1.27	0.62	0.33	0.18	0.09	0.04	0.01	0.00
2	2.24	1.53	0.90	0.52	0.30	0.17	0.09	0.04	0.01	0.00
3	0.94	0.79	0.58	0.39	0.25	0.15	0.09	0.04	0.01	0.00
4	0.47	0.43	0.36	0.28	0.20	0.13	0.08	0.04	0.01	0.00
5	0.25	0.24	0.21	0.18	0.14	0.11	0.07	0.04	0.01	0.00
6	0.13	0.13	0.12	0.11	0.10	0.08	0.05	0.03	0.01	0.00
7	0.06	0.06	0.06	0.06	0.05	0.05	0.04	0.03	0.01	0.00
8	0.03	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.01	0.00
9	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4c. Array $\bar{A}_{ij}$

j \ i	1	2	3	4	5	6	7	8	9	10
0	40.45	4.42	1.54	0.75	0.43	0.27	0.18	0.13	0.10	0.08
1	8.08	3.11	1.36	0.71	0.42	0.26	0.18	0.13	0.10	0.09
2	2.32	1.62	0.99	0.61	0.39	0.26	0.18	0.13	0.10	0.09
3	1.02	0.87	0.67	0.48	0.34	0.24	0.18	0.13	0.10	0.09
4	0.55	0.51	0.44	0.37	0.29	0.23	0.17	0.14	0.11	0.10
5	0.33	0.32	0.30	0.27	0.24	0.20	0.17	0.14	0.12	0.11
6	0.22	0.21	0.21	0.20	0.19	0.18	0.17	0.16	0.14	0.14
7	0.15	0.15	0.15	0.15	0.15	0.16	0.17	0.17	0.18	0.19
8	0.11	0.11	0.11	0.12	0.13	0.14	0.16	0.20	0.26	0.33
9	0.09	0.09	0.09	0.10	0.11	0.13	0.16	0.23	0.42	0.97
10	0.08	0.08	0.09	0.09	0.10	0.12	0.16	0.25	0.57	4.57

On the other hand, the eigenvalues of irregular modes are smaller and because of this, these modes will be damped; only smooth modes will contribute in (6.25).

When the underintegration is used, and when g is zero, the solution $\tilde{u}^h$ is

$$\tilde{u}^h = \sum_{\substack{1\leq i\leq N \\ 0\leq j\leq N}} \tilde{u}_{ij}\, \chi^{ij} \tag{6.29}$$

with

$$\tilde{u}_{ij} = \tilde{A}_{ij}(f, \chi^{ij})_0 \tag{6.30}$$

where

$$\tilde{A}_{ij} = \frac{\alpha_i' \alpha_j}{\alpha_j \beta_i' + \alpha_i' \beta_j} \tag{6.31}$$

and

$$\left.\begin{aligned} \alpha_i' &= \frac{3(1 + \cos(\frac{i\pi}{N} - \frac{\pi}{2N}))}{2(2+ \cos(\frac{i\pi}{N} - \frac{\pi}{2N}))} \\ \alpha_j &= \frac{3(1 + \cos(\frac{j\pi}{N}))}{2(2 + \cos(\frac{j\pi}{N}))} \end{aligned}\right\} \tag{6.32}$$

Again, the values of $\tilde{A}_{ij}$ have been calculated exactly and they are reported in Table 4b. The 20 highest values are in the shaded zone. The comparison between Tables 7a and b shows that these 20 values are approximately the same and they are associated with the same smooth modes. In this case, irregular modes will still be damped , and one can predict that no oscillation will occur.

When a load is only applied on the boundary $(f = 0,\ g \neq 0)$,

$\tilde{u}_{ij}$ is now

$$\tilde{u}_{ij} = \bar{A}_{ij}(g, \chi^{ij})_{0,\partial\Omega} \tag{6.33}$$

where

$$\bar{A}_{ij} = \frac{1}{\alpha_j\beta_i' + \alpha_i'\beta_j} \tag{6.34}$$

Again the values of $\bar{A}_{ij}$ are reported in Table 4.c and the 20 highest values are in the shaded zone. Among these 20 values, three correspond to very irregular modes. In particular, the third value is associated with $\chi^{10,10}$. Therefore, we can predict a strong contribution of irregular modes within the solution $\tilde{u}^h$, which will show oscillations.

Finally, one could wonder if the calculation of the boundary integral can be calculated such that $(g, \chi^{ij})_{0,\partial\Omega}$ is damped for large i and j. Unfortunately, no precise method has been obtained. In particular, if the load g is a concentrated load at (x_0, y_0), then

$$(g, \chi^{ij})_{0,\partial\Omega} = \chi^{ij}(x_0, y_0) \tag{6.35}$$

and this value is not necessarily zero. A procedure, consisting in forcing $(g, \chi^{ij})_{0,\partial\Omega}$ to zero in (6.35), can be obtained in the case in which the load is concentrated. This procedure is discussed in a more general context in the next subsection.

2.6.4. A-priori Orthogonalization of the Data. Our purpose in this subsection is to illustrate how simple orthogonalization consideration can drastically prevent the excitation of spurious modes. In particular, one interprets the existence of artificial anti-hourglass

forces that can be applied to damp spurious oscillations.

In order to solve the underintegrated system

$$\overline{A}\,\overline{U} = \overline{F} \tag{6.36}$$

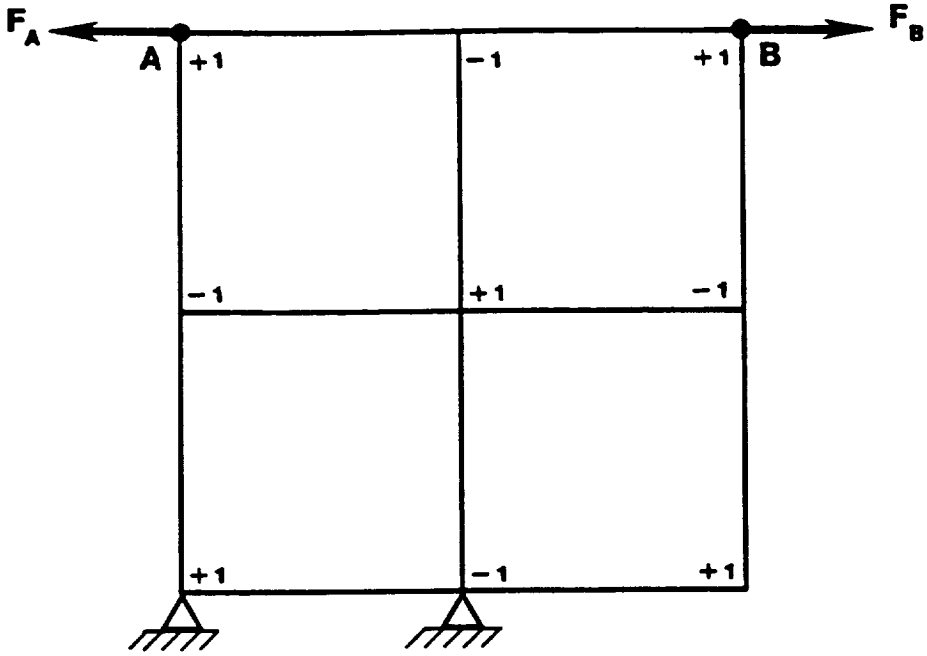
we must have the orthogonality condition

$$\overline{F} \in (\ker \overline{A})^T \tag{6.37}$$

From a practical point of view, the applied load must be orthogonal to the spurious mode (compare with 6.35). If this is not the case, concentrated spurious forces may be induced at the nodes that have been fixed to solve the system *modulo* spurious modes.

This is clearly demonstrated in the following example where Ω is a square partitioned into 4 or 9 4-node elements. For the Laplace equation, we consider two loads concentrated in two opposite points and satisfying the equilibrium condition as shown in Fig. 11. When the number of elements is even (Fig. 11.a), the system of forces are orthogonal to the hourglass mode H(±1-pattern indicated at the nodes) and, therefore, the reactions at the fixed points are zero as desired. Conversely, when the number of elements is odd (Fig. 11.b), the system of forces is not orthogonal to H and two reactions R_1 and R_2 appear at the fixed points such that the system F_A, F_B, R_1 and R_2 is orthogonal to both translation and hourglass mode. Such considerations can explain the peculiar rates of convergence obtained on the unit square discretized with NxN elements with the data

$$f = \begin{cases} -6x(y^2-1)^2 + 4x^3(1-3y^2) & \text{if } x < \tfrac{1}{2} \\ \tfrac{3}{2}(y^2-1)^2 + \tfrac{1}{2}(-12x^2 + 24x - 7)(1-3y^2) & \text{if } x > \tfrac{1}{2} \end{cases} \tag{6.38}$$



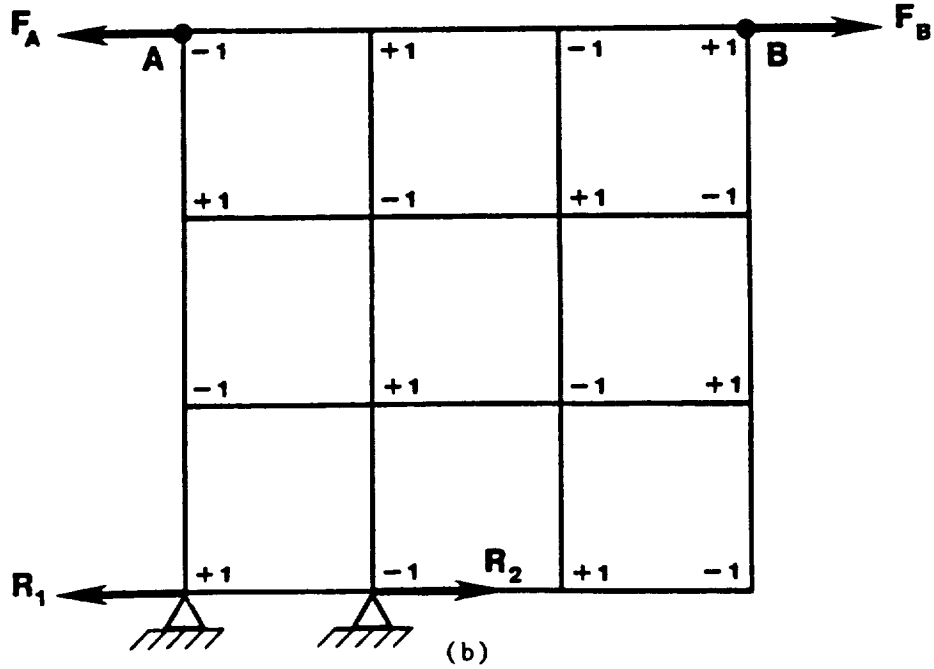


Figure 11: Spurious Reactions: a) Even Mesh and
b) Odd Mesh

Figure 12.a shows the rates of convergence of $\tilde{u}^h$ towards u^h in the $L^2/\mathbb{R}$- and H^1 norms, where $\tilde{u}^h$ is obtained from the underintegrated problem and the projection, and u^h is the solution of the fully integrated problem. We can clearly see the good (more than optimal) behavior of the rates when the mesh is even; in this case, the discontinuity line ($x = \frac{1}{2}$) corresponds to a mesh line. However, the quality of the solution deteriorates when the mesh is odd or when the discontinuity line is across the mesh and coincides with the integration points. Note that when f_1 is averaged on $x = \frac{1}{2}$:

$$f(\tfrac{1}{2}, y) = \tfrac{1}{2}(f(\tfrac{1}{2}^+, g) + f(\tfrac{1}{2}^-, y)) \tag{6.39}$$

then all the points on both plots lay on one straight line. Also note that the knowledge of the exact solution

$$u = \begin{cases} x^3 (y^2 - 1)^2 & \text{if } x < \frac{1}{2} \\ \frac{1}{16}(-12x^2 + 24x - 7)(y^2 - 1) & \text{if } x > \frac{1}{2} \end{cases} \tag{6.40}$$

The interpretation of (6.37) has been possible when Neumann boundary conditions are applied. When Dirichlet conditions are applied, any right-hand side vector can produce a solution, but essentially the same type of behavior is observed (Figure 12.b). These results, the interpretations of (6.35) and (6.37) suggest that for any problem, one must pay attention to the data and make sure that it is orthogonal to the spurious modes. In particular, the procedure consisting in splitting a concentrated load between neighboring elements seems to give good results and will be discussed for elasticity problems in the next section.

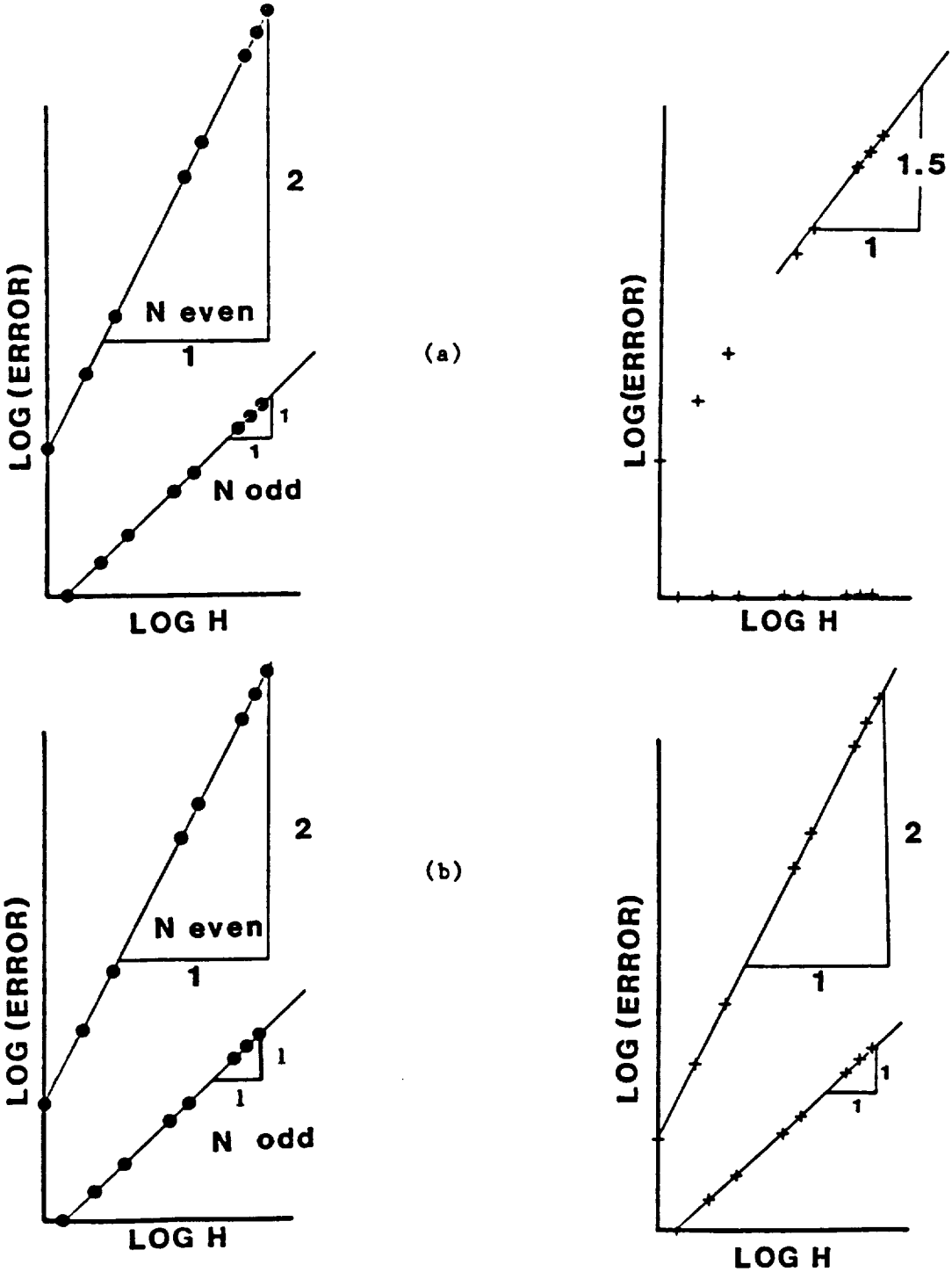


Figure 12 Rates of Convergence Obtained with a Discontinuous Data.

2.7 The Practice of Underintegration in Linear Elasticity

We devote this section to the discussion of the effects of underintegration in linear elasticity. We first exhibit the kernel of the underintegrated operator and obtain a post-processor formula similar to (3.30) to a-posteriori control the spurious modes. This global control can only be used in a very limited number of problems, but it suggests a local control that gives satisfactory results for all the examples considered. We discuss this control, its easy implementation, and several numerical results.

This study is entirely qualitative -- the basis functions obtained in Section 2.4 cannot be used to obtain eigen-functions for the elasticity operator. Both 4- and 9-node elements are discussed with an emphasis on the 9-node element in the numerical studies.

2.7.1. The kernel of the discrete underintegrated linear elasticity operator. We consider the linear elasticity operator defined by

$$\underset{\sim}{A} = \underset{\sim}{\beta}^T \underset{\sim}{C} \underset{\sim}{\beta} \tag{7.1}$$

where

$$\underset{\sim}{\beta} = \begin{pmatrix} \frac{\partial}{\partial x} & 0 & \frac{\partial}{\partial y} \\ 0 & \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{pmatrix}^T \tag{7.2}$$

and $\underset{\sim}{C}$ is a 3x3 symmetric matrix. In the plane strain case we may particularize $\underset{\sim}{C}$:

$$\underset{\sim}{C} = \begin{pmatrix} \lambda+2\mu & \lambda & 0 \\ \lambda & \lambda+2\mu & 0 \\ 0 & 0 & \mu \end{pmatrix} \tag{7.3}$$

In order to exhibit the spurious modes we consider the operator A associated with Neumann boundary conditions. In that case, the kernel of A consists of the usual 2-dimensional rigid body modes denoted by $\underset{\sim}{t}_x$, $\underset{\sim}{t}_y$ and $\underset{\sim}{r}$

$$\text{RBM} = \text{span}\left\{\underset{\sim}{t}_x = \begin{pmatrix}1\\0\end{pmatrix} ; \underset{\sim}{t}_y = \begin{pmatrix}0\\1\end{pmatrix} ; \underset{\sim}{r} = \begin{pmatrix}-y\\x\end{pmatrix}\right\} \tag{7.4}$$

We consider the problem

$$P : \text{Find } \underset{\sim}{u} = \begin{pmatrix}u_1\\u_2\end{pmatrix} \in \left[H^1(\Omega)\right]^2/\text{RBM} \text{ such that}$$

$$\int_\Omega \underset{\sim}{u}^T \underset{\sim}{\beta}^T \underset{\sim}{C} \underset{\sim}{\beta} \underset{\sim}{v} \, dxdy = \int_\Omega \underset{\sim}{f}^T \cdot \underset{\sim}{v} \, dxdy \tag{7.5}$$

where $\underset{\sim}{f} = (f_1, f_2)^T$ is a force satisfying the equilibrium conditions

$$\underset{\sim}{f} \in \text{RBM}^T \tag{7.6}$$

or equivalently

$$\left.\begin{aligned}\int_\Omega f_1 \, dxdy = \int_\Omega f_2 \, dxdy = 0\\ \int_\Omega (f_1 y - f_2 x) \, dxdy = 0\end{aligned}\right\} \tag{7.7}$$

The existence and uniqueness of a solution for P are well known. The construction of finite element approximations of (7.5) involves the calculation of the (2Nx2N) stiffness matrix $\underset{\sim}{K}_e$ for a typical element Ω_e , which is given by the formula

$$\underset{\sim}{K}_e = \int_\Omega \underset{\sim}{N}^T \underset{\sim}{\beta}^T \underset{\sim}{C} \underset{\sim}{\beta} \underset{\sim}{N} \, dxdy \tag{7.8}$$

where $\underset{\sim}{N}$ is a vector representing the bilinear (N=4) or biquadratic (N=9) shape functions in each element Ω_e , $1 \leq e \leq E$. In computational applications K_e is evaluated using an integration rule:

$$\underset{\sim}{K}_e = \sum_{\alpha=1}^{L} W_\alpha \underset{\sim}{B}^{\alpha T} \underset{\sim}{C} \underset{\sim}{B}^{\alpha} \tag{7.9}$$

where, similar to (2.36),

$$\underset{\sim}{B}^{\alpha} = \begin{pmatrix} \underset{\sim}{b}_1^{\alpha} & 0 & \underset{\sim}{b}_2^{\alpha} \\ 0 & \underset{\sim}{b}_2^{\alpha} & \underset{\sim}{b}_1^{\alpha} \end{pmatrix}^T \tag{7.10}$$

and W_α is the weight at the integration point α. Simple rank considerations allow us to predict the rank of $\underset{\sim}{K}_e$. Indeed, since

$$\left.\begin{aligned} &\mathrm{rk}(A \cdot B) \leq \max(\mathrm{rk}\, A, \mathrm{rk}\, B) \\ \text{and}\quad &\mathrm{rk}(A + B) \leq \mathrm{rk}\, A + \mathrm{rk}\, B \end{aligned}\right\} \tag{7.11}$$

we have

$$\mathrm{rk}\, \underset{\sim}{K}_e \leq 3\, L \tag{7.12}$$

When the full integration is used, (7.12) provides no information, but we know that $\underset{\sim}{K}_e^{full}$ has the correct kernel containing only rigid body modes. However, when underintegration (L=1) is used on 4-node elements, we have

$$\mathrm{rk}\, \underset{\sim}{K}_e \leq 3 \tag{7.13}$$

Therefore, the 8x8 matrix K_e possesses at least two spurious modes. In fact, only two are present. Similarly, when underintegration (L=4) is used and 8- or 9-node elements, we have

$$\mathrm{rk}\, \underset{\sim}{K}_e \leq 12 \tag{7.14}$$

This inequality predicts one spurious mode for the 16x16 matrix associ-

ated with 8-node elements, but when the procedure is repeated for two neighboring 8-node elements, the spurious modes can no longer exist in the global matrix [53]. We can also interpret this elimination of the spurious mode by noticing that neighboring element cannot share the mode [14]. Also note that, in spite of failing an element stability test [24], there are no extraneous zero energy modes in the kernel of the underintegrated stiffness matrix and this element is stable.

As far as 9-node elements are concerned, the inequality (7.14) indicates that the 18x18 stiffness matrix has at least three spurious modes. In fact, there are exactly three such modes and they can be shared by adjacent elements. Next the modes will be explicityly described.

2.7.1.a. The Spurious Modes for 4-node Elements. Let $\underset{\sim}{H}_x$ and $\underset{\sim}{H}_y$ be the two hourglass vectors defined as

$$\underset{\sim}{H}_x = \begin{pmatrix} h \\ 0 \end{pmatrix} \qquad \underset{\sim}{H}_y = \begin{pmatrix} 0 \\ h \end{pmatrix} \tag{7.15}$$

where h^* is the hourglass nodal displacement defined in (2.14). Then when $L=1$, we obtain from (2.18) and (2.20)

$$\underset{\sim}{B}^1 \cdot \underset{\sim}{H}_x = \begin{pmatrix} \underset{\sim}{b}_1^T \cdot \underset{\sim}{h} \\ 0 \\ \underset{\sim}{b}_2^T \underset{\sim}{h} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

and similarly

$$\underset{\sim}{B}^1 \cdot \underset{\sim}{H}_y = 0$$

---

* In this section, nodal values and associated functions will be denoted by the same letter, the underlining "~" differentiating them. The nodal values are expressed component by component.

Therefore,

$$\underset{\sim}{K}_{e}^{(1)} \cdot \underset{\sim}{H}_{x} = \underset{\sim}{K}_{e}^{(1)} \cdot \underset{\sim}{H}_{y} = 0 \tag{7.16}$$

These element displacements can be put together to obtain two global spurious modes, also denoted by $\underset{\sim}{H}_{x}$ and $\underset{\sim}{H}_{y}$ and we have :

$$\text{Ker } K^{\text{under}} = \{\text{span } \underset{\sim}{t}_{x}, \underset{\sim}{t}_{y}, \underset{\sim}{r}, \underset{\sim}{H}_{x}, \underset{\sim}{H}_{y}\} \tag{7.17}$$

This defines entirely the kernel of the underintegrated matrix and the spurious modes for 4-node elements.

Remark: In problems where symmetry is used for simplifications, the kernel of $\underset{\sim}{K}^{\text{under}}$ must respect the symmetry. If one axis of symmetry (say, the x-axis) exists, then

$$\text{Ker } \underset{\sim}{K}^{\text{under}} = \text{span } \{\underset{\sim}{t}_{x}, \underset{\sim}{H}_{x}\}$$

If the problem has two axes of symmetry (x- and y-axis)

$$\text{Ker } \underset{\sim}{K}^{\text{under}} = \{\underset{\sim}{0}\}$$

The spurious modes are eliminated by the symmetry conditions.

2.7.1.b. The spurious modes for 9-node elements. Let $\underset{\sim}{H}_{x}$ and $\underset{\sim}{H}_{y}$ be the two vectors defined as

$$\underset{\sim}{H}_{x} = \begin{pmatrix} h \\ 0 \end{pmatrix} \qquad \underset{\sim}{H}_{y} = \begin{pmatrix} 0 \\ h \end{pmatrix} \tag{7.18}$$

where h is the spurious mode of 9-node elements defined in (2.37). Using (2.36), (2.39) and (7.9), we easily get

$$\underset{\sim}{K}_{e}^{(4)} \cdot \underset{\sim}{H}_{x} = \underset{\sim}{K}_{e}^{(4)} \cdot \underset{\sim}{H}_{y} = \underset{\sim}{0} \tag{7.19}$$

Therefore, $\underset{\sim}{H}_{x}$ and $\underset{\sim}{H}_{y}$ are two out of the three spurious modes of

$\underset{\sim}{K}_e^{(4)}$. We remark that the pattern (2.37) defining them does not depend on the geometry of the mesh. As far as the third spurious mode, denoted by

$$\underset{\sim}{W} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \tag{7.20}$$

is concerned, one can show that the equations defining it are

$$\underset{\sim}{b}_1^{\alpha T} \cdot \underset{\sim}{w}_1 = \underset{\sim}{b}_2^{\alpha T} \cdot \underset{\sim}{w}_2 = \underset{\sim}{b}_1^{\alpha T} \cdot \underset{\sim}{w}_2 + \underset{\sim}{b}_2^{\alpha T} \cdot \underset{\sim}{w}_1 = 0 \; ; \; \alpha = 1,4 \tag{7.21}$$

or equivalently :

$$\left.\begin{aligned} \underset{\sim}{y}^T \underset{\sim}{A}^\alpha \underset{\sim}{w}_1 &= 0 \\ \underset{\sim}{x}^T \underset{\sim}{A}^\alpha \underset{\sim}{w}_2 &= 0 \\ \underset{\sim}{x}^T \underset{\sim}{A}^\alpha \underset{\sim}{w}_1 &= \underset{\sim}{y}^T \underset{\sim}{A}^\alpha \underset{\sim}{w}_2 \end{aligned}\right\} \quad \alpha = 1,4 \tag{7.22}$$

Note that for this system of 12 equations, we have 18 unknowns. If we add 5 orthogonality equations between $\underset{\sim}{W}$ and $\underset{\sim}{t}_x$, $\underset{\sim}{t}_y$, $\underset{\sim}{r}$, $\underset{\sim}{H}_x$ and $\underset{\sim}{H}_y$, the system will define only one $\underset{\sim}{W}$ (up to within a multiplicative factor). For a general geometry of Ω_e, one cannot exhibit an explicit form for W ; however, when Ω_e is a quadrilateral (strait-sided), and when $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are of the form

$$\underset{\sim}{x} = (x_1, x_2, x_3, x_4, \tfrac{1}{2}(x_1 + x_2), \tfrac{1}{2}(x_2 + x_3),$$

$$\tfrac{1}{2}(x_3 + x_4), \tfrac{1}{2}(x_4 + x_1), \tfrac{1}{4}(x_1 + x_2 + x_3 + x_4)) \tag{7.23}$$

we can prove that one candidate for $\underset{\sim}{W}$ can be written as

$$\left.\begin{aligned} \underset{\sim}{w}_1 &= \underset{\sim}{T}\, \underset{\sim}{y}' \\ \underset{\sim}{w}_2 &= -\underset{\sim}{T}\, \underset{\sim}{x}' \end{aligned}\right\} \tag{7.24}$$

where

$$\underset{\sim}{T} = \begin{pmatrix} 4 & 2 & 0 & 2 & -1 & 0 & 0 & -1 & 0 \\ -2 & -4 & -2 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 2 & 4 & 2 & 0 & -1 & -1 & 0 & 0 \\ -2 & 0 & -2 & -4 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}^T \tag{7.25}$$

and $\underset{\sim}{x}'$, $\underset{\sim}{y}'$ are the vectors constructed with the first four components of $\underset{\sim}{x}$ and $\underset{\sim}{y}$. An example of $\underset{\sim}{W}$ for a geometry satisfying (7.23) is shown in Figure 13 and can be constructed as follows:

i) the displacement of a mid-side node is normal to the side, alternatively inwardly and outwardly oriented, with magnitude proportional to the length of the side.

ii) the displacement of a corner is obtained by multiplication by -2 of the sum of the two displacements of the closest mid-side nodes.

iii) the displacement of the centroid is zero.

On a square, the pattern of $\underset{\sim}{W}$ is well-known:

$$\underset{\sim}{W} = \begin{pmatrix} -2, & 2, & 2, & -2, & 0, & -1, & 0, & 1, & 0 \\ 2, & 2, & -2, & -2, & -1, & 0, & 1, & 0, & 0 \end{pmatrix} \tag{7.26}$$

Contrary to 8-node elements, and because of the presence of H_x and H_y , this mode can "propagate" from one element to another. For example, on a square mesh, if the nodal displacement vector is $\underset{\sim}{W}$ given by (7.26) on an element Ω_0 , then the displacement vectors $\underset{\sim}{W} + 3\underset{\sim}{H}_x + \underset{\sim}{t}_x$ and $\underset{\sim}{W} - 3\underset{\sim}{H}_y - \underset{\sim}{t}_y$ on the elements to the right of Ω_0 and above Ω_0 allow us to construct a continuous global displacement also denoted $\underset{\sim}{W}$, on the mesh as shown in Figure 14.

We finally have

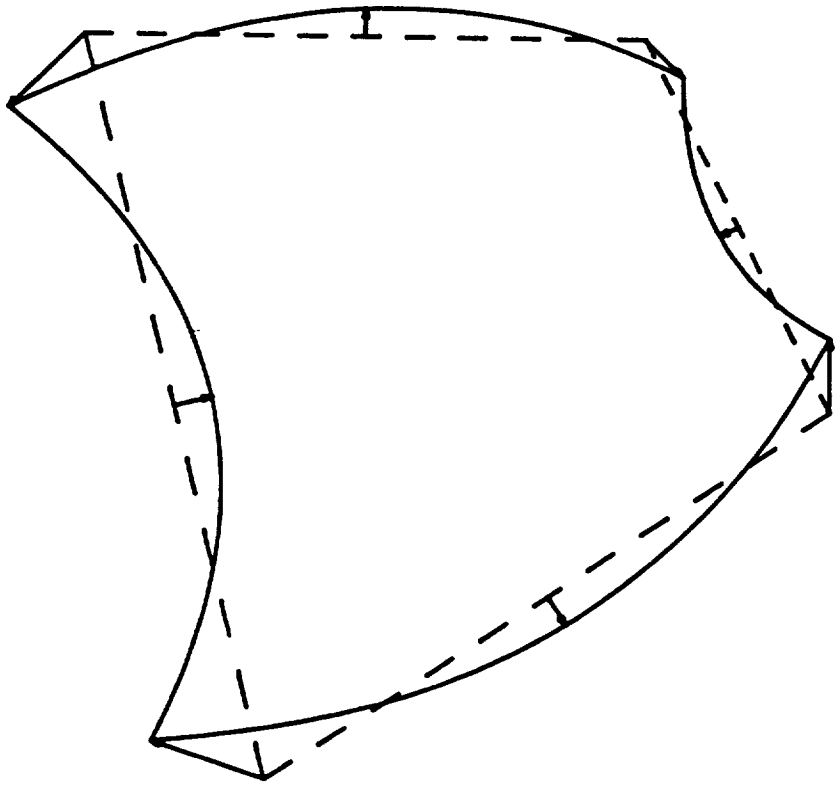


Figure 13. Spurious Mode $\underset{\sim}{W}$ for a General Quadrilateral Element.

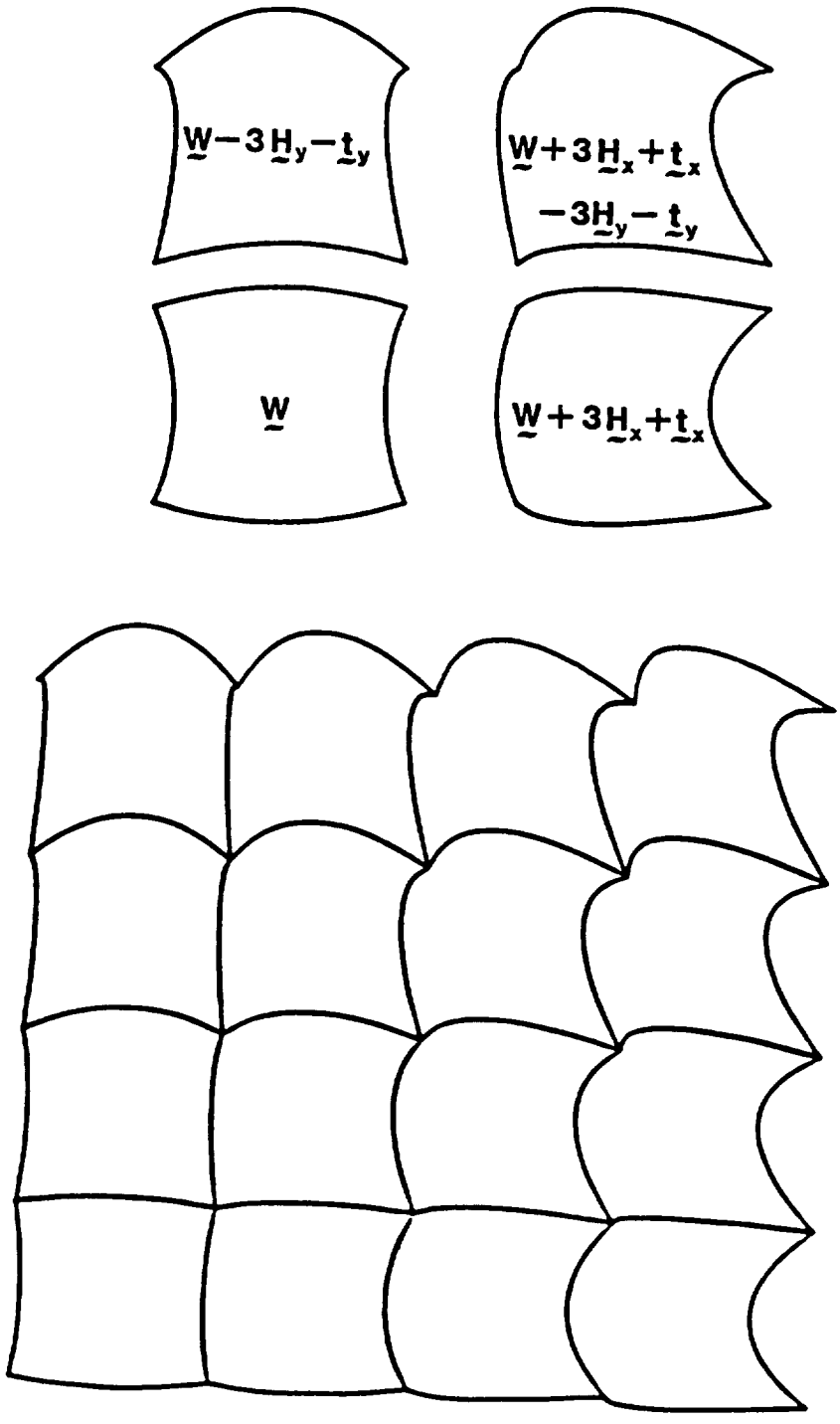


Figure 14. The Spurious Mode $\underset{\sim}{W}$: (a) Construction, (b) The Mode on a 16 Element Mesh.

$$\text{Ker } \underset{\sim}{K}^{\text{under}} = \text{span } \{\underset{\sim}{t}_x, \underset{\sim}{t}_y, \underset{\sim}{r}, \underset{\sim}{H}_x, \underset{\sim}{H}_y, \underset{\sim}{W}\} \tag{7.27}$$

Remark: Similar to what we have with 4-node elements, the existence of one axis of symmetry (say, the x-axis) reduces the kernel of the underintegrated stiffness matrix:

$$\text{Ker } \underset{\sim}{K}^{\text{under}} = \text{span } \{\underset{\sim}{t}_x, \underset{\sim}{H}_1, \underset{\sim}{H}_2 + \underset{\sim}{H}_3\} \tag{7.28}$$

where

$$\left.\begin{aligned} \underset{\sim}{H}_1 &= 3/2(\underset{\sim}{H}_x - \underset{\sim}{t}_x) \\ \underset{\sim}{H}_2 &= -3/2(\underset{\sim}{H}_y - \underset{\sim}{t}_y) \\ \underset{\sim}{H}_3 &= \underset{\sim}{W} + 2(t_x - t_y) \end{aligned}\right\} \tag{7.29}$$

have been chosen such that the displacements of these modes are zero at the intersection of both axes for a square mesh. Contrary to the 4-node case, we still have a spurious mode when two axes of symmetry exist:

$$\text{Ker } \underset{\sim}{K}^{\text{under}} = \text{span } \{\underset{\sim}{H}_1 + \underset{\sim}{H}_2 + \underset{\sim}{H}_3\} \tag{7.30}$$

This mode is shown in Figure 15.

It is also important to point out that whereas the pattern of the spurious modes $\underset{\sim}{H}_x$ and $\underset{\sim}{H}_y$ are independent of both the geometry and the element, the mode $\underset{\sim}{W}$ depends upon both of them. Moreover, we can see by construction on a square mesh that the amplitude varies strongly when we consider successive elements. In fact, the pattern we may observe is a succession of pattern $\underset{\sim}{H}_x$ and $\underset{\sim}{H}_y$ with increasing amplitude.

2.7.2. Global A-Posteriori Control in Linear Elasticity. In this subsection we wish to generalize (3.30) with regard to the discrete operators, using various kernels discussed in the previous subsection.

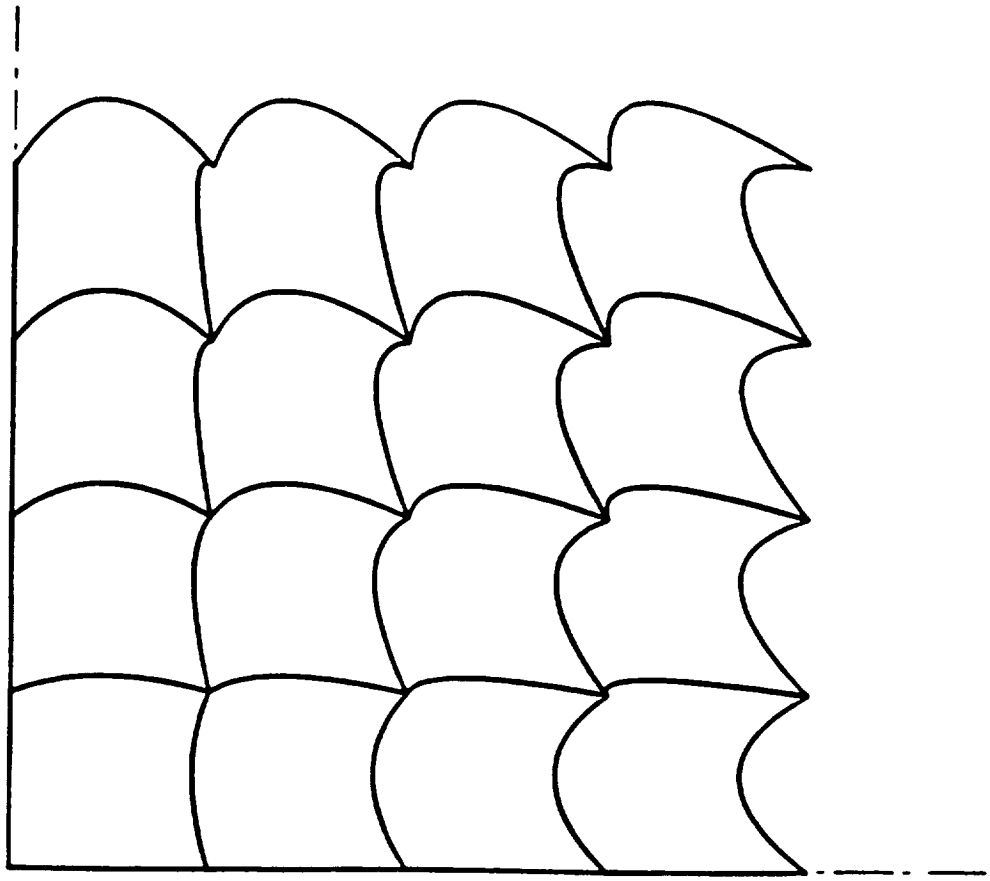


Figure 15. The Spurious Mode $\underset{\sim}{H}_1 + \underset{\sim}{H}_2 + \underset{\sim}{H}_3$.

We consider the general case where

$$\text{Ker } \underset{\sim}{K}^{\text{under}} = \text{RBM} \oplus \text{span}\{H_i,\ i = 1,I\} \tag{7.31}$$

where I may have the values 1, 2 or 3 . We recall that for $I = 1$, we obtained a control formula similar to

$$\tilde{u}^h = \bar{u}^h - \frac{a(\bar{u}^h,\ H_1)}{a(H_1,\ H_1)}\ H_1 \tag{7.32}$$

where the bilinear form $a(\cdot,\cdot)$ was obtained in the variational formulation of the initial problem. This projection satisfies:

$$a(\tilde{u}^h,\ H_1) = 0 \tag{7.33}$$

or, in other words, $\tilde{u}^h$ is orthogonal to the spurious mode. We generalize this property to the elasticity problem by supposing the projection to be orthogonal to all the spurious modes. Therefore the control will consist of looking for I constants λ_i $(i = 1,I)$ such that

$$\begin{cases} \underset{\sim}{\tilde{u}}^h = \underset{\sim}{\bar{u}}^h - \underset{i=1,I}{\Sigma} \lambda_i\ H_i \\ a(\underset{\sim}{\tilde{u}}^h,\ \underset{\sim}{H}_i) = 0 \quad \text{for } i=1,\ I \end{cases} \tag{7.34}$$

This leads to the system of I equations with I unknowns :

Find λ_i , $i = 1,I$ such that

$$\underset{j=1,I}{\Sigma} \lambda_j\ a(\underset{\sim}{H}_i,\ \underset{\sim}{H}_j) = a(\underset{\sim}{\bar{u}}^h,\ H_i)\ ,\ i=1,I \tag{7.35}$$

The computations involved in the control are computations of products of $\bar{u}^h$ and the spurious modes by themselves. The implementation of these computations is discussed in the next section.

2.7.2.a Implementation of the Spurious Modes Control. For the computation of the coefficients in (7.35), we again use the decomposition

$$\underset{\sim}{K}^{full} = \underset{\sim}{K}^{under} + \underset{\sim}{K}' \tag{7.36}$$

where $\underset{\sim}{K}^{under}$ satisfies

$$\underset{\sim}{K}^{under} \cdot \underset{\sim}{H}_i = 0 \tag{7.37}$$

Then

$$a(\underset{\sim}{\bar{u}}^h, \underset{\sim}{H}_i) = \sum_{e=1,E} \underset{\sim}{\bar{U}}^T \cdot \underset{\sim}{K}' \cdot \underset{\sim}{H}_i \tag{7.39}$$

The expressions used for $\underset{\sim}{K}'$ are next given for 4- or 9-node elements.

2.7.2.b. Control for 4-node Elements. For the operator defined in (7.1) and (7.2) with

$$\underset{\sim}{C} = \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{pmatrix} \tag{7.40}$$

we have the *exact* decomposition for *any* geometry of Ω_e :

$$\underset{\sim e}{K}^{exact} = \underset{\sim}{K}^{under} + \begin{pmatrix} \alpha_{11} \underset{\sim}{\gamma} \cdot \underset{\sim}{\gamma}^T & \alpha_{12} \underset{\sim}{\gamma} \cdot \underset{\sim}{\gamma}^T \\ \alpha_{21} \underset{\sim}{\gamma} \cdot \underset{\sim}{\gamma}^T & \alpha_{22} \underset{\sim}{\gamma} \cdot \underset{\sim}{\gamma}^T \end{pmatrix} \tag{7.41}$$

where

$$\left[\underset{\sim}{\alpha}\right] = \begin{pmatrix} C_{11}\bar{\varepsilon}_{xx} + (C_{13}+C_{31})\bar{\varepsilon}_{xy} + C_{33}\bar{\varepsilon}_{yy} ; C_{13}\bar{\varepsilon}_{xx} + (C_{12}+C_{33})\bar{\varepsilon}_{xy} + C_{32}\bar{\varepsilon}_{yy} \\ C_{31}\bar{\varepsilon}_{xx} + (C_{21}+C_{33})\bar{\varepsilon}_{xy} + C_{23}\bar{\varepsilon}_{yy} ; C_{33}\bar{\varepsilon}_{xx} + (C_{32}+C_{23})\bar{\varepsilon}_{xy} + C_{22}\bar{\varepsilon}_{yy} \end{pmatrix} \tag{7.42}$$

The vector $\underset{\sim}{\gamma}$ and the $\bar{\varepsilon}$'s are defined in Section 2 ((2.41) and (2.50)). For practival use, the expressions (2.51) are used for $\bar{\varepsilon}$. For linear isotropic linear material, $\underset{\approx}{C}$ is given by (7.3) and

$$\left[\underset{\sim}{\alpha}\right] = \begin{pmatrix} (\lambda+2\mu)\bar{\varepsilon}_{xx} + \mu\,\bar{\varepsilon}_{yy} & \mu\,\bar{\varepsilon}_{xy} \\ \mu\,\bar{\varepsilon}_{xy} & \mu\,\bar{\varepsilon}_{xx} + (\lambda+2\mu)\bar{\varepsilon}_{yy} \end{pmatrix} \tag{7.43}$$

This expression of $\left[\underset{\approx}{\alpha}\right]$ can be compared to the general strain-stress relationship:

$$\left[\underset{\sim}{\sigma}\right] = \begin{pmatrix} (\lambda+2\mu)\varepsilon_x + \mu\varepsilon_y & \mu\,\varepsilon_{xy} \\ \mu\,\varepsilon_{xy} & \mu\varepsilon_x + (\lambda+2\mu)\varepsilon_y \end{pmatrix} \tag{7.44}$$

An algorithm similar to the one presented in Section 2.5.1 can be constructed. It involves the computation of $\underset{\sim}{\gamma}$, ε and $\underset{\sim}{\alpha}$, then the computation of $a(H_i, H_j)$ and $a(\underset{\sim}{u}^h \cdot H_i)$, and finally the coefficients λ_i are obtained by resolution of a NxN system, N measuring the rank deficiency of K^{under} (N=1 or 2).

2.7.2.c. Control for 9-node Elements. In this subsection, devoted to 9-node elements, we first show why the results obtained by Belytschko are not sufficient to obtain a generalization of the linear elasticity problem, and then we propose an implementation of the control that leads to a stable solution converging to the exact solution with the optimal rate of convergence. However, for 9-node elements, we have not yet been able to obtain a computationally easy way to exhibit the third spurious mode, and the proposed results are only applicable to regular discretizations of a domain.

As far as the stabilization method proposed in [4] is concerned, algebra similar to that in Subsection 2.7.3.a leads to (7.41) where $\underset{\sim}{\gamma}$ was defined in 2.41. But, whereas the stabilization matrix constructed with the submatrix $\underset{\sim}{\gamma}\cdot\underset{\sim}{\gamma}^T$ eliminates $\underset{\sim}{H}_x$ and $\underset{\sim}{H}_y$ from the kernel of the stiffness matrix, it does not take $\underset{\sim}{W}$ into account. Indeed, we have

$$\begin{pmatrix} \alpha_{11}\,\underset{\sim}{\gamma}\cdot\underset{\sim}{\gamma}^T & \alpha_{21}\,\underset{\sim}{\gamma}\cdot\underset{\sim}{\gamma}^T \\ \alpha_{12}\,\underset{\sim}{\gamma}\cdot\underset{\sim}{\gamma}^T & \alpha_{11}\,\underset{\sim}{\gamma}\cdot\underset{\sim}{\gamma}^T \end{pmatrix} \cdot \underset{\sim}{W} = \underset{\sim}{0}$$

Therefore this procedure cannot be used to control $\underset{\sim}{W}$.

In order to obtain an accurate control, we have to consider a generalization of (2.43). Now we have

$$\begin{pmatrix} \alpha_{11}\,\underset{\sim}{s}_i\cdot\underset{\sim}{s}_i & \alpha_{21}\,\underset{\sim}{s}_i\,\underset{\sim}{s}_i \\ \alpha_{12}\,\underset{\sim}{s}_i\cdot\underset{\sim}{s}_i & \alpha_{22}\,\underset{\sim}{s}_i\,\underset{\sim}{s}_i \end{pmatrix} \cdot \underset{\sim}{H}_j \neq \underset{\sim}{0}$$

$$\text{for } \begin{matrix} 7\leq i\leq 9 \\ 1\leq j\leq 3 \end{matrix}$$

where the vectors $\underset{\sim}{s}_i$ are defined in Section 2.2.3. Finally, using (2.43) and (2.52), we have

$$\underset{\sim}{K}_e^{(9)} = \underset{\sim}{K}_e^{(4)} + \frac{4\Omega_e}{135}\begin{bmatrix} (C_{11}+C_{33})\underset{\sim}{s}_9\underset{\sim}{s}_9^T & (C_{13}+C_{32})\underset{\sim}{s}_9\underset{\sim}{s}_9^T \\ (C_{31}+C_{23})\underset{\sim}{s}_9\underset{\sim}{s}_9^T & (C_{22}+C_{33})\underset{\sim}{s}_9\underset{\sim}{s}_9^T \end{bmatrix}$$

$$+\frac{\Omega_e}{45}\left(\begin{bmatrix} C_{11}\,\underset{\sim}{s}_7\underset{\sim}{s}_7^T & C_{13}\underset{\sim}{s}_7\underset{\sim}{s}_7^T \\ C_{31}\underset{\sim}{s}_7\underset{\sim}{s}_7^T & C_{33}\underset{\sim}{s}_7\underset{\sim}{s}_7^T \end{bmatrix} + \begin{bmatrix} C_{33}\underset{\sim}{s}_8\underset{\sim}{s}_8^T & C_{32}\underset{\sim}{s}_8\underset{\sim}{s}_8^T \\ C_{23}\underset{\sim}{s}_8\underset{\sim}{s}_8^T & C_{22}\underset{\sim}{s}_8\underset{\sim}{s}_8^T \end{bmatrix}\right) \qquad (7.45)$$

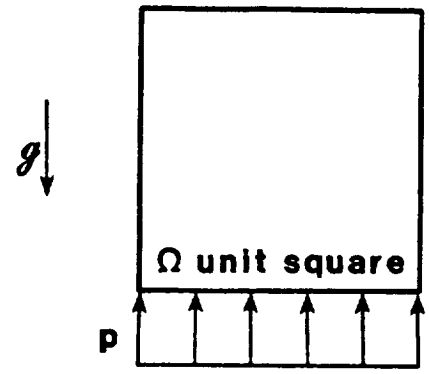
Similarly, for the 4-node case, the algorithm for the computations

of the coefficients in (7.35) has been obtained and implemented. Numerical results agree with our presumptions concerning a "$\underset{\sim}{\gamma}\cdot\underset{\sim}{\gamma}^T$"-type of control and incline in favor of the decomposition (7.45). On a square domain discretized with NxN elements, we have calculated and compared the solutions obtained with full (exact) and underintegration for various boundary and symmetry conditions. Various examples considered are described in Figure 16. The rates of convergence were calculated by comparing the error norms (δ=0: L^2/RBM norm; δ=1: energy norm) obtained with N=5,6 and 7. We consistently obtained the rate $O(h^{2-\delta})$ using a $\underset{\sim}{\gamma}\cdot\underset{\sim}{\gamma}^T$ decomposition and $O(h^{3-\delta})$ with (7.45) for homogeneous materials under the action of gravity ($\underset{\sim}{f} \in C^\infty$) ; the order $3-\delta$ being optimal, we may conclude that the method presented below is accurate. It is also efficient: for one second taken for the fully integrated stiffness matrix, only .61 are taken when the underintegration is used and only .05 seconds are taken for the control Also note that the Example 4 in Figure 30 is also optimal ($u \in H^1/H^2$) .

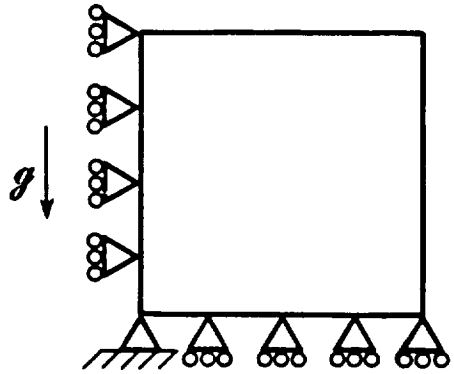
Unfortunately, this control is far from general. In particular, it can only be used when the exact shape of the spurious modes is known, which is the case only when Neumann (traction) boundary conditions are applied on a domain discretized with regular square elements.

2.7.3. A Local Control of the Spurious Modes

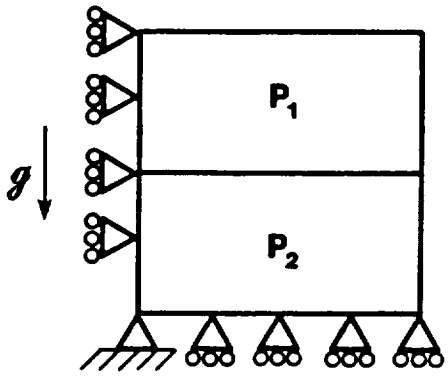
2.7.3.a. Introduction. The major drawbacks of the global control are overcome by considering the procedure consisting of eliminating, element by element, the components of $\underset{\sim}{H}_x$, $\underset{\sim}{H}_y$ and $\underset{\sim}{W}$ and then of averaging the nodal values obtained in neighboring elements. For any element, we choose to do the following simplifications:



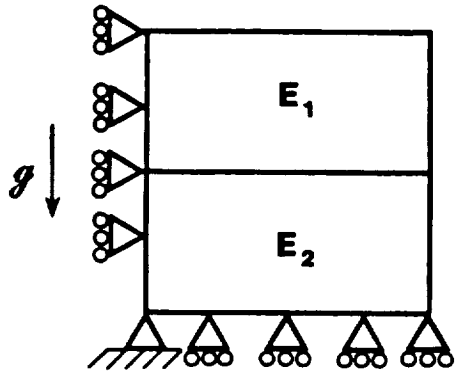
Example 1



Example 2



Example 3



Example 4

Example	Number of Spurious Modes	Rate of Convergence	
		$L^2/\mathbb{R}$	H^1
1	3	3	2
2	1	3	2
3	2	3	2
4	1	2	1

Figure 16. Results Obtained with the Global Control.

i) the nodal values of $\underset{\sim}{W}$ are taken as if the element was strait-sided quadrilateral

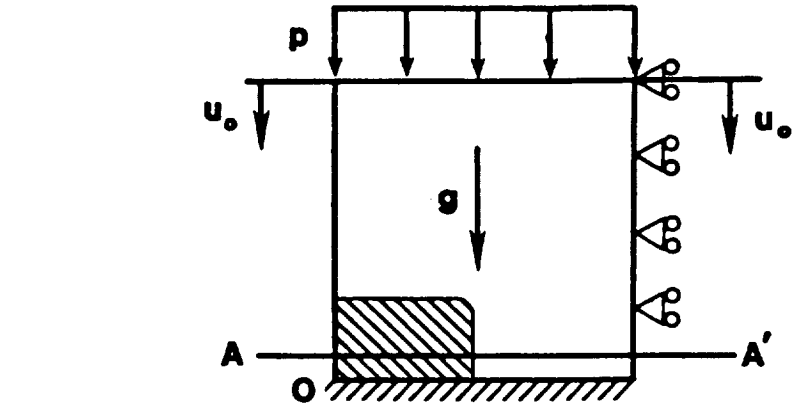
ii) $\underset{\sim}{C}$ is diagonal.

iii) $\underset{\sim}{K}^{stab}$ is given by (7.41) or (7.45).

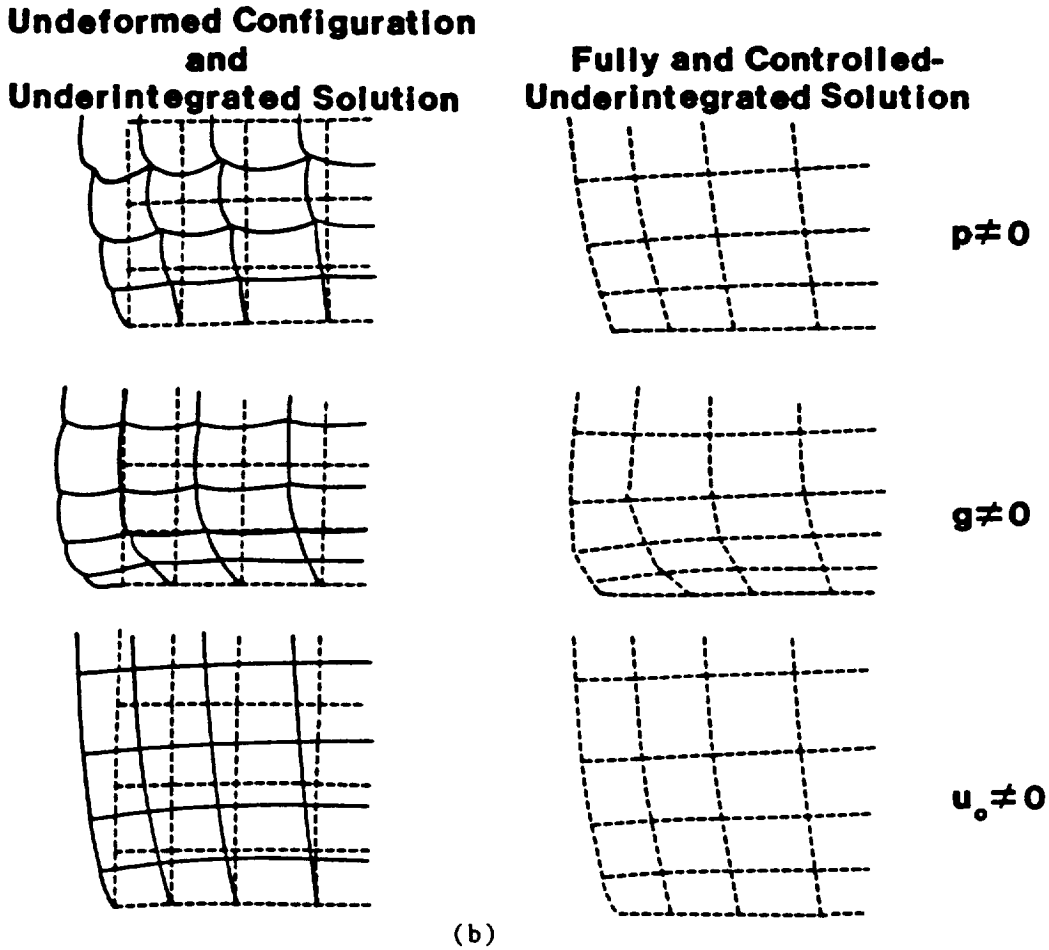
These simplifications lead to simple calculations detailed in Appendix B for the 9-node element and that are easily implemented in the subroutine listed in Appendix C. Note that the expression obtained for the control is uniquely geometric, does not depend upon the material properties and can be used in any linear or nonlinear problem. It only requires that the shape of the element not deviate too much from a quadrilateral.

2.7.3.b. Numerical Results. The examples displayed in this section illustrate the efficiency and the accuracy of the local control previously described. However, we only consider linear elastic material in plane strain, on domains discretized with biquadratic (9-node) elements. Three examples are described.

The first example is defined in Figure 17.a. We consider a square domain with one fixed side, under the action of pressure, gravity or a prescribed compressive displacement. Under any of these loads, a singularity appears at the neighborhood of the origin. Figure 17.b shows how the underintegrated solution behaves in the singularity region and how the control affects the results. Whereas the underintegrated solution shows oscillations, the displacements obtained after control are smooth and similar to those obtained with the full integration of the stiffness matrix. The shear along a line AA' across the singularity region is also shown in Figure 18. Whereas undesirable oscillations



(a)



(b)

Figure 17. Displacement of a Fixed Square :
a) Definition of the Problem; b) Results.

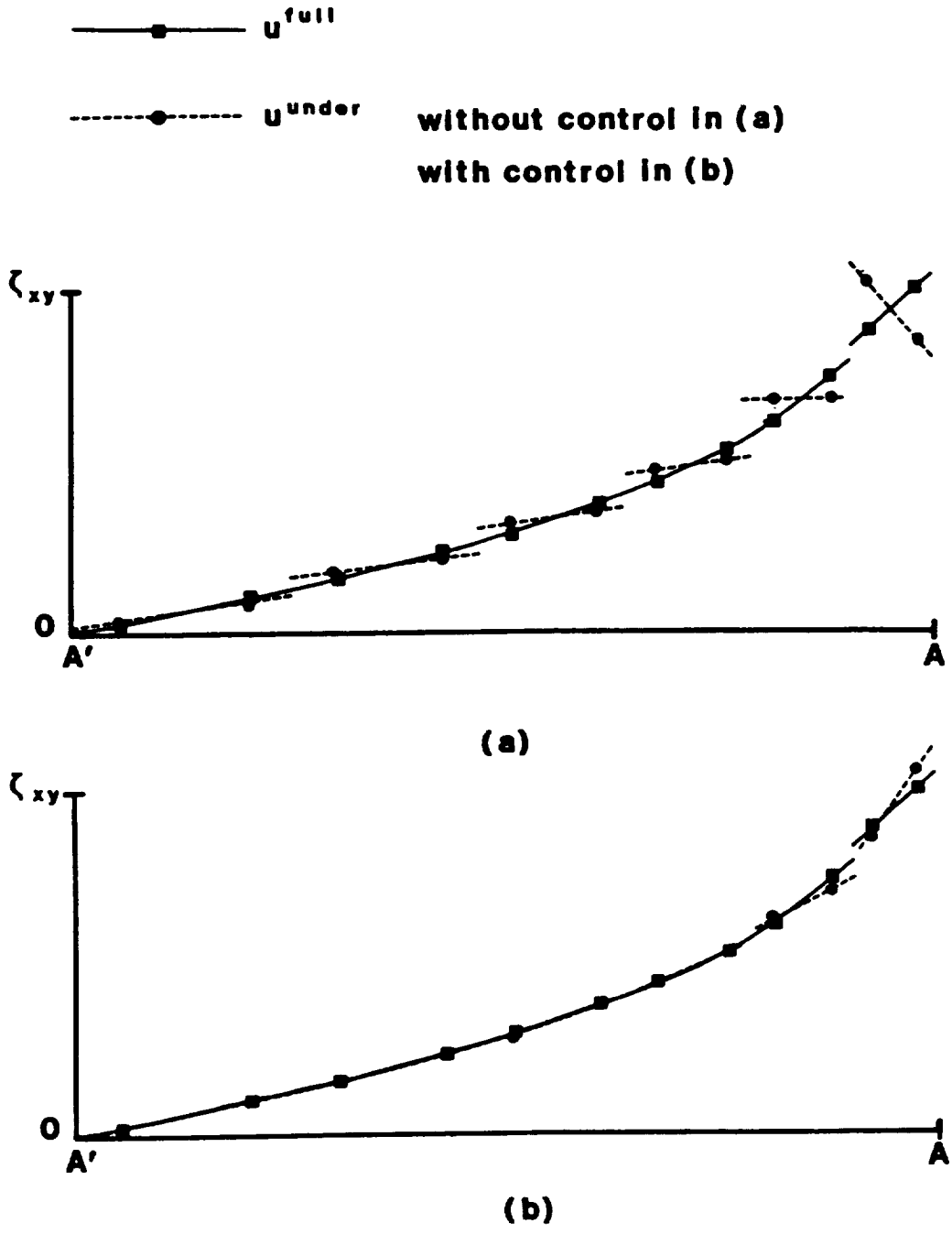


Figure 18. Shear along a Line AA'.

are observed before control, the shear behaves properly after control.

In the second example, we consider a ring under the action of an external pressure (Figure 19). Here again the oscillations generated by the underintegration are damped when the control is applied. Only very slight oscillations remain, not exceeding 5%, and these can be easily interpreted: in the control, the expression taken for the mode $\underset{\sim}{W}$ was obtained for quadrilateral elements. For the mesh considered, the elements are slightly bent and this difference explains these slight oscillations. The same domain (quarter ring) has also been discretized using quadrilateral elements and the control of the underintegrated solution has led to a displacement field without any oscillations and similar to the underintegrated displacements. Calculations of the stress along a radius show behavior identical to the previous example.

The third example involves a concentrated force and illustrates our discussion concerning the excitation of spurious modes and the orthogonalization of the data. A point force is applied at a corner of a fixed side square discretized with a mesh refining in the neighborhood of the singularity. Strong oscillations appear in this region when underintegration is used, whereas the full integration solution is smooth (Figure 20). These oscillations show a pattern similar to the one used to construct the mode $\underset{\sim}{W}$ (Figure 14): amplification of the mode $3\underset{\sim}{H}_x + \underset{\sim}{t}_x$ (respectively $3\underset{\sim}{H}_y + \underset{\sim}{t}_y$) along the x- (resp. y-) direction. Therefore, according to the previous interpretations of (6.35) and (6.37), a way to prevent oscillations is to consider a system of loads similar to the load concentrated in a point A but ortho-

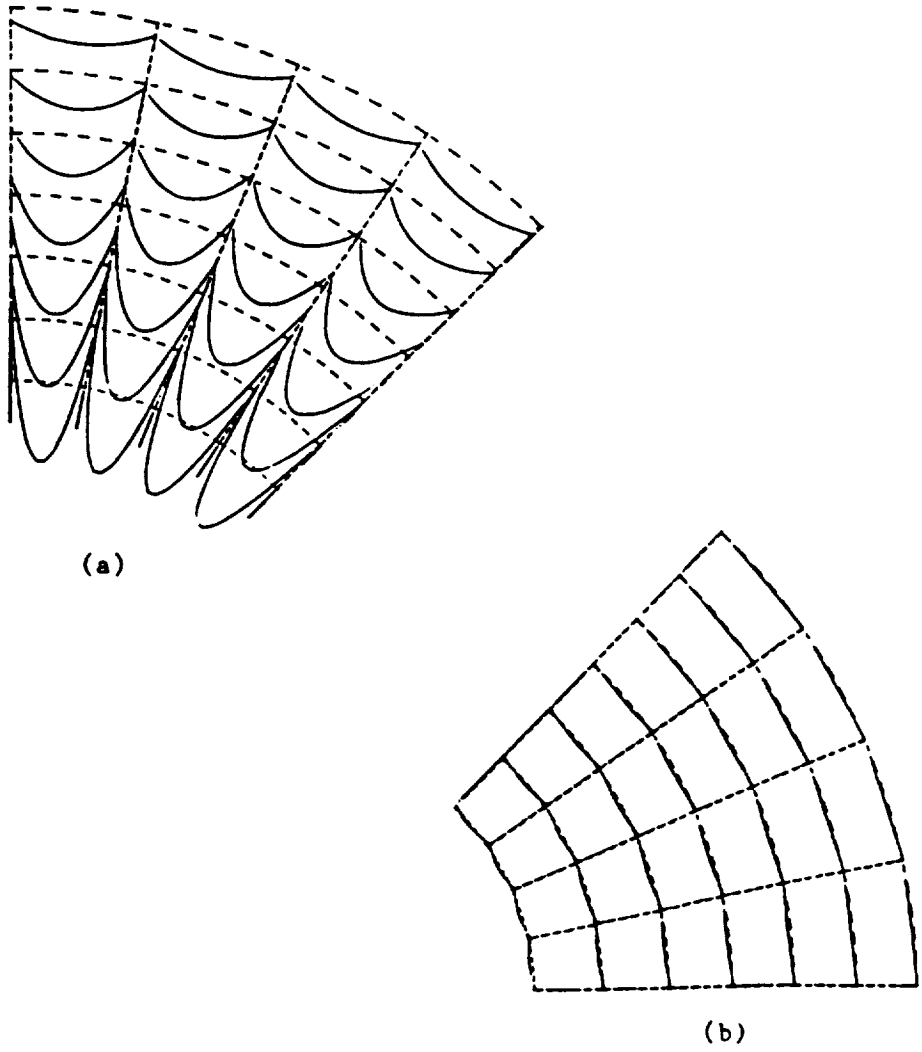
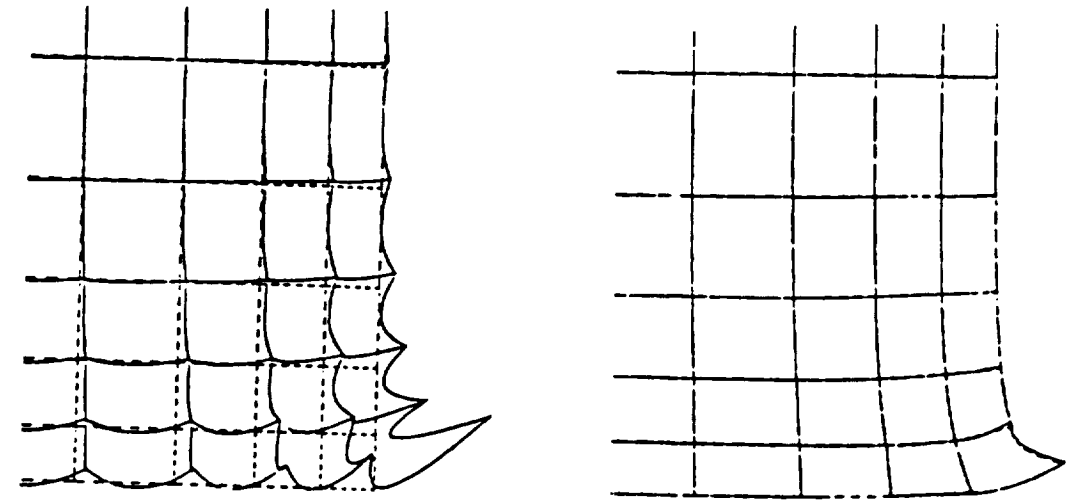


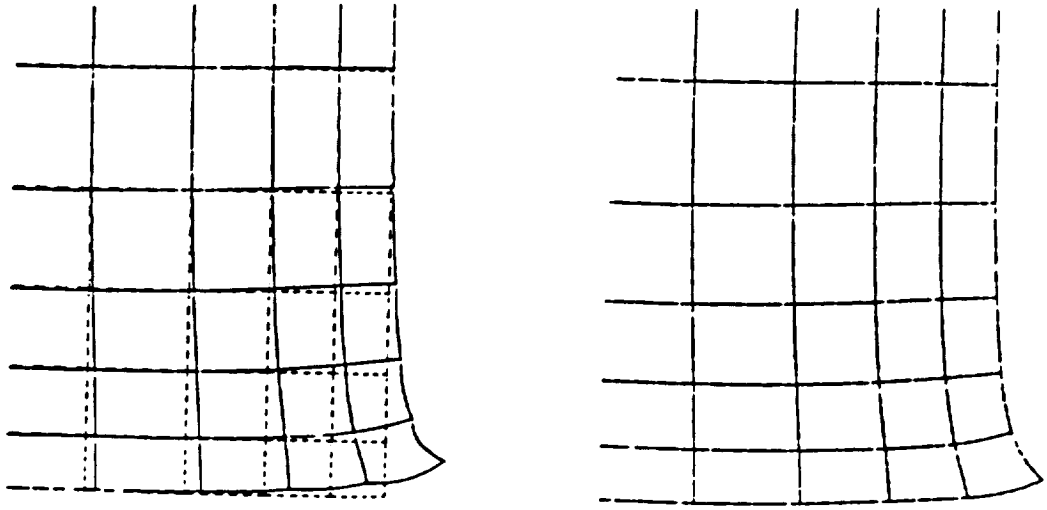
Figure 19. Quarter Ring Under External Pressure

(a) Undeformed Configuration and Underintegrated Solution, and

(b) Fully- and Controlled Underintegrated Solution.



(a)



(b)

Figure 20. Displacements Due to a Concentrated Load: Undeformed, Underintegrated, Controlled and Fully Integrated Solution Due to a) the Force, and b) its Orthogonalized Equivalent.

gonal to $3\underset{\sim}{H} + \underset{\sim}{t}$. This is obtained by splitting the force into 3 equal forces applied at A and its two closest nodes. Indeed, the displacements obtained with this system of load only show slight oscillations. Finally, note that this control produces displacement fields similar to the fields obtained using full integration.

PART III: APPLICATION TO NONLINEAR INCOMPRESSIBLE ELASTICITY

3.1 Introduction

In this concluding part, we analyze instabilities observed in discrete solutions of nonlinear problems in finite elasticity involving incompressible materials. We compare the behavior of the biquadratic (9-node) and isoparametric (8-node) elements associated with linear (P_1) discontinuous pressures. Then we focus on 9-node elements to discuss the efficiency and accuracy of control of the underintegrated solution introduced in Part II for linear operators.

Only Mooney-Rivlin materials are considered. They are characterized by the strain energy function

$$\sigma = C_1(I_1 - 3) + C_2(I_2 - 3) \tag{3.1}$$

where I_i , i=1,3 are the principal invariants of the Cauchy-Green deformation tensor

$$\underset{\sim}{C} = (\underset{\sim}{1} + \underset{\sim\sim}{\nabla u})^T(\underset{\sim}{1} + \underset{\sim\sim}{\nabla u}) \tag{3.2}$$

where $\underset{\sim}{1}$ is the unit tensor and $\underset{\sim\sim}{\nabla u}$ is the displacement gradient. They are:

$$\begin{aligned} I_1 &= \operatorname{tr} \underset{\sim}{C} \\ I_2 &= \tfrac{1}{2}(\operatorname{tr} \underset{\sim}{C}^2 - (\operatorname{tr} \underset{\sim}{C})^2) \\ I_3 &= \det \underset{\sim}{C} \end{aligned} \tag{3.3}$$

The condition of incompressibility can be expressed as

$$I_3 = 1 \tag{3.4}$$

and is taken into account in a mixed formulation of the equilibrium problem by introducing a Lagrange multiplier P. The energy function σ can be replaced by σ^{Lag}:

$$\sigma^{Lag} = C_1(I_1 - 3) + C_2(I_2 - 3) + P(\sqrt{I_3} - 1) \tag{3.5}$$

For this choice of energy function, P is the hydro-static pressure. We consider the usual virtual work equilibrium equation:

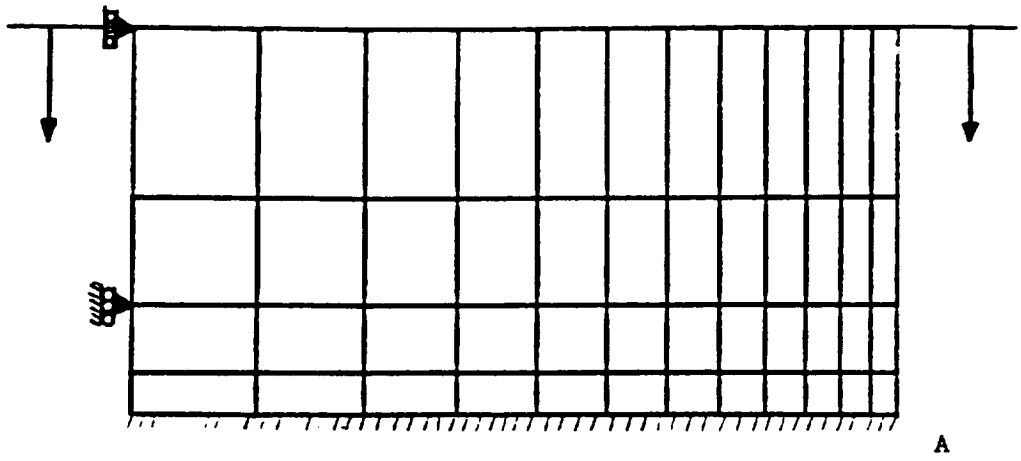
$$\delta \int_{\Omega_0} \sigma^{Lag} \, dv_0 - \int_{\Omega_0} \rho_0 \, \underset{\sim}{f} \cdot \delta \underset{\sim}{u} \, dv_0 - \int_{\partial\Omega_2} \underset{\sim}{t} \cdot \delta \underset{\sim}{u} \, ds = 0 \tag{3.6}$$

The solution of this highly nonlinear problem is accomplished in this work using Newton's method. Details of the finite element method applied to this particular class of problems are discussed at length in the book of Oden [38]. A more recent account is given in Aly [1].

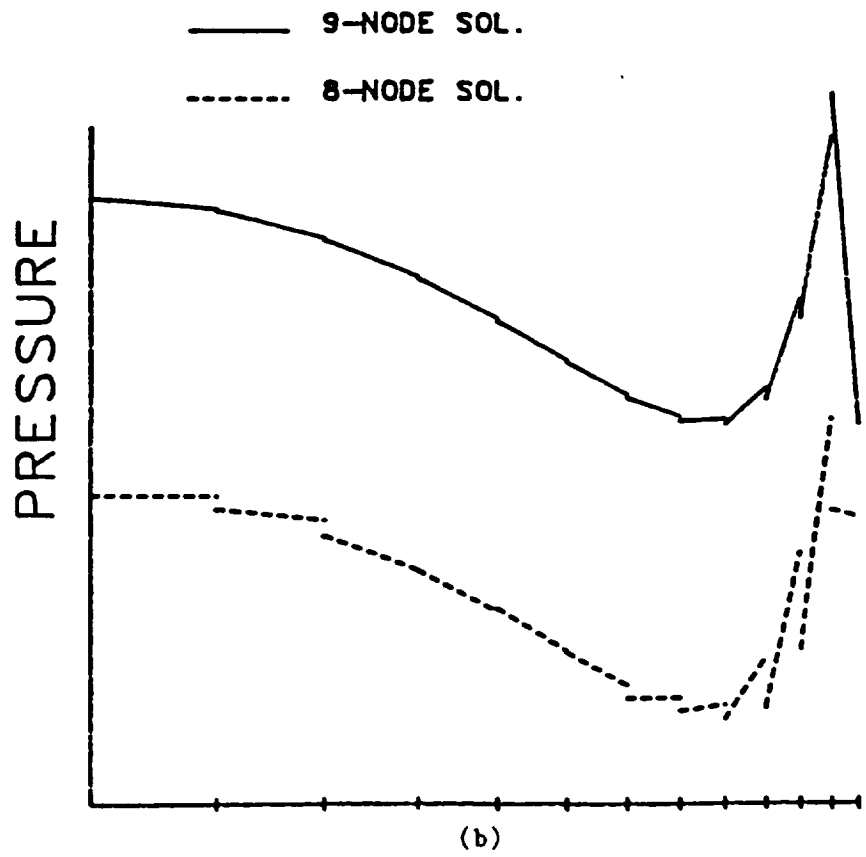
3.2 Behavior of 8- and 9-Node Elements (Full Ingegration)

In this section, we briefly review some observations made by Miller [37] that are now clearly understood with the results obtained by Oden and Jacquotte [42, 43].

We consider a fixed side rectangular domain discretized with the refining mesh shown in Figure 21.a. We compress this domain imposing a displacement on the top side. The displacement increments are 5, 10, 15 and 17.5% of compression with respect to the original shape.



(a)



(b)

Figure 21. Compression of a Square a) Mesh at Definition in Problem; b) Pressure Obtained with 8- and 9-Node Elements (u_0 = 5%).

The solution of this problem is complicated by a stress singularity that occurs in the neighborhood of the corner A . Both 8- and 9-node displacement elements were tested associated with a linear discontinuous Lagrange multiplier. The displacements show similar behaviors. As far as pressure is concerned, oscillations similar to the pattern of ker B_h^* in [44] are observed (Figure 21.b) when 8-node elements are used, whereas the pressure distribution is smoother with 9-node elements. Finally, note that the nodal averaging technique described in Section 1.8 has been tested by Miller [37] for a problem with similar singularity and his conclusions corroborate ours from Part I.

3.3 Control of the 9-Node Underintegrated Element

In this section, we analyze how the displacements and pressures behave when underintegration is used. We noticed in Section 2.7 how the control obtained was only geometric. This is particularly useful for rubber-like materials and makes it very efficient when constitutive relations are numerically expensive to obtain. Also note that the calculation of the projected element solution (Appendix B) involves the knowled e of the mode $\underset{\sim}{W}$ which is computed using the *current* deformed geometry of the element. We may foresee what will be one of the major drawbacks of the method: $\underset{\sim}{W}$ is only exactly known for quadrilateral elements. When the element is too distorted, the approximation we do assuming it to be quadrilateral is too poor and the control is not accurate.

Finally, we point out that the control is applied at each load increment. We present three examples.

3.3.1 Stretch of a Rubber Material Domain. The first example involves the stretch of a rubber square (Fig. 22). The domain is partitioned in 49 elements (450 degrees of freedom). The material is a Mooney Rivlin material with

$$C_1 = 100$$
$$C_2 = 20 \qquad (3.7)$$

and the incremental stretches are 25, 50, 75 and 100% of its undeformed configuration. Whereas the underintegrated solution shows slight oscillations of the displacement in the singularity region, the controlled solution is smooth and similar to a fully integrated solution. As far as pressure is concerned, similar observations to the ones in Section 2.7.3.b for stresses can be made. However, the convergence is obtained slower: for the various displacement increment, 6, 4, 4 and 4 (respectively 6, 5, 5 and 4) Newton iterations have been needed to obtain convergence with the full (resp. under) integration. Nevertheless, the gain in time is almost 40% (673 sec. versus 413 sec.).

3.3.2 Behavior in the Neighborhood of a Concentrated Force. This second example illustrates the ability of the orthogonalization of the data to obtain an underintegrated solution. We consider the same Mooney Rivlin material (1.1, 1.7) and the problem defined in Figure 23.a. The application of the force at only one point leads to the behaviors:

i) Slight oscillations are observed in the singularity region when full integration is performed (Fig. 23.b)

ii) Uncontrollable oscillations appear when underintegration is

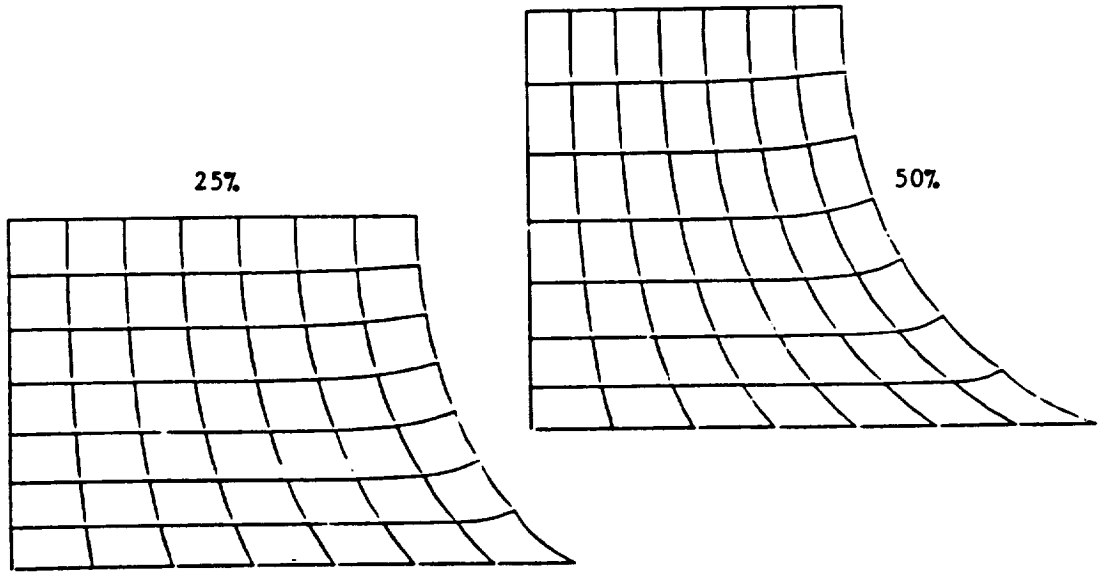


Figure 22.a. Underintegrated Solution.

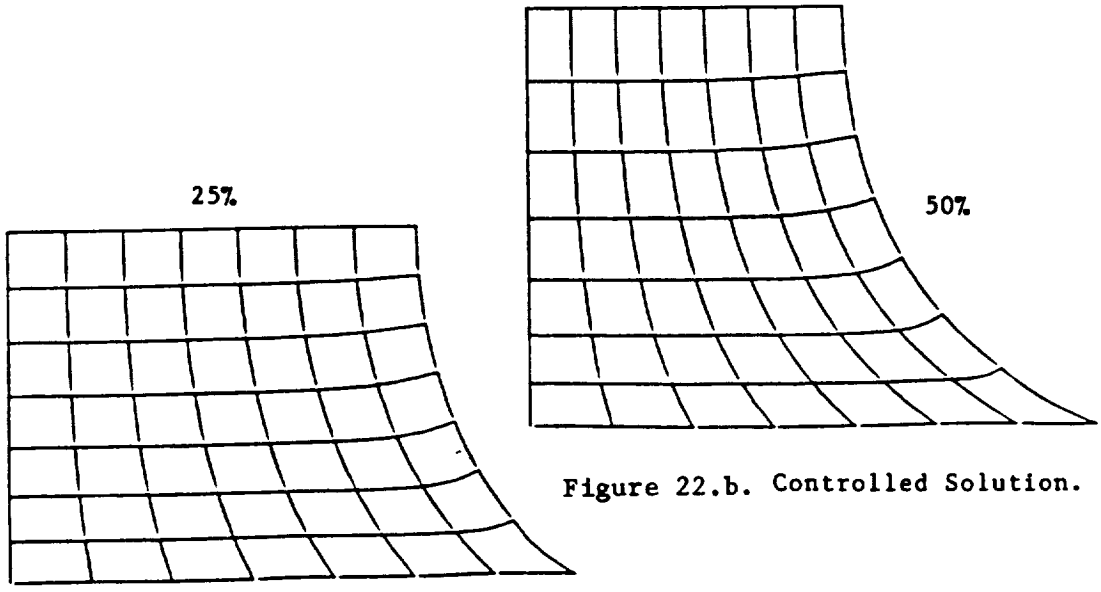
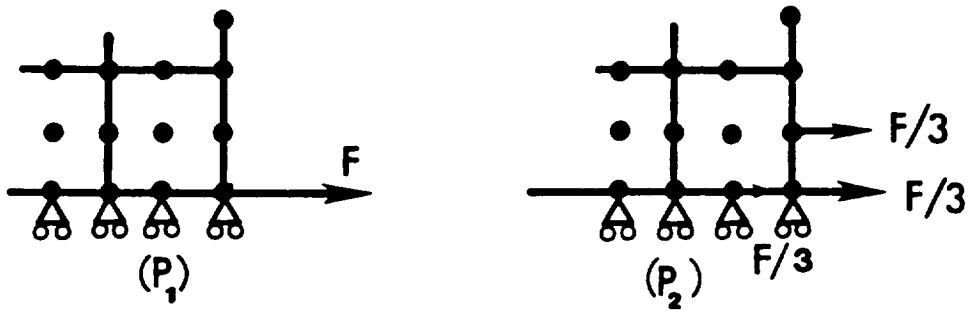
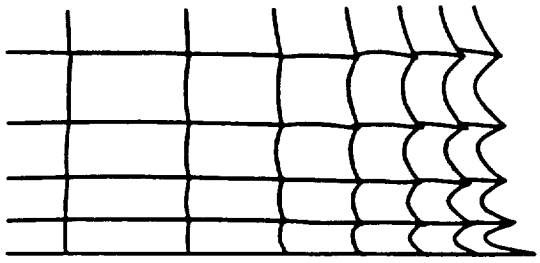


Figure 22.b. Controlled Solution.

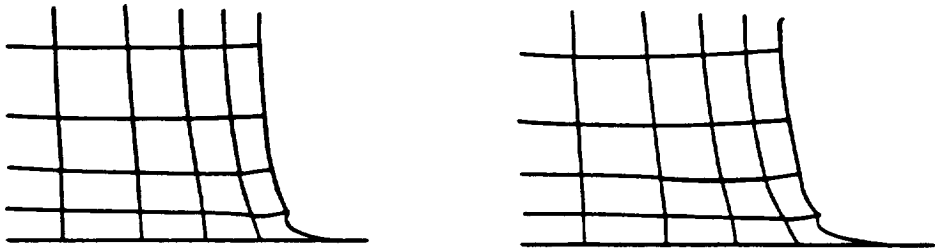
Figure 22. Stretch (25, 50%) of a Rubber Material Domain.



(a)



(b)



(c)

Figure 23. Behavior in the Neighbrohood of a Concentrated Force: a) Definition of the Problems; b) Under-integration of P_1; c) Controlled Solutions at two Increments.

used: Figure 23.c shows the displacements after only 3 iterations for the first load increment. Later the solution diverges.

As in the linear case discussed in Section 2.7.3.b, we split the concentrated force into three equal forces at the closest nodes; we observed that (Fig. 24)

iii) When underintegration is performed, without control, the solution converges, but oscillations still develop (Fig. 24.a).

iv) The control of this solution is smooth and similar to the one obtained with full integration (Figs. 24.b and c).

For this example, one supplemental iteration was needed in the third load increment and the gain in time also approaches 40%.

3.3.3 Compression of a Fixed Rubber Material Domain. This final example reconsiders the problem used to compare the performances of the 8- and 9-node elements (Section 3.2). We consider two discretizations of the domain with 25 and 49 elements, and displacement increments of 5, 10, 15 and 17.5% with respect to the original shape. For the crude mesh, oscillations appear very soon when underintegration is performed, but the control easily corrects the solution and a displacement field close to the fully integrated field is obtained (Fig. 25.a). But when the mesh is refined, the oscillations become more important and deform the element to a degree such that the control is not able to restore the shape of the element corner (Fig. 26.b). We interpret this lack of performance to the fact that the control has been exactly obtained for quadrilateral elements . In this case, the element sides curve and the element is too deformed. Also this lack

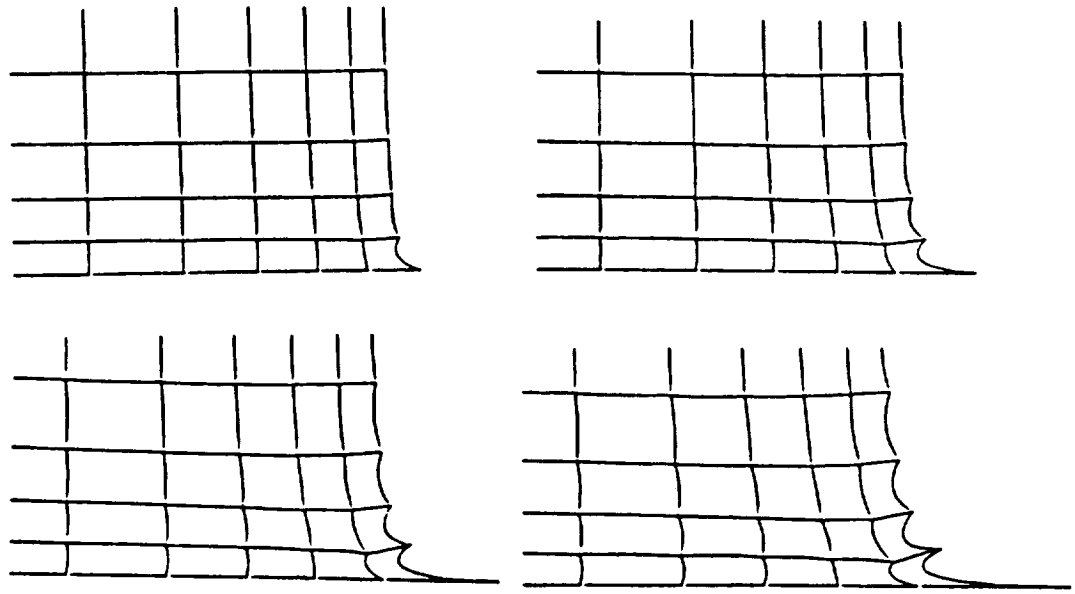


Figure 24.a. Underintegration of P_2 .

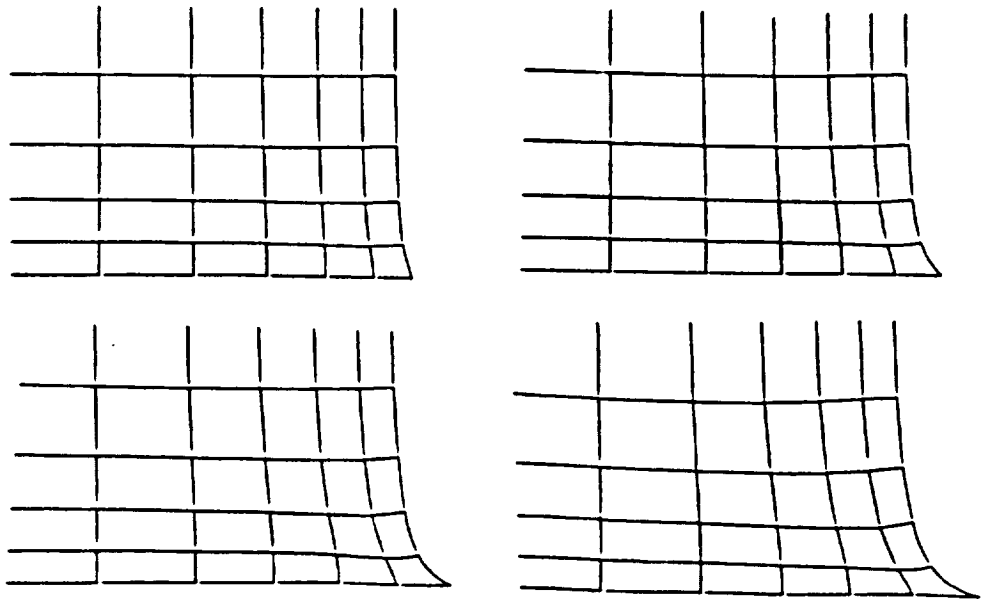


Figure 24.b. Controlled Underintegrated Solution of P_2 .

Figure 24. Behavior in the Neighborhood of a Concentrated Force (Problem P_2).

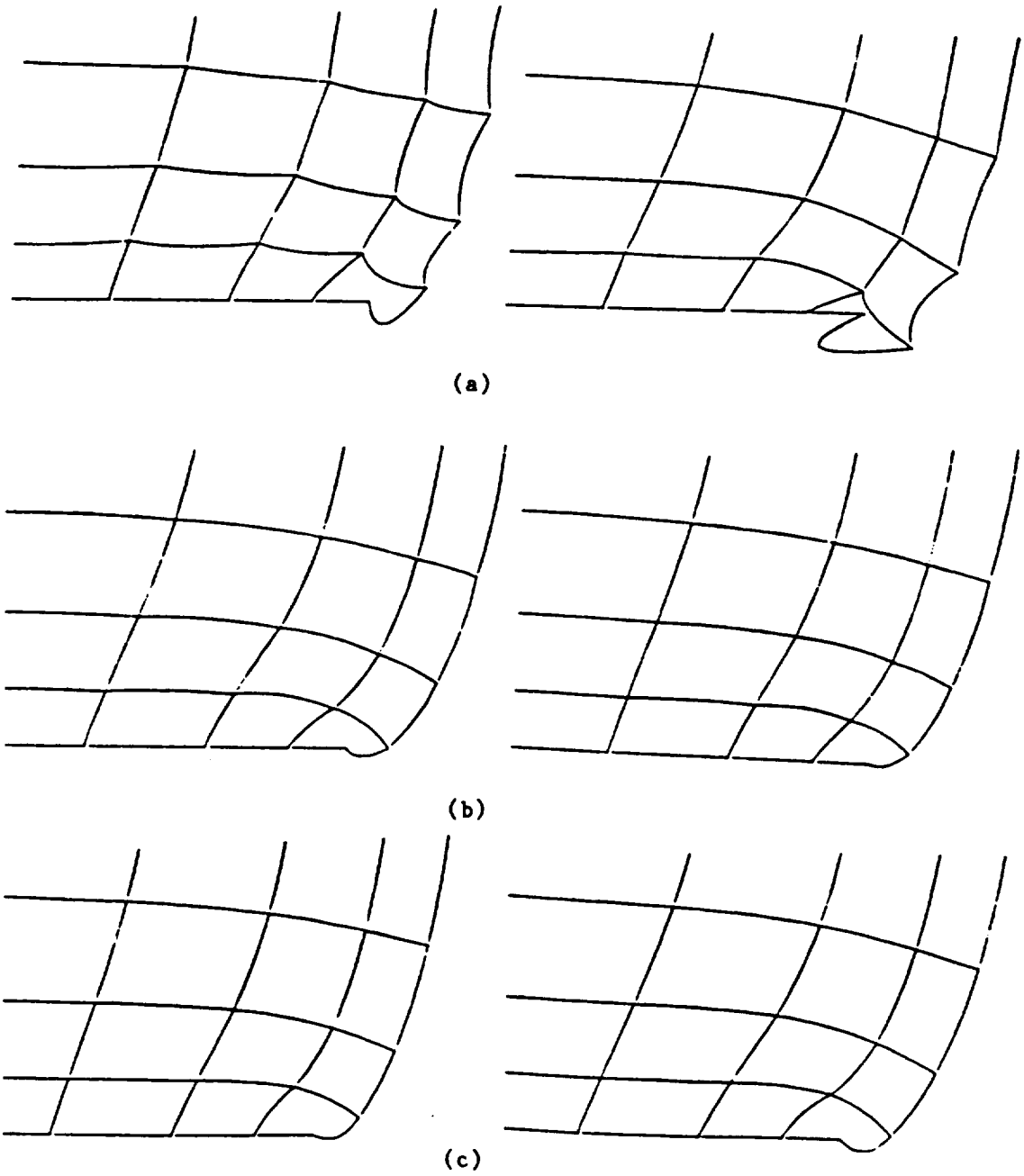


Figure 25. Compression (15%, 17.5%) of a Rubber Material Domain (25 Element Mesh): a) Underintegration, b) Controlled Underintegration, and c) Full Integration.

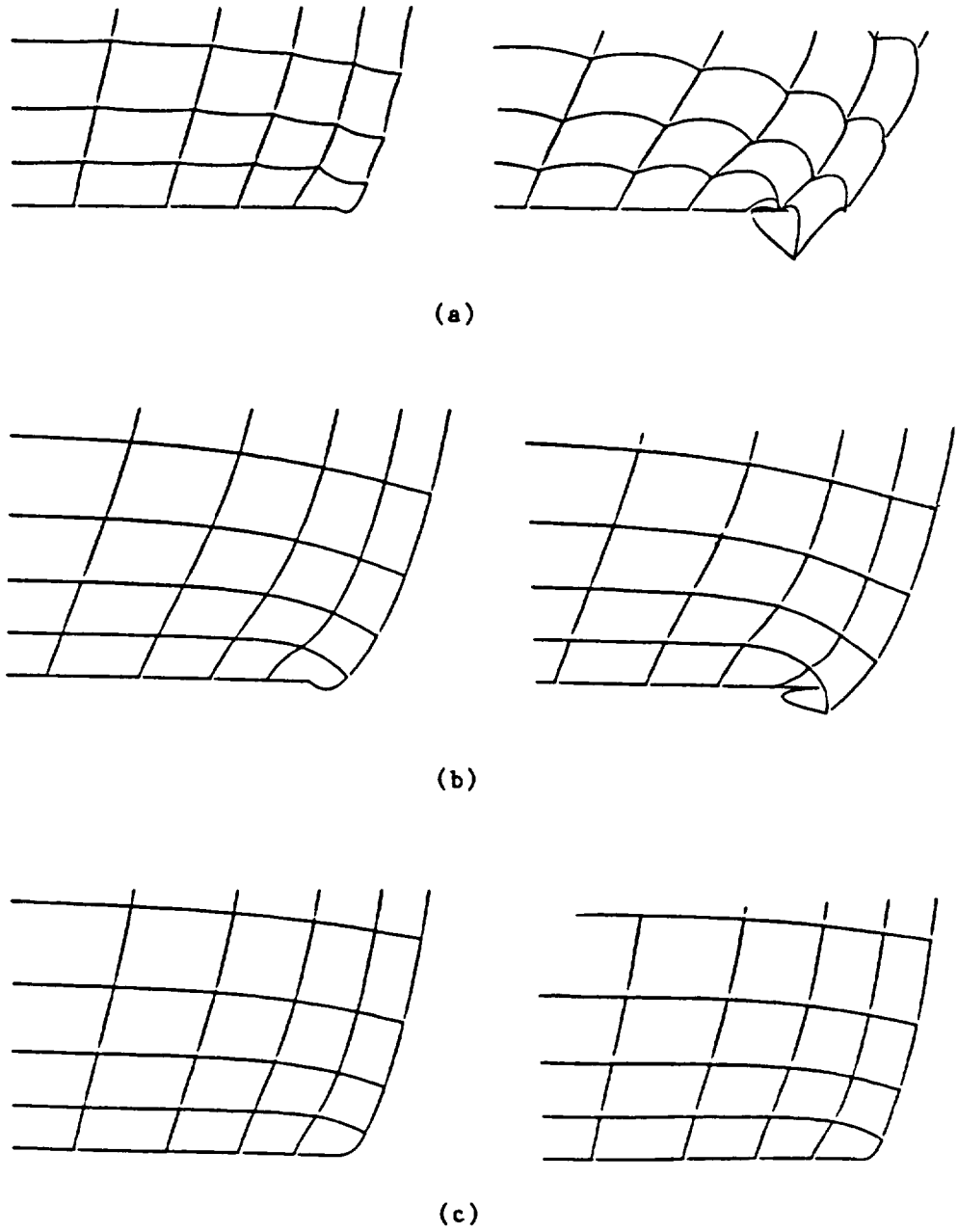


Figure 26. Compression (10, 15%) of a Rubber Material Domain (49 Element Mesh): a) Underintegration, b) Controlled Underintegration, and c) Full Integration.

of performance is observed when looking at the number of iterations required for the convergence of the Newton algorithm: they are 5, 5, 5 and 5 (respectively 5, 5, 6, 7) for the fully (resp. under-) integration in the 25 element mesh and 6, 6, and 6 (resp. 5, 6 and 13) in the 49 element mesh.

PART IV: CONCLUSIONS

In this work, we study instabilities appearing in the finite element resolution of linear and nonlinear, compressible and incompressible elasticity. The study is carried both mathematically and numerically. Some of the principal conclusions are listed as follows:

1) The use of underintegration in the stiffness matrix calculations results in rank-deficient stiffness matrix. These rank-deficiencies correspond to additional modes supplied to the rigid body modes that appropriately belong to the kernel of the operator.

2) There is a significant class of problems in which, with appropriate filtering, it can be shown that an underintegrated solution with hourglass control can yield very satisfactory answers, and produce a finite element method which has the same rate of convergence as the fully integrated method. The fact that this does indeed hold has been rigorously proved in this dissertation for a class of scalar elliptic boundary value problems, and numerically verified for a class of linear and nonlinear elasticity problems.

The method developed in this work seems to give satisfactory results in a broad class of problems. Several questions have, however, arisen:

- When an element is too distorted, the control cannot restore a reasonable shape. The accuracy of the control relies on the

approximation made on the mode $\underset{\sim}{W}$, supposing that the element remains quadrilateral. Does an exact computation of $\underset{\sim}{W}$ (7.21) for very distorted elements give a better answer?

- The lack of stiffness seems to slow the speed of convergence of the Newton's scheme. Can the method be modified in order to increase the speed of convergence?

- Finally, do the results generalize to three dimensional elasticity?

The answer to these questions may provide a tremendous gain in computation time.

APPENDIX A

As far as mixed boundary conditions are concerned, we suppose that a Dirichlet boundary condition is applied at 0 and a Neumann boundary condition at 1 . For the interval $[0,1]$, we consider the NxN matrix

$$D(k) = \begin{pmatrix} 2k & -1 & 0 \\ -1 & & \\ & 2k & -1 \\ 0 & -1 & k \end{pmatrix}$$

The values for which det K(k) vanishes are:

$$k_i = \cos\left(\frac{-\pi}{2N} + \frac{i\pi}{N}\right), \qquad 1 \leq i \leq N$$

and the corresponding vectors (D(ki)vi = 0) are :

$$v_i = \{\sin \frac{ij\pi}{2N}\} \quad 1 \leq j \leq N$$

The corresponding approximation space $V^h_{1,\frac{1}{2}}$ with basis $\{\phi^j\}$ is constructed as in [25]or in Section 2.4. Then, depending upon the sides where the various boundary conditions (D or N) are applied, tensor product of V^h_1 , $V^h_{1,0}$ or $V^h_{1,\frac{1}{2}}$ are to be considered. The results of Theorem II hold for the Mixed Problem.

APPENDIX B

According to the simplification introduced in Section 2.7.3.a we have

$$\underset{\sim}{K}^{stab} = \frac{\Omega_e}{135}\begin{pmatrix} 3(\delta_7\cdot\delta_7^T+\delta_8\cdot\delta_8^T) + \delta_9\cdot\delta_9^T & ; & 0 \\ 0 & ; & 3(\delta_7\cdot\delta_7^T + \delta_8\cdot\delta_8^T) + 4\delta_9\cdot\delta_9^T \end{pmatrix}$$

Note that

$$\delta_7 \cdot h^T = \delta_8 \cdot h^T = 0$$

$$\delta_9 \cdot h^T = 12$$

and that we can choose $\underset{\sim}{w} = (w_1, w_2)^T$ such that

$$\delta_9 \cdot w_1^T = \delta_9 \cdot w_2^T = 0$$

Then the control is

$$\begin{pmatrix}\tilde{u}_1\\ \tilde{u}_2\end{pmatrix} = \begin{pmatrix}u_1\\ u_2\end{pmatrix} - \lambda_1\begin{pmatrix}h\\ 0\end{pmatrix} - \lambda_2\begin{pmatrix}0\\ h\end{pmatrix} - \lambda_3\begin{pmatrix}w_1\\ w_2\end{pmatrix}$$

with

$$\lambda_i = \delta_9 \cdot u_i^T/12 \qquad i=1,2$$

$$\lambda_3 = \frac{(\delta_7\cdot w_1^T)(\delta_7\cdot u_1^T)+(\delta_8\cdot w_1^T)(\delta_8\cdot u_1^T)+(\delta_7\cdot w_2^T)(\delta_7\cdot u_2^T)+(\delta_8\cdot w_2^T)(\delta_8\cdot u_2^T)}{(\delta_7\cdot w_1^T)^2 + (\delta_8\cdot w_1^T)^2 + (\delta_7\cdot w_2^T) + (\delta_8\cdot w_2^T)^2}$$

APPENDIX C

```
      SUBROUTINE PROJ(U,XY)
C
C     THIS SUBROUTINE PROJECTS THE ELEMENT SOLUTION
C     ORTHOGONALLY W.R.T. HX,HY AND W
C     INPUT  :U  SOLUTION
C             XY NODAL COORDINATES(CURRENT CONFIGURATION)
C     OUTPUT :U  PROJECTED SOLUTION
C
      DIMENSION U(2,9),XY(2,9),W(2,9),S7(9),S8(9),S9(9),H(9)
      DIMENSION S7U(2),S8U(2),S9U(2),S7W(2),S8W(2)
      INTEGER SIGN
      DATA S7U,S8U,S9U,S7W,S8W/10*0./
      DATA S7/-1., 1., 1.,-1., 0.,-2., 0., 2., 0./
      DATA S8/-1.,-1., 1., 1., 2., 0.,-2., 0., 0./
      DATA S9/ 1., 1., 1., 1.,-2.,-2.,-2.,-2., 4./
      DATA  H/ 1., 1., 1., 1.,-1.,-1.,-1.,-1., 0./
      SIGN=-1
      DO 1 K=1,2
      K1=3-K
      IF(K.EQ.2) SIGN=1
      W(K,1)=SIGN*(+3.*XY(K1,1)-XY(K1,2)-XY(K1,3)-XY(K1,4))
      W(K,2)=SIGN*(+XY(K1,1)-3.*XY(K1,2)+XY(K1,3)+XY(K1,4))
      W(K,3)=SIGN*(-XY(K1,1)-XY(K1,2)+3.*XY(K1,3)-XY(K1,4))
      W(K,4)=SIGN*(+XY(K1,1)+XY(K1,2)+XY(K1,3)-3.*XY(K1,4))
      W(K,5)=SIGN*(XY(K1,3)-XY(K1,4))
      W(K,6)=SIGN*(XY(K1,1)-XY(K1,4))
      W(K,7)=SIGN*(XY(K1,1)-XY(K1,2))
      W(K,8)=SIGN*(XY(K1,3)-XY(K1,2))
    1 W(K,9)=0.
      DO 2 K=1,9
      DO 2 J=1,2
      S7U(J)=S7U(J)+S7(K)*U(J,K)
      S8U(J)=S8U(J)+S8(K)*U(J,K)
      S9U(J)=S9U(J)+S9(K)*U(J,K)
      S7W(J)=S7W(J)+S7(K)*W(J,K)
    2 S8W(J)=S8W(J)+S8(K)*W(J,K)
      S9U(1)=S9U(1)/12.
      S9U(2)=S9U(2)/12.
      W1=(S7W(1)*S7U(1)+S8W(2)*S8U(2)+S7W(2)*S7U(2)+S8W(1)*S8U(1))/
     .   (S7W(1)*S7W(1)+S8W(2)*S8W(2)+S7W(2)*S7W(2)+S8W(1)*S8W(1))
      DO 3 K=1,9
      DO 3 J=1,2
    3 U(J,K)=U(J,K)-S9U(J)*(H(K)+1./3.)-W1*W(J,K)
      RETURN
      END
```

REFERENCES

1. ALY,A.,S., "A Finite Element Analysis for Problem of Large Strain and Large Displacement, TICOM Report, 81-14, 1981.

2. BABUSKA, I., "The Finite Element Method with Lagrange Multipliers", Num. Math., Vol.20, 1973.

3. BELYTSCHKO, T. and KENNEDY, J.M., "Computer Models for Sub-assembly Simulation", Nuclear Engineering & Design, Vol.49, pp.17-38, July 1978.

4. BELYTSCHKO, ONG, J., S.-J. and LIU, W.K., "A Consistent Control of Spurious Singular Modes in the 9-Node Lagrange Element for the Laplace and Mindlin Plate Equation", Comp. Meth. in Appl. Mech. & Engrg., Vol.44, pp.269-295, 1984.

5. BELYTSCHKO, T. and TSAY, C.S., "A Stabilization Procedure for the Quadrilateral Plate Element with One-Point Quadrature", Comp. Meth. in Appl. Mech. & Engrg., Vol.19, pp.409-419, 1983.

6. BELYTSCHKO, T., TSAY, C.S. and LIU, W.K., "A Stabilization Matrix for the Bilinear Mindlin Plate Element", Comp. Meth. in Appl. Mech. & Engrg., Vol.29, pp.313-327, 1981.

7. BERCOVIER, M., "Perturbation of Mixed Varational Problems - Applications to Mixed Finite Element Methods", R.A.I.R.O., Vol.12., No.3, 1978.

8. BICANIC, N. and HINTON, E., "Spurious Modes in Two-Dimensional Isoparametric Elements", Int. J. Num. Methods in Engrg., Vol.14, pp.1545-1557, 1979.

9. BREZZI, R., "On the Existence,Uniqueness and Approximation of Saddle-Point Probems Arising from Lagrangian Multipliers", RAIRO, Numerical Analysis, 8, 1974.

10. CAREY, G.F. and KRISHNAN, R., "Penalty Approximation of Stokes Flow, Part I: Stability Analysis, Part II: Error Estimates and Numerical Results", TICOM Report, 82-5, June 1982.

11. CIARLET, P., Numerical Analysis of the Finite Element Method for Elliptic Boundary Problems, North Holland, Amsterdam, 1978.

12. CIARLET, P. and RAVIART, P.A., "The Combined Effect of Curved Boundaries and Numerical Integration in Isoparametric Finite Element Methods", The Mathematical Foundations of the Finite Element Method with Application to Partial Differential Equations

13.COOK, R.D., Concepts and Applications of Finite Element Analysis, New-York, Willey, 1981.

14.COOK, R.D. and ZHAO-HUA, F., "Control of Spurious Modes in the Nine-Node Quadrilateral Element", Int. J. Num. Methods in Engrg, Vol. 18, pp.1576-1580, 1982.

15.ENGLEMAN, M., FIDAP User's Manual, Boulder, Colorado, August 1981.

16.ENGLEMAN, M. AND SANI, R., The Mathematics of Finite Elements, Edited by J.R.Whiteman, Academic Press, Ltd., London, 1982.

17.FALK,R.S., "An Analysis of the Penalty Method and Extrapolation for the Stationary Stokes Problem", Advances in Computer Methods for Partial Differential Equations, Edited by R.Vichnevetsky, AICA Publication, pp.66-69, 1975.

18.FLANAGAN, D.P. and BELYTSCHKO, T., "A Uniform Strait Hexahedron and Quadrilateral with Orthogonal Hourglass Control", Int. J. Num. Methods in Engrg., Vol.17, pp.676-706, 1981.

19.FORTIN, M., "A Analysis of Convergence of Mixed Finite Element Methods", R.A.I.R.O., Vol.II, No.4, 1977.

20.FORTIN, M., "Old and New Finite Elements for Incompressible Flows", Int. J. Num. Methods in Fluids, Vol.1, pp.347-364, 1979.

21.GIRAULT, V., "Theory of a Finite Difference Method on Irregular Networks", SIAM J. Numer. Anal., Vol.II, pp.409-474, 1974.

22.GIRAULT, V., "Nonelliptic pprximation of a Class of Partial Differential Equations with Neumann Boundary Conditions", Mathematics of Computation, Vol.30, No.133, pp.68-91, January 1976.

23.GIRAULT, V. and RAVIART, P.A., Lecture Notes in Mathematics 749, Finite Element Approximation of the Navier-Stokes Equations, Springer-Verlag, Berlin, 1979.

24.HAYES, L.J., "Practical Stability Test for Finite Elements with Reduced Integration", Int. J. Num. Methods in Engrg., Vol. 17, pp. 1689-1695, 1981.

25.HUGHES, T.J.R., "Equivalence of Finite Elements for Nearly Incompressible Elasticity", J. Appl.Mech., March 1977.

26.JACQUOTTE, O.-P., "Stability,Accuracy and Efficiency of Some Underintegrated Methods in Finite Element Computations", Comp. Meth. in Appl. Mech. & Engrg., (to appear).

27. JACQUOTTE, O.-P. and ODEN, J.T., "Analysis of Hourglass Instabilities and Control in Underintegrated Finite Element Methods", Comp. Meth. in Appl. Mech. & Engrg., Vol.44, pp.339-363, 1984.

28. JACQUOTTE, O.-P. and ODEN, J.T., "Analysis and Treatment of Hourglass Instabilities in Underintegrated Finite Element Methods", Innovative Methods in Nonlinear Computational Mechanics", Edited by T.Belytschko, K.C.Park and W.K.Liu, Pineridge Press, Swansea, 1984 [presented at the ASEM Winter Annual Meeting, New Orleans, Louisiana, December 9-14, 1984]

29. JACQUOTTE, O.-P., ODEN, J.T. and BECKER, E.B., "Numerical Control of the Hourglass Instabilities", Computers and Structures, (to appear).

30. KOSLOFF, D. and FRAZIER, G.A., "Treatment of Hourglass Patterns in Low Order Finite Element Codes", Int. J. Num. Anal. Meth. in Geo-Mech., Vol.2, pp.57-72, 1978.

31. LADYSZHENSKAYA, O.A., The Mathematical Theory of Viscous Incompressible Flows, Gordon Breach, New York, 1969.

32. LE TALLEC, P., "Numerical Analysis of Equilibrium Problems in Incompressible Nonlinear Elasticity", Ph.D. Dissertation, The University of Texas at Austin, 1980.

33. MALKUS, D.S., "Finite Element Analysis of Incompressible Solids", Dissertation, Department of Mathematics, Boston University, 1972.

34. MALKUS, D.S., "A Finite Element Displacement Model Valid for Any Value of the Compressibility", International Journal of Solids and Structures Vol.12, pp.731-738, 1976.

35. MALKUS, D.S., "Finite Elements with Penalties for Incompressible Elasticity: A Progress Report", Illinois Inst. of Technology, Chicago, 1981.

36. MALKUS, D.S., HUGHES, T.J.R., "Mixed Finite Element Methods - Reduced and Selective Integration Technique - A Unification of Concepts", Comp. Meth. in Appl. Mech. & Engrg, Vol.15, pp.63-81, 1978.

37. MILLER, T., "A Finite Element Study of Instabilities in Rubber Elasticity", Ph.D. Dissertation, The University of Texas at Austin, 1983.

38. ODEN, J.T., Finite Elements in Nonlinear Continua, McGraw-Hill, New York, 1970.

39. ODEN, J.T., "R.I.P. Methods for Stokesian Flows", Finite Elements in Fluids, Vol.IV, Edited by Galaagher et al., John Wiley and Sons, Ltd, London, 1982 [presented at the Third Symposium of Finite Elements in Flow Problems, Bnaff, Canada, June 1980].

40. ODEN,J.T., "Penalty Methods for Constrained Problems in Nonlinear Elasticity", Proceedings, IUTAM Symposium on Finite Elasticity, Edited by D.E. Carlson and R.T. Shield, (Lehigh 1980), Martenus Nijhoff, The Hague, pp.281-300, 1982.

41. ODEN, J.T., "Penalty Method and Reduced Integration for the Analysis of Fluids", Proceedings, Symposium on Penalty Finite Element Methods, ASME Winter Annual Meeting, Phoenix, Arizona, November 14-19, 1982.

42. ODEN, J.T. and JACQUOTTE, O.-P., "Stable and Unstable RIP/Perturbed Lagrangian Methods for Two-Dimensional Viscous Flow Problems", Finite Elements in Fluids, Vol.V,ed. R.H. Gallagher et al, John Wiley and Sons, Ltd., London, 1982.

43. ODEN, J.T. and JACQUOTTE, O.-P., "A Stable Second-Order Accurate, Finite Element Scheme for Analysis of Two-Dimensional Incompressible Viscous Flows", Proceedings of The Fourth International Symposium Finite Element in Flow Problems", Tokyo, Japan, December 19-25, 1982.

44. ODEN, J.T. and JACQUOTTE, O.-P., "Stability of Some Mixed Finite Element Methods for Stokesian Flows", Comp. Meth. in Appl. Mech. & Engrg., Vol. 43, pp. 231-247, 1984.

45. ODEN, J.T. and KIKUCHI, N., "Penalty Methods for Constrained Problems in Elasticity", International Journal for Numerical Methods in Engineering, Vol.18, pp. 701-725, 1982.

46. ODEN, J.T., KIKUCHI, N. and SONG, Y.J., "Penalty-Finite Elements Methods for the Analysis of Stokesian Flows", Comp. Meth. in Appl. Mech. & Engrg., Vol.31, pp.297-329, 1982.

47. ODEN, J.T. and WELLFORD, L.C.,Jr., "Analysis of Flow of Viscous Fluids by the Finite Element Method", AIAA Journal, Vol. 10, No 12, pp. 1591-1599, 1972.

48. REDDY, J.N., "On the Accuracy and Existence of Solutions to Primitive Variable Models of Viscous Incompressible Fluids", International Journal of Engineering Science, 16-21, 921, 1978.

49. STRANG, G. and FIX, G., Analysis of the Finite Element Method, Prentice Hall, Englewood Cliffs, N.J. 1973.

50. TEMAN, R., Navier-Stokes Equations, North-Holland, Amsterdam, New-York, Oxford, 1979.

51. WAHLBIN, L.B., "Maximum Norm ErrorEstimates in the Finite Element Method with Quadratic Elements and Numerical Integration", R.A.I.R.O. Anal. Numer., Vol. 12, pp 173-202, 1978.

52. WAHLBIN, L.B., "A Remark on Parabolic Smoothing and the Finite Element Method", SIAM J.Numer. Anal., Vol. 17,No. 1, 1980.

53. ZIENKIEWICZ, O.C., Finite Element in Engineering, McGraw Hill, London, 1980.

54. ZIENKIEWICZ, O.C., The Finite Element Methods, McGraw Hill, London, 1977.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1994	Final Contractor Report

4. TITLE AND SUBTITLE

Analysis and Control of Hourglass Instabilities in Underintegrated Linear and Nonlinear Elasticity

5. FUNDING NUMBERS

WU-509-10-11
G-NAG3-329

6. AUTHOR(S)

Olivier P. Jacquotte and J. Tinsley Oden

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The University of Texas at Austin
Texas Institute for Computational Mechanics
Austin, Texas 78712

8. PERFORMING ORGANIZATION REPORT NUMBER

E-8659

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

National Aeronautics and Space Administration
Lewis Research Center
Cleveland, Ohio 44135-3191

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

NASA CR-195293

11. SUPPLEMENTARY NOTES

Project Manager, Christos C. Chamis, Structures Division, organization code 5200, NASA Lewis Research Center, (216) 433-3252.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified - Unlimited
Subject Category 39

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 words)***

Methods are described to identify and correct a bad finite element approximation of the governing operator obtained when under-integration is used in numerical code for several model problems: the Poisson problem, the linear elasticity problem, and for problems in the nonlinear theory of elasticity. For each of these problems, the reason for the occurrence of instabilities are given, a way to control or eliminate them are presented, and theorems of existence, uniqueness and convergence for the given methods are established. Finally, numerical results are included which illustrate the theory.

14. SUBJECT TERMS

Error estimates; Poisson problems; Elasticity problems; Nonlinear problems; Existence; Uniqueness; Convergence; Numerical results

15. NUMBER OF PAGES
127

16. PRICE CODE
A07

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	