

LINEAR LIBRARY

C01 0068 0039



MATHEMATICAL AND COMPUTATIONAL ASPECTS
OF THE
ENHANCED STRAIN FINITE ELEMENT METHOD

BY

K. ARUNAKIRINATHAR

M.Sc

Thesis presented for the degree of
DOCTOR OF PHILOSOPHY
in the department of Applied Mathematics
UNIVERSITY OF CAPE TOWN
SEPTEMBER 1994

The University of Cape Town has been given
the right to reproduce this thesis in whole
or in part. Copyright reserved by the author.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

TO PEACE AND HARMONY

DECLARATION

I, Kanagaratnam Arunakirinathar, hereby declare that this thesis is essentially my own work and that no part of it has been submitted for a degree at any other university.

Signed by candidate

signature removed
K Arunakirinathar

September 1994.

At 6pm the well marked $\frac{1}{2}$ inch of water, at nightfall $\frac{3}{4}$ and daybreak $\frac{7}{8}$ of an inch. By noon of the next day there was $\frac{15}{16}$ and on the next night $\frac{31}{32}$ of an inch of water in the hold. The situation was desperate. At this rate of increase few, if any, could tell where it would rise to in a few days.

Stephen Leacock

ABSTRACT

This thesis deals with further investigations of the enhanced strain finite element method, with particular attention given to the analysis of the method for isoparametric elements. It is shown that the results established earlier by B D Reddy and J C Simo for affine-equivalent meshes carry over to the case of isoparametric elements. That is, the method is stable and convergent provided that a set of three conditions are met, and convergence is at the same rate as in the standard method. The three conditions differ in some respects, though, from their counterparts for the affine case. A procedure for recovering the stress is shown to lead to an approximate stress which converges at the optimal rate to the actual stress.

The concept of the equivalent parallelogram associated with a quadrilateral is introduced. The quadrilateral may be regarded as a perturbation of this parallelogram, which is most conveniently described by making use of properties of the isoparametric map which defines the quadrilateral. The equivalent parallelogram generates a natural means of defining a regular family of quadrilaterals; this definition is used together with other properties to obtain in a relatively simple manner estimates, in appropriate seminorms or norms, of the isoparametric map and its Jacobian, for use in the determination of finite element interpolation error estimates.

With regard to computations, a new basis for enhanced strains is introduced, and various examples have been tested. The results obtained are compared with those obtained using other bases, and with those found from an assumed stress approach. Favourable comparisons are obtained in most cases, with the present basis exhibiting an improvement over existing bases. Convergence of the finite element results are verified; it is observed numerically that the improvement of results due to enhancement is as a result of a smaller constant appearing in the error estimates.

ACKNOWLEDGEMENTS

I am greatly indebted to my supervisor, Professor Dayanand Reddy for his invaluable assistance and guidance during the course of this studies and the preparation of this manuscript.

I would like to thank Professor Peter Wriggers for giving us the program PCFEAP.

I would like to thank Conrad Mellin and Kevin Colville for their assistance in proof reading the text. Many thanks to other freinds for their support and encouragement.

I thank my wife Santhini for her unfailing tolerance, support and encouragement.

Finally I would like to thank the Foundation for Research and Development and the FRD/UCT Center for REsearch in Computational and Applied Mechanics(CERECAM) for their financial support.

TABLE OF CONTENTS

Declaration	iii
Abstract	iv
Acknowledgements	v
Table of Contents	vi
1. Introduction	1
1.1 Background	1
1.2 Present work	5
2. Finite Element Approximation and Interpolation	8
2.1 Key definitions and notations	9
2.1.1 Function spaces	12
2.2 Interpolation estimates	17
2.3 Finite element interpolation and approximation	24
2.3.1 Equivalent families of finite elements	27
2.4 The equivalent parallelogram of a quadrilateral	32
2.5 Interpolation error estimates for quadrilateral finite elements	36
3. The Enhanced Strain Finite Element Method	47
3.1 The continuous problem	47
3.2 The discrete problem	54
3.3 Convergence and error estimates	62
3.4 Stress recovery	67
4. Numerical Results	72
4.1 A new basis for enhanced strains	72

4.2 Numerical experiments	77
5. Conclusions	99
5.1 Remarks on work covered in the thesis	99
5.2 Future work	100
References	102

Chapter 1

Introduction

1.1 Background

In the finite element analysis of problems in solid and structural mechanics, it is well-known that the standard low-order elements exhibit rather poor performance in analyses of bending-dominated problems; furthermore, these elements exhibit locking response in the *thin* limit (in beam, plate and shell problems), and in the nearly incompressible limit. However, low-order elements are often preferred due to their economy, and thus there has been a long history of attempts to develop successful finite element schemes for low-order elements to overcome these difficulties.

Most of the methods proposed to improve the performance of the low-order elements can be classified as hybrid-mixed methods; that is, in addition to the displacement field, one or more independent fields are introduced for strains, stresses or non-conforming (incompatible) displacements to enrich the standard (conforming) finite element subspace. In general, these additional fields are chosen to be discontinuous across the element boundaries so that the corresponding degrees of freedom are easily eliminated at the element level. This leads to a system of simultaneous equations with a modified element stiffness matrix.

Solutions are improved significantly for bending-dominated problems with the four-noded rectangular element, by using the method of incompatible modes developed by Wilson *et al.* [59], in which additional nonconforming quadratic terms are included to enhance the standard displacement field. This element is, however, not stable for arbitrary quadrilateral elements, and so it was subsequently modified by Taylor *et al.* [56] in such a way as to be stable in the quadrilateral case.

Recently, Simo and Rifai [48] have proposed a new general methodology, called the *enhanced strain method*, in the context of small strain problems. In this method the strain tensor $\boldsymbol{\varepsilon}$ is expressed as a sum of the conventional strain $\boldsymbol{\varepsilon}(\mathbf{u})$ and an additional tensor $\tilde{\boldsymbol{\varepsilon}}$, called the enhanced strain: that is,

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}(\mathbf{u}) + \tilde{\boldsymbol{\varepsilon}},$$

where $\boldsymbol{\varepsilon}(\mathbf{u}) = \frac{1}{2} (\nabla^T \mathbf{u} + \nabla \mathbf{u})$. This approach has a number of additional advantages:

- it gives high accuracy for coarse meshes in bending-dominated problems
- basis vectors are easily constructed
- it exhibits locking-free response in the nearly incompressible regime
- it is extendable to nonlinear problems.

The rectangular incompatible mode element of Wilson *et al.* [59] and its extension to isoparametric quadrilateral elements by Taylor *et al.* [56] appear as special cases of this method.

Subsequent work in the area of enhanced strain elements has included the development of new four-noded membrane, plate and shell elements, as well as eight-noded brick elements by Andelfinger and Ramm [2], and the extension of the method into the materially and geometrically nonlinear regime by Simo *et al.* [49, 50].

It is also worth mentioning the related work on assumed stress (or hybrid stress) elements, of which the example due to Pian and Sumihara [39] is perhaps the best known; further variants of this class of elements have been studied by Di and Ramm [20], while the relationship between the assumed stress and enhanced strain methods, based respectively on Hellinger-Reissner and Hu-Washizu-type variational formulations, has been addressed by Andelfinger *et al.* [2]. Several other modified versions of Pian's element have been introduced with improved accuracy [40, 41, 60, 61]. Although the assumed stress element shows good performance with low-order elements, the application to nonlinear and plasticity problems leads to rather cumbersome extensions [50].

Recently, an improved incompatible mode element has been introduced and studied by Hueck *et al.* [26, 27]; the basis functions for this element are obtained by making use of truncated Taylor series expansions of the gradients in physical coordinates, and the result is an element which is stable for arbitrary quadrilaterals.

Since exact solutions are not available for many problems in engineering and science the reliability of finite element methods depends on understanding the mathematical framework of the methodology. The nonconforming method of Wilson *et al.* was analysed for a uniform mesh of parallelograms by Lesaint [30], with key contributions by Strang (see [53, 54], and also the treatment by Ciarlet [18, Chapter 4]). The key to the analysis lies in the second Strang Lemma [54], which makes precise the consistency error due to the nonconformity.

For the arbitrary quadrilateral case, it is well known that Wilson's incompatible element is not stable due to the appearance of the bilinear term in the isoparametric map. Lesaint and Zlamal [31] have modified the variational formulation in accordance with Taylor [56], and proved the convergence. Without such a modification of the variational formulation, Shi [46] has proved convergence for Wilson's element, but under the condition that the distance between the midpoints of the diagonals of quadrilateral is of order $O(h^2)$. In other words, Shi has imposed a condition on the coefficient of the bilinear term in order to obtain the convergence. All of these

problems have been considered in the context of linear elasticity. Very recently, application of Wilson's incompatible modes to Mindlin plates is reported in [62]; a stable element similar to that of Bathe and Dvorkin [6] results.

A detailed mathematical analysis of the enhanced strain method was carried out by Reddy and Simo [44] for the case of affine-equivalent meshes, that is, meshes comprising n -rectangles or n -parallelograms (subparametric cases), for problems in \mathbb{R}^n . In this work it was shown that the method is stable and convergent provided that a set of three conditions are met by the enhanced strain basis (*cf.* the conditions (I)–(III) later, in Section 3.2); furthermore, convergence is at the optimal rate. The investigation in [44] also included a detailed treatment of the case of small compressibility and incompressibility. It was shown that, for the basis in two dimensions introduced in [48], no locking occurs; furthermore, the checkerboard mode is the only spurious pressure mode, and this may be filtered out, as in the case of the $Q_1 - P_0$ (bilinear velocity-constant pressure) element, to give an approximation which is convergent at the optimal rate.

Reddy and Simo have shown that meshes comprising triangular elements, with conventional Lagrangian basis functions defined on these elements, are ruled out because imposition of the criteria which have to be satisfied by the enhanced strains results in the problem being entirely equivalent to the standard problem.

1.2 Present work

The purpose of this contribution is to further develop various aspects of the enhanced strain finite element method in respect of linear problems. First, the analysis of [44] is extended to include the important case of isoparametric elements. The task is not as easy as for the affine case because of the appearance of the nonlinear term in the isoparametric transformation. In [46] it was assumed that the coefficient of this nonlinear term is of order $O(h^2)$, in order to obtain convergence. No such

assumptions are made here.

We give a modified version of the conditions introduced in [44], to obtain the required asymptotic estimates for the enhanced strain method. A general estimate is obtained for the error, and asymptotic results are given for plane bilinear and biquadratic elements with the use of the geometrical estimates obtained. The analysis presented here also includes a discussion of the stress recovery process; it is shown that the recovery process proposed in [48] yields stresses which converge at the optimal rate.

No analysis is given for the incompressible case: Reddy and Simo [44] did the analysis for rectangular elements; arbitrary quadrilaterals are very difficult, given that the $Q_1 - P_0$ element for the arbitrary case still resists a complete analysis.

In order to obtain the concrete form of the error estimates, we need some geometrical estimates pertaining to quadrilateral finite elements. Therefore the *second* contribution of this thesis is a study of geometrical estimates for quadrilateral elements. The case of the four-noded quadrilateral differs from that of simplicial elements in an important sense, namely, that whereas simplicial curved elements are conveniently regarded as perturbations of triangular or tetrahedral elements, it is not immediately clear as to how this approach may be carried over to quadrilaterals and their three-dimensional counterparts, especially if the perturbations are to be regarded as small in some sense. This is also commented on in the text by Ciarlet [18] (see page 243). It is therefore necessary to adopt an approach which differs from that employed in the case of curved triangular elements, if the idea of the quadrilateral as a perturbation of a parallelogram is to form the basis of the theory. This issue is addressed by introducing the notion of the equivalent parallelogram associated with a quadrilateral. This parallelogram has the property that the quadrilateral may be regarded as a perturbation, in the sense that the difference between these two figures, as measured in an appropriate way, is $O(h)$, where h is the diameter of the quadrilateral.

The *third* contribution which we make is to propose a new basis for enhanced strains.

This basis differs from those which are in current use, in that the geometry of the actual element is included in the basis as defined on the master element.

The *fourth* contribution is a study of the numerical behaviour of the enhanced strain method and, in particular, of numerical rates of convergence. The asymptotic error estimates predict convergence at the same rate as the standard method, and therefore do not explain the superior performance of the enhanced strain method. The numerical results presented here suggest an explanation for this superior performance, in that the constant C which appears in estimates of the form $\|\text{error}\| \leq C h^k$ appear to be much smaller for the enhanced strain method than for the standard method.

Outline of this work

The plan of the rest of this thesis is as follows. In Chapter 2 some key mathematical definitions and notations are presented, after which we review those aspects of finite element interpolation theory which are required for subsequent developments. We then introduce the notion of the equivalent parallelogram associated with a quadrilateral. We highlight some interesting geometrical properties of quadrilaterals and their associated equivalent parallelograms, which properties allow the estimates necessary in an interpolation theory to be arrived in a simple way.

In Chapter 3, the continuous variational problem for linear elasticity is formulated in a manner appropriate for introduction of the enhanced strain as an additional variable. Well-posedness of the continuous variational problem is discussed. The discrete problem is then posed. Three conditions are placed on the choice of finite element spaces; these are similar to those used in the affine case. These conditions are sufficient to show that the method is stable and convergent, with convergence at the same rate as the standard problem.

The analysis presented here also includes a discussion of the stress recovery process. In [48] a least squares formulation is proposed to recover stresses. We give an

alternative formulation, but equivalent to that of [48]. It is shown that this process yields stresses which converge at the optimal rate.

In Chapter 4 we first introduce a new basis for enhanced strains, and then we report on numerical results which demonstrate the improved performance of this and various other enhanced strain elements. These results also give an idea of actual rates of convergence.

We conclude with Chapter 5 in which we discuss possibilities for future research.

Portions of the work reported in this thesis have been submitted for publication; the work of Section 2.4 appears in [3] while that of Chapters 3 and 4 appear in [4].

Chapter 2

Finite Element Approximation and Interpolation

The purpose of this Chapter is to recall, and to prove, a number of results relating to finite element approximation and interpolation theory in two and three dimensions, of which frequent use will be made in the sequel. In Section 2.1 we give some central definitions and notations concerning function spaces which are required subsequently. In order to make this thesis as self-contained as possible we review in Sections 2.2 and 2.3 the basic ideas of finite element interpolation and finite element approximation in general.

Later developments will be exclusively concerned with quadrilaterals but estimates for the various quantities that arise in interpolation error estimates are not readily available for these elements. With this in mind we introduce in Section 2.4 the new concept of *the equivalent parallelogram* associated with a quadrilateral, and use this concept to derive various estimates for quadrilateral elements.

2.1 Key definitions and notations

In this Section we review, rather briefly, the results, definitions and basic notations and terminology which are required subsequently. Further details can be found in [13, 18, 34]. Throughout the thesis we use the notations c, C, \tilde{C} , etc. to denote positive constants which may have different values in different equations or inequalities.

We deal only with real vector spaces. Let X and Y be two Banach spaces. The vector space consisting of all k -linear mappings from X^k to Y is denoted by $\mathcal{L}_k(X; Y)$ or simply by $\mathcal{L}_k(X)$ if $X = Y$. When $k = 1$ the map is often written as $\mathcal{L}(X; Y)$ instead of $\mathcal{L}_1(X; Y)$. The space $\mathcal{L}_k(X; Y)$ is a normed space under the norm

$$\|A\| = \sup\{|A(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \dots, \boldsymbol{\eta}_k)| : \|\boldsymbol{\eta}_i\| \leq 1; 1 \leq i \leq k, \} \quad (2.1)$$

for all $A \in \mathcal{L}_k(X; Y)$.

If $A \in \mathcal{L}_k(\mathbb{R}^n; Y)$, we set

$$[A] = \max_{1 \leq i \leq n} |A(\mathbf{e}_i)^k|,$$

where $\{\mathbf{e}_i\}_{i=1}^n$ is the canonical basis of \mathbb{R}^n and $A(\mathbf{e}_i)^k = A(\mathbf{e}_i, \mathbf{e}_i, \dots, \mathbf{e}_i)$.

We note that $\|A\| \leq \frac{k^k}{k!} [A]$ for all $A \in \mathcal{L}_k(\mathbb{R}^n; Y)$, so that $[A]$ is a norm equivalent to the norm $\|A\|$ on the space $\mathcal{L}_k(\mathbb{R}^n; Y)$ [16].

Given any subset $\Omega \subset \mathbb{R}^n$, $F : \Omega \rightarrow \mathbb{R}^n$ is said to be an *affine map* if F is of the form

$$F(\mathbf{x}) = B\mathbf{x} + \mathbf{c} \quad \text{for all } \mathbf{x} \in \Omega,$$

where $B \in \mathcal{L}(\mathbb{R}^n)$ and $c \in \mathbb{R}^n$.

We now define the notion of the derivative of a function. Let \mathbf{u} be a function from an open set Ω of a normed vector space X into a normed vector space Y . The function \mathbf{u} is said to be differentiable at a point $\mathbf{a} \in \Omega$ if there exists a linear map $D\mathbf{u}(\mathbf{a}) \in \mathcal{L}(X; Y)$ such that, for all $\mathbf{h} \in X$,

$$\lim_{\|\mathbf{h}\|_X \rightarrow 0} \frac{1}{\|\mathbf{h}\|_X} \|\mathbf{u}(\mathbf{a} + \mathbf{h}) - \mathbf{u}(\mathbf{a}) - D\mathbf{u}(\mathbf{a}) \cdot \mathbf{h}\| = 0.$$

Then $D\mathbf{u}(\mathbf{a})$ is called the Fréchet derivative of \mathbf{u} at \mathbf{a} . If the Fréchet derivative at \mathbf{a} exists it is unique.

We say that *the function \mathbf{u} is differentiable* on Ω if it is differentiable at every point of Ω . It is then possible to define the function

$$D\mathbf{u} : \mathbf{x} \in \Omega \subset X \rightarrow D\mathbf{u}(\mathbf{x}) \in \mathcal{L}(X; Y),$$

which is called the Fréchet derivative. If $D\mathbf{u}$ is continuous in Ω , the function \mathbf{u} is said to be continuously differentiable in Ω , and we write

$$\mathbf{u} \in \mathcal{C}^1(\Omega).$$

Higher order Fréchet derivatives can be defined in the usual way (see Cartan [13] for further details). The k -th Fréchet derivative of \mathbf{u} at a point \mathbf{a} , denoted by $D^k\mathbf{u}(\mathbf{a})$, is a symmetric k -linear mapping of X^k into Y in the sense that $D^k\mathbf{u}(\mathbf{a})(\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_k) = D^k\mathbf{u}(\mathbf{a})(\mathbf{h}_{\sigma_1}, \mathbf{h}_{\sigma_2}, \dots, \mathbf{h}_{\sigma_k})$ for any permutation $\sigma : i \rightarrow \sigma_i$ of the set $\{1, 2, \dots, k\}$. In case $\mathbf{h}_i = \mathbf{h}$ for all $1 \leq i \leq k$, we shall simply write $D^k\mathbf{u}(\mathbf{a})(\mathbf{h})^k$.

The natural norm of the k th derivative $D^k\mathbf{u}(\mathbf{a})$ is defined as the norm of an element of the space $\mathcal{L}_k(X; Y)$.

For the case in which $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$, the first derivative $D\mathbf{u}(\mathbf{a})$ may be represented in matrix form as

$$D\mathbf{u}(\mathbf{a}) = \begin{pmatrix} \partial_1 u_1(\mathbf{a}) & \partial_2 u_1(\mathbf{a}) & \cdots & \partial_n u_1(\mathbf{a}) \\ \partial_1 u_2(\mathbf{a}) & \partial_2 u_2(\mathbf{a}) & \cdots & \partial_n u_2(\mathbf{a}) \\ \vdots & \vdots & & \vdots \\ \partial_1 u_m(\mathbf{a}) & \partial_2 u_m(\mathbf{a}) & \cdots & \partial_n u_m(\mathbf{a}) \end{pmatrix},$$

where $\partial_i u_j(\mathbf{a}) = \frac{\partial u_j}{\partial x_i}(\mathbf{a})$ and $\mathbf{u} = (u_1, u_2, \dots, u_m)$. If $m = n$, the determinant of this matrix is called the *Jacobian* of the function \mathbf{u} at \mathbf{a} , and is denoted by $J_{\mathbf{u}}$.

It may be observed that the partial derivatives $\partial_i \mathbf{u}(\mathbf{a})$ satisfy

$$\partial_i \mathbf{u}(\mathbf{a}) = D\mathbf{u}(\mathbf{a}) \cdot \mathbf{e}_i.$$

Similarly, the higher order partial derivatives are given by

$$\partial_{ij} \mathbf{u}(\mathbf{a}) = D^2 \mathbf{u}(\mathbf{a})(\mathbf{e}_i, \mathbf{e}_j), \quad \partial_{ijk} \mathbf{u}(\mathbf{a}) = D^3 \mathbf{u}(\mathbf{a})(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k), \quad \text{etc.}$$

Throughout this thesis, Ω will denote an open bounded set in \mathbb{R}^n with boundary $\partial\Omega$. We will also assume that Ω is *bounded and simply connected*, and that $\partial\Omega$ is *smooth*. By a smooth boundary we will mean that $\partial\Omega$ is at least *Lipschitzian*: that is, $\partial\Omega$ can be represented as

$$\partial\Omega = \cup_{r=1}^R \{(x_1^r, \mathbf{x}^r) : x_1^r = \Phi^r(\mathbf{x}^r), |\mathbf{x}^r| < \beta; \mathbf{x}^r = (x_2^r, x_3^r, \dots, x_n^r)\}$$

where $\{\Phi^r\}$ is a system of local Lipschitz-continuous coordinate maps, (x_1^r, \mathbf{x}^r) is a local coordinate system and β is a sufficiently large positive number. Here $|\mathbf{x}^r| < \beta$ stands for $|x_i^r| < \beta$, $2 \leq i \leq n$.

2.1.1 Function spaces

A successful finite element analysis depends on the right choice of function spaces appropriate to the given problem. Before we define some function spaces it is necessary to say that we deal with real-valued functions for the sake of convenience.

In what follows we make reference to the following function spaces:

$C^m(\Omega)$: the linear space consisting of all functions u which have continuous partial derivatives (in all variables) up to and including those of order m

$C^\infty(\Omega)$: the space of infinitely differentiable functions on Ω

$\mathcal{D}(\Omega)$: the space of test functions defined on Ω equipped with the usual locally convex topology: $\phi \in \mathcal{D}(\Omega)$ if

(i) ϕ has compact support in Ω ;

(ii) $\phi \in C^\infty(\Omega)$

$\mathcal{D}'(\Omega)$: the space of distributions: that is, the topological dual of $\mathcal{D}(\Omega)$ endowed with the strong dual topology.

We now recall the definition of the space $L_2(\Omega)$ of equivalence classes of (Lebesgue) square-integrable functions u on Ω ; that is

$$u \in L_2(\Omega) \iff \left(\int_{\Omega} |u|^2 \, d\mathbf{x} \right)^{\frac{1}{2}} < \infty$$

Here and henceforth the volume element is denoted by $d\mathbf{x} = dx_1 \cdots dx_n$. Then $L_2(\Omega)$ is a Hilbert space with the inner product

$$(u, v)_{0,\Omega} = \int_{\Omega} u(\mathbf{x})v(\mathbf{x}) \, d\mathbf{x} \quad \text{and the norm } \|u\|_0 = (u, u)_{0,\Omega}^{\frac{1}{2}}.$$

We recall the multi-index notation for partial derivatives. Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ be an n -tuple of nonnegative integers and let $|\alpha| = \sum_{j=1}^n \alpha_j$: then we define the symbol $D^\alpha u$ to mean

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}. \quad (2.2)$$

It is important to recall that in contrast to the ordinary derivative operator, the distributional derivative is a continuous functional on the space of distributions $\mathcal{D}'(\Omega)$. We refer to the elements of the $\mathcal{D}'(\Omega)$, the dual of $\mathcal{D}(\Omega)$, as distributions on Ω . If $u \in \mathcal{D}'(\Omega)$ its derivative $D^j u$ is defined by

$$D^j u(\phi) = -u(D^j \phi) \quad \forall \phi \in \mathcal{D}(\Omega).$$

Differentiation is a continuous operator in $\mathcal{D}'(\Omega)$. Furthermore, the order of successive differentiation is always immaterial, in contrast to the ordinary partial differentiation.

We also note that

$$|D^\alpha u| \leq \|D^k u\| \leq C(k, n) \max_{|\alpha|=k} |D^\alpha u|. \quad (2.3)$$

We are now ready to introduce the (Hilbertian) *Sobolev spaces* which are of great use and significance in any discussion concerning finite element methods.

For each integer $m \geq 0$, let $H^m(\Omega)$ be the space of equivalence classes of functions which, together with their distributional derivatives of order $|\alpha| \leq m$ belong to the space $L_2(\Omega)$. The space $H^m(\Omega)$ is a Hilbert space when endowed with the inner product $(\cdot, \cdot)_m$ and associated norm $\|\cdot\|_m$ given by

$$(u, v)_{m, \Omega} = \sum_{|\alpha| \leq m} \int_{\Omega} D^\alpha u D^\alpha v \, d\mathbf{x}, \quad \|u\|_{m, \Omega}^2 = (u, u)_{m, \Omega}$$

for all $u, v \in H^m(\Omega)$. Here the summation is over all possible combinations of indices such that $|\alpha| \leq m$.

A normed space U is said to be embedded in a normed space V (with norms $\|\cdot\|_U$ and $\|\cdot\|_V$, respectively) if

- (i) U is a linear subspace of V and
- (ii) the injection of U into V is continuous; that is there exists a constant $C > 0$ such that

$$\|u\|_V \leq C\|u\|_U \quad \forall u \in U.$$

To denote such embeddings, we write $U \hookrightarrow V$.

The Sobolev embedding theorem asserts the existence of embeddings of $H^m(\Omega)$ into following spaces:

- (i) $H^j(\Omega), 0 \leq j \leq m$ that is; $H^m(\Omega) \hookrightarrow H^j(\Omega)$, and
- (ii) $\bar{C}^j(\Omega) := \{u \in C^j(\Omega) : D^{|\alpha|}u \text{ is bounded on } \Omega \text{ for } |\alpha| \leq j\}$.

We also define the closed subspace $H_0^m(\Omega)$ of $H^m(\Omega)$, for $m \geq 1$ by

$$H_0^m(\Omega) = \{u : u \in H^m(\Omega) \text{ and } u = Du = \dots = D^\alpha u = 0 \text{ on } \partial\Omega, |\alpha| \leq m - 1\},$$

where $\partial\Omega$ denotes the boundary of Ω and boundary values are in the sense of traces.

We will make use of the seminorm $|\cdot|_{m,\Omega}$ on $H^m(\Omega)$, defined by

$$|u|_{m,\Omega}^2 = \int_{\Omega} \sum_{|\alpha|=m} |D^\alpha u(x)|^2 dx.$$

The seminorm $|\cdot|_1$ is a norm over the space $H_0^1(\Omega)$, equivalent to the norm $\|\cdot\|_1$ by virtue of the following inequality, known as the Poincaré inequality: for a bounded domain Ω , there exists a constant $c(\Omega)$ such that

$$|u|_{0,\Omega} \leq c(\Omega) |u|_{1,\Omega} \quad \text{for all } u \in H_0^1(\Omega). \quad (2.4)$$

For a bounded domain with Lipschitz boundary, this is also true for seminorm $|\cdot|_m$; that is, $|\cdot|_m$ is a norm over the space $H_0^m(\Omega)$, equivalent to the norm $\|\cdot\|_m$ ([18, see page 12]).

In subsequent work we will also make frequent use of the seminorm $[\cdot]_{m,\Omega}$ on $H^m(\Omega)$ which is defined by

$$[u]_{m,\Omega}^2 = \int_{\Omega} \sum_{i=1}^n |D^m u(\mathbf{x})(\mathbf{e}_i^m)|^2 dx, \quad (2.5)$$

where $D^m u(\mathbf{x})(\mathbf{e}_i^m) = \frac{\partial^m u}{\partial x_i^m}$.

We introduce the space $Q_k(\Omega)$ of polynomials q of the form

$$q(\mathbf{x}) = \sum_{i_1, i_2, \dots, i_n \leq k} C(i_1, i_2, \dots, i_n) x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}, \quad (2.6)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ and $n \geq 2$. For example, if $k = 1$ and $n = 2$, $Q_1(\Omega)$ is the space of polynomials of the form $a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2$ (or rather $a_0 + a_1 x + a_2 y + a_3 xy$).

We denote by $H^{k+1}(\Omega)/Q_k(\Omega)$ the quotient space whose members are cosets \dot{v} of functions, that is,

$$\dot{v} = \left\{ w \in H^{k+1}(\Omega) : w - v \in Q_k(\Omega) \right\}.$$

The natural norm on this quotient space is given by

$$\|\dot{v}\|_{k+1,\Omega} = \inf_{q \in Q_k(\Omega)} \|v + q\|_{k+1,\Omega}. \quad (2.7)$$

Remarks.

1. Note that $|v|_{1,\Omega} = [v]_{1,\Omega}$ for all $v \in H_0^1(\Omega)$, and hence the semi-norm $[\cdot]_{1,\Omega}$ satisfies the Poincaré inequality on $H_0^1(\Omega)$.
2. For all $v \in H_0^2(\Omega)$, $|v|_{2,\Omega} \leq \sqrt{n} [v]_{2,\Omega}$.
3. It has been shown in [51] that there exists a constant $C = C(\Omega)$ such that

$$\|v\|_{m,\Omega} \leq C([v]_{m,\Omega} + \|v\|_{0,\Omega}), \quad \text{for all } v \in H^m(\Omega). \quad (2.8)$$

This inequality is useful in showing that the seminorm $[u]_{m,\Omega}$ is a norm on the quotient space $H^{k+1}(\Omega)/Q_k(\Omega)$, equivalent to the natural norm.

The quotient space $H^{k+1}(\Omega)/Q_k(\Omega)$ is a Hilbert space equipped with the quotient norm (2.7). If $\dot{u} \in H^{k+1}(\Omega)/Q_k(\Omega)$ then the mapping $\dot{u} \rightarrow [\dot{u}]_{k+1,\Omega} = [u]_{k+1,\Omega}$ is *a priori* only a seminorm on the quotient space $H^{k+1}(\Omega)/Q_k(\Omega)$.

Since any polynomial $q \in Q_k(\Omega)$ satisfies $D^{k+1}q \cdot (e_i)^{k+1} = 0$ for all i , we have

$$\sum_{i=1}^n \|D^{k+1}u \cdot (e_i)^{k+1}\|_{0,\Omega} = \sum_{i=1}^n \|D^{k+1}(u + q) \cdot (e_i)^{k+1}\|_{0,\Omega} \leq \|u + q\|_{k+1,\Omega} \quad \text{for all } q \in Q_k(\Omega)$$

Taking the infimum over all $q \in Q_k(\Omega)$ we obtain the inequality

$$[u]_{k+1,\Omega} \leq \|\dot{u}\|_{k+1,\Omega} \quad \text{for all } \dot{u} \in H^{k+1}(\Omega)/Q_k(\Omega). \quad (2.9)$$

It can in fact be shown that the seminorm $[u]_{k+1,\Omega}$ is equivalent to the natural norm defined by (2.7), on the quotient space $H^{k+1}(\Omega)/Q_k(\Omega)$. This equivalence can

be established from a result given by Bramble and Hilbert [10, Theorem 1] which allows us to prove the following Theorem (see also Lazarov [29, Lemma 1]).

Theorem 2.1 *There exists a constant $c(\Omega)$ such that*

$$\inf_{q \in Q_k(\Omega)} \|u + q\|_{k+1, \Omega} \leq C(\Omega) [u]_{k+1, \Omega} \quad \text{for all } u \in H^{k+1}(\Omega). \quad (2.10)$$

Proof

See, for example, [10] for the proof of this Theorem. □

2.2 Interpolation estimates

We begin by proving an important estimate which determines the order of accuracy of polynomial interpolations of the type (2.6). This Theorem is according to the exercise given in [18, Section 3.1].

Theorem 2.2 *For some integers $k \geq 0$ and $m \geq 0$, let $H^{k+1}(\Omega)$ and $H^m(\Omega)$ be Sobolev spaces satisfying the inclusion $H^{k+1}(\Omega) \hookrightarrow H^m(\Omega)$, and let $\Pi \in \mathcal{L}(H^{k+1}(\Omega); H^m(\Omega))$ be a mapping such that*

$$\Pi q = q \quad \text{for all } q \in Q_k(\Omega). \quad (2.11)$$

Then there exists a constant $c(\Omega)$ such that

$$\|v - \Pi v\|_{m, \Omega} \leq C(\Omega) [v]_{k+1, \Omega}. \quad (2.12)$$

Proof

To establish the required inequality, we observe that

$$v - \Pi v = v + q - \Pi q - \Pi v$$

$$= (I - \Pi)(v + q),$$

for all $v \in H^{k+1}(\Omega)$ and $q \in Q_k(\Omega)$. Here I , the identity mapping from $H^{k+1}(\Omega)$ into $H^m(\Omega)$, is continuous.

From this identity we deduce that

$$\begin{aligned} |v - \Pi v|_{m,\Omega} &\leq \|I - \Pi\|_L \inf_{q \in Q_k(\Omega)} \|v + q\|_{k+1,\Omega} \\ &\leq c(\Pi, \Omega) [v]_{k+1,\Omega}, \end{aligned}$$

by Theorem 2.1. □

Before proceeding to the next topic, it is appropriate to define another norm which is needed for subsequent developments. If F is a sufficiently smooth one-to-one invertible map from $\hat{\Omega} \subset \mathbb{R}^n$ onto $\Omega \subset \mathbb{R}^n$ we define, for $m \geq 1$,

$$\begin{aligned} |F|_{m,\infty} &= \sup_{\xi \in \hat{\Omega}} \|D^m F(\xi)\|, \\ |F^{-1}|_{m,\infty} &= \sup_{x \in \Omega} \|D^m F^{-1}(x)\|, \end{aligned}$$

and

$$\|F\|_{m,\infty} = \sup_{\xi \in \hat{\Omega}} [D^m F(\xi)]. \quad (2.13)$$

The norm $\|F\|_{m,\infty}$ is equivalent to the standard norm $|\cdot|_{m,\infty}$ on $\mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$ since

$$|F|_{m,\infty} \leq \sqrt{n} \frac{m^m}{m!} \|F\|_{m,\infty}$$

(see [35, page 7]).

The following Theorems [18] summarise relationships between the seminorms $|\cdot|$ and $[\cdot]$ on Ω and $\hat{\Omega}$.

Theorem 2.3 *Let Ω and $\hat{\Omega}$ be two bounded open subsets of \mathbb{R}^n such that $\Omega = F(\hat{\Omega})$, where F is a sufficiently smooth one-to-one mapping with a sufficiently smooth in-*

verse $F^{-1} : \Omega \rightarrow \hat{\Omega}$. Then if a function $\hat{u} : \hat{\Omega} \rightarrow \mathbb{R}$ belongs to the space $H^l(\hat{\Omega})$ for some integer $l \geq 0$, the function $u = \hat{u} \circ F^{-1} : \Omega \rightarrow \mathbb{R}^n$ belongs to the space $H^l(\Omega)$ and, in addition, there exist constants C such that

$$|u|_{0,\Omega} \leq |J|_{0,\infty}^{1/2} |\hat{u}|_{1,\hat{\Omega}} \quad \forall \hat{u} \in H^1(\hat{\Omega}) \quad (2.14)$$

$$|u|_{1,\Omega} \leq C |J|_{0,\infty}^{1/2} \|F^{-1}\|_{1,\infty} |\hat{u}|_{1,\hat{\Omega}} \quad \forall \hat{u} \in H^1(\hat{\Omega}) \quad (2.15)$$

$$|u|_{2,\Omega} \leq C |J|_{0,\infty}^{1/2} (\|F^{-1}\|_{1,\infty}^2 |\hat{u}|_{2,\hat{\Omega}} + \|F^{-1}\|_{2,\infty} |\hat{u}|_{1,\hat{\Omega}}) \quad \forall \hat{u} \in H^2(\hat{\Omega}) \quad (2.16)$$

$$|u|_{3,\Omega} \leq C |J|_{0,\infty}^{1/2} \{ \|F^{-1}\|_{1,\infty}^3 |\hat{u}|_{3,\hat{\Omega}} + \|F^{-1}\|_{1,\infty} \|F^{-1}\|_{2,\infty} |\hat{u}|_{2,\hat{\Omega}} + \|F^{-1}\|_{3,\infty} |\hat{u}|_{1,\hat{\Omega}} \} \quad \forall \hat{u} \in H^3(\hat{\Omega}) \quad (2.17)$$

Here J denotes the Jacobian of the map F .

Proof

The proof of this Theorem may be found in [18]. □

The following Theorem is given as an exercise in [18, Section 4.3].

Theorem 2.4 *Let Ω and $\hat{\Omega}$ be two bounded open subsets of \mathbb{R}^n such that $\Omega = F(\hat{\Omega})$, where F is a map, defined as in Theorem 2.3. If a function $u : \Omega \rightarrow \mathbb{R}$ belongs to the space $H^l(\Omega)$ for some integer $l \geq 0$, the function $\hat{u} = u \circ F : \hat{\Omega} \rightarrow \mathbb{R}^n$ belongs to the space $H^l(\hat{\Omega})$ and, in addition,*

$$|\hat{u}|_{1,\hat{\Omega}} \leq |J^{-1}|_{0,\infty}^{1/2} \|F\|_{1,\infty} |u|_{1,\Omega} \quad \forall u \in H^1(\Omega) \quad (2.18)$$

$$|\hat{u}|_{2,\hat{\Omega}} \leq C |J^{-1}|_{0,\infty}^{1/2} (\|F\|_{1,\infty}^2 |u|_{2,\Omega} + \|F\|_{2,\infty} |u|_{1,\Omega}) \quad \forall u \in H^2(\Omega) \quad (2.19)$$

$$|\hat{u}|_{3,\hat{\Omega}} \leq C |J^{-1}|_{0,\infty}^{1/2} \{ \|F\|_{1,\infty}^3 |u|_{3,\Omega} + \|F\|_{1,\infty} \|F\|_{2,\infty} |u|_{2,\Omega} + \|F\|_{3,\infty} |u|_{1,\Omega} \} \quad \forall u \in H^3(\Omega) \quad (2.20)$$

Proof

For any multi-index α with $|\alpha| = m$, one has

$$D^\alpha \hat{u}(\xi) = D^m \hat{u}(\xi)(e_{\alpha_1}, e_{\alpha_2}, \dots, e_{\alpha_m}).$$

Setting $\alpha_1 = \alpha_2 = \dots = \alpha_m = i$, we have

$$|D^m \hat{u}(\boldsymbol{\xi})(\mathbf{e}_i)^m| \leq \max_{1 \leq i \leq n} |D^m \hat{u}(\boldsymbol{\xi})(\mathbf{e}_i)^m| = [D^m \hat{u}(\boldsymbol{\xi})]. \quad (2.21)$$

By using this result we prove the following: since $\hat{u} = u \circ F_e$, by using the differentiation rule for composition of functions we have

$$D\hat{u}(\boldsymbol{\xi})(\boldsymbol{\eta}) = Du(\mathbf{x}) \circ DF_e(\boldsymbol{\xi})(\boldsymbol{\eta}) \quad \text{for all } \boldsymbol{\xi} = F_e^{-1}(\mathbf{x}) \text{ and } \boldsymbol{\eta} \in \mathbb{R}^n,$$

and thus

$$[D\hat{u}(\boldsymbol{\xi})] \leq [DF_e(\boldsymbol{\xi})] \|Du(F_e(\boldsymbol{\xi}))\|.$$

Consequently,

$$\begin{aligned} \int_{\hat{\Omega}} [D\hat{u}(\boldsymbol{\xi})]^2 d\boldsymbol{\xi} &\leq \int_{\hat{\Omega}} [DF_e(\boldsymbol{\xi})]^2 \|Du(F_e(\boldsymbol{\xi}))\|^2 d\boldsymbol{\xi} \\ &= \|F_e\|_{1,\infty}^2 \int_{\Omega} |J^{-1}| \|Du(\mathbf{x})\|^2 d\mathbf{x} \\ &\leq \|F_e\|_{1,\infty}^2 |J^{-1}|_{0,\infty} \int_{\Omega} \|Du(\mathbf{x})\|^2 d\mathbf{x}, \end{aligned}$$

from which we deduce the inequality (2.18), using the result (2.3) and the inequality (2.21).

By using Taylor expansions as in [13, Section 7.5], we have

$$D^2 \hat{u}(\boldsymbol{\xi})(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2) = Du(\mathbf{x})(D^2 F_e(\boldsymbol{\xi})(\boldsymbol{\eta}_1, \boldsymbol{\eta}_2)) + D^2 u(\mathbf{x})(DF_e(\boldsymbol{\xi})\boldsymbol{\eta}_1, DF_e(\boldsymbol{\xi})\boldsymbol{\eta}_2)$$

for all $\boldsymbol{\xi} \in \hat{\Omega}$ and $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2 \in \mathbb{R}^n$. Since $\hat{u} = u \circ F_e$, we obtain

$$\begin{aligned} [D^2 \hat{u}(\boldsymbol{\xi})] &= \max_{1 \leq i \leq n} \|D^2 \hat{u}(\boldsymbol{\xi})(\mathbf{e}_i)^2\| \\ &= \max_{1 \leq i \leq n} \|Du(\mathbf{x})(D^2 F_e(\boldsymbol{\xi})(\mathbf{e}_i)^2) + D^2 u(\mathbf{x})(DF_e(\boldsymbol{\xi})\mathbf{e}_i, DF_e(\boldsymbol{\xi})\mathbf{e}_i)\| \\ &\leq [D^2 F_e(\boldsymbol{\xi})] \|Du(\mathbf{x})\| + [DF_e(\boldsymbol{\xi})]^2 \|D^2 u(\mathbf{x})\|. \end{aligned}$$

Therefore

$$\begin{aligned}
\| \hat{u} \|_{2, \hat{\Omega}} &= \left(\int_{\hat{\Omega}} \sum_{i=1}^n | D^2 \hat{u}(\xi)(e_i)^2 |^2 d\xi \right)^{\frac{1}{2}} \\
&\leq C_1 \left(\int_{\hat{\Omega}} [D^2 \hat{u}(\xi)]^2 d\xi \right)^{\frac{1}{2}} \quad (\text{by (2.21)}) \\
&\leq C_1 \left\{ \| F_e \|_{2, \infty} \left(\int_{\hat{\Omega}} \| Du(\mathbf{x}) \|^2 d\xi \right)^{\frac{1}{2}} + \| F_e \|_{1, \infty}^2 \left(\int_{\hat{\Omega}} \| D^2 u(\mathbf{x}) \|^2 d\xi \right)^{\frac{1}{2}} \right\}.
\end{aligned}$$

By application of the formula for change of variables in multiple integrals and by the use of inequality (2.3), we derive the inequality (2.19) .

Finally, we use the Taylor expansion again, for all $\xi \in \hat{\Omega}$ and $\eta_1, \eta_2, \eta_3 \in \mathbb{R}^n$, to obtain

$$\begin{aligned}
D^3 \hat{u}(\xi)(\eta_1, \eta_2, \eta_3) &= Du(\mathbf{x})(D^3 F_e(\xi)(\eta_1, \eta_2, \eta_3) + \\
&\quad D^2 u(\mathbf{x})(DF_e(\xi)\eta_1, D^2 F_e(\xi)(\eta_2, \eta_3) + \\
&\quad D^3 u(\mathbf{x})(DF_e(\xi)\eta_1, DF_e(\xi)\eta_2, DF_e(\xi)\eta_3)
\end{aligned}$$

to derive the inequality (2.20). The remaining details are omitted. \square

By manipulating the above theorems we obtain an estimate in the following general form, according to [16, Theorem 2]:

Theorem 2.5 *For some integers $k \geq 0$ and $m \geq 0$ let $H^{k+1}(\hat{\Omega})$ and $H^m(\hat{\Omega})$ satisfy the inclusions*

$$H^{k+1}(\hat{\Omega}) \hookrightarrow C^0(\hat{\Omega}), H^{k+1}(\hat{\Omega}) \hookrightarrow H^m(\hat{\Omega}),$$

and let $\hat{\Pi} \in \mathcal{L}(H^{k+1}(\hat{\Omega}); H^m(\hat{\Omega}))$ be a mapping such that

$$\hat{\Pi} \hat{q} = \hat{q} \quad \text{for all } \hat{q} \in Q_k(\hat{\Omega}) .$$

For any open set Ω which is the image of set $\hat{\Omega}$ under the map F which is as defined in Theorem 2.3, let the mapping $\Pi \in \mathcal{L}(H^{k+1}(\Omega); H^m(\Omega))$ be defined by

$$\widehat{\Pi}u = \hat{\Pi}\hat{u} \quad \text{for all } u \in H^{k+1}(\Omega).$$

Then there exists a constant $C(\hat{\Pi}, \hat{\Omega})$ such that

$$\begin{aligned} \|u - \Pi u\|_{m,\Omega} &\leq C \left(\frac{\sup_{\xi \in \hat{\Omega}} |J(\xi)|}{\inf_{\xi \in \hat{\Omega}} |J(\xi)|} \right)^{\frac{1}{2}} \\ &\times \left(\sum_{l=1}^m \sum_{i \in I(l,m)} |F^{-1}|_{1,\infty}^{i_1} |F^{-1}|_{2,\infty}^{i_2} \cdots |F^{-1}|_{m,\infty}^{i_m} \right) \\ &\times \left(\sum_{l=1}^{k+1} |u|_{l,\Omega} \sum_{i' \in I(l,k+1)} \|F\|_{1,\infty}^{i'_1} \|F\|_{2,\infty}^{i'_2} \cdots \|F\|_{k+1,\infty}^{i'_{k+1}} \right), \end{aligned} \quad (2.22)$$

where $I(l, m) = \{\mathbf{i} = (i_1, i_2, \dots, i_m) \in \mathcal{N}^m; i_1 + i_2 + \dots + i_m = l, i_1 + 2i_2 + \dots + mi_m = m\}$, $1 \leq l \leq m$ and \mathcal{N} is the set of natural numbers.

Proof

The proof can be found in [16]. □

For special cases in which $0 < k \leq 2$ and $0 < m \leq 3$, the above inequality can be established by the application of Theorems 2.2 – 2.4. For example, when $m = 1$ and $k = 1$ the estimate (2.22) becomes

$$\|u - \Pi u\|_{1,\Omega} \leq C \left(\frac{\sup_{\xi \in \hat{\Omega}} |J(\xi)|}{\inf_{\xi \in \hat{\Omega}} |J(\xi)|} \right)^{\frac{1}{2}} |F^{-1}|_{1,\infty} \left(\|F\|_{1,\infty}^2 |u|_{2,\Omega_e} + \|F\|_{2,\infty} |u|_{1,\Omega_e} \right). \quad (2.23)$$

To obtain more concrete estimates we need to find bounds for the various quantities which appear on the righthand sides of (2.22) and (2.23). Such bounds are available for simplicial elements [18]. In [16], a bound for the Jacobian was obtained for quadrilateral elements by interpolating it. Without interpolating the Jacobian, we present an alternative approach in Section 2.4 to obtain the bounds.

Since we are interested in seminorms of matrix-valued functions, Theorem 2.4 can be restated using such functions in a straightforward way from the original Theorem. Here and henceforth we denote by $[P]_{sym}^{n \times n}$ the space of all symmetric matrix-valued functions whose components belong to P .

Theorem 2.6 *Let $\Omega, \hat{\Omega}$ be as defined in the Theorem 2.4, and let ε be a matrix-valued function which belongs to $[H^l(\Omega)]_{sym}^{n \times n}$. Then the function $\hat{\varepsilon}$ defined by $\hat{\varepsilon} = \varepsilon \circ F$ belongs to $[H^l(\hat{\Omega})]_{sym}^{n \times n}$. In addition there exists a constant C such that*

$$[\hat{\varepsilon}]_{0, \hat{\Omega}} \leq |J^{-1}|_{0, \infty}^{1/2} |\varepsilon|_{0, \Omega} \quad \forall \varepsilon \in [L_2(\Omega)]_{sym}^{n \times n}, \quad (2.24)$$

$$[\hat{\varepsilon}]_{1, \hat{\Omega}} \leq |J^{-1}|_{0, \infty}^{1/2} \|F\|_{1, \infty} |\varepsilon|_{1, \Omega} \quad \forall \varepsilon \in [H^1(\Omega)]_{sym}^{n \times n}, \quad (2.25)$$

$$[\hat{\varepsilon}]_{2, \hat{\Omega}} \leq C |J^{-1}|_{0, \infty}^{1/2} (\|F\|_{1, \infty}^2 |\varepsilon|_{2, \Omega} + \|F\|_{2, \infty} |\varepsilon|_{1, \Omega}) \\ \forall \varepsilon \in [H^2(\Omega)]_{sym}^{n \times n}. \quad (2.26)$$

We also record for later use a version of the Bramble-Hilbert lemma [10], in a form which is applicable to matrix-valued functions.

Lemma 2.1 *Suppose that Ω is a bounded Lipschitz domain in \mathbb{R}^n , and suppose that the functional $L : H^{k+1}(\Omega)^{n \times n} \rightarrow \mathbb{R}$ ($k \geq 0$) is bounded and linear, and has the property that*

$$L(\mathbf{q}_k) = 0 \quad \text{for any } \mathbf{q}_k \in [Q_k(\Omega)]^{n \times n},$$

where $Q_k(\Omega)$ denotes the space of polynomials of the form (2.6). Then there exists a constant $C > 0$ such that

$$|L(\varepsilon)| \leq C \|L\|_{k+1}^* [\varepsilon]_{k+1} \quad \text{for all } \varepsilon \in [H^{k+1}(\Omega)]_{sym}^{n \times n}, \quad (2.27)$$

where $\|\cdot\|_m^*$ denotes the norm on the topological dual $[H^m(\Omega)]_{sym}^{n \times n}$ of $[H^m(\Omega)]_{sym}^{n \times n}$.

Remark.

Lemma 2.1 in its standard form applies to the case of linear functionals which are invariant with respect to the space P_k rather than Q_k (see [18], Theorem 4.1.3). The modified version presented here may be obtained by observing that the seminorm $[\cdot]_k$ is a norm on the quotient space $H^{k+1}(\Omega)/Q_k(\Omega)$, equivalent to the standard quotient norm [10].

2.3 Finite element interpolation and approximation

A general theory of finite element interpolation has been developed by Ciarlet and Raviart [16] and further generalised by Ciarlet [17, 18]. Ciarlet [18] gives a comprehensive account of the method for elliptic problems. Reddy [43] gives an introductory treatment. See also Oden and Carey [8, 12], Bathe [7], Zienkiewicz [63] and Hughes [28] for practical aspects of the finite element method.

The foundation of Galerkin finite element approximations is the weak or variational formulation of a problem. A simple example is the following: find a function $u \in V$, where V is a Hilbert space, such that

$$a(u, v) = l(v) \quad \text{for all } v \in V, \tag{2.28}$$

where $a : V \times V \rightarrow \mathbb{R}$ is a bilinear form and $l : V \rightarrow \mathbb{R}$ a linear functional.

The space V is called the space of admissible functions. In order to have existence and uniqueness of a solution to the above problem the bilinear form $a(\cdot, \cdot)$ and linear functional $l(\cdot)$ must be well-defined for all functions $v \in V$. In other words, the bilinear form $a(\cdot, \cdot)$ and linear functional $l(\cdot)$ should be continuous and $a(\cdot, \cdot)$ V -elliptic to have a unique solution (see Strang [54]).

To obtain the Galerkin approximation u_h to u we choose a finite dimensional subspace $V^h \subset V$ and define $u_h \in V^h$ to be the solution of the discrete problem

$$a(u_h, v_h) = l(v_h) \quad \text{for all } v_h \in V^h.$$

If $\{\phi_i\}_{i=1}^N$ denotes a basis for V^h , then the approximate solution u_h can be expressed in the form

$$u_h = \sum_{i=1}^N a_i \phi_i. \tag{2.29}$$

The coefficients a_i are determined on substituting the expression (2.29) in the above discrete problem and solving the resulting linear algebraic system of equations.

The practical implementation of the Galerkin method is crucially dependent on the construction of a suitable basis for the space V^h . The finite element method provides a systematic method for constructing suitable basis functions ϕ_i for V^h . Such construction relies basically on the following:

- discretization of Ω into subdomains Ω_e ($e = 1, \dots, E$),
- local interpolation over each Ω_e ,
- global interpolation over Ω .

The domain Ω is discretized into subdomains $\Omega_1, \Omega_2, \dots, \Omega_E$ which are non-overlapping and cover Ω , in the sense that

$$\Omega_e \cap \Omega_f = \emptyset \quad \text{for } e \neq f, \text{ and } \cup_{e=1}^E \bar{\Omega}_e = \bar{\Omega}.$$

The diameter h_e of the finite element Ω_e is defined by

$$h_e = \max\{|\mathbf{x} - \mathbf{y}| : \mathbf{x}, \mathbf{y} \in \Omega_e\},$$

and the mesh size defined by $h = \max_e h_e$.

Within each element, certain points, called *nodes* or *nodal points*, are identified: these play an important role in constructing the basis functions.

For a given discretization, we define a finite dimensional space \mathcal{P}_e on Ω_e to be the span of linearly independent local interpolation functions $\{N_i^e\}_{i=1}^{n_e}$, which are typically polynomials or *close to polynomials*; that is, $P_k(\bar{\Omega}_e) \subset \mathcal{P}_e$, where $P_k(\bar{\Omega}_e)$ is the space of polynomials of degree $\leq k$, or $Q_k(\bar{\Omega}_e) \subset P_l(\bar{\Omega}_e)$, where Q_k is the space of polynomial defined in (2.6). Locally, we approximate the restriction $u|_{\Omega_e}$ of u , denoted by u_h^e , by linear combinations of the form

$$u_h^e = \sum_{i=1}^{N_e} a_i^e N_i^e; \tag{2.30}$$

the coefficients a_i^e are the *local degrees of freedom* on element $\bar{\Omega}_e$ and usually taken to be the values of u and the values of various partial derivatives of u at the preassigned nodes within the element $\bar{\Omega}_e$. Interpolation involving only function values is usually called *Lagrange interpolation*, while if derivatives are also involved the allusion is to *Hermite interpolation*. Here we restrict ourselves to Lagrange finite elements only.

By appropriate choice of the element geometry, the location of nodes, degrees of freedom, and of local interpolation functions N_i^e , the global basis functions ϕ_i can be constructed so that they have *local support*. By this we mean that the functions ϕ_i vanish outside a subregion of Ω which is adjacent to the i th nodal point. Furthermore, these basis functions can be easily constructed from local basis functions N_i^e .

In view of the properties outlined above, the formal definition, according to [18], of a *finite element* is that of a triple $(\bar{\Omega}_e, D_e, \mathcal{P}_e)$, where

- (i) Ω_e is a nonempty closed subset of Ω with Lipschitzian boundary $\partial\Omega_e$;
- (ii) \mathcal{D}_e is a finite set of linear functionals l_i , $1 \leq i \leq N_e$, called the degrees of freedom of the element;
- (iii) \mathcal{P}_e is a space of functions defined on Ω_e , $\mathcal{P}_e \subset C^\infty(\Omega_e)$, such that for any real scalars α_i , $1 \leq i \leq N_e$, there exists a unique v_h such that

$$l_i(v_h) = \alpha_i, \quad 1 \leq i \leq N_e.$$

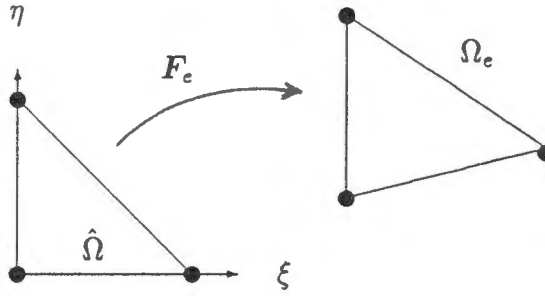
Within the constraints mentioned above, a variety of acceptable elements have been developed. Further details may be found in standard texts on finite element methods, for examples [7, 18, 36].

2.3.1 Equivalent families of finite elements

In one-dimensional problems, the domain is discretized into line elements connected at nodal points at their ends. For two- and three-dimensional cases the discretization is less straightforward. The question arises as how to carry out the discretization. We have to choose a procedure that will be general enough to model irregular domains, but which consists of elements simple enough for computational purposes. It is convenient to carry out interpolation on a simple master or reference element $\hat{\Omega}$, and to generate details of the set $\{\Omega_e\}_{e=1}^E$ via a family of maps from $\hat{\Omega}$. We describe this idea next.

Affine equivalent families of elements

Two finite elements $(\hat{\Omega}, \hat{\mathcal{D}}, \hat{\mathcal{P}})$ and $(\Omega_e, \mathcal{D}_e, \mathcal{P}_e)$ of the same type are equivalent if



$$F_e(\xi) = B\xi + c$$

Figure 2.1: Affine map of a triangle onto a triangle

there exists a unique invertible map F_e mapping points $\xi \in \hat{\Omega}$ onto points $x \in \Omega_e$, such that

$$\Omega_e = F_e(\hat{\Omega}) = (F_{e,1}(\hat{\Omega}), F_{e,2}(\hat{\Omega}), \dots, F_{e,n}(\hat{\Omega})) \subset \mathbb{R}^n,$$

$$\mathcal{P}_e = \{p : \Omega_e \rightarrow \mathbb{R}; p = \hat{p} \circ F_e^{-1}, \hat{p} \in \hat{\mathcal{P}}\},$$

$$\mathcal{D}_e = \{p\{F_e(\xi_i)\}, 1 \leq i \leq \hat{m} = m\}.$$

When F_e is an affine map, we say that two elements are *affine-equivalent*. The mapping from a triangle onto a triangle (see Figure 2.1), a square onto a parallelogram and a tetrahedron onto a tetrahedron are some examples of affine maps.

Isoparametric families of elements

If $F_e(\xi) = \{F_{e,j}(\xi)\}^n \in \mathbb{R}^n$ such that $F_{e,j} \in \hat{\mathcal{P}}$ for all $1 \leq j \leq n$, that is, if the same functions are chosen for obtaining the maps F_e as for the basis functions, then the element $(\Omega_e, \mathcal{D}_e, \mathcal{P}_e)$ is called an *isoparametric element*, and the finite element $(\Omega_e, \mathcal{D}_e, \mathcal{P}_e)$ is said to be isoparametrically equivalent to the reference element $(\hat{\Omega}, \hat{\mathcal{D}}, \hat{\mathcal{P}})$.

Isoparametric elements are more attractive from a practical standpoint as the mapping functions are available as part of the interpolation process and need not be

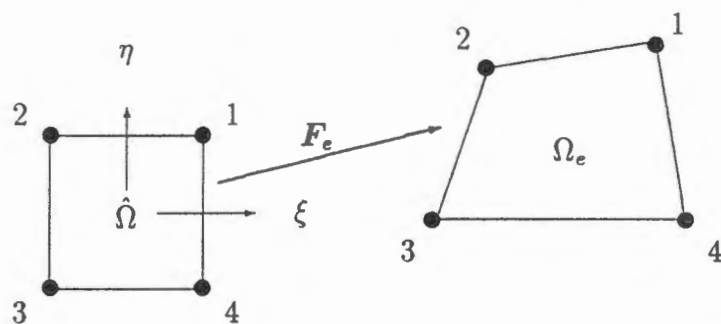


Figure 2.2: Bilinear isoparametric map of a square onto a quadrilateral (type (1))

generated separately. Furthermore, they are useful for interpolation over domains with curved boundaries.

It should be noted that an isoparametric map is invertible, but does not preserve polynomials; thus the inverse map F_e^{-1} is not polynomial even when F_e is polynomial. However, this does not affect the process because all the computations are performed on the reference element $\hat{\Omega}$ rather than on Ω_e . All that is needed is a knowledge of the mapping F_e .

For future reference, since much of the developments will involve quadrilateral and hexahedral elements, we describe some examples of these elements.

Quadrilateral elements. Finite elements with four sides are generally referred to as quadrilateral elements, whether or not the sides are straight. These quadrilaterals are isoparametrically equivalent to a square, as shown in Figures 2.2 and 2.3. An isoparametric quadrilateral with straight sides (quadrilateral of type (1)) is obtained by using four nodes, one at each vertex. The isoparametric bilinear map $F_e(\xi) \in (Q_1(\hat{\Omega}))^2$ is given by

$$F_e : \xi \in \hat{\Omega} \rightarrow \mathbf{x} = \sum_{A=1}^4 \mathbf{x}_A \hat{N}_A \in \Omega_e, \quad (2.31)$$

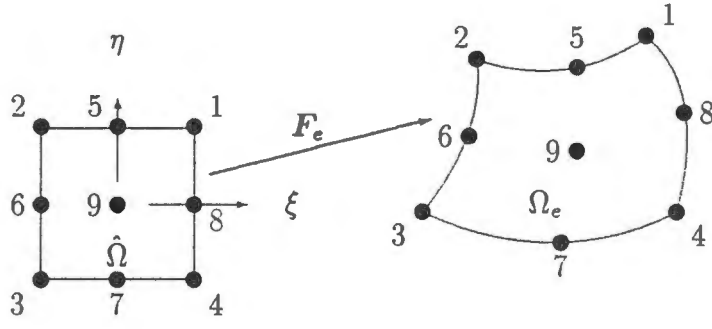


Figure 2.3: Biquadratic isoparametric map of a square onto a quadrilateral (type (2))

where \mathbf{x}_A is the position vector of node A in Ω_e , and the local basis functions N_A are given by

$$\hat{N}_A(\xi, \eta) = \frac{1}{4}(1 + \xi_A \xi)(1 + \eta_A \eta); \quad A = 1, \dots, 4.$$

Here ξ_A is the position vector of node A in $\hat{\Omega}$.

In the case of biquadratic Lagrange quadrilateral elements (Figure 2.3), the sides may be curved. Such elements have corner nodes, a node in each side and a node at the centroid, for a total of 9 nodes. The biquadratic isoparametric transformation is given by

$$F_e : \xi \in \hat{\Omega} \rightarrow \mathbf{x} = \sum_{A=1}^9 \mathbf{x}_A \hat{N}_A \in \Omega_e, \quad (2.32)$$

where

$$\begin{aligned} \hat{N}_1 &= \frac{1}{4}(1 - \xi)(1 - \eta) - \frac{1}{2}(P_8 + P_5) - \frac{1}{4}\hat{N}_9, & \hat{N}_5 &= P_5 - \frac{1}{2}\hat{N}_9, \\ \hat{N}_2 &= \frac{1}{4}(1 + \xi)(1 - \eta) - \frac{1}{2}(P_5 + P_6) - \frac{1}{4}\hat{N}_9, & \hat{N}_6 &= P_6 - \frac{1}{2}\hat{N}_9, \\ \hat{N}_3 &= \frac{1}{4}(1 + \xi)(1 + \eta) - \frac{1}{2}(P_5 + P_6) - \frac{1}{4}\hat{N}_9, & \hat{N}_7 &= P_7 - \frac{1}{2}\hat{N}_9, \\ \hat{N}_4 &= \frac{1}{4}(1 - \xi)(1 + \eta) - \frac{1}{2}(P_5 + P_6) - \frac{1}{4}\hat{N}_9, & \hat{N}_8 &= P_8 - \frac{1}{2}\hat{N}_9, \\ \hat{N}_9 &= (1 - \xi^2)(1 - \eta^2), \end{aligned}$$

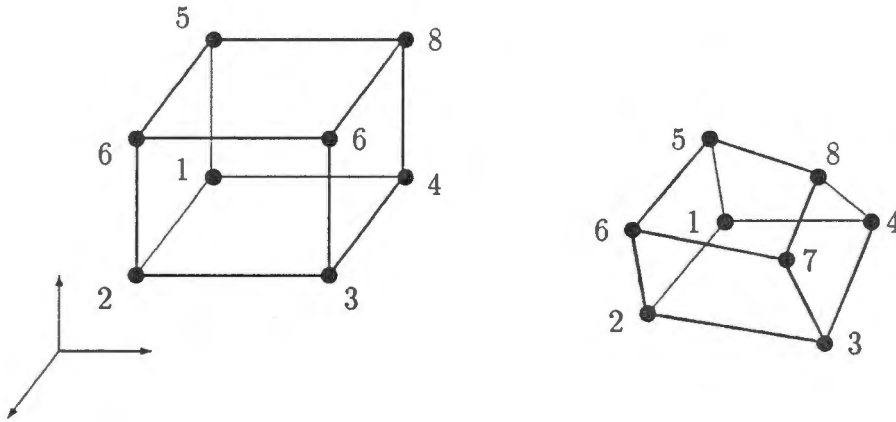


Figure 2.4: Trilinear isoparametric map of a cube onto a hexahedron (type (1))

where $P_5 \cdots P_8$ are given by

$$\begin{aligned} P_5 &= \frac{1}{2}(1 - \xi^2)(1 - \eta) & P_6 &= \frac{1}{2}(1 + \xi)(1 - \eta^2) \\ P_7 &= \frac{1}{2}(1 - \xi^2)(1 + \eta) & P_8 &= \frac{1}{2}(1 - \xi)(1 - \eta^2) \end{aligned}$$

Hexahedral elements. These are three-dimensional elements. Most three-dimensional elements are direct generalizations of elements defined in two dimensions. A hexahedral element has six plane or curved faces, and each face is a quadrilateral with straight or curved sides. The element is isoparametrically equivalent to a cube, as shown in Figure 2.4. An isoparametric hexahedral element with plane faces (type (1)) is obtained by using eight nodes, one at each vertex, and the mapping $F_e(\xi) \in (Q_1(\hat{\Omega}))^3$ is given by

$$F_e : \xi \in \hat{\Omega} \rightarrow \mathbf{x} = \sum_{A=1}^8 \mathbf{x}_A \hat{N}_A \in \Omega_e, \quad (2.33)$$

where the local basis functions \hat{N}_A are given by

$$\hat{N}_A(\xi, \eta, \nu) = \frac{1}{8}(1 + \xi_A \xi)(1 + \eta_A \eta)(1 + \nu_A \nu); \quad A = 1, \dots, 8.$$

The sides and edges of a triquadratic Lagrange solid element may be curved; this element has corner nodes, a node on each edge, midface nodes and a node at the

centroid of the element, for a total of 27 nodes. Basis functions for these elements may be found in [7].

In conclusion, isoparametric elements are very useful particularly for domains with curved boundaries. However, they should be used with caution. Care must be taken to ensure that the Jacobian of the isoparametric map is strictly positive. We avoid this degeneracy by enforcing conditions, which can be found in the following Sections.

2.4 The equivalent parallelogram of a quadrilateral

In general, a finite element solution will be an approximation to the true solution. How close this solution is to the exact solution and whether or not it converges to the exact solution are both important questions. Such analyses are carried out for a wide range of elements in [18, 36]. For convenience, we drop the subscript e denoting a generic element.

To obtain a more concrete form of the error estimates we need to estimate of various norms of geometrical parameters such as F , J etc. appearing in (2.22). Interpolation error estimates for curved elements were first obtained by Ciarlet and Raviart [16], using the notion of the isoparametric element. In that work, the required geometrical estimates for triangles and quadrilaterals were obtained. The case of the four-noded quadrilateral differs from that of triangular elements in an important sense, namely, that whereas curved triangular elements are conveniently regarded as a perturbation of a straight sided triangle, it is not immediately clear how this approach may be carried over to quadrilaterals, especially if the perturbations are to be regarded as small in some sense. This is commented on also in the text by Ciarlet [18, see page 243]. It is therefore necessary to adopt an approach which differs from that employed in the case of curved triangular elements, if the idea of the quad-

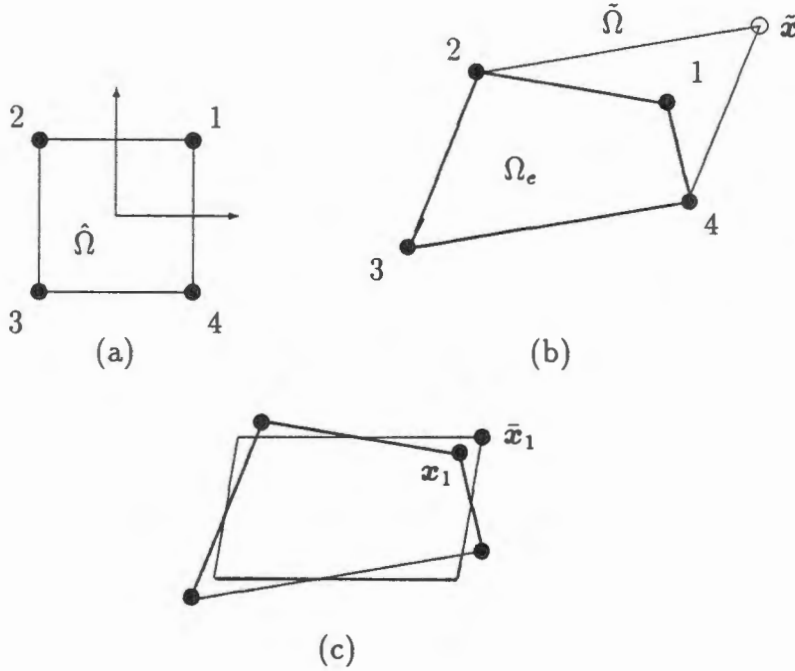


Figure 2.5: How the equivalent parallelogram is constructed

rilateral as a perturbation of a parallelogram is to obtain the required estimates. We first identify such a unique parallelogram associated with a quadrilateral. This parallelogram is called the *equivalent parallelogram*.

Let Ω be a four-noded quadrilateral element of arbitrary shape, and let $\hat{\Omega}$ be the reference element which is mapped to Ω by F . Here F is defined by (2.31). To obtain the equivalent parallelogram we first construct four parallelograms $\tilde{\Omega}_A$ ($A = 1, \dots, 4$), which are generated by affine maps \tilde{F}_A from $\hat{\Omega}$. The parallelogram $\tilde{\Omega}_A$ is constructed by defining as its base the two sides of Ω emanating from the node opposite x_A . Figure 2.5 shows how the parallelogram $\tilde{\Omega}_1$ is constructed. In detail, these four affine maps are

$$\begin{aligned} \tilde{F}_1(\xi) &= \frac{1}{2}(\mathbf{x}_2 + \mathbf{x}_4) + \frac{1}{2}\xi(\mathbf{x}_4 - \mathbf{x}_3) + \frac{1}{2}\eta(\mathbf{x}_2 - \mathbf{x}_3), \\ \tilde{F}_2(\xi) &= \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_3) + \frac{1}{2}\xi(\mathbf{x}_4 - \mathbf{x}_3) + \frac{1}{2}\eta(\mathbf{x}_1 - \mathbf{x}_4), \\ \tilde{F}_3(\xi) &= \frac{1}{2}(\mathbf{x}_2 + \mathbf{x}_4) + \frac{1}{2}\xi(\mathbf{x}_1 - \mathbf{x}_2) + \frac{1}{2}\eta(\mathbf{x}_1 - \mathbf{x}_4), \\ \tilde{F}_4(\xi) &= \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_3) + \frac{1}{2}\xi(\mathbf{x}_1 - \mathbf{x}_2) + \frac{1}{2}\eta(\mathbf{x}_2 - \mathbf{x}_3). \end{aligned}$$

The equivalent parallelogram $\tilde{\Omega}$ (see Fig 2.5 (c)) associated with Ω is now defined

to be the parallelogram generated by the affine map

$$\bar{F} \equiv \frac{1}{4} \sum_{A=1}^4 \tilde{F}_A,$$

with corner nodes

$$\bar{x}_A = \bar{F}(\xi_A) = \frac{1}{4} \sum_{B=1}^4 \tilde{F}_B(\xi_A).$$

In particular,

$$\bar{x}_A = \frac{3}{4} \mathbf{x}_A + \frac{1}{4} (\mathbf{x}_{A+1} - \mathbf{x}_{A+2} + \mathbf{x}_{A+3}) \pmod{4}, \quad (2.34)$$

so, for example,

$$\bar{x}_1 = \frac{3}{4} \mathbf{x}_1 + \frac{1}{4} (\mathbf{x}_2 - \mathbf{x}_3 + \mathbf{x}_4).$$

Thus we see that the node \bar{x}_A , which determines the equivalent parallelogram, lies on the line joined by \tilde{x}_A and \mathbf{x}_A at the ratio 3 : 1. Furthermore, we deduce that

$$(\mathbf{x}_1 - \bar{x}_1) = -(\mathbf{x}_2 - \bar{x}_2) = (\mathbf{x}_3 - \bar{x}_3) = -(\mathbf{x}_4 - \bar{x}_4) = \mathbf{k}, \quad (2.35)$$

so that

$$|\mathbf{x}_A - \bar{x}_A| = |\mathbf{k}|, \quad (2.36)$$

where

$$\mathbf{k} = \frac{1}{4} (\mathbf{x}_1 - \mathbf{x}_2 + \mathbf{x}_3 - \mathbf{x}_4). \quad (2.37)$$

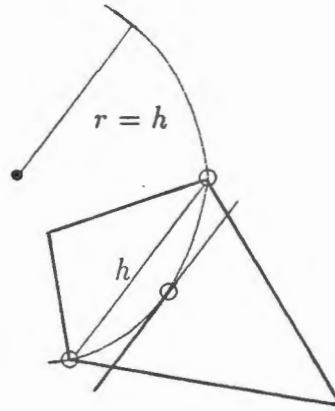


Figure 2.6: A properties of quadrilaterals: O is the centroid.

Now the isoparametric map F can be written in the form

$$\begin{aligned}
 F(\xi) &= \sum_{A=1}^4 \bar{x}_A N_A(\xi) + \sum_{A=1}^4 (x_A - \bar{x}_A) N_A(\xi) \\
 &= \underbrace{\sum_{A=1}^4 \bar{x}_A N_A(\xi)}_{B\xi+c} + \underbrace{k(N_1 - N_2 + N_3 - N_4)}_{k\xi\eta}. \tag{2.38}
 \end{aligned}$$

From (2.38) we see that it is in fact possible to define the equivalent parallelogram in a simpler way, that is, as the image of $\hat{\Omega}$ under the *affine part only* of the map F : the first term on the righthand side ($B\xi+c$) corresponds to the equivalent parallelogram while the second term is the same as the nonlinear part of the map F . We have chosen to present the details in a way which is perhaps less straightforward, but which sheds greater light on the underlying geometry.

Remark

The equivalent parallelogram has some interesting properties:

- From (2.34), it can be shown that the lengths of the corresponding diagonals of the quadrilateral and the equivalent parallelogram are equal and those diagonals are parallel.
- From (2.35) and (2.36), the actual quadrilateral is obtained by translating each node of the equivalent parallelogram by $\pm k$. The magnitude of k is equal to

the distance between the midpoints of the diagonals of the quadrilateral; this magnitude is comparatively smaller than any diagonal of the quadrilateral.

- The corresponding sides of the quadrilateral and the equivalent parallelogram intersect at their midpoints: this is evident in Figure 2.5 (c), and easily proved with the aid of elementary vector algebra.
- The radius of the circle passing through the centroid of the actual quadrilateral and the ends of any diagonal is equal in magnitude to the corresponding diagonal (see Figure 2.6). Note also that the centroid of the actual quadrilateral coincides with that of the associated equivalent parallelogram.

These properties are very useful in establishing the estimates required for interpolation theory.

2.5 Interpolation error estimates for quadrilateral elements

We now obtain estimates for $|J|_{0,\infty}$, $|F|_{m,\infty}$, and so on, by exploiting the notion of the equivalent parallelogram. We continue to omit the subscript e denoting the generic element.

To begin with, we first note from the representation (2.38) that

$$\frac{\partial F}{\partial \xi} = B + \mathbf{k} \otimes \mathbf{l} = B + \mathbf{k} \mathbf{l}^T \quad (2.39)$$

where $\mathbf{l} = M \boldsymbol{\xi}$ and

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\xi} = \begin{pmatrix} \xi \\ \eta \end{pmatrix}.$$

Furthermore,

$$\frac{\partial^2 F}{\partial \xi \partial \xi} = \mathbf{k} \otimes M, \quad \text{that is, } \frac{\partial^2 F_i}{\partial \xi_j \partial \xi_k} = k_i M_{jk}.$$

We next give a definition of a *regular family* of quadrilateral finite elements, based on the concept of the equivalent parallelogram.

A family of four-noded quadrilaterals is said to be regular if

- the matrix B is invertible with $\det B > 0$;
- a constant α exists such that the inequality

$$B^{-1} \mathbf{k} \cdot \mathbf{l} \geq \alpha > -1 \tag{2.40}$$

holds for each element.

In [16] a rather different definition of a regular family is given (see also [18, page 247]): there a family of four-noded quadrilaterals is said to be regular if all quadrilaterals are convex, and if there exist constants σ and ν such that

$$h/h' \leq \sigma \quad \text{and} \quad \nu_e \leq \nu < 1 \tag{2.41}$$

holds for each element, where h and h' are respectively the diameter of Ω_e and the length its shortest side, and ν_e is the maximum of the cosines of the interior angles. It is not a straightforward matter to show that the definition (2.41) implies, or is equivalent to, that given by (2.40). However, we show by means of two examples how (2.40) works in practice.

Example 1. This is an example of a family of quadrilaterals parameterised by a scalar d which determines the length of one of the sides (see Figure 2.7). Clearly we expect problems when $d \rightarrow 0$, and indeed when $d < 0$. We determine the equivalent

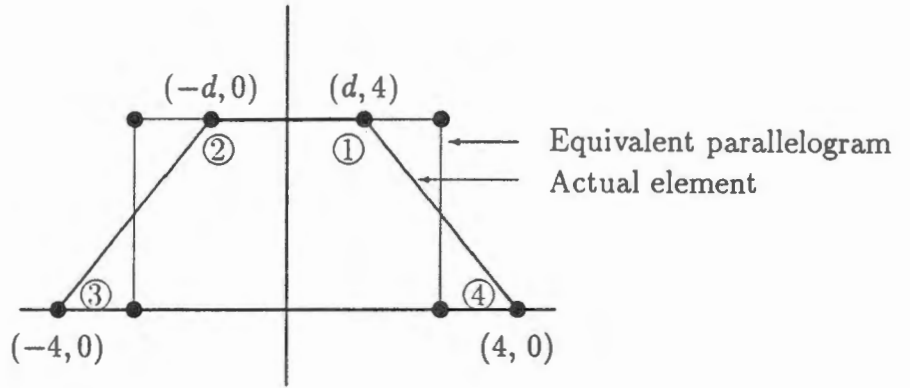


Figure 2.7: The element of *Example 1*

parallelogram by finding \mathbf{k} : from (2.37) we have

$$\mathbf{k} = \frac{1}{4}(\mathbf{x}_1 - \mathbf{x}_2 + \mathbf{x}_3 - \mathbf{x}_4) = \frac{1}{2}(d - 4)\mathbf{e}_1,$$

The nodal points of the equivalent parallelogram are found from (2.36), and are shown in Figure 2.7. The transformation matrix \mathbf{B} is diagonal, and is given by

$$\mathbf{B} = \begin{pmatrix} \frac{4+d}{2} & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow \mathbf{B}^{-1} = \begin{pmatrix} \frac{2}{4+d} & 0 \\ 0 & \frac{1}{2} \end{pmatrix},$$

so that

$$\mathbf{B}^{-1}\mathbf{k} \cdot \mathbf{l} = \frac{d-4}{d+4}\eta.$$

It is sufficient to consider the two cases $\eta = \pm 1$. For $\eta = 1$ we find that

$$\mathbf{B}^{-1}\mathbf{k} \cdot \mathbf{l} > -1 \quad \text{if and only if} \quad d > 0.$$

This guarantees that the quadrilateral is convex and nondegenerate, or equivalently

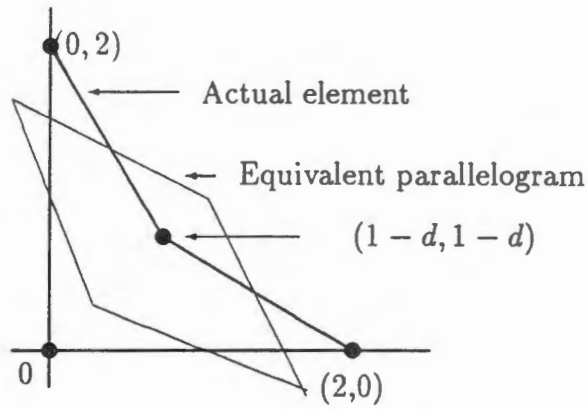


Figure 2.8: The element of *Example 2*

that the constants σ and ν exist. The case $\eta = -1$ does not add anything.

Example 2. Again we have a family of quadrilaterals parameterised by a scalar d , as shown in Figure 2.8. As an alternative to the method used in *Example 1* we obtain the equivalent parallelogram by constructing the map from the reference element to this quadrilateral, and then discarding the bilinear term. This yields the affine map

$$\bar{F}(\xi) = \frac{1}{4} \begin{pmatrix} (3-d) + (3-d)\xi - (1+d)\eta \\ (3-d) - (1+d)\xi + (3-d)\eta \end{pmatrix},$$

from which we obtain

$$B = \frac{1}{4} \begin{pmatrix} (3-d) & -(1+d) \\ -(1+d) & (3-d) \end{pmatrix}.$$

The vector \mathbf{k} is easily found from (2.37), and is given by

$$\mathbf{k} = -\frac{1}{4}(d+1) \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Thus we find that

$$B^{-1}\mathbf{k} \cdot \mathbf{l} = -\frac{1+d}{2(1-d)}(\xi + \eta).$$

The only restriction which we obtain from (2.40) is at node 1; this gives $d < 0$, which constrains Ω to be convex. For any such choice of d , the constant α may be chosen such that $-1 < \alpha \leq -(1+d)/(1-d)$.

Returning to the general case, for a regular family the matrix B has the properties

$$C_1 h \leq \|B\| \leq C_2 h, \quad (2.42)$$

$$C_3 h^{-1} \leq \|B^{-1}\| \leq C_4 h^{-1}, \quad (2.43)$$

$$C_5 h^2 \leq \det B \leq C_6 h^2, \quad (2.44)$$

where C_1, \dots, C_6 are constants independent of h . These properties are readily established by using the definition of the matrix norm, and by observing that B maps $\hat{\Omega}$ (up to a translation) onto the equivalent parallelogram $\bar{\Omega}$ which has the same diameter h as Ω .

We now obtain desired estimates for the quadrilateral element. In what follows, c or C is a generic positive constant whose value will depend on the particular context.

Theorem 2.7 *Let Ω be a quadrilateral finite element, isoparametrically equivalent to the reference square element $\hat{\Omega}$ through a mapping F where F is defined as in (2.38). Then the following estimates hold:*

$$c_1 h^2 \leq |J|_{0,\infty} \leq c_2 h^2, \quad (2.45)$$

$$|F|_{1,\infty} \leq c_3 h, \quad |F|_{2,\infty} \leq c_4 h, \quad (2.46)$$

$$|F^{-1}|_{1,\infty} \leq c_6 h^{-1}, \quad |F^{-1}|_{2,\infty} \leq c_7 h^{-1}, \quad (2.47)$$

where J is the Jacobian.

Proof

To estimate J we recall first the identity for the determinant of the sum of two matrices [14, page 47]:

$$\det(C + D) = \det C + C \cdot D^* + D \cdot C^* + \det D,$$

where the adjugate C^* of C is defined by $C^* = (\det C)C^{-T}$, and the product $C \cdot D = \sum_{i,j=1}^n C_{ij}D_{ij}$. We also know that $\det(\mathbf{a} \otimes \mathbf{b}) = 0$. From these it follows that

$$\begin{aligned}
J &= \det(\mathbf{B} + \mathbf{k} \otimes \mathbf{l}) \\
&= \det(\mathbf{B}(\mathbf{I} + \mathbf{B}^{-1}\mathbf{k} \otimes \mathbf{l})) \\
&= \det \mathbf{B} \det(\mathbf{I} + \mathbf{B}^{-1}\mathbf{k} \otimes \mathbf{l}) \\
&= (1 + \mathbf{B}^{-1}\mathbf{k} \cdot \mathbf{l}) \det \mathbf{B} \\
&\leq c_2 h^2,
\end{aligned} \tag{2.48}$$

since

$$|\mathbf{B}^{-1}\mathbf{k}| \leq c \tag{2.49}$$

from (2.43), and from the fact that $|\mathbf{k}| \leq \bar{c}h$.

To estimate the Jacobian from below we use (2.44) and (2.40), to obtain

$$\begin{aligned}
J &\geq \det \mathbf{B}(1 + \alpha) \\
&\geq c_1 h^2,
\end{aligned} \tag{2.50}$$

so that the first inequality with (2.48).

To estimate $|\mathbf{F}|_{1,\infty}$ we note that

$$D\mathbf{F} = \mathbf{B} + \mathbf{k} \otimes \mathbf{l} = \mathbf{B}(\mathbf{I} + \mathbf{B}^{-1}\mathbf{k} \otimes \mathbf{l})$$

and so

$$\begin{aligned}
\|D\mathbf{F}\| &\leq \|\mathbf{B}\| \|\mathbf{I} + \mathbf{B}^{-1}\mathbf{k} \otimes \mathbf{l}\| \\
&\leq \|\mathbf{B}\| (1 + \|\mathbf{B}^{-1}\mathbf{k} \otimes \mathbf{l}\|) \\
&\leq \|\mathbf{B}\| (1 + \|\mathbf{B}^{-1}\mathbf{k}\| \|\mathbf{l}\|),
\end{aligned}$$

the last inequality resulting from the fact that for any vectors \mathbf{a} and \mathbf{b} ,

$$\|\mathbf{a} \otimes \mathbf{b}\| = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{|\mathbf{a}(\mathbf{b} \cdot \mathbf{x})|}{|\mathbf{x}|} = \|\mathbf{a}\| \|\mathbf{b}\|.$$

Thus

$$|F|_{1,\infty} = \sup_{\xi \in \hat{\Omega}} \|DF(\xi)\| \leq c_3 h$$

from (2.49) and (2.42).

It is readily observed from (2.39) that $D^2F = \mathbf{k} \otimes M$ is in fact a constant depending on the parameter \mathbf{k} , so that

$$|F|_{2,\infty} \leq c_4 h.$$

Finally, since $\sup_{\xi \in \hat{\Omega}} \|B^{-1}\mathbf{k} \otimes \mathbf{l}\| \leq c$, the matrix $(I + B^{-1}\mathbf{k} \otimes \mathbf{l})$ is invertible for all $\xi \in \hat{\Omega}$ and therefore

$$DF^{-1} = (DF)^{-1} = (I + B^{-1}\mathbf{k} \otimes \mathbf{l})^{-1} B^{-1}.$$

By making use of the identity $(I + B^{-1}\mathbf{k} \otimes \mathbf{l})^{-1} = I - \frac{B^{-1}\mathbf{k} \otimes \mathbf{l}}{1 + B^{-1}\mathbf{k} \cdot \mathbf{l}}$, and the inequalities (2.49) and (2.40), notice that

$$\begin{aligned} \left\| I - \frac{B^{-1}\mathbf{k} \otimes \mathbf{l}}{1 + B^{-1}\mathbf{k} \cdot \mathbf{l}} \right\| &\leq 1 + \left\| \frac{B^{-1}\mathbf{k} \otimes \mathbf{l}}{1 + B^{-1}\mathbf{k} \cdot \mathbf{l}} \right\| \\ &\leq 1 + \frac{c}{1+\alpha} \equiv \beta, \end{aligned}$$

and hence we have

$$\begin{aligned} |F^{-1}|_{1,\infty} &= \sup_{\xi \in \hat{\Omega}} \|DF^{-1}(\xi)\| \\ &\leq \beta \|B^{-1}\|, \\ &\leq c_6 h^{-1} \quad (\text{from (2.43)}). \end{aligned}$$

To obtain the estimate for $|F^{-1}|_{2,\infty}$ we will find the function D^2F^{-1} as follows:

For given functions $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$, the composite function

$H = G \circ F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is such that for all vectors $\eta_1, \eta_2 \in \mathbb{R}^n$,

$$D^2 H(\xi)(\eta_1, \eta_2) = DG(\mathbf{x})(D^2 F(\xi)(\eta_1, \eta_2)) + D^2 G(\mathbf{x})(DF(\xi)\eta_1, DF(\xi)\eta_2).$$

If we apply the formula with $G = F^{-1}$, so that $H = I$, we obtain, for all $\mathbf{x} = F(\xi) \in \Omega$,

$$D^2 F^{-1}(\mathbf{x})(DF(\xi)\eta_1, DF(\xi)\eta_2) = -DF^{-1}(\mathbf{x})(D^2 F(\xi)(\eta_1, \eta_2)).$$

Since for each $\mathbf{x} \in \Omega$ the mapping $DF(\xi) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is invertible, we have

$$D^2 F^{-1}(\mathbf{x})(\eta_1, \eta_2) = -DF^{-1}(\mathbf{x})(D^2 F(\xi)(DF^{-1}(\mathbf{x})\eta_1, DF^{-1}(\mathbf{x})\eta_2)),$$

and thus

$$\|D^2 F^{-1}(\mathbf{x})\| = \sup_{\substack{\|\eta_i\| \leq 1 \\ i=1,2}} \|D^2 F^{-1}(\mathbf{x})(\eta_1, \eta_2)\| \leq \|D^2 F(\xi)\| \|DF^{-1}(\mathbf{x})\|^3,$$

so that

$$\|F^{-1}\|_{2,\infty} = \sup_{\xi \in \hat{\Omega}} \|D^2 F^{-1}(\xi)\| \leq \|F\|_{2,\infty} \|F^{-1}\|_{1,\infty}^3 \leq c_7 h^{-2},$$

and the Theorem is proved. □

Lemma 2.2 *Let Ω be a quadrilateral finite element isoparametrically equivalent to the square reference element $\hat{\Omega}$ through the mapping F given by (2.31). Then the following estimates hold:*

$$\|F\|_{1,\infty} \leq c_1 h, \quad \|F\|_{2,\infty} = 0. \tag{2.51}$$

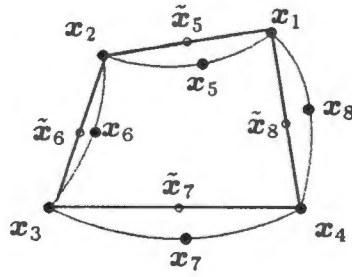


Figure 2.9: quadrilateral element of type (2)

Proof

The above estimates are easy to obtain from the steps we used in the above Theorem.

□

With the use of the above results and Theorem 2.5, we obtain the following interpolation error estimates for bilinear quadrilaterals:

Theorem 2.8 *Let $\hat{\Omega}$ be mapped onto a quadrilateral $\Omega = F(\hat{\Omega}) \subset \mathbb{R}^2$ by bilinear isoparametric map F , and let Ω be regular in the sense of (2.40). Then there exists a constant C such that, for all $u \in H^2(\Omega)$*

$$|u - \Pi u|_{m,\Omega} \leq ch^{2-m} |u|_2, \quad m = 0, 1. \tag{2.52}$$

where Π is an interpolation operator with properties as in Theorem 2.5.

Proof

The proof of this theorem is a direct consequence of Theorem 2.5 with $k = 1$ and Lemma 2.2. □

It is interesting to note that the above results are easily extendable to biquadratic maps. We sketch the general procedure.

For each $\Omega = F(\hat{\Omega})$, where the mapping $F \in (Q_2(\hat{\Omega}))^2$ is uniquely determined by eight points $x_A, 1 \leq A \leq 8$ on the element Ω (see Fig 2.9). Let \tilde{F} denote the bilinear

mapping uniquely determined by the corner nodes \mathbf{x}_A , $1 \leq A \leq 4$. Then we say that the family is regular if the following conditions are satisfied:

- the quadrilateral joined by the corner nodes is regular in the sense of (2.40);
- for each Ω

$$\|\mathbf{x}_A - \tilde{\mathbf{x}}_A\| = O(h^2), \quad 5 \leq A \leq 8, \quad (2.53)$$

$$\text{where } \tilde{\mathbf{x}}_A = \tilde{F}(\boldsymbol{\xi}_A) = \tilde{\mathbf{x}}_A, \quad 5 \leq A \leq 8.$$

Note that the second condition follows that given in [18]. Now, let \tilde{F} denote the affine map for the equivalent parallelogram of the quadrilateral $\tilde{\Omega} = \tilde{F}(\hat{\Omega})$ and $\tilde{\mathbf{x}}_A$ ($1 \leq A \leq 4$) the corner nodes of the equivalent parallelogram. Then the mapping F can be written as

$$\begin{aligned} F &= \sum_{A=1}^8 \mathbf{x}_A N_A \\ &= \sum_{A=1}^8 \tilde{\mathbf{x}}_A N_A + \sum_{A=5}^8 (\mathbf{x}_A - \tilde{\mathbf{x}}_A) N_A. \end{aligned}$$

Note that the first term of the second line generates the mapping \tilde{F} which belongs to $Q_1(\hat{\Omega})$. Therefore, replacing $\tilde{\mathbf{x}}_A$ by $\{\tilde{\mathbf{x}}_A + (\tilde{\mathbf{x}}_A - \tilde{\mathbf{x}}_A)\}$ and recalling the property of the equivalent parallelogram we have

$$F = \sum_{A=1}^8 \tilde{\mathbf{x}}_A N_A + \sum_{A=1}^4 (\tilde{\mathbf{x}}_A - \tilde{\mathbf{x}}_A) N_A + \sum_{A=5}^8 (\mathbf{x}_A - \tilde{\mathbf{x}}_A) N_A, \quad (2.54)$$

since $\tilde{\mathbf{x}}_i = \tilde{\mathbf{x}}_i$, for $5 \leq i \leq 8$.

It can be observed from the preceding representation of the biquadratic map $F(\boldsymbol{\xi})$ that the first term and the second term are respectively linear and bilinear. Therefore, we can obtain the following estimates, with the aid of the assumption (2.53):

$$\begin{aligned} \|F\|_{2,\infty} &= O(h^2) & \|F\|_{3,\infty} &= 0, \\ |F^{-1}|_{2,\infty} &= O(h^{-1}) & \text{and } |F^{-1}|_{3,\infty} &= 0. \end{aligned}$$

We thus have

Theorem 2.9 *Let $\hat{\Omega}$ be mapped onto a quadrilateral $\Omega = \mathbf{F}(\hat{\Omega}) \subset \mathbb{R}^2$, where $\mathbf{F} \in (Q_2(\hat{\Omega}))^2$, and let Ω be regular according to the conditions given above. Then there exists a constant C such that, for all $\mathbf{u} \in H^2(\Omega)$*

$$|\mathbf{u} - \Pi\mathbf{u}|_{m,\Omega} \leq C h^{3-m}; \quad m = 0, 1, 2. \quad (2.55)$$

where Π is an interpolation operator with properties as in Theorem 2.5.

REMARKS.

1. One may readily imagine that it is possible to give the above estimates in more general form, similar to that for affine families. That is, if $\mathbf{u} \in H^{k+1}(\Omega)$ then the following estimate is true for an isoparametric quadrilateral element of k th order:

$$|\mathbf{u} - \Pi\mathbf{u}|_{m,\Omega} \leq C h^{k+1-m}; \quad 0 \leq m \leq k. \quad (2.56)$$

In practice, it is not necessary since we are not interested in elements of higher order than two due to the higher cost. But from a mathematical point of view it is possible to obtain the interpolation estimates for k th order elements with suitable assumptions.

2. For the case of hexahedral elements, it is possible to find such an error estimate by identifying the associated parallelepiped. In fact we have identified such a parallelepiped. However, in contrast to the two-dimensional case, the isoparametric map from a cube to a hexahedron contains four nonlinear terms which makes it extremely difficult to obtain the required geometrical estimates.

Chapter 3

The Enhanced Strain Finite Element Method

In this Chapter we formulate both the continuous and discrete variational problems for the enhanced strain method. The goal is to describe a general framework for the analysis, in which consequences arising from the introduction of enhanced strain and the error estimates may be discussed. The model problem is taken to be that of linear elasticity.

In contrast to most approaches, we introduce the enhanced strain at the outset in the classical formulation, and obtain a weak or variational formulation of the problem by employing standard techniques. Existing approaches [2, 48] begin with a minimization problem.

3.1 The continuous problem

Suppose that a linear elastic body occupies a region Ω in \mathbb{R}^n ($n = 2, 3$) with boundary $\partial\Omega$, and suppose that the body force $\mathbf{f}(\mathbf{x})$ is given on Ω and the displacement field, denoted by $\bar{\mathbf{u}} : \Omega \rightarrow \mathbb{R}^n$, is given by the function $\mathbf{g}(\mathbf{x})$ on $\partial\Omega$. For simplicity

we adopt a Dirichlet boundary condition only.

The governing equations of the problem associated with enhanced strains are:

the equation of equilibrium

$$\operatorname{div} \boldsymbol{\sigma} + \mathbf{f} = \mathbf{0} \quad \text{in } \Omega, \quad (3.1)$$

the elastic constitutive equation

$$\boldsymbol{\sigma} = \mathbf{C}[\boldsymbol{\varepsilon}(\mathbf{u}) + \tilde{\boldsymbol{\varepsilon}}] \quad \text{in } \Omega, \quad (3.2)$$

the enhanced strain condition

$$\tilde{\boldsymbol{\varepsilon}} = \mathbf{0} \quad \text{in } \Omega, \quad (3.3)$$

the strain-displacement relation

$$\boldsymbol{\varepsilon}(\mathbf{u}) = \frac{1}{2} (\nabla \mathbf{u} + \nabla^T \mathbf{u}) \quad \text{in } \Omega, \quad (3.4)$$

and the inhomogeneous essential boundary condition

$$\mathbf{u} = \mathbf{g} \quad \text{on } \partial\Omega. \quad (3.5)$$

Here $\boldsymbol{\sigma}$ is the symmetric stress tensor and $\boldsymbol{\varepsilon}$ is the infinitesimal strain, while \mathbf{C} is the fourth-order tensor of elastic moduli. The quantity $\mathbf{C}[\boldsymbol{\varepsilon}]$ denotes the second-order tensor with components $C_{ijkl}\varepsilon_{kl}$, the tensor $\tilde{\boldsymbol{\varepsilon}}$ is called the enhanced strain, and $\nabla \mathbf{u}$ is the tensor with components $\frac{\partial u_i}{\partial x_j}$.

With (3.3), the modification to the strain is trivial for the continuous problem.

However, it will be seen later that this is not the case for the discrete problem, where the enhanced strain plays an important role.

The above system of equations may be transformed into one with a homogeneous boundary condition. Let $\bar{\mathbf{u}}$ denote a particular displacement field in $\bar{\Omega}$ which satisfies the boundary condition: that is,

$$\bar{\mathbf{u}} = \mathbf{g} \quad \text{on } \partial\Omega.$$

Then the displacement field $\hat{\mathbf{u}} = \mathbf{u} - \bar{\mathbf{u}}$ satisfies the homogeneous boundary value problem

$$\begin{aligned} \operatorname{div} \hat{\boldsymbol{\sigma}} + \hat{\mathbf{f}} &= \mathbf{0} \quad \text{in } \Omega, \\ \hat{\boldsymbol{\sigma}} &= \mathbf{C}[\boldsymbol{\varepsilon}(\hat{\mathbf{u}}) + \tilde{\boldsymbol{\varepsilon}}] \quad \text{in } \Omega, \\ \boldsymbol{\varepsilon}(\hat{\mathbf{u}}) &= \frac{1}{2}(\nabla \hat{\mathbf{u}} + \nabla^T \hat{\mathbf{u}}) \quad \text{in } \Omega, \\ \hat{\mathbf{u}} &= \mathbf{0} \quad \text{on } \partial\Omega, \end{aligned}$$

where $\hat{\mathbf{f}} = \mathbf{f} + \operatorname{div} \bar{\boldsymbol{\sigma}}$ and $\bar{\boldsymbol{\sigma}} = \mathbf{C}[\boldsymbol{\varepsilon}(\bar{\mathbf{u}})]$.

Solving for $\hat{\mathbf{u}}$ in (3.6), we can recover \mathbf{u} from $\mathbf{u} = \hat{\mathbf{u}} + \bar{\mathbf{u}}$. Thus without loss of generality we consider only the homogeneous boundary value problem

$$\operatorname{div} \boldsymbol{\sigma} + \mathbf{f} = \mathbf{0} \quad \text{in } \Omega, \tag{3.6}$$

$$\boldsymbol{\sigma} = \mathbf{C}[\boldsymbol{\varepsilon}(\mathbf{u}) + \tilde{\boldsymbol{\varepsilon}}] \quad \text{in } \Omega, \tag{3.7}$$

$$\boldsymbol{\varepsilon}(\mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u} + \nabla^T \mathbf{u}) \quad \text{in } \Omega, \tag{3.8}$$

$$\tilde{\boldsymbol{\varepsilon}} = \mathbf{0} \quad \text{in } \Omega, \tag{3.9}$$

$$\mathbf{u} = \mathbf{0} \quad \text{on } \partial\Omega. \tag{3.10}$$

The fourth-order tensor of elastic moduli \mathbf{C} has the usual symmetry properties

$$C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klij}. \tag{3.11}$$

The components of C are assumed to be bounded measurable functions; that is,

$$C_{ijkl} \in L^\infty(\Omega) \quad \text{with} \quad \|C_{ijkl}\|_\infty \leq M, \quad M > 0. \quad (3.12)$$

It is furthermore assumed that C is pointwise stable [34] in the sense that a positive constant c_0 exists such that

$$C_{ijkl}M_{ij}M_{kl} \geq c_0M_{ij}M_{ij} \quad (3.13)$$

for almost all $\mathbf{x} \in \Omega$ and for all symmetric matrices M , the summation convention for repeated indices being assumed here and henceforth. Notice that this condition implies (but it is not implied by) the weaker requirement of strong ellipticity.

With the above properties of C , if $\mathbf{f} \in [L_2(\Omega)]^n$ and if Ω is convex, then there exists a unique solution $\mathbf{u} \in [H^2(\Omega)]^n \cap [H_0^1(\Omega)]^n$ to the system (3.6) - (3.10) such that

$$\|\mathbf{u}\|_2 \leq C\|\mathbf{f}\|_0, \quad (3.14)$$

where C is a positive constant independent of \mathbf{f} ; see Fichera [23].

We will also require *Korn's inequality* [23] which states that there exists a positive constant \tilde{C} such that

$$\tilde{C}|\mathbf{v}|_1 \leq \|\boldsymbol{\varepsilon}(\mathbf{v})\|_0. \quad (3.15)$$

In order to construct the weak form of variational formulation of the above problem, we introduce the following function spaces. The space of admissible *displacements* is defined by $V := [H_0^1(\Omega)]^n$. The space V is a Hilbert space with the product norm

$\|\cdot\|_V$ defined by

$$\|v\|_V^2 = \sum_{i=1}^n |v_i|_{1,\Omega}^2, \quad v = (v_1, \dots, v_n) \in V.$$

The space of admissible *stresses* S is defined by

$$S = \{\tau = (\tau_{ij}) : \tau_{ij} = \tau_{ji}, \tau_{ij} \in L_2(\Omega)\},$$

and the space of *enhanced strains* is denoted by Γ ; it suffices to set $\Gamma = S$. The spaces S and Γ are Hilbert spaces equipped with the usual L_2 -inner product and norm; that is,

$$\|\tau\|_S^2 = \sum_{i,j=1}^n \|\tau_{ij}\|_{0,\Omega}^2 \quad \text{for } \sigma \in S,$$

with a similar definition for $\|\cdot\|_\Gamma$. Finally we introduce the product space $\Psi = V \times \Gamma$, which is a Hilbert space with the usual product norm.

To obtain a weak variational problem of (3.6)-(3.10), we take the scalar product of equations (3.6) and (3.7) with arbitrary functions $v \in V$ and $\gamma \in \Gamma$ respectively, to obtain

$$\begin{aligned} \operatorname{div} \sigma \cdot v + f \cdot v &= 0, \\ \sigma \cdot \gamma - C[\varepsilon(u) + \tilde{\varepsilon}] \cdot \gamma &= 0. \end{aligned}$$

The inner product $\beta \cdot \gamma$ of two matrices or second-order tensors is defined by $\beta \cdot \gamma = \beta_{ij} \gamma_{ij}$ and the product of two vectors by the usual scalar product.

By substituting for σ in the first equation, integrating these two equations over the domain Ω , then making use of Green's theorem and the boundary condition (3.10), we have

$$-\int_{\Omega} C[\varepsilon(u) + \tilde{\varepsilon}] \cdot \varepsilon(v) dx + \int_{\Omega} f \cdot v dx = 0,$$

$$\int_{\Omega} \boldsymbol{\sigma} \cdot \boldsymbol{\gamma} \, d\mathbf{x} - \int_{\Omega} \mathbf{C}[\boldsymbol{\varepsilon}(\mathbf{u}) + \tilde{\boldsymbol{\varepsilon}}] \cdot \boldsymbol{\gamma} \, d\mathbf{x} = 0.$$

We obtain a single variational equation by subtracting the second equation from the first; this gives

$$\int_{\Omega} \mathbf{C}[\boldsymbol{\varepsilon}(\mathbf{u}) + \tilde{\boldsymbol{\varepsilon}}] \cdot (\boldsymbol{\varepsilon}(\mathbf{v}) + \boldsymbol{\gamma}) \, d\mathbf{x} - \int_{\Omega} \boldsymbol{\sigma} \cdot \boldsymbol{\gamma} \, d\mathbf{x} = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \, d\mathbf{x}. \quad (3.16)$$

We now define the bilinear forms $a : V \times V \rightarrow \mathbb{R}$, $b : V \times \Gamma \rightarrow \mathbb{R}$ and $c : \Gamma \times \Gamma \rightarrow \mathbb{R}$ by

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) &:= \int_{\Omega} \mathbf{C}[\boldsymbol{\varepsilon}(\mathbf{u})] \cdot \boldsymbol{\varepsilon}(\mathbf{v}) \, d\mathbf{x} \\ b(\mathbf{v}, \boldsymbol{\gamma}) &:= \int_{\Omega} \mathbf{C}[\boldsymbol{\varepsilon}(\mathbf{v})] \cdot \boldsymbol{\gamma} \, d\mathbf{x} \\ c(\tilde{\boldsymbol{\varepsilon}}, \boldsymbol{\gamma}) &:= \int_{\Omega} \mathbf{C}[\tilde{\boldsymbol{\varepsilon}}] \cdot \boldsymbol{\gamma} \, d\mathbf{x} \end{aligned}$$

for $\mathbf{u}, \mathbf{v} \in V$ and $\tilde{\boldsymbol{\varepsilon}}, \boldsymbol{\gamma} \in \Gamma$.

The linear functional $l : V \rightarrow \mathbb{R}$ is defined by

$$l(\mathbf{v}) := \int_{\Omega} \mathbf{f} \cdot \mathbf{v} \, d\mathbf{x} \quad \text{for } \mathbf{v} \in V.$$

We also set $\boldsymbol{\phi} = (\mathbf{u}, \tilde{\boldsymbol{\varepsilon}})$ and $\boldsymbol{\psi} = (\mathbf{v}, \boldsymbol{\gamma})$, and define the bilinear form $A : \Psi \times \Psi \rightarrow \mathbb{R}$ by

$$\begin{aligned} A(\boldsymbol{\phi}, \boldsymbol{\psi}) &= a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, \tilde{\boldsymbol{\varepsilon}}) + b(\mathbf{u}, \boldsymbol{\gamma}) + c(\tilde{\boldsymbol{\varepsilon}}, \boldsymbol{\gamma}) \\ &= \int_{\Omega} \mathbf{C}[\boldsymbol{\varepsilon}(\mathbf{u}) + \tilde{\boldsymbol{\varepsilon}}] \cdot (\boldsymbol{\varepsilon}(\mathbf{v}) + \boldsymbol{\gamma}) \, d\mathbf{x}. \end{aligned}$$

Then (3.16) can be rewritten in the form

$$A(\boldsymbol{\phi}, \boldsymbol{\psi}) - (\boldsymbol{\sigma}, \boldsymbol{\gamma})_0 = l(\mathbf{v}) \quad \text{for all } \boldsymbol{\psi} \in \Psi \quad (3.17)$$

By taking the scalar product of (3.3) with arbitrary $\boldsymbol{\tau} \in S$ and integrating over Ω we obtain

$$\int_{\Omega} \boldsymbol{\tau} \cdot \tilde{\boldsymbol{\varepsilon}} \, d\mathbf{x} = 0.$$

We can now define the variational problem.

Problem P

Given $\mathbf{f} \in [L_2(\Omega)]^n$, find $\phi = (\mathbf{u}, \tilde{\boldsymbol{\varepsilon}}) \in \Psi$ and $\boldsymbol{\sigma} \in S$ such that

$$A(\phi, \psi) - (\boldsymbol{\sigma}, \boldsymbol{\gamma})_0 = l(\mathbf{v}) \quad \text{for all } \psi \in \Psi, \quad (3.18)$$

$$(\boldsymbol{\tau}, \tilde{\boldsymbol{\varepsilon}})_0 = 0 \quad \text{for all } \boldsymbol{\tau} \in S. \quad (3.19)$$

In expanded form, (3.18) and (3.19) is the set of equations

$$a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, \tilde{\boldsymbol{\varepsilon}}) = l(\mathbf{v}) \quad \text{for all } \mathbf{v} \in V \quad (3.20)$$

$$b(\mathbf{u}, \boldsymbol{\gamma}) + c(\tilde{\boldsymbol{\varepsilon}}, \boldsymbol{\gamma}) - (\boldsymbol{\sigma}, \boldsymbol{\gamma})_0 = 0 \quad \text{for all } \boldsymbol{\gamma} \in \Gamma \quad (3.21)$$

$$(\boldsymbol{\tau}, \tilde{\boldsymbol{\varepsilon}}) = 0 \quad \text{for all } \boldsymbol{\tau} \in S. \quad (3.22)$$

Note that in the continuous case the enhanced strain $\tilde{\boldsymbol{\varepsilon}}$ is identically zero, from (3.22).

We note that Problem P is equivalent to finding the stationary point of the functional $\Pi : V \times \Gamma \times S \rightarrow \mathbb{R}$ defined by

$$\Pi(\mathbf{v}, \boldsymbol{\gamma}, \boldsymbol{\tau}) = \int_{\Omega} (\boldsymbol{\varepsilon}(\mathbf{v}) + \tilde{\boldsymbol{\gamma}}) \cdot \mathbf{C}[\boldsymbol{\varepsilon}(\mathbf{v}) + \boldsymbol{\gamma}] d\mathbf{x} - \int_{\Omega} \boldsymbol{\tau} \cdot \boldsymbol{\gamma} d\mathbf{x} - l(\mathbf{v}),$$

which can be viewed as a modified version of the classical Hu-Washizu principle. We refer to [48] for additional details.

Theorem 3.1 [44] *Problem P has a unique solution.*

Proof

The proof of the theorem is approached not by a direct application of the Lax-Milgram theorem, but rather by considering separately the three variational equations (3.20) - (3.22). Equation (3.22) yields directly the result $\tilde{\boldsymbol{\varepsilon}} = \mathbf{0}$. Substituting


this into (3.20), it becomes

$$a(\mathbf{u}, \mathbf{v}) = l(\mathbf{v}) \quad \text{for all } \mathbf{v} \in V.$$

This is a standard variational problem. The V -ellipticity and continuity of the bilinear form $a(\cdot, \cdot)$ may be established by making use of the properties of C and the Korn's inequality, and therefore the solution exists and is unique by Lax-Milgram theorem.

Finally, we observe from equation (3.21) that

$$(\boldsymbol{\sigma} - C[\boldsymbol{\varepsilon}(\mathbf{u})], \boldsymbol{\gamma})_0 = 0 \quad \text{for all } \boldsymbol{\gamma} \in \Gamma;$$

this gives $\boldsymbol{\sigma} = C[\boldsymbol{\varepsilon}(\mathbf{u})]$ in \mathcal{S}  □

3.2 The Discrete Problem

In this Section we formulate the discrete problem and analyse the stability and convergence of solutions. In order to formulate the problem we introduce the finite dimensional spaces $V^h \subset V$, $\Gamma^h \subset \Gamma$ and $S^h \subset S$. Full details of these spaces will be given shortly.

The domain Ω is covered with nonoverlapping convex subdomains $\Omega_1, \dots, \Omega_E$, that is, $\cup_{e=1}^E \bar{\Omega}_e = \bar{\Omega}$; the elements Ω_e are assumed to be generated by *isoparametric maps* from the reference element. In other words, Ω_e is the image of $\hat{\Omega}$ by a bijective mapping

$$\mathbf{F}_e : \boldsymbol{\xi} \in \hat{\Omega} \rightarrow \mathbf{F}_e(\boldsymbol{\xi}) = (F_{e,1}(\boldsymbol{\xi}), \dots, F_{e,n}(\boldsymbol{\xi})) \in \mathbb{R}^n \quad (3.23)$$

where the functions $F_{i,e}$ are such that $F_{i,e}(\boldsymbol{\xi})$ are polynomials which contain either $P_k(\hat{\Omega})$, the space of polynomials of degree $\leq k$ on $\hat{\Omega}$, or $Q_k(\hat{\Omega})$, the space of polynomial of the type (2.6). It is important to note that meshes comprising triangular elements, with conventional Lagrangian basis functions defined on these elements, are ruled out because imposition of criteria which have to be satisfied by the enhanced strains result in the problem being equivalent to the standard problem. Therefore, we are concerned particularly with quadrilateral elements. Therefore $F_{i,e}(\boldsymbol{\xi})$ are chosen to be members of Q_k .

We are now in a position to define V^h . By virtue of the assumption of an isoparametric map V^h is defined by

$$V^h := \{v_h \in [C^0(\bar{\Omega})]^n : v_h|_{\Omega_e} = \hat{v}_h \cdot F_e^{-1} \quad \text{for } \hat{v}_h \in [Q_k(\hat{\Omega})]^n\}. \quad (3.24)$$

In other words, $v_h|_{\Omega_e}(\mathbf{x}) = \hat{v}_h(\boldsymbol{\xi})$ where $\mathbf{x} = F_e(\boldsymbol{\xi})$.

For later use we set

$$\varepsilon(V^h) := \{\boldsymbol{\eta} \in [L_2(\Omega)]_{sym}^{n \times n} : \boldsymbol{\eta} = \varepsilon(v_h) \text{ for } v_h \in V^h\}.$$

In order to define the space Γ^h in a manner which will ensure the stability and consistency of the discrete problem, it is necessary to deviate from the standard relation between functions defined on the reference and actual elements, such as in (3.24) above. We give a preliminary definition of Γ^h according to

$$\Gamma^h \subseteq G^h = \{\boldsymbol{\gamma}_h : \boldsymbol{\gamma}_h^e(\mathbf{x}) = \frac{J_0^e}{J^e} \hat{\boldsymbol{\gamma}}_h(\boldsymbol{\xi}), \quad \hat{\boldsymbol{\gamma}}_h \in [Q_l]_{sym}^{n \times n}\} \quad (3.25)$$

for suitable l , where $\boldsymbol{\gamma}_h^e = \boldsymbol{\gamma}_h|_{\Omega_e}$ and J_0^e is the value of J^e at the centroid of Ω_e . For the case of quadrilateral elements, J_0^e can be interpreted as the Jacobian of the map from a reference element to the associated equivalent parallelogram.

In the definition of the space \mathbf{G}^h it is not essential that the term J_0^e appear; this is included, though, so that the definition of \mathbf{G}^h , and ultimately that of Γ^h , will reduce to the standard form in the case of an affine map, for which J^e is equal to J_0^e .

It is not possible to define Γ^h simply to be the same as \mathbf{G}^h , since Γ^h is required to satisfy a further set of conditions, in order that the discrete problem be stable and consistent. These conditions will be discussed shortly. Details regarding the space S^h will be discussed in the next Section.

We now formulate the discrete problem.

Problem \bar{P}^h

Given $\mathbf{f} \in [L_2(\Omega)]^n$, find $\phi_h = (\mathbf{u}_h, \tilde{\epsilon}_h) \in \Psi^h = V^h \times \Gamma^h$ and $\sigma_h \in S^h$ such that

$$A(\phi_h, \psi_h) - (\sigma_h, \gamma_h)_0 = l(v_h) \quad \text{for all } \psi_h \in \Psi^h, \quad (3.26)$$

$$(\tau_h, \tilde{\epsilon}_h)_0 = 0 \quad \text{for all } \tau_h \in S^h. \quad (3.27)$$

We observe from (3.27) that $\tilde{\epsilon}_h$ is in general nonzero, unlike the continuous case.

Conditions for well-posedness

With V^h given by (3.24), the choice of the subspaces Γ^h and S^h are constrained according to three basic requirements. These are:

- (I) S^h and Γ^h are L_2 -orthogonal to each other.
- (II) $\Gamma^h \cap \varepsilon(V^h) = \{0\}$; Moreover, it is assumed that there exists a constant c_1 independent of h with $0 < c_1 < 1$ such that for any $\gamma_h \in \Gamma^h$, $\|\gamma_h^0\|_\Gamma \leq c_1 \|\gamma_h\|_\Gamma$; where γ_h^0 denotes the L_2 -orthogonal projection of γ_h onto $\varepsilon(V^h)$.
- (III) The space Γ^h is required to be L_2 -orthogonal, in the inner product defined by \mathbf{C} , to the space of functions whose restrictions to $\hat{\Omega}$ are $n \times n$ symmetric matrices of members of $Q_{k-1}(\hat{\Omega})$, where k is as defined in (3.24). In other words,

(the condition) that is:

$$\int_{\hat{\Omega}} \hat{\mathbf{q}}_{k-1} \cdot \mathbf{C}[\hat{\boldsymbol{\gamma}}_h] d\hat{\boldsymbol{\xi}} = 0 \quad \text{or} \quad \int_{\hat{\Omega}} (\hat{\mathbf{q}}_{k-1})_{ij} C_{ijlm}(\hat{\boldsymbol{\gamma}})_{lm} d\hat{\boldsymbol{\xi}} \quad (3.28)$$

must hold for any $\hat{\mathbf{q}}_{k-1} \in [Q_{k-1}(\hat{\Omega})]_{sym}^{n \times n}$ with $k \geq 1$ for all $\hat{\boldsymbol{\gamma}}_h = \frac{J^e}{J_0^e} \boldsymbol{\gamma}_h^e(\mathbf{x})$, and for all Ω_e , where $\boldsymbol{\gamma}_h^e$ denotes the restriction of $\boldsymbol{\gamma}_h \in \Gamma^h$ to Ω_e .

As a result of the condition (I), the stress field may be eliminated from Problem \bar{P}^h which now becomes, with this assumption,

Problem P^h

Given $\mathbf{f} \in [L_2(\Omega)]^n$, find $\boldsymbol{\phi}_h \in \Psi^h = V^h \times \Gamma^h$ such that

$$A(\boldsymbol{\phi}_h, \boldsymbol{\psi}_h) = l(\mathbf{v}_h) \quad \text{for all } \boldsymbol{\psi}_h \in \Psi \quad (3.29)$$

or, equivalently, such that

$$a(\mathbf{u}_h, \mathbf{v}_h) + b(\tilde{\boldsymbol{\varepsilon}}_h, \mathbf{v}_h) = l(\mathbf{v}_h) \quad \forall \mathbf{v}_h \in V^h \quad (3.30)$$

$$b(\mathbf{u}_h, \boldsymbol{\gamma}_h) + c(\tilde{\boldsymbol{\varepsilon}}_h, \boldsymbol{\gamma}_h) = 0 \quad \forall \boldsymbol{\gamma}_h \in \Gamma^h \quad (3.31)$$

The above conditions are either assumed explicitly or appear as particular cases in [48]. Conditions (I) and (II) are unchanged from the set of conditions for the affine case, given in [44].

Condition (III) differs in many ways from that proposed in [44] for affine-equivalent elements. It differs firstly in that it is stated in [44] as a condition involving the space P_k of polynomials of degree $\leq k$, rather than Q_k . The essential difference lies in the fact that the results in [44] apply to affine maps only, whereas here the components of maps \mathbf{F}_e are represented by members of Q_k . For the case $k = 1$ we note that there is no difference, though, in the condition (3.28). Secondly, the orthogonality condition (3.28) is posed on the master element only. In the affine case this would make no difference, of course, but for the isoparametric case and for $k \geq 2$ the condition as stated above does not imply orthogonality with respect

to polynomials on the actual element, since polynomials are not preserved by the isoparametric transformation except for the special case $k = 1$. For this special case we find that

$$\int_{\hat{\Omega}} \hat{\mathbf{q}}_0 \cdot \mathbf{C}[\hat{\boldsymbol{\gamma}}_h] d\hat{\boldsymbol{\xi}} = 0 \iff \mathbf{q}_0 \cdot \int_{\Omega^e} \mathbf{C}[\boldsymbol{\gamma}_h^e] d\mathbf{x} = 0$$

by virtue of the definition (3.25). For the case in which the material is homogeneous, so that \mathbf{C} is constant, this reduces to

$$\int_{\hat{\Omega}} \hat{\boldsymbol{\gamma}}_h d\hat{\boldsymbol{\xi}} = 0 \iff \int_{\Omega^e} \boldsymbol{\gamma}_h^e d\mathbf{x} = 0. \quad (3.32)$$

In [48] the orthogonality condition is enforced on both the actual and reference elements, for the case of four-noded quadrilateral elements, by adopting the alternative definition

$$\boldsymbol{\gamma}_h^e(\mathbf{x}) = \frac{J_0^e}{J^e} \mathbf{M}_0^{-T} \hat{\boldsymbol{\gamma}}_h \mathbf{M}_0 \quad (3.33)$$

where $\mathbf{M}_0 = \left(\frac{\partial \mathbf{x}}{\partial \hat{\boldsymbol{\xi}}} \right)_{\hat{\boldsymbol{\xi}}=0}$. This can be expressed in vectorial form: if $\boldsymbol{\gamma}_h^e = (\gamma_{h,11}, \gamma_{h,22}, 2\gamma_{h,12})^T$, then

$$\boldsymbol{\gamma}_h^e(\mathbf{x}) = \frac{J_0^e}{J^e} \mathbf{T}_0^{-T} \hat{\boldsymbol{\gamma}}_h$$

where

$$\mathbf{T}_0 = \left[\begin{array}{ccc} j_{11}^2 & j_{21}^2 & 2j_{11}j_{21} \\ j_{12}^2 & j_{22}^2 & 2j_{12}j_{22} \\ j_{11}j_{12} & j_{21}j_{22} & (j_{11}j_{22} + j_{12}j_{21}) \end{array} \right]_{\hat{\boldsymbol{\xi}}=0} \quad j_{ij} = \frac{\partial x_i}{\partial \xi_j}. \quad (3.34)$$

It will be clear from the analysis which is to follow that inclusion of the transformation matrix \mathbf{M}_0 in the definition of $\boldsymbol{\gamma}_h$ is not a prerequisite for a stable element.

Note that the assumption (3.28) raises interesting questions about the form which should be taken by the higher-order patch test.

Example [44]

Though the choices of subspaces Γ^h and S^h are constrained according to the requirements listed above, it is not difficult to construct spaces which satisfy these requirements. We consider a concrete example. For the sake of simplicity, consider the case of plane elasticity, for which it is convenient to express the nonzero components of strain and enhanced strain in the vectorial form $\boldsymbol{\varepsilon} = (\varepsilon_{11}, \varepsilon_{22}, 2\varepsilon_{12})$ and $\boldsymbol{\gamma} = (\gamma_{11}, \gamma_{22}, 2\gamma_{12})$ respectively. We consider a uniform mesh of elements generated by isoparametric transformations from a square element $\hat{\Omega} = \{(\xi, \eta) \in \mathbb{R}^2 : -1 \leq \xi \leq 1, -1 \leq \eta \leq 1\}$. It is then most convenient to construct basis functions for $\boldsymbol{\varepsilon}(V^h)$ and Γ^h with requisite properties on $\hat{\Omega}$, and to map these to Ω_e . The space V^h is composed of piecewise bilinear functions, so that the subspace $\boldsymbol{\varepsilon}(V^h)$ is spanned (on $\hat{\Omega}$) by the columns of the matrix

$$\begin{pmatrix} 1 & 0 & 0 & \eta & 0 \\ 0 & 1 & 0 & 0 & \xi \\ 0 & 0 & 1 & \xi & \eta \end{pmatrix}. \quad (3.35)$$

We choose the basis for enhanced strain on the master element as in [48], that is,

$$\begin{pmatrix} \xi & 0 & 0 & 0 \\ 0 & \eta & 0 & 0 \\ 0 & 0 & \xi & \eta \end{pmatrix}. \quad (3.36)$$

Suppose now that the columns of this matrix are labelled as $\gamma_1, \gamma_2, \gamma_3, \gamma_4$. We observe first of all that $\{\gamma_i\}_{i=1}^4$ forms an L_2 -orthogonal set on $\hat{\Omega}$; secondly, γ_1 and γ_2 are all L_2 -orthogonal to $\boldsymbol{\varepsilon}(V^h)$. Thus the only members of the basis which have nonzero L_2 -projections are γ_3 and γ_4 ; these projections onto $\boldsymbol{\varepsilon}(V^h)$ are easily found to be

[44]

$$\gamma_3^0 = \frac{1}{2} \begin{pmatrix} \eta \\ 0 \\ \xi \end{pmatrix} \quad \text{and} \quad \gamma_4^0 = \frac{1}{2} \begin{pmatrix} 0 \\ \xi \\ \eta \end{pmatrix}.$$

Also, from the above it is readily established that

$$\|\gamma_i^0\|_{0,\hat{\Omega}} = \frac{1}{\sqrt{2}} \|\gamma_i\|_{0,\hat{\Omega}} \quad (i = 3, 4).$$

Now set $\hat{\gamma}_h = \frac{J^e}{J_0^e} \gamma_h^e$; then for any γ_h^e we have $\hat{\gamma}_h = \sum_{i=1}^4 \alpha_i \gamma_i$ and it follows, by using the mutual orthogonality of the γ_i , that

$$\begin{aligned} \|\hat{\gamma}_h^0\|_{0,\hat{\Omega}}^2 &= \sum_{i=1}^4 \alpha_i^2 \|\gamma_i^0\|_{0,\hat{\Omega}}^2 \\ &= \alpha_3^2 \|\gamma_3^0\|_{0,\hat{\Omega}}^2 + \alpha_4^2 \|\gamma_4^0\|_{0,\hat{\Omega}}^2 \\ &= \frac{1}{2} (\alpha_3^2 \|\gamma_3\|_{0,\hat{\Omega}}^2 + \alpha_4^2 \|\gamma_4\|_{0,\hat{\Omega}}^2) \\ &\leq \frac{1}{2} \|\hat{\gamma}_h\|_{0,\hat{\Omega}}^2. \end{aligned} \tag{3.37}$$

For the case of isoparametric case it can be shown that the second part of the condition (II) holds, using (3.37) and the fact that $\|\gamma_h^0\|_{0,\hat{\Omega}_e}^2 \leq J_0^2 |J^{-1}|_\infty \|\hat{\gamma}_h^0\|_{0,\hat{\Omega}}^2$.
□

To ensure the existence and uniqueness of a solution of Problem P^h we have to show that the bilinear form $A(\cdot, \cdot)$ is continuous and Ψ^h -elliptic, and that the linear functional $l(\cdot)$ is continuous [54]. To prove the continuity of $A(\cdot, \cdot)$ we need to show that there exists a positive constant M such that

$$|A(\psi_h, \phi_h)| \leq M \|\psi_h\|_\Psi \|\phi_h\|_\Psi. \tag{3.38}$$

By using the properties (3.12) of \mathcal{C} we have

$$\begin{aligned} |A(\psi_h, \phi_h)| &= \left| \int_{\Omega} \mathcal{C}[\varepsilon(\mathbf{u}_h) + \tilde{\varepsilon}_h] \cdot (\varepsilon(\mathbf{v}_h) + \gamma_h) dx \right| \\ &\leq M_0 \|\varepsilon(\mathbf{u}_h) + \tilde{\varepsilon}_h\|_\Gamma \|\varepsilon(\mathbf{v}_h) + \gamma_h\|_\Gamma \\ &\leq M_0 (\|\varepsilon(\mathbf{u}_h)\|_\Gamma + \|\tilde{\varepsilon}_h\|_\Gamma) (\|\varepsilon(\mathbf{v}_h)\|_\Gamma + \|\gamma_h\|_\Gamma). \end{aligned}$$

From the definition of $\boldsymbol{\varepsilon}(\mathbf{u}_h)$ we have

$$\begin{aligned} \|\boldsymbol{\varepsilon}(\mathbf{u}_h)\|_0^2 &\leq \sum_{i,j=1}^n \|\varepsilon_{ij}(\mathbf{u}_h)\|_0^2 \\ &= \sum_{i,j=1}^n \frac{1}{4} \|u_{i,j} + u_{j,i}\|_0^2 \\ &\leq c \sum_{i=1}^n \|u_i\|_1^2 \\ &= c \|\mathbf{u}_h\|_V^2; \end{aligned}$$

whence we obtain (3.38). Here $u_{i,j}$ denotes the differentiation of the component u_i with respect to j th coordinate. Continuity of the linear functional $l(\cdot)$ can be shown in a standard manner.

Next, we show that the bilinear form $A(\cdot, \cdot)$ is Ψ^h -elliptic [44]: from (3.13),

$$\begin{aligned} A(\boldsymbol{\psi}_h, \boldsymbol{\psi}_h) &= \int_{\Omega} \mathbf{C}[\boldsymbol{\varepsilon}(\mathbf{v}_h) + \boldsymbol{\gamma}_h] \cdot (\boldsymbol{\varepsilon}(\mathbf{v}_h) + \boldsymbol{\gamma}_h) \, d\mathbf{x} \\ &\geq c_0 \int_{\Omega} \|\boldsymbol{\varepsilon}(\mathbf{v}_h) + \boldsymbol{\gamma}_h\|^2 \cdot d\mathbf{x} \\ &= c_0 (\|\boldsymbol{\varepsilon}(\mathbf{v}_h)\|_{\Gamma}^2 + \|\boldsymbol{\gamma}_h\|_{\Gamma}^2 + 2(\boldsymbol{\varepsilon}(\mathbf{v}_h), \boldsymbol{\gamma}_h)_{\Gamma}). \end{aligned} \quad (3.39)$$

We next make use of Young's inequality: $2 |(\mathbf{p}, \mathbf{q})_{\Gamma}| \leq d \|\mathbf{p}\|_{\Gamma}^2 + d^{-1} \|\mathbf{q}\|_{\Gamma}^2$ for all $\mathbf{p}, \mathbf{q} \in \Gamma$ and any $d \in \mathbb{R}$, together with the orthogonal decomposition $\boldsymbol{\gamma}_h = \boldsymbol{\gamma}_h^0 + \boldsymbol{\gamma}_h^{\perp}$ (here $\boldsymbol{\gamma}_h^0 \in \boldsymbol{\varepsilon}(V^h)$ and $\boldsymbol{\gamma}_h^{\perp}$ is L_2 -orthogonal to $\boldsymbol{\varepsilon}(V^h)$) to obtain

$$\begin{aligned} 2(\boldsymbol{\varepsilon}(\mathbf{v}_h), \boldsymbol{\gamma}_h)_{\Gamma} &= 2(\boldsymbol{\varepsilon}(\mathbf{v}_h), \boldsymbol{\gamma}_h^0)_{\Gamma} \\ &\geq -d \|\boldsymbol{\varepsilon}(\mathbf{v}_h)\|_{\Gamma}^2 - d^{-1} \|\boldsymbol{\gamma}_h^0\|_{\Gamma}^2 \\ &\geq -d \|\boldsymbol{\varepsilon}(\mathbf{v}_h)\|_{\Gamma}^2 - d^{-1} c_1 \|\boldsymbol{\gamma}_h\|_{\Gamma}^2 \quad \text{by condition (II)}. \end{aligned}$$

Substituting this into (3.39) we obtain

$$A(\boldsymbol{\psi}_h, \boldsymbol{\psi}_h) \geq \alpha \|\boldsymbol{\psi}_h\|_{\Psi}^2$$

where $\alpha = c_0 \{1 - d, 1 - \frac{c_1^2}{d}\}$, with d chosen so that $c_1^2 \leq d \leq 1$.

Thus the bilinear form $A(\cdot, \cdot)$ is Ψ^h -elliptic, and Problem P^h has a unique solution.

From the Ψ^h -ellipticity of $A(\cdot, \cdot)$ and the continuity of $l(\cdot)$ we can show that the solution ϕ_h depends continuously on the data; that is,

$$\|\phi_h\|_{\Psi} \leq C \|f\|_{V^*}, \quad \text{(which means that the approximate solution } (u_h, \tilde{\epsilon}_h) \text{ is stable)}$$

where V^* is the dual space of V .

3.3 Convergence and error estimates

We now return to the task of showing that the problem P^h is consistent, and that the solution $(\mathbf{u}_h, \tilde{\epsilon}_h)$ of the discrete problem converges to the solution $(\mathbf{u}, \mathbf{0})$ of the continuous problem. The proof of consistency hinges on the crucial inequality given in the next Lemma.

Lemma 3.1 *Let $\phi_h = (\mathbf{u}_h, \tilde{\epsilon}_h) \in \Psi^h$ be the solution of the problem P^h and let $\phi = (\mathbf{u}, \mathbf{0})$ be the exact solution. Then we have the estimate*

$$\|\phi - \phi_h\|_{\Psi} \leq C \left\{ \inf_{\psi_h \in \Psi^h} \|\phi - \psi_h\|_{\Psi} + \Delta \right\} \quad (3.39)$$

where

$$\Delta = \sup_{\psi_h \in \Psi^h} \frac{b(\mathbf{u}, \gamma_h)}{\|\psi_h\|_{\Psi}}$$

and the constant $C(> 0)$ is independent of h .

Proof

This inequality is derived in [44] and is independent of the details of the maps \mathbf{F}_e .

From the Ψ^h -ellipticity of $A(\cdot, \cdot)$ it follows that

$$\begin{aligned} \alpha \|\phi_h - \psi_h\|_{\Psi}^2 &\leq A(\phi_h - \psi_h, \phi_h - \psi_h) \\ &= A(\phi - \psi_h, \phi_h - \psi_h) + [A(\phi_h, \phi_h - \psi_h) - A(\phi, \phi_h - \psi_h)] \\ &\leq M \|\phi - \psi_h\|_{\Psi} \|\phi_h - \psi_h\|_{\Psi} - b(\mathbf{u}, \tilde{\epsilon}_h - \gamma_h) \end{aligned}$$

where we have used the continuity of $A(\cdot, \cdot)$ (3.37) and (3.3). We now divide throughout by $\|\phi_h - \psi_h\|_{\Psi}$ and, in the second term on the right-hand side, take the supremum over Ψ^h to obtain the estimate

$$\alpha \|\phi_h - \psi_h\|_{\Psi} \leq M \|\phi - \psi_h\|_{\Psi} + \sup_{\psi_h \in \Psi^h} \frac{b(\mathbf{u}, \gamma_h)}{\|\psi_h\|_{\Psi}}$$

Finally, by the triangle inequality,

$$\|\phi - \phi_h\|_{\Psi} \leq \|\phi - \psi_h\|_{\Psi} + \|\psi_h - \phi_h\|_{\Psi}$$

and if ψ_h is chosen to be the closest element in Ψ^h to ϕ we obtain the required inequality. \square

This result is analogous to the second Strang lemma [54] for nonconforming elements. The first term on the right side of (3.39) may be bounded by making use of the standard interpolation error estimates, but it will be necessary to estimate the second term Δ which arises from the presence of the enhanced strain. We address this issue in the following Lemma.

Lemma 3.2 *Let $\phi = (\mathbf{u}, \mathbf{0})$ be the solution to the continuous problem. Then there is a positive constant c such that*

$$\sup_{\psi_h \in \Psi^h} \frac{b(\mathbf{u}, \gamma_h)}{\|\psi_h\|_{\Psi}} \leq c A_k \sum_{i=1}^k |\mathbf{u}|_{i+1, \Omega} \quad (3.40)$$

where $A_k = \max_{i,e} \{a_{ki}^e\}$ and $a_{11}^e = \|F_e\|_1, a_{21}^e = \|F_e\|_2$ and $a_{22}^e = \|F_e\|_1^2$. The norm $\|\cdot\|_i$ is defined in (2.13).

Proof

From (3.28) we have

$$\int_{\hat{\Omega}} \hat{q}_{k-1} \cdot C[\hat{\gamma}_h] d\xi = 0, \text{ for all } \hat{q}_{k-1} \in [Q_{k-1}(\hat{\Omega})]_{sym}^{n \times n}.$$

Define the bilinear form $\hat{b} : [H^k(\hat{\Omega})]_{sym}^{n \times n} \times [L_2(\hat{\Omega})]_{sym}^{n \times n} \rightarrow \mathbb{R}$ by

$$\hat{b}(\hat{\varepsilon}, \hat{\gamma}_h) = \int_{\hat{\Omega}} \hat{\varepsilon} \cdot \hat{C}[\hat{\gamma}_h] d\xi;$$

then $\hat{b}(\cdot, \hat{\gamma}_h)$ is a bounded linear functional on $[H^k(\hat{\Omega})]_{sym}^{n \times n}$, and it follows from Lemma 2.1 that a constant $C > 0$ exists such that

$$|\hat{b}(\hat{\varepsilon}, \hat{\gamma}_h)| \leq C \|\hat{\varepsilon}\|_{k, \hat{\Omega}} \|\hat{b}(\cdot, \hat{\gamma}_h)\|_{k, \hat{\Omega}}^*, \quad (3.41)$$

where $\|\cdot\|_{k, \hat{\Omega}}^*$ denotes the norm on the topological dual $[H^k(\hat{\Omega})^*]_{sym}^{n \times n}$ of $[H^k(\hat{\Omega})]_{sym}^{n \times n}$.

But

$$\begin{aligned} \|\hat{b}(\cdot, \hat{\gamma}_h)\|_{k, \hat{\Omega}}^* &= \sup_{\hat{\varepsilon} \neq 0} \frac{\hat{b}(\hat{\varepsilon}, \hat{\gamma}_h)}{\|\hat{\varepsilon}\|_k} \\ &\leq \sup_{\hat{\varepsilon} \neq 0} \frac{\|\hat{b}\| \|\hat{\varepsilon}\|_k |\hat{\gamma}_h|_0}{\|\hat{\varepsilon}\|_k} \\ &= \|\hat{b}\| |\hat{\gamma}_h|_0. \end{aligned}$$

Thus

$$|\hat{b}(\hat{\varepsilon}, \hat{\gamma}_h)| \leq \hat{C} \|\hat{\varepsilon}\|_{k, \hat{\Omega}} |\hat{\gamma}_h|_{0, \hat{\Omega}}. \quad (3.42)$$

The next step entails the estimation of $[\hat{\varepsilon}]_{k,\hat{\Omega}}$ and $|\hat{\gamma}_h|_{0,\hat{\Omega}}$ in terms of the quantities on the actual element. First we have, from (3.25),

$$\begin{aligned} |\hat{\gamma}_h|_{0,\hat{\Omega}}^2 &= (J_0^e)^{-2} \int_{\Omega_e} |\gamma_h^e|^2 J^e dx \\ &\leq (J_0^e)^{-2} |J^e|_{0,\infty} |\gamma_h^e|_{0,\Omega_e}^2; \end{aligned}$$

therefore

$$|\hat{\gamma}_h|_{0,\hat{\Omega}} \leq (J_0^e)^{-1} |J^e|_{0,\infty}^{\frac{1}{2}} |\gamma_h^e|_{0,\Omega_e}. \quad (3.43)$$

Next we deal with the term involving $[\hat{\varepsilon}]_{k,\hat{\Omega}}$. For the sake of simplicity, we will do this for the cases $k = 1$ and $k = 2$, as indicated earlier, rather than carry out general analysis.¹ The estimates may be summarised succinctly by writing (see Theorem 2.4)

$$[\hat{\varepsilon}]_{k,\hat{\Omega}} \leq c |(J^e)^{-1}|_{0,\infty}^{\frac{1}{2}} \sum_{i=1}^k a_{k,i}^e |\varepsilon|_{i,\Omega_e} \quad (k = 1, 2). \quad (3.44)$$

Thus we have, from (3.42) - (3.44),

$$\begin{aligned} \left| \int_{\Omega_e} \varepsilon \cdot C[\gamma_h^e] dx \right| &= \left| J_0^e \int_{\hat{\Omega}} \hat{\varepsilon} \cdot \hat{C}[\hat{\gamma}_h] d\xi \right| \\ &\leq c |\gamma_h^e|_{0,\Omega_e} \sum_{i=1}^k a_{ki}^e |\varepsilon|_{i,\Omega_e}, \end{aligned}$$

by making use of the fact that $|(J_0^e)^{-1}|_{0,\infty} = (|J_0^e|_{0,\infty})^{-1}$. Now suppose we choose ε to be the function $\varepsilon(\mathbf{u})$, where $\mathbf{u} \in [H^{k+1}(\Omega_e)]^n$, so that $\varepsilon(\mathbf{u}) \in [H^k(\Omega_e)]_{sym}^{n \times n}$. We then have

$$\begin{aligned} |\varepsilon(\mathbf{u})|_{k,\Omega_e}^2 &\leq \sum_{i,j=1}^n |\varepsilon_{ij}(\mathbf{u})|_k^2 \\ &= \sum_{i,j=1}^n \frac{1}{4} |u_{i,j} + u_{j,i}|_k^2 \\ &\leq \bar{C} \sum_{i=1}^n |u_i|_{k+1}^2 \\ &= \bar{C} |\mathbf{u}|_{k+1}^2; \end{aligned}$$

¹Note that, from Theorem 2.4 this can be done for all $k > 2$ but from a practical point of view, it not necessary to consider the case k higher than 2.

that is

$$|\varepsilon(\mathbf{u})|_{k, \Omega_e} \leq c |\mathbf{u}|_{k+1}. \quad (3.45)$$

Consequently we have

$$\left| \int_{\Omega_e} \varepsilon(\mathbf{u}) \cdot \mathbf{C}[\gamma_h^e] dx \right| \leq C |\gamma_h^e|_{0, \Omega_e} \sum_{i=1}^k a_{ki}^e |\mathbf{u}|_{i+1, \Omega_e} \quad \forall \mathbf{u} \in [H^{k+1}(\Omega_e)]^n, \gamma_h \in \Gamma^h. \quad (3.46)$$

Therefore we find that

$$\begin{aligned} \Delta &= \sup_{\psi_h \in \Psi^h} \frac{b(\mathbf{u}, \gamma_h)}{\|\psi_h\|_{\Psi^h}} \\ &\leq \sup_{\psi_h \in \Psi^h} \frac{C \sum_{e=1}^E |\gamma_h^e|_{0, \Omega_e} \sum_{i=1}^k a_{ki}^e |\mathbf{u}|_{i+1, \Omega_e}}{\|\psi_h\|_{\Psi^h}} \\ &\leq C \left\{ \sum_{e=1}^E \left(\sum_{i=1}^k a_{ki}^e |\mathbf{u}|_{i+1, \Omega_e} \right)^2 \right\}^{\frac{1}{2}}. \end{aligned}$$

Now

$$\begin{aligned} \sum_{e=1}^E \left(\sum_{i=1}^k a_{ki}^e |\mathbf{u}|_{i+1, \Omega_e} \right)^2 &\leq M \sum_{e=1}^E \sum_{i=1}^k (a_{ki}^e)^2 |\mathbf{u}|_{i+1, \Omega_e}^2 \\ &\leq M \sum_{e=1}^E \max_i (a_{ki}^e)^2 \sum_{i=1}^k |\mathbf{u}|_{i+1, \Omega_e}^2 \\ &\leq M \max_{i,e} (a_{ki}^e)^2 \sum_{i=1}^k |\mathbf{u}|_{i+1, \Omega_e}^2. \end{aligned}$$

So the Lemma follows with $A_k = \max_{i,e} a_{ki}^e$. \square

If we adopt the notion of a regular subdivision of Ω as outlined in the previous Chapter, we find that

$$\Delta \leq C \sup_{\psi_h \in \Psi^h} \frac{\sum_{e=1}^E h_e^k |\gamma_h^e|_{0, \Omega_e} \sum_{i=1}^k |\mathbf{u}|_{i+1, \Omega_e}}{\|\psi_h\|_{\Psi^h}}. \quad (3.47)$$

Finally, in order to simplify this result and obtain an estimate in terms of norms on Ω we make use of the Schwarz inequality and the elementary inequality $(a+b)^2 \leq$

$2(a^2 + b^2)$, to obtain

$$\Delta \leq C \begin{cases} h|\mathbf{u}|_{2,\Omega} & \text{for } k = 1, \\ h^2(|\mathbf{u}|_{2,\Omega} + |\mathbf{u}|_{3,\Omega}) & \text{for } k = 2. \end{cases} \quad (3.48)$$

Together with Lemma 3.1 and 3.2 and Theorem 2.5 we arrive at the following result.

Theorem 3.2 *Let $\mathbf{u} \in V \cap [H^{k+1}(\Omega)]^2$ ($k = 1, 2$) be the unique solution to the problem P and suppose that the spaces V^h and Γ^h satisfy the conditions (I) - (III). Then the problem P^h has a unique solution $\phi_h = (\mathbf{u}_h, \tilde{\epsilon}_h) \in \Psi^h$. Furthermore, if V^h is defined by (3.24) and Γ^h by (3.25) then there exists a positive constant c such that*

$$\|\mathbf{u} - \mathbf{u}_h\|_V + \|\tilde{\epsilon}_h\|_\Gamma \leq ch^k. \quad (3.49)$$

Remark.

Theorem 3.2 may in principle be extended to the three-dimensional case. To do this it would be necessary, though, to obtain estimates of the various norms along the lines of those given in Chapter 2, and to extend in an appropriate way the notion of a regular family of elements.

3.4 Stress Recovery

It remains the issue of determining the stresses, which were eliminated from the problem \bar{P}^h in the previous Section. Therefore a different technique is needed to recover the approximate stresses. In [48] a procedure for the recovery of stresses is proposed; this procedure is interpreted there as a least-squares variational approach. We show here that the method proposed in [48] may be formulated in an alternative way, and that this method leads to an approximation for the stresses which converges at the optimal rate.

Consider the problem of finding $\sigma \in S$ which satisfies

$$r(\sigma, \tau) = s(\mathbf{u}, \tau) \quad \text{for all } \tau \in S, \quad (3.50)$$

where the bilinear forms $r(\cdot, \cdot)$ and $s(\cdot, \cdot)$ are defined by

$$\begin{aligned} r : S \times S &\rightarrow \mathbb{R}, & r(\sigma, \tau) &= \int_{\Omega} \sigma \cdot C^{-1}[\tau] \, d\mathbf{x}, \\ s : V \times S &\rightarrow \mathbb{R}, & s(\mathbf{v}, \tau) &= \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{v}) \cdot \tau \, d\mathbf{x}, \end{aligned}$$

and \mathbf{u} is the solution of Problem P (equations (3.18) and 3.19)). The problem defined by (3.50) is clearly a weak statement of the elastic constitutive relation. It is easy to show that the bilinear form $r(\cdot, \cdot)$ is continuous and S -elliptic, while the linear functional $s(\mathbf{u}, \cdot)$ is continuous. Therefore the problem SP is well-posed and there exists a unique solution $\sigma \in S$.

We turn now to the discrete version of (3.50) and pose the following problem:

Problem SP^h

Given $\mathbf{u}_h \in V^h$, find $\sigma_h \in S_0^h$ such that

$$r(\sigma_h, \tau_h) = s(\mathbf{u}_h, \tau_h) \quad \text{for all } \tau_h \in S^h, \quad (3.51)$$

where S_0^h is the subspace of the standard interpolation subspace S^h which is L_2 -orthogonal to Γ^h . Moreover \mathbf{u}_h is the unique solution to the problem P^h . This is the equation used in [48] to determine the stresses, after displacements have been calculated. The authors of that work arrive at (3.51) as one of a pair of equations obtained by minimizing the least-square functional:

$$L(\tilde{\boldsymbol{\varepsilon}}_h, \sigma_h) := \int_{\Omega} (C[\boldsymbol{\varepsilon}(\mathbf{u}_h) + \tilde{\boldsymbol{\varepsilon}}_h] - \sigma_h) C^{-1} (C[\boldsymbol{\varepsilon}(\mathbf{u}_h) + \tilde{\boldsymbol{\varepsilon}}_h] - \sigma_h) \, d\mathbf{x}$$

where \mathbf{u}_h is given.

Alternatively, (3.51) can be obtained by considering the discrete constitutive equation (3.2) in the form

$$\tilde{\epsilon}_h = C^{-1}\sigma_h - \epsilon(u_h).$$

Substitution of this into (3.27) gives

$$(\tau_h, C^{-1}\sigma_h - \epsilon(u_h))_0 = 0$$

or

$$(\tau_h, C^{-1}\sigma_h)_0 = (\tau_h, \epsilon(u_h))_0,$$

which tells us that the above discrete problem may be obtained from (3.27).

Clearly the discrete problem (3.51) is well-posed and hence it has a unique solution $\sigma_h \in S_0^h$, which depends continuously on u_h .

Our goal is now to prove that the stress obtained from (3.51) converges to the exact solution. In order to do so, we first note, from (3.50) and (3.51) that

$$r(\sigma - \sigma_h, \tau_h) = s(u - u_h, \tau_h). \quad (3.52)$$

Thus

$$\begin{aligned} r(\sigma_h - \tau_h, \sigma_h - \tau_h) &= r(\sigma - \overset{\tau_h}{\sigma}_h, \sigma_h - \tau_h) + r(\sigma_h - \sigma, \sigma_h - \tau_h) \text{?} \\ &= r(\sigma - \overset{\tau_h}{\sigma}_h, \sigma_h - \tau_h) + s(u_h - u, \sigma_h - \tau_h) \end{aligned}$$

and so, using the continuity and S -ellipticity of $r(\cdot, \cdot)$ and the continuity of $s(\cdot, \cdot)$ (which follows from the properties of C^{-1}), we find that

$$\|\sigma_h - \tau_h\|_S \leq C(\|\sigma - \tau_h\|_S + \|u - u_h\|_V).$$

Using the triangle inequality we therefore find that

$$\|\sigma - \sigma_h\|_S \leq C \left\{ \inf_{\tau_h \in S^h} \|\sigma - \tau_h\|_S + \|u - u_h\|_V \right\}. \quad (3.53)$$

Thus we have the following estimate from the best approximation property and Theorem 2.5.

Theorem 3.3 *Assume that all the conditions of Theorem 3.1 hold, and assume that the approximate stress is obtained from (3.51). Then if $S^h|_{\Omega_e}$ contains polynomials of degree $k - 1$ where k is defined in (3.24), there exists a constant $C > 0$ such that*

$$\|\sigma - \sigma_h\|_S \leq Ch^k \quad (k = 1, 2). \quad (3.54)$$

Remarks

1. It is interesting to note that, although the bilinear form $s(\cdot, \cdot)$ of the original mixed variational Problem \bar{P}^h appears, no inf-sup condition is required in order to show stability.
2. From Theorem 3.3 we see that it suffices to choose a basis of piecewise constants for the stresses, when the displacements are approximated by bilinear functions. This notwithstanding, it may well be the case that bases for S^h which contain linear terms will give better approximations, particularly in the coarse-mesh regime (note that Theorem 3.3 gives an asymptotic result). In the following Chapter we will present results using the linear basis proposed in [39], and compare these with results obtained using constant terms only.
3. In practice the space S^h will be chosen in such a way that it conforms with the orthogonality requirement of assumption (I). This requirement can be (and is) imposed on the reference element, since we have

$$\int_{\Omega_e} \tau_h^e \cdot \gamma_h^e dx = J_0^e \int_{\hat{\Omega}} \hat{\tau}_h \cdot \hat{\gamma}_h d\xi$$

for all $\boldsymbol{\tau}_h \in S^h$ and $\boldsymbol{\gamma}_h \in \Gamma^h$, bearing in mind the transformation rule in (3.25) and the fact that members of S^h are assumed to transform in conventional way, that is, according to $\boldsymbol{\tau}_h(\boldsymbol{x}) = \hat{\boldsymbol{\tau}}_h(\boldsymbol{\xi})$.

4. It is worth emphasising here that (3.51) allows the stresses to be obtained element by element, since the members of both S^h and Γ^h may be constructed in such a way that they are discontinuous across element boundaries.

Chapter 4

Numerical Results

In this Section, we present a selection of numerical results for two-dimensional problems, with a view to highlighting certain aspects of computational behaviour in practice, and also to comparing computational results against the estimates of Sections 3.2 and 3.3. We focus on four-noded quadrilateral elements. The results are obtained by using PCFEAP.

The choice of a basis for the space Γ^h is not unique, and a crucial factor in the enhanced strain method is to choose one which is optimal in some sense. In [48] a five-dimensional local basis for $\hat{\gamma}_h$ was introduced, while a seven-dimensional local basis was introduced in [2]. In both articles, $\hat{\gamma}_h$ is defined according to (3.33). We propose here a new basis, which will be seen to give promising results. In the next Section we derive this basis.

4.1 A new basis for enhanced strains

The point of departure is an approach closely allied to that for the original Wilson incompatible modes element ([56],[59]; see [28] for a review of the details). We are concerned here with the four-noded bilinear element in the context of plane elasticity,

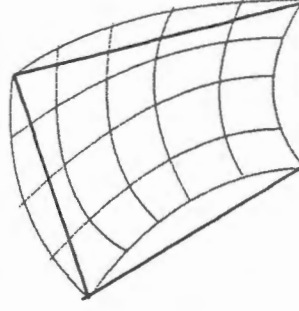


Figure 4.1: Roof function

for which the transformation (2.31) can be written in the form

$$\begin{aligned} x &= a_0 + a_1\xi + a_2\eta + a_3\xi\eta, \\ y &= b_0 + b_1\xi + b_2\eta + b_3\xi\eta. \end{aligned} \quad (4.1)$$

We set $\mathbf{A}_e = \frac{\partial \mathbf{F}_e}{\partial \boldsymbol{\xi}}$ and note the elementary result that

$$\mathbf{A}_e^{-T} = \mathbf{J}_e^{-1} \begin{pmatrix} \frac{\partial y}{\partial \eta} & -\frac{\partial y}{\partial \xi} \\ -\frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \xi} \end{pmatrix}. \quad (4.2)$$

For convenience we henceforth drop the subscript e in this Section. Now suppose that the displacement field \mathbf{v}_h is subdivided into two parts: a compatible part, denoted by \mathbf{v}_h^c , which belongs to V^h , and an incompatible part \mathbf{v}_h^i , which on the reference element is defined by

$$\mathbf{v}_h^i(\mathbf{x}(\boldsymbol{\xi})) = \boldsymbol{\alpha} \left\{ 1 - \frac{1}{2}(\xi^2 + \eta^2) \right\}, \quad (4.3)$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^T \in \mathbb{R}^2$. We are thus adding to the conventional conforming approximation a *roof-like function* (see Figure 4.1) with only two degrees of freedom, in contrast to the four-degree-of-freedom Wilson model [59].

We may associate with the displacement field (4.3) the strain function $\gamma_h = (\gamma_{h,11}, \gamma_{h,22}, 2\gamma_{h,12})^T$ according to

$$\gamma_h(\mathbf{v}_h^i) = \mathbf{D}\mathbf{v}_h^i, \quad (4.4)$$

where \mathbf{D} is the differential matrix operator defined by

$$\mathbf{D} = \begin{pmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{pmatrix}. \quad (4.5)$$

In view of the identity (4.2), the transformation rule

$$\frac{\partial}{\partial \mathbf{x}} = \mathbf{A}^{-T} \frac{\partial}{\partial \boldsymbol{\xi}},$$

and (4.4) we write

$$\gamma_h(\boldsymbol{\xi}) = \mathbf{J}^{-1} \hat{\mathbf{D}}\mathbf{v}_h^i \quad \text{where} \quad \hat{\mathbf{D}} = \begin{pmatrix} y_\eta \frac{\partial}{\partial \xi} - y_\xi \frac{\partial}{\partial \eta} & 0 \\ 0 & -x_\eta \frac{\partial}{\partial \xi} + x_\xi \frac{\partial}{\partial \eta} \\ -x_\eta \frac{\partial}{\partial \xi} + x_\xi \frac{\partial}{\partial \eta} & y_\eta \frac{\partial}{\partial \xi} - y_\xi \frac{\partial}{\partial \eta} \end{pmatrix}$$

and x_ξ, \dots, y_η are partial derivatives of (x, y) with respect to (ξ, η) . To obtain the function $\hat{\gamma}_h$ on the reference element we use the definition (3.25) to write

$$\begin{aligned} \hat{\gamma}_h &= \frac{J}{J_0} \gamma_h \\ &= \frac{1}{J_0} \hat{\mathbf{D}}\mathbf{v}_h^i \\ &= \frac{1}{J_0} \begin{pmatrix} -\alpha_1 \xi (b_2 + b_3 \xi) + \alpha_1 \eta (b_1 + b_3 \eta) \\ \alpha_2 \xi (a_2 + a_3 \xi) - \alpha_2 \eta (a_1 + a_3 \eta) \\ -\alpha_2 \xi (b_2 + b_3 \xi) + \alpha_2 \eta (b_1 + b_3 \eta) + \alpha_1 \xi (a_2 + a_3 \xi) - \alpha_1 \eta (a_1 + a_3 \eta) \end{pmatrix}, \end{aligned}$$

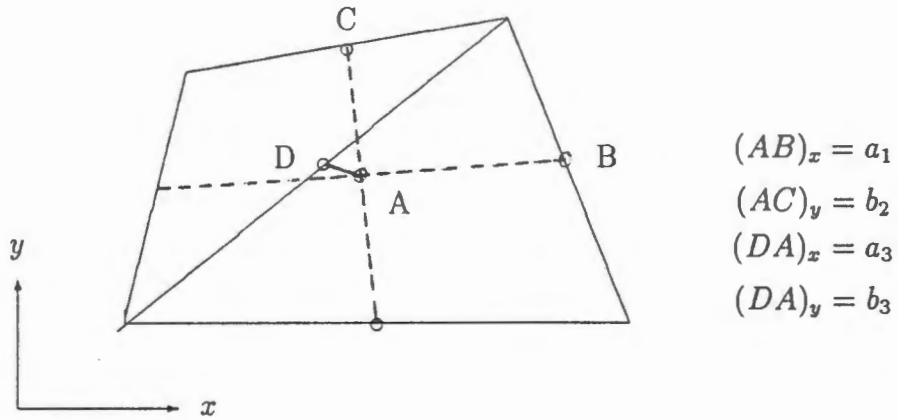


Figure 4.2: Identification of a_3 and b_3 on a quadrilateral

which can be rearranged in matrix form as

$$\hat{\gamma}_h = -\frac{1}{J_0} \begin{pmatrix} \xi - \frac{b_1}{b_2}\eta & 0 & 0 & 0 & \xi^2 - \eta^2 & 0 & 0 \\ 0 & \eta - \frac{a_2}{a_1}\xi & 0 & 0 & 0 & \xi^2 - \eta^2 & 0 \\ 0 & 0 & \xi - \frac{b_1}{b_2}\eta & \eta - \frac{a_2}{a_1}\xi & 0 & 0 & \xi^2 - \eta^2 \end{pmatrix} \times \begin{bmatrix} \alpha_1 b_2 \\ \alpha_2 a_1 \\ \alpha_2 b_2 \\ \alpha_1 a_1 \\ \alpha_1 b_3 \\ -\alpha_2 a_3 \\ \alpha_2 b_3 - \alpha_1 a_3 \end{bmatrix}. \quad (4.6)$$

We note (see Figure 4.2) that a_3 and b_3 are the x and y distances from the centroid of the quadrilateral to the midpoint of any diagonal of the quadrilateral, and are therefore smaller than the parameters a_1 and b_2 (see remarks in Section 2.4). Also, we note that a_3 and b_3 become zero for the case of parallelograms. With this in mind, we can discard the last three columns of the vector containing these parameters. Hence we define (on the reference element) the first four columns of matrix (4.6) as

a local basis for Γ^h , that is;

$$\hat{\Gamma} = \text{span} \begin{pmatrix} \xi - \frac{b_1}{b_2}\eta & 0 & 0 & 0 \\ 0 & \eta - \frac{a_2}{a_1}\xi & 0 & 0 \\ 0 & 0 & \xi - \frac{b_1}{b_2}\eta & \eta - \frac{a_2}{a_1}\xi \end{pmatrix}. \quad (4.7)$$

It is worth comparing this basis with the five-dimensional basis

$$\hat{\Gamma} = \text{span} \begin{pmatrix} \xi & 0 & 0 & 0 & \xi\eta \\ 0 & \eta & 0 & 0 & -\xi\eta \\ 0 & 0 & \xi & \eta & \xi^2 - \eta^2 \end{pmatrix} \quad (4.8)$$

obtained in [48] (Section 4.2.1); for rectangular elements, in which case $b_1 = a_2 = 0$ and $a_3 = b_3 = 0$, the first four columns of these bases coincide. It was observed that in [48] that the fifth column of (4.8) has a negligible effect on results.

In the previous Chapter we eliminated the stress by choosing the space S^h to be L_2 -orthogonal to the space of enhanced strains Γ^h . In accordance with this assumption, we choose the following local basis for S^h , bearing in mind that members of S^h are transformed in the conventional way, that is, according to $\tau_h(\mathbf{x}) = \hat{\tau}_h(\xi)$:

$$S^h|_{\hat{\Omega}} \equiv \hat{S} = \text{span} \begin{pmatrix} 1 & 0 & 0 & \eta + \frac{b_1}{b_2}\xi & 0 \\ 0 & 1 & 0 & 0 & \xi + \frac{a_2}{a_1}\eta \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}. \quad (4.9)$$

We observe that for rectangular meshes, this coincides with the basis

$$\hat{S} = \text{span} \begin{pmatrix} 1 & 0 & 0 & \eta & 0 \\ 0 & 1 & 0 & 0 & \xi \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (4.10)$$

introduced by Pian and Sumihara [39].

These bases will be used, along with alternatives proposed by other authors, in the numerical study carried out in the following Section.

4.2 Numerical Experiments

We discuss here a selection of numerical results, with a view to examining actual rates of convergence, and comparing the performance of the basis introduced in Section 6.1 with existing examples. All examples are two-dimensional, and conditions of isotropy and plane stress are assumed except for the Example 3. Some of the problems included are often used as benchmarks for the numerical behaviour of quadrilateral elements.

At the element level we write

$$\mathbf{u}_h^e(\boldsymbol{\xi}) = \sum_{A=1}^4 \mathbf{u}_A^e \hat{N}_A(\boldsymbol{\xi}) = \hat{\mathbf{N}}^T \mathbf{d}^e$$

where $\hat{N}_A(\boldsymbol{\xi}) = N_A^e(\mathbf{x}(\boldsymbol{\xi}))$ are basis functions defined according to (2.31) and $\hat{\mathbf{N}}^T = (\hat{N}_1, \hat{N}_1, \hat{N}_2, \dots, \hat{N}_4)_{1 \times 8}$. The column vector $\mathbf{d}^e \in \mathbb{R}^8$ is the vector of element nodal displacements. Consequently, the strain vector $\boldsymbol{\varepsilon}(\mathbf{u}_h)$ on element Ω_e becomes

$$\begin{aligned} \boldsymbol{\varepsilon}(\mathbf{u}_h^e) &= \frac{1}{2}(\nabla \mathbf{u}_h^e + \nabla^T \mathbf{u}_h^e) \\ &= \mathbf{B}^e(\boldsymbol{\xi}) \mathbf{d}^e \end{aligned} \tag{4.11}$$

where $\mathbf{B}^e = (\mathbf{B}_1^e, \dots, \mathbf{B}_4^e)$; here

$$\mathbf{B}_A^e = \begin{pmatrix} \frac{\partial \hat{N}_A}{\partial x} & 0 \\ 0 & \frac{\partial \hat{N}_A}{\partial y} \\ \frac{\partial \hat{N}_A}{\partial y} & \frac{\partial \hat{N}_A}{\partial x} \end{pmatrix}.$$

Interpolation of the enhanced strain differs from the standard isoparametric interpolation procedures. In [2, 48] transformation from the master square to the actual element is defined as in (3.33), that is,

$$\boldsymbol{\gamma}_h^e(\boldsymbol{\xi}) = \frac{J_e^c}{J_e^e} \mathbf{T}_0^{-T} \mathbf{G}^e(\boldsymbol{\xi}) \boldsymbol{\alpha}^e, \tag{4.12}$$

where $\alpha^e \in \mathbb{R}^m$ ($m = 4$ in [48] and $m = 7$ in [2]) is the vector of element degrees of freedom for enhanced strain, T_0 is the matrix as defined in (3.34) and $G^e(\xi)$ is the matrix given by

$$G^e(\xi) = \begin{cases} \begin{pmatrix} \xi & 0 & 0 & 0 \\ 0 & \eta & 0 & 0 \\ 0 & 0 & \xi & \eta \end{pmatrix} & \text{in [48]} \\ \begin{pmatrix} \xi & 0 & 0 & 0 & \xi\eta & 0 & 0 \\ 0 & \eta & 0 & 0 & 0 & \xi\eta & 0 \\ 0 & 0 & \xi & \eta & 0 & 0 & \xi\eta \end{pmatrix} & \text{in [2]} \end{cases} \quad (4.13)$$

We also make use of the basis (4.7), with γ_h^e related to $\hat{\gamma}$ by (4.12), but without the presence of T_0 . From a computational point of view, such a rule has some advantages over that introduced in [48] since multiplication with the matrix T_0^{-1} makes all elements of the resulting matrix on the right hand side of (4.12) nonzero, and hence these interpolations require more computer storage and time.

Substitution of the preceding interpolations into equations (3.30) and (3.31) yields the discrete linear system of equations

$$\begin{aligned} \mathring{A}_{e=1}^E \{K_e d^e + H_e \alpha^e\} &= \mathring{A}_{e=1}^E F^e \\ H_e^T d^e + M_e \alpha^e &= 0, \quad (e = 1, 2, \dots, E) \end{aligned}$$

where $\mathring{A}_{e=1}^E$ denotes the standard assembly operator, and element matrices K_e , H_e and M_e are given by

$$\begin{aligned} K_e &= \int_{\hat{\Omega}} J^e B^{eT} C B^e d\xi \\ H_e &= \int_{\hat{\Omega}} J_0 B^{eT} C G^e d\xi \quad \left(= \int_{\hat{\Omega}} J_0 B^{eT} C T_0^{-T} G^e d\xi \text{ in [48, 2]} \right) \\ M_e &= \int_{\hat{\Omega}} \frac{J_0^2}{J^e} G^{eT} C G^e d\xi \quad \left(= \int_{\hat{\Omega}} \frac{J_0^2}{J^e} G^{eT} T_0^{-1} \underline{C} T_0^{-T} G^e d\xi \text{ in [48, 2]} \right), \end{aligned}$$

and the generalized element force vector is given by

$$F_{Ai}^e = \int_{\hat{\Omega}} J^e N_A f_i^e d\xi, \text{ for } i = 1, 2.$$

For homogeneous isotropic elasticity, the elasticity matrix \mathbf{C} is given by

$$\mathbf{C} = \begin{pmatrix} \bar{\lambda} + 2\mu & \bar{\lambda} & 0 \\ \bar{\lambda} & \bar{\lambda} + 2\mu & 0 \\ 0 & 0 & \mu \end{pmatrix},$$

where

$$\bar{\lambda} = \begin{cases} \frac{(1-2\nu)}{(1-\nu)}\lambda & \text{for plane stress} \\ \lambda & \text{for plane strain;} \end{cases}$$

here λ and μ are the *Lamé parameters*. The relationships of λ and μ to E , *Young's modulus*, and ν , *Poisson's ratio*, are given by

$$\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)} \text{ and } \mu = \frac{E}{2(1+\nu)}.$$

We note that \mathbf{M}_e is positive definite since \mathbf{C} is positive-definite. Hence, since α^e is discontinuous across the element boundaries, in other words α^e is independent from the element to element, α^e can be eliminated at the element level.

Example 1

We begin by considering results for the test problem of a tapered cantilever, clamped at one end and subjected to a shearing load on the other, as shown in Figure 4.3. We subdivide the mesh into $n \times n$ elements.

This problem is regarded as a good test of the performance of distorted elements in a bending dominated problem, and is commonly known as the Cook membrane problem (see [19]). The aim of this example is to illustrate the superior coarse mesh accuracy achieved with the various enhanced strain or assumed stress elements.

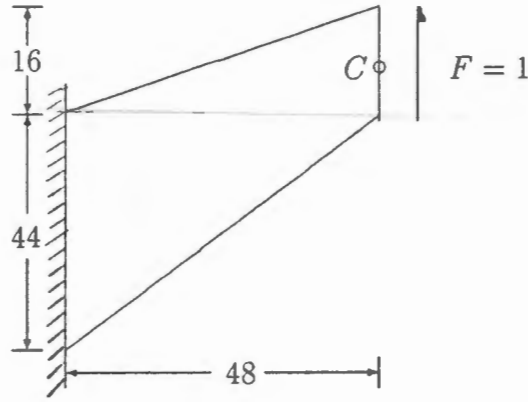


Figure 4.3: Tapered beam subject to the end shear load

Figure 4.4 shows a graph of the vertical tip deflection at the point C obtained using the standard four-node element, and those obtained using the enhanced strain bases of [48] and [2], as well as the basis (4.8). Also shown are results for the assumed stress element of Pian and Sumihara [39]. The values of the material constants used here are Young's modulus $E = 1$ and Poisson's ratio $\nu = 1/3$. The superior accuracy of the enhanced strain elements is quite plain; these all perform well, with the basis (4.8) giving results which are marginally better than those for the other enhanced elements.

Differences in performance between the various enhanced strain elements are more readily appreciated by considering a plot showing the rates of convergence. Figure 4.5 shows this, as a function of mesh size, for the Cook membrane problem. No exact solution is available, so we compare the numerically obtained values of the vertical displacement at C against those obtained using a mesh of 128×128 elements. Here and in the figures that follow, straight lines depict least squares fits to numerical data.

No estimates have been derived for pointwise error estimates, but since error estimates in the energy norm coincide with those for the standard element, one might expect the same to be true of pointwise errors, in which case the order of convergence would be $h^2 |\ln h|^{3/2}$ (see [18], Theorem 3.3.7). In practice one thus has an $O(h^{2-\epsilon})$

convergence for any $\epsilon > 0$. Let assume the error estimate of the form

$$\sup_{\mathbf{x} \in \Omega} |\mathbf{u}(\mathbf{x}) - \mathbf{u}_h(\mathbf{x})| \leq Ch^p$$

for all bases, and take logarithms of both sides to obtain

$$\log|\mathbf{u} - \mathbf{u}_h| \simeq \log C + p \log h. \quad (4.14)$$

Therefore when we plot a log-log graph of $|\mathbf{u}(\mathbf{x}) - \mathbf{u}_h(\mathbf{x})|$ against h in Figure 4.5, the intercept of the line gives an estimate of $\log C$. This graph shows that all elements, including the standard element, yield a rate of convergence of around 1.6. What is interesting, though, and what distinguishes the enhanced strain elements from the standard case, is not the rate of convergence, but rather the constant C appearing in the error estimate; this accounts for the significant difference in the vertical translation of the graphs, between the standard and enhanced cases. Denoting by C_{ST} , C_{SR} and C_{AR} the constants associated with the standard element, the enhanced strain element in [48], and that corresponding to (4.8), respectively, we find that

$$\frac{C_{ST}}{C_{SR}} = 7.4 \quad \text{and} \quad \frac{C_{ST}}{C_{AR}} = 11.7. \quad (4.15)$$

Thus while all three enhanced strain elements render superior results, the basis introduced here gives results which improve on those obtained using the bases of [48] and [2], through a smaller constant in the error estimate.

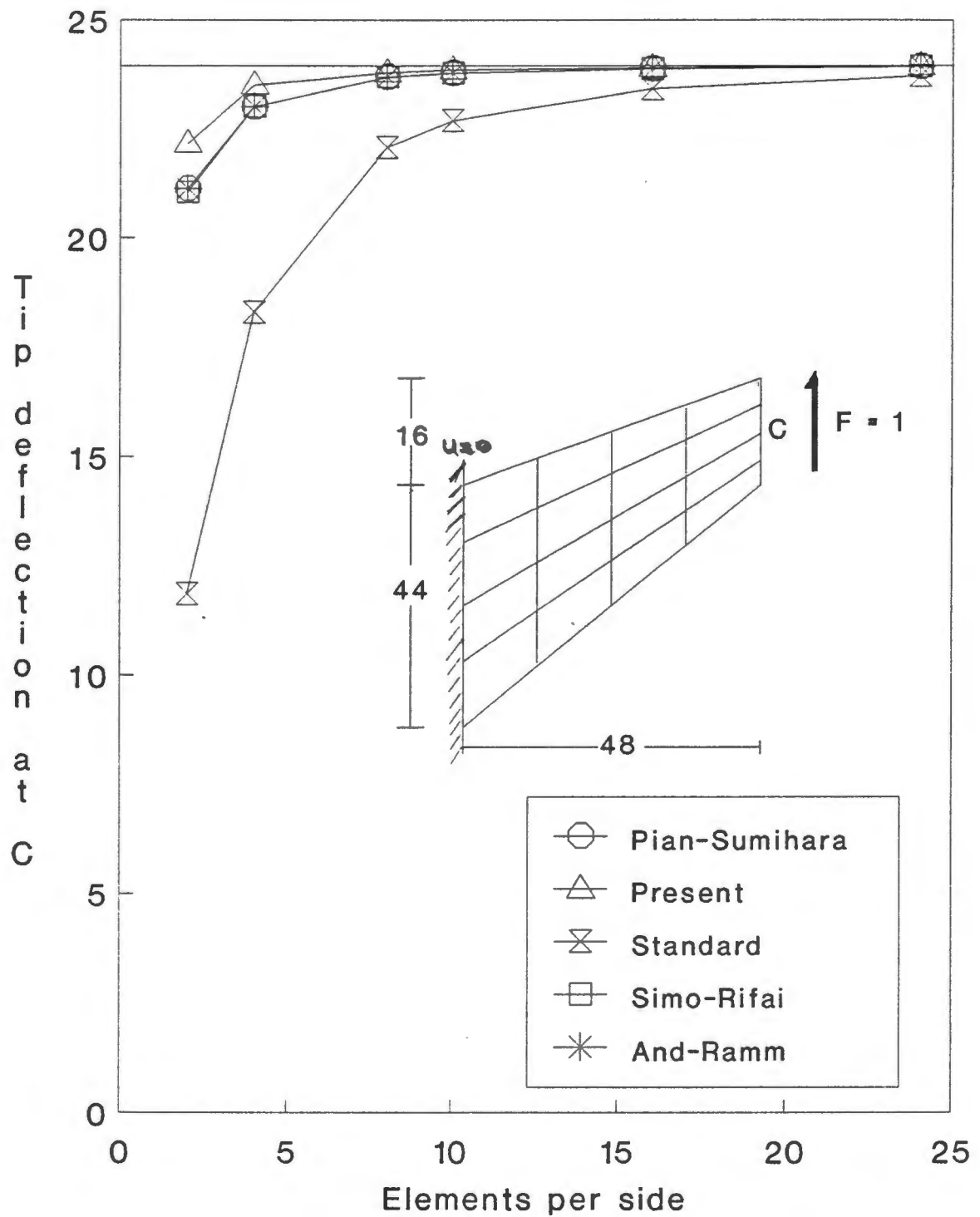


Figure 4.4: Results for the Cook membrane problem

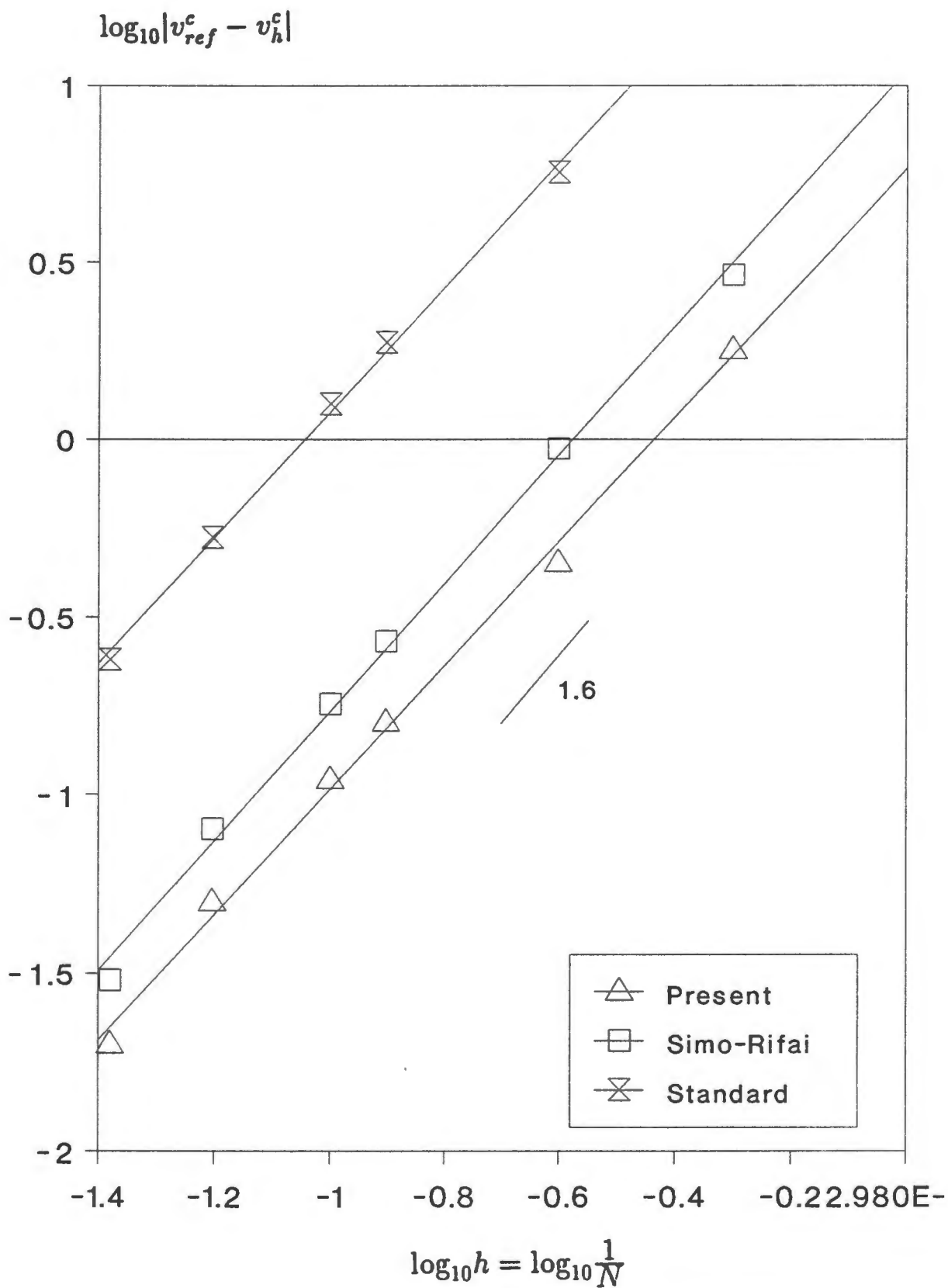


Figure 4.5: Error plot for the Cook membrane problem

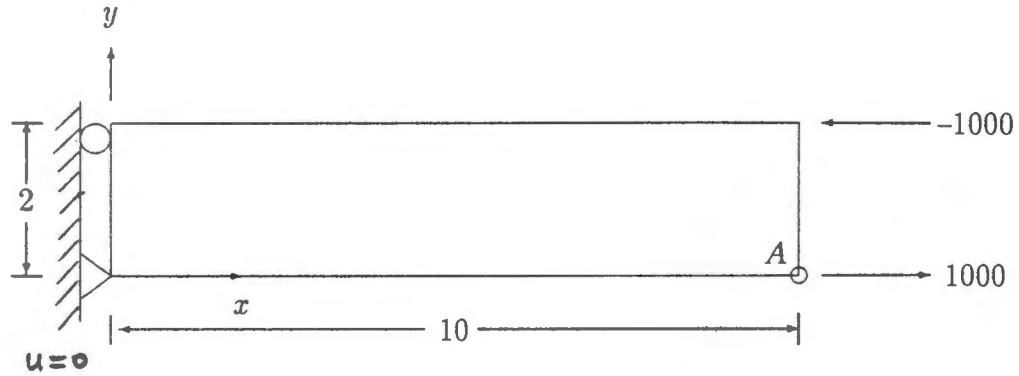


Figure 4.6: Cantilever subject to a couple

Example 2

We give this example to illustrate rates of convergence in the energy norm (*cf.* Theorem 3.2). The test example is a linear, elastic beam, fixed on one end and subjected to a couple on the other end, as shown in Figure 4.6.

With Young's modulus $E = 1500$ and Poisson's ratio $\nu = 0.25$, exact solutions for the displacements are

$$u = 2x(1 - y); \quad v = \frac{1}{4}(y^2 - 2y) + x^2,$$

when referred to the axes shown in Figure 4.6. The corresponding stress field is

$$\sigma_x = 3000(1 - y), \quad \sigma_y = 0, \quad \tau_{xy} = 0.$$

For uniform meshes of rectangular elements, a log-log plot of $(\|u - u_h\|_V + \|\tilde{\epsilon}_h\|_\Gamma)$ against h is shown in Figure 4.7. Results also shown for standard method. All the enhanced methods give the same results, which correspond to a rate of convergence slightly higher (1.2 as against 1.0) than that predicted in Chapter 3, while the standard method gives the rate almost 2. Again we observe the significant difference in the intercepts.

Figure 4.8 shows results obtained for the same problem, but this time for progressive refinements of the mesh of distorted elements shown. Behaviour similar to that depicted in Figure 4.5 is observed here; that is, all of the enhanced elements have approximately the same rate of convergence, approximately 1.4, which exceeds the theoretical rate of 1.0; however, the constant in (3.49) differs from one basis to the next; whereas the bases of [2, 48] appear to have the same constant, that corresponding to (4.8) is lower by a factor of 1.3.

We consider next the stress calculation. In [48] the transformation rule from the reference to the actual element is defined as

$$\sigma_h^e(\xi) = T_0 \hat{S} \beta, \quad (4.16)$$

where \hat{S} is given by (4.11) and $\beta \in \mathbb{R}^5$. We use

$$\sigma_h^e(\xi) = \hat{S} \beta, \quad (4.17)$$

where \hat{S} is given by (4.10) or (4.11) and $\beta \in \mathbb{R}^5$.

The same test example is used to study the stress error, for which an estimate is given in Theorem 3.3. Note that the stress error is zero for uniform meshes. This is expected since the stress is linear here. For distorted meshes we plot the error of the stress against the mesh parameter in Figure 4.9. Again the results corresponding to (4.10) are better. We also plot the stress error with constant stress interpolation on the same graph. We obtain the rate of convergence at the rate predicted by Theorem. The rate of convergence doesn't change for both linear and constant interpolation, but there is a difference in the intercepts.

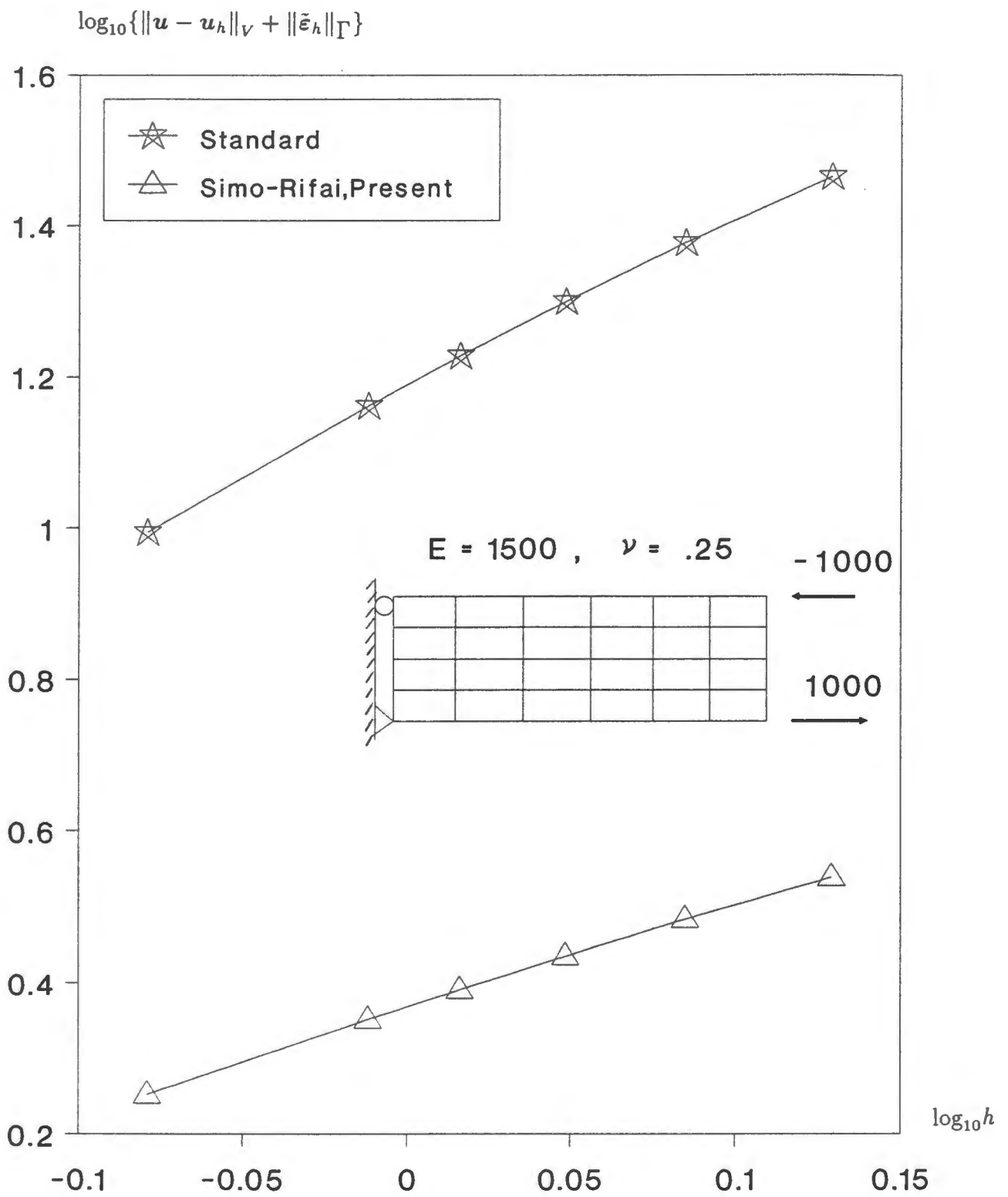


Figure 4.7: Plot of error in energy norm for a uniform mesh of rectangular elements

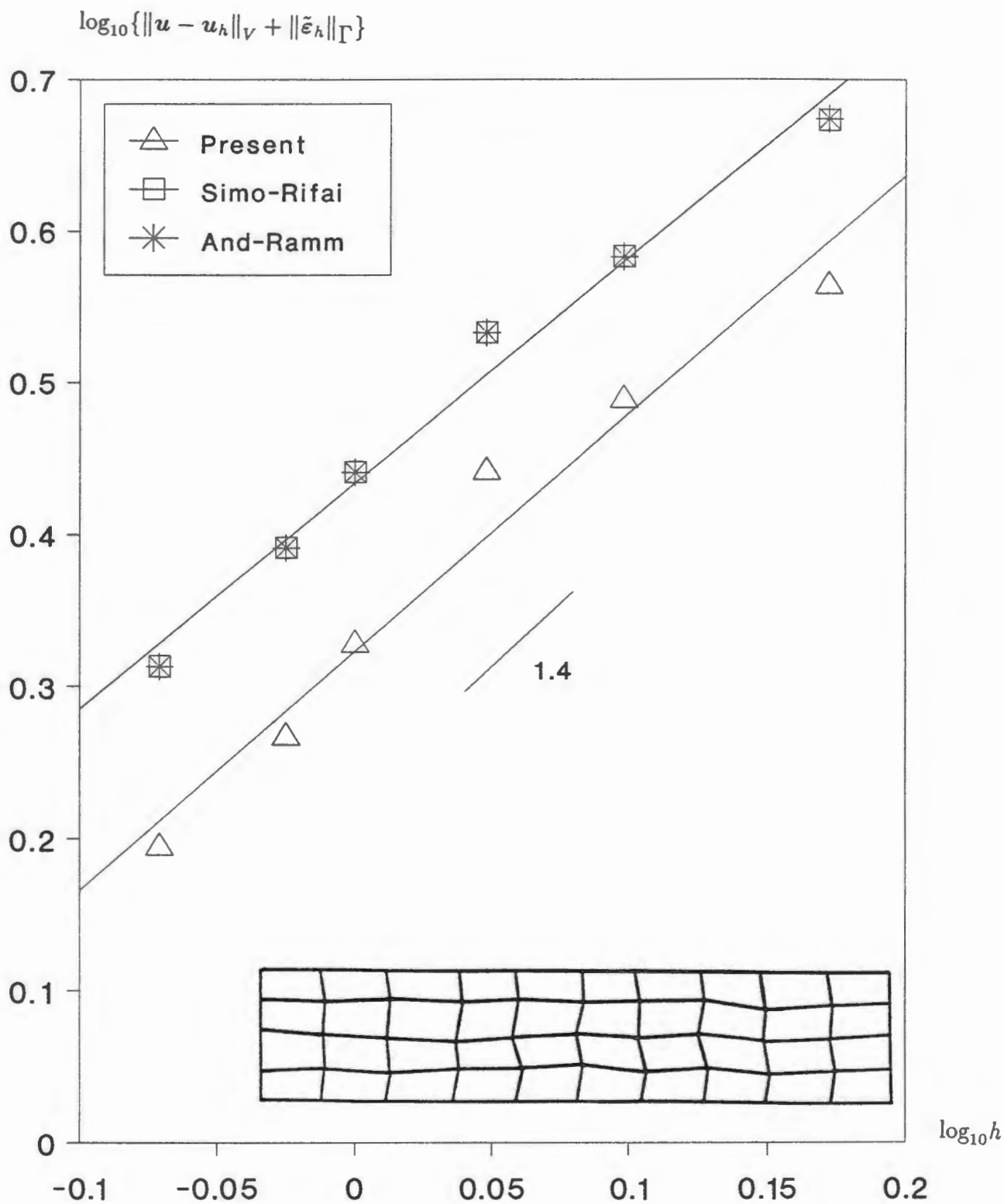


Figure 4.8: Plot of error in energy norm for meshes of distorted quadrilateral elements

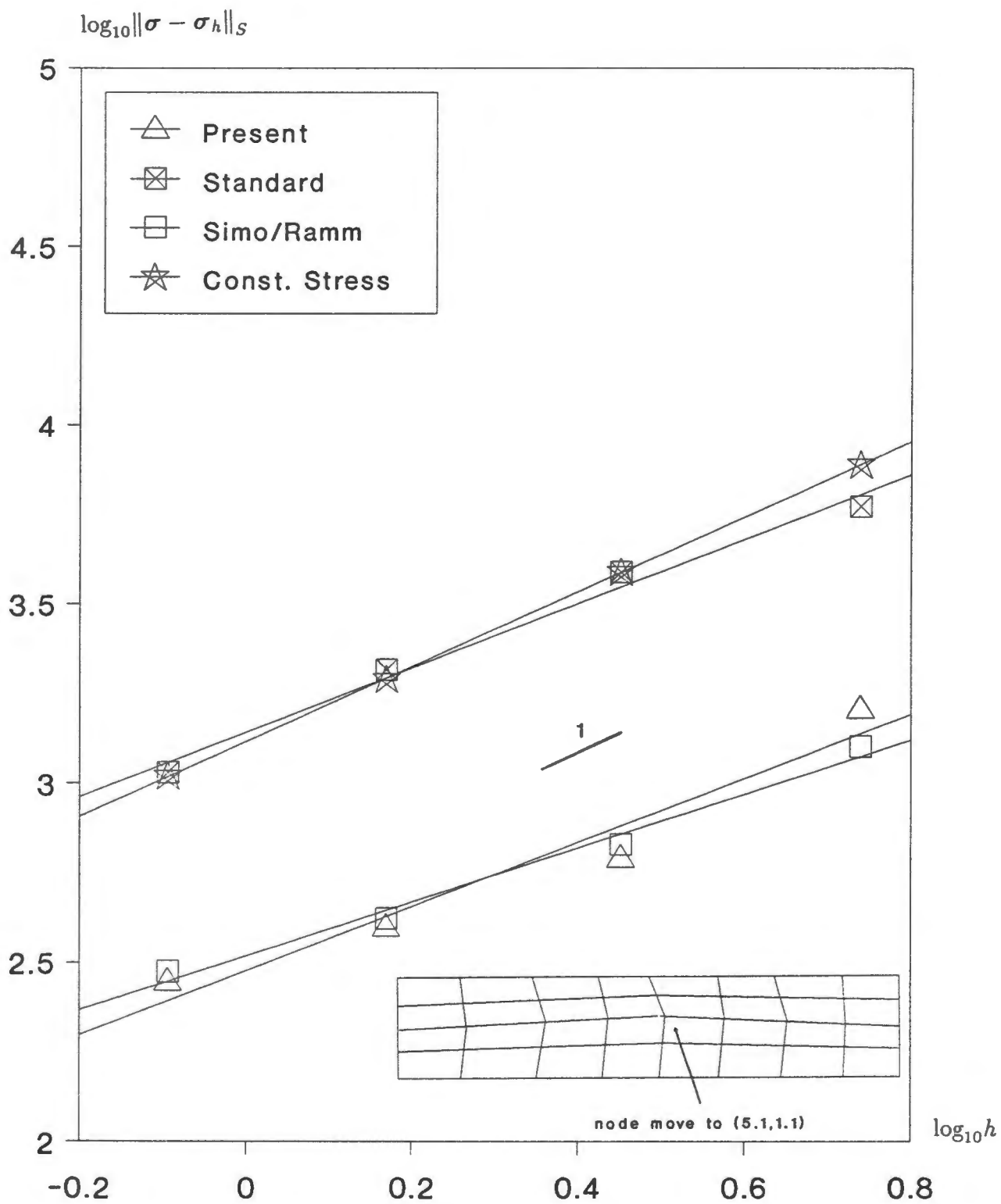


Figure 4.9: Stress error for meshes of distorted quadrilateral elements

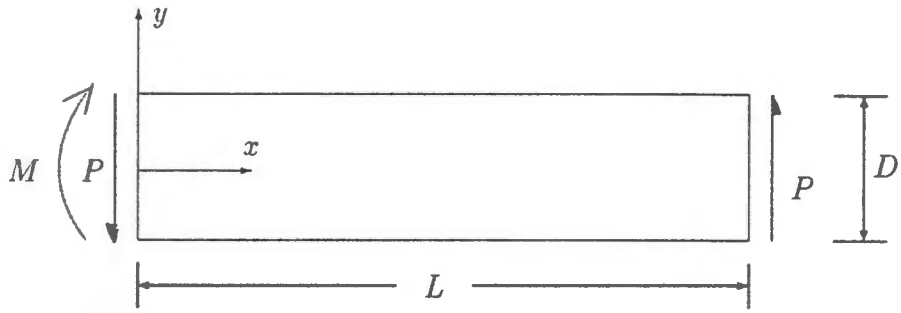


Figure 4.10: Cantilever beam subject to shearing load

Example 3

A third example (given in [9]) is also a linear, elastic cantilever with shearing load at its end as shown in Figure 4.10. M and P are reactions at the support which has placed at the left end. The analytical solution is [58]

$$u = -\frac{Py}{6EI} \left\{ (6L - 3x)x + (2 + \bar{\nu})(y^2 - \frac{1}{4}D^2) \right\},$$

$$v = \frac{P}{6EI} \left\{ 3\bar{\nu}y^2(L - x) + \frac{1}{4}(4 + 5\bar{\nu})D^2x + (3L - x)x^2 \right\},$$

where $I = \frac{1}{12}D^3$,

$$\bar{E} = \begin{cases} E & \text{for plane stress,} \\ E/(1 - \nu^2) & \text{for plane strain;} \end{cases}$$

$$\bar{\nu} = \begin{cases} \nu & \text{for plane stress,} \\ \nu/(1 - \nu) & \text{for plane strain.} \end{cases}$$

Reaction forces are applied at the support based on the stresses which are

$$\sigma_x = -\frac{Py}{I}(L - x), \quad \sigma_y = 0, \quad \tau_{xy} = \frac{P}{2I}(\frac{1}{4}D^2 - y^2).$$

We modelled only the top half of the cantilever since the problem is antisymmetric, and used the following isotropic elastic materials:

- (1) Plane stress, $\nu = .25$,

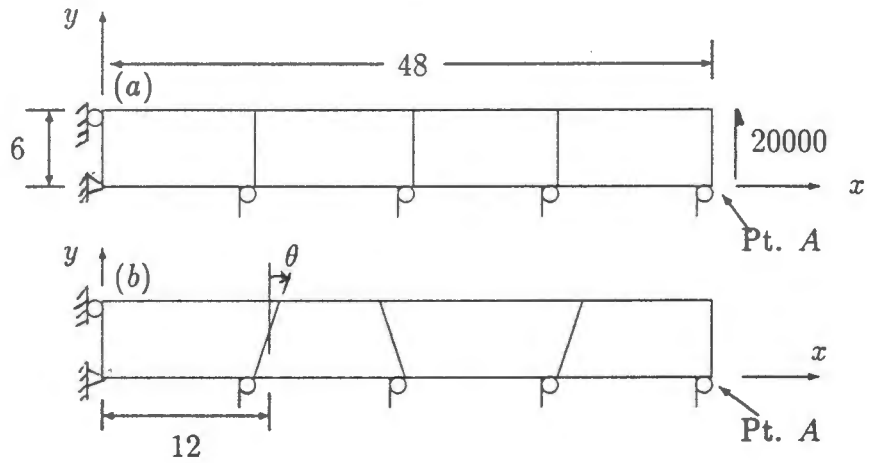


Figure 4.11: (a). Top half of antisymmetric beam mesh. (b) Distorted mesh with $\theta = 9.462^\circ$.

(2) Plane strain, $\nu = .4999$ (near-incompressibility).

The first test to this example is to assess the coarse mesh accuracy of the methods. The normalized end displacements at A are shown in Table 4.1 for rectangular and skewed meshes (see Figure 4.11). For rectangular elements all enhanced strain elements produce the same results. For distorted elements (Figure 4.11 (b)), other enhanced strain elements and Pian-Sumihara's element are slightly better than our element, but the difference is minor. For the plane strain case, the standard method produces very poor results, as expected.

Material	Standard	Present/Pian-S
1	.709	.988
2	.061	.989

Table 4.1 (a): $v^{FEM}/v^{Anal.}$ at point A of mesh in Fig. 4.6(a) (rectangular elements)

Material	Standard	Present	Simo/Pian-S
1	.690	.949	.956
2	.061	.962	.965

Table 4.1 (b): $v^{FEM}/v^{Anal.}$ at point A of mesh in Fig. 4.6(b) (skewed elements)

Using the rectangular meshes we also obtained the errors to study convergence. For plane stress, displacement and stress errors are plotted in Figures 4.12 and 4.13. All methods behave like the previous example, but there is a big difference in intercept, which makes enhanced strain elements superior. For plane strain with $\nu = 0.4999$ (near-incompressibility), displacement and stress errors are shown in Figures 4.14 and 4.15. For the standard case locking is observed, while the enhanced strain element converges almost at the same rate as in the case of plane stress.

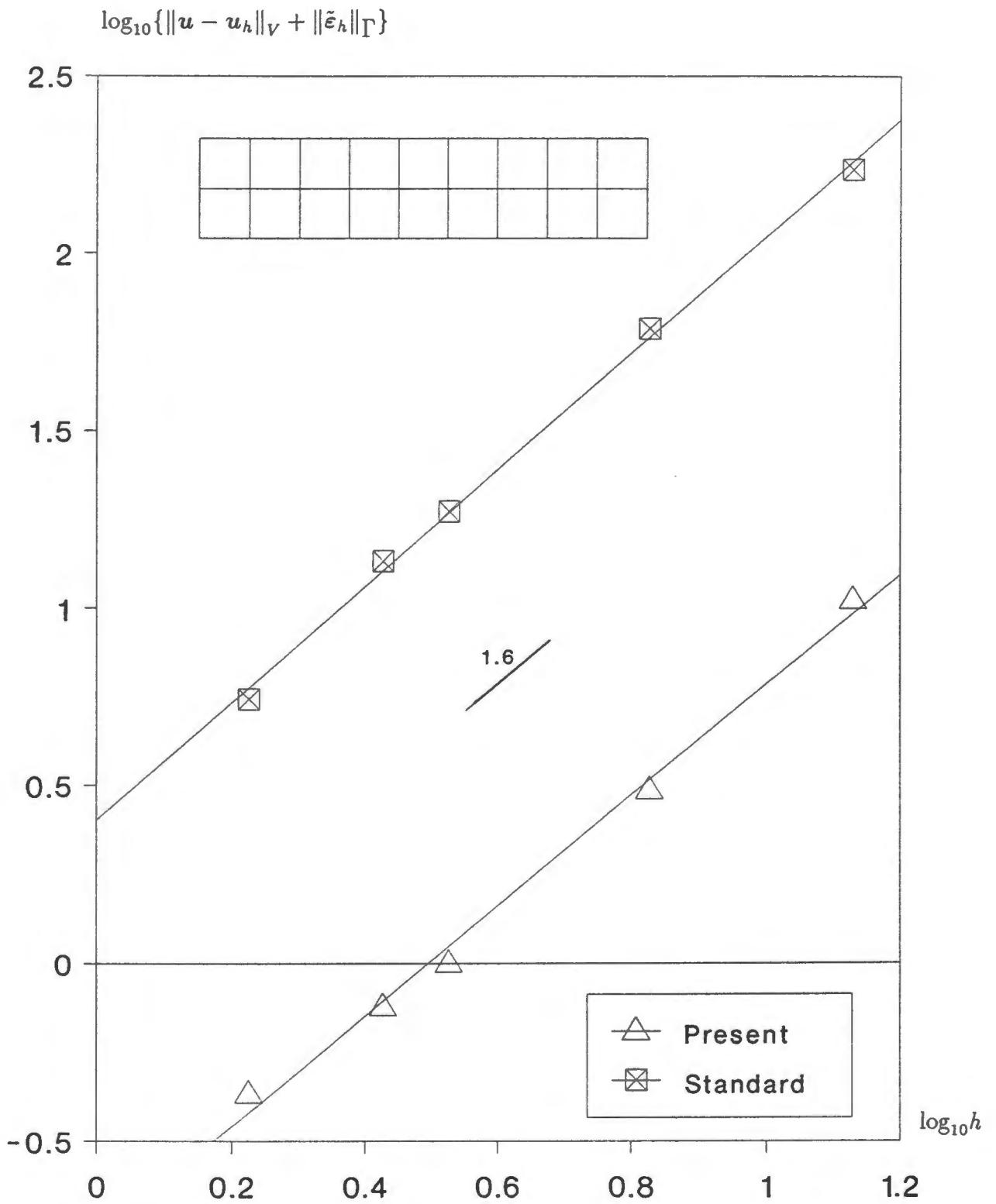


Figure 4.12: Plot of error in energy norm for a uniform mesh of rectangular elements (plane stress)

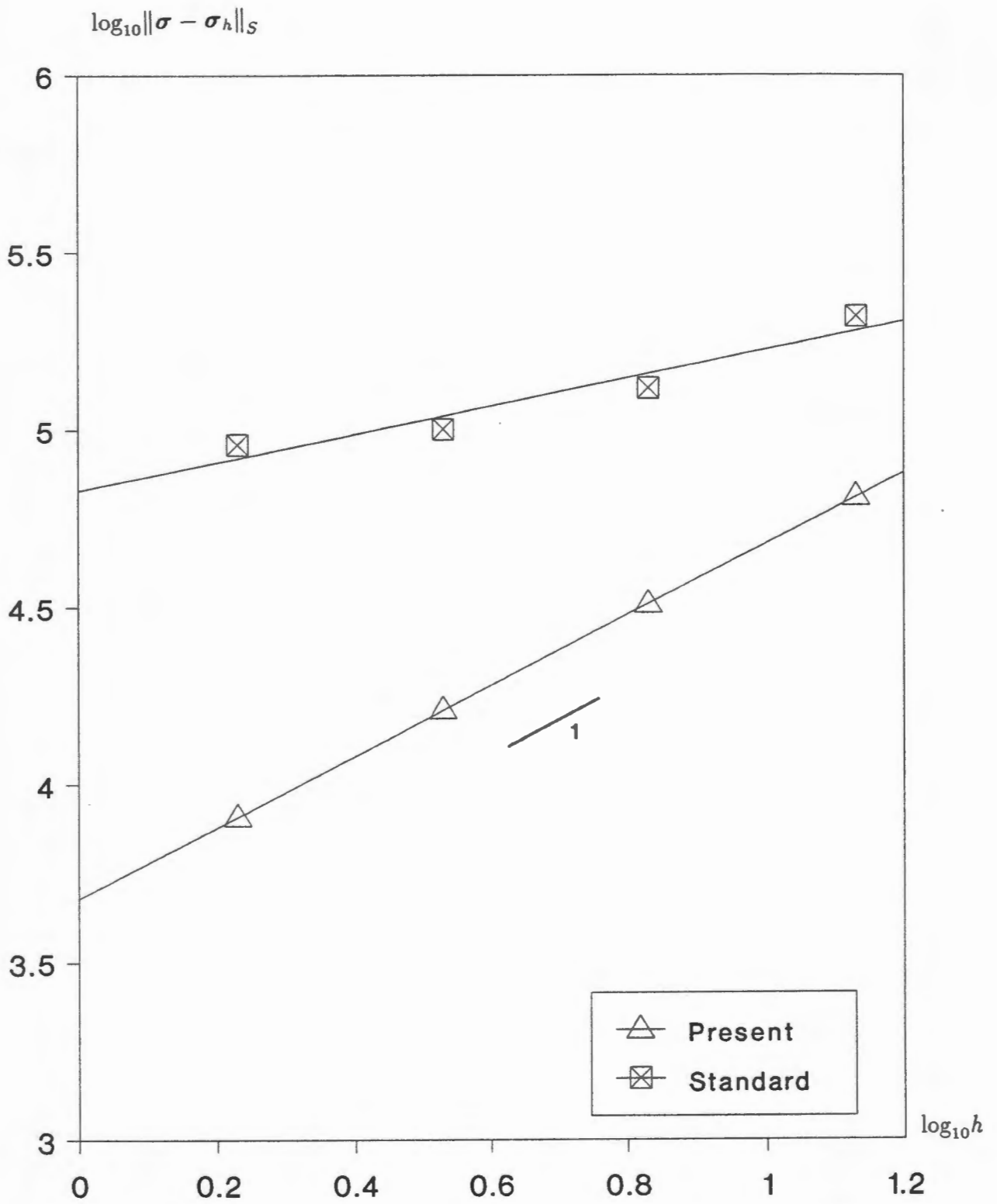


Figure 4.13: Plot of stress error for a uniform mesh of rectangular elements (plane stress)

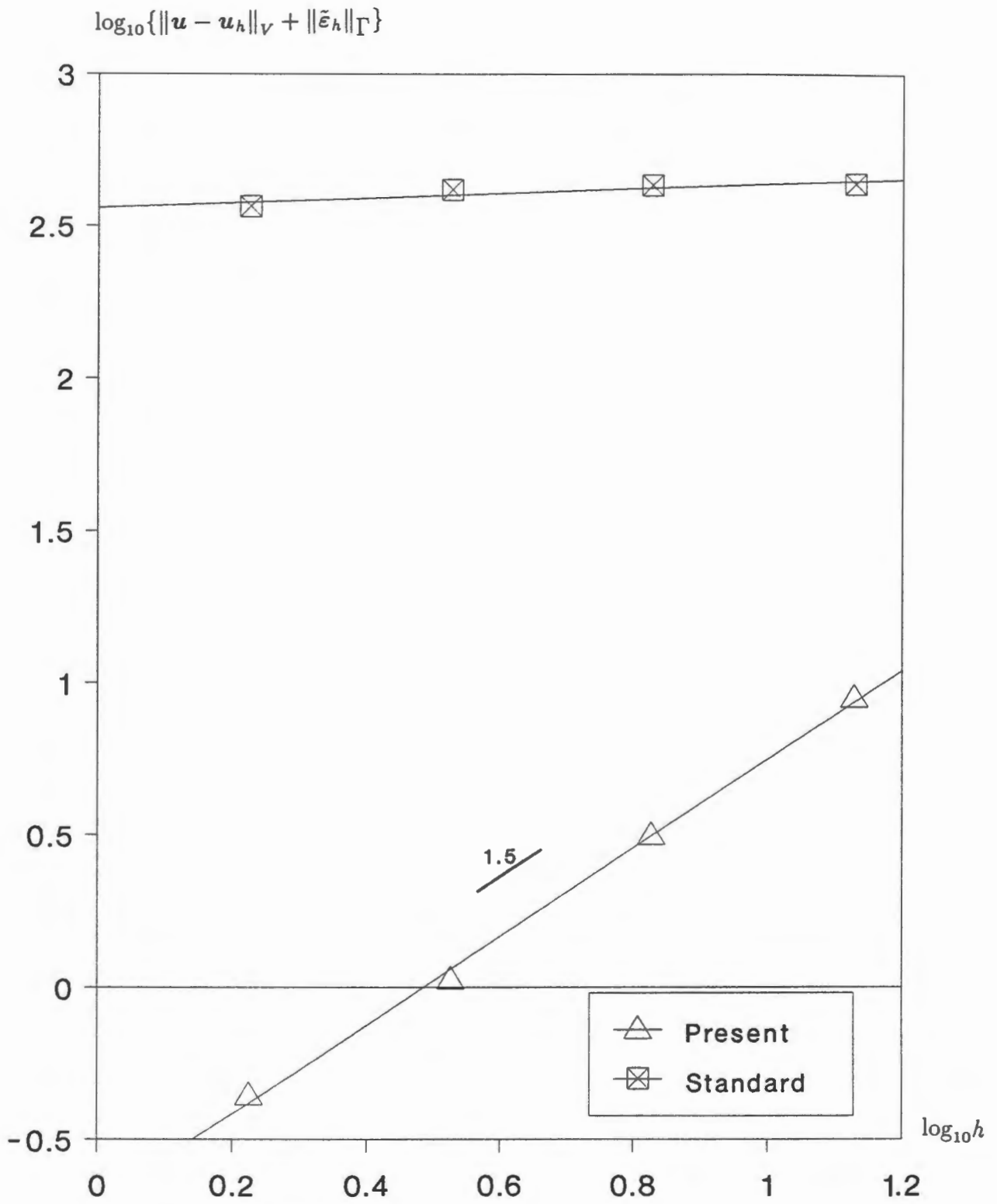


Figure 4.14: Plot of error in energy norm for a uniform mesh of rectangular elements (plane strain)

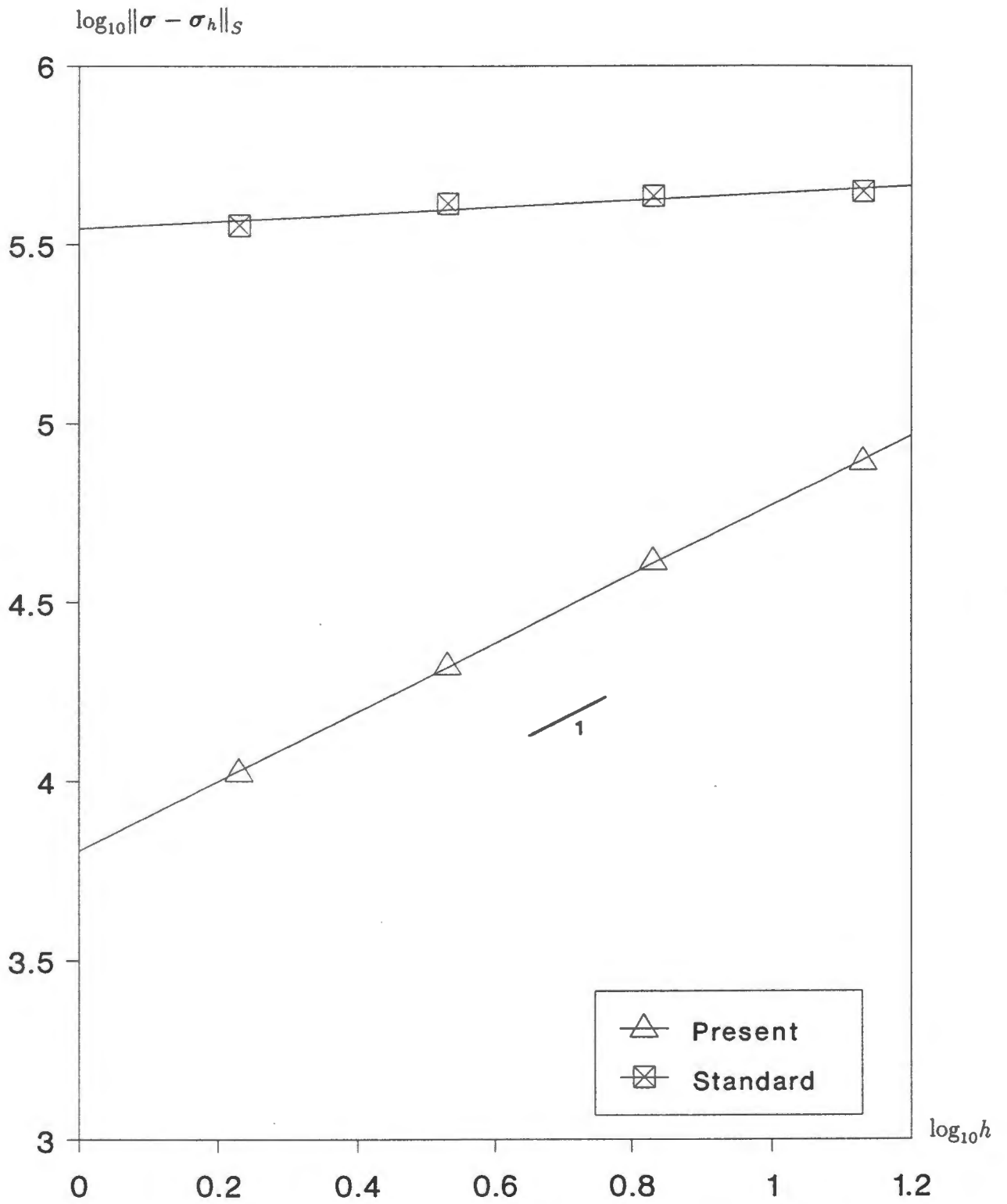


Figure 4.15: Plot of stress error for a uniform mesh of rectangular elements (plane strain)

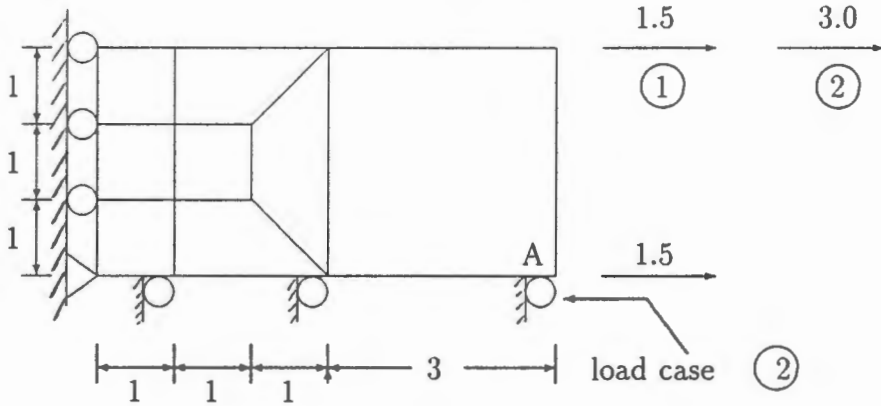


Figure 4.16: Cantilever beam subject to the load

Example 4

This numerical example is illustrated in Figure 4.16, and has also been used in [56]. Two load cases are considered. Load case (1) has the aim of testing whether the enhanced strain method gives the correct representation of a state of constant stress. Table 4.2 shows that all elements pass the test. For load case (2), which contains bending behaviour, the vertical displacement at A is given in the table.

Elements	Case 1	Case 2
	u_A	$-v_A$
Standard	6.00	17.00
And-Ramm (enhanced strain)[2]	6.00	17.64
Simo-Rifai (enhanced strain)[48]	6.00	17.64
Present (enhanced strain)	6.00	17.62
Pian-Sumihara (assumed stress)[39]	6.00	17.64
Exact solution	6.00	18.00

Table 4.2

Example 5

This final example, introduced by Pian and Sumihara [39], addresses the issue of sensitivity to mesh distortions. The test example used here is the same as we used in the Example 2, but contains only two elements as shown in the Figure 4.17.

Figure 4.18 shows the vertical displacement at A. In contrast to the other enhanced

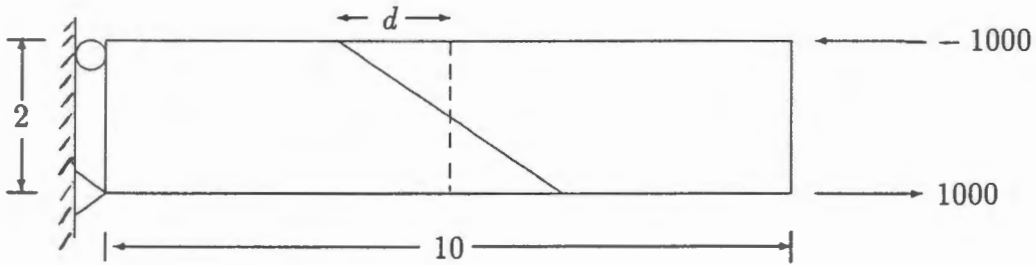


Figure 4.17: Two element problem to test for mesh distortion

strain elements, the basis introduced here exhibits rather extreme sensitivity to mesh distortion.

The sensitivity displayed in Figure 4.18 should be placed in perspective by evaluating it in the light of the results shown in the earlier figures, where it is seen that distortions of meshes which are mild, yet realistic, have no negative effect on the performance of this basis. Indeed, the basis (4.8) consistently gives results which are either equivalent, or (often) superior, to those obtained by existing enhanced strain bases or the assumed strain basis of [39]. This improved behaviour manifests itself most compellingly as an improvement in the constant appearing in the error estimates.

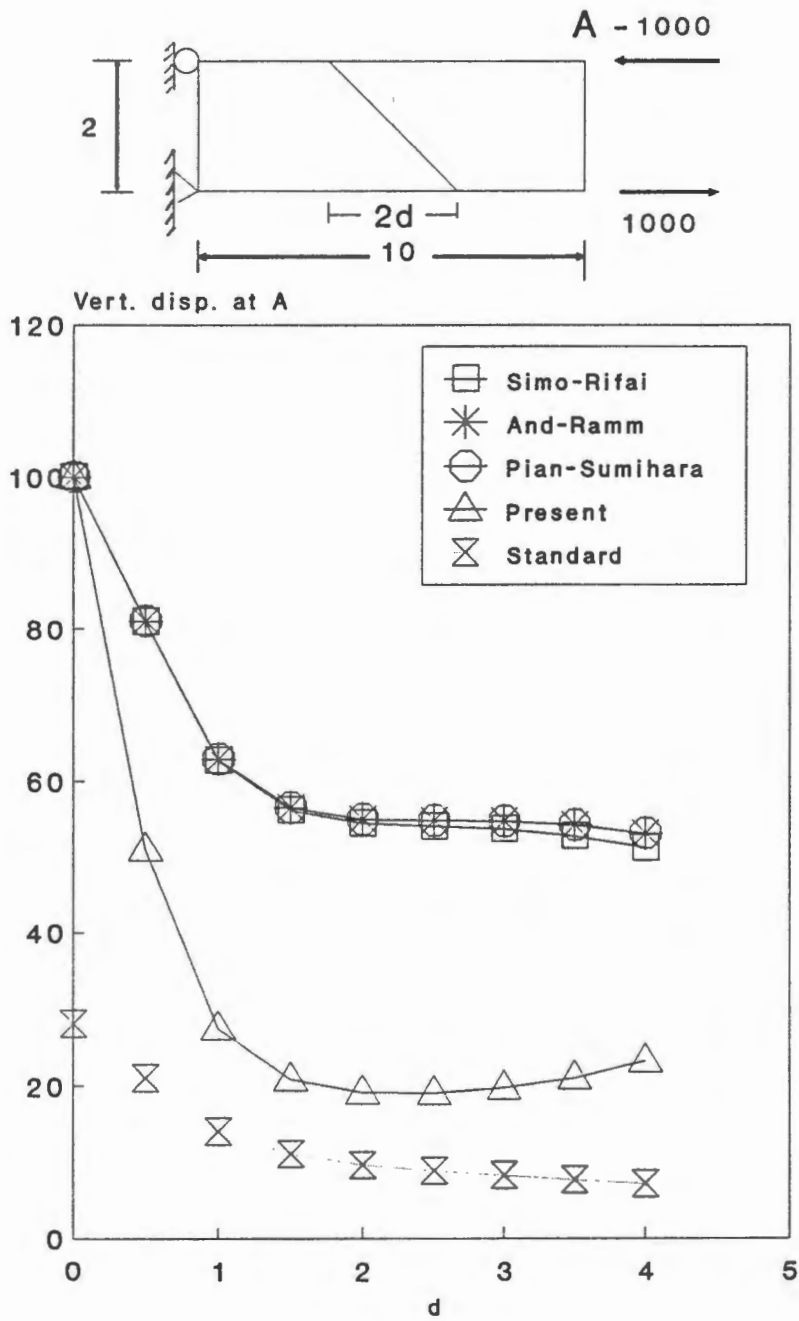


Figure 4.18: Sensitivity to mesh distortion

Chapter 5

Conclusions

5.1 Remarks on work covered in the thesis

The main thrust of this thesis has been to obtain a better understanding of the enhanced strain method, and to carry out a reasonably comprehensive analysis.

In order to obtain various geometrical estimates for quadrilaterals in a systematic way, the notion of the equivalent parallelogram associated with a quadrilateral was introduced in Section 2.4. This parallelogram is found to have geometric properties which are of interest in their own right. It has been shown, furthermore, that the *difference* between the equivalent parallelogram and the quadrilateral, measured in an appropriate way, is $O(h)$, where h is the length of the longer diagonal of the quadrilateral, so that the quadrilateral may be viewed as a small perturbation of the equivalent parallelogram.

The analysis presented in Sections 3.3 and 3.4 goes some way towards providing an overall understanding of the behaviour of enhanced strain elements. The error estimates obtained there are of course asymptotic, and give no clue as to why these elements perform so well in the coarse-mesh regime. An explanation of this phenomenon may instead be found in the numerical results presented in Section

4.2: there is strong evidence which suggests that this good performance is linked to the fact that the constant appearing in the error estimates for the enhanced strain element is smaller than that for the standard case. Furthermore, this difference appears to be more marked for the case of pointwise errors, and it is these errors, rather than errors in Sobolev norms, which tend to be routinely examined when the performance of elements is evaluated. This behaviour also casts some doubt on whether the superior performance of enhanced strain elements is an example of superconvergence.

The numerical results generally exhibit rates of convergence which are higher than those predicted by the theory. This could be for a variety of reasons: for example, the test problems considered are relatively simple in terms of geometry, loading and finite element meshes.

The new basis for enhanced strains introduced in Chapter 4 gives results which are very promising. It also has the advantage of being simple, in that the matrix of basis vectors contains only 4 nonzero members, while the bases used by Simo *et. al* [48] and Ramm *et. al* [2] contain 12 and 21 nonzero members (after being multiplied by \mathbf{T}_0) respectively. Therefore, from a computational point of view, our basis is more efficient than existing bases.

Stress with constant interpolation gives decent results, as predicted, but the additional linear terms suggested by Pian and Sumihara make a big difference in the quality of approximations for coarse meshes.

5.2 Future work

This study has been confined to the analysis of enhanced strain method with isoparametric quadrilateral elements in the context of the linear elasticity problem. The application of this method to small strain plasticity is discussed in [2, 48]. Our ana-

lysis and numerical studies demonstrate further the efficiency and accuracy of the method. Indeed, the enhanced strain approach has opened the door for applications to a wide range of problems such as incompressibility, plates and shells, large strain problems and so on.

As far as mathematical analysis goes there remain some open problems. The incompressibility problem has been studied in [44] for rectangular elements. Extension of this analysis to arbitrary quadrilateral elements is likely to be more difficult since the analysis would parallel that for the problems such as the $Q_1 - P_0$ element which is difficult to analyse, and for which only partial results are known (see Brezzi and Fortin [11]). But results in [44] suggest that the method can be applied successfully to the Navier-Stokes equations.

Since singular loading (point loads) occur in many practical problems, it would be useful to extend the analysis to take account of this type of loading.

Some work has been done in applying the method of enhanced strains to plate and shell problems. The key motivation here is to use the method as a means of overcoming locking in the thin limit, with the context of Mindlin and related models. Some encouraging results are reported in [2, 48, 62], the last of these being concerned with the incompatible modes method. There is, however, no clear analysis which would point the way to using the enhanced strain method as a means of overcoming locking, especially for arbitrary quadrilaterals.

The problems discussed so far are in the context of small strains. The application of the enhanced strain technique to large strain problems is described in [49, 50]. However, a mathematical analysis has yet not been carried out.

Bibliography

- [1] Adams R. A., *Sobolev Spaces*. Academic Press, New York (1975).
- [2] Andelfinger U. and Ramm E., EAS-elements for two dimensional, plate and shell structures and their equivalence to HR-elements. *Int. J. Num. Meth. Engng.*, **36**, 1311-1337 (1993).
- [3] Arunakirinathar K. and Reddy B. D., Some geometrical results and estimates for quadrilateral finite elements. To appear in *Comp. Meth. App. Mech. Engng.*
- [4] Arunakirinathar K. and Reddy B. D., Further results for enhanced strain methods with isoparametric elements. In review, *CERECAM Report No. 238* (1994).
- [5] Aubin J. P., *Approximation of Elliptic Boundary Value Problems*. Wiley, New York (1972).
- [6] Bathe K. J. and Dvorkin E. N., A four-node plate bending element based on Mindlin-Reissner plate theory and a mixed interpolation. *Int. J. Num. Meth. Engng.*, **21**, 367-383 (1985).
- [7] Bathe K. J., *Finite Element Procedures in Engineering Analysis*. Prentice-Hall, Englewood Cliffs, NJ (1987).
- [8] Becker E. B., Carey G. F. and Oden J. T. *Finite Elements, An Introduction (Vol I)*. Prentice-Hall (1981).
- [9] Belytschko T. and Bindemann L. P., Assumed strain stabilization of the 4-node quadrilateral with 1-point quadrature for nonlinear problems. *Comp. Meth. Appl. Mech. Engng.*, **88**, 311-340 (1991).

- [10] Bramble J. H. and Hilbert S. R., Bounds for a class of linear functionals with applications to Hermite interpolation. *Num. Math.*, **16**, 362-369 (1971).
- [11] Brezzi F. and Fortin M., *Mixed and Hybrid Finite Element Methods*. Springer Verlag (1991).
- [12] Carey G. F. and Oden J. T. *Finite Elements: A Second Course Vol II*. Prentice-Hall (1983).
- [13] Cartan H., *Differential Calculus*. Hermann, Paris (1971).
- [14] Chadwick P., *Continuum Mechanics: Concise Theory and Problems*. London, George Allen & unwin Ltd. (1976).
- [15] Ciarlet P. G. and Wagchal C., Multi-point Taylor formulas and applications to the finite element method. *Num. Math.*, **17**, 84-100 (1971).
- [16] Ciarlet P. G. and Raviart P.-A., Interpolation theory over curved elements, with applications to finite element methods. *Comp. Meth. Appl. Mech. Engng.*, **1**, 217-249 (1972).
- [17] Ciarlet P. G. and Raviart P.-A., General Lagrange and Hermite interpolation in \mathbb{R}^n with applications to finite element methods. *Arch. Rat. Mech. Anal.*, **46**, 177-198 (1978).
- [18] Ciarlet P. G. *The Finite Element Method for Elliptic Problems*, North Holland (1978).
- [19] Cook R.D, Malkus D. S. and Plesha M. E., *Concepts and Applications of Finite Element Analysis*, John Wiley & Sons (1989).
- [20] Di S. and Ramm E., On alternative hybrid stress 2D and 3D elements. *Eng. Comp.*, **10** (1993).
- [21] Dupont T. and Scott R., Polynomial approximation of functions in Sobolev spaces. *Math. Comp.*, **34** 150, 441-463 (1980).

- [22] Ergatoudis I., Irons B. M. and Zienkiewicz O. C., Curved, isoparametric *quadrilateral* elements for finite element analysis. *Int. J. Solids Structs.*, **4**, 31-42 (1968).
- [23] Fichera G., *Linear Elliptic Differential Systems and Eigenvalue Problems*, Lecture Notes Series Vol. 8, Springer (1972).
- [24] Harari I. and Hughes T. J. R., What are C and h ?: Inequalities for the Analysis and Design of the Finite Element Methods. *Comp. Meth. App. Mech. Engng.*, **97**, 157-192 (1992).
- [25] Hinton E. and Owen D.R.J. *Finite Element Programming*. Academic Press (1977).
- [26] Hueck U. and Wriggers P., A formulation for the four-node quadrilateral element. *Technical Report*, Institut für Mechanik, Technische Hochschule Darmstadt, No 3/93 (1993).
- [27] Hueck U., Reddy B. D. and Wriggers P., On the stabilization of the rectangular 4-node quadrilateral element. *Commun. Num. Meth. in Engng.*, **10**, 553-663 (1994).
- [28] Hughes T. J. R., *Introduction to the Finite Element Method*. Prentice-Hall, New Jersey (1987).
- [29] Lazarov R. D., Convergence of finite difference schemes for generalized solutions of Poisson's equation. *Diff. Equations*, **17**, 829-837 (1981).
- [30] Lesaint P., On the convergence of Wilson's nonconforming element for solving the elastic problems. *Comp. Meth. App. Mech. Engng.*, **7**, 1-16 (1976).
- [31] Lesaint P. and Zlámal M., Convergence of the nonconforming Wilson element for arbitrary quadrilateral meshes. *Num. Math.*, **36**, 33-52 (1980).
- [32] Love A.E.H. *A Treatise on the Mathematical Theory of Elasticity*. Cambridge University Press (1934).

- [33] MacNeal R. H., Derivation of element stiffness matrices by assumed strain distributions. *Nucl. Engng. Des.*, **70**, 3-12 (1982).
- [34] Marsden J. L. and Hughes T. J. R., *Mathematical Foundations of Elasticity*. Prentice-Hall, New-Jersey (1983).
- [35] Nachbin L., *Topology on Spaces of Holomorphic Mappings*. Springer-Verlag, Berlin (1969).
- [36] Oden J. T. and Reddy J. N. *An Introduction to the Mathematical Theory of Finite Elements*. John Wiley and Sons, Inc. (1976).
- [37] Oden J. T., *Applied Functional Analysis*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1979).
- [38] Pian T. H. H. and Chen D., On the suppression of zero energy deformation modes. *Int. J. Num. Meths. Engng.*, **19**, 1741-1752 (1983).
- [39] Pian T. H. H. and Sumihara K., Rational approach for assumed stress finite elements. *Int. J. Num. Meths. Engng.*, **20**, 1685-1695 (1985).
- [40] Pian T. H. H. and Tong P., Relations between incompatible displacement model and hybrid stress model. *Int. J. Num. Meths. Engng.*, **22**, 173-181 (1986).
- [41] Pian T. H. H. and Wu C. C., A rational approach for choosing stress terms for hybrid finite element formulations *Int. J. Num. Meths. Engng.*, **26**, 2331-2343 (1988).
- [42] Razzaque A. The patch test for elements. *Int. J. Num. Meths. Engng.*, **22**, 63-71 (1986).
- [43] Reddy B. D. *Functional Analysis and Boundary Value Problems: An Introductory Treatment*. Longman, England (1986).
- [44] Reddy B. D. and Simo J. C., Stability and convergence of a class of enhanced strain methods. To appear in *SIAM J. Num. Anal.*

- [45] Scott L. R. and Vogelius M., Norm estimates for a maximal inverse of the divergence operator in spaces of piecewise polynomials. *Math. Modelling Num. Anal.*, **9**, 11-43 (1985).
- [46] Shi Z., A convergence condition for the quadrilateral element. *Num. Math.*, **44**, 349-361 (1984).
- [47] Simo J. C. and Hughes T. J. R., On variational foundations of assumed strain methods. *J. App. Mech.*, **53**, 51-54 (1986).
- [48] Simo J. C. and Rifai M. S., A class of assumed strain methods and the method of incompatible modes. *Int. J. Num. Meths. Engng.*, **29**, 1595-1638 (1990).
- [49] Simo J. C. and Armero F., Geometrically nonlinear enhanced strain mixed methods and the method of incompatible modes. *Int. J. Num. Meths. Engng.*, **33**, 1413-1449 (1992).
- [50] Simo J. C., Armero F. and Taylor R. L., Improved versions of assumed enhanced strain trilinear elements for 3D finite deformation problems. *Comp. Meths. App. Mech. Engng.*, **110**, 359-386 (1993).
- [51] Smith K. T., Inequalities for formally positive integro-differential forms. *Bull. Amer. Math. Soc.*, **67**, 368-370 (1961).
- [52] Stenberg R. A family of mixed finite elements for the elasticity problem. *Num. Math.*, **53**, 513-538 (1988).
- [53] Strang G., Variational crimes in the finite element method. In *Mathematical Foundations of the finite element method* (Ed. Aziz A. K.), Academic press, 689-710 (1972).
- [54] Strang G., Fix G.J. *An Analysis of the Finite Element Method*. Prentice Hall Inc., Englewood Cliffs, New York (1973).
- [55] Stummel F., The limitations of the patch test. *Int. J. Num. Meths. Engng.*, **15**, 177-188 (1980).

- [56] Taylor R. L., Beresford P. J. and Wilson E. L., A nonconforming element for stress analysis. *Int. J. Num. Meths. Engng.*, **10**, 1211-1219 (1976).
- [57] Taylor R. L., Simo J. C., Zienkiewicz O. C. and Chan A. C. H., The Patch Test- A condition for assessing FEM convergence. *Int. J. Num. Meths. Engng.*, **22**, 39-62 (1986).
- [58] Timoshenko S. P. and Goodier J. N., *Theory of Elasticity, 3rd Edition*. McGraw-Hill, New York (1970).
- [59] Wilson E. L., Taylor R. L., Docherty W. P. and Ghaboussi J., Incompatible displacement models. In *Numerical and Computer Models in Structural Mechanics* (ed. S. J. Fenves et al), Academic Press, New York (1973).
- [60] Wu C., Huang M. and Pian T. H. H., Consistency condition and convergence criteria of incompatible elements: General formulation of incompatible functions and its application. *Comp. Struc.*, **27** 5, 639-644 (1987).
- [61] Yuan K., Wen J. and Pian T. H. H., A unified theory for formulations of hybrid stress membrane elements. *Int. J. Num. Meths. Engng.*, **37**, 457-474 (1994).
- [62] Zhang Z. and Zhang S., Wilson's element for the Reissner-Mindlin plate. *Comp. Meth. App. Mech. Engng.*, **113**, 55-65 (1994).
- [63] Zienkiewicz O. C. *The Finite Element Method Volume 1*. McGraw- Hill, London (1989).