A posteriori error estimation in the finite element method

Ainsworth, Mark

A Posteriori Error Estimation in the Finite Element Method

by

Mark Ainsworth

A Thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Mathematics

The University of Durham

December, 1989

ABSTRACT

The work broadly consists of two parts. In the first part we construct a framework for analyzing and developing *a posteriori* error estimators for use in the finite element solution of elliptic partial differential equations which have smooth solutions. The analysis makes use of complementary variational principles and the *superconvergence* phenomenon associated with the finite element method. The second part generalizes these results to the important case when the solution of the boundary value problem contains singularities. It is shown how the classical techniques may be easily modified to perform satisfactorily for the singular case.

ACKNOWLEDGEMENTS

I would like to thank Dr J. Z. Zhu of the Institute for Numerical Methods in Engineering (Swansea) for interesting discussions regarding the engineering applications of the finite element method and for his help in the preparation of the numerical results for Chapter 3. Thanks are due to my supervisor Dr Alan Craig and to Dr John Coleman for words of encouragement when progress seemed slow and for 'sceptism' when things were going too well!

DECLARATION

The work contained in this thesis has not been submitted elsewhere for any other degree or qualification and is all my own work unless referenced to the contrary in the text.

To mum and dad.

CONTENTS

Chapter 1	Introduction	1
Chapter 2	Fundamental Framework for Error Estimation	5
Chapter 3	Error Estimation for Elliptic Systems	66
Chapter 4	Error Estimation for Problems with Singularities	92
Chapter 5	Regularity: Sufficient but unnecessary?	135
Appendix	Lamé-Navier Equations and FEM Discretization	162
References		167

CHAPTER 1

Introduction.

The finite element method has become the standard procedure in the analysis of problems from stationary structural analysis through to transient fluid flow. In spite of the widespread use of the finite element method, it is only relatively recently that the emphasis has shifted towards the assessment of the reliability of the computed solution. Most of the progress has been made over the preceding decade. The first international conference on *Accuracy Estimates and Adaptive Refinements in Finite Element Computations* was held in Lisbon 1984, [20]. At this time the main areas of application were in two spatial dimensions and one time dimension. The prevailing finite element scheme was the *adaptive h version*, with only a small number of talks concerning the *p version* (see [8], [17], [18], [37], [73], [74] for detailed analyses of the *p version*). The advantages which might be expected from combining the two versions *the h-p version* were but a pipe dream (see [7], [9], [37], [38], [64] for detailed analyses of the *h-p version*). However, by the time the *Workshop on Adaptive Computational Methods for Partial Differential Equations* took place in Rensselaer in 1988 [34] several of the presentations discussed three dimensional applications, along with *h, p and h-p versions* of the finite element method. In addition, new areas such as parallelization of the algorithms were included. It seems therefore that the adaptive versions of the finite element method have become the standard versions in spite of the computational difficulties and complexities associated with them. In their introduction to the proceedings [34] of the Renssalaer workshop, the editors point out the shortcomings of the lack of standardisation and a unified approach to the problem of designing and implementing adaptive finite element codes and note that there is now a major effort being devoted to rectifying this situation.

At the heart of all adaptive algorithms lies a means of assessing the accuracy of the computed approximation. Moreover, this means usually performs the dual role of indicating those areas of the domain where the approximation is poor. It is only once the areas in which the solution is unacceptable have been located that the decision of how to improve the accuracy need be taken, giving rise to one or other of the basic types of refinement used to achieve convergence.

In principle the error in the approximation could be assessed using the standard error estimates [7], [17] and [27], for the relevant version. These error estimates can be obtained without having to perform any finite element analysis. Unfortunately, such *a priori* estimates of the error are completely unsuited to the task required of them by an adaptive routine. The main objections being that the estimates necessarily cater for the worse possible cases and therefore tend to be overly pessimistic when applied to a specific problem. Moreover, the *a priori* estimates only take account of the discretization error and ignore the many other sources of error e.g. due to round off, approximate solution techniques (see [78] for a detailed discussion).

One way in which we might hope to obtain more realistic error bounds is by allowing ourselves to make use of more information which is specific to the actual problem in which we are interested. Since we do not really need to have a bound on the error until after we have computed our approximate solution, it seems appropriate to use the computed approximation itself to help in the problem of computing the error. This has the added advantage that the other sources of error are also taken account of to some extent.

The idea of obtaining such *a posteriori* estimates of the error has been exploited throughout numerical analysis generally. The use of *a posteriori* error estimators of the error in the finite element method was pioneered by Babuška and Rheinboldt [12], [13], [14] and [15]. They developed ways of using the finite element approximation to estimate the error and also investigated how to use the estimates to improve the discretization scheme by refinement [14], [16]. Babuška and Rheinboldt were not the first to realize that there were advantages associated with careful mesh design (see Turcke and McNiece [77]) but they were amongst the first to propose concrete and practical methods for this purpose.

The *a posteriori* error estimation techniques of Babuška and Rheinboldt were taken up by both the engineering and mathematical communities and evolved in several directions [11], [22], [30], [45], [46], [47], [57], [71] and [85]. The traditional method of devising error estimators established by Babuška and Rheinboldt is based on using the defects in the equilibrium of the finite element solution. If the finite element approximation were the true solution then there would be no such discrepancy in the equilibrium. The estimators are obtained by solving local auxiliary problems to find an estimate of the discrepancy and hence an estimate of the error on a single element in the mesh [11], [14], [19], [22], [45]. This process can be considerably aided by the use of hierarchical elements which are associated with efficient implementation of the *p version* of the finite element method [8], [28], [35], [37], [75] and [86]. Having obtained estimates of the error on each element, these may be summed to obtain a global estimate of the error. This approach has been used for non-linear problems also, see [14], [22] and [58].

The literature on *a posteriori* error estimation is ever increasing, and one of the main problems is the lack of any standardisation. It is desirable to have such an approach not only on aesthetic grounds but for practical reasons too. It has been noted by several authors (e.g. [56]) that there appears to be a link between *a posteriori* error estimation and the *superconvergence* phenomenon associated with the finite element method. However, such remarks have to our knowledge not been consolidated by exhibiting the connection explicitly or developed any further than a comment in passing. Superconvergence phenomena have been developed to a high level of understanding (see [51] for a recent survey of superconvergence results) and if the link between superconvergence and a *a posteriori*error estimation could be forged, would provide error estimators for large classes of problems.

One of the purposes of the current work is to establish this connection (see [57] in this respect). In Chapter 2, we work towards a main result which proclaims that if there is a *superconvergence* result associated with a particular finite element approximation scheme and an *a posteriori* error estimator is constructed with regard to this *superconvergence* result, then the estimator will be a reliable one. The analysis illustrates the precise meaning of how the estimator should be constructed. In Chapter 3, the basic result of Chapter 2 is extended to Lamé-Navier equations of elasticity and discusses some partial theoretical results for the heuristic estimator of Zienkiewicz and Zhu [87], [88].

The second aim of the work is to attempt to lay down foundations for the construction of estimators which will perform reliably for problems which possess singular solutions. There are currently no estimators which perform reliably in such cases. This state of affairs is somewhat worrying since the problems which are of most interest in practice possess such singularities. In Chapter 4 we restrict our discussion to one spatial dimension and consider how to generalize the existing estimators to cope with singularities. The outcome is a new estimator which performs reliably and yet is extremely simple to incorporate into a code which uses the existing estimator. Finally, in Chapter 5 we admit that our assumptions for the theoretical developments of the earlier chapters have been rather strong and consider to what extent they were really necessary for the analysis to hold.

The current work is really only a start on the processes of unifying estimation techniques and of developing estimators for singular problems. There are many possible ways of generalizing the fundamental framework which is developed, just as there are still many interesting and important questions to be answered regarding *a posteriori* error estimation.

CHAPTER 2

Fundamental Framework for Error Estimation.

2.1 Introduction

The continued popularity of the finite element method in both the engineering and the mathematical communities has led to an increasingly large amount of attention being paid to the problem of assessing the quality of the computed approximation. The *a priori* estimates of the error have proved unsuitable for use in obtaining realistic estimates of the error. An alternative approach is to attempt to use the finite element approximation itself to find such estimates. Many ways of using the approximation to find such *a posteriori* estimates of the error have been suggested and used in practice.

The existing estimators may be roughly classified into two categories. Firstly some estimators have been rigorously analyzed mathematically and shown to estimate the true error increasingly well as the discretization scheme is refined. Conversely, some estimators have been proposed on purely intuitive grounds and justified *heuristically* on the basis of their performance in practical problems.

For some classes of approximation scheme there may be several estimators available, whilst for others there may be none. This situation is unsatisfactory both for the practical numerical analyst; who wishes to know which (if any) of the existing estimators should be used in a given situation, and; to the theoretical numerical analyst, who is interested in understanding the underlying structure which allows several apparently diverse methods to perform effectively.

In this chapter a general approach to error estimation is developed. This will be useful for the following reasons. It will aid the classification of the existing estimators and allow some of the heuristically proposed estimators to be set on a sound theoretical footing. It will help to reveal the underlying framework and enable new estimators to be developed.

The chapter is organized as follows. After establishing the notation and a model problem and its approximation, we develop a result (Theorem 2.3.2) which will be useful in analyzing error estimators. We then consider a particular class of methods for obtaining error estimators, which we shall refer to as recovery based estimators. The approach is kept sufficiently general that it will encompass many types of approximation scheme. By way of example, we shall show how an existing estimator falls within the framework; how an existing estimator may be regarded as a simplified version of a recovery based estimator; and, how a new estimator may be easily developed. Finally, we present numerical examples illustrating the performance of the new estimators.

2.2 Preliminaries.

2.2.1 Notation.

Let Ω be an open, bounded and simply connected domain in $\Re^n$ $(n = 1, 2, 3)$ with boundary $\partial\Omega$ which is Lipschitz continuous. We shall write the point $x \in \bar{\Omega}$ as $(x_1, x_2, \ldots, x_n)$ relative to the canonical basis on $\Re^n$. Let $\mathcal{E}(\Omega)$ denote the space of real valued, infinitely differentiable functions on Ω for which derivatives of all orders have continuous extensions on $\bar{\Omega}$, and let $\mathcal{D}(\Omega)$ denote the subspace of $\mathcal{E}(\Omega)$ consisting of functions which have compact support in Ω.

For $m \geq 0$ and $p \in [1, \infty]$, let $W^{m,p}(\Omega)$ denote the usual Sobolev spaces defined as the completions of $\mathcal{E}(\Omega)$ in the norm on $W^{m,p}(\Omega)$ given by

$$\| v \|_{m,p,\Omega} = \begin{cases} \left\{ \sum_{k \le m} |v|_{k,p,\Omega}^p \right\}^{\frac{1}{p}}, & \text{if } p \in [1,\infty) \\ \max_{k \le m} |v|_{k,p,\Omega}, & \text{if } p = \infty. \end{cases} \tag{2.2.1}$$

where $|\cdot|_{k,p,\Omega}$ is a semi-norm on $W^{k,p}(\Omega)$ given by

$$|v|_{k,p,\Omega} = \begin{cases} \left\{ \sum_{|\alpha|=k} \int_\Omega |D^\alpha v(x)|^p dx \right\}^{1/p}, & \text{if } p \in [1,\infty) \\ \max_{|\alpha|=k} \left\{ \text{ess sup}_{x \in \Omega} |D^\alpha v(x)| \right\}, & \text{if } p = \infty \end{cases} \tag{2.2.2}$$

where α is a multi-index and D^α denotes the derivative in the generalized sense.

The completion of $\mathcal{D}(\Omega)$ in the norm on $W^{m,p}(\Omega)$ is denoted by $W_0^{m,p}(\Omega)$. In the special case of $p = 2$ we shall denote $W^{m,2}(\Omega)$ and $W_0^{m,2}(\Omega)$ by $H^m(\Omega)$ and $H_0^m(\Omega)$ respectively, and in the special case of $m = 0$ we shall denote $W^{0,p}(\Omega)$ by $L^p(\Omega)$, the usual space of Lebesgue p-integrable functions. The completions of $\mathcal{E}(\Omega)$ and $\mathcal{D}(\Omega)$ in the norm on $W^{m,\infty}(\Omega)$ are denoted by $C^m(\Omega)$ and $C_0^m(\Omega)$ respectively.

2.2.2 The model problem.

For simplicity we shall consider the model problem

$$L[u] \equiv -\nabla \cdot [A(x)\nabla u] + a_0(x)u(x) = f(x), \quad x \in \Omega \tag{2.2.3}$$

where

$$\nabla \cdot [A(x)\nabla u] \equiv \sum_{i,j=1}^{n} \frac{\partial}{\partial x_j}\left[a_{ij}(x)\frac{\partial u(x)}{\partial x_i}\right] \tag{2.2.4}$$

with

$$u|_{\partial\Omega} \equiv 0. \tag{2.2.5}$$

We shall assume throughout that the coefficients satisfy

- there exists a constant $\beta > 0$ such that $\forall \xi_i$, $1 \leq i \leq n$

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \beta \sum_{i=1}^{n} \xi_i^2 \quad \forall x \in \bar{\Omega} \tag{2.2.6}$$

- $a_0(x) \geq 0 \quad \forall x \in \bar{\Omega}$
- $a_{ij}(x) \equiv a_{ji}(x)$
- $a_{ij}, a_0 \in L^\infty(\Omega)$
- $f \in L^2(\Omega)$

and that a_{ij}, a_0 and f are defined everywhere on $\bar{\Omega}$. Later we shall have occasion to further strengthen the requirements on the coefficients.

Let $a(\cdot\,,\cdot\,)$ be the bilinear form given by

$$a(u,v) = \int_\Omega \Big\{ \sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial u(x)}{\partial x_i}\frac{\partial v(x)}{\partial x_j} + a_0(x)u(x)v(x)\Big\}\,dx \tag{2.2.7}$$

and let $(f,\cdot\,)$ be the linear form given by

$$(f,v) = \int_\Omega f(x)v(x)dx. \tag{2.2.8}$$

Under the above assumptions on the coefficients it may be shown that $(f,\cdot\,)$ is a continuous linear form on $H_0^1(\Omega)$ and that $a(\cdot\,,\cdot\,)$ is

- continuous bilinear form on $H_0^1(\Omega)\times H_0^1(\Omega)$, *i.e.* there exists a bounded positive constant C_1 such that

$$|a(u,v)| \leq C_1 |u|_{1,2,\Omega} |v|_{1,2,\Omega} \quad \forall u,v \in H_0^1(\Omega) \tag{2.2.9}$$

- $H_0^1(\Omega)$-elliptic, *i.e.* there exists a positive constant C_2 such that

$$a(v,v) \geq C_2 |v|_{1,2,\Omega}^2 \quad \forall v \in H_0^1(\Omega). \tag{2.2.10}$$

Throughout we shall use the letter C to denote generic positive constants which need not necessarily take the same value in any different place.

The Lax-Milgram Lemma (see Ciarlet [27]) guarantees that under conditions (2.2.9) and (2.2.10) there exists a unique solution to the *weak* form of (2.2.3) given by

$$u \in H_0^1(\Omega) : a(u,v) = (f,v) \quad \forall v \in H_0^1(\Omega). \tag{2.2.11}$$

2.2.3 Approximation of the model problem.

Let Ω be such that it may be exactly partitioned into the union of non-empty, closed Lipschitzian subdomains Ω_i, where Ω_i are the images of a standard reference domain $\hat{K}$ under a family of affine transformations $\{F_i\}$. That is

$$\bar{\Omega} = \bigcup_{i=1}^{N} \Omega_i \tag{2.2.12}$$

where Ω_i is the range of the affine mapping

$$F_i : \hat{x} \in \Re^n \to F_i(\hat{x}) = B_i\hat{x} + b_i \tag{2.2.13}$$

and B_i is an invertible $n \times n$ matrix and $b_i \in \Re^n$. The reference domain $\hat{K}$ may be either

- the n-simplex in $\Re^n$ given by

$$\hat{K} = \{(x_1, \ldots, x_n) \in \Re^n : x_i \geq 0 \quad i = 1, \ldots, n \text{ and } \sum_{i=1}^{n} x_i \leq 1.\} \tag{2.2.14}$$

- the n-rectangle in $\Re^n$ given by

$$\hat{K} = \prod_{i=1}^{n} [0, 1]. \tag{2.2.15}$$

Let the partition $\{\Omega_i\}$ be denoted by $\mathcal{T}$. We shall assume that the components Ω_i of the partition satisfy the condition that either

$$\Omega_i \cap \Omega_j = \emptyset \tag{2.2.16}$$

or that

$$\Omega_i \cap \Omega_j \text{ is either } \left\{ \begin{array}{l} \text{an entire face} \\ \text{an entire side} \\ \text{a vertex} \end{array} \right\} \text{ of } \Omega_i \text{ and } \Omega_j. \tag{2.2.17}$$

For any partition $\mathcal{T}$ we follow the usual convention of associating a parameter h with $\mathcal{T}$, defined to be

$$h = \max_{i=1,\ldots,N} h_i \tag{2.2.18}$$

where

$$h_i = \mathrm{diam}(\Omega_i), \quad \text{for } i = 1, \ldots, N. \tag{2.2.19}$$

If we wish to indicate the dependence on h then we shall use the notation

$$\mathcal{T}^h = \{\Omega_i^h\}_{i=1}^{N^h}. \tag{2.2.20}$$

More generally, we shall consider *families* $\mathcal{M} = \{\mathcal{T}^h\}$ of such partitions. If, for all of the partitions $\mathcal{T}^h \in \mathcal{M}$, we have that there exists a constant C which is independent of h, such that

$$\forall \Omega_i^h \in \mathcal{T}^h : 0 < \frac{h_i}{\rho_i} \leq C < \infty \tag{2.2.21}$$

where

$$\rho_i = \sup\{\mathrm{diam}(\mathcal{B}) : \mathcal{B} \text{ is a ball contained in } \Omega_i^h\}, \tag{2.2.22}$$

then we say that $\mathcal{M}$ is a family of *quasi-uniform* meshes. If $h \to 0$ and $\mathcal{M}$ is quasi-uniform, then we say that $\mathcal{M}$ is *regular*. If there exists a constant C which is independent of h such that

$$\forall \Omega_i^h \in \mathcal{T}^h : 0 < \frac{h}{h_i} \leq C < \infty \tag{2.2.23}$$

then the mesh is said to satisfy the *inverse assumption*. We shall assume throughout that the mesh is regular and satisfies the inverse assumption.

In order to characterize the finite element approximation, we follow the formalism of Ciarlet [27] and define an abstract finite element as follows.

DEFINITION 2.2.1 (K, P, Σ) *in* $\Re^n$ *is a finite element if*

- K *is a closed subset of* $\Re^n$ *with non-empty interior and Lipschitzian boundary*
- P *is a space of real valued functions defined on* K
- Σ *is a finite set of linear forms* $\{\phi_i\}$ *on* P.

In order to specify the space P more fully we make some preliminary definitions

DEFINITION 2.2.2 $\mathcal{P}_k(K)$, $K \subset \Re^n$ *is the space of all polynomials of degree* $\leq k$ *in the variables* $x_1, \ldots, x_n$. *That is, any* $p \in \mathcal{P}_k(K)$ *may be written in the form*

$$p : x = (x_1, \ldots, x_n) \in \Re^n \rightarrow p(x) = \sum_{|\alpha| \leq k} \gamma_{\alpha_1, \ldots, \alpha_n} x_1^{\alpha_1} \cdots x_n^{\alpha_n} \tag{2.2.24}$$

where $|\alpha| = \sum_{i=1}^{n} \alpha_i$.

DEFINITION 2.2.3 $\mathcal{Q}_k(K)$, $K \subset \Re^n$ *is the space of all polynomials of degree* $\leq k$ *in each of the variables* $x_1, \ldots, x_n$. *That is, any* $p \in \mathcal{Q}_k(K)$ *may be written in the form*

$$p : x = (x_1, \ldots, x_n) \in \Re^n \rightarrow p(x) = \sum_{\alpha_i \leq k} \gamma_{\alpha_1, \ldots, \alpha_n} x_1^{\alpha_1} \cdots x_n^{\alpha_n}. \tag{2.2.25}$$

A key property we shall ask of the space P is that the inclusions

$$\mathcal{P}_1(K) \subset P \subset \mathcal{P}_m(K), \tag{2.2.26}$$

hold for some $m \geq 1$. Moreover, we shall only consider *Lagrange* finite elements (see Ciarlet [27]), so that Σ only depends on function values and not on derivative values. The finite element approximation space is then defined to be

$$X_h = \{v : v|_{\Omega_i^h} \circ F_i^h \in P \quad \forall \Omega_i^h \in \mathcal{T}^h\} \tag{2.2.27}$$

where $v|_{\Omega_i^h}$ denotes the restriction of v to Ω_i^h, and F_i^h is the affine mapping of $\hat{K}$ onto Ω_i^h. The subspace $V_h = X_h \cap H_0^1(\Omega)$ is called the finite element trial space.

We define the *P-interpolation* operator as follows

DEFINITION 2.2.4 *Let $v \in C(K)$ be given, then the canonical or P- interpolation operator Π is defined so that*

$$\Pi v \in P : \quad \phi_i(\Pi v) = \phi_i(v) \quad \forall \phi_i \in \Sigma. \tag{2.2.28}$$

and the interpolation operator on the finite element subspace is defined so that

DEFINITION 2.2.5 *Let $v \in C(\bar{\Omega})$ be given and denote the P-interpolation operator on the element Ω_i^h by Π_i^h, then the X_h-interpolation operator Π^h is defined so that the restriction of the function $\Pi^h v$ to any element Ω_i^h is $\Pi_i^h v$. That is*

$$\Pi^h v|_{\Omega_i} = \Pi_i^h v \quad \forall \Omega_i^h \in \mathcal{T}^h. \tag{2.2.29}$$

REMARK If the space P satisfies the inclusion

$$\mathcal{P}_p \subset P, \quad p \geq 1$$

and p is the largest such integer then we say that the finite element space has degree p and we shall write Π_p^h to indicate the p dependence of the operator.

The *finite element method* consists of approximating the solution of (2.2.11) by the solution u_h of the problem

$$u_h \in V_h : a(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h. \tag{2.2.30}$$

u_h is referred to as the finite element approximation to u. It is well known that the discretized problem (2.2.30) inherits the properties from the infinite dimensional problem which ensure the existence of a unique solution to (2.2.11).

2.2.4 A priori and a posteriori error estimation.

Of course we are interested in the accuracy of the finite element approximation. Letting $e(x) = u(x) - u_h(x)$ denote the error in this approximation, a natural norm in which to measure the error is the *energy norm*

$$\| e \|_E = a(e, e)^{\frac{1}{2}}. \tag{2.2.31}$$

Under the previous assumptions it may be shown that the energy norm is equivalent to the norm on $H_0^1(\Omega)$. The standard *a priori* estimate for the error in the norm on $H_0^1(\Omega)$, which is valid under the regularity assumptions on the partition and that $u \in H^{p+1}(\Omega)$, is

$$\| u - u_h \|_{1,2,\Omega} \leq Ch^p \| u \|_{p+1,2,\Omega} \tag{2.2.32}$$

where C is a constant which is independent of h and u. This *a priori* estimate of the error tells us the rate of convergence which we can anticipate but is of limited use if we wish to find a numerical estimate of the accuracy. The problem is that either the constant C is unknown explicitly or if bounds are found on C, then the estimate is found to be unduly pessimistic.

One way in which we might hope to enhance the prospects of finding a realistic estimate or bound on the discretization error, is to use the finite element approximation itself in estimating $\| e \|_E$. This idea of using u_h to estimate the error *a posteriori* is not a new one and a variety of methods as to how u_h might be used have appeared in the literature.

The criterion of what constitutes a *good* method of using u_h is quantified by the condition of *asymptotic exactness* of the resulting *a posteriori* error estimator, introduced by Babuška and Rheinboldt [12].

DEFINITION 2.2.6 *Asymptotic Exactness. Let* ϵ *be an a posteriori error estimator, then if under reasonable assumptions on* u, a_{ij}, a_0, f *and the family of meshes* $\{\mathcal{T}^h\}$, *we have that*

$$\| e \|_E \approx \{1 + O(h^\gamma)\}\epsilon \text{ as } h \to 0 \tag{2.2.33}$$

where $\gamma > 0$ *is independent of* h *and the constant in the* $O(h^\gamma)$ *term depends on* u, a_{ij}, a_0 *and* f *only, then we say that* ϵ *is an asymptotically exact a posteriori error estimator.*

The definition means that under favourable conditions, *i.e.* if the coefficients, data and mesh are sufficiently regular, an asymptotically exact error estimator will tend to estimate the true error exactly as the family of partitions becomes increasingly fine. The condition makes no provision for the cases in which the assumed regularity may be lacking, although it is tacitly assumed that the estimator will not be completely unsatisfactory in such cases.

The *a posteriori* error estimators which have been proposed may be roughly divided into two categories. Firstly, some estimators have been rigorously analyzed mathematically and have been shown to be asymptotically exact. Conversely, many estimators have been proposed on purely intuitive grounds and justified *heuristically* on the basis of their performance in a number of specific cases. The latter estimators may be derived from examining the analytical expression for the true error in the case $a_0(x) \equiv 0$:

$$\| e \|_E^2 = \int_\Omega [\nabla(u - u_h)]^t A [\nabla(u - u_h)] dx. \qquad (2.2.34)$$

Naturally, *if* we knew ∇u explicitly then it would be a relatively easy matter to substitute it into this expression and to calculate the true error exactly. The intuitive approach argues that rather than having to know ∇u explicitly, it should be sufficient to use a *good enough* approximation to ∇u in its place.

The case $a_0(x) \not\equiv 0$ is dealt with by arguing that the dominant term in the error is the component containing the derivatives, and so it should be enough to estimate this dominant part only. Essentially this means that the same scheme is used whether or not $a_0(x) \equiv 0$.

The intuitive approach is appealing but does little to provide us with analytical support for the resulting method. Many estimators which are actually *used* in practice are obtained by using such a *heuristic* method based on 'replacing' ∇u by a quantity which is believed to be a good approximation to ∇u.

Conversely, it is found that some rigorously analyzed estimators which have been obtained in quite different ways can be brought within the framework of corresponding to a particular choice of approximation to ∇u obtained from u_h. In the next section we shall develop a result which will have obvious application as a theoretical tool in analyzing estimators which can be viewed within this framework.

2.3 Complementary Variational Principles.

Since we are interested in bounding the error measured in the energy norm, we first of all characterize the error as the solution of a boundary value problem which is analogous to (2.2.11). In fact, making the substitution $u(x) = e(x) + u_h(x)$ in (2.2.11) and rearranging easily gives the following characterization for e

$$e \in H_0^1(\Omega) : a(e, v) = (f, v) - a(u_h, v) \quad \forall v \in H_0^1(\Omega). \tag{2.3.1}$$

The function u_h is regarded as being known explicitly since we envisage using u_h itself in obtaining estimates of the error. In principle we could solve (2.3.1) exactly and hence compute $\| e \|_E$ exactly. Obviously in practice we will be unable to do this, since solving (2.3.1) is equivalent to solving (2.2.11). Equally well, we may characterize e as the solution to a variational problem (since the bilinear form $a(\cdot, \cdot)$ is symmetric)

$$e \in H_0^1(\Omega) : \quad \mathcal{J}(e) \leq \mathcal{J}(w) \quad \forall w \in H_0^1(\Omega) \tag{2.3.2}$$

where $\mathcal{J}$ is the quadratic functional given by

$$\mathcal{J}(w) = \frac{1}{2} a(w, w) - (f, w) + a(u_h, w). \tag{2.3.3}$$

Since $a(\cdot, \cdot)$ and $(f, \cdot)$ are bounded on $H_0^1(\Omega)$, there is a unique solution to (2.3.2).

REMARK We notice *en passant* that using (2.2.11)

$$\begin{aligned}\mathcal{J}(e) &= \frac{1}{2}a(e,e) - (f,e) + a(u_h,e)\\ &= \frac{1}{2}a(e,e) - a(u,e) + a(u_h,e)\\ &= -\frac{1}{2}a(e,e) = -\frac{1}{2}\parallel e \parallel_E^2 .\end{aligned} \tag{2.3.4}$$

Moreover, using (2.3.2) gives

$$\parallel e \parallel_E = \sqrt{-2\mathcal{J}(e)} \geq \sqrt{-2\mathcal{J}(w)} \quad \forall w \in H_0^1(\Omega). \tag{2.3.5}$$

One interesting consequence of (2.3.5) is that if we have any $w \in H_0^1(\Omega)$ then we can calculate a *lower bound* on the error $\parallel e \parallel_E$. In general we expect this lower bound to be poor unless w is chosen suitably, the best choice being $w = e$. ∎

In practice we are interested in finding an upper bound on the error. It is well known [33], [55], [72] that it is possible to associate an alternative variational principle with the *primal* variational problem (2.3.2). Moreover, it is found that this *complementary* variational principle may be used in a similar manner to that in which the *primal* principle was used in the above remark, with the important difference that an *upper* bound rather than a *lower* bound is obtained. The following example illustrates this procedure for a particular case.

EXAMPLE As an example we consider Poisson's equation in $\Re^2$.

$$-\nabla^2 u = f \text{ in } \Omega$$

with

$$u|_{\partial\Omega} \equiv 0.$$

For this special case the *primal* problem for the error (c.f. (2.3.2)) becomes

$$e \in H_0^1(\Omega) : \mathcal{J}(e) \leq \mathcal{J}(w) \quad \forall w \in H_0^1(\Omega)$$

where

$$\mathcal{J}(w) = \frac{1}{2}\int_\Omega |\nabla w(x)|^2 dx - \int_\Omega f(x)w(x)dx + \int_\Omega \nabla u_h(x) \cdot \nabla w(x)dx.$$

The *complementary* problem is to find $\mathbf{p}$ such that

$$\mathbf{p} \in B : \mathcal{H}(\mathbf{p}) \geq \mathcal{H}(\mathbf{q}) \quad \forall \mathbf{q} \in B$$

where $\mathcal{H}$ is the quadratic functional

$$\mathcal{H}(\mathbf{q}) = -\frac{1}{2}\int_\Omega |\mathbf{q} - \nabla u_h(x)|^2 dx$$

and B is the set

$$B = \{\mathbf{q} \in H^1(\Omega) \times H^1(\Omega) : \nabla \cdot \mathbf{q} + f = 0, \text{ in } \Omega\}.$$

It may be shown (see Ekeland and Témam [33]) that the unique solution of the *complementary* problem is $\mathbf{p} = \nabla u$ and further that

$$-2\mathcal{H}(\nabla u) = \| e \|_E^2 .$$

Combining these results gives

$$\| e \|_E \leq \sqrt{-2\mathcal{H}(\mathbf{q})} \quad \forall \mathbf{q} \in B. \tag{2.3.6}$$

This result shows that in order to obtain a computable upper bound on $\| e \|_E$, all we need do is to find a suitable choice of $\mathbf{q}$ to substitute into the functional $\mathcal{H}(\mathbf{q})$. It is in finding a suitable choice (the best choice is $\mathbf{q} = \nabla u$) that the difficulty lies. It is almost as difficult to find an element of the set B as it is to solve the original problem. ∎

The example illustrates the main difficulty in using the *complementary* principle directly as a means of obtaining error bounds. The constraint condition on the choice of functions which we can use is the main drawback. One possibility is to obtain a suitable $\mathbf{q}$ by means of a finite element discretization of the *complementary* problem [6], [69], [79].

Alternatively, it is possible [45], [46], [47], [48] to produce a suitable $\mathbf{q}$ by solving a series of *local* problems of the form

$$\nabla \cdot \mathbf{q}_i + f = 0 \quad \text{on } \Omega_i$$

where $\mathbf{q}_i$ is the restriction of $\mathbf{q}$ to Ω_i. However, in order to satisfy the regularity condition $\mathbf{q} \in H^1(\Omega) \times H^1(\Omega)$ it is necessary to first solve a *global* problem which imposes the necessary inter-element continuity conditions.

These methods entail the solution of a global problem, essentially to satisfy the continuity requirements. Unfortunately the computational effort required in the solution of any global problem is comparable with that of obtaining the finite element approximation itself. We feel that it is unnecessary to carry out any such global computation since there should be sufficient global information in the finite element approximation itself to enable a choice of $\mathbf{q}$ to made which gives a realistic bound on the error. This view is partially justified on noting that for higher order elements, it has been shown (Bramble and Schatz [26], Thomée [76]) that there exist local averaging operators which allow the true solution and its higher derivatives to be recovered to a high degree of accuracy using the finite element approximation.

Yet another difficulty is that the equality constraint on $\mathbf{q}$

$$\nabla\cdot\mathbf{q} + f = 0 \text{ in } \Omega$$

must be satisfied *exactly*. This is a particularly unsatisfactory state of affairs since it rules out any possibility of using a simple function $\mathbf{q}$, unless f is itself simple. Intuitively one would expect that it should be sufficient to satisfy the condition *sufficiently accurately.*

In order to relax the constraint we shall make use of a device used by Babuška and Rheinboldt [12]. Firstly we define a new bilinear form $\hat{a}(\cdot\,,\cdot)$

$$\hat{a}(u,v) = \int_\Omega \Big\{ \sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial u(x)}{\partial x_i}\frac{\partial v(x)}{\partial x_j} + \lambda u(x)v(x)\Big\}\, dx \qquad (2.3.7)$$

where $\lambda > 0$ is a real constant to be specified later. The following problem may be regarded as a perturbed version of (2.3.1)

$$y \in H_0^1(\Omega) : \hat{a}(y,w) = (f,w) - a(u_h,w) \quad \forall w \in H_0^1(\Omega). \qquad (2.3.8)$$

As before the Lax-Milgram Lemma guarantees the existence of a unique solution to (2.3.8). The solution may equally well be characterized as the solution of the *primal* variational problem

$$y \in H_0^1(\Omega) : \hat{\mathcal{J}}(y) \leq \hat{\mathcal{J}}(w) \quad \forall w \in H_0^1(\Omega) \tag{2.3.9}$$

where

$$\hat{\mathcal{J}}(w) = \frac{1}{2}\hat{a}(w,w) - (f,w) + a(u_h,w). \tag{2.3.10}$$

The following Theorem gives the *complementary* principle associated with the perturbed primal problem.

THEOREM 2.3.1 *Let* $\hat{\mathcal{H}}(\mathbf{p})$ *be the quadratic functional on* $[H^1(\Omega)]^n$ *given by*

$$\hat{\mathcal{H}}(\mathbf{p}) = \int_\Omega (\mathbf{p} - A\nabla u_h)^t A^{-1}(\mathbf{p} - A\nabla u_h)dx + \frac{1}{\lambda}\int_\Omega (f + \nabla\cdot \mathbf{p} - a_0 u_h)^2 dx \tag{2.3.11}$$

then the following bound holds

$$\hat{\mathcal{H}}(\mathbf{p}) \geq \hat{\mathcal{H}}[A\nabla(u_h + y)] = \hat{a}(y,y) \quad \forall \mathbf{p} \in [H^1(\Omega)]^n. \tag{2.3.12}$$

Proof. The strong form of the variational problem (2.3.8) is given by

$$\nabla \cdot [A\nabla(u_h + y)] + f - a_0 u_h = \lambda y. \tag{2.3.13}$$

The unique solution y of the weak form (2.3.8) lies in the space $H_0^1(\Omega)$, and so we have that

$$\nabla \cdot \mathbf{p} + f - a_0 u_h \in L^2(\Omega). \tag{2.3.14}$$

where we have let $\mathbf{p} = A\nabla(y + u_h)$. Moreover,

$$\mathbf{p} - A\nabla u_h = A\nabla y \in L^2(\Omega) \tag{2.3.15}$$

so that $\hat{\mathcal{H}}[\mathbf{p}]$ is well defined, and given by

$$\hat{\mathcal{H}}[\mathbf{p}] = \int_\Omega (\nabla y)^t A(\nabla y) dx + \lambda \int_\Omega y^2 dx = \hat{a}(y, y). \tag{2.3.16}$$

Now let $\epsilon \in [0, 1]$ and $\mathbf{q}, \mathbf{r}$ be any two functions for which $\hat{\mathcal{H}}$ exists. It is easily shown that

$$\hat{\mathcal{H}}[(1 - \epsilon)\mathbf{r} + \epsilon\mathbf{q}] \leq (1 - \epsilon)\hat{\mathcal{H}}[\mathbf{r}] + \epsilon\hat{\mathcal{H}}[\mathbf{q}] \tag{2.3.17}$$

so that $\hat{\mathcal{H}}$ is a convex functional. Moreover, we find that with $\mathbf{p} = A\nabla(y + u_h)$

$$\begin{aligned}
&\frac{1}{2}\frac{d}{d\epsilon}\left\{\hat{\mathcal{H}}[(1 - \epsilon)\mathbf{p} + \epsilon\mathbf{q}]\right\}\Big|_{\epsilon=0} \\
&= \int_\Omega (\mathbf{q} - \mathbf{p})^t \nabla y dx + \int_\Omega y\nabla \cdot (\mathbf{q} - \mathbf{p}) dx \\
&= \int_\Omega \nabla \cdot \left[y(\mathbf{q} - \mathbf{p})\right] dx \\
&= \int_{\partial\Omega} y(\mathbf{q} - \mathbf{p}) \cdot \hat{\mathbf{n}} ds \\
&= 0
\end{aligned}$$

where we have used (2.3.13) and that $y \in H_0^1(\Omega)$. Thus $\hat{\mathcal{H}}$ is stationary at $\mathbf{p}$. The result now follows on noting that $\hat{\mathcal{H}}$ is well defined on $[H^1(\Omega)]^n$. ∎

The result in Theorem 2.3.1 shows that the functional $\hat{\mathcal{H}}(\mathbf{p})$ delivers an upper bound on y measured in the perturbed energy norm defined as

$$\| y \|_{\hat{E}} = \hat{a}(y, y)^{1/2}. \tag{2.3.18}$$

In essence the result given in Theorem 2.3.1 is very similar to the result (2.3.6). However there is an important difference in that if we wished to use (2.3.12) to find an upper bound on $\| y \|_{\hat{E}}$ then there is no *equality* constraint to satisfy, merely a *regularity* requirement. This makes (2.3.12) a far more amenable result, but of course it gives us bounds on $\| y \|_{\hat{E}}$ rather than on $\| e \|_E$. However, the fact that y is the solution of a perturbed version of (2.3.1) which characterizes e leads us to suspect that there is a relationship between the functional $\hat{\mathcal{H}}(\mathbf{p})$ and $\| e \|_E$. This idea is the basis of the following result

THEOREM 2.3.2 *Let $\lambda = Mh^{-\alpha}$ where $\alpha \in (0, 2)$ and $M > 0$ are constants, then the following bound holds for any $\mathbf{p} \in [H^1(\Omega)]^n$*

$$\| e \|_E^2 \leq \{1 + O(h^{1-\alpha/2})\}^2 \hat{\mathcal{H}}(\mathbf{p}) \ \text{as } h \to 0 \tag{2.3.19}$$

where the constant in the $O(h^{1-\alpha/2})$ term is independent of u and h.

Proof. From (2.3.8) we have that

$$y \in H_0^1(\Omega) : \hat{a}(y, w) = (f, w) - a(u_h, w) \quad \forall w \in H_0^1(\Omega)$$

and from (2.3.1) we have that

$$e \in H_0^1(\Omega) : a(e,w) = (f,w) - a(u_h,w) \quad \forall w \in H_0^1(\Omega)$$

so that

$$a(e,w) = \hat{a}(y,w) \quad \forall w \in H_0^1(\Omega).$$

From the definition of $\hat{a}$ we obtain that

$$\begin{aligned} 0 &= a(e,w) - \hat{a}(y,w) \\ &= a(e,w) - \{a(y,w) - (a_0 y, w) + Mh^{-\alpha}(y,w)\} \\ &= a(e-y,w) + \Big((a_0 - Mh^{-\alpha})y, w\Big) \quad \forall w \in H_0^1(\Omega). \end{aligned}$$

Making the choice $w = e - y \in H_0^1(\Omega)$, we have for h sufficiently small

$$\begin{aligned} \| e - y \|_E^2 &= a(e-y, e-y) \\ &= -\Big((a_0 - Mh^{-\alpha})y, e-y\Big) \\ &\leq Ch^{-\alpha} \| y \|_{0,2,\Omega} \| e - y \|_{0,2,\Omega} \\ &\leq Ch^{-\alpha} \{ \| y \|_{0,2,\Omega} + \frac{1}{2} \| e \|_{0,2,\Omega} \}^2. \end{aligned}$$

We now make use of the Aubin-Nitsche Method (see e.g. Ciarlet [27]) to bound the L^2-norm of the error in terms of the energy norm of the error. Firstly, define g to be the unique solution of the problem

$$g \in H_0^1(\Omega) : a(g,w) = (e,w) \quad \forall w \in H_0^1(\Omega)$$

and g_h to be the piecewise linear finite element approximation to g. That is

$$g_h \in V_1^h : a(g_h, w_h) = (e, w_h) \quad \forall w_h \in V_1^h.$$

Since $e \in H_0^1(\Omega)$ it now follows that

$$\begin{aligned} \| e \|_{0,2,\Omega}^2 &= a(g, e) \\ &= a(g - g_h, e) + a(g_h, e) \\ &= a(g - g_h, e) \end{aligned}$$

since $V_1^h \subset V^h$ and since

$$a(e, w_h) \equiv 0 \quad \forall w_h \in V^h.$$

The standard *a priori* error estimate implies that

$$\| g - g_h \|_E \leq Ch \| e \|_E \text{ as } h \to 0,$$

and hence we obtain that

$$\| e \|_{0,2,\Omega} \leq Ch \| e \|_E .$$

Similarly we may define $z \in H_0^1(\Omega)$ to be the solution of the problem

$$\hat{a}(z, w) = (y, w) \quad \forall w \in H_0^1(\Omega),$$

and by an analogous argument, show that

$$\| y \|_{0,2,\Omega} \leq Ch \| y \|_{\hat{E}} \text{ as } h \to 0.$$

Combining these results we get that

$$\| e - y \|_{0,2,\Omega}^2 \leq Ch^{2-\alpha}\{ \| y \|_{\hat{E}} + \| e \|_E \}^2$$

and consequently

$$\| e - y \|_{0,2,\Omega} \leq Ch^{1-\alpha/2}\{ \| y \|_{\hat{E}} + \| e \|_E \}$$

Now by the triangle inequality,

$$\| e \|_E \leq \| y \|_E + \| e - y \|_E$$

and also since

$$\begin{aligned} \| y \|_E^2 &= \| y \|_{\hat{E}}^2 + (a_0 y - Mh^{-\alpha} y, y) \\ &\leq \| y \|_{\hat{E}}^2 (1 + Ch^{2-\alpha}). \end{aligned}$$

where we used the bound

$$\| y \|_{0,2,\Omega} \leq Ch \| y \|_{\hat{E}} .$$

We deduce that

$$\| e \|_E \leq \| y \|_{\hat{E}} (1 + Ch^{2-\alpha})^{1/2} + Ch^{1-\alpha/2}(\| y \|_{\hat{E}} + \| e \|_E)$$

or for h sufficiently small

$$\| e \|_E \leq \| y \|_{\hat{E}} (1 + Ch^{1-\alpha/2})$$

and the result follows on using Theorem 2.3.1 . ∎

Theorem 2.3.2 shows that the functional associated with the perturbed primal problem can be used to obtain approximate upper bounds on $\| e \|_E$. More significantly, the only restriction on the choice of $\mathbf{p}$ is one of *regularity.* The use of the perturbed variational formulation has meant that the equality constraint has been removed at the expense of introducing a second term into $\hat{\mathcal{H}}(\mathbf{p})$.

The question which now arises is that of how we should choose $\mathbf{p}$ to obtain a realistic bound on the error. Theorem 2.3.1 shows that $\hat{\mathcal{H}}(\mathbf{p})$ is minimized by taking $\mathbf{p} = A\nabla(u_h + y)$. Of course it is possible to try to solve (2.3.8) to find y and hence obtain a bound.

An alternative way to interpret Theorem 2.3.2 is to regard it as a theoretical tool which may be used to help in the analysis of the various heuristically proposed error estimators. If such an estimator can be shown to be related to a particular choice of $\mathbf{p}$ in (2.3.19), then Theorem 2.3.2 immediately shows that the resulting estimator will be an *asymptotic* upper bound on the error.

There are many heuristically proposed estimators to be found in the literature, yet it is found that many of them may be profitably viewed within the context of corresponding to a particular choice of $\mathbf{p}$ in Theorem 2.3.2 . In addition to the heuristically based estimators, many of the more rigorously analyzed estimators are also found to fit within this framework.

2.4 Recovery Operators.

In the previous section it was shown that computable bounds on the error measured in the energy norm could be obtained provided we could find a good approximation to the gradient of the true solution. Further, Theorem 2.3.2 provides a theoretical tool which is useful in the analysis of the resulting estimator.

In this section we shall define and analyze a class of schemes which tell us how to make use of u_h in finding a suitable approximation to ∇u. In a subsequent section we shall then analyze the properties of the class of *a posteriori* error estimators obtained by using these approximations to ∇u. This will not only illustrate the use of Theorem 2.3.2 , but will also allow *a posteriori* estimators to be developed which will automatically be asymptotically exact.

In order to maintain a degree of generality, the approach we shall adopt is to firstly define certain *abstract recovery* operators G_h which act on the finite element approximation to give an approximation to the gradient ∇u. The approach is made as general as is feasible in order that it will encompass as many types of approximation as possible. In particular we shall try to find a set of conditions for G_h which will mean that $G_h(u_h)$ is a *good* approximation to the true gradient. This idea is based on generalizing the work of Křížek and Neitaanmäki [50].

The question of what constitutes a *good* approximation to the gradient will be quantified by the condition of *asymptotic exactness* of the resulting *a posteriori* error estimator.

In the following we shall discuss on a rather intuitive level, which properties G_h should satisfy in order that the resulting estimator be asymptotically exact. This will lead to a set of conditions some which will be necessary and some of which will be postulated merely for computational convenience and efficiency.

2.4.1 Consistency condition.

Naturally, if we are to have an asymptotically exact *a posteriori* error estimator then we expect to have to use a recovery scheme which will tend to give an approximation consistent with the true gradient under favourable circumstances. The condition we shall impose is that G_h should estimate the true gradient exactly when the true solution is a polynomial of low degree

(R1) Whenever $u \in \mathcal{P}_{p+1}(\Omega)$

$$G_h(\Pi_p^h u) \equiv \nabla u \qquad (2.4.1)$$

where Π_p^h is the X_h-interpolation operator defined previously. The consistency condition does not determine G_h uniquely, nor is it a necessary condition. It is however an amenable condition and one which provides a manageable criterion with which to work in practice.

2.4.2 Localizing condition.

An important practical requirement on G_h is that it should be as inexpensive as possible to calculate. In particular we assume that it is possible to compute G_h without recourse to global computations (*i.e* without having to solve a system of equations whose size is comparable to the system determining u_h) since the cost entailed would always be comparable to the cost involved in obtaining u_h itself. Of course this *localizing assumption* is not a necessary condition but it is a condition which is very attractive computationally, and one which is not found to be unduly restrictive in practice.

The most convenient schemes are those which mean that $G_h[u_h](x^*)$, $x^* \in \Omega$ can be computed by means of a linear combination of values of ∇u_h sampled on elements Ω_i^h which are near to and include the point x^*. In fact, we shall assume that on any element Ω_i^h, $G_h[u_h]$ can be obtained by using the values of ∇u_h from a subdomain $\hat{\Omega}_i^h$, which is defined as

$$\hat{\Omega}_i^h = \bigcup_{j \in \mathrm{adj}(i)} \Omega_j^h \tag{2.4.2}$$

where $\mathrm{adj}(i)$ is an indexing set containing i and the numbers of those elements which are 'local' to Ω_i^h. In order to ensure that the scheme is truly local and that the domains $\hat{\Omega}_i^h$ are small, we shall also make a restriction on the cardinality of the indexing sets. The localizing condition then becomes

(R2) For $x^* \in \Omega_i^h$, $G_h[v](x^*)$ depends only on the values of ∇v on the domain $\hat{\Omega}_i^h$. Further, $i \in \mathrm{adj}(i)$ and there should exist a bounded constant M, which is independent of h such that, for all i

$$\text{card}[\text{adj}(i)] \leq M. \qquad (2.4.3)$$

EXAMPLE Ideally we would like to make the choice

$$\text{adj}(i) = \{i\} \qquad (2.4.4)$$

but it is found that it is not always possible to satisfy $(R1)$ with such a choice. Another possible choice is

$$\text{adj}(i) = \{j : \Omega_i^h \cap \Omega_j^h \neq \emptyset\} \qquad (2.4.5)$$

so that $\hat{\Omega}_i^h$ is the patch of elements consisting of Ω_i^h and the elements adjacent to Ω_i^h.

REMARK In our applications it will usually be sufficient to make the second choice since most of the currently known superconvergence results (see later) are found to fall in with this choice. The reason why we shall develop the theory for more general choices is to pre-empt the discovery of superconvergence results based upon recovery from larger patches of elements.

2.4.3 Boundedness and linearity conditions.

We also postulate that G_h should be a simple function, so that it may be evaluated and integrated easily. It would be particularly convenient if G_h were to be a piecewise polynomial similar to the finite element approximation itself, since then we can use the existing routines within our finite element code to manipulate G_h. Furthermore, it is only necessary that G_h be defined on the finite element subspace X_h, since we shall only need to apply it to $u_h \in X_h$. Finally, we require G_h to be bounded and linear. These considerations lead us to

(R3) $G_h : X_h \to [X_h]^n$ is a linear operator and there exists a constant C, which is independent of h, such that

$$|G_h[v]|_{0,\infty,\Omega_i^h} \leq C|v|_{1,\infty,\hat{\Omega}_i^h} \quad \forall \Omega_i^h \in \mathcal{T}^h, \quad \forall v \in X_h. \tag{2.4.6}$$

EXAMPLE As an example we consider the case of piecewise linear approximation in one dimension $n = 1$, $p = 1$. We assume that

$$\Omega = (0, 1) \tag{2.4.7}$$

and that

$$\bar{\Omega} = \bigcup_{i=0}^{N^h-1} \Omega_i^h \tag{2.4.8}$$

where, for $i = 0, \dots, N^h - 1$

$$\Omega_i^h = [x_i, x_{i+1}] \tag{2.4.9}$$

and $\{x_i\}$ satisfy

$$0 = x_0 < x_1 < \ldots < x_{N^h} = 1. \tag{2.4.10}$$

The reference element is given by $\hat{K} = [0,1]$ and the affine mappings F_i^h, $i = 0, 1, \ldots, N^h - 1$ are given by

$$x = F_i^h(\hat{x}) \equiv x_i + h_i \hat{x} \tag{2.4.11}$$

where $h_i = x_{i+1} - x_i$. The P-interpolation operator on $\hat{K}$ is denoted by $\hat{\Pi}$ and for $\hat{v} \in C[0,1]$

$$(\hat{\Pi}\hat{v})(\hat{x}) = \hat{x}\hat{v}(1) + (1 - \hat{x})\hat{v}(0).$$

The X_h-interpolation operator on Ω_i^h is given by

$$(\Pi_1^h v) = [\hat{\Pi}(v \circ F_i^h)] \circ (F_i^h)^{-1}, \tag{2.4.12}$$

or more simply as

$$(\Pi_1^h v)(x) = [v(x_i)(x_{i+1} - x) + v(x_{i+1})(x - x_i)]/h_i. \tag{2.4.13}$$

With adj[i] taken as in the second case of the previous example we get, for $i = 1, \ldots, N^h - 2$

$$\hat{\Omega}_i^h = [x_{i-1}, x_{i+2}] \tag{2.4.14}$$

with

$$\widehat{\Omega}_0^h = [x_0, x_2] \tag{2.4.15}$$

and

$$\widehat{\Omega}_{N^h-1}^h = [x_{N^h-2}, x_{N^h}]. \tag{2.4.16}$$

One possible choice of $G_h[v] \in X_h$ is to take

$$G_h(v)(x_i) = \begin{cases} \frac{h_{i-1}v'(x_{i+\frac{1}{2}})+h_i v'(x_{i-\frac{1}{2}})}{h_{i-1}+h_i} & \text{if } i = 1, 2, ..., N^h - 1 \\ \frac{(2h_0+h_1)v'(x_{\frac{1}{2}})-h_0 v'(x_{\frac{3}{2}})}{h_0+h_1} & \text{if } i = 0 \\ \frac{(2h_{N^h-1}+h_{N^h-2})v'(x_{N^h-\frac{1}{2}})-h_{N^h-1}v'(x_{N^h-\frac{3}{2}})}{h_{N^h-1}+h_{N^h-2}} & \text{if } i = N^h \end{cases} \tag{2.4.17}$$

where $x_{i+1/2} = (x_i + x_{i+1})/2$. With this choice it is readily shown that

$$\text{card}[\text{adj}(i)] \leq 3 \tag{2.4.18}$$

$$|G_h(v)|_{0,\infty,\Omega_i^h} \leq 3|v|_{1,\infty,\widehat{\Omega}_i^h} \quad \forall v \in X_h \tag{2.4.19}$$

and

$$G_h[\Pi_1^h v] = \frac{dv}{dx} \quad \forall v \in \mathcal{P}_2(0,1), \tag{2.4.20}$$

and consequently that G_h satisfies (R1)-(R3). ∎

2.4.4 Approximation properties of G_h.

In this section we shall make use of the conditions $(R1)$-$(R3)$ to derive some approximation properties of the operator G_h. In particular, we shall show that for any sufficiently smooth function u, $G_h[\Pi_p^h u]$ is a good approximation to ∇u and further that the gradient of $G_h[\Pi_p^h u]$ is a also a good approximation to the second derivatives of u. Before turning to the derivation of these results, it will be useful to collect some preliminary results.

The first lemma combines the Hölder inequality with the Sobolev Embedding Lemma

LEMMA 2.4.1 *Let $u \in H^s(\Omega)$ where $2s > n$ and $\Omega \subset \Re^n$ is an open, bounded, simply connected and non-empty domain with Lispchitzian boundary, then*

$$|u|_{0,\infty,\Omega} \leq Ch^{-n/2} \parallel u \parallel_{s,2,\Omega} \tag{2.4.21}$$

where $\mathrm{diam}(\Omega) = h$ *and C is a constant which is independent of h and u.*

Proof. Omitted. ∎

The next lemma concerns the boundedness of the interpolation operator Π_p^h defined earlier.

LEMMA 2.4.2 *Let $p \geq 1$ be a fixed integer and let $u \in C(\bar{\Omega})$ be taken arbitrarily, then*

$$|\Pi_p^h u|_{0,\infty,\Omega} \leq C(p)|u|_{0,\infty,\Omega} \tag{2.4.22}$$

where $C(p)$ is a positive constant which is bounded for any fixed value of p.

Proof. Omitted. ∎

The final preparatory lemma makes use of the *inverse assumption* on the regularity of the meshes.

LEMMA 2.4.3 *Let X_h denote the finite element subspace and assume that*

$$\hat{P} \subset H^1(\hat{K}),$$

so that

$$X_h \subset C(\bar{\Omega}) \cap H^1(\Omega),$$

then for any $v_h \in X_h$

$$|v_h|_{1,\infty,\Omega} \leq Ch^{-1} \parallel v_h \parallel_{0,\infty,\Omega} \tag{2.4.23}$$

where C is a constant which is independent of h.

Proof. See Ciarlet [27] (3.2.35). ∎

Armed with these preliminary lemmas we may show the following results.

LEMMA 2.4.4 *Suppose that G_h satisfies $(R1)$-$(R3)$ and that $u \in H^{p+2}(\hat{\Omega}_i^h)$, then*

$$\parallel \nabla u - G_h(\Pi_p^h u) \parallel_{0,2,\Omega_i^h} \leq Ch^{p+1} \mid u \mid_{p+2,2,\hat{\Omega}_i^h} \tag{2.4.24}$$

where C is a constant which is independent of h and u.

Proof. Define the functionals $\{F_k\}_{k=1}^{N^h}$ as follows

$$F_k[u] \equiv \left[\nabla u - G_h(\Pi_p^h u)\right]_k(x) \quad x \in \Omega_i^h \tag{2.4.25}$$

where $[v]_k$ denotes the kth component of $v \in \Re^n$. Since G_h and Π_p^h are both linear functionals, it follows that F_k are also linear. Moreover, F_k can be shown to be bounded functionals as follows. Letting $u \in H^{p+2}(\hat{\Omega}_i^h)$ be taken arbitrarily, we have that

$$\begin{aligned} |F_k[u]|_{0,\infty,\Omega_i^h} &= |\left[\nabla u - G_h(\Pi_p^h u)\right]_k|_{0,\infty,\Omega_i^h} \\ &\leq |u|_{1,\infty,\Omega_i^h} + |G_h(\Pi_p^h u)|_{0,\infty,\Omega_i^h} \end{aligned} \tag{2.4.26}$$

and using Lemma 2.4.1 we have that

$$|u|_{1,\infty,\Omega_i^h} \leq Ch^{-n/2} \parallel u \parallel_{p+2,2,\Omega_i^h} \tag{2.4.27}$$

since for $n = 1, 2, 3$ and $p \geq 1$

$$2(p+1) > n.$$

Further, using (R3) we have that

$$|G_h(\Pi_p^h u)|_{0,\infty,\Omega_i^h} \leq C|\Pi_p^h u|_{1,\infty,\hat{\Omega}_i^h} \tag{2.4.28}$$

and using Lemma 2.4.3 and that $\Pi_p^h u \in [X_h]^n$, we obtain that

$$|\Pi_p^h u|_{1,\infty,\hat{\Omega}_i^h} \leq Ch^{-1}|\Pi_p^h u|_{0,\infty,\hat{\Omega}_i^h}. \tag{2.4.29}$$

Lemma 2.4.2 now gives that

$$|\Pi_p^h u|_{0,\infty,\hat{\Omega}_i^h} \leq C|u|_{0,\infty,\hat{\Omega}_i^h} \tag{2.4.30}$$

and using Lemma 2.4.1 once again, we get that

$$|u|_{0,\infty,\hat{\Omega}_i^h} \leq Ch^{-n/2} \parallel u \parallel_{p+2,2,\hat{\Omega}_i^h} . \tag{2.4.31}$$

Combining (2.4.26)–(2.4.31), we find that

$$\begin{aligned} |F_k[u]|_{0,\infty,\Omega_i^h} &\leq Ch^{-n/2} \parallel u \parallel_{p+2,2,\Omega_i^h} + Ch^{-(1+n/2)} \parallel u \parallel_{p+2,2,\hat{\Omega}_i^h} \\ &\leq Ch^{-(1+n/2)} \parallel u \parallel_{p+2,2,\hat{\Omega}_i^h} \end{aligned} \tag{2.4.32}$$

so that F_k are bounded linear functionals on H^{p+2}. Further since G_h satisfies (R1), we have that

$$F_k[u] \equiv 0 \quad \forall u \in \mathcal{P}_{p+1}(\hat{\Omega}_i^h). \tag{2.4.33}$$

Applying the Bramble-Hilbert Lemma (see Bramble and Hilbert [25]), we deduce that

$$|F_k[u]|_{0,\infty,\Omega_i^h} \leq Ch^{p+1-n/2}|u|_{p+2,2,\hat{\Omega}_i^h} \quad \forall u \in H^{p+2}(\hat{\Omega}_i^h). \tag{2.4.34}$$

Finally, we note that since $\text{meas}(\Omega_i^h) \leq Ch^n$

$$\| \nabla u - G_h(\Pi_p^h u) \|_{0,2,\Omega_i^h}^2 \leq Ch^n |\nabla u - G_h(\Pi_p^h u)|_{0,\infty,\Omega_i^h}^2, \tag{2.4.35}$$

so that using the bound in (2.4.34) we obtain

$$\| \nabla u - G_h(\Pi_p^h u) \|_{0,2,\Omega_i^h}^2 \leq Ch^{2(p+1)}|u|_{p+2,2,\hat{\Omega}_i^h}^2 \tag{2.4.36}$$

which is the desired result. ∎

LEMMA 2.4.5 *Suppose that G_h satisfies $(R1)$-$(R3)$ and that $u \in H^{p+2}(\hat{\Omega}_i^h)$, $2p > n$ then*

$$\| \nabla u - G_h(\Pi_p^h u) \|_{1,2,\Omega_i^h} \leq Ch^p |u|_{p+2,2,\hat{\Omega}_i^h} \tag{2.4.37}$$

where C is a constant which is independent of h and u.

Proof. Define the functionals $\{R_{kl}\}$ for $k, l \in \{1, \ldots, n\}$

$$R_{kl}(u)(x) = \frac{\partial^2 u(x)}{\partial x_k \partial x_l} - \frac{\partial}{\partial x_k}[G_h(\Pi_p^h u)]_l(x). \tag{2.4.38}$$

Since G_h and Π_p^h are both linear functionals, R_{kl} are also linear functionals. Moreover, we claim that R_{kl} are bounded since

$$
\begin{aligned}
|R_{kl}[u]|_{0,\infty,\Omega_i^h} &= \left|\frac{\partial^2 u(x)}{\partial x_k \partial x_l} - \frac{\partial}{\partial x_k}[G_h(\Pi_p^h u)]_l\right|_{0,\infty,\Omega_i^h} \\
&\leq |\nabla u - G_h(\Pi_p^h u)|_{1,\infty,\Omega_i^h} \\
&\leq |u|_{2,\infty,\Omega_i^h} + |G_h(\Pi_p^h u)|_{1,\infty,\Omega_i^h} \qquad (2.4.39)
\end{aligned}
$$

By Lemma 2.4.1 we have that

$$
|u|_{2,\infty,\Omega_i^h} \leq Ch^{-n/2} \parallel u \parallel_{p+2,2,\Omega_i^h} \qquad (2.4.40)
$$

since $2p > n$. Moreover, since $G_h \in [X_h]^n$, we may use Lemma 2.4.3 to deduce that

$$
|G_h(\Pi_p^h u)|_{1,\infty,\Omega_i^h} \leq Ch^{-1}|G_h(\Pi_p^h u)|_{0,\infty,\Omega_i^h}. \qquad (2.4.41)
$$

Further, since G_h satisfies (R3) and by Lemma 2.4.3

$$
|G_h(\Pi_p^h u)|_{0,\infty,\Omega_i^h} \leq C|\Pi_p^h u|_{1,\infty,\hat{\Omega}_i^h} \leq Ch^{-1}|\Pi_p^h u|_{0,\infty,\hat{\Omega}_i^h}. \qquad (2.4.42)
$$

Using (2.4.41) and (2.4.42), and proceeding as in the proof of Lemma 2.4.4 , we obtain that

$$
|G_h(\Pi_p^h u)|_{1,\infty,\Omega_i^h} \leq Ch^{-(2+n/2)} \parallel u \parallel_{p+2,2,\hat{\Omega}_i^h} . \qquad (2.4.43)
$$

Thus, combining (2.4.40) and (2.4.43) we have that

$$|R_{kl}[u]|_{0,\infty,\Omega_i^h} \leq Ch^{-(2+n/2)} \parallel u \parallel_{p+2,2,\hat{\Omega}_i^h} . \tag{2.4.44}$$

Thus R_{kl} are bounded linear functionals on $H^{p+2}(\hat{\Omega}_i^h)$, and since G_h satisfies (R1) we have that

$$R_{kl}[u] \equiv 0 \tag{2.4.45}$$

whenever $u \in \mathcal{P}_{p+1}$. Applying the Bramble-Hilbert Lemma we obtain that

$$|R_{kl}[u]|_{0,\infty,\Omega_i^h} \leq Ch^{p-n/2}|u|_{p+2,2,\hat{\Omega}_i^h}, \tag{2.4.46}$$

and finally since $\text{meas}(\Omega_i^h) \leq Ch^n$, we get that

$$\parallel \nabla u - G_h(\Pi_p^h u) \parallel_{1,2,\Omega_i^h}^2 \leq Ch^{2p}|u|_{p+2,2,\hat{\Omega}_i^h}^2 \tag{2.4.47}$$

as required. ∎

LEMMA 2.4.6 *Suppose that G_h satisfies (R1)-(R3) and that $u \in H^{p+2}(\Omega)$, then*

$$|\nabla u - G_h(\Pi_p^h u)|_{0,2,\Omega} \leq Ch^{p+1}|u|_{p+2,2,\Omega} \tag{2.4.48}$$

where C are constants independent of h and u.

Proof. Using Lemma 2.4.4 we deduce that

$$
\begin{aligned}
\| \nabla u - G_h(\Pi_p^h u) \|_{0,2,\Omega}^2 &= \sum_{i=1}^{N^h} \| \nabla u - G_h(\Pi_p^h u) \|_{0,2,\Omega_i^h}^2 \\
&\leq \sum_{i=1}^{N^h} C h^{2(p+1)} |u|_{p+2,2,\hat{\Omega}_i^h}^2 \\
&\leq C h^{2(p+1)} \sum_{i=1}^{N^h} |u|_{p+2,2,\hat{\Omega}_i^h}^2 . \qquad (2.4.49)
\end{aligned}
$$

Since the subdomains $\hat{\Omega}_i^h$ satisfy (R2) we have that

$$
\sum_{i=1}^{N^h} |u|_{p+2,2,\hat{\Omega}_i^h}^2 \leq M \sum_{i=1}^{N^h} |u|_{p+2,2,\Omega_i^h}^2 = M |u|_{p+2,2,\Omega}^2 . \qquad (2.4.50)
$$

Combining (2.4.49) and (2.4.50) yields the desired result. ∎

LEMMA 2.4.7 *Suppose that G_h satisfies $(R1)$-$(R3)$ and that $u \in H^{p+2}(\Omega)$, $2p > n$ then*

$$
|\nabla u - G_h(\Pi_p^h u)|_{1,2,\Omega} \leq C h^p |u|_{p+2,2,\Omega} \qquad (2.4.51)
$$

where C are constants independent of h and u.

Proof. Follows from Lemma 2.4.5 in the same way as Lemma 2.4.6 followed from Lemma 2.4.4 . ∎

2.5 The superconvergence property.

In the previous section we showed that if a recovery operator G_h could be found satisfying the conditions $(R1)$–$(R3)$, then applying the operator to $\Pi_p^h u$ would furnish us with good approximations to $\frac{\partial u}{\partial x_i}$ and to $\frac{\partial^2 u}{\partial x_i \partial x_j}$. In this section we show, using these abstract approximation properties of G_h combined with the *superconvergence* property of finite element approximation, that we can obtain good approximations to $\frac{\partial u}{\partial x_i}$ and to $\frac{\partial^2 u}{\partial x_i \partial x_j}$. Specifically, we show that if a superconvergence property holds then $G_h[u_h]$ will possess approximation properties similar to those of $G_h[\Pi_p^h u]$.

In order to introduce the superconvergence property, we recall the *a priori* error estimates for the discretization error measured in the norm on $H_0^1(\Omega)$

$$\| u - u_h \|_{1,2,\Omega} \leq Ch^p |u|_{p+1,2,\Omega} \tag{2.5.1}$$

where we assume that the inclusions

$$\mathcal{P}_p(\hat{K}) \subset \hat{P} \subset H^1(\hat{K}) \tag{2.5.2}$$

hold, and p is the largest integer for which this is valid. The *a priori* estimate (2.5.1) is optimal in the sense that the exponent of h is the largest possible. In fact, if the true solution $u \notin V_h$ then an *inverse estimate* of the form

$$\| u - u_h \|_{1,2,\Omega} \geq C(u) h^p \tag{2.5.3}$$

is valid [70] for some constant $C(u)$ which depends on u but not on h. Together (2.5.1) and (2.5.3) show that ∇u_h approximates ∇u to $O(h^p)$ in a root mean square sense.

It has been shown under certain conditions regarding the regularity of the partition, the regularity of the true solution and the topology of the mesh, that estimates of the form

$$|u_h - \Pi_p^h u|_{1,2,\Omega} \leq C(u)h^{p+1} \tag{2.5.4}$$

hold.

REMARK The constant $\mathcal{C}(u)$ takes different forms depending on whether exact numerical integration is assumed to be performed throughout the finite element approximation process. If all integrals are evaluated exactly then

$$\mathcal{C}(u) = |u|_{p+1,2,\Omega} + |u|_{p+2,2,\Omega} \tag{2.5.5}$$

whilst if an approximate quadrature rule of sufficiently high precision is used then

$$\mathcal{C}(u) = |u|_{p+1,2,\Omega} + |u|_{p+2,2,\Omega} + |u|_{p+3,2,\Omega}. \tag{2.5.6}$$

Further details may be found in [5], [50], [51], [52], [53], [54], [91], [92]. ∎

Combining (2.5.4) with (2.5.3) gives

$$|u_h - \Pi_p^h u|_{1,2,\Omega} \leq C(u)h \parallel u - u_h \parallel_{1,2,\Omega}, \tag{2.5.7}$$

which shows that ∇u_h is a better approximation to $\nabla \Pi_p^h u$ by a whole order of h than it is to ∇u. This is the superconvergence phenomenon associated with the finite element method, and has been demonstrated for wide classes of finite element approximation schemes. We shall refer to an estimate of the form

$$\text{(SC)} \qquad |u_h - \Pi_p^h u|_{1,2,\Omega} \leq C(u) h^{p+1} \tag{2.5.8}$$

as the superconvergence property.

LEMMA 2.5.1 *Suppose $u \in H^{p+2}(\Omega)$, that (SC) is valid and that G_h satisfies $(R1)$-$(R3)$, then*

$$\| \nabla u - G_h(u_h) \|_{0,2,\Omega} \leq C h^{p+1} \{ |u|_{p+2,2,\Omega} + C(u) \} \tag{2.5.9}$$

holds where the C is a constant independent of h and u.

Proof. By using the Triangle Inequality and the linearity property of G_h we have for any $\Omega_i^h \in \mathcal{T}^h$ that

$$\| \nabla u - G_h(u_h) \|_{0,2,\Omega_i^h} \leq \| \nabla u - G_h(\Pi_p^h u) \|_{0,2,\Omega_i^h} + \| G_h(\Pi_p^h u - u_h) \|_{0,2,\Omega_i^h} . \tag{2.5.10}$$

The boundedness property (R3) of G_h gives

$$\| G_h(\Pi_p^h u - u_h) \|_{0,2,\Omega_i^h} \leq C |\Pi_p^h u - u_h|_{1,2,\hat{\Omega}_i^h} . \tag{2.5.11}$$

Further from Lemma 2.4.4 we have

$$\| \nabla u - G_h(\Pi_p^h u) \|_{0,2,\Omega_i^h} \leq C h^{p+1} |u|_{p+2,2,\hat{\Omega}_i^h} . \tag{2.5.12}$$

Combining (2.5.10), (2.5.11) and (2.5.12) we obtain

$$\| \nabla u - G_h(u_h) \|_{0,2,\Omega_i^h} \leq C|\Pi_p^h u - u_h|_{1,2,\hat{\Omega}_i^h} + Ch^{p+1}|u|_{p+2,2,\hat{\Omega}_i^h}. \tag{2.5.13}$$

Now using (2.5.13) gives

$$\begin{aligned}
\| \nabla u - G_h(u_h) \|_{0,2,\Omega}^2 &= \sum_{i=1}^{N^h} \| \nabla u - G_h(u_h) \|_{0,2,\Omega_i^h}^2 \\
&\leq C \sum_{i=1}^{N^h} \left[|\Pi_p^h u - u_h|_{1,2,\hat{\Omega}_i^h} + h^{p+1}|u|_{p+2,2,\hat{\Omega}_i^h} \right]^2 \\
&\leq C \sum_{i=1}^{N^h} |\Pi_p^h u - u_h|_{1,2,\hat{\Omega}_i^h}^2 + Ch^{2(p+1)} \sum_{i=1}^{N^h} |u|_{p+2,2,\hat{\Omega}_i^h}^2 \\
&\leq CM \sum_{i=1}^{N^h} |\Pi_p^h u - u_h|_{1,2,\Omega_i^h}^2 + CMh^{2(p+1)} \sum_{i=1}^{N^h} |u|_{p+2,2,\Omega_i^h}^2 \\
&\leq CM|\Pi_p^h u - u_h|_{1,2,\Omega}^2 + CMh^{2(p+1)}|u|_{p+2,2,\Omega}^2, \qquad (2.5.14)
\end{aligned}$$

where we have also made use of property (R2) and the inequality

$$(a+b)^2 \leq 2(a^2 + b^2).$$

Finally, using (SC) and (2.5.14) we have that

$$\| \nabla u - G_h(u_h) \|_{0,2,\Omega}^2 \leq Ch^{2(p+1)} [\mathcal{C}(u)^2 + |u|_{p+2,2,\Omega}^2], \tag{2.5.15}$$

and the result follows as claimed. ∎

There is also an analogous estimate for the second derivatives

LEMMA 2.5.2 *Suppose $u \in H^{p+2}(\Omega)$, $2p > n$, that (SC) is valid and that G_h satisfies $(R1)$-$(R3)$, then*

$$\| \nabla u - G_h(u_h) \|_{1,2,\Omega} \leq C h^p \{ |u|_{p+2,2,\Omega} + C(u) \} \tag{2.5.16}$$

holds where the C is a constant independent of h and u.

Proof. Follows in essentially the same way as the proof of Lemma 2.5.1 . ∎

2.5.1 Asymptotic exactness of the estimators.

We now analyze the behaviour of the class of *a posteriori* error estimators obtained by using $G_h[u_h]$ instead of ∇u in the expression for the error. That is, the estimator is taken to be ϵ

$$\epsilon^2 = \sum_{i=1}^{N^h} \epsilon_i^2 \tag{2.5.17}$$

where for $i = 1, \ldots, N^h$

$$\epsilon_i^2 = \int_{\Omega_i^h} (G_h[u_h] - \nabla u_h)^t A (G_h[u_h] - \nabla u_h)\, dx. \tag{2.5.18}$$

The main result we shall show is that the estimator will be asymptotically exact provided that G_h satisfies the recovery conditions $(R1)$–$(R3)$ and provided that the superconvergence property (SC) holds.

THEOREM 2.5.3 *Let ϵ be the a posteriori error estimator defined above, and assume that (SC) and $(R1)$–$(R3)$ hold, then ϵ is an asymptotically exact estimator. That is*

$$\| e \|_E = \epsilon(1 + Ch^{\gamma}) \qquad as\ h \to 0 \tag{2.5.19}$$

where $\gamma > 0$ and C are constants independent of h.

Proof. Let $\hat{\mathcal{H}}[\mathbf{p}]$ be the quadratic functional of Theorem 2.3.2 and make the choice $\mathbf{p} = AG_h(u_h)$ in Theorem 2.3.2 . This is a valid choice since

$$G_h(u_h) \in [X_h]^n \text{ and } X^h \subset C(\bar{\Omega}).$$

With this choice we have that

$$\hat{\mathcal{H}}[AG_h(u_h)] = \epsilon^2 + \frac{h^{\alpha}}{M}\Lambda^2 \tag{2.5.20}$$

where

$$\epsilon^2 = \int_{\Omega} [G_h(u_h) - \nabla u_h]^t A[G_h(u_h) - \nabla u_h] dx \tag{2.5.21}$$

and

$$\Lambda^2 = \int_{\Omega} \{f + \nabla \cdot [AG_h(u_h)] - a_0 u_h\}^2 dx. \tag{2.5.22}$$

Firstly, we consider the term Λ.

$$\begin{aligned}
\Lambda &= \| f + \nabla \cdot [AG_h(u_h)] - a_0 u_h \|_{0,2,\Omega} \\
&= \| -\nabla \cdot [A\nabla u - AG_h(u_h)] + a_0 e \|_{0,2,\Omega} \\
&\leq \| \nabla \cdot [A\nabla u - AG_h(u_h)] \|_{0,2,\Omega} + \| a_0 e \|_{0,2,\Omega} \\
&\leq C(A)\{\| \nabla u - G_h(u_h) \|_{0,2,\Omega} + \| \nabla u - G_h(u_h) \|_{1,2,\Omega}\} \\
&\qquad + C(a_0) \| e \|_{0,2,\Omega} \qquad (2.5.23)
\end{aligned}$$

where $C(A)$ is a constant depending on A, and we assume that A is sufficiently smooth.

Using (2.5.23) and Lemmas 2.5.1 and 2.5.2

$$\begin{aligned}
\Lambda \leq h^p \quad \| A \|_{s,2,\Omega} \, [C_1\{ |u|_{p+2,2,\Omega} + \mathcal{C}(u)\} \\
+ C_2\{ |u|_{r,2,\Omega} + \mathcal{C}(u)\}] + C \| e \|_{0,2,\Omega}, \qquad (2.5.24)
\end{aligned}$$

Using the Aubin Nitsche method (see e.g. Ciarlet [27]) we obtain that

$$\| e \|_{0,2,\Omega} \leq Ch \| e \|_E \, . \qquad (2.5.25)$$

Combining (2.5.24) with (2.5.25) and the inverse estimate (2.5.3) yields the bound

$$\Lambda \leq C(u) \| e \|_E \{1 + Ch\} \qquad (2.5.26)$$

for h sufficiently small.

Considering now the term ϵ we have

$$\begin{aligned}\epsilon &= \left\{ \int_\Omega [G_h(u_h) - \nabla u_h]^t A [G_h(u_h) - \nabla u_h] dx \right\}^{\frac{1}{2}} \\ &\leq \left\{ \int_\Omega (\nabla e)^t A (\nabla e) dx \right\}^{\frac{1}{2}} + \left\{ \int_\Omega [\nabla u - G_h(u_h)]^t A [\nabla u - G_h(u_h)] dx \right\}^{\frac{1}{2}} \\ &\leq \| e \|_E + C \| \nabla u - G_h(u_h) \|_{0,2,\Omega} . \end{aligned} \tag{2.5.27}$$

Applying Lemma 2.5.1 to (2.5.27) we obtain that

$$\epsilon \leq \| e \|_E + C h^{p+1} \{ |u|_{p+2,2,\Omega} + C(u) \} \tag{2.5.28}$$

and applying the inverse estimate (2.5.3) gives

$$\epsilon \leq [1 + C(u)h] \| e \|_E \text{ as } h \to 0. \tag{2.5.29}$$

Now using (2.5.26), (2.5.29) and Theorem 2.3.2 gives

$$\| e \|_E^2 \leq \{1 + C h^{1-\alpha/2}\}^2 [\epsilon^2 + C(u) \| e \|_E^2 (1 + Ch)^2\} \tag{2.5.30}$$

which in turn implies that

$$\| e \|_E^2 \leq (1 + C h^{1-\alpha/2})^2 \epsilon^2 \text{ as } h \to 0. \tag{2.5.31}$$

Together (2.5.29) and (2.5.31) imply that

$$\| e \|_E = \epsilon \{1 + O(h^\gamma)\} \text{ as } h \to 0 \tag{2.5.32}$$

where $\gamma > 0$ is a constant independent of h. ∎

Theorem 2.5.3 reduces the problem of finding *a posteriori* error estimators to that of using the existing superconvergence results to define an appropriate recovery operator G_h. Consequently, whenever we have superconvergence results for a particular finite element scheme, it is then possible to define an *a posteriori* error estimator which is asymptotically exact. Moreover, the cost of computing the error estimator is to all intents negligible and entails a minimal amount of programming effort since most finite element codes have a post processing procedure already implemented.

Throughout we have assumed that the true solution u is regular. In practice these regularity assumptions are seldom satisfied. However it is still possible to use the foregoing framework to develop estimators for problems with singularities. The reader is referred to Ainsworth [1] and to Chapter 4 for further details.

2.6 Examples of recovery based estimators.

In this section we shall illustrate how the foregoing results may be used in the development of *a posteriori* error estimators. For our examples we consider three types of finite element approximation scheme. These will demonstrate; how an existing estimator falls within the framework which we have developed; how an existing estimator may be viewed as a simplified version of a recovery based estimator; and, how an estimator may be easily obtained for a new situation.

2.6.1 The Babuška and Rheinboldt estimator.

In this example we consider the case of piecewise linear finite element approximation in one dimension $n = 1$, $p = 1$. There are many alternative types of *a posteriori* error estimator available for this situation, many of which have been exhaustively analyzed theoretically (e.g. Babuška and Rheinboldt [12]). We do not anticipate our approach to furnish us with new theoretical results. However, the example will show how an existing estimator may be viewed from within our framework.

In an earlier example we defined a recovery operator G_h which satisfied the recovery conditions (R1)-(R3). In order to apply Theorem 2.5.3 we only need to show that the superconvergence property (SC) holds. In fact the superconvergence property has been demonstrated for this case, and may be obtained by using standard arguments similar to those found in Zlámal [91],[92] for example.

We define our estimator to be ϵ,

$$\epsilon^2 = \sum_{i=1}^{N^h} \epsilon_i^2 \tag{2.6.1}$$

where, for $i = 1, \ldots, N^h$

$$\epsilon_i^2 = \int_{\Omega_i^h} A(x)\{G_h(u_h) - \frac{du_h}{dx}\}^2 dx. \tag{2.6.2}$$

It is found that the resulting estimator is precisely the same as an estimator already proposed and analyzed in the literature (Babuška and Rheinboldt, [12]: Definition 6.3). Previously, the estimator was obtained by means of an argument based on locally projecting the error onto a quadratic function which vanishes at the nodes of the partition. For further details the reader is referred to Babuška and Rheinboldt [12] and the references therein. Numerical examples illustrating the effectiveness of this estimator are also given in the aforementioned reference. In addition, we shall return once again to the Babuška and Rheinboldt estimator in Chapter 4 where it will be generalized to cases where the true solution is singular.

2.6.2 The Kelly, Gago, Zienkiewicz and Babuška estimator.

In this example we consider the finite element scheme consisting of piecewise bilinear approximation in two dimensions. That is, the reference element is taken to be

$$\hat{K} = [0,1] \times [0,1] \tag{2.6.3}$$

the approximation space $\hat{P}$ is taken as

$$\hat{P} = Q_1(\hat{K}), \tag{2.6.4}$$

and the linear forms $\{\phi_i\}$ are the standard Lagrange functions based on the vertices of $\hat{K}$ (see e.g. Ciarlet [27]). For the sake of simplicity we shall assume that the model problem reduces to Poisson's equation in two dimensions

$$-\nabla^2 u \equiv f \tag{2.6.5}$$

and also that each of the subdomains Ω_i^h is a square with sides of length h parallel to the x and y axes (see Fig. 2.1). We define the recovery operator G_h to be piecewise bilinear in each component and assign G_h the following values at the nodes of the partition.

If (x_i, y_j) is an *internal* node (*i.e.* does not lie on the boundary of Ω), then

$$G_h[v](x_i, y_j) = \frac{1}{4}[\nabla v|_{i+\frac{1}{2},j+\frac{1}{2}} + \nabla v|_{i-\frac{1}{2},j+\frac{1}{2}} + \nabla v|_{i-\frac{1}{2},j-\frac{1}{2}} + \nabla v|_{i+\frac{1}{2},j-\frac{1}{2}}] \tag{2.6.6}$$

where

$$\nabla v|_{i+\frac{1}{2},j+\frac{1}{2}} \equiv \nabla v(x_i + h/2, y_j + h/2). \tag{2.6.7}$$

If (x_i, y_j) is a *boundary* node then we define $G_h[v](x_i, y_j)$ to be the value at (x_i, y_j) of the bilinear function which interpolates to ∇u_h at the centroids of the elements which are nearest to the point (x_i, y_j). Since the X_h interpolant is uniquely determined by the nodal values, this definition means that G_h is well defined.

It is found that this recovery scheme falls within our framework as follows. The subdomains $\hat{\Omega}_i^h$ are based on the indexing sets adj(i) defined by

$$\text{adj}(i) = \{j : \Omega_i^h \cap \Omega_j^h \neq \emptyset\}. \tag{2.6.8}$$

Moreover, it is found that the cardinality of these sets is bounded independently of h,

$$\text{card}[\text{adj}(i)] \leq 9. \tag{2.6.9}$$

It is easily seen that G_h is linear and is bounded in the sense that

$$|G_h[v]|_{0,\infty,\Omega_i^h} \leq 4|v|_{1,\infty,\hat{\Omega}_i^h} \quad \forall v \in X_h. \tag{2.6.10}$$

Finally, it is easy to verify that

$$G_h[\Pi_1^h v] \equiv \nabla v \quad \forall v \in \mathcal{P}_2(\Omega). \tag{2.6.11}$$

Consequently, the recovery operator is seen to satisfy the conditions (R1)–(R3) and it remains only to demonstrate the superconvergence property.

REMARK In fact the superconvergence property has been shown not only for this case, but also for higher degree approximation on quadrilateral subdomains including *serendipity* elements (see Zlámal [91],[92] and Lesaint and Zlámal [52]. Superconvergence results have also been demonstrated for triangular elements (see Křížek and Neitaanmäki [50] and Levine [53], [54]). It should however be borne in mind however that these results are valid only under restrictions on the regularity of the mesh. ∎

Theorem 2.5.3 now gives us that the estimator ϵ

$$\epsilon^2 = \sum_{i=1}^{N^h} \epsilon_i^2 \tag{2.6.12}$$

where for $i = 1, \ldots, N^h$

$$\epsilon_i^2 = \int_{\Omega_i^h} |G_h[u_h] - \nabla u_h|^2 dx \tag{2.6.13}$$

is asymptotically exact.

It is interesting to compare the new estimator with an existing estimator $\hat{\epsilon}$ used for this approximation scheme and given in Kelly *et al.* [46]

$$\hat{\epsilon}^2 = \sum_{i=1}^{N^h} \hat{\epsilon}_i^2 \tag{2.6.14}$$

where for $i = 1, \ldots, N^h$

$$\hat{\epsilon}_i^2 = \frac{h}{24} \int_{\partial\Omega_i^h} J^2 ds, \tag{2.6.15}$$

and where J is the 'jump' across the element boundary in the finite element approximation to the gradient. Using the midpoint rule for integration, along each side of the element, (2.6.15) may be rewritten as

$$\hat{\epsilon}_i^2 = \frac{h^2}{24} \sum_{k=1}^{4} J_k^2, \tag{2.6.16}$$

where J_k now denotes the 'jump' in the normal derivative across the boundary (it will be noticed that the there is no discontinuity across the element boundary in the tangential component of the gradient). In Kelly *et al.* [46] it is noted that the estimator bears out practical experience that the accuracy of the approximation is related to the discontinuity of the direct approximation to the gradient across the interelement boundaries. In Zienkiewicz *et al.* [86], it is stated that *"the derivation of (2.6.16) is complex and subject to many heuristic arguments"*. Indeed the constant 1/24 appearing in (2.6.16) is obtained by satisfying the condition that the estimator should be exact if the true solution is quadratic and $u_h = \Pi_1^h u$ (see Kelly *et al.* [46]).

The new estimator (2.6.13) is also found to be exact if the true solution u is quadratic and if $u_h = \Pi_1^h u$, but Theorem 2.5.3 shows that this estimator will be asymptotically exact more generally. Moreover, like (2.6.16) this estimator is found (after a lengthy but otherwise straight forward manipulation) to depend on the discontinuities in the finite element approximation to the gradient. However, the dependence is more intricate than in (2.6.16), also involving jumps diagonally across elements (e.g. between elements 1 and 2 in Fig. 2.2), and also differences in tangential components (e.g. between the y components of the gradient at the centroid of elements 1 and 3 in Fig. 2.2).

It is possible to simplify (2.6.13) by approximating the terms involving the jumps in the gradient diagonally across elements and the differences in tangential components of the gradient. In order to obtain these approximations we make use of the fact that averaging the direct approximation ∇u_h at the midpoints of the sides of the element gives a resulting value which is a superconvergent approximation to the gradient of the true solution. For example, we approximate the gradient at the midpoint (x_m, y_m) of the side connecting elements 1 and 3 by

$$\nabla u(x_m, y_m) \approx \frac{1}{2}\Big\{\nabla u_h(x_m, y_m + \xi) + \nabla u_h(x_m, y_m - \xi)\Big\} \qquad (2.6.17)$$

where ξ is arbitrarily small.

Carrying out this averaging at the midpoints of all of the sides of an element gives enough information to approximate (by means of bilinear extrapolation) the gradient at the centroids of the surrounding elements. Now, by applying G_h to these values (rather than to the direct approximation ∇u_h) gives us an approximation $\tilde{\epsilon}_i^2$ to ϵ_i^2.

In fact it is found that $\tilde{\epsilon}_i \equiv \hat{\epsilon}_i$. That is, the estimator derived by Kelly *et al.* may be obtained by simplifying the estimator derived using the above framework. Moreover, we may claim that *the derivation of (2.6.16) is straight forward and rests upon a sound theoretical footing.*

2.6.3 An estimator for quadratic approximation.

In this example we consider the finite element approximation using piecewise quadratic functions in one dimension. The superconvergence property holds true for this situation and may be shown using standard arguments (see e.g. Zlámal [91],[92]). A recovery operator G_h may be defined by exploiting the result that if the true solution is cubic then the true gradient u' and the gradient of the quadratic interpolant $\Pi_2^h u$, coincide at the nodes used in the 2-point Gauss Legendre quadrature rule on the element. That is, for the element $[x_i, x_{i+1}]$ we use the points

$$x_i^{\pm} = \frac{1}{2}\{x_i + x_{i+1} \pm \frac{1}{\sqrt{3}}(x_{i+1} - x_i)\}. \tag{2.6.18}$$

In order to define G_h we need to find a way of recovering the gradient at the nodes and the centroid of each element. Obviously this may be done by suitably extrapolating the gradient recovered at the Gauss Legendre points. There are many possible ways to carry out this process (most of which fall within our framework), but the one we select is to use a *cubic* interpolation process to extrapolate the gradient. In fact a *quadratic* process would meet the recovery criteria (R1)–(R3) and would lead to a more economical but less 'symmetrical process'. The cubic scheme may be summarized as follows:

For $i = 1, \ldots, N^h - 1$,

- at the centroid of element $[x_i, x_{i+1}]$, $G_h[v]$ is taken to be the value at $\frac{1}{2}\{x_i + x_{i+1}\}$ of the cubic interpolating to v' at the points

$$\{x_{i-1}^+, x_i^-, x_i^+, x_{i+1}^-\} \tag{2.6.19}$$

- at the node x_i, $G_h[v]$ is taken to be the value at x_i of the cubic interpolating to v' at the points

$$\{x_{i-1}^-, x_{i-1}^+, x_i^-, x_i^+\}. \tag{2.6.20}$$

For $i = 0$,

- at the centroid of element $[x_0, x_1]$, $G_h[v]$ is taken to be the value at $\frac{1}{2}\{x_0 + x_1\}$ of the cubic interpolating to v' at the points

$$\{x_0^-, x_0^+, x_1^-, x_1^+\} \tag{2.6.21}$$

- at the node x_0, $G_h[v]$ is taken to be the value at x_i of the cubic interpolating to v' at the points

$$\{x_0^-, x_0^+, x_1^-, x_1^+\}. \tag{2.6.22}$$

The operator is defined analogously in the case $i = N^h$.

Having assigned values to $G_h[v]$ at the nodes and at the centroids, we take $G_h[v]$ to be the X_h-interpolant to these values. In order to show that the resulting estimator is asymptotically exact, it remains only to verify that conditions (R1)–(R3) are satisfied. The subdomains $\hat{\Omega}_i^h$ may be defined in the usual way

$$\text{adj}(i) = \{j : \Omega_i^h \cap \Omega_j^h \neq \emptyset\}, \tag{2.6.23}$$

so that

$$\text{card}[\text{adj}(i)] \leq 3. \tag{2.6.24}$$

Since the extrapolation process used to obtain the values at the nodes and at the centroids was based on a cubic, the recovered values obtained using $\Pi_2^h v$ would be the true values of v', whenever v is itself cubic. This means that whenever $v \in \mathcal{P}_3$

$$G_h[\Pi_2^h v] \equiv v' \tag{2.6.25}$$

and hence (R1) is satisfied. Finally, we see that G_h is linear and bounded

$$|G_h[v]|_{0,\infty,\Omega_i^h} \leq \sqrt{3}(1+2C)|v|_{1,\infty,\hat{\Omega}_i^h} \quad \forall v \in X_h. \tag{2.6.26}$$

where C is the constant appearing in (2.2.23). Consequently conditions (R1)–(R3) and (SC) hold and Theorem 2.5.3 guarantees that the estimator is asymptotically exact. It is found that the estimator, when written out explicitly, may be expressed in terms of differences in the direct approximation to the gradient. However, it is entirely unnecessary to derive such an expression since the recovery process combined with a quadrature rule provides a simple method of implementation. Examples showing the performance of the new estimator are presented in the next section.

2.7 Numerical Examples.

In order to demonstrate the behaviour of the new estimators in practice, we present the results obtained in several numerical examples in which the new estimators are employed.

2.7.1 The quadratic estimator.

The model problem in $\Re^1$ reduces to the form

$$-\frac{d}{dx}[A(x)\frac{du}{dx}] + B(x)u(x) = f(x). \tag{2.7.1}$$

We shall specify A, B in each case and choose f so that the true solution is

$$u(x) = e^x + x + \sin 6x + 1 - ex - x\sin 6.$$

The results are presented in Tables 2.1–5 where the notation is as follows

N - number of *uniform* elements *i.e.* $h = h_i \quad \forall i$

$\|e\|_E$ - true value of the error in the energy norm

ϵ - estimated value of the error in the energy norm

θ - effectivity index E/ϵ.

Tables 2.1–5 include results obtained for problems where $A(x)$ has a wide variation and where $A(x)$ is small compared with $B(x)$. In the latter case we might expect the results to be poor because the there is no term in the estimator which corresponds to the term

$$\int_\Omega B(x)e(x)^2dx \tag{2.7.2}$$

in the expression for the true error. Nevertheless the effectivity indices tend towards unity, indicating that the estimator is asymptotically exact.

2.7.2 Kelly, Gago, Zienkiewicz and Babuška estimator.

The model problem we consider is Poisson's equation on the unit square with Dirichlet boundary conditions prescribed on the whole of the boundary. We approximate the problem using a partition consisting of square elements with sides of length h which are parallel to the x or the y axes, and using bilinear basis functions. The results obtained using both the recovery based estimator described in the previous section, and the simplified version of this estimator (*i.e.* the Kelly *et al.* estimator) are given in Tables 2.6–7. The notation is as follows

$$
\begin{aligned}
h &- \text{ side length of element} \\
||e||_E &- \text{ true value of the error in the energy norm} \\
\hat{\epsilon} &- \text{ estimate obtained using Kelly estimator} \\
\theta_{Kelly} &- \text{ effectivity of Kelly estimator } ||e||_E/\hat{\epsilon} \\
\epsilon &- \text{ estimate obtained using new estimator}
\end{aligned}
$$

It is observed that both estimators perform well and converge towards the true error as the partition is refined. However it is seen that whilst the recovery based estimator tends to give an upper bound on the discretization error, the Kelly estimator tends to give a lower bound on the error.

It might be thought that the recovery based estimator is too complicated to be of practical use. In actual fact it is in many ways much simpler than the Kelly estimator. For example, in the case of the Kelly estimator it is not immediately clear how one should define the value of the jump J along an element side which forms part of the boundary of Ω. This difficulty does not arise with the recovery based estimator. In fact the recovery based approach provides the answer to this problem: the value of $|J|$ should be taken to be the same as the jump on the opposite side of the element (since this is the value one would obtain by extrapolating the recovered gradient at the midpoints and the centroids).

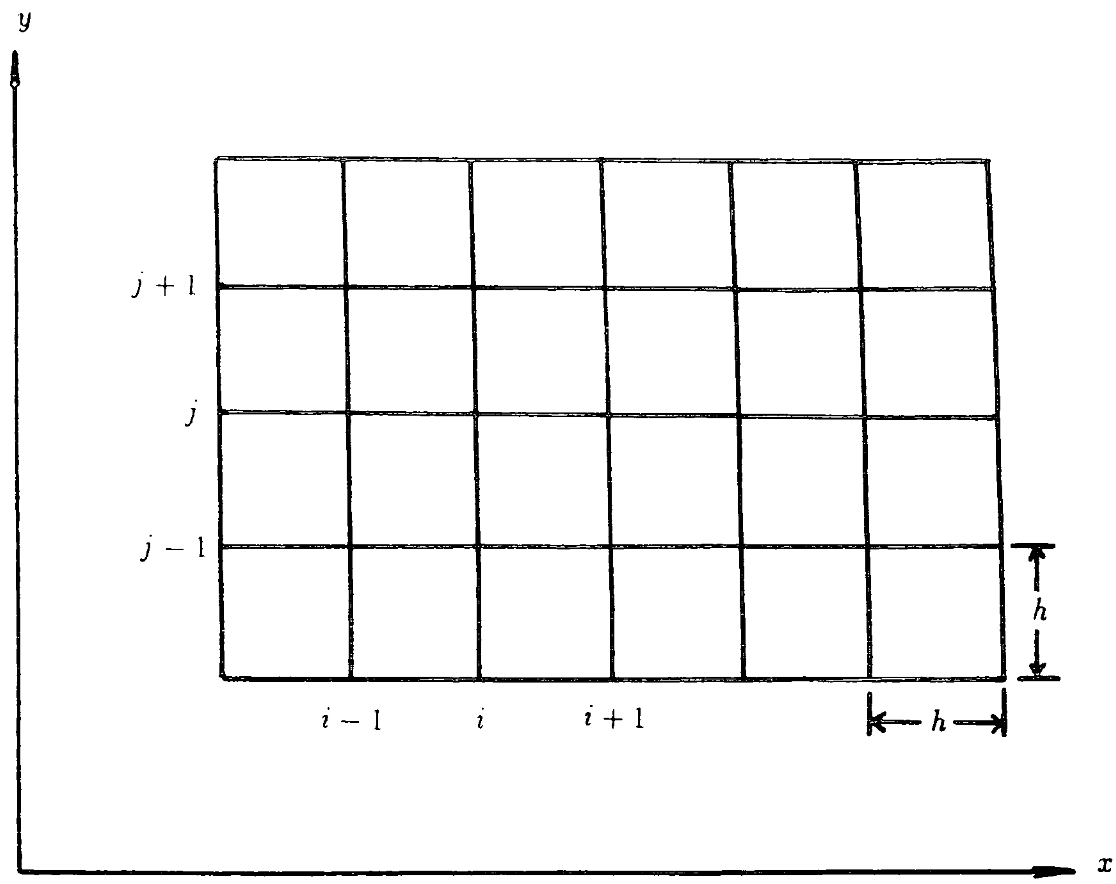


Figure 2.1 Uniform mesh of square elements.

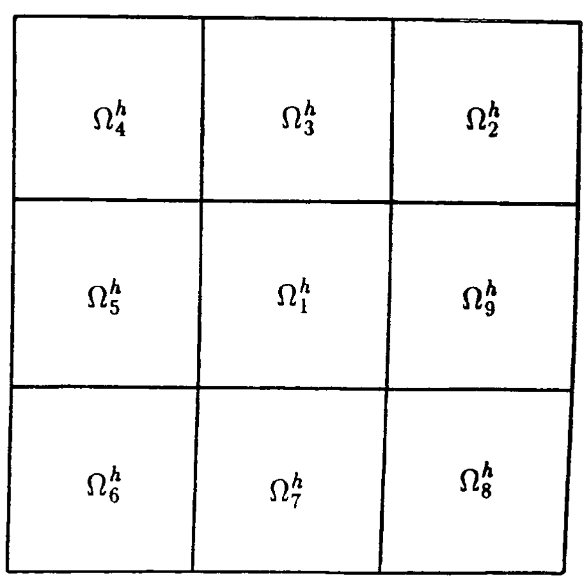


Figure 2.2 $\hat{\Omega}_1^h$ - Patch of elements used for recovery from bilinear approximation.

N	$\|e\|_E$	ϵ	θ
5	$0.3334E+00$	$0.3656E+00$	0.9119
10	$0.6820E-01$	$0.6998E-01$	0.9745
20	$0.1540E-01$	$0.1544E-01$	0.9970
40	$0.3659E-02$	$0.3660E-02$	0.9997
80	$0.8921E-03$	$0.8921E-03$	1.0000
160	$0.2202E-03$	$0.2202E-03$	1.000

Table 2.1 Performance of quadratic estimator.
$A(x) = 1;\ B(x) = 0.$

N	$\|e\|_E$	ϵ	θ
5	$0.3196E+00$	$0.3409E+00$	0.9375
10	$0.6514E-01$	$0.6601E-01$	0.9868
20	$0.1470E-01$	$0.1470E-01$	1.0000
40	$0.3495E-02$	$0.3493E-02$	1.0005
80	$0.8520E-03$	$0.8518E-03$	1.0002
160	$0.2103E-03$	$0.2103E-03$	1.0000

Table 2.2 Performance of quadratic estimator.
$A(x) = 1 - (x - \frac{1}{2})^2;\ B(x) = 0.$

N	$\|e\|_E$	ϵ	θ
5	$0.5421E+00$	$0.5431E+00$	0.9983
10	$0.1102E+00$	$0.1090E+00$	1.0107
20	$0.2491E-01$	$0.2476E-01$	1.0060
40	$0.5923E-02$	$0.5911E-02$	1.0020
80	$0.1444E-02$	$0.1443E-02$	1.0005
160	$0.3565E-03$	$0.3565E-03$	1.0001

Table 2.3 Performance of quadratic estimator.
$A(x) = 1 + 10x(1 - x);\ B(x) = 0.$

N	$\|e\|_E$	ϵ	θ
5	$0.3358E+00$	$0.3685E+00$	0.9113
10	$0.6829E-01$	$0.7000E-01$	0.9756
20	$0.1540E-01$	$0.1544E-01$	0.9972
40	$0.3660E-02$	$0.3660E-02$	0.9998
80	$0.8921E-03$	$0.8921E-03$	1.0000
160	$0.2202E-03$	$0.2202E-03$	1.0000

Table 2.4 Performance of quadratic estimator.
$A(x)=1;\ B(x)=10.$

N	$\|e\|_E$	ϵ	θ
5	$0.3543E+00$	$0.3881E+00$	0.9129
10	$0.6918E-01$	$0.7019E-01$	0.9854
20	$0.1545E-01$	$0.1545E-01$	1.0000
40	$0.3662E-02$	$0.3660E-02$	1.0004
80	$0.8923E-03$	$0.8921E-03$	1.0002
160	$0.2203E-03$	$0.2202E-03$	1.0001

Table 2.5 Performance of quadratic estimator.
$A(x)=1;\ B(x)=100.$

h^{-1}	$\|e\|_E$	$\hat{\varepsilon}$	θ^{-1}_{Kelly}	ϵ	θ^{-1}_{New}
2	$1.27203E+00$	$1.22900E+00$	0.9662	$1.23463E+00$	0.9706
4	$6.39078E-01$	$6.29153E-01$	0.9845	$6.37037E-01$	0.9968
8	$3.19921E-01$	$3.18401E-01$	0.9953	$3.20064E-01$	1.0004
16	$1.60008E-01$	$1.59806E-01$	0.9987	$1.60058E-01$	1.0003
32	$8.00100E-02$	$7.99829E-02$	0.9997	$8.00185E-02$	1.0001

Table 2.6 Comparison of bilinear estimator with Kelly estimator.
True Solution: $u(x,y) = (1+x+y)(1+x^2+y^2)$.

h^{-1}	$\|e\|_E$	$\hat{\varepsilon}$	θ^{-1}_{Kelly}	ϵ	θ^{-1}_{New}
2	$2.92060E-01$	$2.08417E-01$	0.7136	$2.94746E-01$	1.0092
4	$1.39574E-01$	$1.31975E-01$	0.9456	$1.52488E-01$	1.0925
8	$6.89620E-02$	$6.77710E-02$	0.9827	$7.16257E-02$	1.0386
16	$3.43783E-02$	$3.41993E-02$	0.9948	$3.48398E-02$	1.0134
32	$1.71763E-03$	$1.71518E-03$	0.9986	$1.72432E-02$	1.0039

Table 2.7 Comparison of bilinear estimator with Kelly estimator.
True Solution: $u(x,y) = x(1-x)\sin \pi y$.

CHAPTER 3

Error Estimation for Elliptic Systems.

3.1 Introduction

The results of Chapter 2 were shown for a model elliptic boundary value problem satisfying homogeneous Dirichlet conditions on the whole of the boundary. In this chapter we shall generalize the results of Chapter 2. In particular, we shall deal with second order elliptic systems, rather than a single equation. Moreover, we shall consider the more general case of mixed non-homogeneous boundary conditions of the Dirichlet and Neumann type. The ideas are illustrated for the practically important case of the Lamé-Navier equations of linear elasticity in $\Re^2$. A short Appendix discussing the weak formulation and finite element approximation of this system is included where several basic results needed in the main text are collected together.

We shall again discuss the *a posteriori* estimation of the error in the energy norm. For the Lamé-Navier equations this takes the form (see (3.2.15) below)

$$B(\mathbf{e},\mathbf{e}) = \int_\Omega (\sigma - \sigma_h)^t \mathbf{D}^{-1} (\sigma - \sigma_h) d\mathbf{x}$$

where σ and σ_h are the true stress and the finite element approximation to the stress respectively. Error estimators have been proposed for this problem based on the intuitive argument that replacing σ by a good approximation to the stress should give a good approximation to the error. This leaves us with the question of how to find a suitable choice of approximation to the true stress to replace σ in (3.2.15). We follow the approach of Chapter 2 based on using superconvergence phenomena. However, instead of using recovery schemes based on an averaging process we shall consider the use of projection based recovery schemes. Of course, it is also desirable to develop analytical evidence regarding the behaviour of the resulting projection based error estimators. Following the approach of Chapter 2 we shall quantify the effectiveness of the estimators by *asymptotic exactness.* It is worth noting that there are suitable superconvergence results available for the development of averaging based estimators [41], [42], [43], [80] but we shall deal with projection techniques primarily in this chapter.

The chapter is organized as follows. Firstly, we develop a generalization of Theorem 2.3.1 of Chapter 2 to aid in the analysis of the estimators. Secondly, we consider the class of *projection* based error estimators including as a special case the estimator proposed by Zienkiewicz and Zhu [82],[87],[88]. Finally, a comparison is made as to the relative merits of *averaging* versus *projection* based error estimation. Numerical results are included to support the claims made.

3.2 Estimation Framework.

We denote the solution domain by Ω and denote its boundary by Γ. The problem which we shall consider is the Lamé-Navier equations in two dimensions

$$-(\lambda+\mu)\nabla(\nabla\cdot\mathbf{u})-\mu\Delta\mathbf{u}=\mathbf{f} \quad in \quad \Omega \tag{3.2.1}$$

where

- $\mathbf{u}$ is the displacement vector
- $\mathbf{f}$ is the body force
- λ and μ are the Lamé coefficients given by

$$\lambda=\frac{E\nu}{(1+\nu)(1-2\nu)} \tag{3.2.2}$$

and

$$\mu = \frac{E}{2(1+\nu)}, \qquad (3.2.3)$$

with E and ν Young's modulus and Poisson's ratio respectively.

The boundary conditions are prescribed displacements on part of Γ and prescribed tractions on the remainder of Γ,

$$\mathbf{u} = \hat{\mathbf{u}} \quad on \quad \Gamma_d \qquad (3.2.4)$$

$$\mathbf{H}\sigma = \hat{\mathbf{t}} \quad on \quad \Gamma_n. \qquad (3.2.5)$$

If we define the differential operator **S** to be

$$\mathbf{S} = \begin{pmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{pmatrix} \qquad (3.2.6)$$

and **D** to be the elasticity matrix, in the case of plane strain

$$\mathbf{D} = \frac{E}{(1+\nu)(1-2\nu)} \begin{pmatrix} 1-\nu & \nu & 0 \\ \nu & 1-\nu & 0 \\ 0 & 0 & \frac{1}{2}(1-2\nu) \end{pmatrix} \qquad (3.2.7)$$

then the stress is given by

$$\sigma = \mathbf{DSu}. \qquad (3.2.8)$$

The weak form of this problem is (see Appendix)

$$\mathbf{u} \in A_{\hat{\mathbf{u}}} : B(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) + \langle \hat{\mathbf{t}}, \mathbf{v} \rangle_{\Gamma_n} \quad \forall \mathbf{v} \in A_0 \qquad (3.2.9)$$

where

$$A_{\hat{\mathbf{u}}} = \{\mathbf{u} \in [H^1(\Omega)]^2 : \mathbf{u} = \hat{\mathbf{u}} \quad on \quad \Gamma_d\} \qquad (3.2.10)$$

$$B(\mathbf{u},\mathbf{v}) = \int_\Omega (\mathbf{S}\mathbf{u})^t \mathbf{D}\mathbf{S}\mathbf{v}d\mathbf{x} \tag{3.2.11}$$

$$(\mathbf{f},\mathbf{v}) = \int_\Omega \mathbf{f}^t \mathbf{v}d\mathbf{x} \tag{3.2.12}$$

and

$$\langle \hat{\mathbf{t}}, \mathbf{v}\rangle_{\Gamma_n} = \int_{\Gamma_n} \mathbf{v}^t \hat{\mathbf{t}} ds. \tag{3.2.13}$$

The finite element approximation $\mathbf{u}_h$ is characterized as the solution of

$$\mathbf{u}_h \in A^h_{\hat{\mathbf{u}}} : B(\mathbf{u}_h, \mathbf{v}_h) = (\mathbf{f}, \mathbf{v}_h) + \langle \hat{\mathbf{t}}, \mathbf{v}_h\rangle_{\Gamma_n} \quad \forall \mathbf{v}_h \in A^h_0. \tag{3.2.14}$$

Letting $\mathbf{e} = \mathbf{u} - \mathbf{u}_h$ denote the error in the Galerkin approximation, we shall be interested in obtaining computable bounds on the energy norm of the error B($\mathbf{e}$,$\mathbf{e}$), which may be rewritten using (3.2.8) as

$$B(\mathbf{e},\mathbf{e}) = \int_\Omega (\sigma - \sigma_h)^t \mathbf{D}^{-1} (\sigma - \sigma_h) d\mathbf{x} \tag{3.2.15}$$

where

$$\sigma_h = \mathbf{D}\mathbf{S}\mathbf{u}_h \tag{3.2.16}$$

is the approximation to the stress obtained from the Galerkin approximation.

Following the approach of Chapter 2, we characterize the error in the approximation as the solution of an elasticity problem as follows. Substituting $\mathbf{u} = \mathbf{e} + \mathbf{u}_h$ into (3.2.9) and assuming that $\mathbf{u}_h$ satisfies the displacement conditions exactly we obtain

$$\mathbf{e} \in A_0 : B(\mathbf{e}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) - B(\mathbf{u}_h, \mathbf{v}) + \langle \hat{\mathbf{t}}, \mathbf{v}\rangle_{\Gamma_n} \quad \forall \mathbf{v} \in A_0. \tag{3.2.17}$$

We define a perturbed bilinear form $B_\lambda(.,.)$, $\lambda > 0$, as

$$B_\lambda(\mathbf{u},\mathbf{v}) = B(\mathbf{u},\mathbf{v}) + \lambda(\mathbf{u},\mathbf{v}) \tag{3.2.18}$$

and define $\mathbf{y}$ to be the solution of the problem

$$\mathbf{y} \in A_0 : B_\lambda(\mathbf{y},\mathbf{v}) = (\mathbf{f},\mathbf{v}) - B(\mathbf{u}_h,\mathbf{v}) + \langle \hat{\mathbf{t}},\mathbf{v}\rangle_{\Gamma_n} \quad \forall \mathbf{v} \in A_0. \tag{3.2.19}$$

Equally well we may reformulate this as a primal variational principle

$$\mathbf{y} \in A_0 : \mathcal{J}_\lambda(\mathbf{y}) \leq \mathcal{J}_\lambda(\mathbf{v}) \quad \forall \mathbf{v} \in A_0 \tag{3.2.20}$$

where

$$\mathcal{J}(\mathbf{v}) = \frac{1}{2}B_\lambda(\mathbf{v},\mathbf{v}) - (\mathbf{f},\mathbf{v}) + B(\mathbf{u}_h,\mathbf{v}) - \langle \hat{\mathbf{t}},\mathbf{v}\rangle_{\Gamma_n}. \tag{3.2.21}$$

The following result gives the complementary variational principle associated with this primal problem

THEOREM 3.2.1

Let $\mathcal{K}_\lambda[\mathbf{p}]$ be the quadratic functional

$$\mathcal{K}_\lambda[\mathbf{p}] = \int_\Omega \left\{(\mathbf{p} - \mathbf{DSu}_h)^t\mathbf{D}^{-1}(\mathbf{p} - \mathbf{DSu}_h)\right\} d\mathbf{x} + \frac{1}{\lambda}\int_\Omega (\mathbf{f} + \mathbf{S}^t\mathbf{p})^2 d\mathbf{x} \tag{3.2.22}$$

defined on the set

$$\mathcal{S}_{\mathbf{f}} = \{\mathbf{p} : \mathbf{S}^t\mathbf{p} + \mathbf{f} \in [L^2(\Omega)]^2 \textit{ and } \mathbf{H}\mathbf{p} = \hat{\mathbf{t}} \textit{ on } \Gamma_n\} \tag{3.2.23}$$

then the following bound holds

$$B_\lambda(\mathbf{y},\mathbf{y}) = \mathcal{K}_\lambda[\mathbf{DS}(\mathbf{y}+\mathbf{u}_h)] \leq \mathcal{K}_\lambda[\mathbf{p}] \quad \forall \mathbf{p} \in \mathcal{S}_{\mathbf{f}} \tag{3.2.24}$$

and $\mathbf{DS}(\mathbf{y}+\mathbf{u}_h) \in \mathcal{S}_{\mathbf{f}}$.

Proof. Firstly we show that $\mathbf{DS}(\mathbf{y}+\mathbf{u}_h) \in \mathcal{S}_{\mathbf{f}}$. Now (3.2.19) gives

$$B_\lambda(\mathbf{y},\mathbf{v}) + B(\mathbf{u}_h,\mathbf{v}) = (\mathbf{f},\mathbf{v}) + \langle \hat{\mathbf{t}},\mathbf{v}\rangle_{\Gamma_n} \quad \forall \mathbf{v} \in A_0. \tag{3.2.25}$$

and (3.2.18) implies that

$$B_\lambda(\mathbf{y}+\mathbf{u}_h,\mathbf{v}) = (\mathbf{f}-\lambda\mathbf{y},\mathbf{v}) + \langle \hat{\mathbf{t}},\mathbf{v}\rangle_{\Gamma_n} \quad \forall \mathbf{v} \in A_0. \tag{3.2.26}$$

Comparing (3.2.26) with the weak form of the Lamé-Navier equations derived in the Appendix, we deduce the formal relationships

$$-\mathbf{S}^t\mathbf{DS}(\mathbf{y}+\mathbf{u}_h) = \mathbf{f} - \lambda\mathbf{y} \text{ in } \Omega \tag{3.2.27}$$

$$\mathbf{y}+\mathbf{u}_h = \hat{\mathbf{u}} \text{ on } \Gamma_d \tag{3.2.28}$$

and

$$\mathbf{HDS}(\mathbf{y} + \mathbf{u}_h) = \hat{\mathbf{t}} \text{ on } \Gamma_n. \tag{3.2.29}$$

Hence, we conclude that $\mathbf{DS}(\mathbf{y} + \mathbf{u}_h) \in \mathcal{S}_{\mathbf{f}}$. Moreover, we see from (3.2.27) that

$$\mathcal{K}_\lambda[\mathbf{DS}(\mathbf{y} + \mathbf{u}_h)] = \int_\Omega (\mathbf{S}\mathbf{y})^t \mathbf{D}(\mathbf{S}\mathbf{y}) dx + \lambda \int_\Omega \mathbf{y}^t \mathbf{y} dx = B_\lambda(\mathbf{y}, \mathbf{y}). \tag{3.2.30}$$

It therefore remains only to show that $\mathcal{K}_\lambda$ is convex and stationary at $\mathbf{p} = \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)$. To see this, let $\epsilon > 0$ be taken arbitrarily, and let $\mathbf{p}$ be any element of the (non-empty) convex set $\mathcal{S}_{\mathbf{f}}$. Then,

$$\begin{aligned}
&\mathcal{K}_\lambda[\mathbf{DS}(\mathbf{y} + \mathbf{u}_h) + \epsilon\{\mathbf{p} - \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)\}] - \mathcal{K}_\lambda[\mathbf{DS}(\mathbf{y} + \mathbf{u}_h)] \\
&= 2\epsilon\Big\{ \int_\Omega (\mathbf{DSy})^t \mathbf{D}^{-1} \{\mathbf{p} - \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)\} d\mathbf{x} + \\
&\quad \frac{1}{\lambda} \int_\Omega \{\mathbf{f} + \mathbf{S}^t \mathbf{DS}(\mathbf{y} + u_h)\}^t \mathbf{S}^t \{\mathbf{p} - \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)\} d\mathbf{x} \Big\} + O(\epsilon^2) \\
&= 2\epsilon\Big\{ \int_\Omega (\mathbf{Sy})^t \{\mathbf{p} - \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)\} d\mathbf{x} + \\
&\quad \frac{1}{\lambda} \int_\Omega \lambda \mathbf{y}^t \mathbf{S}^t \{\mathbf{p} - \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)\} d\mathbf{x} \Big\} + O(\epsilon^2) \text{ using (3.2.27).}
\end{aligned} \tag{3.2.31}$$

Thus

$$\begin{aligned}
&\frac{1}{2} \lim_{\epsilon \to 0} \frac{1}{\epsilon} \{\mathcal{K}_\lambda[\mathbf{DS}(\mathbf{y} + \mathbf{u}_h) + \epsilon\{\mathbf{p} - \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)\}] - \mathcal{K}_\lambda[\mathbf{DS}(\mathbf{y} + \mathbf{u}_h)]\} \\
&= \Big\{ \int_\Omega (\mathbf{Sy})^t \{\mathbf{p} - \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)\} d\mathbf{x} + \int_\Omega \lambda \mathbf{y}^t \mathbf{S}^t \{\mathbf{p} - \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)\} d\mathbf{x} \Big\} \\
&= \int_\Gamma \mathbf{y}^t \mathbf{H} \{\mathbf{p} - \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)\} ds
\end{aligned} \tag{3.2.32}$$

where we made use of (A.16) and (A.17), from the Appendix. From (3.2.28) we have that

$$\mathbf{y} = \mathbf{0} \text{ on } \Gamma_d$$

and from (3.2.18) we have

$$\mathbf{H}\{\mathbf{p} - \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)\} = \mathbf{0} \text{ on } \Gamma_n$$

and consequently, the line integral in (3.2.32) vanishes giving

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon}\{\mathcal{K}_\lambda[\mathbf{DS}(\mathbf{y} + \mathbf{u}_h) + \epsilon\{\mathbf{p} - \mathbf{DS}(\mathbf{y} + \mathbf{u}_h)\}] - \mathcal{K}_\lambda[\mathbf{DS}(\mathbf{y} + \mathbf{u}_h)]\} = 0$$

so that $\mathcal{K}_\lambda$ is stationary at $\mathbf{DS}(\mathbf{y} + \mathbf{u}_h)$. To see that $\mathcal{K}_\lambda$ is convex, let $\mu \in (0,1)$ be taken arbitrarily and let $\mathbf{p}$, $\mathbf{q}$ be any elements of $\mathcal{S}_\mathbf{f}$. Firstly, we notice that

$$\begin{aligned} &\mathcal{K}_\lambda[\mu\mathbf{p} + (1-\mu)\mathbf{q}] \\ &= \mu^2\mathcal{K}_\lambda[\mathbf{p}] + (1-\mu)^2\mathcal{K}_\lambda[\mathbf{q}] + 2\mu(1-\mu)\int_\Omega (\mathbf{p} - \mathbf{DSu}_h)^t\mathbf{D}^{-1}(\mathbf{q} - \mathbf{DSu}_h)dx \\ &\qquad + \frac{2}{\lambda}\mu(1-\mu)\int_\Omega (\mathbf{f} + \mathbf{S}^t\mathbf{p})^t(\mathbf{f} + \mathbf{S}^t\mathbf{q})dx. \end{aligned} \tag{3.2.33}$$

Now, since $\mathbf{D}^{-1}$ is positive definite and symmetric it has a Cholesky factorization L^tL, and therefore

$$\begin{aligned} &\int_\Omega (\mathbf{p} - \mathbf{DSu}_h)^t\mathbf{D}^{-1}(\mathbf{q} - \mathbf{DSu}_h)dx \\ &= \int_\Omega \{L(\mathbf{p} - \mathbf{DSu}_h)\}^t\{L(\mathbf{q} - \mathbf{DSu}_h)\}dx \\ &\le \frac{1}{2}\int_\Omega \{L(\mathbf{p} - \mathbf{DSu}_h)\}^2 dx + \frac{1}{2}\int_\Omega \{L(\mathbf{q} - \mathbf{DSu}_h)\}^2 dx, \end{aligned} \tag{3.2.34}$$

where we have made use of the inequality

$$\sum_i a_i b_i \le \frac{1}{2}\sum_i a_i^2 + \frac{1}{2}\sum_i b_i^2.$$

The same inequality may used to show that

$$\int_\Omega (\mathbf{f} + \mathbf{S}^t\mathbf{p})^t(\mathbf{f} + \mathbf{S}^t\mathbf{q})d\mathbf{x} \le \frac{1}{2}\int_\Omega (\mathbf{f} + \mathbf{S}^t\mathbf{p})^2 d\mathbf{x} + \frac{1}{2}\int_\Omega (\mathbf{f} + \mathbf{S}^t\mathbf{q})^2 d\mathbf{x}. \qquad (3.2.35)$$

Substituting (3.2.34) and (3.2.35) into (3.2.33) yields the desired result

$$\mathcal{K}_\lambda[\mu\mathbf{p} + (1-\mu)\mathbf{q}] \le \mu\mathcal{K}_\lambda[\mathbf{p}] + (1-\mu)\mathcal{K}_\lambda[\mathbf{q}].$$

This completes the proof of the Theorem. ∎

Having obtained this preliminary result we have

THEOREM 3.2.2 *Let $\mathcal{K}_\lambda[\mathbf{p}]$ be the quadratic functional given by (3.2.22) and $\mathcal{S}_\mathbf{f}$ be the set given by (3.2.23). If we take $\lambda = Mh^{-\alpha}$ where $\alpha \in (0,2)$ and $K > 0$ are real constants then the following bound holds for any $\mathbf{p} \in \mathcal{S}_\mathbf{f}$*

$$\| \mathbf{e} \|_E \le \{1 + O(h^{1-\alpha/2})\}\mathcal{K}_\lambda[\mathbf{p}] \text{ as } h \to 0 \qquad (3.2.36)$$

where the constant in the $O(h^{1-\alpha/2})$ is independent of h and $\mathbf{u}$.

Proof. The proof follows from Theorem 3.2.1 using arguments similiar to those used to obtain Theorem 2.3.1 from Theorem 2.3.2 , and is therefore omitted. ∎

REMARK We have given a full proof of Theorem 3.2.1 because the method of proof is virtually the same for any elliptic system, and not just specific to the Lamé-Navier equations.

3.3 Projection based error estimation.

Although the Galerkin approximation $\mathbf{u}_h$ is C^0-continuous, the approximation σ_h to the stress will be discontinuous across the interelement boundaries. A smooth approximation to the true stress σ may be obtained by projecting the direct approximation σ_h on to a C^0-continuous space. This space is usually constructed from the same basis functions used to construct the finite element approximation $\mathbf{u}_h$, and the inner product used for the projection is usually the standard $L^2(\Omega)$ inner product. It has been noted in practice (Hinton and Campbell [40]) that the smoothed approximation to the stress is generally more accurate than σ_h. This technique forms the basis of an estimator proposed and numerically investigated by Zienkiewicz and Zhu [88],[87]. In addition to the projection used by Zienkiewicz and Zhu we shall consider also the projection onto basis functions of different degree to that of the basis functions used to calculate the approximation to the displacement. Moreover, we shall also consider projections in *weighted* $L^2(\Omega)$ inner products.

It must be stated at the outset that the results we shall derive regarding the reliability of the resulting estimators will be incomplete. However, in view of the current interest in projection based error estimation, especially amongst the engineering fraternity, it is desirable to obtain whatever results are accessible. One of the chief difficulties in the analysis is the paucity of rigorous results on the superconvergence of the projection recovery technique (see Rachowicz and Oden [63] for some recent results in this area and also [23], [24]). In fact our analysis will not improve this situation since the approach sidesteps the main questions relating to superconvergence proper and concentrates on the problem of analyzing the error estimator. We shall return to the subject of superconvergence of projection techniques when discussing the relative merits of projection versus averaging based error estimation.

Let $\{M_i\}$ denote a set of basis functions consisting of piecewise polynomials of degree $\leq q$. The finite element approximation to the displacement is supposed to have been based on basis functions of degree $\leq p$, where p need not necessarily be equal to q. The smoothed approximation to the stress has the form

$$\sigma^* = \begin{pmatrix} \sum_i \lambda_i M_i \\ \sum_i \mu_i M_i \\ \sum_i \nu_i M_i \end{pmatrix} \tag{3.3.1}$$

and we shall assume that the traction conditions can be satisfied exactly. This leads to the conditions on $\{\lambda_i, \mu_i, \nu_i\}$

$$\mathbf{H}\sigma^*(\mathbf{x}_k) = \hat{\mathbf{t}}(\mathbf{x}_k) \quad \forall \mathbf{x}_k \in \Gamma_n \tag{3.3.2}$$

or

$$\begin{pmatrix} \lambda_k n_x + \nu_k n_y \\ \mu_k n_y + \nu_k n_x \end{pmatrix} = \begin{pmatrix} \hat{t}_x(\mathbf{x}_k) \\ \hat{t}_y(\mathbf{x}_k) \end{pmatrix}. \tag{3.3.3}$$

We shall denote the set of functions of the form (3.3.1) which satisfy the conditions (3.3.3) by $\mathcal{W}_{\hat{\mathbf{t}}}$. The projection used by Zienkiewicz and Zhu [88] is the unique $\sigma^* \in \mathcal{W}_{\hat{\mathbf{t}}}$ which satisfies the projection condition

$$\| \sigma^* - \sigma_h \|_{L^2(\Omega)} \leq \| \tilde{\sigma} - \sigma_h \|_{L^2(\Omega)} \quad \forall \tilde{\sigma} \in \mathcal{W}_{\hat{\mathbf{t}}}. \tag{3.3.4}$$

Letting

$$\| \mathbf{v} \|^2_{\mathbf{D}^{-1}} = \int_\Omega \mathbf{v}^t \mathbf{D}^{-1} \mathbf{v} d\mathbf{x} \tag{3.3.5}$$

denote the *weighted* norm, we shall also consider the projections $\sigma^{\mathbf{D}}$ given by

$$\| \sigma^{\mathbf{D}} - \sigma_h \|_{\mathbf{D}^{-1}} \leq \| \tilde{\sigma} - \sigma_h \|_{\mathbf{D}^{-1}} \quad \forall \tilde{\sigma} \in \mathcal{W}_{\mathfrak{t}}. \tag{3.3.6}$$

Since the basis functions M_i are continuous and satisfy the boundary conditions on the stress automatically, all of the projections defined by (3.3.4) and (3.3.6) will satisfy the conditions of Theorem 3.2.2 .

We now turn to the problem of showing that the estimators obtained via these smoothed stresses are reasonably tight bounds. We have that

THEOREM 3.3.1 *Let $\sigma^{\mathbf{D}}$ denote the stress obtained by the projection (3.3.6), then*

$$\| \sigma^{\mathbf{D}} - \sigma_h \|_{\mathbf{D}^{-1}} \leq \{1 + C(u) h^{q-p+1}\} \| \mathbf{e} \|_E \tag{3.3.7}$$

provided that the true stress σ is sufficiently smooth.

Proof. Applying the Triangle Inequality gives

$$\| \tilde{\sigma} - \sigma_h \|_{\mathbf{D}^{-1}} \leq \| \tilde{\sigma} - \sigma \|_{\mathbf{D}^{-1}} + \| \mathbf{e} \|_E . \tag{3.3.8}$$

Now by the standard approximation properties we have that there exists a $\tilde{\sigma} \in \mathcal{W}_{\mathfrak{t}}$ such that

$$\| \tilde{\sigma} - \sigma \|_{\mathbf{D}^{-1}} \leq C(\sigma) h^{q+1} \tag{3.3.9}$$

Further, if the true solution cannot be written as a linear combination of the shape functions (if it can then there is no error at all) then we have an inverse estimate of the form

$$\| \mathbf{e} \|_E \geq C(u)h^p. \tag{3.3.10}$$

Together (3.3.6), (3.3.8), (3.3.9) and (3.3.10) give

$$\| \sigma^{\mathbf{D}} - \sigma_h \|_{\mathbf{D}^{-1}} \leq \{1 + C(u)h^{q-p+1}\} \| \mathbf{e} \|_E$$

as required. ∎

We may also show an analogous result for the projections (3.3.4).

THEOREM 3.3.2 *Let σ^* denote the stress obtained by the projection (3.3.4), then we have that*

$$\| \sigma^* - \sigma_h \|_{\mathbf{D}^{-1}} \leq \sqrt{\frac{2}{1-2\nu}} \{1 + C(u)h^{q-p+1}\} \| \mathbf{e} \|_E \tag{3.3.11}$$

provided that the true stress σ is sufficiently smooth.

Proof. The following equivalence between the norms denoted by $\| \cdot \|_{\mathbf{D}^{-1}}$ and $\| \cdot \|_{L^2(\Omega)}$ is easily shown

$$\frac{(1+\nu)(1-2\nu)}{E} \| \mathbf{v} \|^2_{L^2(\Omega)} \leq \| \mathbf{v} \|^2_{\mathbf{D}^{-1}} \leq \frac{2(1+\nu)}{E} \| \mathbf{v} \|^2_{L^2(\Omega)} . \tag{3.3.12}$$

Using the right hand inequality in (3.3.12) we obtain

$$\| \sigma^* - \sigma_h \|_{\mathbf{D}^{-1}} \leq \sqrt{\frac{2(1+\nu)}{E}} \| \sigma^* - \sigma_h \|_{L^2(\Omega)} . \tag{3.3.13}$$

By the projection condition (3.3.4) we have for any $\tilde{\sigma} \in \mathcal{W}_{\mathfrak{k}}$ that

$$\| \sigma^* - \sigma_h \|_{L^2(\Omega)} \leq \| \tilde{\sigma} - \sigma_h \|_{L^2(\Omega)} \tag{3.3.14}$$

and by the Triangle Inequality

$$\| \tilde{\sigma} - \sigma_h \|_{L^2(\Omega)} \leq \| \tilde{\sigma} - \sigma \|_{L^2(\Omega)} + \| \sigma - \sigma_h \|_{L^2(\Omega)},$$

so that

$$\| \tilde{\sigma} - \sigma_h \|_{L^2(\Omega)} \leq \| \tilde{\sigma} - \sigma \|_{L^2(\Omega)} + \sqrt{\frac{E}{(1+\nu)(1-2\nu)}} \| \mathbf{e} \|_E \tag{3.3.15}$$

where we used the left hand inequality in (3.3.12). By the approximation properties of polynomials we have for sufficiently smooth stresses that there exists $\tilde{\sigma}$ such that

$$\| \tilde{\sigma} - \sigma \|_{L^2(\Omega)} \leq C(u) h^{q+1}$$

and so by (3.3.10) we find

$$\| \tilde{\sigma} - \sigma \|_{L^2(\Omega)} \leq C(u) h^{q-p+1} \| \mathbf{e} \|_E . \tag{3.3.16}$$

Now the result follows immediately from (3.3.13), (3.3.14), (3.3.15) and (3.3.16). ∎

Letting

$$\epsilon^2 = \int_\Omega (\hat{\sigma} - \sigma_h)^t \mathbf{D}^{-1} (\hat{\sigma} - \sigma_h) d\mathbf{x} \tag{3.3.17}$$

and

$$\Lambda^2 = \int_\Omega (\mathbf{f} + \mathbf{S}^t \sigma)^2 d\mathbf{x} \tag{3.3.18}$$

where $\hat{\sigma}$ is the smoothed stress obtained via one of the above projections, our results may be summarized as

$$\{1 + C_1(u) h^{q-p+1}\}\epsilon^2 \leq \| \mathbf{e} \|_E^2 \leq \{1 + C_2(u) h^\delta\}(\epsilon^2 + \frac{h^\alpha}{K}\Lambda^2) \tag{3.3.19}$$

for the projections characterised by (3.3.6) and as

$$\frac{(1-2\nu)}{2}\{1 + C_1(u) h^{q-p+1}\}\epsilon^2 \leq \| \mathbf{e} \|_E^2 \leq \{1 + C_2(u) h^\delta\}(\epsilon^2 + \frac{h^\alpha}{K}\Lambda^2) \tag{3.3.20}$$

for the projections characterised by (3.3.4), provided the stress is sufficiently smooth. For problems other than the plane strain problem the elasticity matrix $\mathbf{D}$ is different. Nevertheless, the results shown above continue to hold, the only modification being that the term $(1-2\nu)/2$ in (3.3.20) is replaced by the ratio of the moduli of the minimum to the maximum eigenvalues of the elasticity matrix (this is related to the *condition number* of the elasticity matrix). In particular for the plane stress problem the elasticity matrix is

$$\mathbf{D} = \frac{E}{1-\nu^2}\begin{pmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{1}{2}(1-\nu) \end{pmatrix} \tag{3.3.21}$$

and the quantity corresponding to the term $\frac{1}{2}(1-2\nu)$ is given by $(1-\nu)/[2(1+\nu)]$, whereas for axisymmetric problems the matrix is

$$\mathbf{D} = \frac{E}{(1+\nu)(1-2\nu)}\begin{pmatrix} 1-\nu & \nu & \nu & 0 \\ \nu & 1-\nu & \nu & 0 \\ \nu & \nu & 1-\nu & 0 \\ 0 & 0 & 0 & \frac{1}{2}(1-2\nu) \end{pmatrix} \tag{3.3.22}$$

and the multiplier is given by $(1-2\nu)/[2(1+\nu)]$.

The term Λ^2 is the L^2 norm of the residual for the smoothed stress and so is expected to be small compared to the term ϵ^2. This is indeed found to be the case for smooth solutions. The results show that we should choose $q > p-1$, i.e. $q \geq p$ (since p, q are integers). Moreover, by choosing $q > p$ we will not achieve a substantially more effective error estimator in return for the extra cost expended on projecting onto higher degree basis functions. Thus we conclude that the method used by Zienkiewicz and Zhu [88], which corresponds to choosing $q = p$, is the most economical from the class of projection based estimators which they considered. However, the results tend to suggest that the estimator obtained using an unweighted projection may not be *asymptotically exact* since the constants in (3.3.20) are not both equal to one. Correspondingly, since the constants in (3.3.19) are both equal to one, it may be worthwhile to use this *weighted* projection rather than the standard $L^2(\Omega)$ projection, especially since there is little difference in the overall costs of producing each projection.

REMARK Earlier we stated that the results were incomplete. The problem is that the above results do not show that the estimators are *asymptotically exact.* In order to show *asymptotic exactness* we would need to show that the term Λ^2 was sufficiently small. For averaging based estimators in Chapter 2, we were able to do this but as yet no proof is known for the case of projection based estimators.

3.4 Numerical Examples.

We present three numerical examples taken from two dimensional linear elasticity in order to test the reliability of projection based error estimators when used as the basis of an adaptive process. The *effectivity index* θ is defined to be

$$\theta = \frac{\epsilon}{\| \mathbf{e} \|_E} \tag{3.4.1}$$

where ϵ and $\| \mathbf{e} \|_E$ are the estimated and true errors respectively. The stopping criterion for the adaptive process was that the relative error η given by

$$\eta = \frac{\| \mathbf{e} \|_E}{\| \mathbf{u} \|_E} \approx \frac{\epsilon}{\sqrt{\| \mathbf{u}_h \|_E^2 + \epsilon^2}} \tag{3.4.2}$$

should not exceed five percent.

For the examples given, there is no analytic expression available for the true solution and so the true error is estimated using numerical results obtained from an *adaptive h-p algorithm* proposed by Zienkiewicz, Zhu and Gong [90]. The stopping criterion for the *adaptive h-p algorithm* the criterion was one percent, meaning that the estimates of the true error are sufficiently close to give a valid estimate of the effectivity index.

In all of the examples Poisson's ratio and Young's modulus were taken to be 0.3 and 1.0 respectively.

In the first example, a short cantilever beam under plane strain conditions was approximated using bilinear basis functions on quadrilateral elements. Six adaptive refinements were performed giving the meshes shown in Fig. 3.1. The performance of the estimators is shown in Table 3.1. It can be seen that the effectivity index improves as the mesh is refined.

A machine part under plane stress conditions was analyzed using triangular elements in conjunction with linear basis functions (see Fig. 3.2). In Table 3.2, we give the effectivity indices for both projection and simple averaging based error estimation. It is seen that the averaging based estimator performs better than the projection based estimator.

Finally, an axisymmetric problem shown in Fig. 3.3 was considered. Triangular elements and quadratic basis functions were used to find an approximate solution. The error was estimated using both types of error estimation. Table 3.3 shows the results obtained, and in particular that the averaging estimator is superior to the projection estimator.

REMARK Further numerical results comparing averaging and projection based estimators are presented in Chapter 5.

3.5 Some comments on averaging versus projection based estimation.

The results tend to suggest that averaging is a better estimation technique than projection. In this section we briefly remark on the relative merits of averaging versus projection based error estimation.

The cost of using projection is vast when compared with the cost of using averaging. The main reason for this is the need to solve the system of linear equations which arises from the projection conditions. In practice, this cost may be reduced by not requiring the smoothed approximation to satisfy the traction conditions exactly. This causes the matrix to decouple into three separate systems corresponding to each component of the stress. Moreover, since the matrix involved in each case is the same, and also symmetric and positive definite, further savings can be made by using a banded Cholesky decomposition along with forward and back substitution to solve the systems. In spite of these improvements, we shall always be faced with the solution of a large system of equations. With averaging, there is no such need for the solution of a global system once the finite element approximation has been computed.

One of the main advantages of projection is the ease with which it can be implemented. The reason for this is that it entails the use of routines which are similar to those used to implement the finite element method itself. To implement the averaging based method efficiently gives rise to the need for efficient data structures. For example, it is necessary to be able to gather information regarding which elements surround any given element so that an appropriate averaging of the gradients can be carried out. Data structures which allow this sort of information to be rapidly accessed and efficiently stored are already available, having been developed for adaptive finite element analysis and multigrid solution procedures in the first instance. Further details can be found in [31], [32],[65],[66],[67],[68].

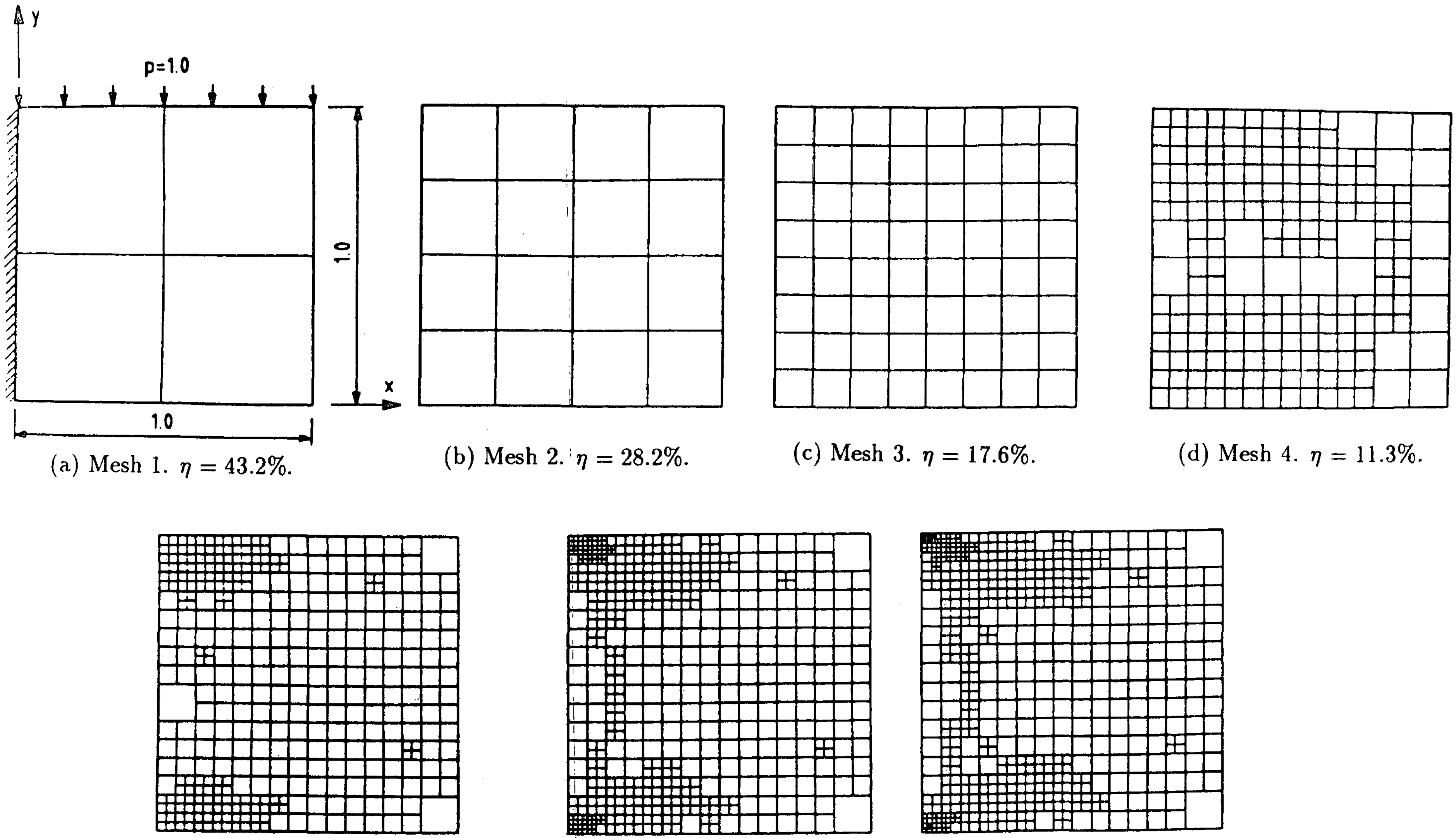
It has been observed that superconvergence results lie at the heart of *a posteriori* error estimation. The superconvergence results relating to averaging are rather better understood than in the case of projection. This is one reason why in Chapter 2, we were able to analyse the second term corresponding to the residual and show that it is insignificant in the case of averaging (provided the true solution is sufficiently smooth). For projection, as we have already mentioned, the results are less well established. This is one reason that we were unable to analyse the term Λ^2 of (3.3.18). Another deficiency of projection as a recovery technique is that it is only superconvergent on the interior of the domain (Rachowicz and Oden [63]). This observation has serious implications regarding how well the error is estimated on elements which have a side lying on the boundary of the domain. Of course, as the partition is refined the measure of the set of elements which have a side in common with the boundary tends to zero. Certainly in the case of smooth problems this means that the global error is estimated increasingly well as the mesh is refined. However, if the elemental estimators are to be used to refine the mesh adaptively then it is imperative that the error is found as accurately as possible. This problem is even more acute for algorithms which seek to generate the final mesh which achieves the prescribed accuracy in at most two adaptive sweeps (see Zhu and Zienkiewicz [82]).

The averaging technique seems to be more favourable than the projection technique. However, the widespread use of projection amongst the engineering community tends to suggest that projection based estimation will become increasingly popular.

Mesh	N_{dof}	$\|u_h\|_E^2$	ϵ_{proj}	$\|e\|_E$	θ	η ($\times 100$)
1	12	1.54842	0.266650	0.596051	0.45	43.3
2	40	1.75234	0.221895	0.389047	0.57	28.2
3	144	1.84507	0.150575	0.242130	0.62	17.6
4	378	1.87945	0.102059	0.155716	0.66	11.3
5	718	1.89289	0.074022	0.103958	0.71	7.53
6	964	1.89763	0.061626	0.077894	0.79	5.65
7	1094	1.89919	0.057077	0.067137	0.85	4.87

Table 3.1 Short Cantilever under plane strain conditions.

Exact solution $\|u\|_E^2 = 1.903697$.



(a) Mesh 1. $\eta = 43.2\%$.

(b) Mesh 2. $\eta = 28.2\%$.

(c) Mesh 3. $\eta = 17.6\%$.

(d) Mesh 4. $\eta = 11.3\%$.

(e) Mesh 5. $\eta = 7.5\%$.

(f) Mesh 6. $\eta = 5.7\%$.

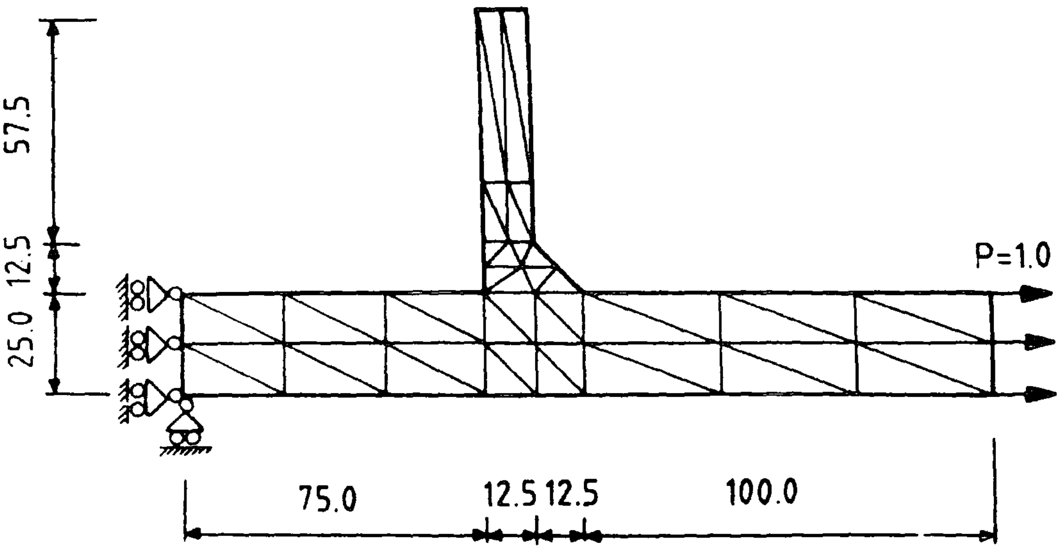
(g) Mesh 7. $\eta = 4.9\%$.

Figure 3.1 Adaptively designed meshes for short cantilever beam.

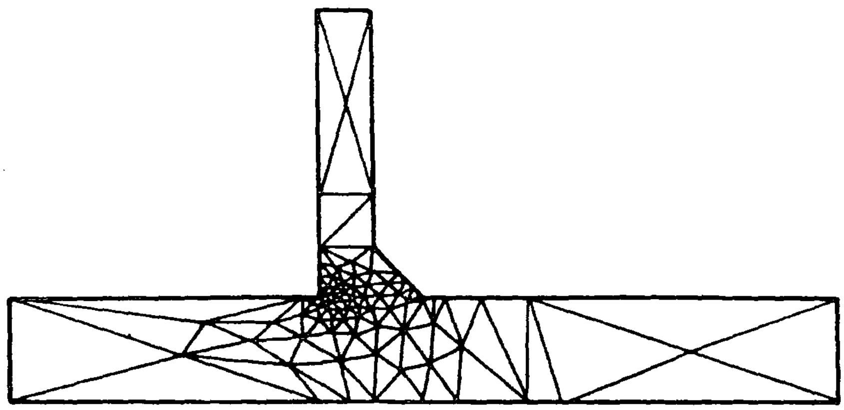
Mesh	N_{dof}	$\|u_h\|_E^2$	$\|e\|_E$	ϵ_{proj} ϵ_{aver}	θ_{proj} θ_{aver}	η ($\times 100$)
1	74	$4.92924E-02$	$1.88653E-02$	$1.24579E-02$	0.66	8.47
				$1.44749E-02$	0.77	
2	179	$4.95569E-02$	$9.56033E-03$	$6.18858E-03$	0.65	4.29
				$7.83910E-03$	0.82	
3	175	$4.95808E-02$	$8.21583E-03$	$5.90213E-03$	0.72	3.69
				$7.10570E-03$	0.86	

Table 3.2 Machine part under plane stress conditions.

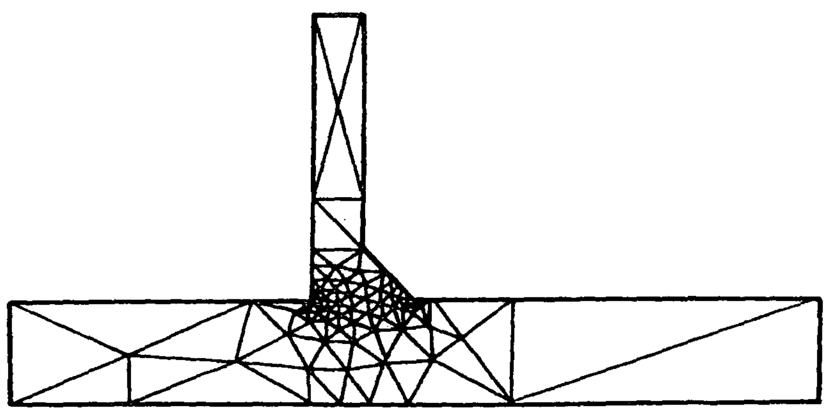
Exact solution $\| u \|_E^2 = 4.964835E-02$.



(a) Mesh 1. $\eta = 8.5\%$.



(b) Mesh 2. $\eta = 4.3\%$.



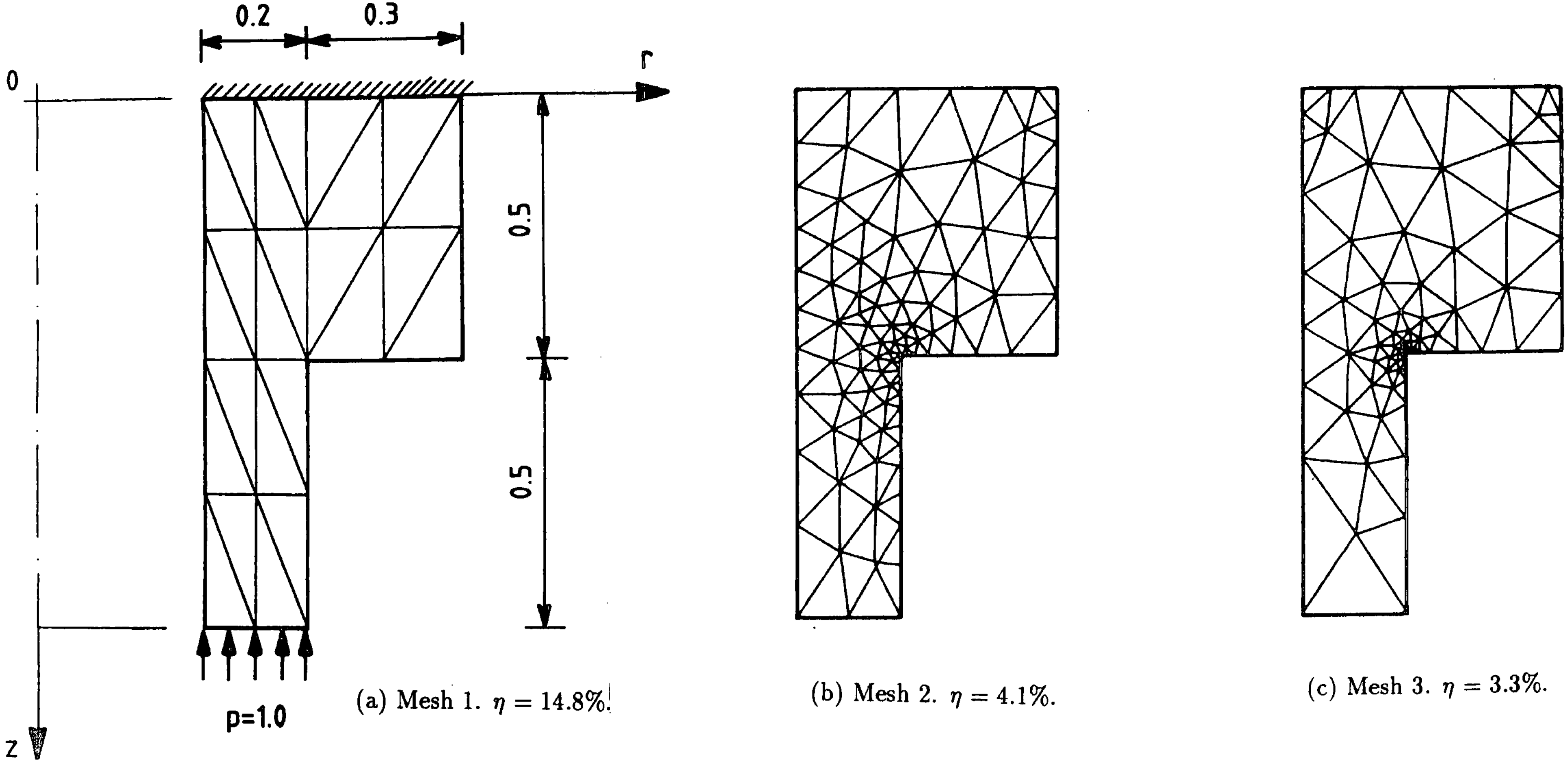
(c) Mesh 3. $\eta = 3.7\%$.

Figure 3.2 Adaptively designed meshes for machine part.

Mesh	N_{dof}	$\|u_h\|_E^2$	$\|e\|_E$	ϵ_{proj} ϵ_{aver}	θ_{proj} θ_{aver}	η ($\times 100$)
1	112	$5.55652E-01$	$1.11617E-01$	$7.216E-02$ $7.703E-02$	0.65 0.69	14.8
2	658	$5.67158E-01$	$3.0861E-02$	$2.073E-02$ $2.317E-02$	0.67 0.75	4.09
3	587	$5.67479E-01$	$2.5127E-02$	$1.448E-02$ $1.613E-02$	0.58 0.64	3.33

Table 3.3 Axisymmetric problem.

Exact solution $\| u \|_E^2 = 5.68110E-01$.



(a) Mesh 1. $\eta = 14.8\%$.

(b) Mesh 2. $\eta = 4.1\%$.

(c) Mesh 3. $\eta = 3.3\%$.

Figure 3.3 Adaptively designed meshes for axisymmetric problem.

CHAPTER 4

Error Estimation for Problems with Singularities.

4.1 Introduction.

The results obtained so far all depend on the true solution possessing a high degree of regularity. This assumption was necessary in part to ensure that the superconvergence results were valid and to allow the use of various technical devices such as the Bramble Hilbert Lemma. It might be thought that the assumptions on the regularity were needed only for such technicalities and that the techniques would perform even if the assumptions were not satisfied. In fact the methods fail to perform satisfactorily in cases where the assumed regularity is lacking.

One particularly important area where existing methods fail to perform satisfactorily is the case of problems for which the true solution possesses some kind of singularity. In fact, most problems which are of practical interest fall into this category. For example, an engineer may wish to obtain information as to the likely behaviour of a structure should a crack develop. The stress field at the tip of the crack is found to be singular. To the best of our knowledge there are no rigorously analyzed estimators which can cope satisfactorily with such singularities. Thus the existing estimators fail to perform in precisely the situations in which they are most needed.

In this chapter we deal with the problem of generalising error estimators to problems with singularities and also the extension of the foregoing analysis to such cases. The framework developed in Chapter 2 allowed us to understand why an estimator should perform well for smooth problems. The philosophy behind the investigations of this chapter is that this framework should also provide insight as to why an estimators should fail to perform for problems with singularities.

We shall illustrate our ideas by dealing with a model one dimensional two point boundary value problem, introduced in section 4.3, whose solution we shall approximate by using piecewise linear basis functions. In section 4.5 we use the insight gained to propose a new estimator capable of coping with singular problems. Although the method of deriving this new estimator is somewhat indirect, the final result is found to represent a natural generalisation of the existing estimator and is particularly simple to implement in practice. Illustrative numerical results are also included.

4.2 Notation.

The notation will be similar to that of Chapter 2 except that we shall deal with one spatial dimension only. This will be reflected by the fact that we shall use I rather than Ω to denote an open bounded interval (α, β) of the real line.

We shall also need to extend the definition of the Sobolev spaces $H^m(I)$ of Chapter 2 to nonintegral and negative values of m. For m positive and nonintegral let $m = M + \sigma$, where M is an integer and $0 < \sigma < 1$. The norm on $H^m(I)$ is given by

$$\| v \|_{m,2,I} = \left(\| v \|_{M,2,I}^2 + J_{M,\sigma,I}[v^{(M)}] \right)^{\frac{1}{2}} \tag{4.2.1}$$

where

$$J_{M,\sigma,I}[v] = \int_I \int_I \frac{|D^M v(x) - D^M v(y)|^2}{|x-y|^{1+2\sigma}} dx dy. \tag{4.2.2}$$

For negative m the norm on $H^m(I)$ is given by

$$\| v \|_{m,2,I} = \sup_{w \in H_0^{-m}(I)} \frac{| \int_I v(x) w(x) dx |}{\| w \|_{-m,2,I}}. \tag{4.2.3}$$

The standard reference domain in $\Re^1$ is taken to be the interval $\hat{K} = (0,1)$, and the finite elements are open intervals of the form

$$I_i^h = (x_i^h, x_{i+1}^h) \quad i = 0, 1, \ldots, m^h - 1$$

where

$$\alpha = x_0^h < x_1^h < \ldots < x_{m^h}^h = \beta,$$

$\{x_i^h\}$ are referred to as the nodes of the partition. When there is no danger of confusion we shall omit the superscript h. The diameter h_i of I_i is given by

$$h_i = \text{diam}(I_i) = x_{i+1} - x_i, \text{ for } i = 0, 1, \ldots, m - 1$$

and we define

$$h = \max_{i \in \{0,1,\ldots,m-1\}} h_i \text{ and } \underline{h} = \min_{i \in \{0,1,\ldots,m-1\}} h_i.$$

We shall assume that the meshes are *regular* and satisfy the *inverse assumption* (see Chapter 2). The finite element space $S^h(I)$ is taken to consist of piecewise linear functions, *i.e.* it has degree 1. The subspace $H_0^1(I) \cap S^h(I)$ is denoted by $S_0^h(I)$.

4.3 The model problem and its approximation.

We shall, without loss of generality, take $I = (0,1)$ and for simplicity we shall consider the two point boundary value problem

$$L[u] \equiv -\frac{d}{dx}\left[a(x)\frac{du}{dx}\right] + b(x)u(x) = f(x) \text{ on } I \tag{4.3.1}$$

with the homogeneous essential boundary conditions

$$u(0) = 0 \text{ and } u(1) = 0. \tag{4.3.2}$$

We shall make the following assumptions about the coefficients $a(x)$ and $b(x)$

- $a(x) \in C^1(\bar{I})$ and that there exist constants $\underline{a}$ and $\bar{a}$ such that

$$0 < \underline{a} \leq a(x) \leq \bar{a} < \infty \quad \forall x \in \bar{I}. \tag{4.3.3}$$

- $b(x) \in C(\bar{I})$ and that there exists a constant $\bar{b}$ such that

$$0 \leq b(x) \leq \bar{b} < \infty \quad \forall x \in \bar{I}. \tag{4.3.4}$$

Regarding the data $f(x)$, we shall initially assume that f is smooth but later we shall consider cases where f has somewhat less regularity. We shall state the assumed regularity of f in each instance.

The weak or variational form of the problem (4.3.1) is to find

$$u \in H_0^1(I) : a(u,v) = (f,v) \quad \forall v \in H_0^1(I) \tag{4.3.5}$$

where

- $a(.,.)$ is the bilinear form given by

$$a(u,v) = \int_I \left\{ a(x)\frac{du}{dx}\frac{dv}{dx} + b(x)u(x)v(x) \right\} dx. \tag{4.3.6}$$

- $(f,.)$ is the linear form given by

$$(f,v) = \int_I f(x)v(x)dx. \tag{4.3.7}$$

Under the above assumptions on a and b it may be shown that $a(.,.)$ is a continuous and coercive bilinear form on $H_0^1(I) \times H_0^1(I)$. That is

$$|a(u,v)| \leq C_1(\underline{a}, \bar{a}, \bar{b}, I) \parallel u \parallel_{1,2,I} \parallel v \parallel_{1,2,I} \quad \forall u, v \in H_0^1(I) \tag{4.3.8}$$

and

$$a(v,v) \geq C_2(\underline{a}, \bar{a}, \bar{b}, I) \parallel v \parallel_{1,2,I}^2 \quad \forall v \in H_0^1(I). \tag{4.3.9}$$

Throughout we shall use the letter C to denote generic positive constants which need not necessarily assume the same value or be related in any two places.

In this chapter we shall be concerned with the case where the true solution u is less regular than we have assumed so far. In particular, we assumed that the true solution $u \in H^3 \cap H_0^1$. The regularity of the solution in $\Re^1$ is related to the regularity of the data. The following standard result guarantees the existence of a unique solution to the variational problem (4.3.5) and also relates the regularity of the solution to the regularity of the data f

THEOREM 4.3.1 *Suppose that $f \in H^s(I)$ $s > -1$, then there exists a unique solution $u \in H_0^1(I) \cap H^{s+2}(I)$ to the problem (4.3.5). Further,*

$$\| u \|_{s+2,2,I} \leq C(\underline{a}, \bar{a}, \bar{b}, I) \| f \|_{s,2,I}, \tag{4.3.10}$$

for some constant C.

Proof. See for example Oden and Reddy [59]. ∎

The finite element approximation is defined to be the solution of

$$u_h \in S_0^h(I) : a(u_h, v_h) = (f, v_h) \quad \forall v_h \in S_0^h(I). \tag{4.3.11}$$

We shall be concerned with estimating the error in the *energy norm* given by

$$\| v \|_E = a(v,v)^{\frac{1}{2}}. \tag{4.3.12}$$

(4.3.8) and (4.3.9) show that the energy norm is equivalent to the norm on $H^1(I)$.

Letting $e(x) = u(x) - u_h(x)$ denote the error in the approximation we have the following *a priori* estimate on the energy norm of the error in terms of the regularity of the data f

THEOREM 4.3.2 *Suppose that $f \in H^s(I)$ $s > -1$, then for h sufficiently small*

$$\| e \|_E \leq Ch^\nu \| f \|_{s,2,I} \tag{4.3.13}$$

where the constant C is independent of h, u and f and

$$\nu = \min(1, s+1). \tag{4.3.14}$$

Proof. See for example Oden and Reddy [59]. ∎

4.4 An a posteriori error estimator for singular problems.

One drawback of the analysis of the previous chapter, and for all *a posteriori* estimators in general, is that the theoretical results rely heavily on a high assumed regularity of the solution. In the case of piecewise linear finite element approximation we had to assume that the true solution $u \in H^3(I)$. Such high regularity assumptions fail to be satisfied in almost all practical applications. Moreover, the regularity requirements are not just necessary to compensate for any limitations on the part of the analyst. In practice it is found that very often unless the solution is regular, the error estimator will not tend towards the true error as the partitions are refined. Naturally this constitutes a serious problem since it means that our estimators fail to behave satisfactorily in precisely the situations in which we wish to apply them. At the moment there exist no really satisfactory means of estimating the error in the approximation to a problem for which the solution is singular.

In this chapter we shall for simplicity consider the one dimensional two point boundary value problem (4.3.5). However, in practice we are more concerned with the higher dimensional cases. There are important differences between the one dimensional and higher dimensional cases. One such difference is that whereas for a one dimensional problem regular data f necessarily gives a regular solution (see Theorem 4.3.1), the same result is not true in two (or more) dimensions.

In two dimensions the domain Ω may give rise to singularities, even if the data f is very regular. Let $\Omega \in \Re^2$ be a bounded, simply connected domain with boundary $\partial\Omega$ consisting of a finite number of straight line segments meeting at vertices $\{V^i\}_{i=1}^M$, which have interior angles ϑ_i. Introduce polar coordinates (r, ϑ) at a vertex V^i of the domain Ω, so that the interior of the wedge is specified by $0 < \vartheta < \vartheta_i$, and for $i = 1, 2, ..., M$ let

$$\alpha_i = \frac{\pi}{\vartheta_i}. \tag{4.4.1}$$

Near the vertex V^i the true solution of the two dimensional analogue of (4.3.1) behaves like [36], [44], [49]

$$u(r, \vartheta) = c_i r^{\alpha_i} \Big[\ln \frac{1}{r} \Big]^{\beta_i} \sin(\alpha_i \vartheta) + w(r, \vartheta) \tag{4.4.2}$$

where c_i is a constant, $\beta_i \equiv 0$ unless $\alpha_i = 2, 3, ...$ and w is a smoother function than the first term in the representation (4.4.2). Moreover, letting Ω_i be the intersection between the domain Ω and a disc centred at the vertex V^i, it is possible to show [36], [44], [49] that if f is smooth then

$$u \in C^\infty \Big(\Omega \setminus [\cup_{i=1}^M \bar{\Omega}_i] \Big). \tag{4.4.3}$$

These results are important for several reasons. Firstly, they show that the even if f is very smooth the solution may not be. Secondly, they show that the singularity is only local in nature, and *cuts-off* outside the neighbourhood of the vertex.

In order to try and simulate this behaviour with our model one dimensional problem we shall assume that the true solution takes the form

$$u(x) = s(x) + w(x) \tag{4.4.4}$$

where $w(x) \in H^3(I)$ represents the smooth component of the solution and $s(x)$ represents the singular component of the solution. Moreover we shall further assume that $s(x)$ takes the particular form

$$s(x) = kx^\alpha \quad \alpha \in \left(\frac{1}{2}, 2\right) \tag{4.4.5}$$

in a neighbourhood of x_0^h and *cuts-off* outside this neighbourhood. The condition that $\alpha > \frac{1}{2}$ must be imposed to ensure that the error measured in the energy norm is finite, whilst the condition $\alpha < 2$ is imposed since if $\alpha \geq 2$, the major contribution to the error comes from smoother components.

4.4.1 An abstract a posteriori error estimator.

As one of our examples of the general results of Chapter 2 we showed that the Babuška-Rheinboldt estimator could be obtained by means of a simple intuitive argument based on the availability of a *superconvergent* approximation to the gradient. Consequently, we might attribute the failure of this estimator when the true solution is less smooth to the degradation of the superconvergence results when the solution is singular.

The problem of finding an error estimator which will perform well for singular problems is thus equivalent to that of finding a way of recovering the gradient of the singular solution. Once this point has been realised, the road to defining an estimator which will be valid for singular problems is clear, if not altogether straight forward. Unfortunately, the lack of any existing superconvergence results for singular problems means that we cannot simply modify the estimator by using an alternative recovery scheme for the gradient.

REMARK We pause to mention that there are superconvergence techniques which allow the coefficient of the leading term in the expansion of the singular components to be estimated [10]. This could, in principle, be used to form the basis of an error estimation algorithm, but we shall not pursue the idea here since the cost and extra effort which such a method would entail is prohibitive. ∎

The approach we shall adopt is as follows. Firstly we shall define an *abstract error estimator* (by which we shall mean an estimator which exists in theory but which may not be available to us explicitly), paying little regard as to how it may be obtained in practice. Once we have analysed the abstract estimator (and in particular shown that it is *asymptotically exact* even for singular problems), we shall then turn to the problem of how the *abstract estimator* itself, rather than the error, may be estimated in practice.

We define $s_h, w_h \in S_0^h(I)$ to be the solutions of the problems:

$$s_h \in S_0^h(I) : a(s_h, v_h) = a(s, v_h) \quad \forall v_h \in S_0^h(I) \tag{4.4.6}$$

and

$$w_h \in S_0^h(I) : a(w_h, v_h) = a(w, v_h) \quad \forall v_h \in S_0^h(I). \tag{4.4.7}$$

REMARK It is easily demonstrated (using the uniqueness property for solutions to (4.3.11), and (4.4.4)) that $u_h = s_h + w_h$. Thus, defining s_h and w_h in this way gives us a means of partitioning the approximation u_h into components corresponding to the singular and smooth components of the true solution. Of course, we would not be able to carry out this partitioning in practice. □

Having carried this partitioning of u_h we may define an approximation to the gradient of the true solution u as follows:

- taking w_h we calculate $G_h(w_h)$ where G_h is the recovery operator defined previously
- taking s_h we assume that there is some way of recovering s' exactly, which of course we use to find s'.

REMARK The assumption that we can find s' exactly using s_h may seem rather strong. However, in the sequel we shall find that it is unnecessary to actually find s' explicitly. It should be borne in mind that in this section we are merely trying to define an *abstract estimator* which will estimate the error well even if the true solution is singular. If, in order to do this, we wish to assume that we have s' explicitly then this will not detract from the overall result, but will only make the task of estimating the estimator more difficult. ■

We now take $s' + G_h(w_h)$ to be our approximation to u', and define an abstract estimator $\tilde{\epsilon}$, where

$$\tilde{\epsilon}^2 = \sum_{i=0}^{m-1} \tilde{\eta}_i^2 \tag{4.4.8}$$

with

$$\tilde{\eta}_i^2 = \int_{I_i^h} a(x)\{s'(x) + G_h(w_h) - u_h'(x)\}^2 dx. \tag{4.4.9}$$

4.4.2 Analysis of the abstract estimator.

Before discussing the question of *asymptotic exactness* of the estimator, we give the following generalisation of Theorem 2.3.1 .

THEOREM 4.4.1 *Let $K[p]$ be the quadratic functional*

$$K[p] = \int_I \frac{1}{a(x)}\left\{p - a(x)\frac{du}{dx}\right\}^2 + \frac{h^\tau}{M}\int_I \left\{f(x) + \frac{dp}{dx} - b(x)u_h(x)\right\}^2 dx$$

and let F be the set

$$F = \{p \in L_2(I) : f + \frac{dp}{dx} \in L_2(I)\}, \tag{4.4.10}$$

then for any $p \in F$ the following bound holds

$$\| e \|_E^2 \leq \{1 + O(h^{1-\frac{1}{2}\tau})\}K[p] \text{ as } h \to 0 \tag{4.4.11}$$

where the constant in the $O(h^{1-\frac{1}{2}\tau})$ is independent of h and u.

Proof: Omitted. ∎

REMARK Theorem 4.4.1 is true for the more general problems discussed in earlier chapters, but here we shall restrict attention to the one dimensional two point boundary value problem.

We shall show in the proof of Theorem 4.4.2 that the choice

$$p(x) = a(x)\{s'(x) + G_h(w_h)\} \tag{4.4.12}$$

satisfies the conditions of Theorem 4.4.1 . By making this choice of p we obtain

THEOREM 4.4.2 *Suppose that the family $\{T^h\}$ is quasi-uniform and that the true solution u is of the form (4.4.4), then the estimator $\tilde{\epsilon}$ is asymptotically exact. That is,*

$$\| e \|_E = \tilde{\epsilon}\{1 + O(h^\gamma)\} \text{ as } h \to 0. \tag{4.4.13}$$

Proof. Firstly, note that

$$p(x) = a(x)[s'(x) + G_h(w_h)] \in F \tag{4.4.14}$$

since

$$\begin{aligned} & f + \frac{dp}{dx} \\ =& -\frac{d}{dx}\Big[au'\Big] + bu + \frac{d}{dx}\Big[a(s' + aG_h(w_h))\Big] \\ =& -\frac{d}{dx}\Big[a(w' + G_h(w_h))\Big] + bu \in L^2(I). \end{aligned}$$

Making this choice of p in Theorem 4.4.1 yields the bound

$$\| e \|_E^2 \leq \Big[\sum_{i=0}^{m-1} \tilde{\eta}_i^2 + \frac{h^\tau}{M} \| f + p' - bu_h \|_{0,2,I}^2 \Big] \{1 + O(h^{1-\frac{\tau}{2}})\}. \tag{4.4.15}$$

We deal with each term in (4.4.15) separately.

Analysis of the second term:

$$\begin{aligned} & \| f + p' - bu_h \|_{0,2,I} \\ =& \| be - [a(w' - G_h(w_h))]' \|_{0,2,I} \\ \leq& \sqrt{\bar{b}} \| e \|_{0,2,I} + \sqrt{\bar{a}}\, |w' - G_h(w_h)|_{1,2,I} + \sqrt{\bar{a}'}\, |w' - G_h(w_h)|_{0,2,I}, \end{aligned} \tag{4.4.16}$$

where $|a'(x)| \leq \bar{a}', x \in I$. Using the estimates (4.5.5) and (4.5.6) applied to w and the inverse estimate (4.5.8) we obtain

$$|w' - G_h(w_h)|_{0,2,I} \leq C(u)h \parallel e \parallel_E \tag{4.4.17}$$

and

$$|w' - G_h(w_h)|_{1,2,I} \leq C(u) \parallel e \parallel_E . \tag{4.4.18}$$

Further, using the Aubin-Nitsche method gives

$$\parallel e \parallel_{0,2,I} \leq Ch \parallel e \parallel_E . \tag{4.4.19}$$

Using (4.4.17), (4.4.18) and (4.4.19) in (4.4.16) now gives

$$\parallel f + p' - bu_h \parallel_{0,2,I} \leq \bar{C}(u) \parallel e \parallel_E . \tag{4.4.20}$$

Analysis of the first term:

Moreover,

$$\Big[\sum_{i=0}^{m-1} \tilde{\eta}_i^2\Big]^{\frac{1}{2}}$$

$$= \Big[\int_I a(x)(s' + G_h(w_h) - u_h')^2 dx\Big]^{\frac{1}{2}}$$

$$
\begin{aligned}
&\leq \bar{a} \parallel w' - G_h(w_h) \parallel_{0,2,I} + \Big[\int_I a(x) e'^2 dx \Big]^{\frac{1}{2}} \\
&\leq \bar{a} \parallel w' - G_h(w_h) \parallel_{0,2,I} + \parallel e \parallel_E \\
&\leq C(a,u) h \parallel e \parallel_E + \parallel e \parallel_E, \qquad (4.4.21)
\end{aligned}
$$

where we have used (4.4.17).

It now follows from (4.4.15) and (4.4.20), that for h sufficiently small

$$
\parallel e \parallel_E^2 \leq \{1 + O(h^\gamma)\} \sum_{i=0}^{m-1} \tilde{\eta}_i^2, \qquad (4.4.22)
$$

where $\gamma = \min(\epsilon, 1 - \frac{\tau}{2})$. Finally, combining (4.4.21) and (4.4.22) gives the result as claimed. ∎

This result shows that it is indeed possible to estimate the error asymptotically exactly even if the true solution is singular, provided that we can recover information about the gradient of the singular component.

4.4.3 Approximation of the abstract estimator.

In the following we shall, in order to simplify the presentation, assume that the mesh is uniform *i.e.* $\underline{h} = h$. Later we shall show how the results and conclusions may be modified to the more general quasi-uniform meshes.

Before continuing, we pause to collect some preliminary lemmas:

LEMMA 4.4.3 *Let $w \in H^3(\alpha,\beta)$ then there exists a constant $C > 0$ which is independent of h and w such that*

$$|w(\beta) - 2w(\frac{1}{2}\alpha + \frac{1}{2}\beta) + w(\alpha)| \leq h^2|w|_{2,\infty,(\alpha,\beta)} \tag{4.4.23}$$

where $h = \beta - \alpha$.

Proof. Follows from an application of the Peano Kernel Theorem, see e.g. Davis [29]. ∎

LEMMA 4.4.4 *Let $\tilde{u}$ denote the solution of the problem*

$$\tilde{u} \in H_0^1(I) : a(\tilde{u}, v) = (\tilde{f}, v) \quad \forall v \in H_0^1(I) \tag{4.4.24}$$

and let $\tilde{u}_h$ denote the solution of the problem

$$\tilde{u}_h \in S_0^h(I) : a(\tilde{u}_h, v_h) = (\tilde{f}, v_h) \quad \forall v_h \in S_0^h(I). \tag{4.4.25}$$

Further, assume that $\tilde{u}$ has the form

$$\tilde{u}(x) = kx^\alpha + r(x)$$

where $r \in H^2(I)$ and $\alpha > \frac{1}{2}$. Let $\{x_i\}$ be the nodes of the mesh then

$$|\tilde{u}(x_i) - \tilde{u}_h(x_i)| \leq Ch^\mu \parallel \tilde{u} \parallel_{\alpha+\frac{1}{2}-\epsilon,2,I} \quad \forall i \tag{4.4.26}$$

where $\epsilon > 0$ and $\mu = \min(2, \alpha + \frac{1}{2} - \epsilon)$.

Proof. This is a slight generalisation of a standard superconvergence result and may be obtained using arguments found in e.g. Oden and Reddy [59] pp.348–349. ∎

Before giving the next lemma we introduce some new notation. The weighted L^2 norm with weighting function $a(x)$ is denoted by

$$_a \| f \|^2_{0,2,(\alpha,\beta)} = \int_\alpha^\beta a(x)|f(x)|^2 dx. \tag{4.4.27}$$

We recall from Chapter 2 that the intervals $\hat{I}_i^h \subset I$ are obtained from a partition $\{T^h\}$ by defining

$$\hat{I}_i^h = \cup_{j \in \mathrm{adj}(i)} \bar{I}_j^h \tag{4.4.28}$$

where, for $i = 0, 1, 2, ..., m-1$

$$\mathrm{adj}(i) = \{j : \bar{I}_i^h \cap \bar{I}_j^h \neq \emptyset\}. \tag{4.4.29}$$

It is easy to see that this definition gives

$$\hat{I}_i^h = [x_{i-1}^h, x_{i+1}^h] \text{ for } i = 1, \dots, m^h - 1 \tag{4.4.30}$$

and

$$\hat{I}_0^h = [x_0^h, x_2^h] \text{ and } \hat{I}_{m-1}^h = [x_{m-2}^h, x_m^h], \tag{4.4.31}$$

so that $\hat{I}_i^h$ consists of the element I_i^h and the elements 'near' to it.

LEMMA 4.4.5 *Let $w \in H^3(I)$ and let G_h be the recovery operator defined previously, then*

$$_a \parallel G_h(w_h) - w_h' \parallel_{0,2,I_i^h} \leq Ch^{\frac{3}{2}}(|w|_{2,\infty,I_i^h} + \parallel w \parallel_{2,2,I}) \qquad (4.4.32)$$

where C is a constant which depends on a but is independent of h and w.

Proof. Case $i = 0$:

Since

$$_a \parallel G_h(w_h) - w_h' \parallel^2_{0,2,I_0^h} \leq \bar{a} \parallel G_h(w_h) - w_h' \parallel^2_{0,2,I_0^h} \qquad (4.4.33)$$

and

$$\parallel G_h(w_h) - w_h' \parallel^2_{0,2,I_0^h} = \frac{h}{12}|w_h'(x^h_{\frac{3}{2}}) - w_h'(x^h_{\frac{1}{2}})|^2, \qquad (4.4.34)$$

we get that

$$\begin{aligned} &_a \parallel G_h(w_h) - w_h' \parallel^2_{0,2,I_0^h} \\ &\leq \frac{h\bar{a}}{12}|w_h'(x^h_{\frac{3}{2}}) - w_h'(x^h_{\frac{1}{2}})|^2 \\ &= \frac{\bar{a}}{12h}|w_h(x^h_2) - 2w_h(x^h_1) + w_h(x^h_0)|^2. \end{aligned} \qquad (4.4.35)$$

Moreover,

$$\begin{aligned} &\frac{1}{h}|w_h(x_2) - 2w_h(x_1) + w_h(x_0)| \\ &\leq \frac{1}{h}|w(2h) - 2w(h) + w(0)| + \frac{4}{h}\max_{i=0,1,2}|e_w(x_i)|, \end{aligned} \qquad (4.4.36)$$

where $e_w(x) = w(x) - w_h(x)$. Using Lemma 4.4.3 gives

$$\frac{1}{h}|w(2h) - 2w(h) + w(0)| \leq Ch|w|_{2,\infty,I_0^h} \tag{4.4.37}$$

and Lemma 4.4.4 gives

$$\frac{4}{h}\max_{i=0,1,2}|e_w(x_i)| \leq Ch \parallel w \parallel_{2,2,I} . \tag{4.4.38}$$

Now (4.4.36) ,(4.4.37) and (4.4.38) give

$$\frac{1}{h}|w_h(x_2) - 2w_h(x_1) + w_h(x_0)| \leq Ch(|w|_{2,\infty,I_0^h} + \parallel w \parallel_{2,2,I}), \tag{4.4.39}$$

and using (4.4.35) and (4.4.39) gives

$$_a \parallel G_h(w_h) - w_h' \parallel^2_{0,2,I_0^h} \leq C(a)h^3(|w|_{2,\infty,I_0^h} + \parallel w \parallel_{2,2,I})^2, \tag{4.4.40}$$

as required.

Case $i = 1, 2, ..., m-2$:

It is easily shown that

$$_a \parallel G_h(w_h) - w_h' \parallel^2_{0,2,I_i^h} \leq \frac{\bar{a}h}{12}[J_+^2 - J_+J_- + J_-^2], \tag{4.4.41}$$

where

$$J_+ = \frac{1}{h}\{ w_h(x_{i+2}) - 2w_h(x_{i+1}) + w_h(x_i)\} \tag{4.4.42}$$

and

$$J_- = \frac{1}{h}\{ w_h(x_{i+1}) - 2w_h(x_i) + w_h(x_{i-1})\} . \tag{4.4.43}$$

By using a similar argument to that used to obtain (4.4.39) we obtain the bounds

$$|J_+| \le Ch(|w|_{2,\infty,\hat{I}_i^h} + \| w \|_{2,2,I}) \tag{4.4.44}$$

and

$$|J_-| \le Ch(|w|_{2,\infty,\hat{I}_i^h} + \| w \|_{2,2,I}) . \tag{4.4.45}$$

Finally (4.4.41), (4.4.44) and (4.4.45) give

$$_a \| G_h(w_h) - w_h' \|_{0,2,I_i^h}^2 \le C(a)h^3(|w|_{2,\infty,\hat{I}_i^h} + \| w \|_{2,2,I})^2 \tag{4.4.46}$$

as required. The proof in the case $i = m - 1$ is similar to that in the case $i = 0$. ∎

Armed with these preliminary lemmas we may consider how the abstract estimator may be estimated in practice. The key result is contained in the following theorem, which shows that the abstract estimator may be easily estimated in practice by multiplying the classical local estimator by a certain scalar correction factor, and the correction factor is given explicitly.

THEOREM 4.4.6 *Let η_i and $\tilde{\eta}_i$ be the local error estimators defined in (2.6.2) and (4.4.9) respectively, then*

$$\lim_{h \to 0} \frac{\tilde{\eta}_i^2}{\eta_i^2} = A(\alpha, i)^2 \tag{4.4.47}$$

where for $\alpha \geq 2$ or $\alpha = 1$,

$$A(\alpha, i)^2 \equiv 1 \tag{4.4.48}$$

and for other values of $\alpha > \frac{1}{2}$,

$$A(\alpha,i)^2 = \begin{cases} \frac{3}{2\alpha-1}\left|\frac{\alpha-1}{2^{\alpha-1}-1}\right|^2, & \text{if } i = 0 \\ 1, & \text{if } I_i^h \cap \mathrm{supp}(s) = \emptyset \\ \frac{12}{2\alpha-1}\left\{\frac{\alpha^2[(1+i)^{2\alpha-1}-i^{2\alpha-1}]-(2\alpha-1)[(1+i)^\alpha-i^\alpha]^2}{\Delta(i+1)^2-\Delta(i+1)\Delta(i)+\Delta(i)^2}\right\}, & i \text{ otherwise} \end{cases}$$

where supp(s) denotes the support of s and

$$\Delta(i) \equiv (i+1)^\alpha - 2i^\alpha + (i-1)^\alpha. \tag{4.4.49}$$

Proof. *Case $i = 0$:*

Letting $a_{\frac{1}{2}} = a(x_{\frac{1}{2}})$ and recalling that $a \in C^1(\bar{I})$ we have that

$$\begin{aligned}\eta_0^2 &= {}_a \| G_h(u_h) - u_h' \|^2_{0,2,I_0^h} \\ &= \frac{h}{12}\{ a_{\frac{1}{2}} + O(h)\} |u_h'(x_{\frac{1}{2}}) - u_h'(x_{\frac{3}{2}})|^2 \\ &= \frac{1}{12h}\{ a_{\frac{1}{2}} + O(h)\} |u_h(x_2) - 2u_h(x_1) + u_h(x_0)|^2. \end{aligned} \tag{4.4.50}$$

Further,

$$\begin{aligned} & |u_h(x_2) - 2u_h(x_1) + u_h(x_0)| \\ \leq & |s(2h) - 2s(h) + s(0)| + |w(2h) - 2w(h) + w(0)| \\ & \qquad + |e(2h) - 2e(h) + e(0)| \\ \leq & 2kh^\alpha |2^{\alpha-1} - 1| + C_1 h^2 |w|_{2,\infty,\hat{I}_0^h} + 4 \max_{j=0,1,2} |e(x_j)| \\ \leq & 2kh^\alpha |2^{\alpha-1} - 1| + C_1 h^2 |w|_{2,\infty,\hat{I}_0^h} + C_2 h^\mu \| u \|_{\alpha+\frac{1}{2}-\epsilon,2,I} \end{aligned} \tag{4.4.51}$$

with $\mu = \min(2, \alpha + \frac{1}{2} - \epsilon)$ and where we have used Lemmas 4.4.3 and 4.4.4 . We can show similarly (using the reverse triangle inequality) that

$$\begin{aligned} |u_h(x_2) - 2u_h(x_1) + u_h(x_0)| \geq & 2kh^\alpha |2^{\alpha-1} - 1| - C_1 h^2 |w|_{2,\infty,\hat{I}_i^h} \\ & - C_2 h^\mu \| u \|_{\alpha+\frac{1}{2}-\epsilon,2,I} . \end{aligned} \tag{4.4.52}$$

Since $\alpha \in (\frac{1}{2}, 2)$, (4.4.51) and (4.4.52) imply that

$$|u_h(x_2) - 2u_h(x_1) + u_h(x_0)| \sim 2kh^\alpha |2^{\alpha-1} - 1| \text{ as } h \to 0, \tag{4.4.53}$$

and using (4.4.50) and (4.4.53) gives us that

$$\eta_0^2 = {}_a \parallel G_h(u_h) - u_h' \parallel_{0,2,I_0^h}^2 \sim \frac{1}{3}k^2 h^{2\alpha-1} a_{\frac{1}{2}} |2^{\alpha-1} - 1|^2 \text{ as } h \to 0. \tag{4.4.54}$$

Turning now to $\tilde{\eta}_0$ we find that

$$\begin{aligned}\tilde{\eta}_0 &= {}_a \parallel s' + G_h(w_h) - u_h' \parallel_{0,2,I_0^h} \\ &\leq {}_a \parallel s' - s_h' \parallel_{0,2,I_0^h} + {}_a \parallel G_h(w_h) - w_h' \parallel_{0,2,I_0^h} .\end{aligned} \tag{4.4.55}$$

Now,

$$\begin{aligned}&{}_a \parallel s' - s_h' \parallel_{0,2,I_0^h} \\ &\leq {}_a \parallel s' - (\Pi s)' \parallel_{0,2,I_0^h} + {}_a \parallel (\Pi s - s_h)' \parallel_{0,2,I_0^h} \\ &\leq {}_a \parallel s' - (\Pi s)' \parallel_{0,2,I_0^h} + C(a) h^{-\frac{1}{2}} \max_{i=0,1,2} |e_s(x_i^h)| \\ &\leq {}_a \parallel s' - (\Pi s)' \parallel_{0,2,I_0^h} + C(a) h^{\mu-\frac{1}{2}} \parallel s \parallel_{\alpha+\frac{1}{2}-\epsilon,2,I}\end{aligned} \tag{4.4.56}$$

where $e_s(x) = s(x) - s_h(x)$ and we have used Lemma 4.4.4 . Further,

$$\begin{aligned}{}_a \parallel s' - (\Pi s)' \parallel_{0,2,I_0^h}^2 &= \{a_{\frac{1}{2}} + O(h)\} |s - \Pi s|_{1,2,I_0^h}^2 \\ &= \frac{1}{2\alpha - 1}\{a_{\frac{1}{2}} + O(h)\} h^{2\alpha-1} |\alpha - 1|^2.\end{aligned} \tag{4.4.57}$$

(4.4.55), (4.4.56) and (4.4.57) now give on applying Lemma 4.4.5

$$\tilde{\eta}_0 \leq \frac{1}{\sqrt{2\alpha-1}}\{a_{\frac{1}{2}} + O(h)\}^{\frac{1}{2}} h^{\alpha-\frac{1}{2}}|\alpha-1| + C(a)h^{\mu-\frac{1}{2}} \parallel s \parallel_{\alpha+\frac{1}{2}-\epsilon,2,I}$$
$$+ C(a)h^{\frac{3}{2}}(|w|_{2,\infty,I_0^h} + \parallel w \parallel_{2,2,I}). \qquad (4.4.58)$$

By using the reverse triangle inequality and following the same steps used to obtain (4.4.58), we get the lower bound

$$\tilde{\eta}_0 \geq \frac{1}{\sqrt{2\alpha-1}}\{a_{\frac{1}{2}} + O(h)\}^{\frac{1}{2}} h^{\alpha-\frac{1}{2}}|\alpha-1| - C(a)h^{\mu-\frac{1}{2}} \parallel s \parallel_{\alpha+\frac{1}{2}-\epsilon,2,I}$$
$$- C(a)h^{\frac{3}{2}}(|w|_{2,\infty,I_0^h} + \parallel w \parallel_{2,2,I}). \qquad (4.4.59)$$

Recalling that $\alpha \in (\frac{1}{2}, 2)$ and that $\mu = \min(2, \alpha + \frac{1}{2} - \epsilon)$, (4.4.58) and (4.4.59) give

$$\tilde{\eta}_0^2 \sim \frac{1}{2\alpha-1} a_{\frac{1}{2}} h^{2\alpha-1}|\alpha-1|^2 \text{ as } h \to 0. \qquad (4.4.60)$$

Finally we obtain from (4.4.54) and (4.4.60)

$$\lim_{h\to 0} \frac{\tilde{\eta}_0^2}{\eta_0^2} = \left|\frac{\alpha-1}{2^{\alpha-1}-1}\right|^2 \frac{3}{2\alpha-1}, \qquad (4.4.61)$$

as required. The proof in the other cases follows using similar arguments and bounds to those in the case $i = 0$ and is therefore omitted. ∎

If we define modified local estimators $\{\hat{\eta}_i\}$ as

$$\hat{\eta}_i = A(\alpha, i)\eta_i, \quad i = 0, 1, ..., m \qquad (4.4.62)$$

and a modified global estimator $\hat{\epsilon}$ as

$$\hat{\epsilon}^2 = \sum_{i=0}^{m-1} \hat{\eta}_i^2 \tag{4.4.63}$$

then Theorems 4.4.2 and 4.4.6 immediately give us that

THEOREM 4.4.7 *Suppose that the family $\{T^h\}$ is quasi-uniform and that the true solution u is of the form (4.4.4), then*

$$\hat{\epsilon} \approx \tilde{\epsilon} \tag{4.4.64}$$

and consequently

$$\hat{\epsilon} \approx \| e \|_E . \tag{4.4.65}$$

This result shows that the modified estimator $\hat{\epsilon}$ is

- easily computable since it may be obtained from the classical error estimator of Babuška and Rheinboldt by simply multiplying the classical local estimators by a scalar which depends on α
- simple to incorporate into existing codes which use classical estimators
- asymptotically exact even if the solution has a singular component.

REMARK We are of course assuming that the value of α is known exactly, which is sometimes the case for the higher dimensional problems since α can be obtained from looking at the geometry of the domain Ω. If α is not known explicitly, then a lower bound on the value of α may be used instead.

4.4.4 Behaviour of $A(\alpha, i)$.

In spite of the ease with which the modified estimator of the previous section may be incorporated into existing codes, there is still the slight problem that the definition of the correction factors $A(\alpha, i)$ depends on where the *cut-off* of $s(x)$ occurs. In fact through mollification we can force the cut off to occur wherever we wish, thus in order to decide how many of the classical estimators should be corrected using the correction factor $A(\alpha, i)$, we shall examine the behaviour of $A(\alpha, i)$ more closely.

Intuitively, we expect to find for elements further away from the singularity that the influence of the singularity is relatively small and consequently there is no need to correct the classical estimator on these elements. Thus we expect to find that

$$\lim_{i \to \infty} A(\alpha, i) = 1. \tag{4.4.66}$$

In fact we find that for large values of i

$$A(\alpha, i)^2 = 1 + \frac{2(\alpha - 2)^2}{(\alpha - 1)} \frac{1}{i} + O\left(\frac{1}{i^2}\right), \tag{4.4.67}$$

which confirms our expectation. However, this form suggests that $A(\alpha, i)$ may tend to 1 slowly. In order to see how rapid the convergence to unity is, we tabulate $A(\alpha, i)$ for several values of α and i. Table 4.1 shows that since $A(\alpha, i)$ is approximately equal to unity for $i > 1$, it is really only necessary to correct the classical estimators on the element containing the singularity and the elements immediately adjacent to the singular element.

REMARK This conclusion is further supported by the intuition gained from our knowledge of the *pollution effect* of a singularity. Experience shows that the effect of a singularity affects the accuracy of a finite element approximation not only in the element containing the singularity, but also in the elements neighbouring the singular element. However, this pollution effect does not seriously affect the accuracy in other elements (when the error is measured in the energy norm).

4.4.5 Numerical Examples.

In order to illustrate the preceding results we present some numerical examples in which the error is estimated using the classical estimator (4.5.3) and the modified estimator (4.4.62). The function $f(x)$ is chosen so that the true solution for each problem is of the form

$$u(x) = x^{\alpha} - x + \sin(6x) - x\sin(6) \tag{4.4.68}$$

with values of $\alpha = 1.5, 0.9, 0.6$.

The modified estimator is obtained by correcting the classical estimator on the singular element and its neighbour only, since examination of the behaviour of $A(\alpha, i)$ suggested that it should be unnecessary to have to correct on any other elements.

The notation we shall use is as follows

N – number of uniform elements

$\| e \|_E$ – true error in energy norm

ϵ – classical estimator

θ – effectivity index $\epsilon / \| e \|_E$

$\hat{\epsilon}$ – modified estimator

$\hat{\theta}$ – effectivity index $\hat{\epsilon} / \| e \|_E$.

The *effectivity indices* θ and $\hat{\theta}$ measure how well the estimators are performing. If the estimator is asymptotically exact then we expect to find that the effectivity index tends towards unity as the number of elements is increased.

Numerical results for the cases

1. $a(x) \equiv 1; b(x) \equiv 10$
2. $a(x) = \sqrt{x + \frac{1}{10}}; b(x) \equiv 0$
3. $a(x) = (x + \frac{1}{10})^2; b(x) \equiv 0$

are presented in Tables 4.2, 4.3 and 4.4 respectively.

The results show that when $\alpha = 1.5$ the classical estimator performs satisfactorily with $\theta \approx 1$ for all of the above cases. However, as the singularity becomes increasingly severe, the performance of the classical estimator becomes unsatisfactory with $\theta \approx 0.95$ when $\alpha = 0.9$ and $\theta \approx 0.2$ when $\alpha = 0.6$.

The modified estimator performs extremely well and in all cases $\hat{\theta} \to 1$, as the theory developed above would lead us to expect.

4.5 Generalizations of basic results.

In this section we generalize the basic results and conclusions already reached in two directions. Firstly, the correction factors for the practically important case of non-uniform meshes are obtained. Secondly, the correction factors needed for singularities of more general type are discussed.

4.5.1 Correction factors on non-uniform meshes.

A careful examination of the proof of Theorem 4.4.6 reveals that the source of the correction factor $A(\alpha, i)$ was the quantities

$$\eta_0 \approx \| G[\Pi^h s] - (\Pi^h s)' \|_{0,2,I_0^h} \qquad (4.5.1)$$

from (4.4.50),(4.4.51) and (4.4.54); and

$$\tilde{\eta}_0 \approx \| s' - \Pi^h s' \|_{0,2,I_0^h} \tag{4.5.2}$$

from (4.4.55), (4.4.56) and (4.4.57); so that

$$A(\alpha, 0) = \frac{\| s' - \Pi^h s' \|_{0,2,I_0^h}}{\| G_h[\Pi^h s] - (\Pi^h s)' \|_{0,2,I_0^h}}. \tag{4.5.3}$$

In fact, generalizing the proof of Theorem 4.4.6 to the case of *quasi-uniform* meshes would lead to the result that the correction factor on the element (x_e, x_{e+1}) is given by

$$A(e; e+1; \alpha) = \frac{\| s' - \Pi^h s' \|_{0,2,I_e^h}}{\| G_h[\Pi^h s] - (\Pi^h s)' \|_{0,2,I_e^h}}. \tag{4.5.4}$$

REMARK The proof of this result for linear elements follows in precisely the same way as in Theorem 4.4.6 , except that the manipulations become rather complicated owing to the non-uniformity of the mesh. We have therefore omitted the details in this case, preferring instead to give full details for the case of uniform meshes and to extract the key feature which determines the corrective constant in the more general case. This approach is justifiable in the sense that the proof of the general case does not provide any further insight which cannot be extracted from the uniform case and ultimately it is the corrective constant which is of interest. ∎

We may simplify the denominator of (4.5.4) slightly by recalling that G satisfies the consistency condition $(R1)$ of Chapter 2

$$\forall u \in \mathcal{P}_{p+1} \quad G_h[\Pi_p^h u] \equiv u'. \tag{4.5.5}$$

Consequently, we have that

$$G_h[\Pi^h s] - \left(\Pi_s^h\right)' = G_h[\Pi^h s - \Pi^h u] - \left(\Pi_s^h - u\right)' \quad \forall u \in \mathcal{P}_2 \tag{4.5.6}$$

and, by choosing u to be the linear function which agrees with $\Pi^h s$ on (x_e, x_{e+1}) ($i.e.$ we choose u to be the piecewise linear interpolant to s at x_e and x_{e+1}), which we denote by $\tilde{\Pi}^h s$, we find that

$$\| G_h[\Pi^h s] - (\Pi^h s)' \|_{0,2,I_e^h} = \| G_h[\Pi^h s - \tilde{\Pi}^h s] \|_{0,2,I_e^h} . \tag{4.5.7}$$

(4.5.4) may then be rewritten as

$$A(e; e+1; \alpha) = \frac{\| s' - \bar{\Pi}^h s' \|_{0,2,I_e^h}}{\| G_h[\Pi^h s - \tilde{\Pi}^h s] \|_{0,2,I_e^h}}. \tag{4.5.8}$$

This expression may be used to derive the corrective constant for non-uniform meshes.

Case (i): Element (x_0, x_1).

The standard divided difference quotients are defined as

$$s[x_k] = s(x_k) \tag{4.5.9}$$

and then inductively as

$$s[x_i, x_{i+1}, \ldots, x_{i+k}] = \frac{s[x_{i+1}, x_{i+2}, \ldots, x_{i+k}] - s[x_i, x_{i+1}, \ldots, x_{i+k-1}]}{x_{i+k} - x_i}. \quad (4.5.10)$$

With this notation we find that

$$(\Pi^h s)'(x) = \begin{cases} s[x_0, x_1], & x \in (x_0, x_1) \\ s[x_1, x_2], & x \in (x_1, x_2) \end{cases} \quad (4.5.11)$$

so that

$$(\tilde{\Pi}^h s)'(x) = s[x_0, x_1], \quad (4.5.12)$$

and

$$(\tilde{\Pi}^h s - \Pi^h s)'(x) = \begin{cases} 0, & x \in (x_0, x_1) \\ -(x_2 - x_0)s[x_0, x_1, x_2], & x \in (x_1, x_2) \end{cases}. \quad (4.5.13)$$

REMARK This shows why it is profitable to carry out the simplification (4.5.4)–(4.5.8). The recovery operator will be easier to compute due to the zero data on the element under consideration. ∎

It is now a simple matter to compute

$$\| G[\Pi^h s - \tilde{\Pi}^h s] \|^2_{0,2,I_0^h} = \frac{1}{3}(x_1 - x_0)^3 s^2[x_0, x_1, x_2] \quad (4.5.14)$$

and

$$\| s' - (\Pi^h s)' \|^2_{0,2,I_0^h} = \frac{(\alpha-1)^2(x_1-x_0)s^2[x_0,x_1]}{(2\alpha-1)}. \tag{4.5.15}$$

Consequently we deduce that the correction factor on the first element of a non-uniform mesh is given by

$$A^2(0;1;\alpha) = \frac{3(\alpha-1)^2 s^2[x_0,x_1]}{(2\alpha-1)s^2[x_0,x_1,x_2](x_1-x_0)^2}. \tag{4.5.16}$$

The apparent dependence of the correction factor on the absolute mesh spacing can be removed by introducing the ratio of the sizes of the first two elements given by

$$\lambda = \frac{x_2-x_1}{x_1-x_0}. \tag{4.5.17}$$

The correction factor may then be rewritten as

$$A^2(0;1;\alpha) = \frac{3(\alpha-1)^2 s^2[0,1]}{(2\alpha-1)s^2[0,1,1+\lambda]}. \tag{4.5.18}$$

This may be further simplified on noting that

$$s[0,1] = 1 \text{ and } s[0,1,1+\lambda] = \{(1+\lambda)^{\alpha-1}-1\}/\lambda \tag{4.5.19}$$

whereupon it becomes

$$A^2(0;1;\alpha) = \frac{3}{2\alpha-1}\left\{\frac{(\alpha-1)\lambda}{(1+\lambda)^{\alpha-1}-1}\right\}^2. \tag{4.5.20}$$

By putting $\lambda = 1$ we recover the earlier expression derived for corrective factor on uniform meshes.

Case (ii): Element (x_e, x_{e+1}).

The derivative of the piecewise linear interpolant to s at the nodes is given by

$$(\Pi^h s)'(x) = \begin{cases} s[x_{e-1}, x_e], & x \in (x_{e-1}, x_e) \\ s[x_e, x_{e+1}], & x \in (x_e, x_{e+1}) \\ s[x_{e+1}, x_{e+2}], & x \in (x_{e+1}, x_{e+2}) \end{cases} \tag{4.5.21}$$

so that

$$(\tilde{\Pi}^h s)'(x) = s[x_e, x_{e+1}], \tag{4.5.22}$$

and

$$(\tilde{\Pi}^h s - \Pi^h s)'(x) = \begin{cases} (x_{e+1} - x_{e-1}) s[x_{e-1}, x_e, x_{e+1}], & x \in (x_{e-1}, x_e) \\ 0, & x \in (x_e, x_{e+1}) \\ (x_{e+2} - x_e) s[x_e, x_{e+1}, x_{e+2}], & x \in (x_{e+1}, x_{e+2}) \end{cases} . \tag{4.5.23}$$

The recovery operator $G[\Pi^h s - \tilde{\Pi}^h s]$ is linear on each element with nodal values at x_e and x_{e+1} being obtained from $\Pi^h s - \tilde{\Pi}^h s$ by interpolating the gradient values at the element midpoints (see Chapter 2 for the precise definition of G for piecewise linear approximation). This process yields

$$G[\Pi^h s - \tilde{\Pi}^h s](x_e) = (x_{e+1} - x_e) s[x_{e-1}, x_e, x_{e+1}] \tag{4.5.24}$$

and

$$G[\Pi^h s - \tilde{\Pi}^h s](x_{e+1}) = -(x_{e+1} - x_e) s[x_e, x_{e+1}, x_{e+2}]. \tag{4.5.25}$$

Carrying out the integrations we find that

$$
\begin{aligned}
&\| G[\Pi^h s - \tilde{\Pi}^h s] \|^2_{0,2,I^h_e} \\
&= \frac{1}{3}(x_{e+1} - x_e)^3 \Big\{ s^2[x_{e-1}, x_e, x_{e+1}] \\
&\qquad\qquad - s[x_{e-1}, x_e, x_{e+1}] s[x_e, x_{e+1}, x_{e+2}] \\
&\qquad\qquad\qquad + s^2[x_e, x_{e+1}, x_{e+2}] \Big\}
\end{aligned}
\tag{4.5.26}
$$

and

$$
\| s' - \Pi^h s' \|^2_{0,2,I^h_e} = \alpha^2 (x^{2\alpha-1}_{e+1} - x^{2\alpha-1}_e)/(2\alpha - 1) - (x_{e+1} - x_e) s^2[x_e, x_{e+1}]. \tag{4.5.27}
$$

Once again the apparent dependence of the correction factor on the absolute mesh spacing can be removed by introducing the quantity

$$
X_e = \frac{x_e - x_0}{x_{e+1} - x_e} \tag{4.5.28}
$$

and the ratio of the sizes of the adjoining two elements given by

$$
\mu = \frac{x_e - x_{e-1}}{x_{e+1} - x_e} \tag{4.5.29}
$$

and

$$
\lambda = \frac{x_{e+2} - x_{e+1}}{x_{e+1} - x_e}. \tag{4.5.30}
$$

The correction factor may then be rewritten as

$$A^2(e; e+1; \alpha)$$
$$= \frac{3}{2\alpha - 1}\Big\{\alpha^2[(X_e+1)^{2\alpha-1} - X_e^{2\alpha-1}] - (2\alpha-1)s^2[X_e, X_e+1]\Big\}$$
$$\div \Big\{s^2[X_e, X_e+1, X_e+\lambda+1] +$$
$$- s[X_e, X_e+1, X_e+\lambda+1]s[X_e-\mu, X_e, X_e+1]$$
$$+ s^2[X_e-\mu, X_e, X_e+1]\Big\}. \qquad (4.5.31)$$

This expression again reduces to the expression for uniform meshes if we put $\mu = \lambda = 1$. We have given the expression for a general element rather than the first two elements because in the case of non-uniform meshes it may be necessary to correct on more than just the first two elements. Later, we present numerical evidence for the validity of the corrective constants which we have obtained for non-uniform meshes.

REMARK The expression for the correction factors has been proven for the case of piecewise linear finite element approximation only. However, we conjecture that the same expression represents the correction constant for more general types of approximation scheme. ∎

4.5.2 More general types of singularity.

The analysis has so far been carried out for a model problem possessing one singularity of the form

$$u(x) = kx^\alpha + w(x). \qquad (4.5.32)$$

We now consider what modifications should be made if the true solution has the form

$$u(x) = \sum_{i=1}^{p} s_i(x) + w(x) \tag{4.5.33}$$

where $w \in H^3(I)$ and

$$s_i(x) = k_i x^{\alpha_i} \tag{4.5.34}$$

with

$$\frac{1}{2} < \alpha_1 < \alpha_2 < \cdots < \alpha_p < 2. \tag{4.5.35}$$

The estimator on element (x_e, x_{e+1}) is

$$\begin{aligned}
\eta_e &= {}_a \| \, G[u_h] - u_h' \, \|_{0,2,I_e^h} \\
&= {}_a \| \, G[w_h] - w_h' + \sum_i \left(G[s_{h,i}] - s_{h,i}' \right) \|_{0,2,I_e^h} \\
&\leq {}_a \| \, G[w_h] - w_h' \, \|_{0,2,I_e^h} + \sum_i {}_a \| \, G[s_{h,i}] - s_{h,i}' \, \|_{0,2,I_e^h} \, .
\end{aligned} \tag{4.5.36}$$

Manipulations similar to those used to obtain Lemma 4.4.5 and (4.4.54) show that the dominant term is that corresponding to the first component in the expression for the singularity

$$\eta_e \approx {}_a \| \, G[s_{h,1}] - s_{h,1}' \, \|_{0,2,I_e^h} \quad \text{as } h \to 0. \tag{4.5.37}$$

Moreover, the abstract estimator for the more general type of singularity is given by

$$
\begin{aligned}
\tilde{\eta}_e &= {}_a\| \sum_i s_i' + G[w_h] - u_h' \|_{0,2,I_e^h} \\
&= {}_a\| G[w_h] - w_h' + \sum_i (s_i' - s_{h,i}') \|_{0,2,I_e^h} \\
&\leq {}_a\| G[w_h] - w_h' \|_{0,2,I_e^h} + \sum_i {}_a\| s_i' - s_{h,i}' \|_{0,2,I_e^h} \, .
\end{aligned} \tag{4.5.38}
$$

Equally well, manipulations similar to those used to derive Lemma 4.4.5 and (4.4.60) show that the dominant term is again the one corresponding to the first component of the singularity

$$
\tilde{\eta}_e \approx {}_a\| s_1' - s_{h,1}' \|_{0,2,I_e^h} \ \text{as } h \to 0. \tag{4.5.39}
$$

Combining (4.5.38) and (4.5.40), we have that

$$
\begin{aligned}
&\lim_{h\to 0} \frac{\tilde{\eta}_e}{\eta_e} \\
&= \lim_{h\to 0} \frac{{}_a\| s_1' - s_{h,1}' \|_{0,2,I_e^h}}{{}_a\| G[s_{h,1}] - s_{h,1}' \|_{0,2,I_e^h}} \\
&= \lim_{h\to 0} \frac{\| s_1' - s_{h,1}' \|_{0,2,I_e^h}}{\| G[s_{h,1}] - s_{h,1}' \|_{0,2,I_e^h}} \\
&= \lim_{h\to 0} \frac{\| s_1' - (\Pi s_1)' \|_{0,2,I_e^h}}{\| G[s_{h,1}] - s_{h,1}' \|_{0,2,I_e^h}}, \ \text{using (4.4.56)} \\
&= \lim_{h\to 0} \frac{\| s_1' - (\Pi s_1)' \|_{0,2,I_e^h}}{\| G[\Pi s_1] - (\Pi s_1)' \|_{0,2,I_e^h}}, \ \text{following (4.4.50)–(4.4.51)} \\
&= A(e; e+1; \alpha_1).
\end{aligned} \tag{4.5.40}
$$

That is,

$$\lim_{h\to 0} \frac{\tilde{\eta}_e}{\eta_e} = A(e; e+1; \alpha_1) \tag{4.5.41}$$

so that for the more general singularity of (4.5.34), we employ the same correction factor as if there were only one term in the singularity with exponent α_1 (as we might expect on purely intuitive grounds).

Multiple Singularities.

Suppose now that the true solution has multiple singularities so that

$$u(x) = \sum_{j=1}^{q} s_j(x) + w(x) \tag{4.5.42}$$

where $w \in H^3$ and each s_j is of the form

$$s_j(x) = \sum_{i=1}^{p} k_{ij}(x - \beta_j)^{\alpha_i}. \tag{4.5.43}$$

We are concerned with obtaining the correction factors as the mesh becomes increasingly refined and so without loss of generality we may assume that the refinement has proceeded sufficiently to ensure that each element contains at most one singular point β_j. Further, we shall assume that the nodes of the mesh are chosen to coincide with the singular points. In practice this is possible since the singular points manifest themselves by being located at the vertices of the domain or through the form of the data (in $\Re^1$). In addition, we would normally try to place the nodes in this way since an improved rate of convergence is obtained in this way.

We have commented previously on the *cut-off* nature of the singularity and the need to only correct on elements near to the singularity. We can see therefore that it will *be sufficient to apply the correction factor corresponding to the singularity* *which is nearest to the element with which we are concerned.* We shall present numerical evidence to support this conclusion in Chapter 5.

α	$i = 0$	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$i = 6$	$i = 7$	$i = 8$	$i = 9$
0.60	6.397879	0.321502	0.793971	0.893432	0.935025	0.956311	0.968636	0.976401	0.981606	0.985262
0.70	4.376002	0.380907	0.815959	0.905720	0.942759	0.961596	0.972466	0.979300	0.983874	0.987084
0.80	3.454736	0.441557	0.837419	0.917469	0.950088	0.966581	0.976068	0.982021	0.986001	0.988792
0.90	2.891710	0.503059	0.858199	0.928625	0.956989	0.971255	0.979437	0.984563	0.987985	0.990384
1.10	2.202957	0.626308	0.897098	0.948961	0.969432	0.979634	0.985457	0.989095	0.991519	0.993215
1.20	1.968890	0.686488	0.914900	0.958047	0.974936	0.983322	0.988099	0.991080	0.993065	0.994453
1.30	1.777210	0.744406	0.931397	0.966353	0.979940	0.986666	0.990491	0.992876	0.994462	0.995572
1.40	1.616230	0.798947	0.946441	0.973840	0.984430	0.989660	0.992629	0.994479	0.995710	0.996570
1.50	1.478400	0.848959	0.959898	0.980474	0.988393	0.992296	0.994511	0.995889	0.996806	0.997447
1.60	1.358590	0.893350	0.971651	0.986224	0.991818	0.994572	0.996133	0.997105	0.997751	0.998202
1.70	1.253190	0.931178	0.981600	0.991067	0.994697	0.996482	0.997495	0.998124	0.998543	0.998835
1.80	1.159540	0.961756	0.989671	0.994985	0.997023	0.998025	0.998593	0.998947	0.999182	0.999346
1.90	1.075660	0.984708	0.995812	0.997964	0.998791	0.999198	0.999429	0.999572	0.999668	0.999734

Table 4.1 Values of $A(\alpha, i)$.

N	$\|e\|_E$	ϵ	θ	$\hat{\epsilon}$	$\hat{\theta}$
(a) $\alpha = 1.5$					
25	2.96751(−1)	3.00018(−1)	1.01101	2.99846(−1)	1.01043
50	1.48456(−1)	1.48824(−1)	1.00248	1.48810(−1)	1.00239
100	7.42557(−2)	7.42762(−2)	1.00028	7.42962(−2)	1.00054
200	3.71398(−2)	3.71292(−2)	0.99972	3.71447(−2)	1.00013
(b) $\alpha = 0.9$					
25	3.04899(−1)	3.06832(−1)	1.00634	3.16059(−1)	1.03660
50	1.53468(−1)	1.52400(−1)	0.99305	1.55341(−1)	1.01220
100	7.79133(−2)	7.63137(−2)	0.97947	7.83186(−2)	1.00520
200	4.02865(−2)	3.84499(−2)	0.95441	4.03961(−2)	1.00272
(c) $\alpha = 0.6$					
25	7.23721(−1)	3.49242(−1)	0.48257	8.35880(−1)	1.15498
50	6.26389(−1)	2.05235(−1)	0.32765	6.47667(−1)	1.03397
100	5.70774(−1)	1.44878(−1)	0.25383	5.74711(−1)	1.00690
200	5.28850(−1)	1.19893(−1)	0.22671	5.29626(−1)	1.00147

Table 4.2 Classical and corrected estimators for one singular term.

$s(x) = x^{\alpha}$; $a(x) \equiv 1$; $b(x) \equiv 10$

N	$\|e\|_E$	ϵ	θ	$\hat{\epsilon}$	$\hat{\theta}$
(a) $\alpha = 1.5$					
25	2.62445(−1)	2.65260(−1)	1.01073	2.65168(−1)	1.01038
50	1.31327(−1)	1.31671(−1)	1.00262	1.31665(−1)	1.00257
100	6.56831(−2)	6.57172(−2)	1.00052	6.57243(−2)	1.00063
200	3.28470(−2)	3.28465(−2)	0.99999	3.28521(−2)	1.00016
(b) $\alpha = 0.9$					
25	2.64861(−1)	2.67125(−1)	1.00855	2.70659(−1)	1.02189
50	1.32879(−1)	1.32686(−1)	0.99854	1.33783(−1)	1.00680
100	6.68783(−2)	6.63199(−2)	0.99165	6.70606(−2)	1.00273
200	3.39161(−2)	3.32595(−2)	0.98064	3.39835(−2)	1.00199
(c) $\alpha = 0.6$					
25	4.56962(−1)	2.85139(−1)	0.62399	5.16760(−1)	1.13086
50	3.67726(−1)	1.54839(−1)	0.42107	3.79054(−1)	1.03081
100	3.25476(−1)	9.66298(−2)	0.29689	3.27471(−1)	1.00613
200	2.98722(−1)	7.23825(−2)	0.24231	2.99053(−1)	1.00111

Table 4.3 Classical and corrected estimators for one singular term.
$s(x) = x^{\alpha}$; $a(x) = \sqrt{x + \frac{1}{10}}$; $b(x) \equiv 0$.

N	$\|e\|_E$	ϵ	θ	$\hat{\epsilon}$	$\hat{\theta}$
(a) $\alpha = 1.5$					
25	2.08215(−1)	2.10680(−1)	1.01184	2.10670(−1)	1.01179
50	1.04175(−1)	1.04491(−1)	1.00304	1.04491(−1)	1.00303
100	5.20962(−2)	5.21358(−2)	1.00076	5.21360(−2)	1.00076
200	2.60493(−2)	2.60541(−2)	1.00018	2.60543(−2)	1.00019
(b) $\alpha = 0.9$					
25	2.04348(−1)	2.06773(−1)	1.01187	2.06956(−1)	1.01276
50	1.02256(−1)	1.02547(−1)	1.00285	1.02597(−1)	1.00334
100	5.11538(−2)	5.11687(−2)	1.00029	5.12002(−2)	1.00091
200	2.55982(−2)	2.55754(−2)	0.99911	2.56058(−2)	1.00030
(c) $\alpha = 0.6$					
25	2.15103(−1)	2.07536(−1)	0.96482	2.23498(−1)	1.03903
50	1.19304(−1)	1.03565(−1)	0.86808	1.21002(−1)	1.01423
100	7.65154(−2)	5.27293(−2)	0.68913	7.68108(−2)	1.00386
200	5.87748(−2)	2.80726(−2)	0.47763	5.87921(−2)	1.00029

Table 4.4 Classical and corrected estimators for one singular term.
$s(x) = x^{\alpha}$; $a(x) = \left(x + \frac{1}{10}\right)^2$; $b(x) \equiv 0$.

CHAPTER 5

Regularity: Sufficient but unnecessary?

5.1 Introduction.

In our development we have made rather strong assumptions regarding the regularity of the mesh and of the true solution. These were *sufficient* for the estimators to be asymptotically exact. In this Chapter we shall consider to what extent these assumptions are *necessary* for the satisfactory performance of the estimators in more practical circumstances.

In practice, the regularity of the meshes fail to meet our assumptions. This is most acute in the case of adaptively designed meshes in which the sequence of partitions may even violate the assumption $h \to 0$. Adaptive algorithms are rapidly becoming the standard method of analyzing problems with smooth solutions as well as those with singularities.

In Chapter 4 we have shown theoretically and numerically that the classical error estimators fail to perform satisfactorily for problems with non-smooth solutions even if the mesh is very regular. The use of adaptively designed meshes could conceivably have one of two possible effects:

- The resulting non-uniformity of the meshes could cause the error estimators to behave in a completely unsatisfactory manner. Indeed we may well expect this to be the case since the superconvergence results fail to hold unless the mesh satisfies the regularity requirements.
- Conversely, the adaptively designed meshes take account of the specific nature

of the true solution and we may therefore find that the estimators' performance is actually enhanced due to the improved mesh design.

The second point might also be expanded upon by arguing that in an adaptive algorithm the singular elements (by which we mean those on which a corrective factor would be applied) are the ones which are most refined. Therefore their combined measure tends to zero and the effect of the singular elements on the global effectivity index becomes smaller. Consequently it is tempting to suggest that corrective constants are a theoretical nicety which can be dispensed with in practice.

In this Chapter we shall attempt to resolve these questions. We shall also take this opportunity to make further comparisons between the relative merits of averaging based estimation (Chapter 2) and projection based estimation (Chapter 3) and to test the corrective constants for non-uniform meshes (Chapter 4).

5.2 Estimation on adaptively designed meshes.

The adaptive algorithm we shall use is similar to that proposed by Rivara [68]. This uses a refinement method for triangular elements which delivers a sequence of meshes in which the minimal angle of any element is bounded away from zero [66], [67] thus ensuring that the elements do not become degenerate.

The system of equations arising from the finite element discretization will be solved using a multigrid method. The sequence of meshes obtained using Rivara's algorithm is nested which means that the restriction and prolongation operators needed by the multigrid solver to transfer between grids are easy to define (see *e.g.* Hackbusch [39] for further details and references). We shall use both V- and W-cycle versions of the multigrid method based on the multigrid algorithm of Bank and Dupont [21]. The smoothing iteration is taken to be Gauss Seidel, for which we carry out only pre-smoothing steps (since these are more efficient than carrying out post-smoothing steps, see [39]).

The choice of multigrid as a method of solution is a natural one to make with an adaptive algorithm since we always have available (with the exception of the initial mesh!) a good choice of starting iterate, namely the solution on the next coarser mesh in the hierarchy. With this scheme we essentially have a nested iteration version of the multigrid method which is known [21] to give the solution to an accuracy equal to that of the discretization error in $O(n)$ operations (which is of course optimal order for a system containing n degrees of freedom).

This work estimate can be shown to also hold for problems with singular solutions provided the sequence of meshes is designed *properly*, by which we roughly mean that there should be refinement around singularities, but not over-refinement (see Yserentant [81] for a more precise discussion). This provides a useful means of assessing whether the refinement process guided by the error estimators is effective. This being the case we should observe that the *contraction factor* γ (which gives the amount by which the error is reduced by each multigrid cycle) is independent of the number of levels of refinement and satisfies

$$\gamma \leq r < 1 \tag{5.2.1}$$

for some constant r. Furthermore, we should find that the number of multigrid cycles needed to solve on each mesh should be roughly the same for all meshes. Yet another way in which to assess the efficiency of the refinement is that the optimal rate of convergence should be attained in spite of the presence of singularities which would lead to a degradation in the rate of convergence should uniform refinement be performed. For piecewise linear approximation we expect $O(n^{-\frac{1}{2}})$ in two dimensions.

The decision as to which elements in any given mesh should be refined is based on an averaging estimator of the type discussed in Chapter 2. Estimates of the error ϵ_e in each element e are computed and refinement is performed for any elements e satisfying

$$\epsilon_e \geq \mu \bar{\epsilon}_e \tag{5.2.2}$$

where

$$\bar{\epsilon}_e = \max_e \epsilon_e \tag{5.2.3}$$

and $\mu \in (0, 1)$. Typically we choose $\mu \approx 0.3$. In addition to presenting the results of the averaging based error estimator we give the true error on each mesh and the result obtained by using a projection based error estimator of the type discussed in Chapter 3. We give the true accuracy of the solution

$$\eta = \frac{\| e \|_E}{\| u \|_E} \times 100\% \tag{5.2.4}$$

and also the estimated accuracy obtained using averaging and projection

$$\eta_{aver} = \frac{\epsilon_{aver}}{E_{aver}} \times 100\% \tag{5.2.5}$$

where

$$E_{aver}^2 = \epsilon_{aver}^2 + \| u_h \|_E^2 \tag{5.2.6}$$

is the approximation to the energy of the true solution. It is vital to examine the estimated accuracy since this is the quantity which is used as a stopping criterion within the adaptive algorithm.

5.2.1 Example 1. L–shaped domain.

The classic examples used to test adaptive algorithms are the L-shaped domain and the slit domain. Both of these possess re-entrant corners and the derivative is singular at these points. The first term in the expansion for the singular part of the solution to the L-shaped domain is given by

$$u(r,\theta) = r^{2/3}\sin 2\theta/3. \tag{5.2.7}$$

The solution domain is shown in Fig. 5.1 and we arrange the boundary conditions so that the exact solution $u(r,\theta)$ is given by (5.2.7). The sequence of meshes obtained are shown in Fig. 5.2(a)–(h). The multigrid solver was used in V-cycle mode with 3 pre-smoothing steps. The results are shown in Table 5.1 where it can be seen that 9 meshes were used to obtain an accuracy of 3.75 percent (which is typical of the type of accuracy demanded in engineering applications).

It is seen that the multigrid solver takes roughly 4 cycles to converge with a contraction factor of about 1/4. The rate of convergence is $O(n^{-\frac{1}{2}})$ which is optimal for piecewise linear approximation on triangles. The error estimators become increasingly accurate as the refinement proceeds with averaging performing rather better than projection. The averaging method also delivers better approximations to both the accuracy and the energy of the true solution, and as such appears to be superior to projection in all respects.

5.2.2 Example 2. Slit domain.

The slit domain (or prototype of a cracked panel) has a singularity around the tip of the crack. The geometry is shown in Fig. 5.3 and the singular part of the solution may be expanded in terms of singular functions of which the first two are given by

$$s_1(r,\theta) = r^{1/4} \sin \theta/4 \quad (5.2.8)$$

$$s_2(r,\theta) = r^{3/4} \sin 3\theta/4. \quad (5.2.9)$$

We arrange the boundary conditions so that the true solution is given by (5.2.9). The sequence of meshes generated are shown in Fig. 5.4(a)–(h). This time the multigrid solver was employed in W-cycle mode with 2 pre-smoothing steps. The numerical results are shown in Table 5.2 where it is seen that 9 meshes were needed to give an overall accuracy of 2.8 percent. The results indicate the success of the refinement process in terms of the characteristics of the multigrid solver and the optimal rate of convergence. The error estimators perform satisfactorily with averaging once again proving superior to projection based estimation.

5.3 Corrected Estimator on Non-Uniform Meshes.

The two dimensional results presented in the previous section leave a certain amount of doubt as to the question of whether corrected estimators are really necessary if the mesh is properly designed. In a sense this is because the singular behaviour is obscured by the presence of other factors. If we continued the refinement process sufficiently then we would be able to assess the true asymptotic behaviour.

In this section we return to one dimension where we can solve on sufficiently graded meshes as to be sure of the asymptotic behaviour. This provides an opportunity to observe the performance of the correction scheme proposed in Chapter 4 for the case several singular terms on non-uniform meshes.

The model problem will be the same as in Chapter 4, which we shall approximate using piecewise linear functions. Results are obtained on three types of mesh.

1. Uniform. This is the case dealt with in Chapter 4 but is included so that the effects of having several singular terms in the true error can be assessed in the light of the numerical results in Chapter 4 obtained with one singular term on uniform meshes.

2. Geometric. This is the prototype for adaptively refined meshes in which the element nearest the singularity is progressively subdivided into two elements. The mesh points on a mesh containing n elements are given by

$$\frac{x_{r+1} - x_r}{x_r - x_{r-1}} = \frac{1}{2}. \tag{5.3.1}$$

   with $x_0 = 0$ and $x_n = 1$.

3. Fibonacci. The previous cases do not test the corrective constants on non-uniform meshes to their full since the ratio of the sizes of adjacent elements is constant for both uniform and geometric meshes. In this case we make the ratio between the sizes of successive elements equal to the ratio of two successive members of the Fibonacci sequence given by

$$t_{r+1} = t_r + t_{r-1}, \quad t_1 = 2, \quad t_2 = 3 \tag{5.3.2}$$

   so that

$$\frac{x_{r+1} - x_r}{x_r - x_{r-1}} = \frac{t_{r+1}}{t_r} \tag{5.3.3}$$

   with $x_0 = 0$ and $x_n = 1$.

   Results are presented for each of the cases

(i) $a(x) = 1$ and $b(x) = 10$

(ii) $a(x) = (1.1 - x)^2$ and $b(x) = 0$.

and for when the true solution has the form

$$u(x) = s(x) - xs(1) - (1 - x)s(0) \tag{5.3.4}$$

where $s(x)$ takes the forms

1. $x^{3/5}$
2. $x^{3/5} + x^{8/5}$
3. $x^{3/5} + x^{8/5} + (1 - x)^{9/10}$.

Tables 5.3–9 show the results obtained for each combination. The notation used is

$$\begin{aligned} \| e \|_E &- \text{ the actual error} \\ \epsilon &- \text{ uncorrected estimator} \\ \hat{\epsilon} &- \text{ corrected estimator} \\ \theta &= \epsilon / \| e \|_E \\ \hat{\theta} &= \hat{\epsilon} / \| e \|_E . \end{aligned}$$

The results show that the proposed correction method for several singular terms performs well with the effectivity index $\hat{\theta}$ converging to unity in all cases. Meanwhile, even on geometrically refined meshes the effectivity index θ of the classical estimator does not tend towards unity.

5.4 Final Remarks.

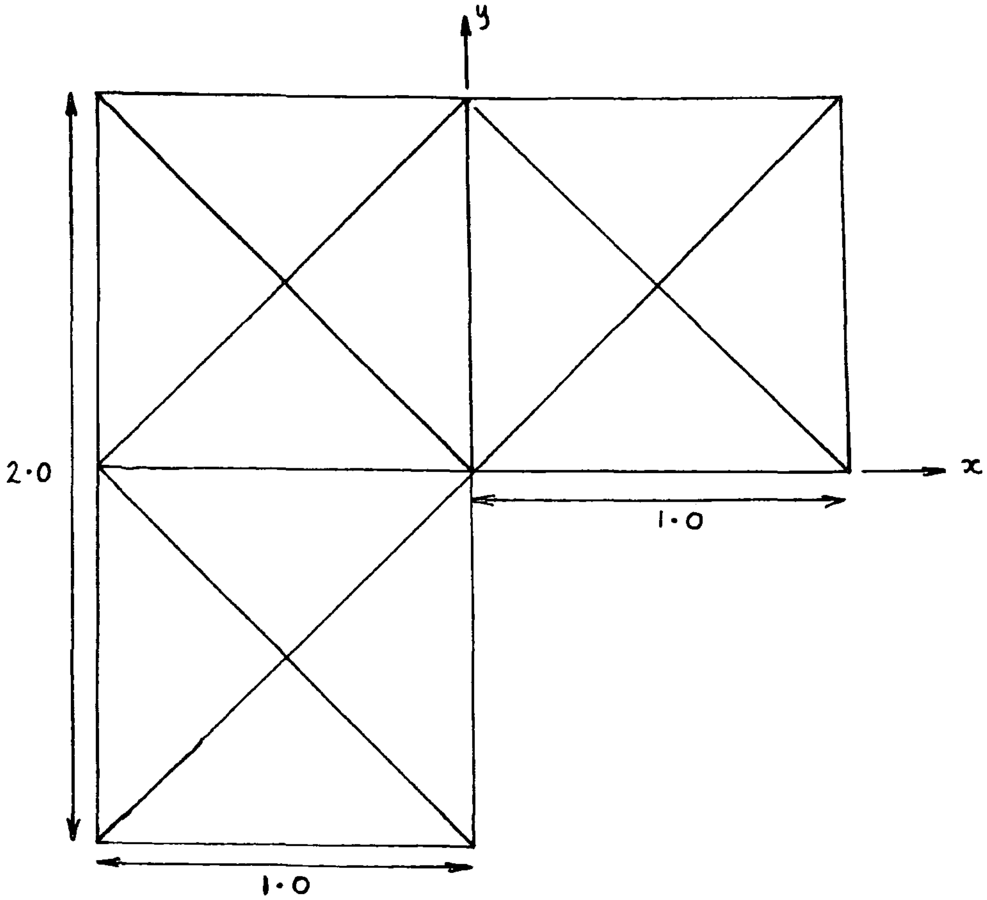
In this Chapter we have considered to what extent the assumptions made in the earlier theoretical work were necessary for the estimators to perform reliably. In particular the very important case of adaptively refined meshes was singled out for closer attention. An intuitive argument suggested that the use of corrective factors was unnecessary if an adaptively designed mesh was used. The argument revolved around the observation that the measure of the set of singular elements tends to zero and therefore should not have much global influence.

This argument strikes us as being somewhat akin to *sweeping the dust under the carpet*, in so far as it proposes ignoring the essential problem of obtaining estimators which are not purely attractive on a global level but which give a real estimate of the error at an element level. We carried out numerical examples in two dimensions for which the results are slightly inconclusive (but once again demonstrated the superiority of averaging over projection based estimation). Accordingly we returned to the one dimensional case. Here we could virtually guarantee that the results were free of peripheral disturbances. For example the use of a Kronrod Patterson routine [60] to find the true error to high accuracy removes one source of worry present in the higher dimensional case. These results show convincingly that the use of corrective factors is absolutely essential even with strongly graded meshes.

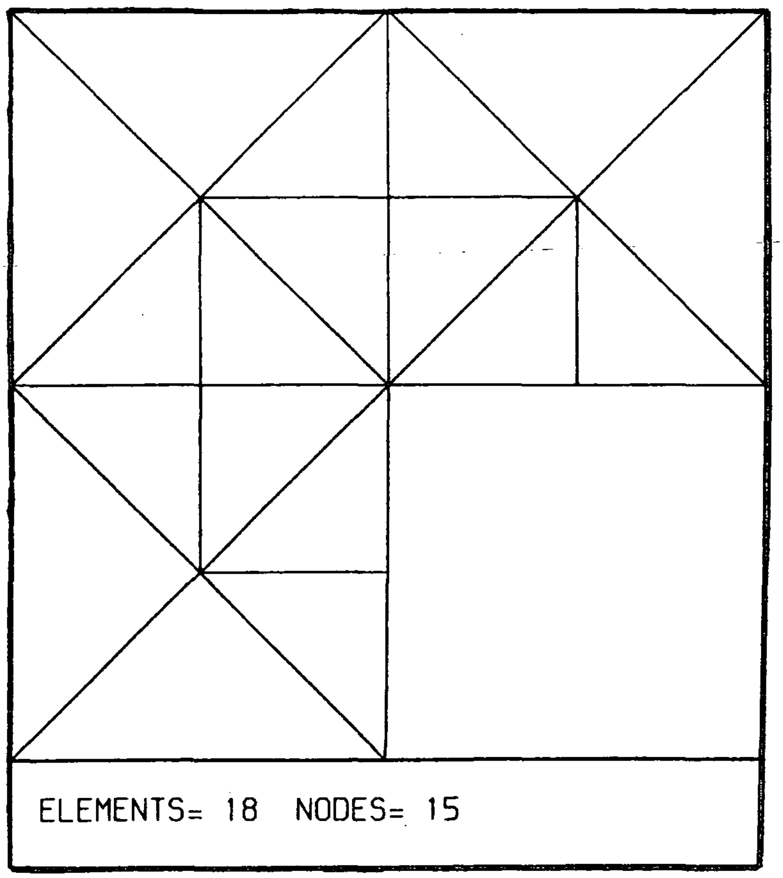
Lest the reader should still have doubts regarding the need for correction factors we might note that the current trend in adaptive remeshing is towards algorithms which attempt to generate the final mesh in at most 2-3 adaptive steps, *e.g.* [61], [62], [82], [87], [88], [90] and the numerical examples of Chapter 3. If we decide to forego the use of corrective constants in estimating the error on the first (and typically coarse) mesh then the information on which the first regeneration is carried out is inferior to that using corrective constants. Moreover, it is inferior in precisely those areas (*i.e.* near the singularities) where it needs to be most accurate making the need for multiple remeshings a real possibility. Further, the cost of performing the correction estimator is minimal when viewed with regard to the computational costs involved in a finite element code and as such seems to be a small price to pay for getting the best information available.

Mesh	Nodes Elements	MG-Cycles γ	$\|\|e\|\|_E$ $\|\|u\|\|_E$ η ($\times$100)	ϵ_{proj} E_{proj} η_{proj}($\times$100) θ_{proj}	ϵ_{aver} E_{aver} η_{aver}($\times$100) θ_{aver}
2	15	3	$0.328905E+00$	$0.281796E+00$	$0.315534E+00$
	18	*	$0.130807E+01$	$0.140848E+01$	$0.137561E+01$
			$0.251443E+02$	$0.200071E+02$	$0.229378E+02$
				$0.856770E+00$	$0.959348E+00$
3	28	4	$0.230542E+00$	$0.198859E+00$	$0.219101E+00$
	40	0.100	$0.133665E+01$	$0.138864E+01$	$0.136998E+01$
			$0.172478E+02$	$0.143204E+02$	$0.159930E+02$
				$0.862572E+00$	$0.950375E+00$
4	43	5	$0.172787E+00$	$0.152036E+00$	$0.166076E+00$
	66	0.165	$0.134782E+01$	$0.138037E+01$	$0.136726E+01$
			$0.128197E+02$	$0.110142E+02$	$0.121466E+02$
				$0.879905E+00$	$0.961159E+00$
5	66	4	$0.127466E+00$	$0.114982E+00$	$0.124515E+00$
	108	0.207	$0.135224E+01$	$0.136575E+01$	$0.135669E+01$
			$0.942631E+01$	$0.841895E+01$	$0.917787E+01$
				$0.902061E+00$	$0.976847E+00$
6	92	4	$0.103463E+00$	$0.942724E-01$	$0.100424E+00$
	158	0.214	$0.135396E+01$	$0.136387E+01$	$0.135699E+01$
			$0.764153E+01$	$0.691212E+01$	$0.740050E+01$
				$0.911166E+00$	$0.970622E+00$
7	136	4	$0.812562E-01$	$0.763114E-01$	$0.801423E-01$
	236	0.265	$0.135463E+01$	$0.136152E+01$	$0.135650E+01$
			$0.599840E+01$	$0.560486E+01$	$0.590804E+01$
				$0.939145E+00$	$0.986292E+00$
8	212	4	$0.631237E-01$	$0.597936E-01$	$0.623566E-01$
	378	0.294	$0.135490E+01$	$0.135882E+01$	$0.135551E+01$
			$0.465892E+01$	$0.440040E+01$	$0.460024E+01$
				$0.947245E+00$	$0.987848E+00$
9	306	3	$0.508396E-01$	$0.487071E-01$	$0.504513E-01$
	558	0.274	$0.135501E+01$	$0.135726E+01$	$0.135493E+01$
			$0.375198E+01$	$0.358863E+01$	$0.372354E+01$
				$0.958056E+00$	$0.992364E+00$

Table 5.1 Analysis of L shaped domain.

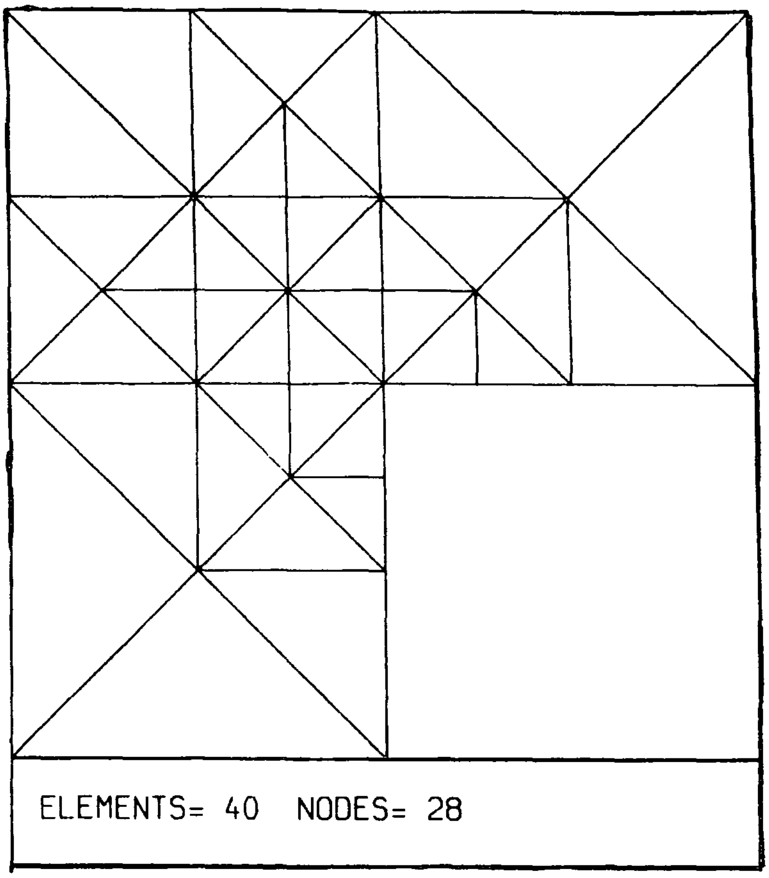


(a) Initial Mesh.

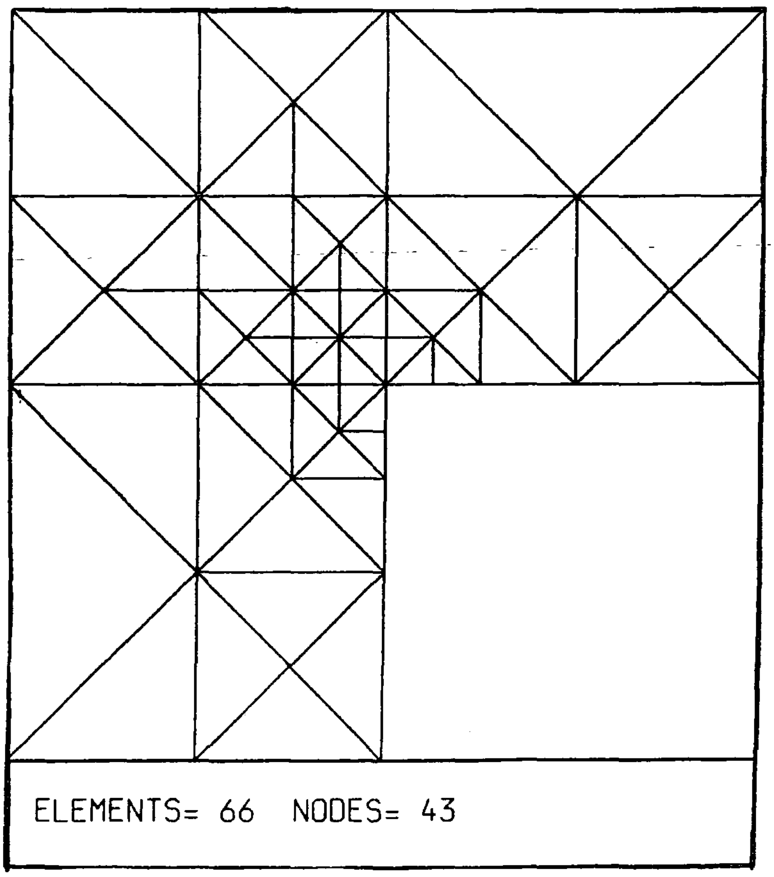


(b) Mesh 2. $\eta = 25.1\%$.

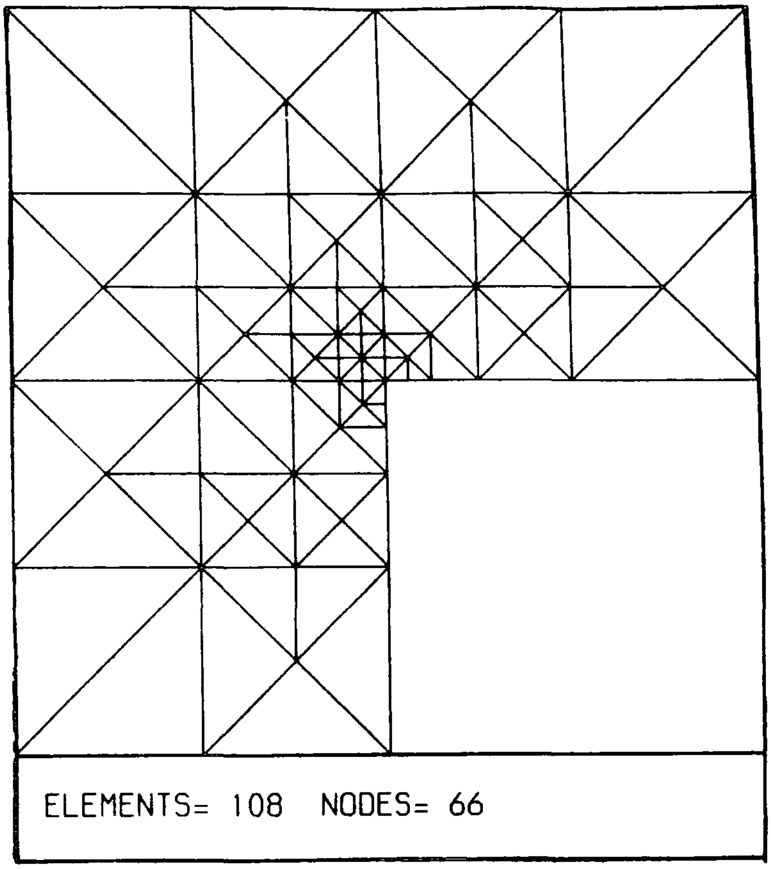
Figure 5.1 Sequence of adaptively designed meshes for L shaped domain.



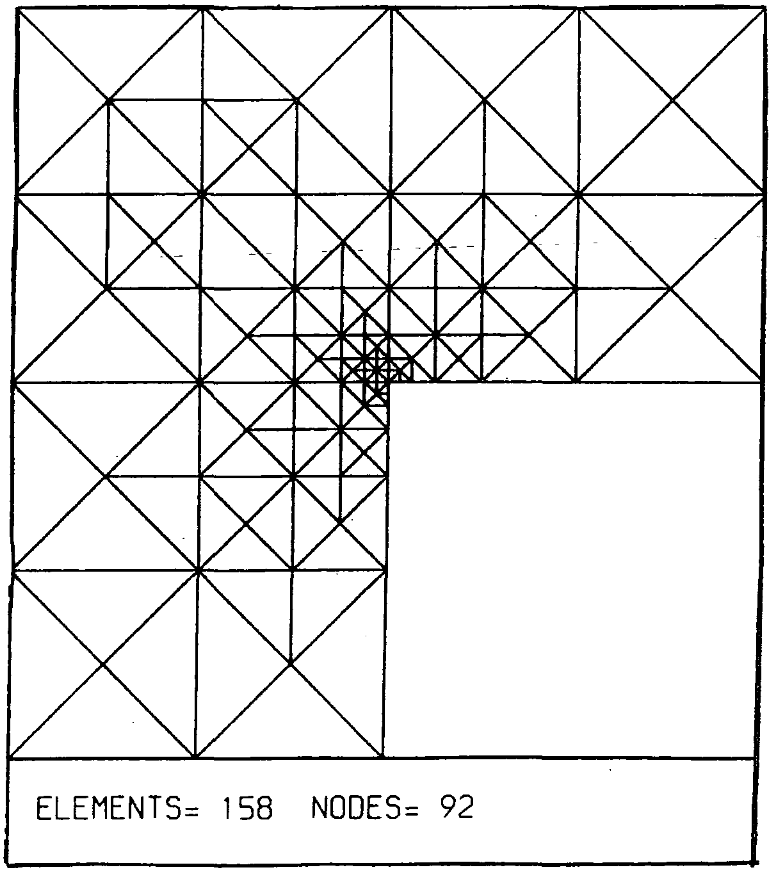
(c) Mesh 3. $\eta = 17.2\%$.



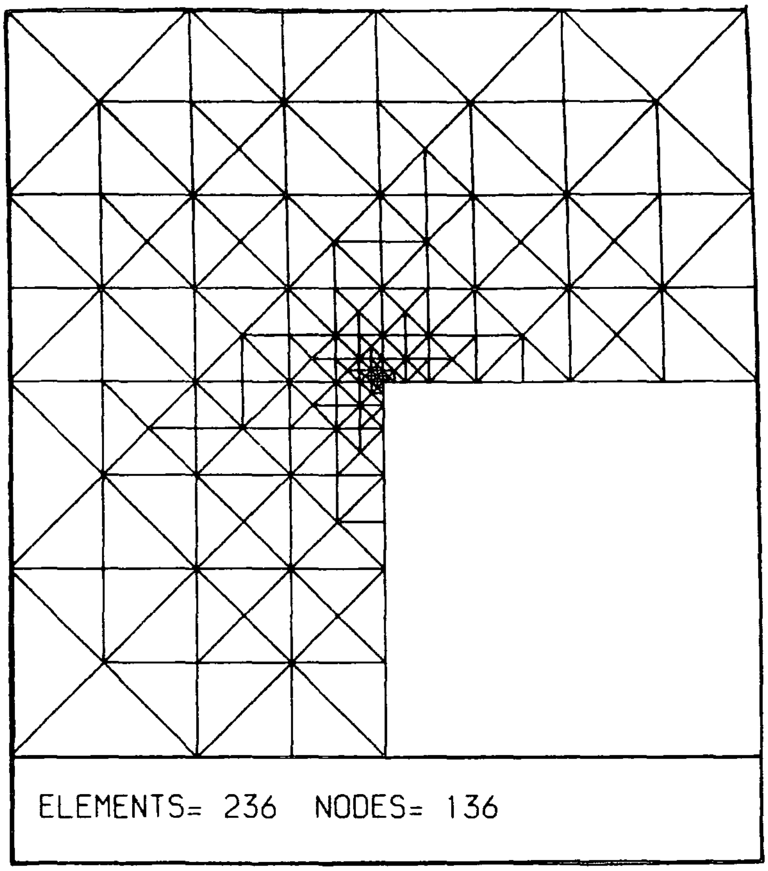
(d) Mesh 4. $\eta = 12.8\%$.



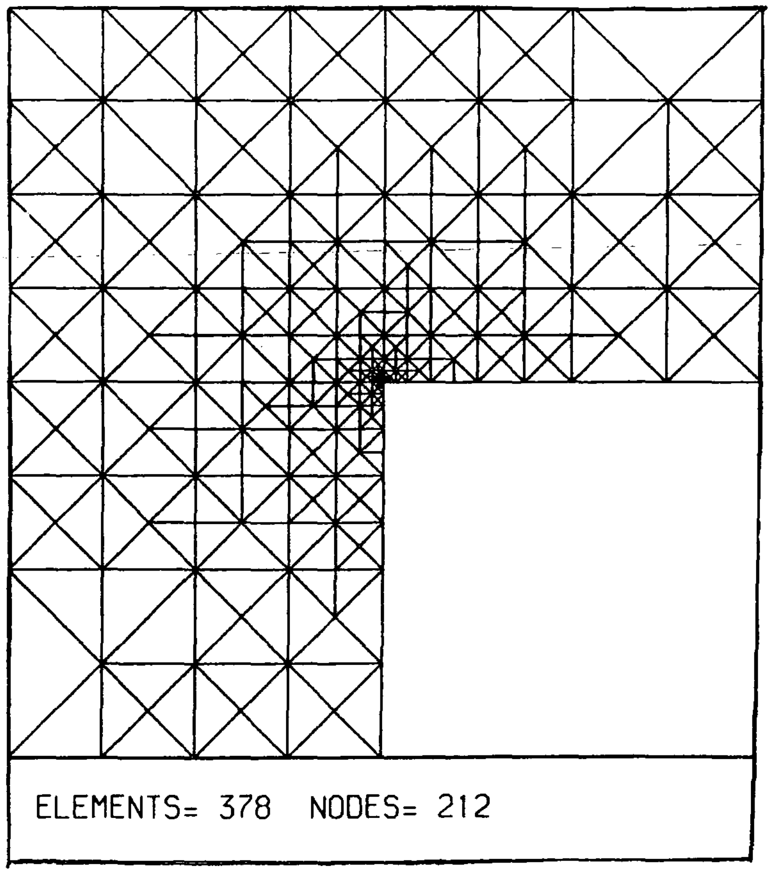
(e) Mesh 5. $\eta = 9.4\%$.



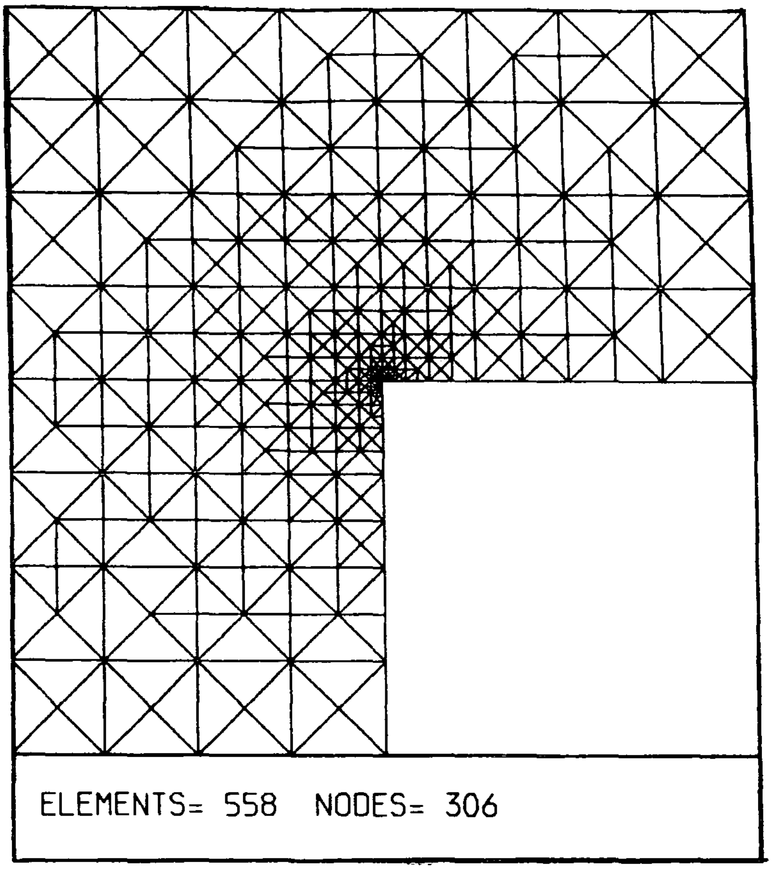
(f) Mesh 6. $\eta = 7.6\%$.



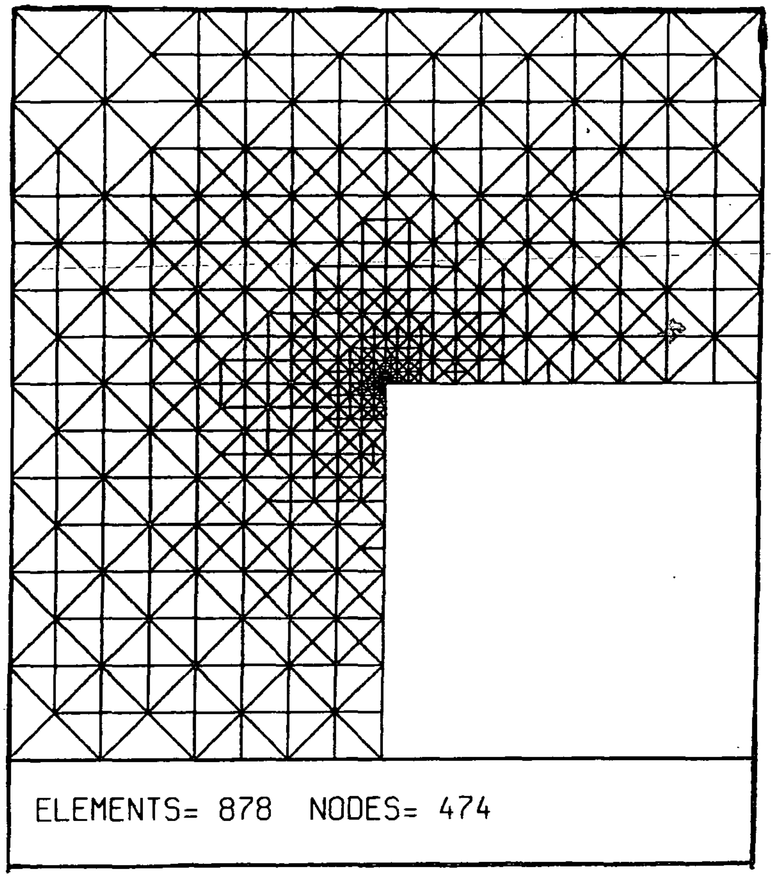
(g) Mesh 7. $\eta = 6.0\%$.



(h) Mesh 8. $\eta = 4.7\%$.



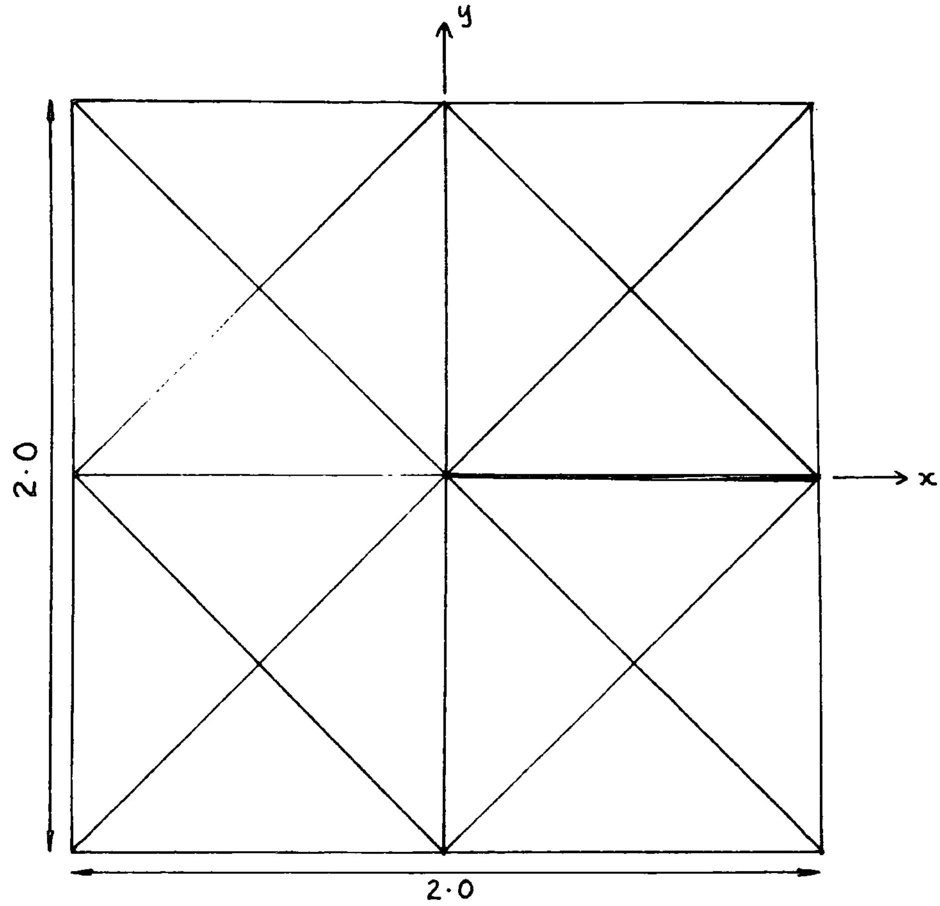
(i) Mesh 9. $\eta = 3.8\%$.



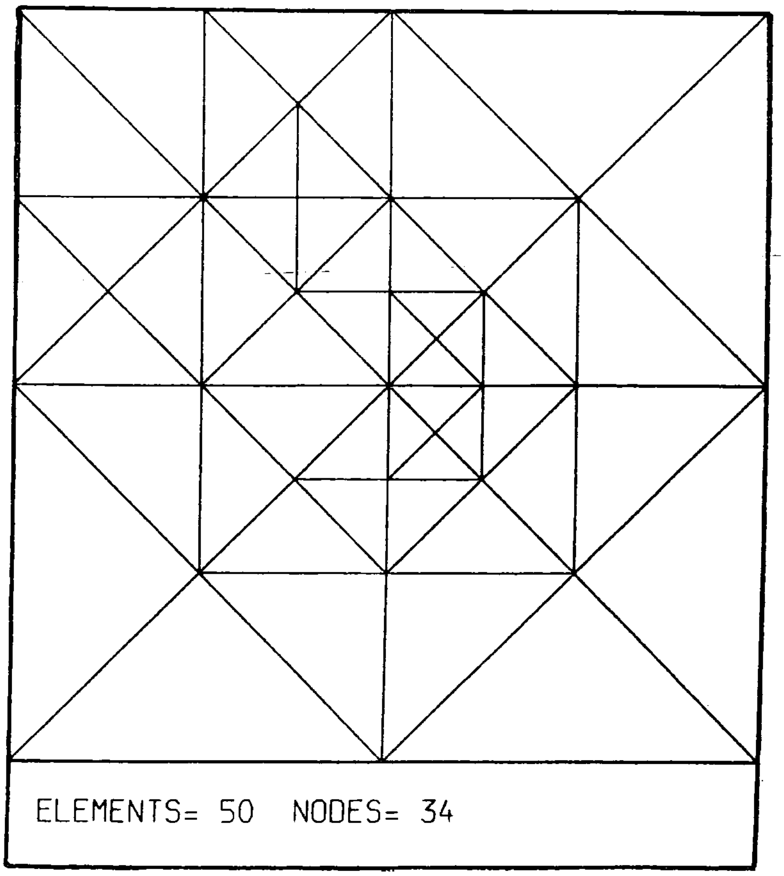
(j) Final predicted mesh.

Mesh	Nodes Elements	MG-Cycles γ	$\|\|e\|\|_E$ $\|\|u\|\|_E$ η ($\times 100$)	ϵ_{proj} E_{proj} $\eta_{proj}(\times 100)$ θ_{proj}	ϵ_{aver} E_{aver} $\eta_{aver}(\times 100)$ θ_{aver}
3	34 50	4 0.186	$0.217530E+00$ $0.166339E+01$ $0.130775E+02$	$0.167802E+00$ $0.170501E+01$ $0.984170E+01$ $0.771398E+00$	$0.182553E+00$ $0.169162E+01$ $0.107916E+02$ $0.839210E+00$
4	47 73	5 0.211	$0.167213E+00$ $0.167202E+01$ $0.100007E+02$	$0.137283E+00$ $0.170130E+01$ $0.806932E+01$ $0.821007E+00$	$0.147813E+00$ $0.169088E+01$ $0.874177E+01$ $0.883980E+00$
5	66 104	4 0.209	$0.128856E+00$ $0.167393E+01$ $0.769780E+01$	$0.113101E+00$ $0.168603E+01$ $0.670813E+01$ $0.877734E+00$	$0.122389E+00$ $0.167819E+01$ $0.729293E+01$ $0.949815E+00$
6	108 186	4 0.255	$0.974793E-01$ $0.167573E+01$ $0.581714E+01$	$0.863249E-01$ $0.168441E+01$ $0.512492E+01$ $0.885571E+00$	$0.918675E-01$ $0.167936E+01$ $0.547038E+01$ $0.942430E+00$
7	171 299	4 0.277	$0.806971E-01$ $0.167630E+01$ $0.481401E+01$	$0.720297E-01$ $0.168172E+01$ $0.428309E+01$ $0.892594E+00$	$0.758638E-01$ $0.167825E+01$ $0.452041E+01$ $0.940106E+00$
8	213 377	3 0.0697	$0.682281E-01$ $0.167652E+01$ $0.406963E+01$	$0.628254E-01$ $0.168105E+01$ $0.373728E+01$ $0.920814E+00$	$0.655650E-01$ $0.167811E+01$ $0.390707E+01$ $0.960968E+00$
9	424 778	3 0.129	$0.464466E-01$ $0.167659E+01$ $0.277030E+01$	$0.437755E-01$ $0.167846E+01$ $0.260808E+01$ $0.942491E+00$	$0.451733E-01$ $0.167700E+01$ $0.269370E+01$ $0.972586E+00$

Table 5.2 Analysis of slit domain.

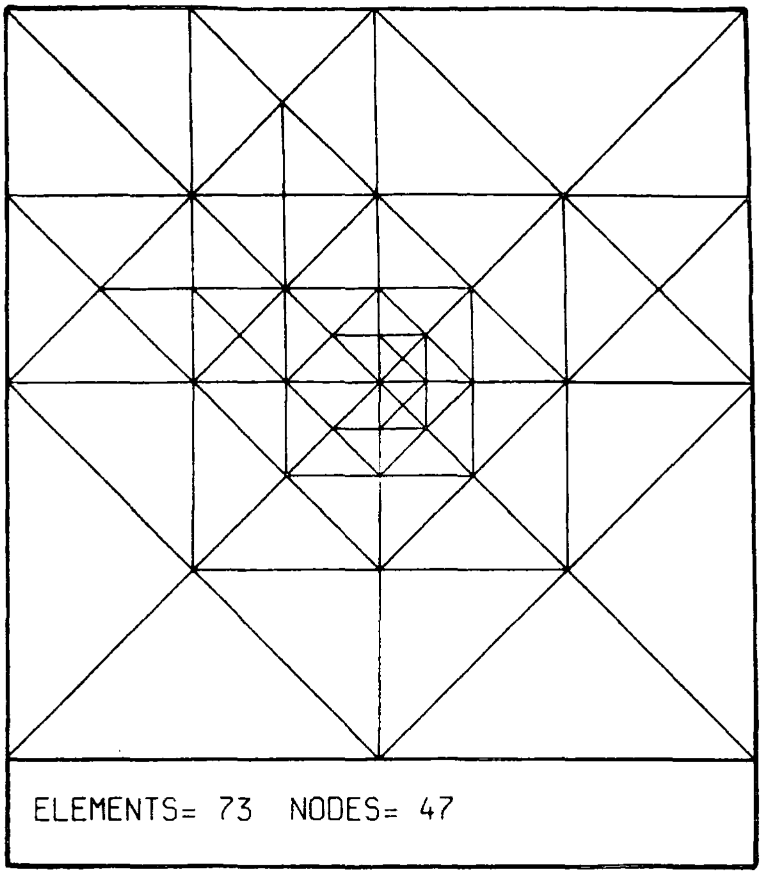


(a) Initial Mesh.

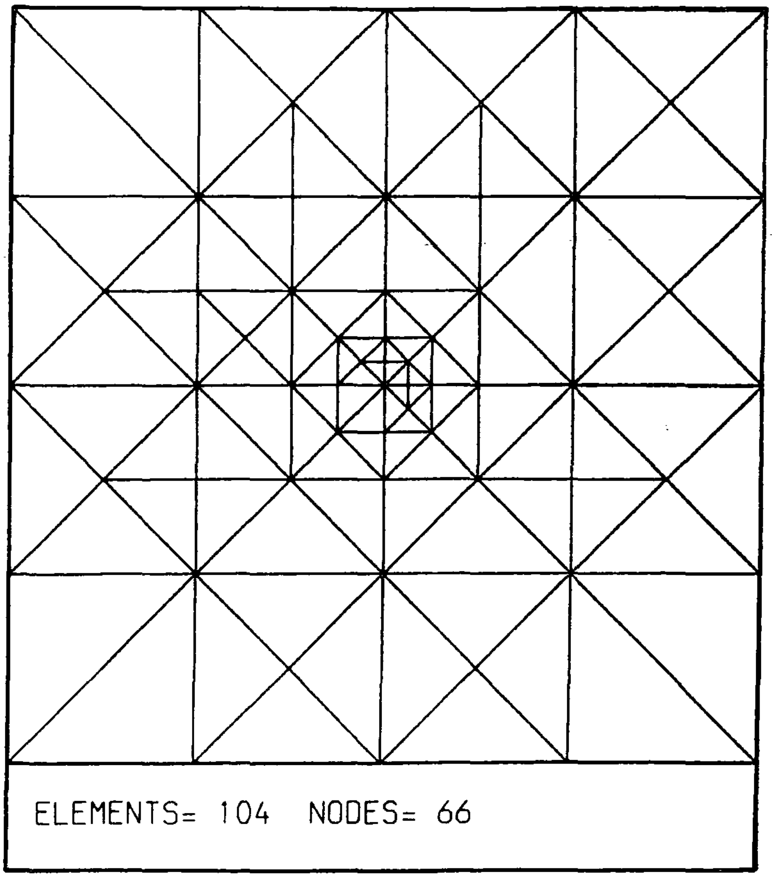


(b) Mesh 3. $\eta = 13.1\%$.

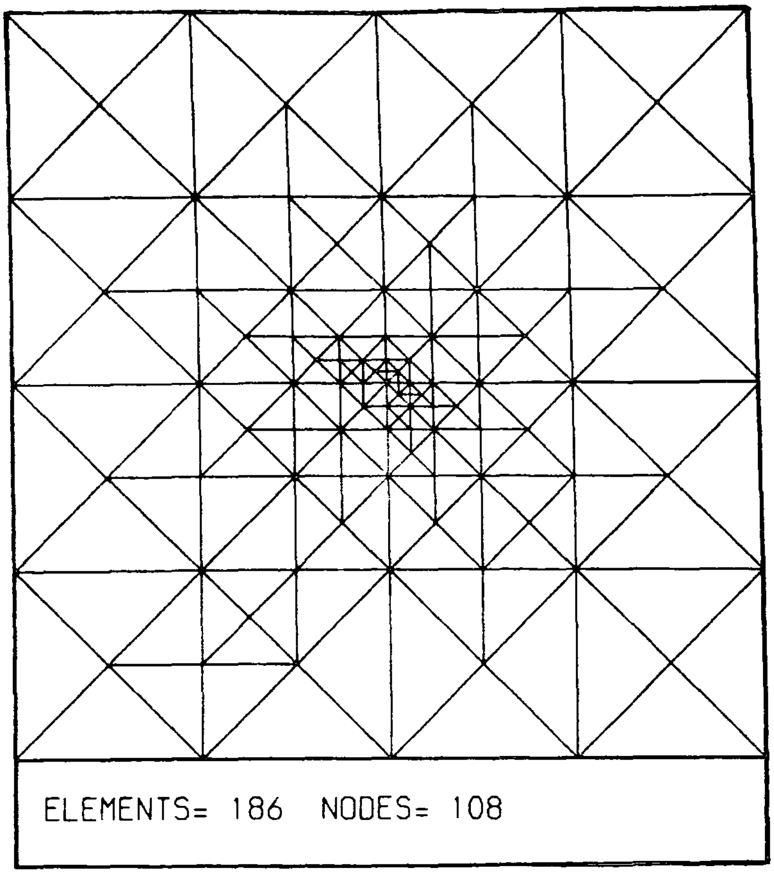
Figure 5.2 Sequence of adaptively designed meshes for slit domain problem.



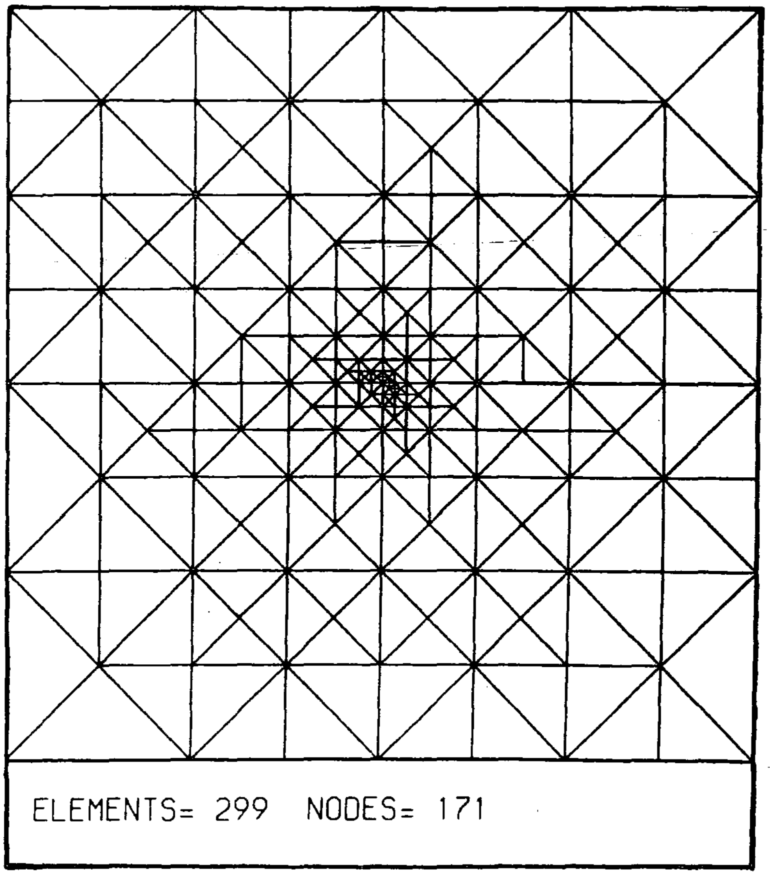
(c) Mesh 4. $\eta = 10.0\%$.



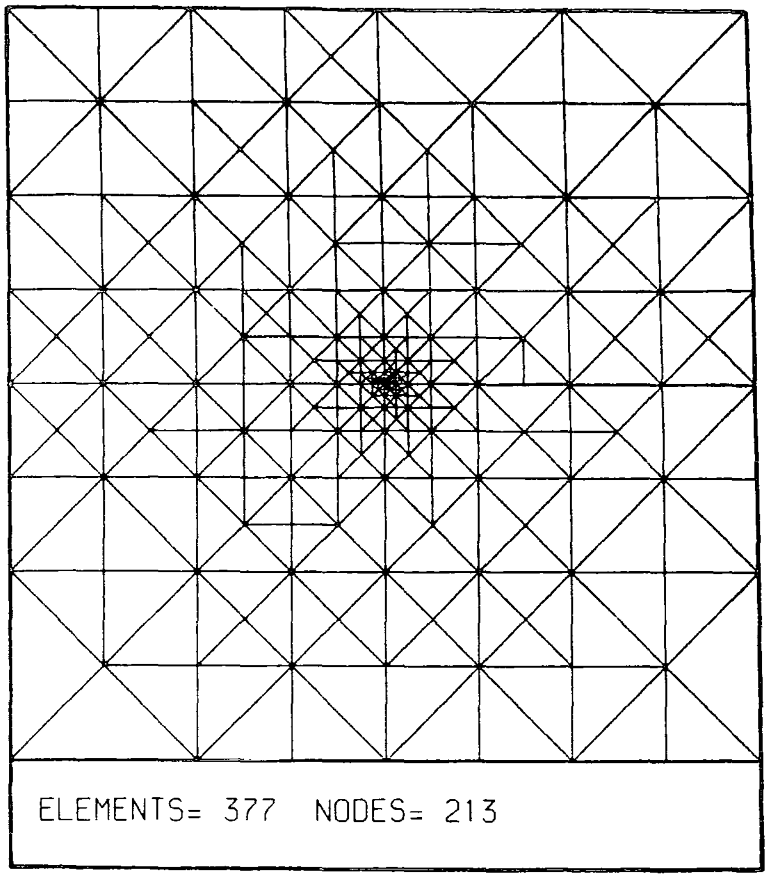
(d) Mesh 5. $\eta = 7.7\%$.



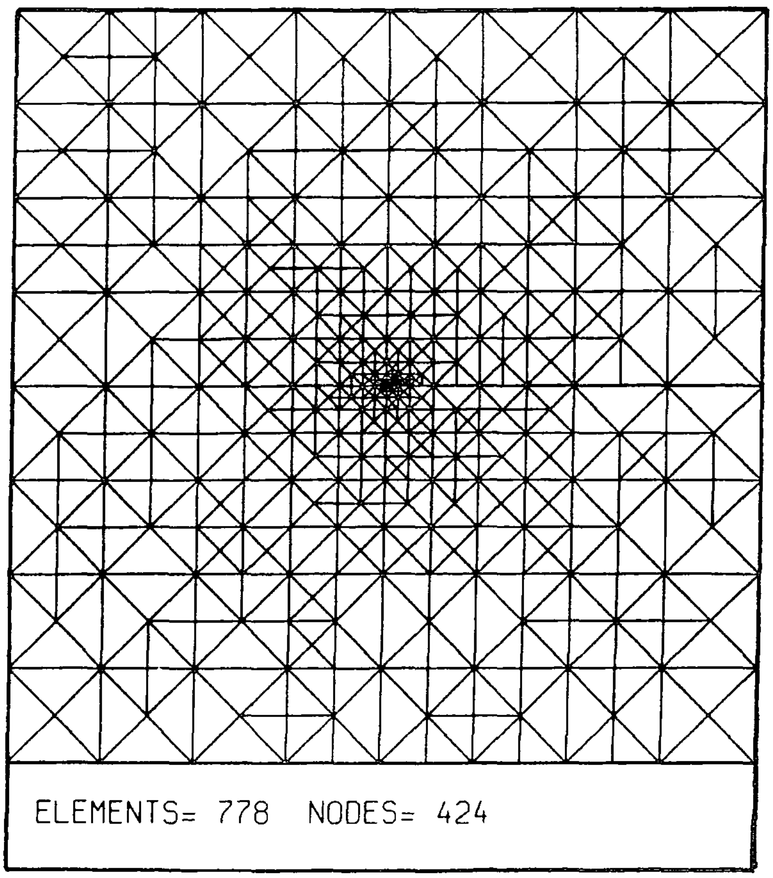
(e) Mesh 6. $\eta = 5.8\%$.



(f) Mesh 7. $\eta = 4.8\%$.



(g) Mesh 8. $\eta = 4.1\%$.



(h) Mesh 9. $\eta = 2.8\%$.

Elements	$\| e \|_E$	ϵ $\widehat{\epsilon}$	θ $\widehat{\theta}$
(a) Uniform Meshing.			
12	$.699245E+00$	$.151037E+00$ $.704114E+00$	$.216000E+00$ $.100696E+01$
24	$.652035E+00$	$.140159E+00$ $.653220E+00$	$.214957E+00$ $.100182E+01$
48	$.608278E+00$	$.130588E+00$ $.608560E+00$	$.214684E+00$ $.100046E+01$
96	$.567525E+00$	$.121799E+00$ $.567590E+00$	$.214614E+00$ $.100011E+01$
(b) Geometric Meshing.			
3	$.743390E+00$	$.264231E+00$ $.758764E+00$	$.355441E+00$ $.102068E+01$
6	$.599094E+00$	$.227678E+00$ $.599215E+00$	$.380038E+00$ $.100020E+01$
12	$.402971E+00$	$.187134E+00$ $.402861E+00$	$.464386E+00$ $.999727E+00$
24	$.200653E+00$	$.156561E+00$ $.200429E+00$	$.780257E+00$ $.998880E+00$
(c) Fibonacci Meshing.			
3	$.767866E+00$	$.217078E+00$ $.793453E+00$	$.282702E+00$ $.103332E+01$
6	$.649897E+00$	$.185868E+00$ $.650971E+00$	$.285997E+00$ $.100165E+01$
12	$.486880E+00$	$.151033E+00$ $.486871E+00$	$.310207E+00$ $.999982E+00$
24	$.280154E+00$	$.112773E+00$ $.280136E+00$	$.402539E+00$ $.999937E+00$

Table 5.3 Corrected and classical estimators for one singular term.

$s(x) = x^{3/5}$; $a(x) = 1$; $b(x) = 10$.

Elements	$\| e \|_E$	ϵ $\widehat{\epsilon}$	θ $\widehat{\theta}$
(a) Uniform Meshing.			
24	$.644840E+00$	$.127952E+00$ $.595175E+00$	$.198424E+00$ $.922982E+00$
48	$.604859E+00$	$.124694E+00$ $.581474E+00$	$.206154E+00$ $.961338E+00$
96	$.565909E+00$	$.118987E+00$ $.554945E+00$	$.210258E+00$ $.980626E+00$
192	$.528753E+00$	$.112301E+00$ $.523627E+00$	$.212389E+00$ $.990306E+00$
(b) Geometric Meshing.			
3	$.719365E+00$	$.194583E+00$ $.479268E+00$	$.270493E+00$ $.666237E+00$
6	$.597423E+00$	$.211671E+00$ $.572860E+00$	$.354307E+00$ $.958886E+00$
12	$.404720E+00$	$.177621E+00$ $.400080E+00$	$.438875E+00$ $.988536E+00$
24	$.204237E+00$	$.145261E+00$ $.195307E+00$	$.711239E+00$ $.956277E+00$
(c) Fibonacci Meshing.			
3	$.732774E+00$	$.150488E+00$ $.422618E+00$	$.205368E+00$ $.576738E+00$
6	$.643656E+00$	$.170261E+00$ $.591125E+00$	$.264522E+00$ $.918386E+00$
12	$.487272E+00$	$.148848E+00$ $.483941E+00$	$.305472E+00$ $.993165E+00$
24	$.281452E+00$	$.111083E+00$ $.279182E+00$	$.394679E+00$ $.991936E+00$

Table 5.4 Corrected and classical estimators for two singular terms.

$s(x) = x^{3/5} + x^{8/5}$; $a(x) = 1$; $b(x) = 10$.

Elements	$\Vert e \Vert_E$	ϵ $\hat{\epsilon}$	θ $\hat{\theta}$
(a) Uniform Meshing.			
25	$.643258E+00$	$.128621E+00$ $.596922E+00$	$.199953E+00$ $.927966E+00$
48	$.605354E+00$	$.125206E+00$ $.582409E+00$	$.206832E+00$ $.962097E+00$
96	$.566214E+00$	$.119309E+00$ $.555405E+00$	$.210713E+00$ $.980911E+00$
192	$.528943E+00$	$.112499E+00$ $.523869E+00$	$.212687E+00$ $.990407E+00$
(b) Geometric Meshing.			
3	$.715245E+00$	$.177517E+00$ $.497710E+00$	$.248191E+00$ $.695860E+00$
6	$.593861E+00$	$.200432E+00$ $.576625E+00$	$.337506E+00$ $.970976E+00$
12	$.399727E+00$	$.164363E+00$ $.404470E+00$	$.411189E+00$ $.101187E+01$
24	$.194166E+00$	$.128719E+00$ $.204137E+00$	$.662930E+00$ $.105135E+01$
(c) Fibonacci Meshing.			
3	$.730998E+00$	$.132644E+00$ $.449300E+00$	$.181456E+00$ $.614639E+00$
6	$.642782E+00$	$.161792E+00$ $.593422E+00$	$.251706E+00$ $.923209E+00$
12	$.486467E+00$	$.139968E+00$ $.484319E+00$	$.287724E+00$ $.995586E+00$
24	$.280095E+00$	$.989450E-01$ $.279754E+00$	$.353255E+00$ $.998784E+00$

Table 5.5 Corrected and classical estimators for three singular terms.

$s(x) = x^{3/5} + x^{8/5} + (1-x)^{9/10}$; $a(x) = 1$; $b(x) = 10$.

Elements	$\lVert e \rVert_E$	ϵ $\hat{\epsilon}$	θ $\hat{\theta}$
(a) Uniform Meshing.			
24	$.715211E+00$	$.151390E+00$ $.716169E+00$	$.211672E+00$ $.100134E+01$
48	$.668196E+00$	$.142399E+00$ $.668619E+00$	$.213110E+00$ $.100063E+01$
96	$.623859E+00$	$.133408E+00$ $.624051E+00$	$.213844E+00$ $.100031E+01$
192	$.582273E+00$	$.124732E+00$ $.582360E+00$	$.214215E+00$ $.100015E+01$
(b) Geometric Meshing.			
3	$.805652E+00$	$.233828E+00$ $.799091E+00$	$.290235E+00$ $.991856E+00$
6	$.655417E+00$	$.231575E+00$ $.654789E+00$	$.353324E+00$ $.999042E+00$
12	$.439310E+00$	$.188353E+00$ $.439466E+00$	$.428748E+00$ $.100036E+01$
24	$.212687E+00$	$.150981E+00$ $.213023E+00$	$.709873E+00$ $.100158E+01$
(c) Fibonacci Meshing.			
3	$.828933E+00$	$.192220E+00$ $.830097E+00$	$.231888E+00$ $.100140E+01$
6	$.711609E+00$	$.191311E+00$ $.711045E+00$	$.268843E+00$ $.999207E+00$
12	$.533959E+00$	$.158261E+00$ $.533996E+00$	$.296391E+00$ $.100007E+01$
24	$.305554E+00$	$.113874E+00$ $.305661E+00$	$.372680E+00$ $.100035E+01$

Table 5.6 Corrected and classical estimators for one singular term.

$s(x) = x^{3/5}$; $a(x) = (1.1 - x)^2$.

Elements	$\Vert e \Vert_E$	ϵ $\hat{\epsilon}$	θ $\hat{\theta}$
(a) Uniform Meshing.			
24	$.707519E+00$	$.138465E+00$ $.653431E+00$	$.195705E+00$ $.923552E+00$
48	$.664494E+00$	$.136081E+00$ $.639083E+00$	$.204789E+00$ $.961758E+00$
96	$.622099E+00$	$.130367E+00$ $.610204E+00$	$.209560E+00$ $.980879E+00$
192	$.581442E+00$	$.123283E+00$ $.575883E+00$	$.212030E+00$ $.990439E+00$
(b) Geometric Meshing.			
3	$.777861E+00$	$.164972E+00$ $.516951E+00$	$.212084E+00$ $.664580E+00$
6	$.650624E+00$	$.213991E+00$ $.626113E+00$	$.328902E+00$ $.962327E+00$
12	$.436860E+00$	$.177053E+00$ $.436633E+00$	$.405285E+00$ $.999480E+00$
24	$.207696E+00$	$.136856E+00$ $.207709E+00$	$.658921E+00$ $.100006E+01$
(c) Fibonacci Meshing.			
3	$.792034E+00$	$.120908E+00$ $.464521E+00$	$.152655E+00$ $.586491E+00$
6	$.703468E+00$	$.171664E+00$ $.645925E+00$	$.244026E+00$ $.918201E+00$
12	$.532600E+00$	$.152384E+00$ $.530346E+00$	$.286113E+00$ $.995768E+00$
24	$.303859E+00$	$.107034E+00$ $.303852E+00$	$.352249E+00$ $.999977E+00$

Table 5.7 Corrected and classical estimators for two singular terms.

$s(x) = x^{3/5} + x^{8/5}$; $a(x) = (1.1 - x)^2$.

Elements	$\parallel e \parallel_E$	ϵ $\widehat{\epsilon}$	θ $\widehat{\theta}$
(a) Uniform Meshing.			
24	$.707650E+00$	$.138768E+00$ $.655041E+00$	$.196096E+00$ $.925657E+00$
48	$.664546E+00$	$.136201E+00$ $.639647E+00$	$.204953E+00$ $.962533E+00$
96	$.622117E+00$	$.130414E+00$ $.610403E+00$	$.209629E+00$ $.981171E+00$
192	$.581450E+00$	$.123302E+00$ $.575954E+00$	$.212059E+00$ $.990549E+00$
(b) Geometric Meshing.			
3	$.777919E+00$	$.164813E+00$ $.535811E+00$	$.211864E+00$ $.688774E+00$
6	$.650120E+00$	$.213069E+00$ $.627211E+00$	$.327738E+00$ $.964762E+00$
12	$.436084E+00$	$.175803E+00$ $.437351E+00$	$.403140E+00$ $.100291E+01$
24	$.206060E+00$	$.135234E+00$ $.209212E+00$	$.656282E+00$ $.101530E+01$
(c) Fibonacci Meshing.			
3	$.792744E+00$	$.121362E+00$ $.491912E+00$	$.153091E+00$ $.620518E+00$
6	$.703359E+00$	$.171184E+00$ $.647896E+00$	$.243380E+00$ $.921145E+00$
12	$.532344E+00$	$.151592E+00$ $.530473E+00$	$.284763E+00$ $.996485E+00$
24	$.303412E+00$	$.105895E+00$ $.304029E+00$	$.349015E+00$ $.100203E+01$

Table 5.8 Corrected and classical estimators for three singular terms.
$s(x) = x^{3/5} + x^{8/5} + (1-x)^{9/10}$; $a(x) = (1.1-x)^2$.

APPENDIX

Lamé-Navier Equations and FEM Discretization.

In this section we collect together some basic results on the Lamé-Navier equations of linear elasticity and their finite element discretization. The approach we follow is absolutely standard and may be found in any standard reference on the subject e.g. Zienkiewicz [83]. However, we shall repeat the manipulations here in order to establish the notation and assumptions used in Chapter 3, and also to emphasize those results which play a key role in the analysis of Chapter 3.

We denote the solution domain by Ω and denote its boundary by Γ. The problem which we shall consider is the Lamé-Navier equations in two dimensions

$$-(\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) - \mu \Delta \mathbf{u} = \mathbf{f} \quad in \quad \Omega \tag{A.1}$$

where

- $\mathbf{u}$ is the displacement vector
- $\mathbf{f}$ is the body force
- λ and μ are the Lamé coefficients given by

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \tag{A.2}$$

and

$$\mu = \frac{E}{2(1+\nu)}, \tag{A.3}$$

with E and ν Young's modulus and Poisson's ratio respectively.

If we define the differential operator $\mathbf{S}$ to be

$$\mathbf{S} = \begin{pmatrix} \frac{\partial}{\partial x} & 0 \\ 0 & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{pmatrix} \tag{A.4}$$

and $\mathbf{D}$ to be the elasticity matrix, in the case of plane strain

$$\mathbf{D} = \frac{E}{(1+\nu)(1-2\nu)} \begin{pmatrix} 1-\nu & \nu & 0 \\ \nu & 1-\nu & 0 \\ 0 & 0 & \frac{1}{2}(1-2\nu) \end{pmatrix} \tag{A.5}$$

then the stress is given by

$$\sigma = \mathbf{DSu}. \tag{A.6}$$

With this notation, the Lamé-Navier equations (A.1) may be written in the form

$$-\mathbf{S}^t\mathbf{DSu} = \mathbf{f} \text{ in } \Omega \tag{A.7}$$

The boundary conditions are prescribed displacements on part of Γ and prescribed tractions on the remainder of Γ,

$$\begin{aligned} \mathbf{u} &= \hat{\mathbf{u}} \text{ on } \Gamma_d \\ \mathbf{H}\sigma &= \hat{\mathbf{t}} \text{ on } \Gamma_n \end{aligned} \tag{A.8}$$

where $\mathbf{H}$ is the linear operator formed from the components of the unit outward normal vector $\mathbf{n} = (n_x, n_y)^t$ to the boundary and given by

$$\mathbf{H} = \begin{pmatrix} n_x & 0 & n_y \\ 0 & n_y & n_x \end{pmatrix}. \tag{A.9}$$

(A.7) and (A.8) represent the strong form of the boundary value problem. It is necessary to transform this to a weak form. Firstly, we define the function spaces

$$A_{\hat{\mathbf{u}}} = \{\mathbf{u} \in [H^1(\Omega)]^2 : \mathbf{u} = \hat{\mathbf{u}} \quad on \quad \Gamma_d\} \tag{A.10}$$

and

$$A_0 = \{\mathbf{u} \in [H^1(\Omega)]^2 : \mathbf{u} = 0 \quad on \quad \Gamma_d\}. \tag{A.11}$$

Letting $\mathbf{v} \in A_0$ be taken arbitrarily, from (A.7) we have

$$-\mathbf{v}^t\mathbf{S}^t\mathbf{D}\mathbf{S}\mathbf{u} = \mathbf{v}^t\mathbf{f} \text{ in } \Omega \tag{A.12}$$

or, using $\sigma = \mathbf{DSu}$ we may rewrite this as

$$-\mathbf{v}^t\mathbf{S}^t\sigma = \mathbf{v}^t\mathbf{f} \text{ in } \Omega \quad \forall \mathbf{v} \in A_0. \tag{A.13}$$

Now, noting that

$$\nabla \cdot \begin{pmatrix} \sigma_x v_x + \sigma_{xy} v_y \\ \sigma_y v_y + \sigma_{xy} v_x \end{pmatrix} = \mathbf{v}^t\mathbf{S}^t\sigma + \sigma^t\mathbf{S}\mathbf{v} \tag{A.14}$$

we obtain, using the Divergence Theorem, that

$$\int_\Gamma \begin{pmatrix} \sigma_x v_x + \sigma_{xy} v_y \\ \sigma_y v_y + \sigma_{xy} v_x \end{pmatrix} \cdot \mathbf{n} ds = \int_\Omega \{\mathbf{v}^t\mathbf{S}^t\sigma + \sigma^t\mathbf{S}\mathbf{v}\}\, d\mathbf{x}, \tag{A.15}$$

using (A.13) now yields

$$\int_{\Gamma} \begin{pmatrix} \sigma_x v_x + \sigma_{xy} v_y \\ \sigma_y v_y + \sigma_{xy} v_x \end{pmatrix} \cdot \mathbf{n} ds = \int_{\Omega} \{ -\mathbf{v}^t \mathbf{f} + \sigma^t \mathbf{S} \mathbf{v} \} d\mathbf{x}. \qquad (A.16)$$

Considering the integral along Γ and noting that $\mathbf{v} = 0$ on Γ_d, we have

$$\begin{aligned} \int_{\Gamma} \begin{pmatrix} \sigma_x v_x + \sigma_{xy} v_y \\ \sigma_y v_y + \sigma_{xy} v_x \end{pmatrix} \cdot \mathbf{n} ds &= \int_{\Gamma_n} \begin{pmatrix} \sigma_x v_x + \sigma_{xy} v_y \\ \sigma_y v_y + \sigma_{xy} v_x \end{pmatrix} ds \\ &= \int_{\Gamma_n} [v_x, v_y] \begin{pmatrix} \sigma_x n_x + \sigma_{xy} n_y \\ \sigma_y n_y + \sigma_{xy} n_x \end{pmatrix} \cdot \mathbf{n} ds \\ &= \int_{\Gamma_n} \mathbf{v}^t \mathbf{H} \sigma ds \\ &= \int_{\Gamma_n} \mathbf{v}^t \hat{\mathbf{t}} ds. \qquad (A.17) \end{aligned}$$

Combining (A.16) and (A.17) gives, after some rearrangement,

$$B(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) + \langle \hat{\mathbf{t}}, \mathbf{v} \rangle_{\Gamma_n} \quad \forall \mathbf{v} \in A_0 \qquad (A.18)$$

where

$$A_{\hat{\mathbf{u}}} = \{\mathbf{u} \in [H^1(\Omega)]^2 : \mathbf{u} = \hat{\mathbf{u}} \quad on \quad \Gamma_d\} \qquad (A.19)$$

$$B(\mathbf{u}, \mathbf{v}) = \int_{\Omega} (\mathbf{S}\mathbf{u})^t \mathbf{D} \mathbf{S} \mathbf{v} d\mathbf{x} \qquad (A.20)$$

$$(\mathbf{f}, \mathbf{v}) = \int_{\Omega} \mathbf{f}^t \mathbf{v} d\mathbf{x} \qquad (A.21)$$

and

$$\langle \hat{\mathbf{t}}, \mathbf{v} \rangle_{\Gamma_n} = \int_{\Gamma_n} \mathbf{v}^t \hat{\mathbf{t}} ds. \qquad (A.22)$$

(A.18) is the weak form of the Lamé-Navier equations and incorporates the boundary conditions (A.8). The standard (Galerkin) finite element approximation to (A.18), gives an approximation of the form

$$\mathbf{u}_h = \hat{\mathbf{u}} + \begin{bmatrix} u_h^x \\ u_h^y \end{bmatrix} \tag{A.23}$$

where $u_h^x = \sum_i u_i^x N_i(x,y)$ and $u_h^y = \sum_i u_i^y N_i(x,y)$, and $N_i(x,y)$ are basis functions. The conditions on the displacement may be imposed by constraining the coefficients $\{u_i^x, u_i^y\}$ appropriately (see below). The remaining degrees of freedom are then determined from the condition

$$B(\mathbf{u}_h, \mathbf{N}_i) = (\mathbf{f}, \mathbf{N}_i) + \langle \hat{\mathbf{t}}, \mathbf{N}_i \rangle_{\Gamma_n} \tag{A.24}$$

where

$$\mathbf{N}_i = \begin{bmatrix} N_i & 0 \\ 0 & N_i \end{bmatrix} \tag{A.25}$$

and i ranges over all values of i whose associated node $\mathbf{x}_i$ does not lie on the portion Γ_d of the boundary. The space spanned by $\mathbf{N}_i$ for such values of i is denoted by A_0^h and we let $A_{\hat{\mathbf{u}}}^h = \hat{\mathbf{u}} + A_0^h$. We may then rewrite (A.24) in the form

$$\mathbf{u}_h \in A_{\hat{\mathbf{u}}}^h : B(\mathbf{u}_h, \mathbf{v}_h) = (\mathbf{f}, \mathbf{v}_h) + \langle \hat{\mathbf{t}}, \mathbf{v}_h \rangle_{\Gamma_n} \quad \forall \mathbf{v}_h \in A_0^h. \tag{A.26}$$

References.

[1] M. Ainsworth, *An asymptotically exact a posteriori error estimator for the finite element approximation of problems with singular solutions.* Durham University Numerical Analysis Report NA89/02, 1989.

[2] M. Ainsworth and A. W. Craig, *A posteriori error estimators in the finite element method.* Durham University Numerical Analysis Report NA88/02, Revised Version NA89/03, 1989.

[3] M. Ainsworth, A. W. Craig, J. Z. Zhu and O. C. Zienkiewicz, *h and h-p version error estimation and adaptive procedures from theory to practice.* Durham University Numerical Analysis Report NA89/04, 1989. To appear in a special edition of *Engineering with Computers.*

[4] M. Ainsworth, J. Z. Zhu, A. W. Craig and O. C. Zienkiewicz, *Analysis of the Zienkiewicz Zhu a posteriori error estimator in the finite element method.* Int. J. Num. Meth. Eng., **28**, 2161–2174, 1989.

[5] A. B. Andreev and R. D. Lazarov, *Superconvergence of the gradient for quadratic triangular elements.* Num. Meth. for PDEs, **4** , pp.15–32, 1988.

[6] J. P. Aubin and H. G. Burchard, *Some aspects of the method of the hypercircle applied to elliptic variational problems.* in *Numerical Solution of Partial Differential Equations-II SYNSPADE 1970.* ed. B.Hubbard, Academic Press 1971.

[7] I. Babuška and M. R. Dorr, *Error estimates for the combined h and p versions of the finite element method.* Numer. Math. **27**, pp.257–277, 1987.

[8] I. Babuška, M. Griebel and J. Pitkäranta, *The problem of selecting shape functions for a p type finite element method..* Int. J. Num. Meth. Eng. **28**,

pp.1891–1908, 1989.

[9] I. Babuška and B. Q. Guo, *The h-p version of the finite element method for domains with curved boundaries.* SIAM J. N. A. **25**, pp.837–861, 1988.

[10] I. Babuška and A. Miller, *The post-processing approach in the finite element method. Part II: The calculation of stress intensity factors.* Int. J. Num. Meth. Eng., **20**, pp.1111–1129, 1984.

[11] I. Babuška and A. Miller, *The post-processing approach in the finite element method. Part III: A posteriori error estimates and adaptive mesh selection.* Int. J. Num. Meth. Eng., **20**, pp.2311–2324, 1984.

[12] I. Babuška and W. C. Rheinboldt, *A posteriori error analysis of finite element solutions for one dimensional problems.* Siam J. Numer. Anal. **18**, No.3, pp.565–589, 1981.

[13] I. Babuška and W. C. Rheinboldt, *A posteriori error estimates for the finite element method.* Int. J. Num. Meth. Eng., **12**, pp.1597–1615, 1978.

[14] I. Babuška and W. C. Rheinboldt, *Error estimates for adaptive finite element computations.* SIAM J. Num. Anal., **15**, pp.736–754, 1978.

[15] I. Babuška and W. C. Rheinboldt, *Adaptive approaches and reliability estimations in finite element analysis.* Comp. Meth. App. Mech. Eng., **17**, pp.519–540, 1979.

[16] I. Babuška and W. C. Rheinboldt, *Analysis of optimal finite element meshes in* $\Re^1$. Math. Comp., **33**, pp.435–463, 1979.

[17] I. Babuška and M. Suri, *The optimal convergence rate of the p version of the finite element method.* SIAM J. N. A. **24**, pp.750–776, 1987.

[18] I. Babuška, B. A. Szabo and I. N. Katz, *The p version of the finite element method.* SIAM J. N. A. **18**, pp.512–545, 1981.

[19] I. Babuška and D. Yu, *Asymptotically exact a posteriori error estimator for biquadratic elements.* Tech. Note BN–1050, Inst. Phys. Sci. and Tech., Univ. of Maryland, June 1986.

[20] I. Babuška, O. C. Zienkiewicz, J. Gago and E. R. de A. Oliveira, *Accuracy Estimates and Adaptive Refinements in Finite Element Computations.* Wiley, New York, 1986.

[21] R. E. Bank and T. F. Dupont, *An optimal order process for solving finite element equations.* Math. Comp., **36**, pp.35–51, 1985.

[22] R. E. Bank and A. Weiser, *Some a posteriori error estimators for elliptic partial differential equations.* Math. Comp., **44**, pp.283–301, 1985.

[23] J. W. Barrett, G. Moore and K. W. Morton, *Optimal Recovery in the Finite Element Method. Part I: Recovery from weighted L^2 fits.* IMA J. Num. Anal., **8**, pp.149–184, 1988.

[24] J. W. Barrett, G. Moore and K. W. Morton, *Optimal Recovery in the Finite Element Method. Part II: Defect correction for ordinary differential equations.* IMA J. Num. Anal., **8**, pp.527–540, 1988.

[25] J. H. Bramble and S. R. Hilbert, *Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolation.* Siam J. Num. Anal. **7**, pp.112–124, 1970.

[26] J. H. Bramble and A. H. Schatz, *Higher order local accuracy by averaging in the finite element method.* Math. Comp. **31**, pp.94–111, 1977.

[27] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems.* North–Holland, 1978.

[28] A. W. Craig, J. Z. Zhu and O. C. Zienkiewicz, *A posteriori error estimation, adaptive mesh refinement and multigrid methods using hierarchical finite element bases.* in J. R. Whiteman (ed), *Mathematics of Finite Element Applications V.* Academic Press, New York, pp.587–594, 1985.

[29] P. Davis, *Interpolation and Approximation.* Blaisdell, New York, 1963.

[30] L. Demkowicz, Ph. Devloo and J. T. Oden, *On an h-type mesh refinement strategy based on minimization of interpolation errors.* Comp. Meth. Appl. Mech. Eng., **53**, pp.67–89, 1985.

[31] L. Demkowicz and J. T. Oden, *A review of mesh refinement techniques and corresponding data structures in h-type adaptive finite element methods.* TICOM Report 88-02, The University of Texas at Austin, 1988.

[32] L. Demkowicz, J. T. Oden, W. Rachowicz and O. Hardy, *Toward a universal h-p adaptive finite element strategy. Part 1: Constrained approximation and data structure.* TICOM Report 89-01, The University of Texas at Austin, March 1989.

[33] I. Ekeland and R. Témam, *Convex Analysis and Variational Problems.* North-Holland, 1976.

[34] J. E. Flaherty, P. J. Paslow, M. S. Shephard and J. D. Vasilakis, *Adaptive methods for partial differential equations.* SIAM 1989.

[35] J. P. de S. R. Gago, D. W. Kelly, O. C. Zienkiewicz and I. Babuška, *A posteriori error analysis and adaptive processes in the finite element method. Part 2: Adaptive mesh refinement.* **19**, pp.1621–1656, 1983.

[36] P. Grisvard, *Behavior of the solution of an elliptic boundary value problem in a polygonal or polyhedral domain.* in *Numerical Solution of Partial Differential Equations III: SYNSPADE 1975.* ed. B. Hubbard, Academic Press.

[37] W. Gui and I. Babuška, *The h, p and h-p versions of the finite element method in one dimension. Part 1: The error analysis of the p version. Part 2: The error analysis of the h and h-p versions. Part 3: The adaptive h-p version.* Numer. Math., **48**, pp.557–612, pp.613–657, pp.658–683, 1986.

[38] B. Q. Guo and I. Babuška, *The h-p version of the finite element method. Part 1: The basic approximation results. Part 2: General results and applications.* Comp. Mech. **1**, pp.21–41, pp.203–226, 1986.

[39] W. Hackbusch, *Multi-Grid Methods and Applications.* Springer-Verlag, 1985.

[40] E. Hinton and J. Campbell *Local and global smoothing of discontinuous finite element functions using least squares methods.* Int. J. Numer. Meth. Eng. **8**, pp.461–480, 1974.

[41] I. Hlaváček and M. Křížek, *On a superconvergent finite element scheme for elliptic systems. Part I: Dirichlet Boundary Conditions.* Aplikace Matematiky, **32**, pp.131–153, 1987.

[42] I. Hlaváček and M. Křížek, *On a superconvergent finite element scheme for elliptic systems. Part II: Boundary Conditions of Newton's or Neumann's Type.* Aplikace Matematiky, **32**, pp.200–213, 1987.

[43] I. Hlaváček and M. Křížek, *On a superconvergent finite element scheme for elliptic systems. Part III: Optimal Interior Estimates.* Aplikace Matematiky, **32**, pp.276–289, 1987.

[44] R. B. Kellogg, *Higher order singularities for interface problems.* in *Numerical Solution of Partial Differential Equations III: SYNSPADE 1975.* ed. B. Hubbard, Academic Press.

[45] D. W. Kelly, *The self-equilibration of residuals and complementary a posteriori error estimates in the finite element method.* Int. J. Numer. Meth. Eng. **20**, pp.1491–1506, 1984.

[46] D. W. Kelly, R. Gago, O. C. Zienkiewicz and I. Babuška, *A posteriori error analysis and adaptive processes in the finite element method. Part I – Error Analysis.* Int. J. Num. Meth. Eng. **19**, pp.1593–1619, 1983.

[47] D. W. Kelly and J. D. Isles, *A procedure for a posteriori error analysis for the finite element method which contains a bounding measure.* Comp. and Struct., **31**, pp.63–71, 1989.

[48] D. W. Kelly, R. J. Mills, J. A. Reizes and A. D. Miller, *A posteriori error estimates of the solution error caused by discretization in the finite element, finite difference and boundary element methods.* Int. J. Numer. Meth. Eng. **24**, pp.1921, 1987.

[49] V. A. Kondratév, *Boundary value problems for elliptic equations in domains with conical or angular points.* Trans. Moscow Math. Soc., pp.227–313, **16**, 1967.

[50] M. Křížek and P. Neitaanmäki, *On a global superconvergence of the gradient of linear triangular elements.* J. Comp. App. Math. **18**, pp.221-233, 1987.

[51] M. Křížek and P. Neitaanmäki, *On superconvergence techniques.* Acta Applicandae Mathematicae, **9**, pp.175–198, 1987.

[52] P. Lesaint and M. Zlámal, *Superconvergence of the gradient of finite element solutions.* RAIRO Anal. Numér., **13**, pp.139–166, 1979.

[53] N. D. Levine, *Superconvergent estimation of the gradient from linear finite element approximations on triangular elements.* Reading University Numerical Analysis Report 3/85, 1985.

[54] N. D. Levine, *Superconvergent recovery of the gradient from piecewise linear finite element approximations.* IMA J. Num. Anal., **5**, pp.407–427, 1985.

[55] S. G. Mikhlin, *Variational Methods in Mathematical Physics.* Pergamon Press, 1964.

[56] A. K. Noor and I. Babuška, *Quality assessment and control of finite element solutions.* Finite Elements and Design, **3**, pp.1–26, 1987.

[57] J. T. Oden, L. Demkowicz, W. Rachowicz and T. A. Westermann, *Toward a universal h-p adaptive finite element strategy. Part 2: A posteriori error estimation.* TICOM Report 89–02, The University of Austin at Texas, March 1989.

[58] J. T. Oden, L. Demkowicz, T. Strouboulis and Ph. Devloo, *Adaptive Methods for Problems in Solid and Fluid Mechanics.* in Babuška*et al.* (eds), *Accuracy Estimates and Adaptive Refinements in Finite Element Computations.* Wiley, New York, pp.249–280, 1986.

[59] J. T. Oden and J. N. Reddy, *An introduction to the mathematical theory of finite elements.* Wiley 1976.

[60] T. Patterson, *Algorithm 468: Algorithm for numerical integration over a finite interval.* ACM Trans. Math. Soft., **16**, pp.694–699, 1973.

[61] J. Peraire, J. Peiro, L. Formaggia, K. Morgan and O. C. Zienkiewicz, *Finite element Euler computations in three dimensions.* Int. J. Num. Meth. Eng. **26**, pp.2135–2159, 1988.

[62] J. Peraire, M. Vahdati, K. Morgan and O. C. Zienkiewicz, *Adaptive remeshing for compressible flow computations.* J. Comp. Phys. **72**, pp.449–466, 1987.

[63] W. Rachowicz and J. T. Oden, *On the accuracy and convergence of conjugate flux approximations.* Num. Meth. for PDEs, **5** , pp.143–156, 1988.

[64] W. Rachowicz, J. T. Oden and L. Demkowicz, *Toward a universal h-p adaptive finite element strategy. Part 3: Design of h-p meshes.* TICOM Report 89–03, The University of Austin at Texas, March 1989.

[65] W. C. Rheinboldt and C. K. Mesztenyi, *On a data structure for adaptive finite element mesh refinements.* ACM Trans. Math. Software, **6**, pp.166–187, 1980.

[66] M. C. Rivara, *Algorithms for refining triangular grids suitable for adaptive and multigrid techniques.* Int. J. Num. Meth. Eng., **20**, pp.745–756, 1984.

[67] M. C. Rivara, *Mesh refinement processes based on the generalized bisection of simplices.* SIAM J. Num. Anal., **21**, pp.604–613, 1984.

[68] M. C. Rivara, *Design and data structure of fully adaptive multigrid finite element software.* ACM Trans. Math. Soft., **10**, pp.242–264, 1984.

[69] G. Sander, *Application of the dual analysis principle.* in *High Speed Computing of Elastic Structures.* ed. B. F. de Veubeke, Université de Liége, 1971.

[70] L. L. Schumaker, *Spline functions: Basic theory.* Wiley, 1981.

[71] B. Specht, *A general construction of local error estimators for conforming finite element methods.* Comp. and Struct., **19**, pp.815–822, 1984.

[72] J. L. Synge, *The hypercircle method in mathematical physics.* Cambridge University Press, 1957.

[73] B. A. Szabo, *Estimation and control of error based on p convergence.* in I.

Babuška *et al.* (eds), *Accuracy Estimates and Adaptive Refinements in Finite Element Computations.* Wiley, New York, pp.61–70, 1986.

[74] B. A. Szabo, *Mesh design for the p version of the finite element.* Comp. Meth. Appl. Mech. Eng. **55**, pp.181–197, 1986.

[75] B. A. Szabo, *PROBE: Theoretical Manual.* Noetic Technologies Corporation, St Louis, MO, 1985.

[76] V. Thomée , *High order local approximations to derivatives in the finite element method.* Math. Comp. **31**, pp.652–660, 1977.

[77] D. J. Turcke and G. M. McNiece, *Guidelines for selecting finite element grids based on optimisation studies.* Comp. and Struct., **4**, pp.449–519, 1974.

[78] S. Utku and R. J. Melosh, *Solution errors in finite element analysis.* Comp. and Struct., **18**, pp.379–393, 1984.

[79] B. Fraeijs de Veubeke, *Displacement and equilibrium models in the finite element method.* in *Stress Analysis.* ed. Zienkiewicz and Holister, Wiley, London, 1965.

[80] J. R. Whiteman and G. Goodsell, *Lectures on gradient superconvergence of finite element approximation to elliptic problems.* Numerical Analysis Summer School, Lancaster 1987. To appear in P. Turner (ed), Springer Lecture Notes in Mathematics, Springer-Verlag, Berlin.

[81] H. Yserentant, *The convergence of multi-level methods for solving finite element equations in the presence of singularities.* Math. Comp., **47**, pp.399–409, 1986.

[82] J. Z. Zhu and O. C. Zienkiewicz *Adaptive techniques in the finite element method.* Comm. Appl. Num. Meth. **4**, pp.197–204, 1988

[83] O. C. Zienkiewicz, *The Finite Element Method.* 3rd Edition, McGraw-Hill 1977.

[84] O. C. Zienkiewicz and A. W. Craig, *Adaptive mesh refinement and a posteriori error estimation for the p version of the finite element method.* in I. Babuška

et al. (eds.) *Accuracy Estimates and Adaptive Refinements in Finite Element Computations.* Wiley, New York, 1986, pp.25–59, 1986.

[85] O. C. Zienkiewicz and A. W. Craig, *A posteriori error estimation and adaptive mesh refinement in the finite element method.* in *The mathematical basis of finite element methods with applications to partial differential equations.* ed. D. F. Griffiths, Clarendon Press, Oxford, pp.71–89, 1984.

[86] O. C. Zienkiewicz, J. P. de S. R. Gago and D. W. Kelly, *The hierarchical concept in the finite element method.* Comp. and Struct. **16**, pp. 53-65, 1983.

[87] O. C. Zienkiewicz and J. Z. Zhu *Further research on the simple error estimators.* Report of Inst. Num. Meth. Eng. C/R/615/88, Swansea, UK. 1988.

[88] O. C. Zienkiewicz and J. Z. Zhu *A simple error estimator and adaptive procedure for practical engineering analysis.* Int. J. Numer. Meth. Eng. **24**, pp.337–357, 1987

[89] O. C. Zienkiewicz, J. Z. Zhu, A. W. Craig and M. Ainsworth, *Simple and practical error estimation and adaptivity: h and h-p version procedures.* Chapter 8, pp.100–115 in *Adaptive Methods for PDEs.* ed. J.E. Flaherty *et al.*, SIAM 1989.

[90] O. C. Zienkiewicz, J. Z. Zhu and N. G. Gong, *Effective and practical h-p version adaptive analysis procedures for the finite element method.* Int. J. Num. Eng. **28**, pp.879–891, 1989.

[91] M. Zlámal, *Some superconvergence results in the finite element method.* in *Mathematical Aspects of Finite Element Methods: Proceedings of the conference held in Rome 1975* Springer Lecture Notes in Mathematics, **606**, 1975.

[92] M. Zlámal, *Superconvergence and reduced integration in the finite element method.* Math. Comp. **32**, pp.663–685, 1978.