Verification in computational structural mechanics: recovery-based a posteriori error estimation

by

Giovanni Castellazzi

Advisor: Prof. Ing. Francesco Ubertini

co-Advisor: Ing. Stefano de Miranda

Dottorato di Ricerca in Meccanica delle Strutture - XIX Ciclo

Coordinatore: Prof. Ing. Erasmo Viola

Settore scientifico disciplinare:

Area 08 - Ingegneria civile e Architettura

ICAR/08 - Scienza delle Costruzioni

ALMA MATER STUDIORUM - University of Bologna

Bologna, March 12^{th}, 2007

"Verification in computational structural mechanics: recovery-based a posteriori error estimation," a dissertation prepared by Giovanni Castellazziin partial fulfillment of the requirements for the degree of Doctor of Philosophy, has been approved and accepted by the following:

Prof. Elio Sacco

Prof. Alessandro De Stefano

Prof. Francesco Ubertini

April 2007

Abstract

Verification in computational structural mechanics: recovery-based a posteriori error estimation

by

Giovanni Castellazzi

Dottorato di Ricerca in Meccanica delle Strutture - XIX Ciclo

Coordinatore: Prof. Ing. Erasmo Viola

Prof. Francesco Ubertini

Ing. Stefano de Miranda

ALMA MATER STUDIORUM - University of Bologna

Computational engineering, the discipline concerned with the use of computational methods and devices to simulate physical events and engineering systems, is being heralded by many as one of the most important developments in recorded history. Computer predictions of physical events, it is argued, can be of enormous importance in making critical decisions that affect every facet of human existence. As the speed and capacity of computer systems continue to grow, the expectations of users of computer models in decision making continues to grow in kind. Today, some look toward computer-based predictions as a means to obtain vital quantitative information on events that influence the security, health, and well being of much of mankind and many nations and that influence the success of major businesses and other enterprises worldwide.

"Homo qui erranti comiter monstrat viam
Quasi lumen de suo lumine accendat facit.
Nihilo minus ipsi lucet cum illi accenderit."

[maxim ascribed to Ennio by Cicerone
in "De Officiis", I, chapter XVI]

Aknowledgements

I owe a debt of gratitude to Professor Francesco Ubertini who gave me the opportunity to write the present work and to Ing. Stefano de Miranda who was always helpful. I would say that I feel very lucky to have worked in this group because they embody the maxim reported on this page. I wish that it would never change, because there are no better research mates for me.

I would like to thank Professor Erasmo Viola for giving me the opportunity to attend the Ph.D. program in Structural Mechanics, and for his enthusiasm, passion and greatest care to improve at the highest level the school that I'm leaving in reaching this goal.

I would like to thank all the guys of the LAMC laboratory for the friendship and the support during these years.

I would like to thank John Barlow for his kindly encouragement and suggestion for historical information about pioneers in FE, and Richard Haigh of Rolce Royce Heritage Trust who put me in contact with John Barlow.

I would like to thank Giuseppe Miccoli of NAFEMS Italian Division, Trevor Hellen and Tim Morris of NAFEMS Ltd for some interesting advice about the bibliography of geometry distortions topic.

I would like to thank my family: my sister Sara and my parents Rosanna and Dario who always supported me in all situations and in all my choices.

Bologna, March 12^{th}, 2007

Contents

Sommario

Questa tesi riguarda la stima a posteriori dell'errore commesso nella discretizzazione in elementi finiti di problemi di analisi strutturale e sulle tecniche di adattività ad essa correlate. L'attenzione è rivolta ai metodi di ricostruzione, di recovery in dizione anglosassone, che stimano l'errore per confronto con una soluzione più accurata che viene ricostruita con opportune procedure a partire da quella per elementi finiti. In questo contesto è stata sviluppata una nuova procedura di ricostruzione locale degli sforzi, ottenuta rilassando la compatibilità cinematica, attraverso la minimizzazione di un funzionale tipo energia complementare, appositamente definito su un piccolo insieme di elementi. La procedura è stata impiegata con risultati molto soddisfacenti su vari modelli strutturali, discretizzati sia con elementi compatibili che con elementi finiti agli sforzi. In particolare, essa è stata applicata con successo alle piastre laminate, descritte attraverso la cosiddetta "First order Shear Deformation Theory", con l'obiettivo di ricostruire la distribuzione delle tensioni tangenziali lungo lo spessore della piastra. In aggiunta a ciò, infine, è stato analizzato il ruolo dello stimatore d'errore quale parametro guida per la ricostruzione della discretizzazione in analisi adattative, mostrando l'efficacia della procedura sviluppata anche in questo contesto.

La tesi si articola in sei capitoli.

Nel primo capitolo, introduttivo, vengono illustrate le motivazioni che hanno spinto e spingono gli analisti alla ricerca di strumenti di verifica che consentano di attribuire alle analisi numeriche un determinato grado di affidabilità.

Nel secondo capitolo, viene presentato il metodo degli elementi finiti, definito il problema modello ed introdotto il concetto di errore commesso nella discretizzazione per elementi finiti, inquadrando la tesi nello stato dell'arte sull'argomento.

Nel terzo capitolo, vengono illustrati alcuni risultati classici sulla stima a priori dell'errore e presentati alcuni aspetti originali inerenti gli effetti delle distorsioni di geometria sulle prestazioni degli elementi finiti isoparametrici.

Nel quarto capitolo, viene discussa la stima a posteriori dell'errore. Vengono analizzati e confrontati i principali metodi di stima a posteriori, ponendo l'attenzione sui metodi basati sulla ricostruzione del campo di sforzi (ovvero i cosiddetti recovery based methods). In questo contesto, viene sviluppata una nuova procedura che si dimostra essere versatile, robusta e superconvergente. Alcune

applicazioni numeriche sono presentate a conclusione di tale capitolo.

Nel quinto capitolo, viene discusso il ruolo dello stimatore d'errore quale parametro di controllo per la ricostruzione della discretizzazione in analisi adattative. Alcune applicazioni numeriche mostrano l'efficacia della procedura proposta come guida del processo di analisi.

Nel sesto capitolo, conclusivo, vengono presentate alcune applicazioni avanzate, quali l'impiego della tecnica proposta con elementi finiti di tipo misto agli sforzi, l'applicazione alle piastre e l'estensione alla ricostruzione del profilo delle tensioni tangenziali lungo lo spessore di compositi laminati.

Chapter 1

What is the error?

Questo capitolo introduce il concetto di verifiche in analisi strutturali per elementi finiti. A tale scopo, vengono illustrati i passaggi fondamentali che conducono ad una analisi agli elementi finiti, mettendo n luce i punti in cui vengono introdotte le approssimazioni e gli errori che ne conseguono. In particolare, vengono illustrati gli errori di modellazione, discretizzazione e soluzione, e viene definito il processo di verifica come quello che conduce alla quantificazione della somma degli ultimi due.

Nella parte conclusiva del capitolo, inoltre, viene riportato un piccolo estratto storico sull'evoluzione del metodo degli elementi finiti, cercando di evidenziare la crescente sensibilità degli analisti verso i procedimenti di verifica e controllo dell'errore in analisi strutturali agli elementi finiti.

1.1 A point of view

The word error has different meanings in different fields. The concrete meaning of the Latin word error means "wandering" or "straying", although the metaphorical meaning "mistake, misapprehension" is actually more common. To the contrary of an illusion, an error or a mistake can be dispelled through knowledge (knowing that one is looking at a mirage and not at real water doesn't make the mirage disappear). Looking at the Figure 1.1 is simple to understand that something was wrong wen the train broke down the safety structure, passing through the restaurant railway, destroying a newspaper kiosk and concluding its trip on the

Figure 1.1: A famous image about error: train wreck at Montparnasse, Paris, France, 1895, (photgrapher unknown). From Wikipedia, the free encyclopedia, *www.wikipedia.org*

road. Let us start with some point of view, in some disciplines.

In statistics an error is a difference between a computed, estimated, or measured value and the true, specified, or theoretically correct value.

In experimental science an error is a bound on the precision and accuracy of the result of a measurement. These can be classified into two types: statistical error (see above) and systematic error. Statistical error is caused by random (and therefore inherently unpredictable) fluctuations in the measurement apparatus, whereas systematic error is caused by an unknown but nonrandom fluctuation. If the cause of the systematic error can be identified, then it can usually be eliminated. Such errors can also be referred to as uncertainties.

In engineering an error is a difference between desired and actual performance. Engineers often seek to design systems in such a way as to mitigate or preferably avoid the effects of error, whether unintentional or not. One type of error is the human error which includes cognitive bias. Human factors engineering is often applied to designs in an attempt to minimize this type of error by making systems more forgiving or error-tolerant. Errors in a system can also be latent design errors that may go unnoticed for years, until the right set of circumstances arises that cause them to become active.

In computer programming (i.e. in software engineering), the term error refers to an incorrect action or calculation performed by software as a result of a fault (a fault being incorrect code or data in the software or hardware). If, as a result of the error, the software performs an undesired action or fails to perform a desired action, then this is referred to as a failure. Note that a fault can exist in software and not cause an error, and an error can occur without causing a failure.

Looking again at Figure 1.1 is clear that the desired performance was not reproduced in the practice and that the safety structures and the restaurant furniture were not forgiving system as described above.

In science, the word error does not carry the usual connotations of the terms mistake or blunder. Error in a scientific measurement means the inevitable uncertainty that attends all measurements. As such, errors are not mistakes, you cannot eliminate them by being very careful. The best you can hope to do is to ensure that errors are as small as reasonably possible and to have a reliable estimate of how large they are. For now, error is used exclusively in the sense of uncertainty, and the two words are used interchangeably.

1.2 Computational engineering

Computational engineering, the discipline concerned with the use of computational methods and devices to simulate physical events and engineering systems, is being heralded by many as one of the most important developments in recorded history. Computer predictions of physical events, it is argued, can be of enormous importance in making critical decisions that affect every facet of human existence. As the speed and capacity of computer systems continue to grow, the expectations of users of computer models in decision making continues to grow in kind. Today, some look toward computer-based predictions to obtain vital quantitative information on events that influence the security, health, and well being of much of mankind and of many nations and that influence the success of major businesses and other enterprises worldwide.

Somebody speaks of the coming crisis in computational engineering arising from three major challenges: a) performance, b) programming, and c) prediction. The performance and programming challenges have been met or will be met soon, but the prediction challenge will require considerable advancement and maturity in the way that simulation is done and interpreted.

The prediction challenge, which is at the heart of computational engineering, is viewed as the most difficult challenge to be met in the future, and stands as a major bottleneck, perhaps a crisis, in computational engineering. Again, the

major issue is the reliability of computer predictions and their use as a basis for important decisions [Babuška & Oden, 2005].

It is understandable that the reliability of computer predictions has been an issue of growing concern for several decades. Major questions that arise with increasing frequency are "can computer predictions be used as a reliable base for crucial decisions? How can one assess the accuracy or validity of a computer prediction? What confidence can be assigned to a computer prediction of a complex event?". Professor E. Wilson, who has been responsible for the development of several computer programs, talking about structural engineering, collected all these questions giving a nice definition: Structural Engineering is:

The art of using materials
That have properties which can only be estimated
To build real structures
That can only be approximately analyzed
To withstand forces
That are not accurately known
So that our responsibility with respect to Public safety is satisfied

This little aphorism collects all the common greater uncertain about structural engineering. It is possible to extend it by including the idea that the computer nowadays is an essential powerful tool to use in structural design and modelling. So Computational Engineering is:

The art of simulate physical events
That have properties which can only be estimated
To develop mathematical models
That can only be approximately analyzed
To improve the solution reliability
That are not accurately known
So that our responsibility with respect to Public safety is satisfied

The collection of processes, philosophical issues, and procedures connected with answering these questions has become known as the subject of *Verification & Validation* (V&V), the verification process addressing the quality of the numerical treatment of the model used in the prediction and the validation process addressing the quality of the model.

V&V has emerged in recent years as the intellectual and technological discipline that addresses the prediction challenge. Both are processes, verification being the processes addressing the quality of the numerical approximation of the mathematical model used as the basis for a prediction, and validation being the process

addressing the reliability of the mathematical model as a faithful abstraction of reality. V&V has been the focus of much study and debate in recent years and a relatively large literature exists and is expanding.

Verification: the process of determining if a computational model obtained by discretizing a mathematical model of a physical event and the code implementing the computational model can be used to represent the mathematical model of the event with sufficient accuracy.

Validation: the process of determining if a mathematical model of a physical event represents the actual physical event with sufficient accuracy.

A typical example of a mathematical model is the set of equations and conditions characterizing a boundary value problem involving deterministic or stochastic differential equations together with functionals defining quantities of interest. These quantities of interest are the goals of the computer predictions, which, in turn, are the basis for decisions. Let us start with some primitive notions, depicted by Babuška and Oden in [Babuška & Oden, 2004].

Physical event: an occurrence in nature or in a physical system; a fundamental entity of a physical reality; a physical phenomenon. The dictionary (Merriam-Webster Collegiate Dictionary, 10th edition) indicates that an event is "something that happens". Thus, we are interested in something that happens as a physical reality; not for example, in behavioral aspects or trends in, for instance, sociological or economical systems.

Simulation: a process that build a likeness; in our case, a likeness produced by an interpretation of output from a computer or computational device.

Mathematical model (of a physical event): a collection of mathematical constructions that provide abstractions of a physical event consistent with a scientific theory proposed to cover that event.

Data of a mathematical model (of a physical event): factual information that defines the values or ranges of values of parameters in the mathematical model of a physical event.

Discretization: the process that transforms a mathematical model into a finite number of discrete components that can be processed by a digital computer.

Computational model: the discretized version of a mathematical model that has been designed to be implemented on (or to be processed by) a computer or computational device.

Code: A computer program designed (in the present context) to implement a computational model.

Prediction: something that is predicted, declared or indicated in advance; fore-

told on the basis of observation, experience, or scientific reason. A prediction is not simply a deduction or a consequence of a theory of something that mayor may not be known. It is the indication of an event not already known.

The mathematical problem is described by its structure and its input data. The structure of the mathematical problem comprises functional relations between the input and the output. For example, the structure can be expressed by a system of conservation laws. The input data is the set of all admissible data needed for solving the mathematical problem. For example, input data includes the classic boundary conditions and parameters used in the structure of the problem. The data also include the characterization of the uncertainty as when it is part of the mathematical problem.

In the literature the term mathematical problem and mathematical model are not distinguished, although the term model is often used in a more generic sense than mathematical, which involves all input data. We will not distinguish between the problem and the model either.

Verification therefore is the process of determining if a computational problem and the code implementing the computational problem leads to a prediction of sufficient accuracy, i.e., the difference between the exact and computed quantity of interest is sufficiently small. Hence, verification has two aspects, the approximation aspect and the verification of the correctness of the code, i.e., the program developed to implement the computational model can faithfully produce the intended results. Although code verification is obviously essential, we will not address it in the present thesis. The first part of verification, so-called solution verification, is essentially a problem of a-posteriori error estimation. It addresses not only classical methods of error estimation of standard approximation methods (such as finite elements) but also errors due to simplification of the problem. For example, errors due to the linearization of a nonlinear problem or of dimensional reduction.

A posteriori error estimation is a fairly mature subject and many techniques for developing a-posteriori error estimates have been proposed in the literature. A posteriori error estimation is a purely mathematical process and, while many open problems remain, effective methods exist for addressing solution verification for a large class of computational models.

Validation is the process of determining if the mathematical model of a physical event (the prediction) represents the actual physical event with sufficient reliability. In contrast to the verification, validation addresses the problem how well the theory describes reality. This question is related to a major problem in philosophy, especially in the philosophy of science. A first question is what is actually meant by validation and whether validation is even possible. In the validation procedure,

we consider a set of validation problems. These problems are specific mathematical problems for which some of the input data are the same as those in the prediction problem, but others may be different. For example, in the elasticity problem, the domain and the boundary conditions are different for the validation and the prediction problem, but the constitutive law is the same. These validation problems are simpler than the prediction problems and can, in general, be experimentally studied.

1.2.1 Idealization of a physical event

The goal of computer simulation is thus to make predictions of physical events using computational models implemented with appropriate codes. It is important to underscore the relationship between a mathematical model of an event and the scientific theory or theories used to characterize the event. As mathematics is, in a fundamental way, the language of science and technology, the mathematical model can be viewed as a transcription into a concrete and precise format of the theoretical framework in which the modeler intends the event to be depicted. A mathematical model may not necessarily be equivalent to a complete theory covering events in nature. Mathematical models may be the mathematical constructions representing the result of collections of assumptions and constraints on theoretical arguments and, therefore, may not describe all physical processes responsible for an event we are interested in predicting. This relationship between the mathematical model and the scientific theory provides an indirect connection between the philosophical underpinnings of V&V and major tenants of contemporary philosophy of science.

If the computational model describes the mathematical model well and the mathematical model relates to the theory well, then the computational model also relates well to the theory.

The models referred to in the definitions of verification and those in validation are, in general, quite different, as are, in some respects, the events they seek to emulate. Note that both the verification and validation processes involve determining if the respective processes lead to results of "sufficient accuracy", leaving open both the meaning of "sufficient" and of how "accuracy" is to be quantified.

1.2.1.1 Discrete model: some implications

The limitations of the human mind are such that it cannot grasp the behaviour of its complex surroundings and creations in one operation. Thus the process of subdividing all systems into their individual components or elements, whose behaviour is readily understood, and then rebuilding the original system from such components to study its behaviour is a natural way in which the engineer, or even the scientist proceeds.

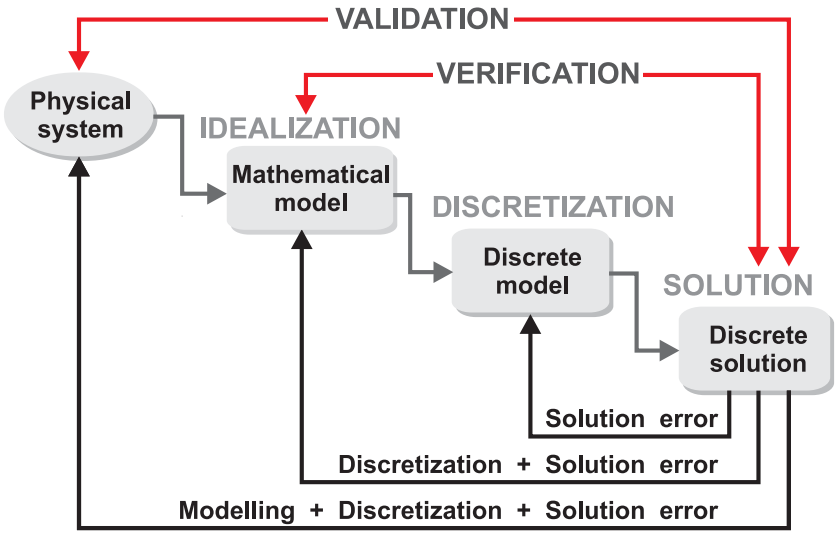


Figure 1.2: Verification and Validation processes

In many situations an adequate model is obtained using a finite number of well-defined components. We shall term such models discrete. In other situations the subdivision is continued indefinitely and the problem can only be defined using the mathematical fiction of an infinitesimal. This leads to differential equations or equivalent statements which imply an infinite number of elements. We shall term such models continuous.

With the advent of digital computers, discrete problems can generally be solved readily even if the number of elements is very large. As the capacity of all computers is finite, continuous problems can only be solved exactly by mathematical manipulation.

To overcome the intractability of realistic types of continuous problems, various methods of discretization have from time to time been proposed both by engineers and mathematicians. All involve an approximation which, hopefully, approaches in the limit the true continuous solution as the number of discrete variables increases.

The discretization of continuous problems has been approached differently by mathematicians and engineers. Mathematicians have developed general techniques applicable directly to abstract differential equations, such as finite difference approximations, various weighted residual procedures or approximate techniques for determining the stationarity of properly defined functionals. Engineers, on the other hand, often approach the problem more intuitively by creating analogies between real discrete elements and finite portions of a continuum domain. Figure 1.3 shows an interesting discretization study of a car, which is a very common physi-

cal system, for performing crashworthiness simulations. The funny thing for this case study is that for first the discretization has begun directly on the car, by a preliminary vehicle tear-down. Actually in this study the non-structural trim and interior components have been removed. Tape on vehicle is a coordinate system used to measure the positions of removed components. The resulting measured surfaces were then modified to an appropriate format for use in the vehicle model mesh generation program. This is also an attempt of model validation using both the component tests and full vehicle crash tests as shown in [Kirkpatrick *et al.* , 1998].

1.2.1.2 Discretization methods

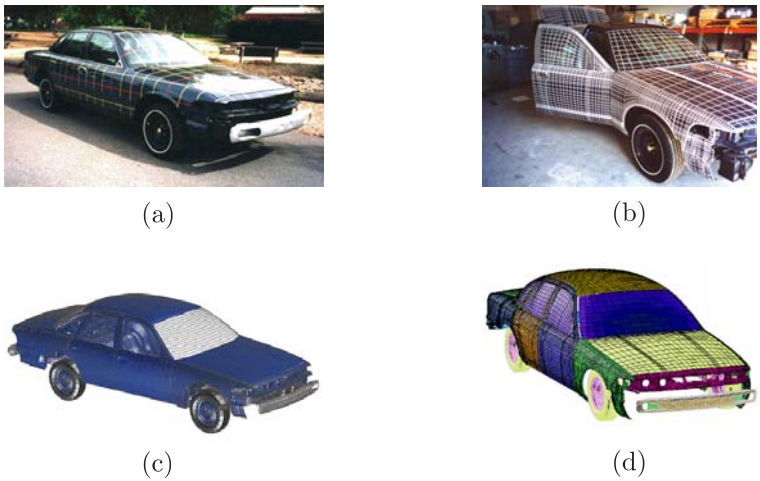
A typical classification of computational analysis is based on the discretization method by which the continuum mathematical model is discretized (in space or/and in time), i.e., converted to a discrete model with a finite number of degrees of freedom. Some of the most famous methods are:

(a) Finite Element Method (FEM),
(b) Boundary Element Method (BEM),
(c) Finite Difference Method (FDM),
(d) Finite Volume Method (FVM),
(e) Spectral Method,
(f) Meshfree Method.

In computational solid and structural mechanics the finite element method currently dominates the scene as regards space discretization of linear problems. Boundary element method posts a strong second choice in specific application areas. For nonlinear problems the dominance of the finite element method is overwhelming.

Finite difference methods for space discretization in solid and structural mechanics have virtually disappeared from practical use. This statement is not true, however, for fluid mechanics, where finite difference discretization methods are still important. Finite volume methods, which directly address the discretization of conservation laws, are important in difficult problems of fluid mechanics, for example high-Reynolds gas dynamics. Spectral methods are based on transforms that map space and/or time dimensions to domains (for example, the frequency domain) where the problem is easier to solve.

A recent newcomer to the scene is the brood class of meshfree methods. These combine techniques and tools of finite element methods, such as variational formulation and interpolation, with finite difference features, such as non-local support.



(a) (b)

(c) (d)

Figure 1.3: Model of a Ford Crown Victoria for performing crashworthiness simulations: (a) physical system; (b) discretization process; (c) discrete model; (d) subdomain subdivision (http://www.arasvo.com/crown_victoria/crown_vic.htm)

1.2.2 Analysis process by the finite element method

Processes using the Finite Element Method involve carrying out a sequence of steps in some way. This sequence takes two canonical configurations, depending on (i) the environment in which FEM is used and (ii) the main objective of the analysis: model-based simulation of physical systems, or numerical approximation to mathematical problems. Both are reviewed below to introduce terminology used in the sequel. Historically the model based was the first one to be developed to model complex physical systems. The second came later and, among other things, provided the necessary theoretical underpinnings to extend the method beyond structural analysis.

A glance at the schematics of a train wreck shown in Figure 1.1 makes obvious the reasons behind the necessity of simplifications. There is no simple differential equation that captures, at a continuum mechanics level, the train structure, coal, cargo, and passengers eating dinner [Felippa, 2004]. There is no reason for despair, however. The time honored *divide and conquer* strategy [1], coupled with abstrac-

[1] *Divide et impera*: the phrase is attributed to Philip II, king of Macedon (382-336 BC), describing his policy toward the Greek city-states. In computer science, divide and conquer (D&C) is an important algorithm design paradigm. It works by recursively breaking down a problem into two or more sub-problems of the same (or related) type, until these become simple enough to be solved directly. The solutions to the sub-problems are then combined to give a solution to the original problem. A divide and conquer algorithm is closely tied to a type of recurrence relation between functions of the data in question; data is "divided" into smaller portions and the result calculated thence.

tion, comes to the rescue. First, separate the structure out and view the rest as masses and forces, most of which are time-varying and nondeterministic. Second, consider the train structure as built of substructures (a part of a structure devoted to a specific function): wheel, wheel bars, steam boiler, coal box and so on. Take each substructure, and continue to break it down into components: rings, ribs, spars, cover plates, etc., continuing through as many levels as necessary.

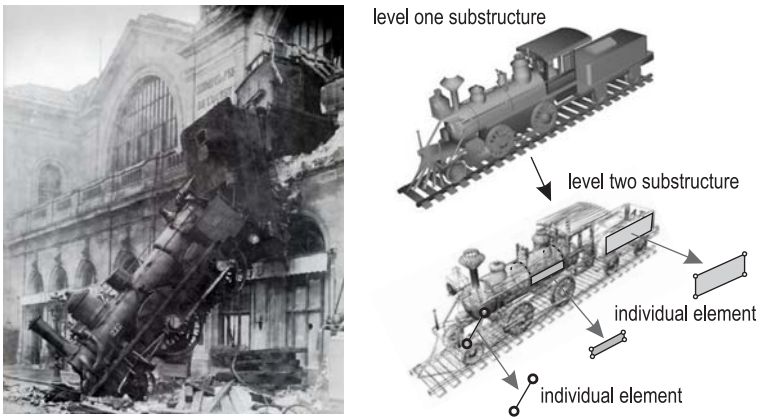


Figure 1.4: *Divide and conquer* (divide et impera) strategy: (a) physical system; (b) idealization process

Eventually those components become sufficiently simple in geometry and connectivity that they can be reasonably well described by the continuum mathematical models provided, for instance, by Mechanics of Materials or the Theory of Elasticity. At that point, stop. The component level discrete equations are obtained from a finite element method library based on the mathematical model.

The system model is obtained by going through the reverse process: from component equations to substructure equations, and from those to the equations of the complete locomotive.

This system assembly process is governed by the classical principles of Newtonian mechanics, which provide the necessary "component glue." The multilevel decomposition process is diagramed in Figure 1.4.

1.2.3 Interpretations of the finite element method

As shown in the previous section there are two practicable way to study and use the finite element method. Model based simulation has been shaped by the

discovery and extensive use of the method in the field of structural mechanics. This historical connection is reflected in the use of structural terms such as "stiffness matrix", "force vector" and "degrees of freedom". This terminology carries over to non-structural applications.

The basic concept in the physical interpretation is the breakdown (that mean disassembly, tearing, partition, separation, decomposition) of a complex mechanical system into simpler, disjoint components called finite elements, or simply elements. The mechanical response of an element is characterized in terms of a finite number of degrees of freedom. These degrees of freedoms are represented by the values of the unknown functions in a set of node points. The element response is defined by algebraic equations constructed from mathematical or experimental arguments. The response of the original system is considered to be approximated by that of the discrete model constructed by connecting or assembling the collection of all elements.

The breakdown-assembly concept occurs naturally when an engineer considers many artificial and natural systems. For example, it is easy and natural to visualize a building, bridge, or a metallic frame as being fabricated from simpler parts.

As discussed in the previous paragraph, the underlying theme is divide and conquer. If the behavior of a system is too complex, the recipe is to divide it into more manageable subsystems. If these subsystems are still too complex the subdivision process is continued until the behavior of each subsystem is simple enough to fit a mathematical model that represents well the knowledge level the analyst is interested in. In the finite element method such "primitive pieces" are called elements. The behavior of the total system is that of the individual elements plus their interaction. A key factor in the initial acceptance of the FEM was that the element interaction can be physically interpreted and understood in terms that were eminently familiar to structural engineers.

On the other hand there are the mathematical interpretations. The method is viewed as a procedure for obtaining numerical approximations to the solution of boundary value problems (BVPs) posed over a domain. This domain is replaced by the union of disjoint subdomains called finite elements.

The unknown function (or functions) is locally approximated over each element by an interpolation formula expressed in terms of values taken by the function(s), and possibly their derivatives, at a set of node points generally located on the element boundaries. The states of the assumed unknown function(s) determined by unit node values are called shape functions. The union of shape functions "patched" over adjacent elements form a trial function basis for which the node values represent the generalized coordinates. The trial function space may be

inserted into the governing equations and the unknown node values determined by the Ritz method (if the solution extremizes a variational principle) or by the Galerkin, least-squares or other weighted-residual minimization methods if the problem cannot be expressed in a standard variational form.

1.2.4 Historical sketch

This historical sketch is intended to outline an ideal common way about the finite elements method, by thinking back the experience of the most famous author. A first idea about finite element method may be associate to method of exhaustion developed by Eudoxus from Cnido (408–355 A.C.) (a very far-off start point) calculating an area by approximating it by the areas of a sequence of polygons.

J.W. Strutt (Lord Rayleigh)

Walter Ritz

Lately Archimedes from Syracuse (287-212 A.C.) generalized the method and went on to use Eudoxus's method of exhaustion to prove a remarkable collection of theorems. A first important contribution was brought by Lord Rayleigh, whose real name was John William Strutt, after revised by W. Ritz in 1909. He developed an effective method [Ritz, 1909] for the approximate solution of problems in the mechanics of deformable solids. It includes an approximation of energy functional by the known functions with unknown coefficients. Minimization of functional in relation to each unknown leads to the system of equations from which the unknown coefficients may be determined. One of the main restrictions in the Ritz method is that functions used should satisfy to the boundary conditions of the problem.

In 1910 Richardson [Richardson, 1910] published the first paper for practical computations utilizing finite differences and which presents a general procedure by which errors can be estimated. The original Richardson method starts from the fact that errors in the solution of any finite difference scheme usually depend of some power of the mesh size used (or today the size of finite elements). The order of the error is thus usually given and depends as well on the solution method adopted. In 1915 Boris Grigorievich Galerkin published an article, in which he put forward an idea of differential equations boundary problems approximate solution method. He has applied his method to a big number of pivots and plates analysis problems. Some time before I.G. Bubnov developed a similar approach for the variational problems solution, which he interpreted as a variant of the Ritz method algorithm. The Galerkin method (or Bubnov-Galerkin method) with Galerkin's (or "weak")

differential equations problem statement form are known all over the world.

Boris Grigorievich Galerkin

Richard Courant

In 1943 Richard Courant considerably increased the possibilities of the Ritz method by introducing of the special linear functions defined over triangular regions and applied the method for the solution of torsion problems [Courant, 1943]. The values of functions in the node points of triangular regions were chosen as unknowns. Thus, the main restriction of the Ritz functions - a satisfaction to the boundary conditions was eliminated. The name "finite element method" was not due to Courant, however, but appears only in the 1960s. The Ritz method together with the Courant modification is similar with the method proposed independently by Clough many years later introducing for the first time in 1960 the term "finite element" in the paper "The finite element method in plane stress analysis" [Clough, 1960]. Clough baptized the method and went on to form at Berkeley the first research group that expanded the idea into Civil Engineering applications.

John H. Argyris

Ray Clough

Edward L. Wilson

Now, for exposition convenience, structural "finitelementology" may be divided into four generations that span 10 to 15 years each, as suggested by C. Felippa [Felippa, 2004].

The First generation, which spans 1950 through 1962, formed by the pioneers, starts from the coming of word finite element method. The 1956 paper by Turner, Clough, Martin and Topp [Turner *et al.* , 1956], henceforth abbreviated to TCMT, is recognized as the start of the current FEM, as used in the overwhelming majority of commercial codes. Indeed the finite element procedure as we know it today has its origin in this work, which differs from the previous concepts in establishing arbitrary-shaped elements directly as pieces of the original domain. The original process suggested by them assumed specific stress distribution in each element equilibrating nodal forces.

The main reason of wide spreading of FEM in 1960 is the possibility to use computers for the big volume of computations required by FEM. However, Courant

did not have such possibility in 1943, in fact he employ his students to compute the calculus of its methods: by disposing the students on a mesh, so that each students working as a node can spread the calculus, he can compute the simple example presented in his work. It interesting to note that each student was paid proportionally to each calculus.

Finally, it should be remembered that an important contribution was brought into FEM development by the papers of Argyris [Argyris, 1954]-[Argyris, 1955].

The pioneers were structural engineers, schooled in classical mechanics. They followed a century of tradition in regarding structural elements as a device to transmit forces. This "element as force transducer" was the standard view in pre-computer structural analysis. It explains the use of flux assumptions to derive stiffness equations in TCMT. Element developers worked in, or interacted closely with, the aircraft industry. Accordingly they focused on thin structures built up with bars, ribs, spars, stiffeners and panels. Although the Classical Force Method dominated stress analysis during the 1950s, stiffness methods were kept alive by use in dynamics and vibration. It is a pity that, not being an engineer, Courant did not link the idea with networks of discrete elements and that, for this reason maybe, his work had to lie in obscurity for so many years.

Olgierd C. Zienkiewicz

Robert L. Taylor

The next period, second generation, spans the golden age of FEM: 1962-1972. This is the *variational generation.* Melosh showed that conforming displacement models are a form of Rayleigh-Ritz based on the minimum potential energy principle. This influential paper marks the confluence of three lines of research: Argyris's dual formulation of energy methods, the Direct Stiffness Method (DSM) of Turner, and early ideas of interelement compatibility as basis for error bounding and convergence. From 1962 onward a two-step interpretation emerges: discrete elements approximate continuum models, which in turn approximate real structures.

In 1963 E. L. Wilson and R. Clough developed, using the FORTRAN language, the SMIS software (Symbolic Matrix Interpretive System) based on the matrix analysis, for static and dynamic structural analysis. After, in 1969, Wilson was responsible for the development of several computer programs which are extensively used by professionals in the Civil, Mechanical and Aerospace engineering. The general three-dimensional finite element analysis program SAP and the TABS (after called ETABS) series of programs for the static and dynamic analysis of

three-dimensional building systems are examples of programs initially developed by Wilson. He was the first to develop computational methods and practical computer programs based on matrix notation.

Bruce Irons

John Barlow

Olek Zienkiewicz, originally an expert in finite difference methods who learned the trade from Southwell, was convinced in 1964 by Clough to try FEM. He went on to write with Cheung the first textbook [Zienkiewicz & Cheung, 1967] on the subject and called "The finite element method in structural and continuum mechanics". This book presents the broad interpretation of the method and its applicability to any general field problem. Zienkiewicz organized another important Civil Engineering research group at the University of Wales at Swansea.

In 1967 John Barlow at Rolls-Royce Aero Engine Division, discovered the optimal stress sampling points, after usually called Barlow points, but he published his work only in 1976, (as reported by his dear friend Bruce Irons in [Irons & Ahmadn, 1980]). This was the first UK FE system to exploit features such as isoparametric elements, the front solution, eigenvalue economisation, etc. After, in 1968, B. Irons and O.C. Zienkiewicz presented the isoparametric formulation of finite element stiffness matrices and this work had an immediate and significant impact on the finite element research being conducted at Berkeley.

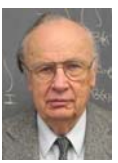
Thomas J. Hughes

Klaus J. Bathe

Professor Taylor was the first to program this new formulation at Berkeley and to demonstrate the power of this new type of element formulation. In a short time software and structural code were available: STRESS (1964) [Fenves *et al.* , 1964], STRUDL (1966), ASTRAN (1969), SAP (1970) in the United States of America, SAMCEF (1965) and ASKA (1969) in Europe. Low order displacement models often exhibit disappointing performance. Thus there was a frenzy to develop higher order elements. Other variational formulations, notably hybrids, mixed and equilibrium models emerged. In this second generation can be placed the monograph of Strang and Fix [Strang & Fix, 1973], the first book to focus on the mathematical

foundations.

A third generation, based on consolidation of the method, put substantial effort into improving the stock of second generation displacement elements by tools initially labeled *variational crimes*, but later justified. Textbooks by Hughes [Hughes, 1987] and Bathe [Bathe, 1996] reflect the technology of this period. Hybrid and mixed formulations record steady progress. Assumed strain formulations appear. The residual based error estimators were introduced by the pioneering work of Babuška and Rheinboldt in 1978 [Babuška & Rheinboldt, 1978], and further developed by many other authors. In fact, during the early 1980s the search for effective adaptive methods led to a wide variety of ad hoc error estimators. Many of these were based on a priori or interpolation estimates, that provided crude but effective indications of features of error sufficient to drive adaptive processes. A large activity in error estimation and mesh adaptivity is fostered by better understanding of the mathematical foundations.



Ivo M. Babuška

J. Tinsley Oden

Commercial FEM codes gradually gain importance. They provide a reality check on what works in the real world and what does not. By the mid-1980s there was gathering evidence that complex and high order elements were commercial flops. Exotic gadgetry interweaved amidst millions of lines of code easily breaks down in new releases. Complexity is particularly dangerous in nonlinear and dynamic analyses conducted by novice users. A trend back toward simplicity starts.

The fourth generation begins by the early 1980s. More approaches come on the scene, notably the Free Formulation, orthogonal hourglass control, Assumed Natural Strain methods, stress hybrid models in natural coordinates, as well as variants and derivatives of those approaches. Although technically diverse the fourth generation approaches share two common objectives:

1. elements must fit into DSM-based programs since that includes the vast majority of production codes, commercial or otherwise;
2. elements are kept simple but should provide answers of engineering accuracy with relatively coarse meshes (these were collectively labeled "high performance elements" in 1989).

Zienkiewicz and Zhu in 1987 developed a simple error estimation technique that is effective for some classes of problems and types of finite element approximations.

Gradients of solutions obtained on a particular partition are smoothed and then compared with the gradients of the original solution to assess error.

This four generations reflects so the main steps done by scientist until 1990s. After, several develops have been done. In particular the dedication posed in adaptive procedures is to be noted. Actually, with the coming of faster computers, the adaptive finite element methods have become an important tool for the analysis of engineering structures,.and a very important feature of any adaptive code is a posteriori error estimation, i.e. the ability of the code to estimate the error in the computed engineering quantity of interest. By the early 1990s the basic techniques of a posteriori error estimation were established. Briefly two types of procedures are currently well-known, by engineering point of view, for deriving error estimators. They are either Residual based or Recovery based.

Here, substantial progress was made in 1993 with the introduction of so called residual equilibration by Ainsworth and Oden [Ainsworth & Oden, 1993].

B. Boroomand

Mark Ainsworth

On the other hand, the recovery based error estimators are more recent, having been first introduced by Zienkiewicz and Zhu in 1987 [Zienkiewicz & Zhu, 1987]. Again, these were extensively improved by the introduction of new recovery processes. Here, in particular, the so called SPR (or Superconvergent Patch Recovery) method introduced in 1992 by the same authors [Zienkiewicz & Zhu, 1992c]-[Zienkiewicz & Zhu, 1992a]-[Zienkiewicz & Zhu, 1992b] has produced a very significant improvement of performance of the Recovery based methods. Many others attempted further improvement [Wiberg & Abdulwahab, 1993]-[Blacker & Belytschko, 1994] but the simple procedure originally introduced remains still most effective. A viable recovery alternative, known by the acronym REP (Recovery by Equilibrium in Patches), was first presented by Boroomand and Zienkiewicz [Boroomand & Zienkiewicz, 1997b] and later improved by same authors [Boroomand & Zienkiewicz, 1997a].

In 2002 a new effective stress recovery procedure, called RCP (Recovery by Compatibility in Patches), and based on complementary energy has been presented by Ubertini in [Ubertini, 2004]. This procedure can be viewed as dual to REP procedure in a certain sense.

So we reached now the present days and the error estimation topic. We have quickly flown over the later improvements of finite element procedures, because

this will be the main topic of the present work, and will be discussed widely in the next chapters.

1.3 The main focus of this work

Nowadays the combination of this powerful modern computers with effective numerical procedures, particularly finite element method, have transformed what were once purely qualitative theories of mechanics and physics into effective bases for simulation of physical phenomena important in countless engineering applications and scientific predictions.

No matter how sophisticated and appropriate the mathematical models as an event, all computational results obtained using them involve numerical error. So the presence of numerical error in calculations has been a principal source of concern since the beginning of computer simulations of physical phenomena. The discretization process of transforming a continuum model of physical phenomena into one manageable by digital computers cannot capture all of the information embodied in models characterized by partial differential equations or integral equations. What is the approximation error in such simulations? How can the error be measured, controlled, and effectively minimized? These questions have confronted computational mechanicians, practitioners, and theorists alike since the earliest applications of numerical methods to problems in engineering, and science. The reliability of computer simulations has become one of the most critical subjects in computational science. Lately, much attention has been posed on procedures aimed at improving accuracy of finite element solutions. Recovered solutions can be directly used for most practical purposes and, in particular, for evaluating the finite element discretization errors, that is a verification of solution.

The subject of a posteriori error estimation for finite element approximation has now reached maturity. The emphasis has now shifted from the development of new techniques to the study of robustness of existing estimators and identifying limits on their performance.

This work is intended to provide a contribution to give an answer to these questions through a study of error estimation for finite element approximations of boundary value problems, from *a priori* to *a posteriori* error estimation.In particular the work put particular care on a recent a posteriori error estimator, RCP procedure, and the role that this can assume as an independent measure of the quality of the simulation under study. Actually solution verification of a computational model must, in general, be based on a posteriori estimates of the error. By definition, methods of a posteriori error estimation are post-processing techniques

that attempt to determine quantitative information on the actual numerical error in computed approximations of the target quantities of interest. Effective methods for such calculations exist and are used with increasing frequency in solution verification processes.


F. Ubertini


S. de Miranda


G. Castellazzi

We expect to obtain quantitative information on system performance and outputs, and this expectation put stringent demands on the sophistication and accuracy of computer simulations.

Chapter 2

Finite element analysis

In questo capitolo vengono illustrati gli aspetti fondamentali del metodo degli elementi finiti, presentato il problema modello (assunto lineare) e introdotta la definizione di errore di discretizzazione. Tutto ciò getta le basi per la stima dell'errore, la quale può essere affrontata a priori, ovvero prima di eseguire l'analisi per elementi finiti, oppure a posteriori, ovvero dopo aver eseguito l'analisi. Tali approcci verranno affrontati separatamente nei capitoli successivi.

2.1 The basic ideas of the finite element method

The finite element method is a computational technique for obtaining approximate solutions to the partial differential equations that arise in scientific and engineering applications. Rather than approximating the partial differential equation directly as with, e.g., the finite difference methods, the finite element method utilizes a variational problem that involves an integral of the differential equation over the problem domain.

The finite element method can be described in a few words. Suppose that the problem to be solved is in variational form, it may be required to find the function which minimizes a given expression of potential energy. This minimizing property leads to a differential equation (the Euler equation), but normally an exact solution is impossible and some approximation is necessary. The Rayleigh-Ritz-Galerkin idea is to choose a finite number of trial functions and among all their linear combinations to find the one which is minimizing. This is the Ritz approxi-

mation. The unknown weights are determined, not by a differential equation, but by a system of discrete algebraic equations which the computer can handle. The theoretical justification for this method is simple, and compelling: the minimizing process automatically seeks out the combination which is closest to the solution. Therefore, the goal is to choose trial functions which are convenient enough for the potential energy to be computed and minimized, and at the same time general enough to approximate closely the unknown solution.

The real difficulty is the first one, to achieve convenience and computability. In theory there always exists a set of trial functions which is complete, their linear combinations fill the space of all possible solutions as the trial functions increases, and therefore the Ritz approximations converge, but to be able to compute with them is another matter. This is what finite elements have accomplished. The underlying idea is simple. It starts by a subdivision of the structure, or the region of physical interest, into smaller pieces. These pieces must be easy for the computer to record and identify; they may be triangles or quadrilateral. Then within each piece the trial functions are given an extremely simple form (normally they are polynomials). Boundary conditions are infinitely easier to impose locally, along the edge of a triangle or quadrilateral, than globally along a more complicated boundary. The accuracy of the approximation can be increased, if that is necessary, but not by the classical Ritz method of including more and more complex trial functions. Instead, the same polynomials are retained, and the subdivision is refined. The computer follows a nearly identical set of instructions, and just takes longer to finish. In fact, a large scale finite element system can use the power of the computer, for the formulation of approximate equations as well as their solution, to a degree never before achieved in complicated physical problems.

The method was created by structural engineers, and it was not recognized at the start as an instance of the Rayleigh-Ritz principle. The subdivision into simpler pieces, and the equations of equilibrium and compatibility between the pieces, were initially constructed on the basis of physical reasoning. The later development of more accurate elements happened in a similar way; it was recognized that increasing the degree of the polynomials would greatly improve the accuracy, but the unknowns computed in the discrete approximation have always retained a physical significance. In this respect the computer output is much easier to interpret than the weights produced by the classical method.

The whole procedure became mathematically respectable at the moment when the unknowns were identified as the coefficients in a Ritz approximation and the discrete equations were seen to be exactly the conditions for minimizing the po-

tential energy. The effect was instantly to provide a sound theoretical basis for the method. As the techniques of constructing more refined elements have matured, the underlying theory has also begun to take shape [Strang & Fix, 1973].

The fundamental problem is to discover how closely piecewise polynomials can approximate an unknown solution. In other words, we must determine how well finite elements, which were developed on the basis of computational simplicity, satisfy the second requirement of good trial functions, to be effective in approximation. Intuitively, any reasonable function can be approached to arbitrary accuracy by piecewise linear functions. The mathematical task is to estimate the error as closely as possible and to determine how rapidly the error decreases as the number of pieces (or the degree of the polynomial within each piece) is increased.

2.1.1 Variants of the finite element method

In the analysis of problems of a discrete nature, a standard methodology has been developed over the years. The civil engineer, dealing with structures, first calculates force-displacement relationships for each element of the structure and then proceeds to assemble the whole by following a well-defined procedure of establishing local equilibrium at each node or connecting point of the structure. The resulting equations can be solved for the unknown displacements. Similarly, the electrical or hydraulic engineer, dealing with a network of electrical components (resistors, capacitances, etc.) or hydraulic conduits, first establishes a relationship between currents (flows) and potentials for individual elements and then proceeds to assemble the system by ensuring continuity of flows. The term Finite Element Method actually identifies a broad spectrum of techniques that share common features. Two subclassifications that fit well applications to structural mechanics are:

1. **FEM Formulation**: (a) Displacement; (b) Equilibrium; (c) Mixed; (d) Hybrid;
2. **FEM Solution**: (a) Stiffness; (b) Flexibility; (c) Mixed;

The existence of a unified treatment of *standard discrete problems* leads us to the first definition of the finite element process as a method of approximation to continuum problems such that

1. the continuum is divided into a finite number of parts (elements), the behaviour of which is specified by a finite number of parameters,
2. the solution of the complete system as an assembly of its elements follows precisely the same rules as those applicable to standard discrete problems.

It will be found that most classical mathematical approximation procedures as

well as the various direct approximations used in engineering fall into this category.

2.1.2 Discretization of the domain

Since the problem is usually defined over a continuous domain, the governing equations, with the exception of the essential boundary conditions, are valid for the entirety of, as well as for any portion of, that domain. This allows idealization of the domain in the form of interconnected finite-sized domains (elements) of different size and shape (see Figure 2.1). By doing this, certain approximations are introduced (e.g., cutting the corners, making curved lines straight and curved elements flat). Putting enough numbers of nodes between the elements (higher-order elements, etc.) also comes into the picture at this stage of the method. Here, one should be concerned with how well the idealized discrete domain represents the actual continuous domain. To a certain extent, it is true that the smaller elements (finer mesh) produce better results. But it is also true that the finer mesh results in a larger number of equations to be solved. The question then arises: What is the most efficient element type, size, and pattern? A partial answer to this question is given in the literature under the key word modeling. Adaptive processes or mesh refinements and automatic mesh generation are also techniques relevant to the discretization of the domain.

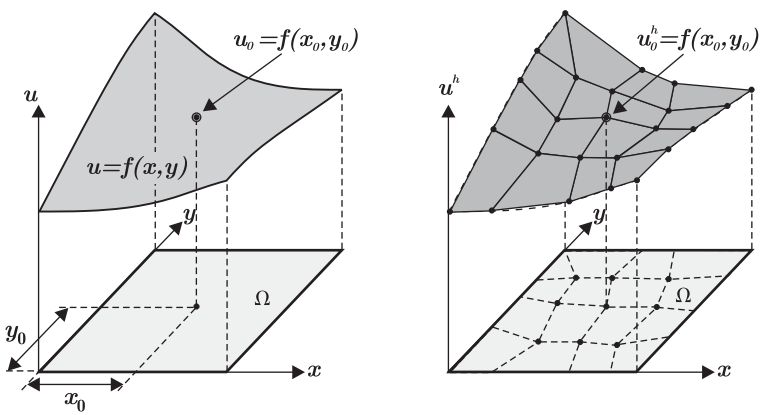


Figure 2.1: Discretization of a domain Ω

2.1.2.1 Discretization error

The real issue for us to grapple with now is that the computational model prepared to simulate the mathematical model may be faulty and can lead to errors. In the process of replacing the continuum region by finite elements, errors originate in

many ways. From physical intuition, we can argue that this will depend on the type, shape and number of elements that we use, the grading or density of the mesh used, the way distributed loads are assigned to nodes, the manner in which boundary conditions are modelled by specification of nodal degrees of freedom etc. These are the discretization errors that can occur.

We can assume that the approximation for the displacement $\mathbf{u}$ will yield the exact solution in the limit as the size h of elements decreases. The arguments for this are simple: if the expansion is capable, in the limit, of exactly reproducing any displacement form conceivable in the continuum, then, as the solution of each approximation is unique, it must approach, in the limit of $h \to 0$, the unique exact solution. In some cases the exact solution is indeed obtained with a finite number of subdivisions (or even with one element only) if the approximation used in that element fits exactly the correct solution. Thus, for instance, if the exact solution is of the form of a quadratic polynomial and the shape functions include all the polynomials of that order, the approximation will yield the exact answer.

The last argument helps in determining the order of convergence of the finite element procedure as the exact solution can always be expanded in the vicinity of any point (or node) i of coordinates $\mathbf{x}_0$ as a polynomial:

$$\mathbf{u}(\mathbf{x}) = \mathbf{u}(\mathbf{x}_0) + \nabla \mathbf{u}|_{\mathbf{x}=\mathbf{x}_0} (\mathbf{x} - \mathbf{x}_0) + \dots \tag{2.1}$$

If within an element of *size* h, a polynomial expansion of degree p is employed for $\mathbf{u}$, this can fit locally the Taylor expansion up to that degree and, as $(\mathbf{x} - \mathbf{x}_0)$ is of the order of magnitude h, the error in $\mathbf{u}$ will be of the order h^{p+1}.

2.1.3 Identification of the state variable(s)

Until this step, no reference has been made to the physical nature of the problem. Whether it is a heat-transfer problem, fluid- or solid-mechanics problem, etc., comes into the picture at this stage. The mathematical description of steady-state physical phenomena, for instance, leads to an elliptic boundary-value problem in which the formula contains the state variable and the flux. These variables are related to each other by a constitutive equation representing a mathematical expression of a particular physical law. Table 2.1 presents various physical problems with associated state variables and constitutive equations. Once the state variable and the flux have been identified, the formulation can take place containing either or both. The choice is usually dictated by the problem.

2.1.4 Formulation of the problem

Posing the problem to be solved in its most general terms we find that we seek an

Table 2.1: Classification of various physical problems

Physical Problem	State variable	Flux	Constitutive Equation
Elasticity	Displacements or forces	Stress or strain	Hooke's law
Torsion	Warping function	Rate of twist	Hooke's law
Heat Transfer	Temperature	Heat Flux	Fourier's law
Fluid Flow	Velocity	Shear stress	Stokes's law
Flow through porous media	Hydraulic head	Flow rate	Darcy's law
Electrostatics	Electric potential	Electric flux	Couloumb's law
Magnetostatic	Magnetic potential	Magnetic Flux	Maxwell's law

unknown function $\mathbf{u}$ such that it satisfies a certain differential equation set

$$\mathbf{A}(\mathbf{u}) = \begin{Bmatrix} A_1(\mathbf{u}) \\ A_2(\mathbf{u}) \\ \vdots \end{Bmatrix} = \mathbf{0}, \tag{2.2}$$

in a domain (volume, area, etc.) Ω (Figure 2.2), together with certain boundary conditions

$$\mathbf{B}(\mathbf{u}) = \begin{Bmatrix} B_1(\mathbf{u}) \\ B_2(\mathbf{u}) \\ \vdots \end{Bmatrix} = \mathbf{0}, \tag{2.3}$$

on the boundaries $\partial\Omega$ of the domain.

The function sought may be a scalar quantity or may represent a vector of several variables. Similarly, the differential equation may be a single one or a set of simultaneous equations and does not need to be linear. It is for this reason that we have resorted to matrix notation in the above.

The finite element process, being one of approximation, will seek the solution in the approximate form

$$\mathbf{u} \approx \mathbf{u}^h = \sum_{i=1}^{n} U_i d_i = \mathbf{U}\mathbf{d}, \tag{2.4}$$

where U_i are shape functions prescribed in terms of independent variables (such as the coordinates x_1, x_2, *etc.*) and all or most of the parameters d_i are unknown. As the set of differential equations (2.2) has to be zero at each point of the domain Ω, it follows that

$$\int_\Omega \mathbf{v}^{\mathrm{T}} \mathbf{A}(\mathbf{u})\, \mathrm{d}V \equiv \int_\Omega \left[v_1 A_1(\mathbf{u}) + v_2 A_2(\mathbf{u}) + \ldots\right] \mathrm{d}V = 0, \tag{2.5}$$

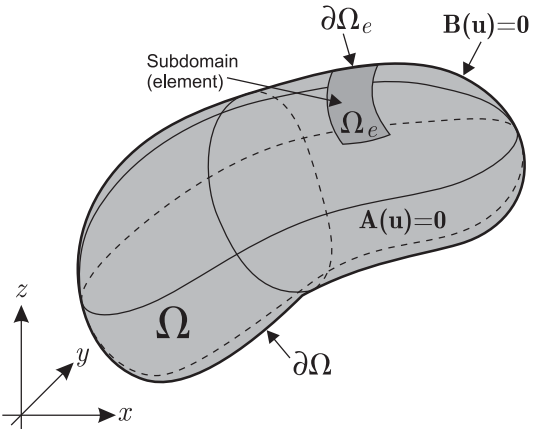


Figure 2.2: Problem domain Ω and boundary $\partial\Omega$

where

$$\mathbf{v} = \left\{ \begin{array}{c} v_1 \\ v_2 \\ \vdots \end{array} \right\}, \tag{2.6}$$

is a set of arbitrary functions equal in number to the number of equations (or components of $\mathbf{u}$) involved.

The statement is, however, more powerful. We can assert that if (2.5) is satisfied for all $\mathbf{v}$, then the differential equations (2.2) must be satisfied at all points of the domain. The proof of the validity of this statement is obvious if we consider the possibility that $\mathbf{A}(\mathbf{u}) \neq \mathbf{0}$ at any point or part of the domain. Immediately, a function $\mathbf{v}$ can be found which makes the integral of (2.5) non-zero, and hence the point is proved. If the boundary conditions are to be simultaneously satisfied, then we require that

$$\int_\Omega \bar{\mathbf{v}}^{\mathrm{T}} \mathbf{B}\,(\mathbf{u})\ \mathrm{d}S \equiv \int_\Omega \left[\bar{v}_1 B_1\,(\mathbf{u}) + \bar{v}_2 B_2\,(\mathbf{u}) + \ldots\right] \mathrm{d}S = 0, \tag{2.7}$$

for any set of functions $\bar{\mathbf{v}}$. Indeed, the integral statement that

$$\int_\Omega \mathbf{v}^{\mathrm{T}} \mathbf{A}\,(\mathbf{u})\ \mathrm{d}V + \int_{\partial\Omega} \bar{\mathbf{v}}^{\mathrm{T}} \mathbf{B}\,(\mathbf{u})\ \mathrm{d}S = 0, \tag{2.8}$$

is satisfied for all $\mathbf{v}$ and $\bar{\mathbf{v}}$ is equivalent to the satisfaction of the differential equations (2.2) and their boundary conditions (2.3).

In the above discussion it was implicitly assumed that integrals such as those in Equation (2.8) are capable of being evaluated. This places certain restrictions on the possible families to which the functions $\mathbf{v}$ or $\mathbf{u}$ must belong. In general we shall seek to avoid functions which result in any term in the integrals becoming infinite.

Thus, in Equation (2.8) we generally limit the choice of $\mathbf{v}$ and $\bar{\mathbf{v}}$ to bounded functions without restricting the validity of previous statements.

What restrictions need to be placed on the functions? The answer obviously depends on the order of differentiation implied in the equations $\mathbf{A}(\mathbf{u})$ [or $\mathbf{B}(\mathbf{u})$]. Consider, for instance, a function $\mathbf{u}$ which is continuous but has a discontinuous slope in a particular direction. We imagine this discontinuity to be replaced by a continuous variation in a very small distance. It is easy to see that although the first derivative is not defined here, it has finite value and can be integrated easily, but the second derivative tends to infinity. This therefore presents difficulties if integrals are to be evaluated numerically by simple means, even though the integral is finite. If such derivatives are multiplied by each other the integral does not exist and the function is known as non-square integrable. Such a function is said to be C^0 continuous. In a similar way it is easy to see that if nth-order derivatives occur in any term of $\mathbf{A}$ or $\mathbf{B}$ then the function has to be such that its $n-1$ derivatives are continuous (C^{n-1} continuity).

On many occasions it is possible to perform an integration by parts on Equation (2.8) and replace it by an alternative statement of the form

$$\int_{\Omega} \overline{\mathbf{A}}\,(\mathbf{v})^{\mathrm{T}}\,\overline{\overline{\mathbf{A}}}\,(\mathbf{u})\;\mathrm{d}V + \int_{\partial\Omega} \overline{\mathbf{B}}\,(\bar{\mathbf{v}})^{\mathrm{T}}\,\overline{\overline{\mathbf{B}}}\,(\mathbf{u})\;\mathrm{d}S = 0. \tag{2.9}$$

In this, the operators $\overline{\mathbf{A}}$, $\overline{\overline{\mathbf{A}}}$, $\overline{\mathbf{B}}$ and $\overline{\overline{\mathbf{B}}}$ usually contain lower order derivatives than those occurring in operators $\mathbf{A}$ or $\mathbf{B}$. Now a lower order of continuity is required in the choice of the function $\mathbf{u}$ at a price of higher continuity for $\mathbf{v}$ and $\bar{\mathbf{v}}$.

The statement (2.9) is now more *permissive* than the original problem posed by Equations (2.2), (2.3), or (2.8) and is called a weak form of these equations. It is a somewhat surprising fact that often this weak form is more realistic physically than the original differential equation which implied an excessive *smoothness* of the true solution. Integral statements of the form of (2.8) and (2.9) will form the basis of finite element approximations.

The finite element method is one particular Galerkin method, named after the Russian engineer Galerkin. It is a method for finding approximate solutions to partial differential equations. It is closely related to the Rayleigh-Ritz method which involves choosing functions (a basis) for the solution and finding the amplitude of each function by minimizing the energy. The Galerkin method is however more general, being able to solve a greater range of problems.

The essence of the Galerkin method involves taking the weak form of the governing Equation (2.8), and finding the best solution to a problem given a collection of functions.

First, consider the approximate solution to some problem, $\mathbf{u}^h \in \mathcal{S}^h$, where

$\mathcal{S}^h$ is a finite-dimensional space. This means that there is a limited number of possibilities. For example, $\mathbf{u}^h$ could be a combination of low order polynomial functions. The Galerkin problem for Equation (2.9) involves: find $\mathbf{u}^h \in \mathcal{S}^h$ such that

$$\int_\Omega \overline{\mathbf{A}}\left(\mathbf{v}^h\right)^{\mathrm{T}} \overline{\overline{\mathbf{A}}}\left(\mathbf{u}^h\right)\, \mathrm{d}V + \int_{\partial\Omega} \overline{\mathbf{B}}\left(\mathbf{v}^h\right)^{\mathrm{T}} \overline{\overline{\mathbf{B}}}\left(\mathbf{u}^h\right)\, \mathrm{d}S = 0, \qquad \forall \mathbf{v}^h \in \mathcal{V}^h, \tag{2.10}$$

where $\mathcal{V}^h$ is a finite dimensional space.

The Galerkin method (more specifically, the Bubnov-Galerkin method) requires that the weight and trial functions, $\mathbf{v}^h$ and $\mathbf{u}^h$ respectively, come from the same finite-dimensional space, taking into account the special requirements on the weight and trial functions where Dirichlet boundary conditions are applied. In Petrov-Galerkin method, the weight functions come from a different function space than the trial functions. This method is used for special applications, often in fluid mechanics.

Different Galerkin-based methods are defined by how the unknown field $\mathbf{u}^h$ is represented. In the finite element method, $\mathbf{u}^h$ and $\mathbf{v}^h$ will be simple continuous, piecewise low-order polynomials defined on *finite elements*. Spectral Galerkin methods for example use a truncated Fourier series as the basis. A basic question which arises when computing an approximate solution is how $\mathbf{u}^h$ relates to the exact solution $\mathbf{u}$. Given a finite number of possibilities in $\mathcal{S}^h$, which solution does the method seek? Understanding this requires some basic error analysis. The error analysis will tell how the computed solution $\mathbf{u}^h$ differs from the actual solution $\mathbf{u}$ and why the Galerkin method works (or for problems not considered here, why it does not work). This is examined in the following chapter.

2.2 The model problem

Consider the model problem governed by the following equations:

$$\mathbf{D}^*\boldsymbol{\sigma} = \mathbf{b}, \tag{2.11}$$

$$\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\epsilon}, \tag{2.12}$$

$$\boldsymbol{\epsilon} = \mathbf{D}\mathbf{u}, \tag{2.13}$$

in Ω, together with the following boundary conditions

$$\mathbf{u} = \bar{\mathbf{u}} \quad \text{on } \partial\Omega_u, \tag{2.14}$$

$$\mathbf{N}^{\mathrm{T}}\boldsymbol{\sigma} = \bar{\mathbf{t}} \quad \text{on } \partial\Omega_t. \tag{2.15}$$

In the above equations, $\mathbf{u}$, $\boldsymbol{\epsilon}$ and $\boldsymbol{\sigma}$ are the unknown vectors, $\mathbf{D}$ and $\mathbf{D}^*$ are linear differential operators, $\mathbf{C}$ is a symmetric positive definite matrix of physical

parameters, $\mathbf{b}$ is a prescribed source term, $\mathbf{N}$ is a matrix which contains the direction cosines of the outward unit normal vector on $\partial\Omega$, Ω is the domain and $\partial\Omega$ is the domain boundary. Operators $\mathbf{D}$ and $\mathbf{D}^*$ are assumed to be related by the Gauss-Green formula

$$\int_\Omega \mathbf{w}^{\mathrm{T}}\mathbf{D}\mathbf{v}\,\mathrm{d}V = \int_\Omega \mathbf{v}^{\mathrm{T}}\mathbf{D}^*\mathbf{w}\,\mathrm{d}V + \int_{\partial\Omega} \mathbf{v}^{\mathrm{T}}\mathbf{N}^{\mathrm{T}}\mathbf{w}\,\mathrm{d}S, \tag{2.16}$$

where $\mathbf{v}$ and $\mathbf{w}$ are vectors of sufficiently regular functions defined on Ω. Combining Equations (2.11)-(2.15), the boundary value problem can be put in the form

$$\mathbf{A}(\mathbf{u}) \equiv \mathbf{D}^*\mathbf{C}\mathbf{D}\mathbf{u} - \mathbf{b} = \mathbf{0}, \qquad \text{in } \Omega, \tag{2.17}$$

$$\mathbf{B}(\mathbf{u}) \equiv \mathbf{N}^{\mathrm{T}}\mathbf{C}\mathbf{D}\mathbf{u} - \bar{\mathbf{t}} = \mathbf{0}, \qquad \text{on } \partial\Omega_t. \tag{2.18}$$

The model problem stated above is typical of many self-adjoint problems in engineering. To fix the ideas, we may refer to a problem of elasticity so that $\mathbf{u}$ can be interpreted as the displacement vector, $\boldsymbol{\epsilon}$ as the strain vector, $\boldsymbol{\sigma}$ as the stress vector, $\mathbf{D}$ and $\mathbf{D}^*$ as the compatibility and equilibrium operators and $\mathbf{C}$ as the elasticity matrix. Then, Equations (2.11)-(2.13) are the equilibrium equation, the constitutive equation and the compatibility equation, respectively. More details about the structural model used in the present work are given in Appendix A.

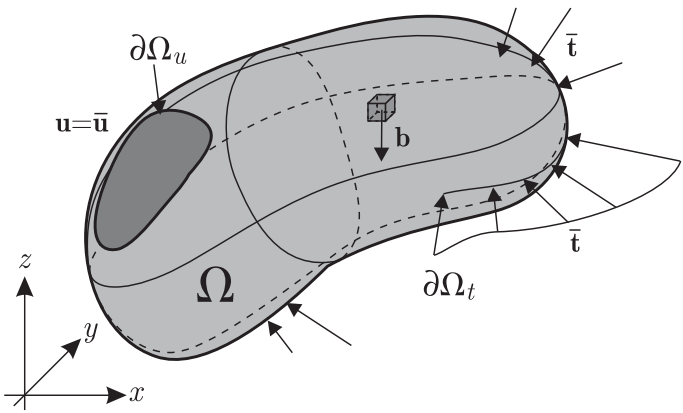


Figure 2.3: Problem domain Ω and boundary $\partial\Omega_t$ and $\partial\Omega_u$

The model problem admits weak formulations in terms of stationary of single or multifield functionals. In elasticity, the variational formulation in terms of displacements stems from the principle of minimum potential energy

$$\Pi(\mathbf{u}) = \frac{1}{2}\int_\Omega (\mathbf{D}\mathbf{u})^{\mathrm{T}}\,\mathbf{C}\mathbf{D}\mathbf{u}\mathrm{d}V - \int_\Omega \mathbf{u}^{\mathrm{T}}\mathbf{b}\mathrm{d}V - \int_{\partial\Omega_u} \mathbf{u}^{\mathrm{T}}\mathbf{t}\mathrm{d}S, \tag{2.19}$$

whose domain is the set of the functions $\mathbf{u}$ continuous in Ω, which satisfy the displacement boundary condition (2.14).

Assuming the stress as an additional variable leads to the Hellinger-Reissner functional

$$\Pi_{\mathrm{HR}}(\mathbf{u},\boldsymbol{\sigma}) = \int_{\Omega}\left(\boldsymbol{\sigma}^{\mathrm{T}}\mathbf{D}\mathbf{u} - \frac{1}{2}\boldsymbol{\sigma}^{\mathrm{T}}\mathbf{C}^{-1}\boldsymbol{\sigma} + \mathbf{b}^{\mathrm{T}}\mathbf{u}\right)\mathrm{d}V - \int_{\partial\Omega_t}\mathbf{u}^{\mathrm{T}}\mathbf{t}\mathrm{d}S, \tag{2.20}$$

whose domain is the set of functions $\mathbf{u}$ continuous in Ω, and functions $\boldsymbol{\sigma}$ fulfilling a priori the stress-strain relationship. If $\boldsymbol{\sigma}$ satisfies a priori the equilibrium equation (2.11), then the Hellinger-Reissner functional reduces to the hybrid stress functional

$$\Pi_{\mathrm{HY}}(\mathbf{u},\boldsymbol{\sigma}) = -\frac{1}{2}\int_{\Omega}\boldsymbol{\sigma}^{\mathrm{T}}\mathbf{C}^{-1}\boldsymbol{\sigma}\mathrm{d}V - \int_{\partial\Omega_t}\mathbf{u}^{\mathrm{T}}\mathbf{t}\mathrm{d}S. \tag{2.21}$$

Functionals (2.19) and (2.20) are the variational supports in developing compatible and mixed finite element models, respectively. While functional (2.21) is the variational support in developing hybrid stress model as well as the support of the so called Equilibrium Model II [de Veubeke, 1965]–[Pian & Tong, 1972].

Let's focus attention on the potential energy functional (2.19). Displacement $\mathbf{u}$, is approximated over each element Ω_e as follows

$$\mathbf{u} \simeq \mathbf{u}^h = \mathbf{U}(\mathbf{x})\,\bar{\mathbf{u}}^h, \tag{2.22}$$

where $\mathbf{U}$ is the matrix of shape functions and $\mathbf{d}$ is the vector of nodal displacement values. Once nodal displacements are determined, stresses (or gradients) are usually computed by means of Equations (2.12) and (2.13):

$$\boldsymbol{\sigma}^h = \mathbf{C}\mathbf{B}\mathbf{d}, \tag{2.23}$$

where

$$\mathbf{B} = \mathbf{D}\mathbf{U}. \tag{2.24}$$

It is well known that finite element stresses $\boldsymbol{\sigma}^h$ are in general discontinuous at inter-elements and have a low accuracy at nodes and element boundaries. The objective of a stress recovery procedure is to compute improved stresses by post-processing the finite element solution.

In the case of the mixed and hybrid finite element models based on the Hellinger-Reissner's functional (2.20) and the hybrid functional (2.21), respectively, $\mathbf{u}$ is represented, over each element, according to Equation (2.22), and stress $\boldsymbol{\sigma}$ is represented as

$$\boldsymbol{\sigma}^h(\mathbf{x}) = \mathbf{S}(\mathbf{x})\,\mathbf{s},$$

where $\mathbf{S}$ is a matrix of square integrable functions and $\mathbf{s}$ a vector of unknown stress parameters.

2.3 Definitions of error

Before proceeding further it is necessary to define what we mean by error. In principle, the error is the difference between the exact solution and the approximate one. This can apply to the basic function, such as displacement which we have called $\mathbf{u}$ and can be given as

$$\mathbf{e} = \mathbf{u} - \mathbf{u}^h. \tag{2.25}$$

In a similar way, however, we could focus on the error in the strains $\boldsymbol{\epsilon}$ or stresses $\boldsymbol{\sigma}$ (i.e., gradients of the solution), and define an error in those quantities as

$$\mathbf{e}_\varepsilon = \boldsymbol{\epsilon} - \boldsymbol{\epsilon}^h, \tag{2.26}$$

$$\mathbf{e}_\sigma = \boldsymbol{\sigma} - \boldsymbol{\sigma}^h. \tag{2.27}$$

The specification of local errors in the manner given in Equations (2.25)-(2.27) is not always convenient and occasionally misleading. For instance, under a point load both errors in displacements and stresses will be locally infinite but the overall solution may well be acceptable. Similar situations will exist near re-entrant corners where, as is well known, stress singularities exist in elastic analysis and gradient singularities develop in field problems. For this reason various *norms* representing some integral scalar quantity are often introduced to measure the error.

We can define an energy norm written for the error as

$$\|\mathbf{e}\| = \left[\int_\Omega \left(\boldsymbol{\epsilon} - \boldsymbol{\epsilon}^h\right)^{\mathrm{T}} \left(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h\right) \mathrm{d}V\right]^{\frac{1}{2}}, \tag{2.28}$$

$$= \left[\int_\Omega \left(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h\right)^{\mathrm{T}} \mathbf{C}^{-1} \left(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h\right) \mathrm{d}V\right]^{\frac{1}{2}}, \tag{2.29}$$

$$= \left[\int_\Omega \left(\mathbf{u} - \mathbf{u}^h\right)^{\mathrm{T}} \mathbf{D}^* \mathbf{C} \mathbf{D} \left(\mathbf{u} - \mathbf{u}^h\right) \mathrm{d}V\right]^{\frac{1}{2}}. \tag{2.30}$$

$$= \left[\int_\Omega \left(\boldsymbol{\epsilon} - \boldsymbol{\epsilon}^h\right)^{\mathrm{T}} \mathbf{C} \left(\boldsymbol{\epsilon} - \boldsymbol{\epsilon}^h\right) \mathrm{d}V\right]^{\frac{1}{2}}, \tag{2.31}$$

and its relation to strain energy is evident.

Other scalar norms can easily be devised. For instance, the L_2 norm of displacement and stress error can be written as

$$\|\mathbf{e}\|_{\mathrm{L}_2} = \left[\int_\Omega \left(\mathbf{u} - \mathbf{u}^h\right)^{\mathrm{T}} \left(\mathbf{u} - \mathbf{u}^h\right) \mathrm{d}V\right]^{\frac{1}{2}}, \tag{2.32}$$

$$\|\mathbf{e}_\sigma\|_{\mathrm{L}_2} = \left[\int_\Omega \left(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h\right)^{\mathrm{T}} \left(\boldsymbol{\sigma} - \boldsymbol{\sigma}^h\right) \mathrm{d}V\right]^{\frac{1}{2}}. \tag{2.33}$$

It is possible to evaluate root mean square (RMS) values of error. For instance,

the RMS error in displacement, $\Delta\mathbf{u}$,

$$|\Delta\mathbf{u}| = \left(\frac{\|\mathbf{e}\|^2_{\mathrm{L}_2}}{\Omega} \right)^{\frac{1}{2}} . \tag{2.34}$$

Similarly, the RMS error in stress, $\Delta\boldsymbol{\sigma}$, becomes for the domain Ω

$$|\Delta\boldsymbol{\sigma}| = \left(\frac{\|\mathbf{e}_\sigma\|^2_{\mathrm{L}_2}}{\Omega} \right)^{\frac{1}{2}} . \tag{2.35}$$

Any of the above norms can be evaluated over the whole domain or over subdomains or even individual elements.

We note that

$$\|\mathbf{e}\|^2 = \sum_{e=1}^{nelem} \|\mathbf{e}\|^2_e \,, \tag{2.36}$$

where e refers to individual elements Ω_e such that their sum (union) is Ω. We note further that the energy norm given in terms of the stresses, the L_2 stress norm and the RMS stress error have a very similar structure and that these are similarly approximated.

To describe the behaviour of stress analysis problems we define the variation of the relative energy norm error (percentage) as

$$\eta = \frac{\|\mathbf{e}\|}{E} \times 100 \quad (\%) \,, \tag{2.37}$$

where

$$E = \left(\int_\Omega \boldsymbol{\epsilon}^{\mathrm{T}} \mathbf{C} \boldsymbol{\epsilon} \mathrm{d}V \right)^{\frac{1}{2}} , \tag{2.38}$$

is the (exact) energy norm of the solution.

At the end of this chapter is clear that the analysis of the finite element method lead to a natural definition of error, by the attempt to compare the approximated quantity to the exact one. It appears by this definition, that this comparison is not always possible, due on the fact that the exact or true solution does not exist for all problems. Still actually the necessity to quantify this error by using different tools.

Commonly two are the practicable ways: the a priori error estimation and the a posteriori error estimation. In the next two chapters this two complementary ways, that can sometime interact, will be shown.

Chapter 3

A-priori error estimation

Nella prima parte di questo capitolo vengono brevemente illustrati alcuni risultati classici sulla stima a priori dell'errore di discretizzazione in analisi agli elementi finiti.

La seconda parte del capitolo si incentra, invece, sullo studio delle prestazioni che diversi elementi finiti mostrano in funzione del loro grado di approssimazione. In particolare, viene presentato un nuovo metodo di indagine [Castellazzi & Ubertini, 2004] per la valutazione del ruolo che le distorsioni di geometria hanno sulla prestazioni degli elementi finiti. L'indagine prende in esame elementi finiti parametrici piani, e ne quantifica la sensibilità alle distorsioni di geometria. L'indagine é inoltre accompagnata da un 'ampia sperimentazione numerica in cui si evidenzia come la progressiva distorsione della geometria influisca sulle prestazioni dell'elemento. Alcune considerazioni conclusive terminano il capitolo.

3.1 Analysis of the finite element method

Most of the errors introduced by the finite element methods are difficult to quantify analytically or determine in a logically coherent way. We can only rely on heuristic judgement (based on intuition and experience) to understand how best to minimize errors. However, we shall now look only at that category of discretization error that appears because the computational or discretized model uses trial functions which are an approximation of the true solution to the mathematical model. It seems possible that to some extent, analytical quantification of these errors is possible.

At the same time, a large body of literature has been created by mathematical analysts using the concepts of variational calculus and functional analysis to derive projection theorems and energy error rules on an a priori basis. These were first reported more than thirty years ago by Strang and Fix in [Strang & Fix, 1973]. Unfortunately, these statements on error analysis of the finite element method are derived using rigorous mathematical abstractions which are sometimes difficult for the engineer to grasp. It is indeed possible to re-derive these using the energy, virtual work and least action principles that the engineer or physicist is more familiar with, especially for linear elastostatics.

As shown in the previous chapter, the finite element method is one particular Galerkin method, and finds approximate solutions for partial differential equations. The essence of the Galerkin method involves taking the weak form of the governing equation, as developed in the previous chapter, and finding the best solution to a problem given a collection of functions. It is interesting to check how the solution computed using the Galerkin procedure compares to the exact solution.

Following the nomenclature used in Strang and Fix [Strang & Fix, 1973], we write the weak form in terms of the energy inner product for the exact solution $\mathbf{u}$ to the problem. Let be $\mathcal{V}$ a space of functions, which has an infinite number of members, and $\mathbf{u} \in \mathcal{V}$, then consider the following:

$$\int_{\Omega} (\mathbf{Du})^{\mathrm{T}} \mathbf{CDu} \mathrm{d}V = \int_{\Omega} \mathbf{u}^{\mathrm{T}} \mathbf{b} \mathrm{d}V \qquad \forall \mathbf{u} \in \mathcal{V}, \tag{3.1}$$

where the first term of Equation (3.1) is a measure of the total elastic strain energy. The vector $\mathbf{u}$ describes the degrees of freedom and the second term of the Equation (3.1) describes the potential of the applied loads. This virtual work statement refers to the exact solution of the elastostatic problem: the trial function and test function are taken as $\mathbf{u} \in \mathcal{V}$ and the virtual work argument establishes that (3.1) is truly satisfied only when $\mathbf{u}$ is the exact solution at the point of equilibrium. Consider now a subspace $\mathcal{V}^h \subset \mathcal{V}$ which has a finite number of members, then we can rewrite the Equation (3.1) as

$$\int_{\Omega} \left(\mathbf{Du}^h\right)^{\mathrm{T}} \mathbf{CDu} \mathrm{d}V = \int_{\Omega} \left(\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{b} \mathrm{d}V \qquad \forall \mathbf{u}^h \in \mathcal{V}^h. \tag{3.2}$$

In (3.2), we take note of the fact that the test function $\mathbf{u}^h \in \mathcal{V}^h$ need not be the exact displacement function for the virtual work principle to be true. For convenience, we take this to be the discrete finite element displacement field, as long as it is admissible (i.e. it satisfies all the geometric boundary conditions).

By using $\mathbf{u}^h$ for both the trial and test function, we get the actual finite element

equations, or we pose the Galerkin formulation, finding $\mathbf{u}^h \in \mathcal{V}^h$ such that

$$\int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}^h \mathrm{d}V = \int_\Omega \left(\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{b}\mathrm{d}V \quad \forall \mathbf{u}^h \in \mathcal{V}^h, \tag{3.3}$$

with the right hand side leading to the consistent load vector and the left hand side representing the stiffness matrix. This equation will now reflect the error due to the finite element discretization. We are now in a position to see how the error $\mathbf{e} = \mathbf{u} - \mathbf{u}^h$ can be assessed. Comparing (3.2) and (3.3) and noting that the energy inner product is bilinear, we can arrive at

$$\int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}\mathrm{d}V = \int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}^h \mathrm{d}V \quad \forall \mathbf{u}^h \in \mathcal{V}^h, \tag{3.4}$$

and from this we obtain the projection, or Cea's, theorem:

$$\int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\left(\mathbf{u} - \mathbf{u}^h\right) \mathrm{d}V = 0 \quad \forall \mathbf{u}^h \in \mathcal{V}^h. \tag{3.5}$$

This important result is commonly known also as *Galerkin orthogonality*. This means that the approximate solution $\mathbf{u}^h$ is a projection of the exact solution $\mathbf{u}$ onto the space of the weight functions. In the Bubnov-Galerkin, the weight functions $\mathbf{v}^h$ come from the same space as the trial functions $\mathbf{u}^h$, hence the solution $\mathbf{u}^h$ is the projection of the exact solution onto the finite dimensional space of trial functions. In simple linear elastostatics cases, this would imply that the strains or stresses are obtained in a best-fit sense and that there would be points in the element domain where these stresses or strains are very accurately computed. Consider the following:

$$\begin{aligned}
\int_\Omega \left[\mathbf{D}\left(\mathbf{u} - \mathbf{u}^h\right)\right]^{\mathrm{T}} \mathbf{C}\mathbf{D}\left(\mathbf{u} - \mathbf{u}^h\right) \mathrm{d}V &= \int_\Omega \left(\mathbf{D}\mathbf{u}\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}\mathrm{d}V + \\
&+ \int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}^h \mathrm{d}V - 2\int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}\mathrm{d}V \\
&= \int_\Omega \left(\mathbf{D}\mathbf{u}\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}\mathrm{d}V - \int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}^h \mathrm{d}V + \\
-2\left[\int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}\mathrm{d}V - \int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}^h \mathrm{d}V\right] \\
&= \int_\Omega \left(\mathbf{D}\mathbf{u}\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}\mathrm{d}V - \int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}^h \mathrm{d}V + \\
&\underbrace{- \int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\left(\mathbf{u} - \mathbf{u}^h\right) \mathrm{d}V}_{=\mathbf{0}},
\end{aligned}$$

for any $\mathbf{u}^h \in \mathcal{V}^h$. So we get an energy error theorem which can be expressed as:

$$\text{Energy of the error} = \text{Error of the energy}, \tag{3.6}$$

$$\left\|\mathbf{u} - \mathbf{u}^h\right\|_E = \|\mathbf{u}\|_E - \left\|\mathbf{u}^h\right\|_E, \tag{3.7}$$

where the subscript E indicates that the type of norm is in energy sense:

$$\left\|\mathbf{u}-\mathbf{u}^h\right\|_E = \int_\Omega \left[\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h\right)\right]^{\mathrm{T}} \mathbf{C}\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h\right) \mathrm{d}V, \tag{3.8}$$

$$\left\|\mathbf{u}\right\|_E = \int_\Omega \left(\mathbf{D}\mathbf{u}\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u} \mathrm{d}V, \tag{3.9}$$

$$\left\|\mathbf{u}^h\right\|_E = \int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}^h \mathrm{d}V. \tag{3.10}$$

As the left hand side of (3.6) is always positive definite, we get the useful statement that:

$$\int_\Omega \left(\mathbf{D}\mathbf{u}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u}^h \mathrm{d}V < \int_\Omega \left(\mathbf{D}\mathbf{u}\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{u} \mathrm{d}V. \tag{3.11}$$

Thus, in a variationally correct approach, the energy inner product of the approximate (Ritz or finite element) solution will always be a lower bound of the exact energy.

From these statements, we can set simple estimates for the error in a finite element discretization and, therefore, for the rate of convergence of the solution. It can be shown that the Galerkin finite element method is optimal in terms of the energy. The finite element solution is therefore seen to be a best-fit or best approximation solution. This error analysis tells something of what the Galerkin method calculates. Given some approximate functions, the Galerkin method will yield the best fit to the exact solution in terms of energy. Consider the following:

$$\begin{aligned}
\left\|\mathbf{u}-\mathbf{u}^h+\mathbf{v}^h\right\|_E &= \int_\Omega \left[\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h+\mathbf{v}^h\right)\right]^{\mathrm{T}} \mathbf{C}\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h+\mathbf{v}^h\right) \mathrm{d}V \\
&= \int_\Omega \left[\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h+\mathbf{v}^h\right)\right]^{\mathrm{T}} \mathbf{C}\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h\right) \mathrm{d}V + \\
&\quad + \int_\Omega \left[\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h\right)\right]^{\mathrm{T}} \mathbf{C}\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h+\mathbf{v}^h\right) \mathrm{d}V \\
&= \int_\Omega \left[\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h\right)\right]^{\mathrm{T}} \mathbf{C}\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h\right) \mathrm{d}V + \\
&\quad \underbrace{+2\int_\Omega \left[\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h\right)\right]^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{v}^h \mathrm{d}V}_{=\mathbf{0}} + \int_\Omega \left(\mathbf{D}\mathbf{v}^h\right)^{\mathrm{T}} \mathbf{C}\mathbf{D}\mathbf{v}^h \mathrm{d}V
\end{aligned}$$

for any $\mathbf{v}^h \in \mathcal{V}^h$. As seen before this last Equation was simplified by the Equation (3.5). Denoting now $\mathbf{w}^h = \mathbf{u}^h - \mathbf{v}^h$, the above result leads to the conclusion that:

$$\left\|\mathbf{u}-\mathbf{u}^h\right\|_E = \int_\Omega \left[\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h\right)\right]^{\mathrm{T}} \mathbf{C}\mathbf{D}\left(\mathbf{u}-\mathbf{u}^h\right) \mathrm{d}V, \tag{3.12}$$

$$\left\|\mathbf{u}-\mathbf{w}^h\right\|_E = \int_\Omega \left[\mathbf{D}\left(\mathbf{u}-\mathbf{w}^h\right)\right]^{\mathrm{T}} \mathbf{C}\mathbf{D}\left(\mathbf{u}-\mathbf{w}^h\right) \mathrm{d}V, \tag{3.13}$$

$$\left\|\mathbf{u}-\mathbf{u}^h\right\|_E \leq \left\|\mathbf{u}-\mathbf{w}^h\right\|_E \qquad \forall \mathbf{w}^h \in \mathcal{V}^h. \tag{3.14}$$

This implies that the solution $\mathbf{u}^h$ is closer to $\mathbf{u}$ than any other element of $\mathcal{V}^h$ in terms of energy. Given a choice of functions, the Galerkin method therefore chooses those which minimize the error in terms of the strain energy.

There is a close relationship between the Rayleigh-Ritz method and the Galerkin method. It has been shown that the Galerkin method minimizes the error in terms of the energy, which is the principle behind the Rayleigh-Ritz method. For many problems in solid mechanics, the two are equivalent.

3.2 Convergence of the finite element method

A priori error estimation does not provide quantitative information about error in a finite element simulation. What it does provide is the order of convergence. The order of convergence indicates how quickly the error reduces upon mesh refinement. The order depends on several factors. For finite element analysis, the order of convergence depends heavily on the type of element. We want to know how much smaller the error will be if we halve the element size (leading to at least a doubling in computational effort).

To give an indication of the size of the error $\mathbf{e} = \mathbf{u} - \mathbf{u}^h$, it is necessary to consider a norm of the error, $\|\mathbf{e}\|_n$. The subscript n indicates the type of norm. The energy norm, which is the type of norm which will be used here, is given by:

$$\|\mathbf{e}\|_n = \left(\sum_{\alpha=0}^{n} \int_\Omega \left(\mathbf{D}^\alpha \mathbf{e}\right) \mathbf{C} \left(\mathbf{D}^\alpha \mathbf{e}\right) \mathrm{d}V \right)^{1/2}, \tag{3.15}$$

where $\mathbf{D}^\alpha$ denotes the α derivative. This scalar value is a measure of the *magnitude* of the error. Two norms of particular interest involve $n = 0$ and $n = 1$. For example, setting $n = 0$ gives:

$$\|\mathbf{e}\|_0 = \left(\int_\Omega \mathbf{e}^{\mathrm{T}} \mathbf{C} \mathbf{e} \mathrm{d}V \right)^{1/2}, \tag{3.16}$$

which is a measure of the error in the displacements. Setting $n = 1$,

$$\|\mathbf{e}\|_1 = \left(\int_\Omega \left(\mathbf{e}^{\mathrm{T}} \mathbf{C} \mathbf{e} + \left(\mathbf{D}\mathbf{e}\right)^{\mathrm{T}} \mathbf{C} \mathbf{D} \mathbf{e} \right) \mathrm{d}V \right)^{1/2}, \tag{3.17}$$

which now includes the operator $\mathbf{D}$ on $\mathbf{e}$, thereby introducing a measure of the error in the strain.

Consider the nodal interpolate $\mathbf{u}^I$,

$$\mathbf{u}^I = \sum_i U_i a_i. \tag{3.18}$$

The nodal interpolate is equal to the exact solution at the nodes ($\mathbf{u} - \mathbf{u}^I = \mathbf{0}$ at the nodes), hence it interpolates the exact solution using the provided shape functions.

Given that a finite solution is the best possible solution from the finite-dimensional space $\mathcal{V}^h$, it holds that

$$\|\mathbf{e}\|_E \leq \left\|\mathbf{u} - \mathbf{u}^I\right\|_E .$$

This means that the finite element solution is at least as accurate as the nodal interpolate in terms of energy. In this way, the question of determining an error estimate is no longer a finite element question, rather a question which can be answered from interpolation theory. Using the Equation (3.5) we can state that:

$$\|\mathbf{e}\|_r \leq \alpha \left\|\mathbf{u} - \mathbf{u}^I\right\|_r , \tag{3.19}$$

where α is a positive constant. Interpolation theory provides the inequality

$$\left\|\mathbf{u} - \mathbf{u}^I\right\|_m \leq ch^{k+1-m} \|\mathbf{u}\|_E , \tag{3.20}$$

where k is the polynomial order, c is an unknown constant, and h is a *measure* of the element size. Note that to ensure convergence it should be: $k + 1 > m$. Inserting now Equation (3.19) into the above expression,

$$\|\mathbf{e}\|_r \leq Ch^{k+1-r} \|\mathbf{u}\|_E , \tag{3.21}$$

where C is a constant, independent of h. For an elasticity problem, $r = 1$, hence

$$\|\mathbf{e}\|_1 \leq Ch^k \|\mathbf{u}\|_E . \tag{3.22}$$

This implies that the solution converges in this norm at a rate equal to the polynomial order. For practical purposes, this norm is dominated by the error in the strain, hence it represents the error in the energy. If the exact solution is not sufficiently smooth, there maybe no gain in accuracy through increasing the polynomial order.

It would be useful to find an estimate in terms of lower norms, such as the error in the displacements, $\|\mathbf{e}\|_0$. An estimate can be obtained using *Nitsche's trick* (see [Strang & Fix, 1973] for details). It leads to the estimate:

$$\|\mathbf{e}\|_s \leq Ch^\beta \|\mathbf{u}\|_E , \tag{3.23}$$

where $\beta = \min\{k + 1 - s, 2(k + 1 - r)\}$.

In elasticity, the two quantities of special interest are the displacements and the strains (and hence the stresses). For $k \geq 1$, the displacement error is:

$$\|\mathbf{e}\|_0 \leq Ch^{k+1} \|\mathbf{u}\|_E , \tag{3.24}$$

and the strain error is:

$$\|\mathbf{e}\|_1 \leq Ch^k \|\mathbf{u}\|_E , \tag{3.25}$$

which implies that the approximate strains converge to the exact result slower than

the approximate displacements. For linear elements $k = 1$. Therefore:

$$\|\mathbf{e}\|_0 \leq Ch^2 \|\mathbf{u}\|_E , \tag{3.26}$$

$$\|\mathbf{e}\|_1 \leq Ch \|\mathbf{u}\|_E . \tag{3.27}$$

Hence linear elements are known as being $O\left(h^2\right)$ in terms of displacements, and $O\left(h\right)$ in terms of strains. The notation $O\left(h^a\right)$ means that the error is *of the order* h^a. Of course, as h becomes small, the greater a the smaller the error. Larger a implies faster convergence.

3.2.1 Richardson's theorem

This procedure may be applied to a sequence of displacement models only. Thus, for instance, if the displacement converges at $O(h^2)$ and we have two approximate solutions $\mathbf{u}_1$ and $\mathbf{u}_2$ obtained with meshes of size h and $\frac{h}{2}$, we can write, with $\mathbf{u}$ being the exact solution,

$$\frac{\mathbf{u}_1 - \mathbf{u}}{\mathbf{u}_2 - \mathbf{u}} = \frac{O\left(h^2\right)}{O\left(\frac{h}{2}\right)^2} = 4 \tag{3.28}$$

From the above, an (almost) exact solution $\mathbf{u}$ can be predicted and, then, an estimation $\|\mathbf{u}_R\|_E$ of the exact strain energy can be obtained..

Another method to obtain an estimation of the exact strain energy of the structure is to perform dual analyses [de Veubeke, 1965] in which the same problem is solved by using both displacement models and equilibrium models. Then if the displacement boundary conditions are homogeneous and consistent loads are applied, a displacement model gives a lower bound to the exact strain energy while an equilibrium model gives an upper bound. Richardson's extrapolation may then be applied to both sequences of displacement and equilibrium models so as to obtain two estimates which bound the exact strain energy.

3.3 Completeness requirements of shape functions

Convergence of finite element results with the mesh refinement is an important requirement to be satisfied by any successful element. The continuity and completeness requirements (e.g., see, Zienkiewicz [Zienkiewicz & Taylor, 1989], Bathe [Bathe, 1996]) must be satisfied so as to ensure convergence. The continuity requirement demands that the shape functions must be chosen so that the assumed interpolation function of the field variable, the displacement in structural analysis, is sufficiently continuous within and across the boundary of the elements. The completeness requirement demands that the shape functions must be such that the interpolation function is capable of representing rigid body motions and constant

strain states correctly.

The continuity requirement is usually satisfied by ensuring that the shape functions (and hence the displacement model) are polynomials of an appropriate order. Typically if the stiffness integrands involve derivatives of order m, the shape functions must be at least C^m continuous within the element and C^{m-1} continuous across the element interface. The completeness requirements are satisfied for a two dimensional isoparametric element, for examples if the element shape functions, U_i satisfy the conditions

$$\sum_i U_i = 1, \tag{3.29}$$

$$\sum_i U_i x_{1,i} = x_1, \tag{3.30}$$

$$\sum_i U_i x_{2,i} = x_2, \tag{3.31}$$

In the literature, the completeness requirements are frequently associated with the ability of the element to represent an arbitrary linear polynomial displacement field. For this reason, we will refer to these completeness requirements as the *linear completeness requirements*. In the limit of mesh refinement, the element size becomes so small that the displacement field within each element can be considered to be linear, and,. hence an element satisfying the linear completeness requirements is expected to converge to an exact solution without any problem.

It was once believed that elements satisfying the continuity and completeness requirements would indeed, in the limit of mesh refinement, converge to the actual solution without any difficulty. However, there have been instances which have contradicted this belief. Locking problems (e.g. see, [Prathap, 1993] and [Wilson *et al.* , 1973]) and the deterioration of solution accuracy with geometric distortion are two examples. To avoid the problem of locking, in addition to the continuity and completeness requirements, certain other conditions have to be satisfied by the displacement model [Prathap, 1993]. Nonconforming elements (e.g. see, [Zienkiewicz & Taylor, 1989], [Wilson *et al.* , 1973] and [Taylor *et al.* , 1976]) which violate the inter-element continuity requirements have been used successfully. However, such formulations are sensitive to geometric distortion of the element, and the patch test (see [Irons & Razzaque, 1972]–[Irons & Ahmadn, 1980]) bas been introduced to cheek the ill effects of the violation of inter-element continuity and completeness requirements in a global sense, i.e., for a patch. These observations are the source of motivation for several investigations into the continuity and completeness requirements.

It is frequently inconvenient to refine the mesh successively in order to check

for convergence. As an alternative to successive mesh refinement, higher order elements are sometimes used. In the context of the higher order elements, however, the completeness requirements need to be interpreted carefully. These elements use a higher order interpolation model, and attempt to obtain convergence using fewer elements. With fewer elements, the elements are larger and. hence, the linear completeness requirements are no longer sufficient as they ensure convergence only in the limit of mesh refinement. Thus, for the effective use of higher order elements, it is necessary to investigate the higher order completeness requirements of shape functions.

3.3.1 Constructing approximate functions for the elements

Once the state variable(s) and the local coordinate system have been chosen, the state variable(s) can be approximated in numerous ways, i.e. different choices for expression (2.4) are possibile. We mention here only the approximation in terms of algebraic functions. There are two entities that need to be approximated. The first is physical (the state variable) and the second is geometrical (the shape of the element). If the element is actually made of straight lines or planes, the coordinates of primary nodes (those at the extremes of the elements) will define the element shape accurately. In this case, the geometric approximation does not enter into the picture. Because of this, discretization of the entire domain is most often made by straight-line (linear) elements. For some problems, however, linear elements (i.e., flat elements in shells) may introduce unacceptable errors, and discretization must be done by using isoparametric elements (see Figure 3.1). A similar argument is, of course, valid for the approximation of the state variable. It can be approximated in the form of a linear function or a higher-order function (i.e., quadratic, cubic, etc.). The analyst then must decide whether to approximate physics (state variable) and geometry (element shape) equally or to give preference to one or the other in various regions of the domain.

This leads to three different categories of elements with m and n representing the degree of approximation for element shape and for the state variable, respectively. Figure 3.1 illustrates examples, related to plane case, for (a) superametric ($m > n$), (b) isoparametric ($m = n$), and (c) subparametric ($m < n$) elements. Approximation of any function can be made in the form of a family of algebraic equations, using the well-known Pascal's Triangle. At this point one is tempted to say that the finite element solution may converge to the exact solution either by increasing the degrees of polynomials (these often correspond to the number of nodes in the element) or by decreasing the element size. Each has advantages and disadvantages. Since, however, the finite element method is a procedure for

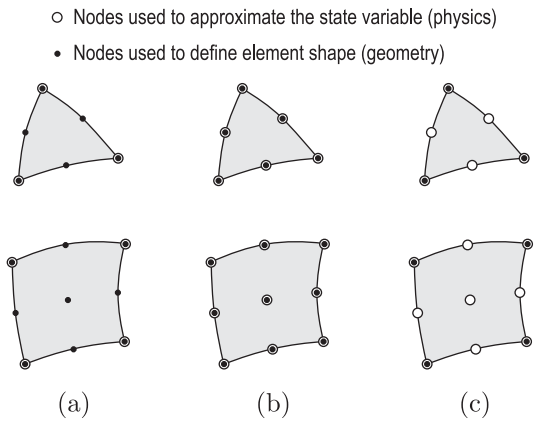


Figure 3.1: Example of superametric, isoparametric and subparametric elements

constructing the solution for the entire domain from local approximations, the convergence can be attained if the functional itself converges as the size of the element diminishes. It is therefore to be noted that:

1. the local functions have to be constructed so that their discontinuities (in terms of their derivatives as well) should not make the functional itself undefined over the entire domain. In other words, not only the local functions but the derivatives of one order less than that occurring in the functional must be continuous;
2. the integrand of the functional must be single-valued and represent a constant as the element size approaches zero.

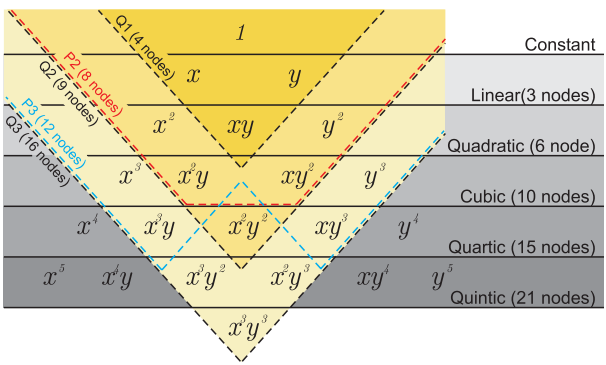


Figure 3.2: Pascal's triangle to generate various trial functions and corresponding elements

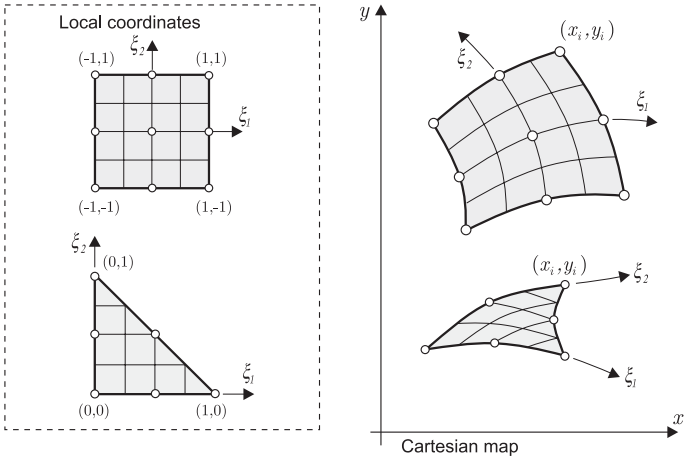


Figure 3.3: Two dimensional transformation between local or theory space (left) and the real or user space (right) of quadrilateral and triangular elements

3.3.2 Isoparametric quadrilateral elements

Is useful to start from the most simple isoparametric quadrilateral finite element, and extend the result to other elements after some study. The shape functions are polynomial representations of behaviour inside each element. This behavior is written in terms of local coordinates. Consider 2D behaviour: 3D is a straightforward extension. Although there are differences in detail in the theoretical derivations between triangles and quadrilaterals, the end result are equivalent.

The basic theory is developed in a square of side length 2 units, which can be termed the fundamental shape. This space is the theory space and has dimensionless coordinates (ξ_1, ξ_2). The real, or user space, has the usual (x_1, x_2) cartesian coordinates. These spaces are shown in Figure 3.3 for the quadrilateral and the triangular elements.

If the degree of freedom at each node i are the displacement components $\left(\bar{u}^h_{1,i}, \bar{u}^h_{2,i}\right)$, as used in stress and dynamics analysis, then the shape functions can express $\left(u^h_1, u^h_2\right)$ at any point in the element based on these nodal values as:

$$u^h_1 = \sum U_i \bar{u}^h_{1,i}, \qquad u_2 = \sum U_i \bar{u}^h_{2,i}, \tag{3.32}$$

$$u^h_1 = U_1 \bar{u}^h_{1,1} + U_2 \bar{u}^h_{1,2} + \ldots + U_n \bar{u}^h_{1,n} \tag{3.33}$$

$$u^h_2 = U_1 \bar{u}^h_{2,1} + U_2 \bar{u}^h_{2,2} + \ldots + U_n \bar{u}^h_{2,n} \tag{3.34}$$

where the summations are over the nodes on the element. A requirement of each of these shape functions is that they equal unity at their reference node and zero at every other node. Also the Equation (3.29) have to be verified throughout the element.

Table 3.1: Shape functions of 4, 8 and 9 noded quadrilaterals elements

coords		shape function	include only if node i exists on the element				
			$i=5$	$i=6$	$i=7$	$i=8$	$i=9$
-1	-1	$U_1{=}\frac{1}{4}(1-\xi_1)(1-\xi_2)$	$-\frac{1}{2}U_5$			$-\frac{1}{2}U_8$	$\frac{1}{4}U_9$
1	-1	$U_2{=}\frac{1}{4}(1+\xi_1)(1-\xi_2)$	$-\frac{1}{2}U_5$	$-\frac{1}{2}U_6$			$\frac{1}{4}U_9$
1	1	$U_3{=}\frac{1}{4}(1+\xi_1)(1+\xi_2)$		$-\frac{1}{2}U_6$	$-\frac{1}{2}U_7$		$\frac{1}{4}U_9$
-1	1	$U_4{=}\frac{1}{4}(1-\xi_1)(1+\xi_2)$			$-\frac{1}{2}U_7$	$-\frac{1}{2}U_8$	$\frac{1}{4}U_9$
0	-1	$U_5{=}\frac{1}{4}\left(1-\xi_1^2\right)(1-\xi_2)$					$-\frac{1}{2}U_9$
1	0	$U_6{=}\frac{1}{4}(1+\xi_1)\left(1-\xi_2^2\right)$					$-\frac{1}{2}U_9$
0	1	$U_7{=}\frac{1}{4}\left(1-\xi_1^2\right)(1+\xi_2)$					$-\frac{1}{2}U_9$
-1	0	$U_8{=}\frac{1}{4}(1-\xi_1)\left(1-\xi_2^2\right)$					$-\frac{1}{2}U_9$
0	0	$U_9{=}\frac{1}{4}\left(1-\xi_1^2\right)\left(1-\xi_2^2\right)$					

The shape functions for the 4, 8, and 9 noded quadrilateral elements are given in Table 3.1.

The performance of those elements will be evaluated in the next sections, for now it is interesting to pose the attention on Equation (2.24). After some manipulation, which in detail is not necessary here, it turns out that $\mathbf{B}$ can be written out in full as:

$$\mathbf{B} = \left[\begin{array}{cccc|c} g_{11} & 0 & g_{12} & 0 & 0 \\ 0 & g_{21} & 0 & g_{22} & g_{2n} \\ g_{21} & g_{11} & g_{22} & g_{12} & g_{1n} \end{array}\right] \tag{3.35}$$

where the three rows correspond to the three components of strain, and the columns cover the n nodes. A typical term applying to node k is:

$$g_{1k} = \left[\left(x_{2,i}\frac{\partial U_i}{\partial \xi_2}\right)\frac{\partial U_k}{\partial \xi_1} - \left(x_{2,i}\frac{\partial U_i}{\partial \xi_1}\right)\frac{\partial U_k}{\partial \xi_2}\right]\frac{1}{\det J} \tag{3.36}$$

$$g_{2k} = \left[-\left(x_{1,i}\frac{\partial U_i}{\partial \xi_2}\right)\frac{\partial U_k}{\partial \xi_1} + \left(x_{1,i}\frac{\partial U_i}{\partial \xi_1}\right)\frac{\partial U_k}{\partial \xi_2}\right]\frac{1}{\det J} \tag{3.37}$$

An important factor here is the inverse $\det J$ term. The scalar $\det J$ is the Jacobian transformation, effectively a numerical representation, or scaling, between the theory space and the real space of the element, over which the integration is performed.

$$\mathbf{J} = \left[\begin{array}{cc} \dfrac{\partial x_1}{\partial \xi_1} & \dfrac{\partial x_2}{\partial \xi_1} \\ \dfrac{\partial x_1}{\partial \xi_2} & \dfrac{\partial x_2}{\partial \xi_2} \end{array}\right] \tag{3.38}$$

$$\det J = \frac{\partial x_1}{\partial \xi_1}\frac{\partial x_2}{\partial \xi_2} - \frac{\partial x_1}{\partial \xi_2}\frac{\partial x_2}{\partial \xi_1} \tag{3.39}$$

A last note about the Jacobian lies on the fact that the most common finite element method is based on the principle of minimum potential energy, which means that when the structure is loaded and suitably fixed, the displacements and stresses are those which render the potential energy a minimum. Then the approximating equation system will yield a set of linear equations of the form

$$\mathbf{Kd} + \mathbf{f} = \mathbf{0}, \tag{3.40}$$

here, $\mathbf{K}$ is the global stiffness matrix and $\mathbf{u}$ assume the role of the total degree of freedom vector, and $\mathbf{f}$ is the corresponding vector of equivalent nodal loads. Each element contributes its own stiffness matrix $\mathbf{K}^e$, added into the global stiffness matrix, plus any contributions to the load vector

$$\mathbf{K}_{ij} = \sum_{e=1}^{nelem} \mathbf{K}_{ij}^e \qquad \mathbf{f}_i = \sum_{e=1}^{nelem} \mathbf{f}_i^e. \tag{3.41}$$

The element stiffness matrix is given by

$$\mathbf{K}^e = \int_{-1}^{1}\int_{-1}^{1} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{B} \det J \mathrm{d}\xi \mathrm{d}\eta \tag{3.42}$$

in terms of the element's 2D theory space.

When $\det J$ is no longer constant, the stiffness matrix polynomial becomes of infinite order, and a strictly accurate numerical integration is not available. This can be seen from Equation (3.42), whose right-hand side contains $\det J$ in the integrand.

The product which gives $\mathbf{K}^e$ is seen therefore to contain $1/\det J$ in the integrand, which is an infinite polynomial. We can term any element whose shape is such that $\det J$ is not constant over the element as distorted. Only undistorted elements can be accurately evaluated numerically.

If singularities exist in the mapping, $\det J$ becomes zero, corresponding to certain types of gross distortion, discussed in the next section. Large numerical errors could result. Another requirement of the mapping is that it should not be both positive and negative in the same element. If negative throughout the element, the mapping sign can be reversed by software (e.g. by reversing the topology order).

Over an element, if $\det J$ is not constant but a polynomial, then a useful check on the shape is to compare the maximum and minimum values of $\det J$, or their ratio. Hence, checking $\det J$ is a valuable element shape check.

To understand what are distortions it is useful to analyze the isoparametric maps of 4 and 8 node elements.

Any change in shape in the user space from the original square is called a dis-

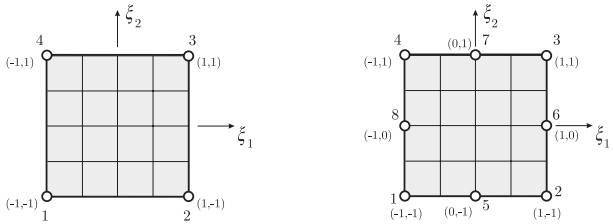


Figure 3.4: Reference coordinates for four-node and eight node element used for give physical meaning of coefficients $\alpha_{i,j}$

tortion. Several investigators have described the various shapes and their significance, see, for example, Burrows [Burrows & Enderby, 1993], Robinson [Robinson, 1987b]–[Robinson, 1988] and the references therein.

3.3.2.1 Four-node element in detail

The isoparametric map

$$x : [-1,1] \times [-1,1] \rightarrow B \subset \mathbb{R}^2$$

of a biunit on $\mathbb{R} \times \mathbb{R}$ onto a quadrilateral is defined for an 4-node element by

$$\mathbf{x} = \left\{ \begin{array}{c} x_1 \\ x_2 \end{array} \right\} = \sum_{i=1}^{4} U_i\left(\xi_1, \xi_2\right) \left\{ \begin{array}{c} x_{1i} \\ x_{1i} \end{array} \right\} = \tag{3.43}$$

$$= \left\{ \begin{array}{c} \alpha_{1,0} + \alpha_{1,1}\xi_1 + \alpha_{1,2}\xi_2 + \alpha_{1,3}\xi_1\xi_2 \\ \alpha_{2,0} + \alpha_{2,1}\xi_1 + \alpha_{2,2}\xi_2 + \alpha_{2,3}\xi_1\xi_2 \end{array} \right\} \tag{3.44}$$

$x_{1,i}$, $x_{2,i}$ are nodal coordinate of an element. Shape functions $U_i\left(\xi_1, \xi_2\right)$ for each node are given as follows:

$$U_i = \frac{1}{4}\left(1 + \xi_{1,i}\xi_1\right)\left(1 + \xi_{2,i}\xi_2\right) \qquad \text{for } i = 1,2,3,4$$

the coefficients $\alpha_{i,j}$, for $i = 1,2$ and $j = 0-3$, in terms of the nodal coordinates $x_{1,k}$, $k = 1-4$, are:

$$\alpha_{1,0} = \frac{1}{4}\left(x_{1,1} + x_{1,2} + x_{1,3} + x_{1,4}\right) \tag{3.45a}$$

$$\alpha_{1,1} = \frac{1}{4}\left(-x_{1,1} + x_{1,2} + x_{1,3} - x_{1,4}\right) \tag{3.45b}$$

$$\alpha_{1,2} = \frac{1}{4}\left(-x_{1,1} - x_{1,2} + x_{1,3} + x_{1,4}\right) \tag{3.45c}$$

$$\alpha_{1,3} = \frac{1}{4}\left(x_{1,1} - x_{1,2} + x_{1,3} - x_{1,4}\right) \tag{3.45d}$$

and similar $\alpha_{2,j}$, which are determined in terms of the coordinates $x_{2,k}$.

As show in [Robinson, 1987b] and whit respect to Equations (3.45) the physical significance of the $\alpha_{i,j}$ coefficients is shown in Figure 3.5. From Figure 3.5(a) it is clear that $\alpha_{1,0}$ and $\alpha_{2,0}$ define an origin (*translation* of axes), Figure 3.5(b)

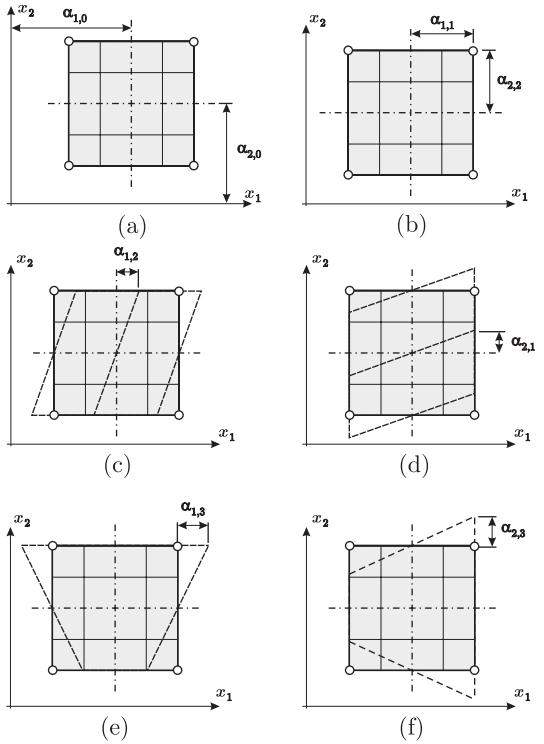


Figure 3.5: Physical meaning of the coefficients $\alpha_{1,i}$ and $\alpha_{2,i}$ for $i = 1, 2, 3, 4$

shows that $\alpha_{1,1}$ and $\alpha_{2,2}$ define the size of rectangle (*aspect ratio*), Figure 3.5(c) demonstrates that $\alpha_{1,2}$and $\alpha_{2,1}$ give two rotations (*skew* and *rotation* of axes), and Figure 3.5(d), (e) shows that $\alpha_{1,3}$ and $\alpha_{2,3}$ give two *tapers*. If the local axes are defined as in Figure **??** the shape parameter are:

$$\text{aspect ratio} = \frac{\max(\alpha_{1,1}, \alpha_{2,2})}{\min(\alpha_{1,1}, \alpha_{2,2})} \tag{3.46}$$

$$\text{skew} = \frac{\alpha_{1,2}}{\alpha_{2,2}} \tag{3.47}$$

$$\text{taper in the } x-\text{direction} = \frac{\alpha_{2,3}}{\alpha_{2,2}} \tag{3.48}$$

$$\text{taper in the } y-\text{direction} = \frac{\alpha_{1,3}}{\alpha_{1,2}} \tag{3.49}$$

3.3.2.2 Eight-node element in detail

The isoparametric map

$$x : [-1, 1] \times [-1, 1] \to B \subset \mathbb{R}^2 \tag{3.50}$$

of a biunit on $\mathbb{R} \times \mathbb{R}$ onto a quadrilateral is defined for an 8-node element by

$$\mathbf{x} = \left\{ \begin{array}{c} x_1 \\ x_2 \end{array} \right\} = \sum_{i=1}^{8} U_i(\xi_1, \xi_2) \left\{ \begin{array}{c} x_{1i} \\ x_{1i} \end{array} \right\} = \tag{3.51}$$

$$= \left\{ \begin{array}{l} \alpha_{1,0} + \alpha_{1,1}\xi_1 + \alpha_{1,2}\xi_2 + \ldots + \alpha_{1,5}\xi_2^2 + \alpha_{1,6}\xi_1^2\xi_2 + \alpha_{1,7}\xi_1\xi_2^2 \\ \alpha_{2,0} + \alpha_{2,1}\xi_1 + \alpha_{2,2}\xi_2 + \ldots + \alpha_{2,6}\xi_2^2 + \alpha_{2,6}\xi_1^2\xi_2 + \alpha_{2,7}\xi_1\xi_2^2 \end{array} \right\} \tag{3.52}$$

$x_{1,i}$, $x_{2,i}$ are nodal coordinate of an element. Shape functions $U_i(\xi_1, \xi_2)$ for each node are given as follows, for corner nodes:

$$U_i = \frac{1}{4}\left(1 + \xi_{1,i}\xi_1\right)\left(1 + \xi_{2,i}\xi_2\right)\left(\xi_{1,i}\xi_1 + \xi_{2,i}\xi_2 - 1\right) \qquad \text{for } i = 1, 2, 3, 4 \tag{3.53}$$

for boundary nodes:

$$U_i = \frac{1}{2}\left(1 - \xi_1^2\right)\left(1 + \xi_{2,i}\xi_2\right) \qquad \text{for } i = 5, 7 \tag{3.54}$$

$$U_i = \frac{1}{2}\left(1 + \xi_{1,i}\xi_1\right)\left(1 - \xi_2^2\right) \qquad \text{for } i = 6, 8 \tag{3.55}$$

the coefficients $\alpha_{i,j}$, for $i = 1, 2$ and $j = 0 - 7$, in terms of the nodal coordinates $x_{1,k}$, $k = 1 - 8$, are:

$$\alpha_{1,0} = -\frac{1}{4}\left(x_{1,1} + x_{1,2} + x_{1,3} + x_{1,4}\right) + \frac{1}{2}\left(x_{1,5} + x_{1,6} + x_{1,7} + x_{1,8}\right) \tag{3.56}$$

$$\alpha_{1,1} = \frac{1}{2}\left(x_{1,6} + x_{1,8}\right) \tag{3.57}$$

$$\alpha_{1,2} = \frac{1}{2}\left(-x_{1,5} + x_{1,7}\right) \tag{3.58}$$

$$\alpha_{1,3} = \frac{1}{2}\left(x_{1,1} - x_{1,2} + x_{1,3} - x_{1,4}\right) \tag{3.59}$$

$$\alpha_{1,4} = \frac{1}{2}\left(x_{1,1} + x_{1,2} + x_{1,3} + x_{1,4}\right) - \frac{1}{2}\left(-x_{1,5} + x7\right) \tag{3.60a}$$

$$\alpha_{1,5} = \frac{1}{2}\left(x_{1,1} + x_{1,2} + x_{1,3} + x_{1,4}\right) - \frac{1}{2}\left(-x_{1,6} + x_{1,8}\right) \tag{3.61}$$

$$\alpha_{1,6} = \frac{1}{2}\left(-x_{1,1} - x_{1,2} + x_{1,3} + x_{1,4}\right) + \frac{1}{2}\left(x_{1,5} - x_{1,7}\right) \tag{3.62}$$

$$\alpha_{1,7} = \frac{1}{2}\left(-x_{1,1} + x_{1,2} + x_{1,3} - x_{1,4}\right) + \frac{1}{2}\left(-x_{1,6} + x_{1,8}\right) \tag{3.63a}$$

and similarly $\alpha_{2,j}$, which are determined in terms of the coordinates $x_{2,k}$. An interpretation of the eight coefficients $\alpha_{1,i}$ and $\alpha_{2,i}$ for $i = 1, 2, 3, 4$, is shown in Figure 3.6.

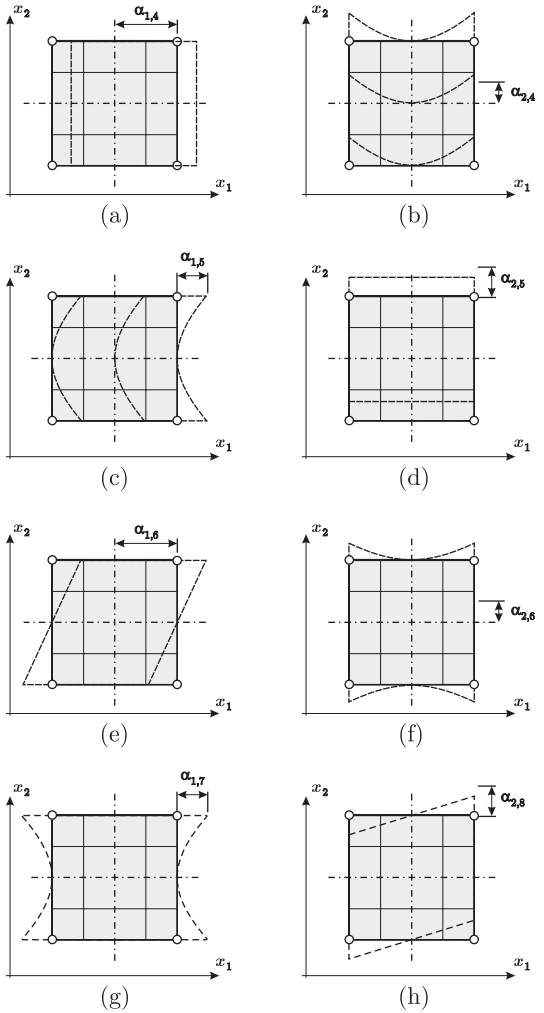


Figure 3.6: Physical meaning of the coefficients $\alpha_{1,i}$ and $\alpha_{2,i}$ for $i = 4, 5, 6, 7$

3.4 Numerical integration and optimal sampling points

The integration in Equation (3.42) cannot be performed analytically, because the equations are too complicated also for the simplest types of elements. Hence the famous numerical integration scheme called Gaussian quadrature is used.

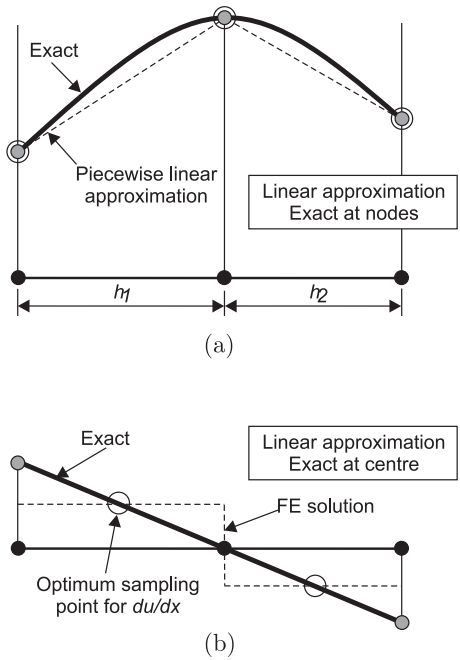


Figure 3.7: Optimal sampling points for the function (a) and its gradient (b) in one dimension (linear elements)

The scheme integrates a polynomial curve in 1D, which effectively calculates the area under the curve over the required interval. The rule order is the number of integration points, or Gauss points, needed in the interval to perform an accurate numerical integration. For each order, the Gauss points have to be situated at particular locations within the interval. For a rule with n such points, polynomials of up to and including order $2n-1$ are integrated accurately. In numerical analysis terms, Gaussian quadrature is very efficient, requiring less points than other methods of integration. In the limit of a straight line, $n = 1$, the integration calculates the area of a simple quadrilateral. Polynomials higher than $2n-1$ cannot be integrated accurately, which in finite element terms results in instabilities such

as zero energy modes. If a given polynomial is integrated accurately by a rule of order k, fit is also integrated accurately by all orders higher than k.

In our finite element calculation we often have a need for accurate estimates of the derivatives of the primary variable. For example, in a plane stress or plane strain analysis, the primary unknowns which we compute are the displacement components of the nodes. However, we often are equally concerned about the strains and stresses which are computed from derivatives of the displacements. Likewise, when we model an ideal fluid with a velocity potential, we actually have little or no interest in the computed potential, but we are very interested in the velocity components which are the derivatives of the potential. A logical question at this point is: what location in the element will give me the most accurate estimate of derivatives?

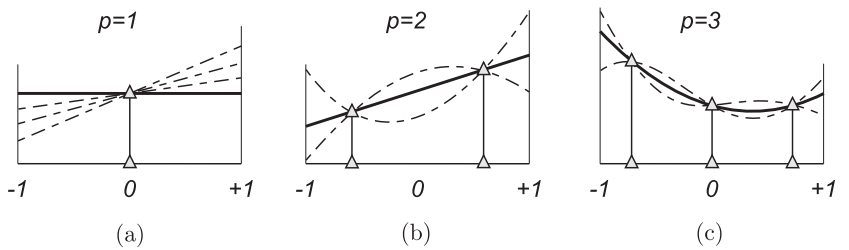


Figure 3.8: Integration property of gauss points: (a) $p = 1$; (b) $p = 2$; (c) $p = 3$ which guarantees superconvergence

At this stage, it is appropriate to restate the general problems discussed earlier. In the displacement method, the stresses are discontinuous between elements because of the nature of the assumed displacement variation. In analysis involving numerically integrated elements such an isoparametric elements experience has shown that the integration points are best stress sampling points. The nodes, which are the most useful output locations for stresses, appear to be worst sampling points. Reasons for this phenomenon are not immediately apparent even if they could be somehow explained basing on the arguments exposed in Section 3.1. However, it is well known that interpolation functions tend to behave badly near the extremities of the interpolation region. It is therefore reasonable to expect the shape function derivatives (and hence stresses) sampled in the interior of elements would be more accurate than those sampled on the element periphery. Indeed in one dimension at least we shall find that such points often exhibit the quality known as *superconvergence* (i.e., the values sampled at these points show an error which decreases more rapidly than elsewhere). Obviously, the user of fi-

nite element analysis should be encouraged to employ such points but at the same time note that the errors overall may be much larger. To clarify ideas we shall start with a typical problem of second order in one dimension.

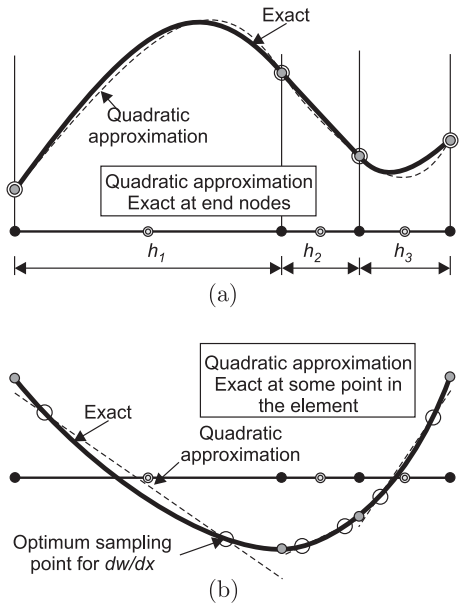


Figure 3.9: Optimal sampling points for the function (a) and its gradient (b) in one dimension (quadratic elements)

3.4.1 A one-dimensional example

Here we consider a problem of a second-order equation which may be typical of either one-dimensional heat conduction or the equilibrium of a thin membrane subjected to a transverse load b This equation can readily be written as

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(k\frac{\mathrm{d}u}{\mathrm{d}x}\right) = b \tag{3.64}$$

with the boundary conditions either defining the values of the function u or of its gradients at the ends of the domain.

Let us consider a typical problem shown in Figure 3.7. Here we show an exact solution for u and $\mathrm{d}u/\mathrm{d}x$ for a span of several elements and indicate the type of solution which will result from a finite element calculation using linear elements. We have already noted that on occasions we shall obtain exact solutions for u at nodes. This will happen when the shape functions contain the exact solution of the homogeneous differential equation, a situation which happens for Equation (3.64)

when $b = 0$ and polynomial shape functions are used. In all cases, even when b is non-zero and linear shape functions are used, the nodal values generally will be much more accurate than those elsewhere, Figure 3.7(a). For the gradients shown in Figure 3.7(b) we observe large discrepancies of the finite element solution from the exact solution but we note that somewhere within each element the results are nearly exact. It would be useful to locate such points and indeed we have already remarked in the context of two-dimensional analysis that values obtained within the elements tend to be more accurate for gradients (strains and stresses) than those values calculated at nodes. Clearly, for the problem illustrated in Figure 3.7(b) we should sample somewhere near the centre of each element.

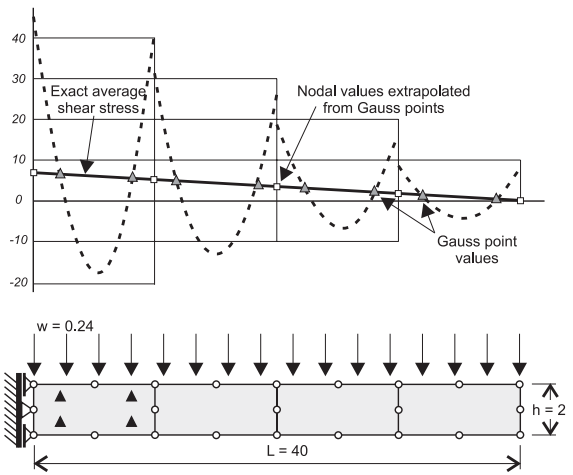


Figure 3.10: Cantilever beam with four 8-node elements. Stress sampling at cubic order (2x2) Gauss points with extrapolation to nodes

Pursuing this problem further in a heuristic manner we shall note that, if higher order elements (e.g., quadratic elements) are used, the solution still remains exact or nearly exact at the end nodes of an element but may depart from exactness at the interior nodes, as shown in Figure 3.9(a). The stresses, or gradients, in this case will be optimal at points which correspond to the two Gauss quadrature points for each element as indicated in Figure 3.9(b). This fact was observed experimentally by Barlow in [Barlow, 1976]-[Barlow, 1977]-[Barlow, 1989], and such points are frequently referred to as *Barlow points*.[2]

We shall now state in an axiomatic manner that:

[2] The waste bin is not the right place for our failures, when an intelligent response can generate ideas like the Barlow points. (by B. Irons from [Irons & Ahmadn, 1980])

1. the displacements are best sampled at the nodes of the element, whatever the order of the element is,
2. the best accuracy is obtainable for gradients or stresses at the Gauss points corresponding, in order, to the polynomial used in the solution.

At such points the order of the convergence of the function or its gradients is one order higher than that which would be anticipated from the appropriate polynomial and thus such points are known as *superconvergent*.

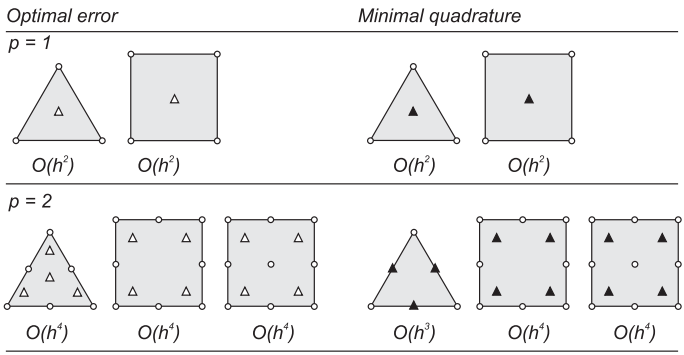


Figure 3.11: Optimal superconvergent sampling and minimum integration points for some elements

3.4.2 Optimal sampling points

The single point at the centre of an element integrates precisely all linear functions passing through that point and, hence, if the stresses are exact to the linear form they will be exact at that point of integration. For any higher order polynomial of order p, the Gauss-Legendre points numbering p will provide points of superconvergent sampling. We see this from Figure 3.8 directly. Here we indicate one, two, and three point Gauss-Legendre quadrature showing why exact results are recovered there for gradients and stresses. For points based on rectangles and products of polynomial functions it is clear that the exact integration points will exist at the product points as shown in Figure 3.11 for various rectangular elements. In the same figure we show, however, some triangles and what appear to be *good* but not necessarily superconvergent sampling points. Though we find that superconvergent points do not exist in triangles, the points shown in Figure 3.11 are optimal. In Figure 3.11 we contrast these points with the minimum number of quadrature points necessary for obtaining an accurate (though not always stable)

Table 3.2: Basic element properties: the optimal points are also the reduced rule, althoug the latter may give mechanism for some of the element types

		Number of:			number of intgn rules:		stress var'n:	
	el. type	nodes	dofs	el. order	complete	optimal points	complete	pseudo
2D	TRI3	3	6	linear	1	1	const	-
2D	TRI6	6	12	quadr	3×3	2×2	linear	-
2D	Q4	4	8	linear	2×2	1×1	const	linear
2D	P8	8	16	quadr	3×3	2×2	linear	quadr
2D	Q9	9	18	quadr	3×3	2×2	linear	quadr

stiffness representation and find these to be almost coincident at all times.

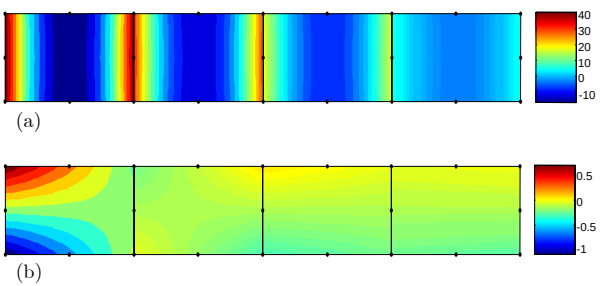


Figure 3.12: Bending of a slender beam using 3×3 integration rules points: (a) shear stress map, (b) longitudinal stress

The Gaussian quadrature integration rules suitable for different members of the isoparametric element families are shown in Table 3.2.

In Figure 3.10, representing an analysis of a cantilever by four rectangular elements with 8 nodes, we see how well the stresses sampled at superconvergent points behave compared to the overall stress pattern computed in each element. It is from results like this that many suggestions have been made to obtain improved nodal values and one method proposed by Hinton and Campbell [Hinton & Campbell, 1974] has proved to be quite widely used. However, we shall discuss better recovery procedures later.

The extension of the idea of superconvergent points from one-dimensional elements to two-dimensional rectangles was fairly obvious. However, the full superconvergence is lost when isoparametric distortion occurs.

3.5 Effects of geometry distortion on the element performance

It is well known that many existing finite elements exhibit poor performance when the element geometry is distorted (see for example [Carey & Oden, 1983]–[Zienkiewicz & Taylor, 1989]–[Striklin *et al.* , 1977]–[Backlund, 1978]–[Gifford, 1979]–[Cannarozzi *et al.* , 1997]–[Macneal & Harder, 1992]. A set of basic distor-

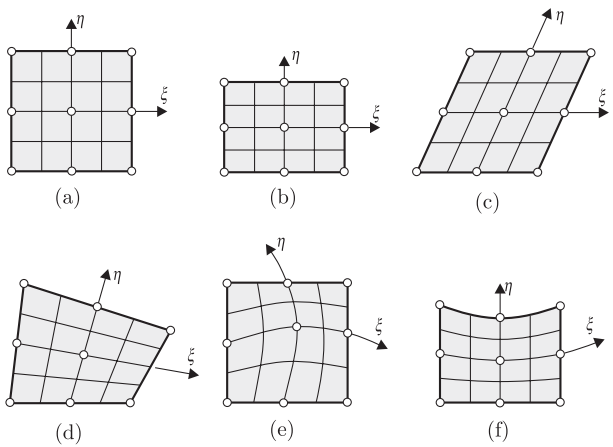


Figure 3.13: Classification of element distortions: (a) undistorted element; (b) ratio distortion; (c) parallelogram distortion; (d) skew distortion; (e) mid-side node distortion; (f) curved edge distortion.

tion shapes is shown in Figure 3.13. In practice, most distortions are a combination of these.

It has been a common knowledge that the interpolation precision of quadrangular finite elements deteriorates if the element geometry is not a parallelogram. This is because for more general element configurations, the global coordinates are no more linear functions of local coordinates in the geometry mapping. Accordingly, the local coordinates are no more polynomials of the global coordinates, but irrational functions. Thus, the interpolated field is badly dependent on the real shape of the element. Then, as hint before, there immediately occur two issues: the first is how to determinate the interpolation precision that can be exactly achieved by the shape functions for the element configuration, the second is whether we can find an approach to improve the element performance. A qualitative method is available to investigate the first issue: the polynomial we are interested in understanding when correctly captured (say x^2) is expanded in terms of the local

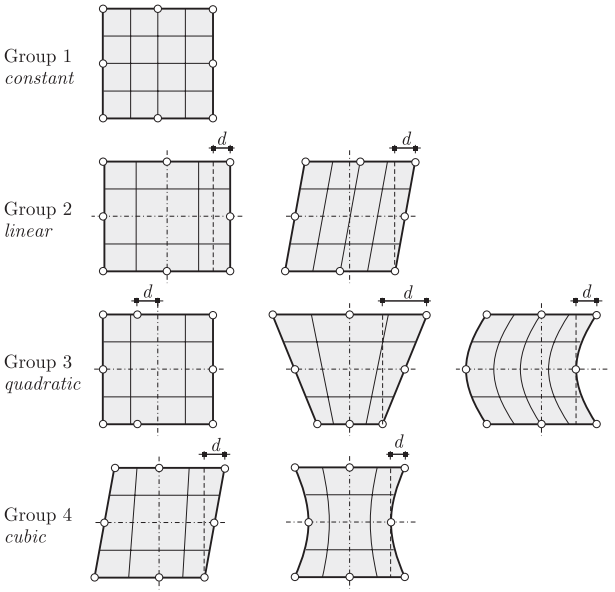


Figure 3.14: Distortion groups for serendipity eigth node elements

coordinates by using the geometry mapping and then we check if the shape functions contain every term needed to describe it [Lee & Bathe, 1993]. This method has been widely used, but due to the indirect analysis it gives no useful information to evaluate, control or predict the resulting error as well as no indication on how this trouble can be cured.

To study element distortions, it could be convenient to divide the shapes into 5 groups. Groups 1 to 4 follow the convention of Barlow [Barlow, 1989] and are shown in Figure 3.14, while group 5 is covered by Figure 3.15. Each group has geometric distortions in the user space, which contain certain terms from the (ξ_1, ξ_2) polynomial in the theory space. The higher the group, the more terms in this polynomial to represent the cartesian geometry. In ascending order of distortion severity, the groups are:

- group 1: constant (ξ_1, ξ_2) terms, a square shape,
- group 2: linear (ξ_1, ξ_2) terms, a rectangle or a parallelepiped,
- group 3: quadratic (ξ_1, ξ_2) terms, displaced midside nodes, tapered rectangles, simple curved sides,
- group 4: quasi-cubic (ξ_1, ξ_2) terms, more complicated distortion shapes,
- group 5: extreme distortions, where singularities exist such that the Jacobian determinant, Equation (3.39) , is zero at one or more points.

Most practical distortions will be combinations of the above, when the group-

Table 3.3: Total response errors from distributios for quadrilaterals 8-node serendipity elements, (see [Barlow, 1989])

	2×2 Rule			3×3 Rule		
	Distortion groups:					
Order of strains	2	3	4	2	3	4
constant	0	0	0	0	0	0
linear	0	d^2	d	0	d	d
quadratic	1	1	1	1	1	1

Table 3.4: Stress calculation errors from distortions for quadrilaterals 8-node serendipity elements, (see [Barlow, 1989])

	General points			Optimal 2×2 points		
	Distortion groups:					
Order of strains	2	3	4	2	3	4
constant	0	0	0	0	0	0
linear	0	d	d	0	d^2	d
quadratic	1	1	1	0	d	d

ing is decided by the most severe component. In figure 3.14, the amount of the distortion is indicated by a single measure of magnitude d (the α parameters of Figure 3.5 and 3.6). Note that as the length of the side is 1, then $d < 1$ and so $d^2 < d$. The significance of these groupings is that errors due to the distortions depend on the group, the magnitude of d and the nature of the strain field experienced by the element. The higher the group number, the worse the error can be.

The nature of the strain experienced by the element is an important consideration, since a highly distorted element may well model a certain component of strain accurately and another (such as a transverse strain) badly. Also, for elastic isotropic elements, strain is proportional to stress, so that we can consider either strain or stress in these arguments. The exact strain distribution across an element is not known a priori, so different strain fields are considered in turn. The highest strain field that a single 8-node element can accommodate is quadratic (albeit incomplete), hence we consider the effects of constant, linear and quadratic strain variations. In practice, they appear across the element as a combined effect, but in this assessment they are considered in this separated form.

Distortion errors do not occur for group 1 shapes, since there is no distortion. For groups 2 to 4, Barlow [Barlow, 1989] has deduced for each distortion group the error variations occurring in each element due to the applied strain variations. Table 3.3 shows the order of magnitude of the errors for each distortion group, under different orders of applied strain field. These errors are the combined errors of the calculated stresses (or strains) and the resulting stiffness matrix. Table 3.4 shows the corresponding orders of magnitude of the error in the calculated stresses (or strains) only. Here, general points are considered to be any locations in the element other than the optimal points. In these tables, the error of 1 denotes the same magnitude as the individual components of stiffness or stress, i.e. very large, whilst d is the order of the distortion, which in this analysis is assumed to be small, with $d^2 << d$. In each entry, the error is for the worst case, so that some cases in each category may show much less error.

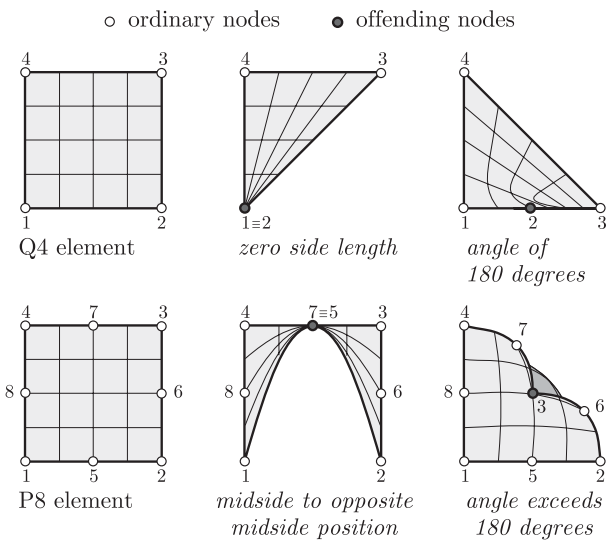


Figure 3.15: Types of extreme distortions in four and eigth node elements

In Table 3.3, it is seen that the total response errors due to any order of distortion are negligible under constant strain fields, but are of order of unity under quadratic strain fields, for both complete and reduced integration. The linear strain field errors are slightly more favourable for reduced integration, mainly of order d. Table 3.4 shows very small errors for stresses evaluated at the optimal points, never more than d even for quadratic strain fields. At other points, the errors are of order of unity under quadratic strain fields.

The above groups have distortions measured by the geometric quantity d for each particular shape. An alternative way to view element distortion is the Jacobian determinant, as discussed in previous section. For groups 1 and 2, $\det J$ is constant over the element. For the other groups, it is not constant and has been observed to vary: in general larger variations are observed for larger distortions. Hence, the ratio of the maximum to minimum values of $\det J$ over an element gives an a priori guide to distortion. Many points over the element need to be sampled, particularly on sides and faces; just sampling at the Gauss points is not sufficient. The ratio is thus easily calculated for the mesh at generation time and is available for diagnostic warnings in some commercial software. A value of typically 2 is often taken as a guide for the upper limit of this ratio, such that when exceeded the distortion may be too severe, and the user should check this part of the mesh.

It should be emphasized that, in real life, some applied load cases may only give rise to insignificant distortion errors in elements where the $\det J$ ratio is high, and, conversely, significant errors where the ratio is low. Evaluating the ratio therefore has to be considered only as an indicative a priori guide to distortion.

Distortion effects can be assessed by calculating the individual distortion shapes of Figure 3.13 for each element. These can be checked, along with $\det J$, by automatic software procedures, establishing which group is relevant to each element. Then, knowledge of the amount of distortion in each element, by diagnostic warnings of elements in certain of the above groups, or by the above $\det J$ ratios, enables suitable warnings to be highlighted. The effects are, however, very dependent on the applied loading, so that particular caution would be required if multiple load cases were being run concurrently.

3.5.1 A simple distortion test for quadratic elements

For elements that have complete polynomial basis functions of at least degree 2 (8 and 9-node elements), we use the constant-bending-moment problem shown in Figure 3.16 to demonstrate the effects of element distortions [Lee & Bathe, 1993]. The Young's modulus is $E = 1.0 \times 10^7$, Poisson's ratio is $\nu = 0.3$, thickness is $t = 1.0$ and we are in a plane stress condition. We solve this problem using 8-node, 9-node isoparametric elements and the meshes of Figure 3.16 (b)-(c). The corresponding stress band plots are given in Figures 3.16 (d)-(e).

The results show that the 9-node Lagrangian element is not affected by angular distortions, no matter how severe they may be. The 8-node serendipity element, on the other hand, is badly affected by angular distortions. We also see that the 8- and 9-node elements are significantly affected by the curved-edge distortions, and by the mid-side node distortion Figure 3.17.

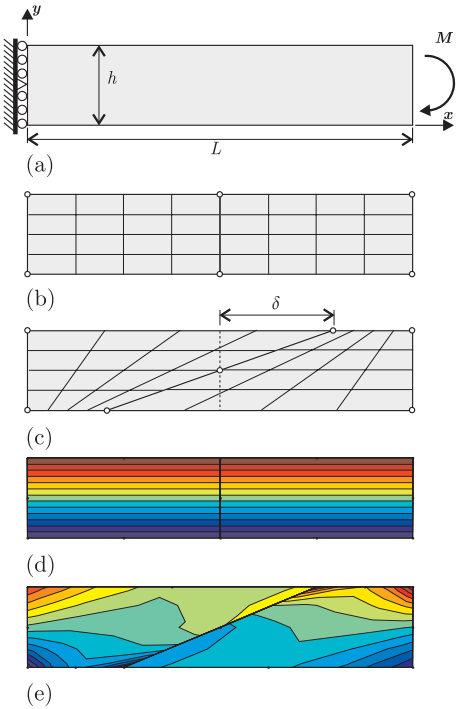


Figure 3.16: The two element test: (a) the geometry and loading; (b) undistorted configuration; (c) progressive distortion due to δ parameter; (d) stress band plots of longitudinal stresses (σ_{xx}) in constant-bending-moment problem using undistorted mesh; (e) stress band plots of longitudinal stresses (σ_{xx}) using a distorted mesh

These findings have several important practical implications, such as:

1. for a reliable analysis, the use of the serendipity elements is best restricted to cases where rectangular or parallelogram elements can be used;
2. Lagrangian elements can be freely used: one need not be concerned about awkward-looking meshes as long as the elements have straight sides and evenly spaced nodes;
3. elements with curved sides should be used to represent only curved boundaries since the performance of such elements is inferior to those that have straight sides;
4. Lagrangian elements are preferred in large-displacement, large-strain analyses since the elements in these analyses can quickly become distorted;
5. adaptive schemes in large deformation analysis should be based on the use of completely new meshes when mesh/interpolation refinement at a particular step is required.

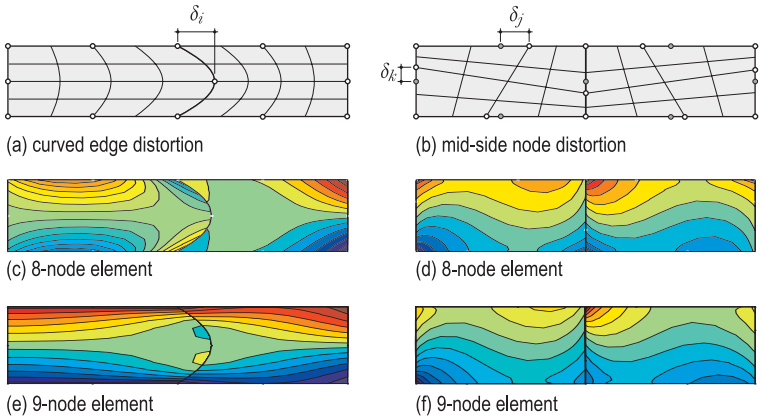


Figure 3.17: Stress band plots of longitudinal stresses (σ_{xx}) in constant-bending-moment problem for a curved edge distortion and a mid-side node distortion

Angular distorted element usually appear in the mesh generation of curved geometries and in transition regions from coarse to fine mesh, and also in the nonlinear problems involving large deformations.

In some cases, extreme distortions actually model certain singularities correctly, such as the elastic and elasto-plastic strain fields around sharp crack tips.

Automatic mesh generators would normally avoid generating such extreme distortions, although in some awkward geometrical shapes they could well arise. It is therefore sensible to utilize any pre-processor checks for zero $\det J$. However, checking this quantity has to be made at many locations over each element, including on its boundaries, since $\det J$ has a polynomial variation which is of higher order the more the distortion, and so which requires a lot of sampling points to detect trends towards zero. This check is usefully made with the ratio of maximum to minimum $\det J$ over each element, to give some guidance to distortion effects.

Table 3.5 collect the polynomial displacement fields that can be solved exactly by the various serendipity and Lagrangian elements in their undistorted and distorted configurations, as shown by Lee and Bathe in [Lee & Bathe, 1993].

3.6 A new procedure to evaluate element distortions

In this section an analytical procedure to evaluate element distortions is presented [Castellazzi, 2001]-[Castellazzi & Ubertini, 2004].

Table 3.5: Polynomial displacement fields that can be solved exactly by the various serendipity and Lagrangian elements in their undistorted and distorted configurations

Type of element	Undistorted configuration - - - - - - - - - - - - - - - - Additional fields if also unrotated								Angular distortion		Quadratic curved-edge distortion	
8-node serendipity element		x^2	x	1	y	y^2			x	y	x	y
12-node serendipity element	x^3	x^2	x	1	y	y^2	y^3		x	y	x	y
		x^3y	x^2y	xy	xy^2	xy^3						
9-node Lagrangian element		x^2	x	1	y	y^2			x^2 x	1 y y^2	x	y
			x^2y	xy	xy^2				xy			
				x^2y^2								
16-node Lagrangian element	x^3	x^2	x	1	y	y^2	y^3		x^3 x^2	x 1 y y^2 y^3	x	y
		x^3y	x^2y	xy	xy^2	xy^3			x^2y xy xy^2			
			x^3y^2	x^2y^2	x^2y^3							
				x^3y^3								
25-node Lagrangian element	x^4	x^3	x^2	x	1	y	y^2	y^3	y^4	x^3 x^2	x 1 y y^2 y^3	x	y
		x^4y	x^3y	x^2y	xy	xy^2	xy^3	xy^4		x^2y xy xy^2			
			x^3y^2	x^2y^2	x^2y^3		xy^3						
				x^3y^3									

Consider Figure 3.3 the coordinate transformation defined by

$$T: \ \mathbf{x} = \mathbf{x}(\boldsymbol{\xi}), \tag{3.65}$$

over the master element $\hat{B}$. The following condition on the determinant of the Jacobian matrix

$$j(\boldsymbol{\xi}) = \det\left[\nabla_{\boldsymbol{\xi}}\mathbf{x}\right] > 0, \tag{3.66}$$

ensures locally the existence and uniqueness of the inverse map:

$$T^{-1} : \boldsymbol{\xi} = \boldsymbol{\xi}(\mathbf{x}). \tag{3.67}$$

Transformation T is called direct transformation, transformation T^{-1} is called inverse transformation. For convenience, in standard finite elements interpolation functions (e.g. shape functions) are defined on the master element, $\hat{F} = \hat{F}(\boldsymbol{\xi})$. However, the interpolation quality should be evaluated over the physical domain. Obviously, to determine the actual dependence of interpolation functions on $\mathbf{x}$ the inverse map is needed:

$$F(\mathbf{x}) = F(\boldsymbol{\xi}(\mathbf{x})) \overset{T^{-1}}{=} \hat{F}(\boldsymbol{\xi}).$$

Except for very simple element geometry, determining the inverse transformation is an hard task, due to nonlinearity of Equation 3.65. In the following, the inverse mapping is determined trough the Gröbner basis method [Cox *et al.* , 1992]. This means to compute a Gröbner basis of the ideal generated by polynomials $\mathbf{x} - \mathbf{x}(\boldsymbol{\xi})$. The Gröbner basis is a new set of polynomials which has the same roots of the original one $\mathbf{x} - \mathbf{x}(\boldsymbol{\xi})$ and, in addition has a triangular form. Thus, one of the new polynomials contains just one component of $\boldsymbol{\xi}$ and can be solved for in. Then by back substitution into the other polynomial the inverse mapping can be found. This procedure has no theoretical limitations, since Gröbner bases can always be computed in a finite number of steps [Buchberger, 2001]–[Tran, 2000], using for example a commercial software for algebraic manipulation such as MAPLE, although it may be a very time consuming. However, in practice the one component equation can be explicitly solved only up to the fourth degree. Moreover, the resulting expressions could be heavy to be handled. In the following some basic transformations are analyzed and discussed.

The analysis is conducted basing on the fact that if the direct transformation is defined by polynomial $\hat{P}$:

$$\mathbf{x} \overset{def}{=} \hat{P}(\boldsymbol{\xi}), \tag{3.68}$$

the inverse map, can be represented as:

$$\boldsymbol{\xi}(\mathbf{x}) = P(\mathbf{x}) + NP(\mathbf{x}), \ \text{with } P \subseteq \hat{P}, \tag{3.69}$$

where NP is the non-polynomial part. Indeed these equations have a completely

polynomial form only if T is linear. Based on the above remark and taking into account that the finite element generally crucially depends on the polynomial part of the interpolation functions, the analysis is focused on this part. In particular we consider vector $\hat{\boldsymbol{\Xi}}$ collecting monomials in $\boldsymbol{\xi}$

$$\hat{\boldsymbol{\Xi}} = \left[1,\ \xi_1,\ \xi_2,\ \xi_1\xi_2,\ \xi_1^2 \ldots,\ \xi_2^n\right], \tag{3.70}$$

which can be viewed as the basis of the shape functions. Then through the computed inverse mapping we could determine the transformation basis $\boldsymbol{\Xi}$, whose structure can be analyzed based on the following matrix:

$$\mathbf{J}' = \nabla\boldsymbol{\Xi}, \tag{3.71}$$

where derivations carried out with respect to

$$\mathbf{X} = \left[1,\ \ x_1,\ x_2,\ x_1x_2,\ x_1^2, \ldots,\ x_2^m\right]. \tag{3.72}$$

It is to be intended as a formal derivation capable to extract from lengthy expressions the polynomial part of the transformation, as shown by the following example:

$$\hat{\boldsymbol{\Xi}} = [1,\ \xi_1,\ \xi_2,\ \xi_1\xi_2], \tag{3.73}$$

$$\mathbf{X} = \left[1,\ \ x_1,\ x_2,\ x_1x_2,\ x_1^2, \ldots\right], \tag{3.74}$$

a generic the inverse map, can be represented as:

$$\begin{aligned} \xi_1 &= x_1 + \sqrt{x_1} \\ \xi_2 &= x_2 \end{aligned} \tag{3.75}$$

$$\boldsymbol{\Xi} = \left[1,\ x_1 + \sqrt{x_1},\ x_2,\ x_1x_2 + x_2\sqrt{x1}\right] \tag{3.76}$$

$$\mathbf{J}' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.77}$$

By inspecting matrix $\mathbf{J}'$ and, in particular, linear independence of rows and columns we can extract useful information on the polynomial part of the transformation.

3.6.1 Some basic distortions of plane elements

For later convenience the transformation is written as

$$T: \quad \mathbf{x} = \mathbf{A}_{2\times1} + \mathbf{B}_{2\times2}\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \mathbf{C}_{2\times3}\begin{bmatrix} \xi_1^2 \\ \xi_1\xi_2 \\ \xi_2^2 \end{bmatrix} + \mathbf{D}_{2\times i}\begin{bmatrix} \xi_1^3 \\ \vdots \\ \xi_2^3 \end{bmatrix} + \ldots \tag{3.78}$$

Table 3.6: Relation between elements of $\mathbf{X}$ and elements of $\mathbf{\Xi}$

		1	x_1	x_2	x_1x_2	x_1^2	$\cdots$	
		$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$		
1	$\leftrightarrow$	■	0	0	0	0	$\cdots$	
ξ_1	$\leftrightarrow$	0	■	0	0	0	$\cdots$	
ξ_2	$\leftrightarrow$	0	0	■	■	0	$\cdots$	$\Longrightarrow J^{'}$
$\xi_1\xi_2$	$\leftrightarrow$	0	0	0	■	■	$\cdots$	
$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	

Table 3.7: Trasformation and relation

$T:\ \mathbf{x}=\mathbf{A}_{2\times1}$	$\rightarrow$ translation	
$T:\ \mathbf{x}=\mathbf{B}_{2\times2}\cdot\boldsymbol{\xi}'_{2\times1}$	$\rightarrow$ simple rotation if $\mathbf{B}_{2\times2}=\mathbf{R}$	$\mathbf{R}\in Orth$
$T:\ \mathbf{x}=\mathbf{A}_{2\times1}+\mathbf{B}_{2\times2}\cdot\boldsymbol{\xi}'_{2\times1}$	$\rightarrow$ generic linear transformation	

Where $\mathbf{A}_{2\times1}$ is 0^{th} order matrix , $\mathbf{B}_{2\times2}$ is the first order matrix and so on. Table 3.7 collects some basic transformations.

In the following section we will denote by the letter Q quadrilateral Lagrangian elements plus a number to identify the degree of completeness, that, Q1 denotes four node Lagrangian elements, Q2 nine node Lagrangian element and Q3 sixteen node Lagrangian element. We will use also the letter P for serendipity elements : P2 denotes eigth-node serendipity element, P3 the twelve-node element.

3.6.2 A simple rotation

The simpler linear transformation is a rotation. With respect to Figure 3.18(a) the transformation (3.78) reduces to

$$T:\quad \mathbf{x}=\mathbf{R}\cdot\boldsymbol{\xi}, \tag{3.79}$$

with $\mathbf{R}^{-1}=\mathbf{R}^{\mathrm{T}}$, and in extended form

$$T:\quad \begin{cases} x_1=\cos\alpha\xi_1-\sin\alpha\xi_2, \\ x_2=\sin\alpha\xi_1+\cos\alpha\xi_2. \end{cases} \tag{3.80}$$

In this simple case, the inverse transformation can be easily computed and given by

$$T^{-1}:\quad \begin{cases} \xi_1=\cos\alpha x_1+\sin\alpha x_2, \\ \xi_2=-\sin\alpha x_1+\cos\alpha x_2. \end{cases} \tag{3.81}$$

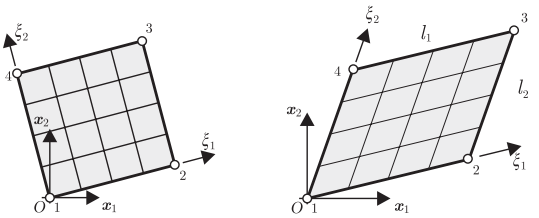


Figure 3.18: Linear transformation: the local normal co-ordinate system for a simple rotation (a) and for a generic lianear transformation (b).

Table 3.8: Relation between of the monomials

	1	x_1	x_2	x_1x_2	x_1^2	x_2^2	x_1^3	$x_1^2x_2$	$x_1x_2^2$	x_2^3	x_1^4	$x_1^3x_2$	$x_1^2x_2^2$	$x_1x_2^3$	x_2^4
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ξ_1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
ξ_2	0	−1	0	0	0	0	0	0	0	0	0	0	0	0	0
$\xi_1\xi_2$	0	0	0	−1	0	0	0	0	0	0	0	0	0	0	0
ξ_1^2	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
ξ_2^2	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
$\xi_1^2\xi_2$	0	0	0	0	0	0	0	0	−2	0	0	0	0	0	0
$\xi_1\xi_2^2$	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
$\xi_1^2\xi_2^2$	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0

Following the outlined procedure we can compute matrix $\mathbf{J}'$, given by Table 3.8.

The analysis of this matrix can be summarized in the following three points.

1. The sub-matrix established by the first 4 rows and columns, corresponding to bilinear Lagrange element Q1, have independent rows, resulting in the following polynomial basis:
$$[1 \quad x_2 \quad -x_1 \quad -x_1x_2]. \tag{3.82}$$
2. The sub-matrix corresponding to Serendipity element P2 has independent rows, resulting in the following polynomial basis:
$$[1 \quad x_2 \quad -x_1 \quad -x_1x_2 \quad 2x_2^2 \quad 2x_1^2 \quad -2x_1x_2^2 \quad 2x_1^2x_2]. \tag{3.83}$$
3. The sub-matrix corresponding to Lagrangian element Q2 has independent

Table 3.9: Coordinate's point, see Figure 3.18

	x	y
P_1	0	0
P_2	$l_1 \cos\alpha_1$	$l_1 \sin\alpha_1$
P_3	$l_1 \cos\alpha_1 + l_2 \cos\alpha_2$	$l_1 \sin\alpha_1 + l_2 \sin\alpha_2$
P_4	$l_2 \cos\alpha_2$	$l_2 \sin\alpha_2$

rows, resulting in the following polynomial basis:

$$[1 \quad x_2 \quad -x_1 \quad -x_1x_2 \quad 2x_2^2 \quad 2x_1^2 \quad -2x_1x_2^2 \quad 2x_1^2x_2 \quad 4x_1^2x_2^2]. \tag{3.84}$$

This practical, and very simple, example show as rotation does not deteriorate element performance, as expected.

3.6.3 A generic linear transformation

A generic linear transformation can be written as

$$T: \quad \mathbf{x} = \mathbf{A} + \mathbf{B} \cdot \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} \alpha_{1,0} \\ \alpha_{2,0} \end{bmatrix} + \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}, \tag{3.85}$$

and represents a quadrilateral with parallel straight edges, eventually rotated and translated. For simplicity we set $\mathbf{A} = \mathbf{0}$ ($\alpha_{i,0} = 0$) that rules out the translation. Making reference to Figure 3.18(b) and Table 3.1, the direct and inverse transformations are

$$T: \begin{cases} x_1 = (l_1 \cos\alpha_1)\xi_1 + (l_2 \cos\alpha_2)\xi_2, \\ x_2 = (l_1 \sin\alpha_1)\xi_1 + (l_2 \sin\alpha_2)\xi_2, \end{cases}, \tag{3.86}$$

$$T^{-1}: \begin{cases} \xi_1 = \frac{\sin\alpha_2}{l_1 d} x_1 - \frac{\cos\alpha_2}{l_1 d} x_2, \\ \xi_2 = \frac{\sin\alpha_1}{l_2 d} x_1 - \frac{\cos\alpha_1}{l_2 d} x_2, \end{cases}, \tag{3.87}$$

where

$$d = \cos\alpha_1 \sin\alpha_2 - \cos\alpha_2 \sin\alpha_1. \tag{3.88}$$

Setting: $c_1 = \cos\alpha_1$, $s_1 = \sin\alpha_1$, $c_2 = \cos\alpha_2$, $s_2 = \sin\alpha_2$, and making reference to Q2 element we have

$$\hat{\boldsymbol{\Xi}} = [1 \ \xi_1 \ \ \xi_2 \ \ \xi_1\xi_2 \ \ {\xi_1}^2 \ \ {\xi_2}^2 \ \ {\xi_1}^2\xi_2 \ \ \xi_1{\xi_2}^2 \ \ {\xi_1}^2{\xi_2}^2], \tag{3.89}$$

$$\mathbf{X}_i = [1 \ x_1 \ x_2 \ x_1x_2 \ \ x_1^2 \ \ x_2^2 \ \ x_1^3 \ \ x_1^2x_2 \ \ x_1x_2^2 \ \ x_2^3 \ \ \ x_1^4 \ \ x_1^3x_2 \ \ x_1^2x_2^2 \ \ x_1x_2^3 \ \ x_2^4]. \tag{3.90}$$

Table 3.10: Coefficients of a generic linear transformation

$e_1 = \frac{s_2}{l_1 d}$,	$e_{10} = \frac{{c_2}^2}{l_1^2 d^2}$,	$e_{19} = \left(\frac{2c_2 s_1 c_1}{l_1 d^3 {l_2}^2} + \frac{s_2 {c_1}^2}{l_1 d^3 {l_2}^2}\right)$,
$e_2 = -\frac{c_2}{l_1 d}$,	$e_{11} = \frac{s_1^2}{l_2^2 d^2}$,	$e_{20} = \frac{s_2 {s_1}^2}{l_1 d^3 {l_2}^2}$,
$e_3 = -\frac{s_1}{l_2 d}$,	$e_{12} = -2\,\frac{s_1 c_1}{l_2^2 d^2}$,	$e_{21} = -\frac{c_2 {c_1}^2}{l_1 d^3 {l_2}^2}$,
$e_4 = \frac{c_1}{l_2 d}$,	$e_{13} = \frac{c_1^2}{l_2^2 d^2}$,	$e_{22} = \frac{4 s_2 c_2 s_1 c_1 + {s_2}^2 {c_1}^2 + {c_2}^2 {s_1}^2}{{l_1}^2 d^4 {l_2}^2}$,
$e_5 = -\frac{s_2 s_1}{l_1 d^2 l_2}$,	$e_{14} = \left(\frac{s_2^2 c_1}{l_1^2 d^3 l_2} + \frac{2 s_2 c_2 s_1}{l_1^2 d^3 l_2}\right)$,	$e_{23} = \left(-\frac{2{s_2}^2 s_1 c_1}{{l_1}^2 d^4 {l_2}^2} - \frac{2 s_2 c_2 {s_1}^2}{{l_1}^2 d^4 {l_2}^2}\right)$,
$e_6 = \left(\frac{s_2 c_1}{l_1 d^2 l_2} + \frac{c_2 s_1}{l_1 d^2 l_2}\right)$,	$e_{15} = \left(-\frac{{c_2}^2 s_1}{{l_1}^2 d^3 l_2} - \frac{2 s_2 c_2 c_1}{{l_1}^2 d^3 l_2}\right)$,	$e_{24} = \left(-\frac{2 s_2 c_2 {c_1}^2}{{l_1}^2 d^4 {l_2}^2} - \frac{2{c_2}^2 s_1 c_1}{{l_1}^2 d^4 {l_2}^2}\right)$,
$e_7 = -\frac{c_2 c_1}{l_1 d^2 l_2}$,	$e_{16} = -\frac{{s_2}^2 s_1}{l_1^2 d^3 l_2}$,	$e_{25} = \frac{{c_2}^2 {c_1}^2}{{l_1}^2 d^4 {l_2}^2}$,
$e_8 = \frac{{s_2}^2}{l_1^2 d^2}$,	$e_{17} = \frac{{c_2}^2 c_1}{l_1^2 d^3 l_2}$,	$e_{26} = \frac{{s_2}^2 {s_1}^2}{{l_1}^2 d^4 {l_2}^2}$
$e_9 = -2\,\frac{s_2 c_2}{l_1^2 d^2}$,	$e_{18} = \left(-\frac{2 s_2 s_1 c_1}{l_1 d^3 {l_2}^2} - \frac{c_2 {s_1}^2}{l_1 d^3 {l_2}^2}\right)$,	

Then the functional matrix in Equation (3.71) takes the form:

$$\mathbf{J}' = \left[\begin{array}{cccccc|cccc|ccccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & e_1 & e_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & e_3 & e_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & e_6 & e_5 & e_7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & e_9 & e_8 & e_{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & e_{12} & e_{11} & e_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & e_{16} & e_{14} & e_{15} & e_{17} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & e_{20} & e_{18} & e_{19} & e_{21} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & e_{26} & e_{23} & e_{22} & e_{24} & e_{25}
\end{array}\right] \tag{3.91}$$

where coefficients e_i are collected in Table 3.10. This matrix has full rank, the first six columns are linear independent, just two columns are linear independent from 7^{th} to 10^{th} and just one among the last five columns. Therefore the polynomial part admits the following representation

$$\left[1 \; x_1 \;\; x_2 \;\; x_1 x_2 \;\; x_1^2 \; x_2^2 \; p_1 \; p_2 \; p_3\right], \tag{3.92}$$

where p_1 and p_2 are independent polynomials that are linear combinations of third degree monomials: $p_1 = p_1(x_1^3,\, x_1^2 x_2,\, x_1 x_2^2,\, x_2^3)$, $p_2 = p_2(x_1^3,\, x_1^2 x_2,\, x_1 x_2^2,\, x_2^3)$, whereas p_3 is a linear combination of fourth degree monomials: $p_3 = p_3(x_1^4,\, x_1^3 x_2,\, x_1^2 x_2^2,\, x_1 x_2^3,\, x_2^4)$. For the eight-node serendipity element, $\hat{\mathbf{\Xi}}$ is the same except for $\xi_1 \xi_2$ and $\mathbf{x}$ takes the form

$$\mathbf{X} = \left[1 \; x_1 \; x_2 \; x_1 x_2 \;\; x_1^2 \;\; x_2^2 \;\; x_1^3 \;\; x_1^2 x_2 \;\; x_1 x_2^2 \;\; x_2^3 \;\right]. \tag{3.93}$$

The corresponding functional matrix can be obtained by the first eight rows and ten columns of Equation (3.91). This new sub-matrix has full rank, and the most

Table 3.11: Effects of a generic linear transformation

degree	Q1	P2	Q2	P3	Q3
0°	1	1	1	1	1
1°	$x_1 \;\; x_2$	$x_1 \;\; x_2$	$x_1 \;\; x_2$	$x_1 \;\; x_2$	$x_1 \;\; x_2$
2°	p_1	$x_1x_2 \; x_1^2 \; x_2^2$	$x_1x_2 \; x_1^2 \; x_2^2$	$x_1x_2 \; x_1^2 \; x_2^2$	$x_1x_2 \; x_1^2 \; x_2^2$
3°	-	p_1, p_2	p_1, p_2	$x_1^3 \quad x_1^2x_2 \quad x_2^2x_1 \quad x_2^3$	$x_1^3 \quad x_1^2x_2 \quad x_2^2x_1 \quad x_2^3$
4°	-	-	p_3	p_4, p_5	p_6, p_7, p_8
5°	-	-	-	-	p_9, p_{10}
6°	-	-	-	-	p_{11}

general algebraic polynomials represented by $\hat{F}$ belong now to whole described by

$$\begin{bmatrix} 1 & x_1 & x_2 & x_1x_2 & x_1^2 & x_2^2 & p_1 \; p_2 \end{bmatrix}, \tag{3.94}$$

where p_1 and p_2 are already known. The functional matrix of Q1 element can be obtained by first four rows and first six columns of matrix contents in Equation (3.91). This new sub-matrix has full rank, and the most general algebraic polynomials is described by

$$\begin{bmatrix} 1 & x_1 & x_2 & p_0 \end{bmatrix}, \tag{3.95a}$$

where p_0 is a linear combination of second degree monomial:

$$p_0 = p_0(x_1^2, x_1x_2, x_2^2)$$

The same approach can be repeated also for other types of elements. The results are summarized in Table 3.11.

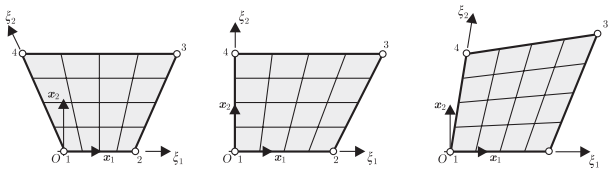
3.6.4 A bilinear transformation

A generic bilinear transformation is given by

$$T: \quad \mathbf{x} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \alpha_{1,3} \\ \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_1\xi_2 \end{bmatrix}. \tag{3.96}$$

It transforms $\hat{B}$ into a generic quadrilateral element. For simplicity we can analyze three bilinear transformation considering to progressively distorted element configuration as described in Figure 3.19. Whit respect to Figure 3.19(a), transformation reduces to

$$T: \begin{cases} x_1 = \alpha_{1,1} \cdot \xi_1 + \alpha_{1,3} \cdot \xi_1\xi_2 \\ x_2 = \alpha_{2,2} \cdot \xi_2 \end{cases}, \tag{3.97}$$



(a) (b) (c)

Figure 3.19: Bilinear transformation: (a) taper distortion), (b) taper distortion with skew, (c) generic bilinear distortion

and the inverse mapping is

$$T^{-1}: \begin{cases} \xi_1 = -\dfrac{-x_2}{\alpha_{2,2}} \\ \xi_2 = -\dfrac{\alpha_{2,2}\left(-x_1+\alpha_{1,1}\right)}{\alpha_{1,1}\alpha_{2,2}+\alpha_{1,3}x_2} \end{cases}, \tag{3.98}$$

The T^{-1} are therefore established from the following basis functions:

$$[1 \quad x_2 \quad r_1 \quad r_2], \tag{3.99}$$

where r_1 and r_2 are rational functions. With respect to Figure 3.19(b) transformation is:

$$T: \begin{cases} x_1 = \alpha_{1,1}\cdot\xi_1 + \alpha_{1,2}\cdot\xi_2 + \alpha_{1,3}\cdot\xi_1\xi_2 \\ x_2 = \alpha_{2,2}\cdot\xi_2 \end{cases}, \tag{3.100}$$

and the inverse takes the form

$$T^{-1}: \begin{cases} \xi_1 = \dfrac{x_1\alpha_{2,4} - \alpha_{1,1}\alpha_{2,4} - \alpha_{1,4}x_2 + \alpha_{1,4}\alpha_{2,1}}{\alpha_{1,2}\alpha_{2,4}} \\ \xi_2 = \dfrac{\alpha_{1,2}\left(x_2-\alpha_{2,1}\right)}{x_1\alpha_{2,4} - \alpha_{1,1}\alpha_{2,4} - \alpha_{1,4}x_2 + \alpha_{1,4}\alpha_{2,1}} \end{cases},$$

Now the quadrilateral element is not a parallelogram, and results from previous case let to take off from (3.96) rotation terms, but without loosing generality.

With respect to Figure 3.19(c) the transformation is:

$$T: \begin{cases} x_1 = \alpha_{1,1}\cdot\xi_1 + \alpha_{1,2}\cdot\xi_2 + \alpha_{1,3}\cdot\xi_1\xi_2 \\ x_2 = \alpha_{2,2}\cdot\xi_2 + \alpha_{2,3}\cdot\xi_1\xi_2 \end{cases}. \tag{3.101}$$

In this case the elimination process based in Gröbner basis leads to a second order equation which yields to two possible solutions for the inverse map, but just one is acceptable:

$$T^{-1}: \begin{cases} \xi_1 = \beta_{1,0} + \beta_{1,1}x_1 + \beta_{1,2}x_2 + \beta_{1,3}\sqrt{r\left(1,x_1,x_2,x_1x_2,x_1^2,x_2^2\right)} \\ \xi_2 = \beta_{2,0} + \beta_{2,1}x_1 + \beta_{2,2}x_2 + \beta_{2,3}\sqrt{r\left(1,x_1,x_2,x_1x_2,x_1^2,x_2^2\right)} \end{cases}, \tag{3.102}$$

where $r\left(1,x_1,x_2,x_1x_2,x_1^2,x_2^2\right)$ is a polynomial, and $\beta_{i,j}$ are coefficients depending

Table 3.12: Bilinear transformation

degree	Q1	P2	Q2	P3	Q3
0°	1	1	1	1	1
1°	x_1 x_2	x_1 x_2	x_1 x_2	x_1 x_2	x_1 x_2
2°	-	$h_1(x_1x_2,\ldots)$, $h_2(x_1x_2,\ldots)$	x_1x_2 x_1^2 x_2^2	x_1x_2 x_1^2 x_2^2	x_1x_2 x_1^2 x_2^2
3°	-	-	-	$h_3(x_1^3,\ldots)$	x_1^3 $x_1^2x_2$ $x_2^2x_1$ x_2^3

on element geometry:

$$\begin{aligned}
\beta_{1,0} &= \frac{(a_{1,2}a_{2,1} - a_{1,1}a_{2,2})}{2\,(a_{1,1}a_{2,3} - a_{1,3}a_{2,1})}, \qquad \beta_{1,1} = \frac{a_{2,3}}{2\,(a_{1,1}a_{2,3} - a_{1,3}a_{2,1})}, \\
\beta_{1,2} &= -\frac{a_{1,3}}{2\,(a_{1,1}a_{2,3} - a_{1,3}a_{2,1})}, \quad \beta_{1,3} = -\frac{1}{2a_{1,1}a_{2,3} - a_{1,3}a_{2,1}}, \\
\beta_{2,0} &= \frac{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}}{2\,(a_{1,2}a_{2,3} - a_{1,3}a_{2,2})}, \qquad \beta_{2,1} = \frac{a_{2,3}}{2\,(a_{1,2}a_{2,3} - a_{1,3}a_{2,2})}, \\
\beta_{2,2} &= -\frac{a_{1,3}}{2\,(a_{1,2}a_{2,3} - a_{1,3}a_{2,2})}, \quad \beta_{2,3} = \frac{1}{2\,(\alpha_{1,3}\alpha_{2,2} - \alpha_{2,3}\alpha_{1,2})},
\end{aligned} \tag{3.103}$$

The T^{-1} is therefore based on the following:

$$[1 \quad x_1 \quad x_2 \quad \sqrt{r\,(1, x_1, x_2, x_1x_2, x_1^2, x_2^2)}]. \tag{3.104}$$

Based on the above results on inverse mapping, the analysis of the polynomial part of interpolation functions can be carried out. The main results are summarized in Table 3.12, when h_i are linear combination of n-degree monomial, being n the table row.

3.6.5 Higher degree transformations: quadratic curved-edge

Here element geometry with quadratic curved edge is discussed. This most general quadratic transformation is given by

$$T: \quad \mathbf{x} = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} \\ \alpha_{2,1} & \alpha_{2,2} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \alpha_{1,3} & \alpha_{1,4} & \alpha_{1,5} \\ \alpha_{2,3} & \alpha_{2,4} & \alpha_{2,5} \end{bmatrix} \begin{bmatrix} \xi_1^2 \\ \xi_1\xi_2 \\ \xi_2^2 \end{bmatrix}. \tag{3.105}$$

Among this class we focus the attention on two cases with an increasing level of distortion as shown in Figure 3.20. Consider the semirectangulars [Frey *et al.* , 1978] of Figure 3.20(a) and (b), whose transformation are given by

$$T: \begin{cases} x_1 = \alpha_{1,1}\xi_1 + \alpha_{1,4}\xi_1\xi_2 + \alpha_{1,7}\xi_1\xi_2^2 \\ x_2 = \alpha_{2,2}\xi_2 \end{cases}, \tag{3.106}$$

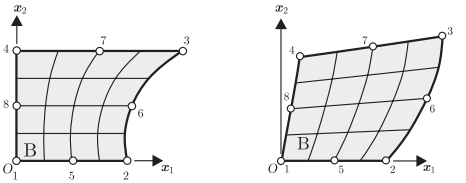


Figure 3.20: Higher degree transformations

and

$$T : \begin{cases} x_1 = \alpha_{1,1}\xi_1 + \alpha_{1,2}\xi_2 + \alpha_{1,3}\xi_1\xi_2 + \alpha_{1,4}\xi_2^2 \\ x_2 = \alpha_{2,2}\xi_2 + \alpha_{2,3}\xi_1\xi_2 + \alpha_{2,4}\xi_2^2 \end{cases} . \tag{3.107}$$

Using the Gröbner basis method the inverse mappings of the first transformation can be computed in the form

$$T^{-1} : \begin{cases} \xi_1 = \frac{x_1 {\alpha_{2,2}}^2}{\alpha_{1,1}{\alpha_{2,2}}^2 + \alpha_{1,3}x_2\alpha_{2,2} + \alpha_{1,5}{x_2}^2} \\ \xi_2 = \frac{x_2}{\alpha_{2,2}} \end{cases} . \tag{3.108}$$

In the second most severe case the Gröbner basis method leads to a fourth order triangular system. The only acceptable solution can be written as

$$T^{-1} : \begin{cases} \xi_1 = \gamma_{1,0} + \gamma_{1,1}x_1 + \gamma_{1,2}x_2 + \frac{\sqrt[3]{p_1^{(3)} + \gamma_{1,3}\sqrt{r_1}}}{\gamma_{1,4}} + \frac{p_2^{(2)}}{\gamma_{1,5}\cdot\sqrt[3]{p_3^{(3)} + \gamma_{1,6}\sqrt{r_1}}} \\ \xi_2 = \gamma_{2,0} + \frac{\sqrt[3]{p_4^{(1)} + \gamma_{2,1}\sqrt{r_2}}}{\gamma_{2,2}} + \frac{p_5^{(1)}}{\gamma_{2,3}\cdot\sqrt[3]{p_6^{(3)} + \gamma_{2,4}\sqrt{r_2}}} \end{cases} ; \tag{3.109}$$

were r_1 and r_2 are polynomials of second degree, while $\gamma_{i,j}$ are coefficients depending on element geometry, while $p_j^{(k)}$ are polynomials, of degree k. Then, the T^{-1} has the following basis

$$\left[1\ , x_1, x_2, \sqrt[3]{p_1 + t_1\sqrt{r_1}}, \frac{p_2}{\sqrt[3]{p_3 + t_2\sqrt{r_1}}}, \sqrt[3]{p_4 + t_3\sqrt{r_2}}, \frac{p_2}{\sqrt[3]{p_6 + t_4\sqrt{r_2}}}\right], \tag{3.110}$$

where t_i are coefficients depending on $\gamma_{i,j}$. Now the analysis can be carried out and the main results are summarized in Table 3.13, where w_i are polynomials whose degree is indicated in the first column.

3.7 Some numerical tests

To numerically evaluate the effect of coordinate transformations on the element interpolation capability and, so, verify the analytical results discussed in the previous section, single element test are used. The idea is the same of the Continuum Region Element (CRE) method of Robinson [Robinson, 1985]–[Robinson, 1987a] (Figure 3.21). A single element of any shape is constructed and displacement are

Table 3.13: Higher degree transformations: quadratic curved-edge

degree	P2	Q2	P3	Q3
0°	1	1	1	1
1°	$x_1 \;\; x_2$	$x_1 \;\; x_2$	$x_1 \;\; x_2$	$x_1 \;\; x_2$
2°	$w_1(x_1x_2,\ldots)$	$w_2(x_1x_2,\ldots)$	$w_3(\ldots),\; w_4(\ldots)$	$x_1^2 \; x_1x_2 \; x_2^2$
3°	-	-	$w_5(\ldots)$	$w_6(\ldots),\; w_7(\ldots)$
4°	-	-	-	$w_8(\ldots)$

prescribed in all the nodes according to (a priori selected) analytical solution. Then the numerical solution is computed and checked against the analytical solution.

In the following the basic element geometries discussed in the previous section are considered, together with linear, quadratic and cubic prescribed displacement fields. The results are presented by plotting the stress field and the finite element stress field, and comparing locally the two solutions overlapping the contours of the exact solution (red line) and the contour of the FEM solution (blue line) with the local error. This plots confirm also that unfortunately the Barlow points not always remain superconvergent points as already pointed out in [Barlow, 1989]. As, experienced, when linear displacement field are prescribed, no errors were recorded, independently of the element type used.

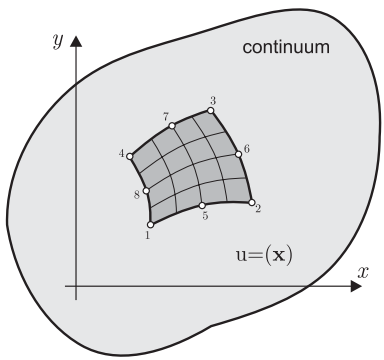


Figure 3.21: The Continuum Region Element test: displacements $u = (x)$ are prescribed

3.7.1 Bilinear transformations

Consider the example shown in Figure 3.16 where a simple distortion test is performed on a cantilever beam. Here we reproduce this distortion case using a CRE approach. As shown by Table 3.12 all the elements represent correctly the linear

displacement field, and so we do not report them. Differently for a quadratic displacement field some difference occurs as shown by Table 3.12. Actually P2 and Q2 elements have some difference due quadratic terms.

Consider first a quadratic displacement field. Figures 3.22 and 3.23 show that a P2 element is not capable to represent it correctly. Taking a cubic displacement field, Figures 3.24 and 3.25 shows the results for P2 elements. Figures 3.26 and 3.27 show that in this case also Q2 element is not capable to represent a quadratic displacement field correctly and suffer of the same error distributions of P2 element.

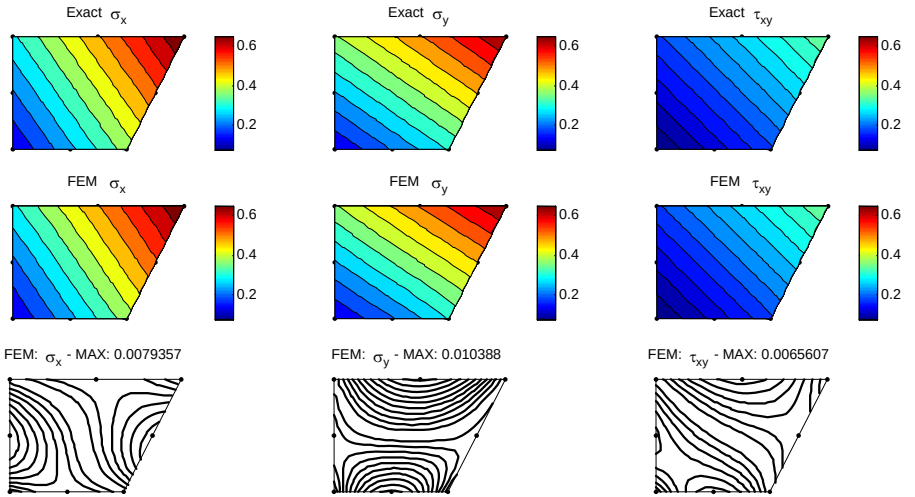


Figure 3.22: Element test for P8 elements: stress band plots stress for quadratic displacement field and error distributions with contour line step: 0.001

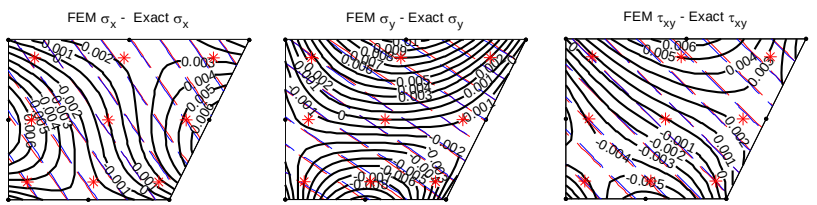


Figure 3.23: Element test for P8 elements: local error distributions for quadratic displacement field plotted on stress contour plots for exact solution (red line) and FEM solution (blu line). The red star markers show the positions of integrations points (contour interval: 0.001)

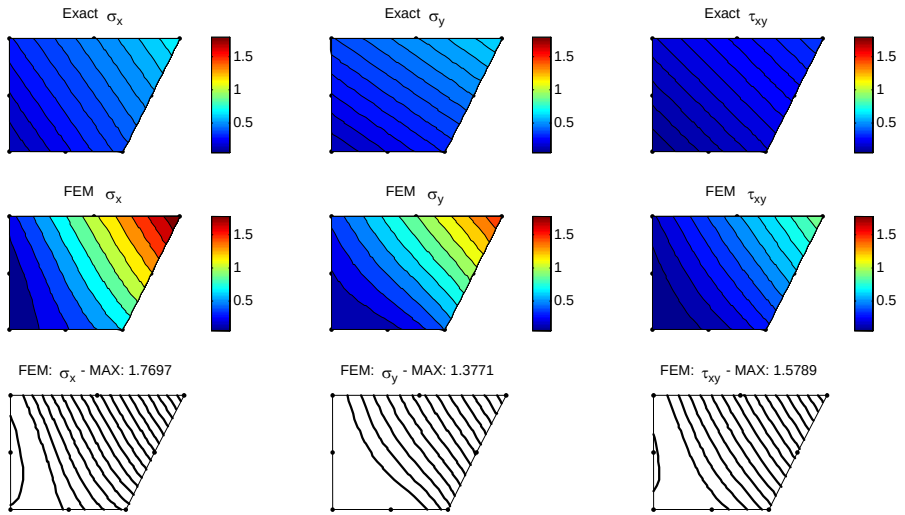


Figure 3.24: Element test for P8 elements: stress band plots stress for cubic displacement field and error distributions with contour line step: 0.1

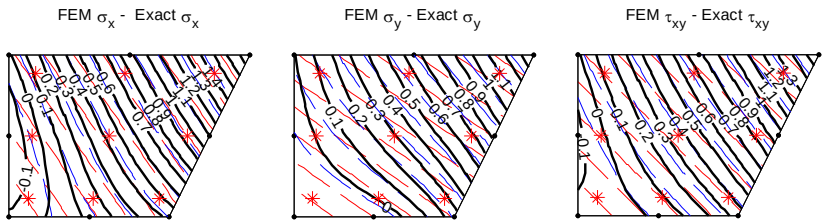


Figure 3.25: Element test for P8 elements: local error distributions for cubic displacement field plotted on stress contour plots for exact solution (red line) and FEM solution (blu line). The red star markers show the positions of integrations points (contour interval: 0.1)

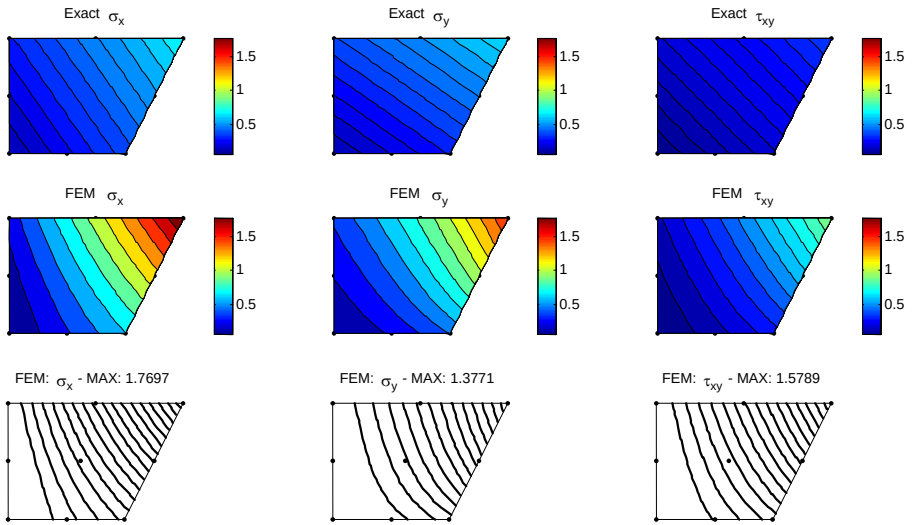


Figure 3.26: Element test for Q9 elements: stress band plots stress for cubic displacement field and error distributions with contour line step: 0.1

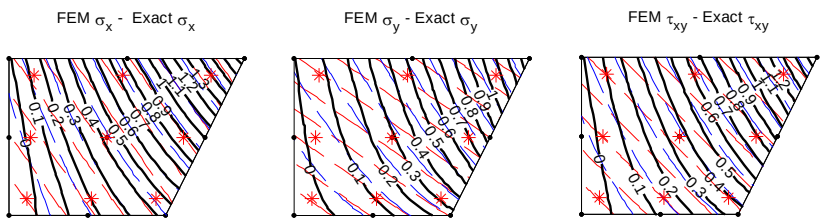


Figure 3.27: Element test for Q9 elements: local error distributions for cubic displacement field plotted on stress contour plots for exact solution (red line) and FEM solution (blu line). The red star markers show the positions of integrations points (contour interval: 0.1)

3.7.2 Higher order transformations

Here we intend to reproduce distortions of Figure 3.17 using the CRE approach. As shown by Table 3.13 and Figure 3.17 all the "curved edge" elements represent correctly the linear displacement field, and so we do not report them. On the contrary for a bilinear displacement field some problems occur as shown by the same table. Figures 3.28 and 3.29 show the results corresponding to a bilinear

displacement field, which confirm that a P2 element is not capable to correctly capture it. Moreover, Figures 3.30 and 3.31 confirm that Q2 element yields in this case a similar response. Summing up, the numerical results presented above confirm the results of the analysis presented in the previous section.

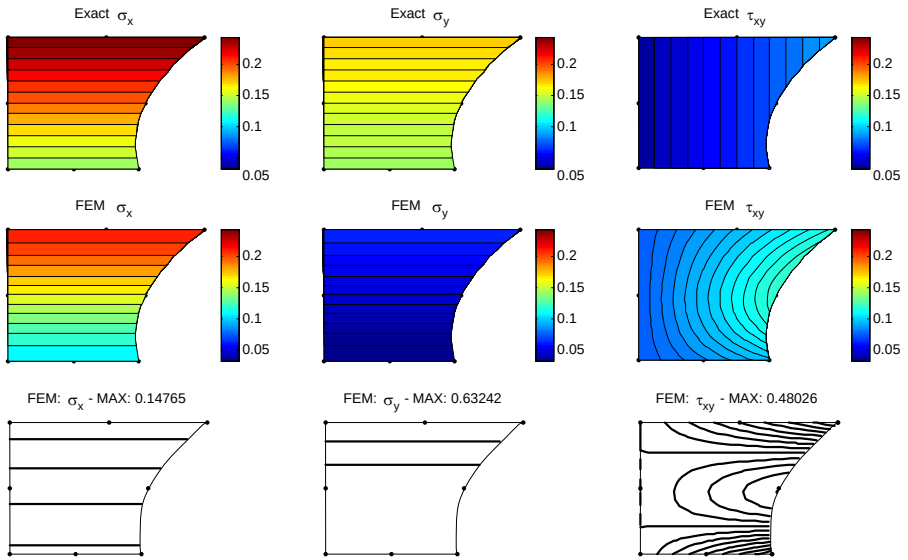


Figure 3.28: Element test for P8 elements: stress band plots stress for cubic displacement field and error distributions with contour line step: 0.02

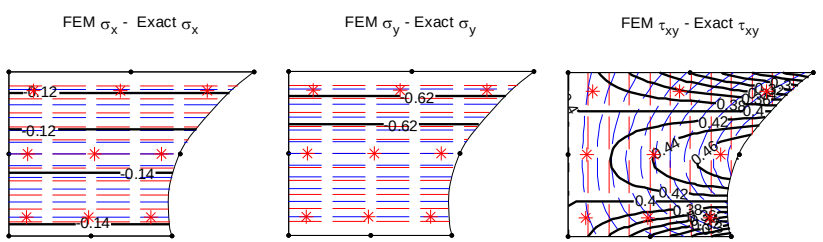


Figure 3.29: Element test for P8 elements: local error distributions for cubic displacement field plotted on stress contour plots for exact solution (red line) and FEM solution (blu line). The red star markers show the positions of integrations points (contour interval: 0.02)

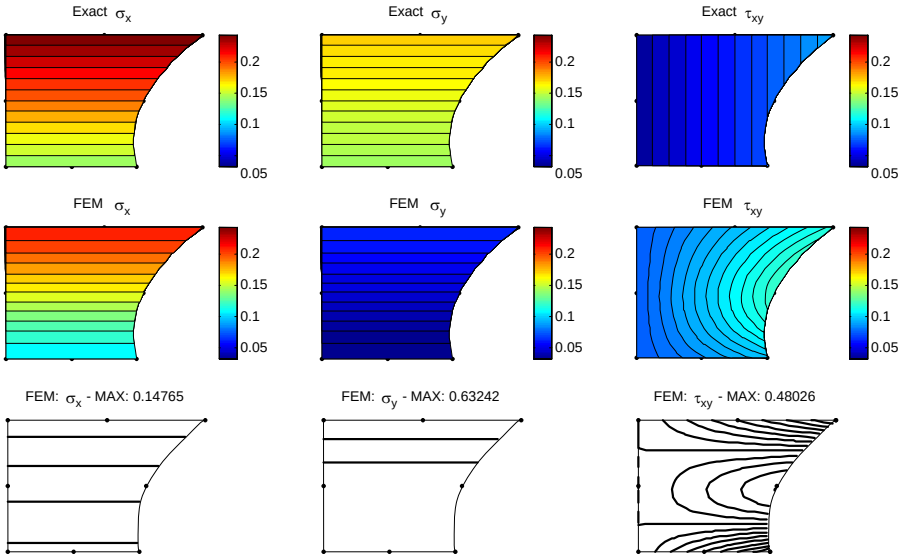


Figure 3.30: Element test for Q9 elements: stress band plots stress for cubic displacement field and error distributions with contour line step: 0.02

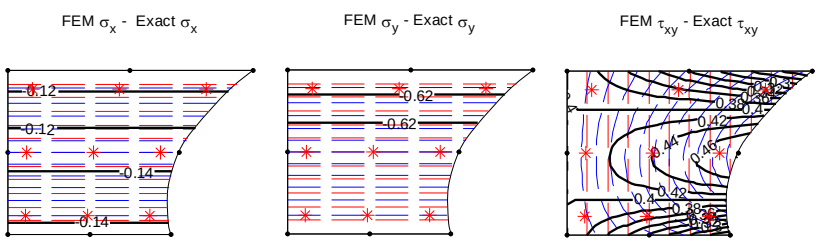


Figure 3.31: Element test for Q9 elements: local error distributions for cubic displacement field plotted on stress contour plots for exact solution (red line) and FEM solution (blu line). The red star markers show the positions of integrations points (contour interval: 0.02)

3.8 Distortion measures

One of the main concerns when generating a new mesh is the quality of elements produced. In particular, the efficiency of adaptive refinement analysis depends

on the shape of the elements, and so estimating the quality of element shape is a requisite during the adaptive analysis being performed. Unfortunately, most posterior error estimates can not evaluate the shape error of an element, so that some difficulties remain in the application of an adaptive analysis.

The use of element distortion factors (computable *a priori*) is, therefore, a necessary step in order to ensure that the elements shapes do not act as an extra cause of deterioration in finite element analysis accuracy. Element shape distortion can be defined in many forms, most notable of which are aspect ratios, skew, taper and warp. Robinson [Robinson, 1987b] used three of the latter properties to classify the distortion of four-noded quadrilaterals. A sensible choice of distortion measure would be the determinant of the element's Jacobian matrix. Robinson expressed the shape parameters in terms of the Jacobian determinant for four-noded quadrilaterals. However, the difficulty arising in linking each element's Jacobian determinant to such parameters (e.g., the aspect ratio or internal angles) for elements of displacement order higher than one with curved sides was pointed out. Eight further parameters in the form of the tangential and normal deviations for each of the four edges were later added [Robinson, 1988]. Although these 12 parameters for eight-noded curved quadrilaterals provide a high degree of accuracy, a higher computational workload is needed when calculating all twelve factors for each element.

In [El-Hamalawi, 2000] a simple element distortion factor is proposed, based only on interior angles, and here will be used as a comparison factor. In [Yuan & Pian, 1994] a set of distortion measuring parameters for the quadrilaterals hybrid stress membrane element is presented, based on Robinson's distortion parameters modified by geodesic coordinates, subsequently used in [Lautersztajn-S & Samuelsson, 1997b] and [Lautersztajn-S & Samuelsson, 1998b]. Thus, significant reduction of the number of these parameters, with respect to previous work, from 12 to 4, from is obtained.

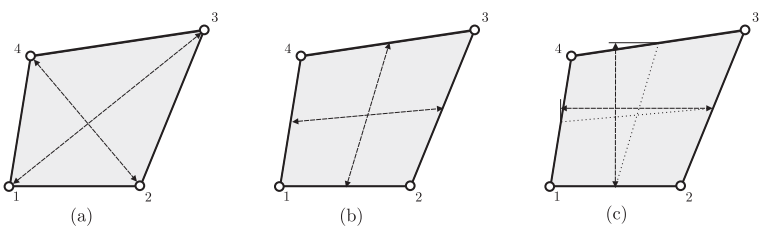


Figure 3.32: Simple distortione measure: (a) diagonal ratio c_1; (b) middle point distance ratio c_2; (c) middle point pojection ratio (on reference axes) c_3

Table 3.14: Test No.1: two element test.

			Distortion measure			
$\hat{\delta}$	η	e	c_0	c_1	c_2	c_3
0.0000	-16.1778	0.0000	0.0000	1.0000	1.0000	1.0000
0.1000	-15.6361	0.0335	0.2792	0.9050	0.9950	1.0000
0.2000	-14.2400	0.1198	0.5381	0.8198	0.9806	1.0000
0.3000	-12.5666	0.2232	0.7643	0.7442	0.9578	1.0000
0.4000	-11.0137	0.3192	0.9542	0.6778	0.9285	1.0000
0.5000	-9.7113	0.3997	1.1107	0.6202	0.8944	1.0000
0.6000	-8.6383	0.4660	1.2389	0.5708	0.8575	1.0000
0.7000	-7.7203	0.5228	1.3443	0.5293	0.8192	1.0000
0.8000	-6.8674	0.5755	1.4315	0.4953	0.7809	1.0000
0.9000	-5.9574	0.6318	1.5043	0.4681	0.7433	1.0000

3.8.1 Distortion measures for four node elements

A quality measure for quadrilaterals can be depend as [El-Hamalawi, 2000]

$$c_0 = \sqrt{\sum_{i=1}^{4} (\delta\theta_i)^2} \tag{3.111}$$

where $\delta\theta_i = \left|\frac{\pi}{2} - \theta_i\right|$ is the deviation of interior angles θ_i from the optimum value of $\frac{\pi}{2}$, corresponding to the rectangular shape. In addition to the above indicator, three more measures can be included [Castellazzi & Ubertini, 2004]: (a) diagonal ratio c_1; (b) middle point distance ratio c_2; (c) middle point pojection ratio (on reference axes) c_3 see Figure 3.32. Making reference to the notation defined in the previous section, the first parameter is given by

$$c_1 = \frac{\sqrt{\alpha_{1,1}^2 + 2\alpha_{1,1}\alpha_{1,2} + \alpha_{1,2}^2 + \alpha_{2,1}^2 + 2\alpha_{2,1}\alpha_{2,2} + \alpha_{2,2}^2}}{\sqrt{\alpha_{1,1}^2 - 2\alpha_{1,1}\alpha_{1,2} + \alpha_{1,2}^2 + \alpha_{2,1}^2 - 2\alpha_{2,1}\alpha_{2,2} + \alpha_{2,2}^2}}, \tag{3.112}$$

the second is given by

$$c_2 = \frac{\sqrt{\alpha_{1,2}^2 + \alpha_{2,2}^2}}{\sqrt{\alpha_{1,1}^2 + \alpha_{2,1}^2}}, \tag{3.113}$$

and the third parameter by

$$c_3 = \frac{|\alpha_{1,1}|}{|\alpha_{2,2}|}. \tag{3.114}$$

3.8.2 Test No.1

The test in Figure 3.16 is considered, with $h = 1$, $L = 2$, $M = 1$ and results for the tip deflection are given for the cases $v = 0.3$ and Young's modulus equal to one.

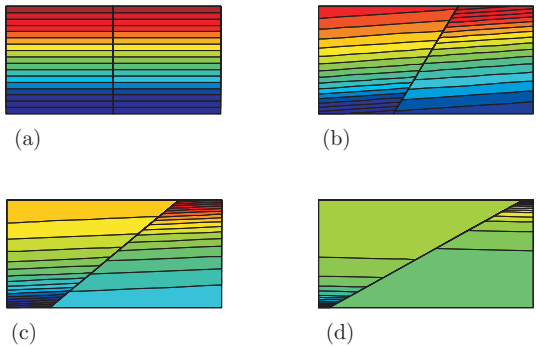


Figure 3.33: Contour bands of stress for a four-node element: (a) undistorted configuration; (b) angular distortion with $\hat{\delta} = 0.3$; (c) angular distortion with $\hat{\delta} = 0.6$; (d) angular distortion with $\hat{\delta} = 0.9$

In Table 3.14 the first column indicates the dimensionless distortion coefficient $\hat{\delta}$, (which varies from 0 to 1 and defines the mesh) and the second column the tip deflection η, the third column the percentage relative deviation

$$e = \frac{\bar{\eta} - \eta}{\bar{\eta}} \tag{3.115}$$

where $\bar{\eta}$ is the tip deflection when $\hat{\delta} = 0$. In addition the distortion measures c_0, c_1, c_2, c_3 are reported. In Figure 3.33 the stress bands plot of the σ_x stress component, are plotted for four different value of $\hat{\delta}$. It is clear the effect of progressive deterioration.

3.8.3 Test No.2

The cantilever beam of Figure 3.34 is analyzed with four node elements and different meshes. Young's modulus is set to 1, Poisson's coefficient to 0.3 and F to 1. The results are collected in Tables 3.15–3.18.

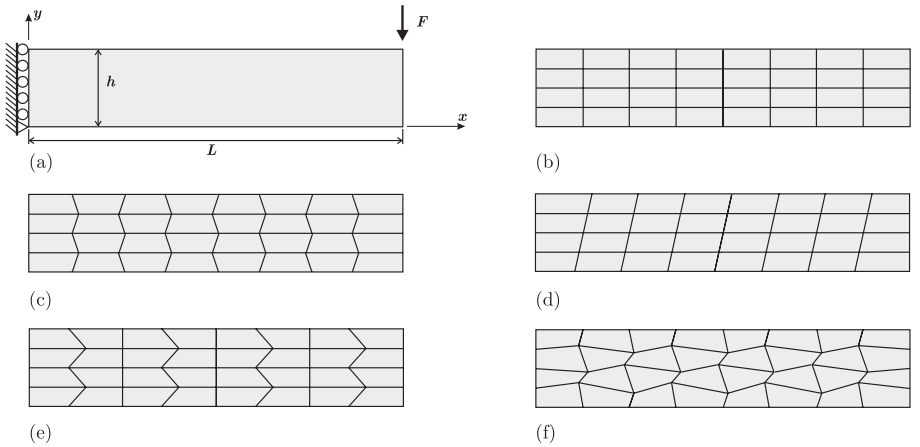


Figure 3.34: A beam problem: (a) the geometry and load; (b) discretization for undistorted configuration, reference solution; (c) linear distortion; (d) first bilinear distortion; (e) second bilinear distortion; (f) taper distortion; (g) arbitrary bilinear distortion;

Table 3.15: A beam problem: clamped beam along one end and vertically loaded at the free end with distorted element as shown in Figure (3.34)(c)

			Distortion measure.			
$\hat{\delta}$	η	e	c_0	c_1	c_2	c_3
0.0000	-3791.1856	0.0000	0.0000	1.0000	1.0000	1.0000
0.1000	-3786.8955	0.0011	0.0926	0.9938	0.9999	1.0000
0.2000	-3780.5241	0.0028	0.1847	0.9878	0.9997	1.0000
0.3000	-3772.0722	0.0050	0.2760	0.9820	0.9993	1.0000
0.4000	-3761.5454	0.0078	0.3659	0.9762	0.9988	1.0000
0.5000	-3748.9549	0.0111	0.4541	0.9707	0.9981	1.0000
0.6000	-3734.3160	0.0150	0.5402	0.9653	0.9972	1.0000
0.7000	-3717.6481	0.0194	0.6240	0.9601	0.9963	1.0000
0.8000	-3698.9729	0.0243	0.7053	0.9550	0.9951	1.0000
0.9000	-3678.3134	0.0298	0.7838	0.9500	0.9939	1.0000

Table 3.16: A beam problem: clamped beam along one end and vertically loaded at the free end with distorted element as shown in Figure (3.34)(c)

			Distortion measure.			
$\hat{\delta}$	η	e	c_0	c_1	c_2	c_3
0.0000	-3791.1856	0.0000	0.0000	1.0000	1.0000	1.0000
0.1000	-3786.8033	0.0012	0.0463	0.9784	0.9936	0.9938
0.2000	-3781.9247	0.0024	0.0926	0.9573	0.9869	0.9879
0.3000	-3776.5385	0.0039	0.1388	0.9367	0.9799	0.9820
0.4000	-3770.6324	0.0054	0.1847	0.9166	0.9727	0.9764
0.5000	-3764.1936	0.0071	0.2305	0.8971	0.9650	0.9708
0.6000	-3757.2084	0.0090	0.2760	0.8781	0.9571	0.9654
0.7000	-3749.6631	0.0110	0.3211	0.8596	0.9489	0.9601
0.8000	-3741.5434	0.0131	0.3659	0.8417	0.9404	0.9549
0.9000	-3732.8347	0.0154	0.4102	0.8243	0.9315	0.9498

Table 3.17: A beam problem: clamped beam along one end and vertically loaded at the free end with distorted element as shown in Figure (3.34)(c)

			Distortion measure.			
$\hat{\delta}$	η	e	c_0	c_1	c_2	c_3
0.0000	-3791.1856	0.0000	0.0000	1.0000	1.0000	1.0000
0.1000	-3786.7882	0.0012	0.0707	0.9753	0.9997	1.0000
0.2000	-3780.9705	0.0027	0.1410	0.9512	0.9988	1.0000
0.3000	-3773.7398	0.0046	0.2106	0.9278	0.9972	1.0000
0.4000	-3765.1218	0.0069	0.2792	0.9050	0.9950	1.0000
0.5000	-3755.1588	0.0095	0.3465	0.8828	0.9923	1.0000
0.6000	-3743.9076	0.0125	0.4122	0.8612	0.9889	1.0000
0.7000	-3731.4362	0.0158	0.4761	0.8402	0.9850	1.0000
0.8000	-3717.8206	0.0194	0.5381	0.8198	0.9806	1.0000
0.9000	-3703.1413	0.0232	0.5980	0.8001	0.9756	1.0000

Table 3.18: A beam problem: clamped beam along one end and vertically loaded at the free end with distorted element as shown in Figure (3.34)(c)

			Distortion measure.			
$\hat{\delta}$	η	e	c_0	c_1	c_2	c_3
0.0000	-3791.1856	0.0000	0.0000	1.0000	1.0000	1.0000
0.1000	-3788.0883	0.0008	0.1027	0.9861	0.9861	0.9861
0.2000	-3782.5727	0.0023	0.2071	0.9725	0.9726	0.9725
0.3000	-3774.5421	0.0044	0.3150	0.9591	0.9595	0.9593
0.4000	-3763.8307	0.0072	0.4279	0.9460	0.9468	0.9463
0.5000	-3750.1957	0.0108	0.5468	0.9332	0.9345	0.9336
0.6000	-3733.3006	0.0153	0.6724	0.9205	0.9227	0.9212
0.7000	-3712.6732	0.0207	0.8049	0.9081	0.9113	0.9090
0.8000	-3687.5855	0.0273	0.9437	0.8959	0.9004	0.8970
0.9000	-3656.5848	0.0355	1.0876	0.8839	0.8900	0.8852

Chapter 4

A posteriori error estimation

Nella prima parte di questo capitolo, vengono introdotti i concetti di base su cui si fonda la stima a posteriori dell'errore di discretizzazione in analisi agli elementi finiti.

In seguito, l 'attenzione é incentrata sui cosiddetti "patch recovery methods", ovvero sui metodi che stimano l 'errore per confronto della soluzione agli elementi finiti con una più accurata ottenuta previa opportuna ricostruzione (recovery) del campo di sforzi operata localmente su piccoli insiemi di elementi contigui (patch). In questo contesto, sono illustrate alcune delle più note tecniche presenti in letteratura e viene sviluppata una nuova procedura di ricostruzione locale degli sforzi, denominata RCP (Recovery by Compatibility in Patches), ottenuta rilassando la compatibilità cinematica, attraverso la minimizzazione di un funzionale tipo energia complementare, definito sulla patch.

Nell'ultima parte del capitolo, viene condotta un 'ampia sperimentazione numerica che mostra come la nuova procedura risulti versatile, robusta e superconvergente.

4.1 Aims and concepts of a posteriori error estimation

The remarkable success of some a posteriori error estimators has opened a new chapter in computational mechanics. By effectively estimating error, the possibility of controlling the entire computational process through new adaptive algo-

rithms emerges. Fresh criteria for judging the performance of algorithms become apparent. Most importantly, the analyst can use a posteriori error estimates as an independent measure of the quality of the simulation under study.

The a priori estimation of errors in numerical methods has long been an enterprise of numerical analysts. Such estimates give information on the convergence and stability of various solvers and can give rough information on the asymptotic behavior of errors in calculations as mesh parameters are appropriately varied. Traditionally, the practitioner using numerical simulations, while aware that errors exist, is rarely concerned with quantifying them. The quality of a simulation is generally assessed by physical or heuristic arguments based on the experience and judgment of the analyst. Frequently such arguments are later proved to be flawed.

Some of the earliest a posteriori error estimates used in computational mechanics were developed for the solution of ordinary differential equations. These are typified by *predictor-corrector* algorithms in which the difference in solutions obtained by schemes with different orders of truncation error is used as rough estimates of the error. This estimate can in turn be used to adjust the discretization step. It is notable that the original a posteriori error estimation schemes for elliptic problems had many features that resemble those for ordinary differential equations.

Modern interest in a posteriori error estimation for finite element methods for two point elliptic boundary value problems began with the pioneering work of Babuška and Rheinboldt [Babuška & Rheinboldt, 1978]-[Babuška & Rheinboldt, 1979]. A posteriori error estimation techniques were developed by investigating the residuals occurring in a patch of elements or even in a single element what makes possible to estimate the errors in energy, or an energy norm, which arise locally. The original approaches for error estimation and indeed for correcting these by adaptive refinement started by simply looking at the elements with largest error and dividing these to achieve some acceptable accuracy. These formed the basis of adaptive meshing procedures designed to control and minimize the error.

The use of complementary energy formulations for obtaining a posteriori error estimates was put forward by de Veubeke [de Veubeke, 1965]. However, the method failed to gain popularity being based an a global computation. The idea of solving element by element complementary problems together with the important concept of constructing equilibrated boundary data to obtain error estimates was advanced by Ladeveze and Leguillon [Ladeveze & Leguillon, 1983]. Related ideas are found in the work of Kelly [Kelly, 1984].

During the early 1980s the search for effective adaptive methods led to a wide variety of ad hoc error estimators. Many of these were based on a priori or interpolation estimates, that provided crude but effective indications of features of error sufficient to drive adaptive processes.

Since these early works many players entered the field and today the procedures available for error estimation are essentially reduced to two kinds. These are, first, the residual error estimators, continuing the original suggestions of Babuška but the emphasis is now on use of self-equilibrating patches. Here the work of Ainsworth and Oden [Ainsworth & Oden, 1992], [Ainsworth & Oden, 1993], and others is important. Such residual approaches have gained universal mathematical aplomb and for some years formed the basis of acceptable error estimators.

The second kind of procedures was initially suggested by Zienkiewicz and Zhu in 1987 [Zienkiewicz & Zhu, 1987] and consist of using a recovery process to obtain a more accurate representation of the unknowns. In the original procedure, simple averaging and the so-called L_2 projection recovery of Brauchli and Oden [Brauchli & Oden, 1971] were used to estimate errors. This is a simple error estimation technique that is effective for some classes of problems and types of finite element approximations. Indeed, their method started the new category of recovery-based methods: gradients of solutions obtained on a certain domain partition are smoothed and then compared with the gradients of the original solution to assess error. This approach was later modified [Zienkiewicz & Zhu, 1992a]-[Zienkiewicz & Zhu, 1992b], leading to the well-known Superconvergent Patch Recovery (SPR) method, which provides very accurate recovered values that generally converge at a higher rate than the original finite element solution so giving a solid basis for error estimation.

In the SPR, the standard process was to find, in a patch formed by the elements surrounding a node or a element, a set of points at which the accuracy of the finite element solution was higher, such as the superconvergence points, and then to pass a least squares approximation of the function of the stresses or gradients (as in the case of thermal problems) through such values. With the knowledge of these more accurate stresses or gradients of the solution, the energy norm could easily be estimated and, indeed, so could the local values of stresses and gradients [Lee *et al.* , 1997]–[Lo & Lee, 1998].

Today the methodology is widely used, industrially, following the same practice, but some variants of it have been introduced. In one of these variants the superconvergent values are taken not as the stresses or gradients of the solution, but on the values of the original functions themselves. It is known for instance that superconvergence will always occur in these values for triangular elements

at the vertex nodes, whereas in the same triangular elements such superconvergence for the stresses or gradients does not exist. Therefore for triangular elements it is preferable to include in the patch superconvergence of the displacements or the basic function itself and then to find its gradients by establishing a suitable polynomial of higher order. This procedure of obtaining the recovered values is currently widely used and many papers by Wiberg have contributed very much to this recovery methodology [Wiberg & Abdulwahab, 1993]-[Wiberg, 1997]-[Wiberg & Abdulwahab, 1997a]-[Wiberg & Abdulwahab, 1997b]-[Wiberg & Li, 1999].

The subject of a posteriori error estimation for finite element approximation has now reached maturity. In fact, study of robustness of existing estimators and identifying limits on their performance has gained importance. Noteworthy in this respect is the work of Babuška et al. [Babuška *et al.* , 1994a]-[Babuška *et al.* , 1994b] who conducted an extensive study on the performance and robustness of the main error estimation techniques. This methodology is essentially a sort of patch test (*Babuška patch test*) in which arbitrary fields of finite elements, assumed to be repeatable in space, are subjected to the next higher order basic polynomial contained in the original shape functions. Obviously if the finite elements used include quadratics any solution which is quadratic terms will be exactly returned through the numerical computation and the main errors will be concentrated only in the cubic terms. Thus it is necessary, in a repeatable sense, to impose only cubic polynomials on the patch and test whether errors arise and, if so, their magnitude. By doing this it is easy to establish whether the procedures capture errors exactly or simply show the errors for inexact solutions. Here Babuška proposed to value each error estimator with a robustness index and after many tests concluded that the robustness index was definitively optimal for the recovery methods, with the residual methodology giving much poorer results.

4.1.1 Residual based approaches

The residual type of error estimator was first introduced by Babuška and Rheinboldt [Babuška & Rheinboldt, 1978] and their research was later followed by many others [Ladeveze & Leguillon, 1983]–[Babuška & Miller, 1987]–[Ainsworth & Oden, 1992]–[Ainsworth & Oden, 1993]–[Ainsworth & Oden, 1997]. This type of error estimator is computed by using the residuals of the finite element solution implicitly or explicitly. The term *explicit* or *implicit* is used in a context similar to the characterization of explicit and implicit difference schemes for time-marching algorithms: explicit schemes involve a direct computation using available data while implicit schemes involve the solution of an algebraic system of equations. The implicit residual error estimators are determined by solving local boundary value

problems for the error, which use the residuals of the finite element solution as data. The explicit residual error estimators are, on the other hand, expressed directly by the residuals of the discretization solution. As might be expected, explicit schemes generally require less computational effort than implicit schemes but involve compromises in robustness and in utility as means for accurate, quantitative error estimation. The accuracy of the implicit residual error estimator depends on how the local problem is solved. In particular, how the boundary conditions of the local problem are set. Several techniques have been developed to split the flux on the boundary of the element by satisfying some equilibrium conditions. The accuracy of the explicit residual error estimator depends on the accuracy of the constants contained in the error estimator. The robustness of this type of error estimator, as shown by Babuška and coworkers [Babuška *et al.* , 1994a]–[Babuška *et al.* , 1994b]–[Babuška *et al.* , 1995b]–[Babuška *et al.* , 1995a]–[Babuška *et al.* , 1997a]–[Babuška *et al.* , 1997c]–[Babuška *et al.* , 1998], depends on the regularity of the problem and the mesh used in the finite element analysis.

4.1.2 Recovery based approaches

The finite element method produces the optimal approximation from the finite element subspace. However, it is frequently the case that the finite element analyst is more interested in the gradient of the finite element approximation, than in the approximation itself. For instance, in computational elasticity, the stresses and strains are the primary concern, rather than the displacements. Furthermore, the normal component of the gradient of the approximation is generally discontinuous across the element boundaries, meaning that the practitioner is presented with a discontinuous approximation to the main quantity of interest. For this reason, many finite element packages incorporate a post-processing procedure whereby the discontinuous approximation to the gradient is smoothed before being presented to the user.

The reasons for performing such a post-processing are not purely cosmetic: under certain circumstances it is found that the accuracy of the smoothed gradient is superior to the approximation provided by the untreated gradient of the original finite element approximation. A rather natural approach to a posteriori error estimation is based on measuring the difference between the direct and post-processed approximations to the gradient. This ostensibly rather crude approach can result in astonishingly good estimates for the true error. In this approach, the error estimator is computed by first using some recovery techniques to achieve a more accurate, or even superconvergent solution from the finite element approximation. The recovered solution is then used in place of the exact solution to compute the

error:

$$\textit{exact error} = \textit{FEM solution - exact solution,} \tag{4.1}$$

$$\textit{estimated error} = \textit{FEM solution - recovered solution.} \tag{4.2}$$

This type of error estimator is very easy to implement and is computationally very efficient. Obviously the accuracy of the error estimator depends on the accuracy of the recovered solution.

The appeal of such procedures to practitioners is easy to appreciate. The ready availability of a post-processed gradient in the finite element package means that the estimator may be easily implemented. This approach allows considerable leeway in the selection of the post-processed gradient, and there are as many different estimators as there are post-processing techniques.

One class of techniques that is particularly popular with the engineering community is averaging methods. The gradient of the finite element approximation provides a discontinuous approximation to the true gradient. This may be used to construct an approximation at each node by averaging contributions from each of the elements surrounding the node. These values may then be interpolated to obtain a continuous approximation over the whole domain. The specific steps used to construct the averaged gradient at the nodes distinguish the various estimators and have a major influence on the accuracy and robustness of the ensuing estimator.

Nodal averaging of derivatives (or stresses) has been practised since the inception of the finite element method to present the user with more realistic results. Although the basis of the process was initially intuitive its success is closely linked with the existence of superconvergent points in many finite element solutions at which the normal rate of theoretical convergence is exceeded by one order. The existence of such optimal sampling points has been suggested quite early [Barlow, 1976]-[Barlow, 1977] and has been subject to much research since. Extrapolation from such superconvergent points and subsequent averaging at nodes (local projection) gives clearly a sound basis for improved nodal values obtainable for linear elements in which the superconvergent point is that of the element centroid for quadrilaterals and that of a mid side of triangles. A widely used technique of Hinton and Campbell [Hinton & Campbell, 1974] for quadrilateral quadratic elements follows a similar procedure using the 2×2 Gauss points sampling and bilinear extrapolation.

Different strategies were developed in recent years, namely, simple nodal averaging, L_2-projections (see [Brauchli & Oden, 1971]), and the so called Superconvergent Patch Recovery (SPR) technique, introduced by Zienkiewicz and Zhu

[Zienkiewicz & Zhu, 1992c], [Zienkiewicz & Zhu, 1992a], [Zienkiewicz & Zhu, 1992b]-[Zienkiewicz *et al.* , 1993]-[Zienkiewicz & Zhu, 1995]. Meanwhile, various enhancements of the SPR technique have been presented (see e.g. Wiberg and Abdulwahab [Wiberg & Abdulwahab, 1993]-[Wiberg, 1997]-[Wiberg & Abdulwahab, 1997a]-[Wiberg & Abdulwahab, 1997b], Blacker and Belytschko [Blacker & Belytschko, 1994], Kvamsdal and Okstad [Kvamsdal & Okstad, 1998], Duarte [Duarte & Dutra Do Carmo, 2000] and Gu [Gu *et al.* , 2004]).

4.2 Some recovery methods

In this section, two of the most popular recovery based methods, that is the SPR method, the REP method are illustrated.

4.2.1 Superconvergent Patch Recovery: SPR

The Superconvergent Patch Recovery procedure was introduced by Zienkiewicz and Zhu in 1992 [Zienkiewicz & Zhu, 1992c] and is characterized by single and continuous polynomial expansion of the function describing the derivatives used on an element patch surrounding the nodes at which recovery is desired. This expansion can be made to fit locally the superconvergent points in a least square manner or simply be an L_2 projection of the consistent finite element derivatives. The first of these processes will be shown always to lead to superconvergent recovery of nodal derivatives or stresses. The second, although not generally superconvergent, will still always result in a considerable improvement of recovered nodal values.

In the previous chapter we have shown that sampling of the gradients and stresses at some particular points is generally optimal and possesses a higher order accuracy when such points are superconvergent. However, we would also like to have similarly accurate quantities elsewhere within each element for general analysis purposes, and in particular we need such highly accurate gradients and stresses when the energy norm or other similar norms have to be evaluated in error estimates. We have already shown how with some elements very large errors exist beyond the superconvergent point and attempts have been made from the earliest days to obtain a complete picture of stresses which is more accurate overall. Here attempts are generally made to recover the nodal values of stresses and gradients from those sampled internally and then to assume that throughout the element the recovered stresses $\boldsymbol{\sigma}^*$ are obtained by interpolation in the same manner as the displacements by, first, recovering nodal values $\bar{\boldsymbol{\sigma}}^*$ and, then, interpolating these values by standard shape functions:

$$\boldsymbol{\sigma}^* = \mathbf{U}\bar{\boldsymbol{\sigma}}^*, \tag{4.3}$$

where $\mathbf{U}$ are the same basis functions as for the interpolation of displacements in Equation (2.22). Thus, the key step is to determine the nodal values $\bar{\boldsymbol{\sigma}}^*$. This is a process used almost from the beginning of finite element calculations for triangular elements, where elements are sampled at the centroid (assuming linear shape functions have been used) and then the stresses are averaged at nodes. However this is not the best for triangles and for higher order elements such averaging is inadequate.

4.2.1.1 Implementation of the SPR procedure

Local patches of elements are considered, assembled around a central node. If we accept the superconvergence of $\boldsymbol{\sigma}^h$ at certain points s in each element, then it is a simple matter to compute $\boldsymbol{\sigma}^*$ which is superconvergent at all points within the element. The procedure is illustrated for two dimensions in Figure 4.1.

At the superconvergent points the values of $\boldsymbol{\sigma}^h$ are accurate to order p and not $p-1$ as is true elsewhere (p is the order of the polynomials in the dispacement shape functions). However, we can easily obtain an approximation given by a polynomial of degree p, with identical order to these occurring in the shape function for displacement, which has superconvergent accuracy everywhere if this polynomial is made to fit the superconvergent points in a least square manner.

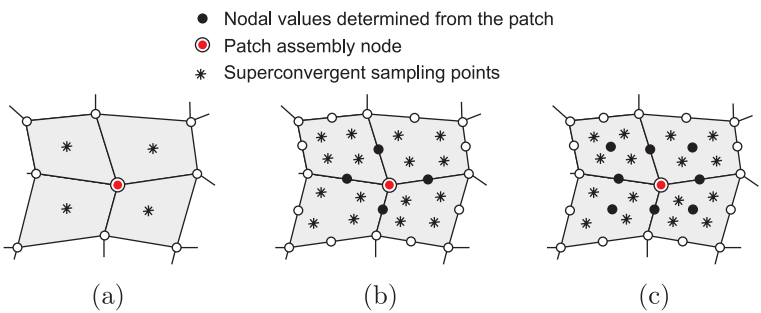


Figure 4.1: Interior patches for quadrilateral elements: (a) 4-nodes elements; (b) 8-node elements; (c) 9-node elements

Thus we proceed by introducing the following approximation for the recovered solution over the patch

$$\boldsymbol{\sigma}_p^* = \mathbf{P}^* \mathbf{a}. \tag{4.4}$$

where $\mathbf{P}^*$ is a matrix of approximation functions, $\mathbf{a}$ is a vector of unknown parameters and subscript $\cdot_p$ emphasizes that the recovery is done over the patch.

For instance, in two dimensional elasticity $\mathbf{P}^*$ and $\mathbf{a}$ are given by

$$\boldsymbol{\sigma}^* = \begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \tau_{12} \end{bmatrix}, \ \mathbf{P}^* = \begin{bmatrix} \tilde{\mathbf{P}}(\mathbf{x}) & & \\ & \tilde{\mathbf{P}}(\mathbf{x}) & \\ & & \tilde{\mathbf{P}}(\mathbf{x}) \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix}, \tag{4.5}$$

where $\tilde{\mathbf{P}}(\mathbf{x})$ is a $1 \times m$ vector that contains the polynomials terms in $\mathbf{x}$ coordinates, $\mathbf{a}_i$ is a $m \times 1$ vector of unknown coefficients and m is the number of terms used to form the polynomial bases of order p

$$\tilde{\mathbf{P}}(x, y) = \left[1, \ x_1, \ x_2, \ \ldots, \ x_1^p, \ x_1^{p-1}x_2, \ \ldots, \ x_1x_2^{p-1}, \ x_2^p\right]. \tag{4.6}$$

To evaluate the vector of unknown parameters $\mathbf{a}$, we minimize, the following functional

$$\begin{aligned} F(\mathbf{a}) &= \sum_{i=1}^{n} \left[\boldsymbol{\sigma}^h(\mathbf{x}_i) - \boldsymbol{\sigma}_p^*(\mathbf{x}_i)\right]^2 = \\ &= \sum_{i=1}^{n} \left[\boldsymbol{\sigma}^h(\mathbf{x}_i) - \mathbf{P}^*(\mathbf{x}_i)\mathbf{a}\right]^2, \end{aligned} \tag{4.7}$$

over an assigned patch Ω_p

$$\Omega_p = \bigcup_{j=1}^{nep} \Omega_j, \tag{4.8}$$

formed by nep elements, where $\mathbf{x}_i$ is the i-th superconvergent point, and $n = nep \times k$ is the amount of sampling points being k the number of sampling points for element.

Minimization of $F(\mathbf{a})$ yields:

$$\frac{\partial F(\mathbf{a})}{\partial \mathbf{a}} = \sum_{i=1}^{n} \mathbf{P}^{*\mathrm{T}}(\mathbf{x}_i)\mathbf{P}^*(\mathbf{x}_i)\mathbf{a} - \sum_{i=1}^{n} \mathbf{P}^{*\mathrm{T}}(\mathbf{x}_i)\boldsymbol{\sigma}^h(\mathbf{x}_i) = \mathbf{0}, \tag{4.9}$$

or in compact form

$$\mathbf{H}\mathbf{a} = \mathbf{g}, \tag{4.10}$$

where

$$\mathbf{H} = \sum_{i=1}^{n} \mathbf{P}^{*\mathrm{T}}(\mathbf{x}_i)\mathbf{P}^*(\mathbf{x}_i), \tag{4.11}$$

and

$$\mathbf{g} = \sum_{i=1}^{n} \mathbf{P}^{*\mathrm{T}}(\mathbf{x}_i)\boldsymbol{\sigma}^h(\mathbf{x}_i). \tag{4.12}$$

Equation (4.10) can now be solved to compute the coefficients $\mathbf{a}$

$$\mathbf{a} = \mathbf{H}^{-1}\mathbf{g}. \tag{4.13}$$

The availability of $\boldsymbol{\sigma}_p^*$ allows the values of $\bar{\boldsymbol{\sigma}}^*$ to be determined at all nodes. As some nodes belong to more than one patch, average values of $\bar{\boldsymbol{\sigma}}^*$ are best obtained. Finally, the recovered solution $\boldsymbol{\sigma}^*$ throughout each element is achieved by Equation (4.3).

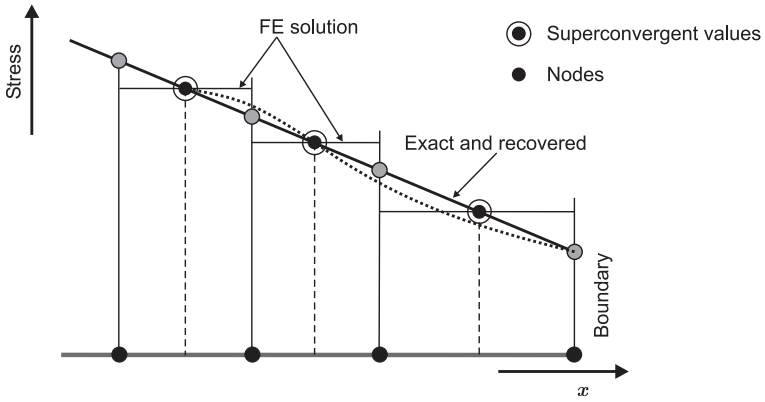


Figure 4.2: Recovery of exact σ of degree p by linear elements ($p = 1$)

In Figure 4.2 and 4.3 we show, in a one-dimensional example, how the Superconvergent Patch Recovery works. The Superconvergent Patch Recovery can be extended to produce superconvergent displacements. The procedure for the displacements is quite simple if we assume the superconvergent points to be at nodes of the patch. However, as we have already observed it is always necessary to have more data than the number of coefficients in the particular polynomial to be able to execute a least square minimization. Here of course we occasionally need a patch which extends further than before, particularly since the displacements will be given by a polynomial one order higher than that used for the shape functions.

The recovered solution $\boldsymbol{\sigma}^*$ has on occasion been used in dynamic problems, because in dynamic problems the displacements themselves are often important.

4.2.1.2 L_2 recovery variant

An alternative procedure lies on the local L_2 projection to obtain the polynomial solution locally on each patch. The functional $F(\mathbf{a})$ to be minimized is given by

$$F(\mathbf{a}) = \int_{\Omega_p} (\boldsymbol{\sigma}^h - \boldsymbol{\sigma}_p^*)^2 \mathrm{d}V = \int_{\Omega_p} (\boldsymbol{\sigma}^h - \mathbf{P}^*\mathbf{a})^2 \mathrm{d}V, \tag{4.14}$$

where, similar to (4.8), the patch domain is defined by

$$\Omega_p = \bigcup_{j=1}^{nep} \Omega_j. \tag{4.15}$$

Minimizing $F(\mathbf{a})$ yields

$$\int_{\Omega_p} \mathbf{P}^{*\mathrm{T}}\mathbf{P}^* \mathrm{d}V \mathbf{a} = \int_{\Omega_p} \mathbf{P}^{*\mathrm{T}} \boldsymbol{\sigma}^h \mathrm{d}V, \tag{4.16}$$

and again

$$\mathbf{a} = \mathbf{H}^{-1}\mathbf{g}, \tag{4.17}$$

where

$$\mathbf{H} = \int_{\Omega_p} \mathbf{P}^{*\mathrm{T}} \mathbf{P}^* \mathrm{d}V \quad \text{and} \quad \mathbf{g} = \int_{\Omega_p} \mathbf{P}^{*\mathrm{T}} \boldsymbol{\sigma}^h \mathrm{d}V. \tag{4.18}$$

It is interesting to note that if we use the numerical integration to compute the Equation (4.18), we obtain an analogous expression to Equation (4.11); in this case each terms of the sum will be multiplied by a weight coefficient, and we obtain the same results obtained before if and only if the integration points are the same of the superconvergence points.

The L_2 recovery gives much inferior answers, showing superconvergence only for odd values of p and almost no improvement for even values of p.

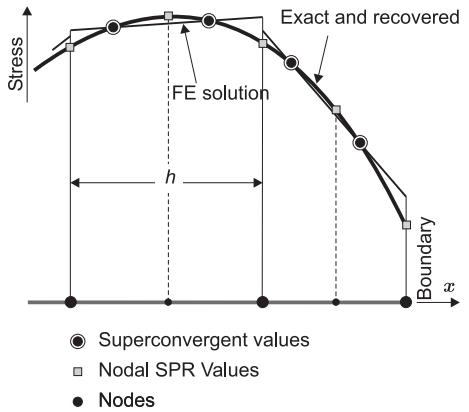


Figure 4.3: Recovery of exact σ of degree p by quadratic elements ($p = 2$)

4.2.2 Recovery by Equilibrium in Patches: REP

Although SPR has proved to work well generally, the reason behind its capability of producing an accurate recovered solution even when superconvergent points do not in fact exist remains an open question. An alternative, known by the acronym REP (Recovery by Equilibrium in Patches), was first presented in [Boroomand & Zienkiewicz, 1997b] and later improved in [Boroomand & Zienkiewicz, 1997a]-[Boroomand & Zienkiewicz, 1999]-[Boroomand *et al.* , 1999].

To some extent the motivation is similar to that of Ladeveze et al. [Ladeveze & Leguillon, 1983] who sought to establish (for somewhat different reasons) a fully equilibrating stress field which can replace that of the finite element approximation. The method operates in the some way of SPR. It is aimed at recovering a new solution $\boldsymbol{\sigma}_p^*$ over patches of elements. These solutions are used to compute the

nodal values $\bar{\boldsymbol{\sigma}}_p^*$ which are formally interpolated by Equation (4.3) to determine the recovered stresses over the whole domain.

4.2.2.1 Implementation of the REP procedure

Consider a patch, as shown in Figure 4.4. The finite element solution derives by imposing the discrete equilibrium of all elements. Thus also every isolated patch of elements Ω_p will be in equilibrium. This can be shown by writing the equilibrium equations for such a patch as

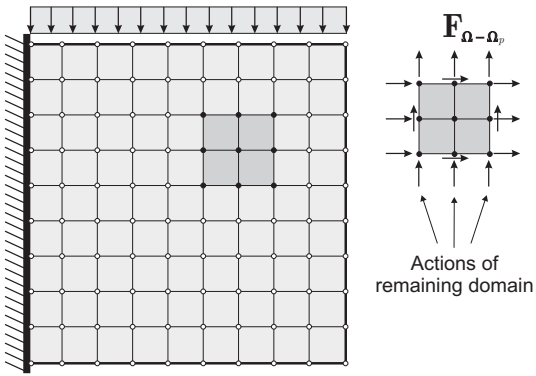


Figure 4.4: Example of a patch of elements

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}^h \mathrm{d}V - \int_{\Omega_p} \mathbf{U}^{\mathrm{T}} \mathbf{b} \mathrm{d}V - \int_{\partial\Omega_t^p} \mathbf{U}^{\mathrm{T}} \bar{\mathbf{t}} \mathrm{d}V + \mathbf{F}_{\Omega-\Omega_p} = \mathbf{0}, \tag{4.19}$$

where $\mathbf{F}_{\Omega-\Omega_p}$ represents the forces from the domain $\Omega - \Omega_p$ and its exterior boundary (see Figure 4.4) and $\mathbf{B}$ is given by Equation (2.24). Thus, we can write for any patch

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}^h \mathrm{d}V = \mathbf{g}, \tag{4.20}$$

with

$$\mathbf{g} = -\mathbf{F}_{\Omega-\Omega_p} - \mathbf{F}_{\Omega_p}, \tag{4.21}$$

where $\mathbf{F}_{\Omega_p}$ are the equilibrant nodal forces:

$$\mathbf{F}_{\Omega_p} = -\int_{\Omega_p} \mathbf{U}^{\mathrm{T}} \mathbf{b} \mathrm{d}V - \int_{\partial\Omega_t^p} \mathbf{U}^{\mathrm{T}} \bar{\mathbf{t}} \mathrm{d}S. \tag{4.22}$$

Now assuming the forces on the right-hand side of Equation (4.22) are known, the patch of elements can be considered as a separate system on which these forces are acting as nodal forces. Thus, again applying the virtual work principle, we can

write for any stress system $\boldsymbol{\sigma}_p^*$ the discrete equilibrium equation as

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}_p^* \mathrm{d}V = \mathbf{g}, \tag{4.23}$$

and

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}_p^* \mathrm{d}V = \int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}^h \mathrm{d}V, \tag{4.24}$$

and consequently

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}_p^* \mathrm{d}V = \int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{C} \mathbf{B} \bar{\mathbf{u}}^h \mathrm{d}V. \tag{4.25}$$

At this stage we shall assume a continuous form of stresses over the patch. This can be done in several ways:

1. smoothing the stresses by defining these in a polynomial form:

$$\boldsymbol{\sigma}_p^* = \mathbf{P}^* \mathbf{a}, \tag{4.26}$$

resulting in

$$\left(\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{P}^* \mathrm{d}V \right) \mathbf{a} = \left(\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{C} \mathbf{B} \mathrm{d}V \right) \bar{\mathbf{u}}^h; \tag{4.27}$$

2. smoothing the strains:

$$\boldsymbol{\epsilon}_p^* = \mathbf{P}^* \mathbf{a}, \tag{4.28}$$

resulting in

$$\left(\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{C} \mathbf{P}^* \mathrm{d}V \right) \mathbf{a} = \left(\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{C} \mathbf{B} \mathrm{d}V \right) \bar{\mathbf{u}}^h; \tag{4.29}$$

3. or smoothing the displacements:

$$\mathbf{u}_p^* = \mathbf{P}^* \mathbf{a}, \tag{4.30}$$

resulting in

$$\left(\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{C} \mathbf{D} \mathbf{P}^* \mathrm{d}V \right) \mathbf{a} = \left(\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{C} \mathbf{B} \mathrm{d}V \right) \bar{\mathbf{u}}^h. \tag{4.31}$$

In all the above cases the number of unknowns should be equal to or less than the number of equations implied in Equation (4.25). In the third case the rigid-body motion must be omitted from series of unknowns. One convenient way for this is to consider the displacements of all boundary nodes of the patch and minimize the square difference of recovered values and those of finite element method. As already mentioned, to achieve the recovery we proceed in an exactly analogous way to that used in the SPR procedure, first approximating the stress in each patch by a polynomial of appropriate order $\boldsymbol{\sigma}_p^*$, second using this approximation to obtain nodal values of $\bar{\boldsymbol{\sigma}}^*$ and finally interpolating these values by standard

shape functions.

Therfore, in what follows we shall only use the first type of smoothing and assume that

$$\boldsymbol{\sigma}_p^* = \mathbf{P}^* \mathbf{a}, \tag{4.32}$$

and so

$$\left(\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{P}^* \mathrm{d}V \right) \mathbf{a} = \left(\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{C} \mathbf{B} \mathrm{d}V \right) \bar{\mathbf{u}}^h. \tag{4.33}$$

The stress is taken as a vector of appropriate components, and for equilibrium we shall attempt to ensure that the smoothed stress $\boldsymbol{\sigma}_p^*$ satisfies in the least square sense the same patch equilibrium conditions as the finite element solution. For instance, in plane elasticity the stress approximation is taken as

$$\boldsymbol{\sigma}_p^* = \mathbf{P}^* \mathbf{a} = \begin{bmatrix} \tilde{\mathbf{P}}(\mathbf{x}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{P}}(\mathbf{x}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \tilde{\mathbf{P}}(\mathbf{x}) \end{bmatrix} \begin{Bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{Bmatrix}. \tag{4.34}$$

Obvious modifications are made for more or less components. It has been found in practice that the constraints provided in this way are not sufficient to always produce non-singular least square minimization. Accordingly, the equilibrium constraints are split into an alternative form in which each component of stress $\boldsymbol{\sigma}_{pi}^*$ is subjected to equilibrium requirements. This may be achieved by expressing the stress as

$$\boldsymbol{\sigma}^* = \sum_i \mathbf{1}_i \boldsymbol{\sigma}_{pi}^*, \tag{4.35}$$

where $\mathbf{1}_i$ represents a zero vector having unit value at the position of i-th stress (i.e. $\mathbf{1}_2 = [0,1,0]^{\mathrm{T}}$ for the second stress component), and imposing the set of constraints

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}_{pi}^* \mathrm{d}V = \int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}_{pi}^h \mathrm{d}V. \tag{4.36}$$

The imposition of the above equation allows each set of coefficients $\mathbf{a}_i$ to be solved independently. The least square minimization is based on the following functional

$$F = \left(\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}_{pi}^* \mathrm{d}V - \int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}_i^h \mathrm{d}V \right)^{\mathrm{T}} \left(\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}_{pi}^* \mathrm{d}V - \int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}_i^h \mathrm{d}V \right), \tag{4.37}$$

providing the number of parameter $\mathbf{a}_i$ is smaller (or equal) than the number of equations implied. The above functional can be rewritten as

$$F(\mathbf{a}_i) = (\mathbf{H}_i \mathbf{a}_i - \mathbf{g}_i)^{\mathrm{T}} (\mathbf{H}_i \mathbf{a}_i - \mathbf{g}_i), \tag{4.38}$$

where

$$\mathbf{H}_i = \int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{1}_i \tilde{\mathbf{P}} \mathrm{d}V, \tag{4.39}$$

and

$$\mathbf{g}_i = \int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}_i^h \mathrm{d}V = \left(\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{1}_i \otimes \mathbf{1}_i^{\mathrm{T}} \mathbf{C}\mathbf{B} \mathrm{d}V \right) \bar{\mathbf{u}}^h. \tag{4.40}$$

The minimization condition results in

$$\mathbf{a}_i = \left[\mathbf{H}_i^{\mathrm{T}} \mathbf{H}_i \right]^{-1} \mathbf{H}_i^{\mathrm{T}} \mathbf{g}_i. \tag{4.41}$$

Sometimes the application of this minimization to the patch alone is not sufficient to provide non-singular solution due to a large number of parameters $\mathbf{a}$. Generally, this may be eliminated by modifying the patch requirement to the minimization of

$$F^*(\mathbf{a}_i) = (\mathbf{H}_i \mathbf{a}_i - \mathbf{g}_i)^{\mathrm{T}} (\mathbf{H}_i \mathbf{a}_i - \mathbf{g}_i) + \sum_e^{nep} \alpha (\mathbf{H}_i^e \mathbf{a}_i - \mathbf{g}_i^e)^{\mathrm{T}} (\mathbf{H}_i^e \mathbf{a}_i - \mathbf{g}_i^e), \tag{4.42}$$

where the added terms represent modification on individual elements and α is a parameter.

$$\mathbf{H}_i^e = \int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \mathbf{1}_i \tilde{\mathbf{P}} \mathrm{d}V \mathbf{a}_i, \tag{4.43}$$

$$\mathbf{g}_i^e = \int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma}_i^h \mathrm{d}V. \tag{4.44}$$

Minimization now gives

$$\mathbf{a}_i = \left[\mathbf{H}_i^{\mathrm{T}} \mathbf{H}_i + \alpha \sum_e \mathbf{H}_i^{e\mathrm{T}} \mathbf{H}_i^e \right]^{-1} \left[\mathbf{H}_i^{\mathrm{T}} \mathbf{g}_i + \alpha \sum_e \mathbf{H}_i^{e\mathrm{T}} \mathbf{g}_i^e \right]. \tag{4.45}$$

The REP procedure gives overall an approximation which does not require knowledge of any superconvergent points. The accuracy of REP and SPR is comparable.

4.2.2.2 Some remark on superconvergence

Numerical tests show that the REP procedure exhibits superconvergence. The reason lies on the formulation that is based on using nodal forces. Referring to Equation (4.19) it can be seen that these forces are directly related to integrals of the displacement shape functions. Therefore, the rate of convergence of these forces is the same as that of displacements. This means that the rate of convergence of stresses in Equation (4.23) will, at least, be that of displacements if proper smoothing function is used.

Alternatively, we can prove this by using the principle of minimum potential energy. We must show that the answer, after smoothing, will approach the exact solution with higher order of error than that of the standard finite element solution. For this we shall use the equivalence of minimization of error of potential energy and minimization of total potential energy. The equivalence has been proved by Herrmann in [Herrmann, 1972]. Starting from Equation (4.20) we can consider

this equation as an answer of minimization of total potential energy over a patch provided that the nodal forces of patch are known:

$$\Pi_p = \int_{\Omega_p} \left(\mathbf{B}\bar{\mathbf{u}}^h\right)^{\mathrm{T}} \mathbf{C} \left(\mathbf{B}\bar{\mathbf{u}}^h\right) \mathrm{d}V - \bar{\mathbf{u}}^{h\mathrm{T}} \mathbf{F}_{\Omega-\Omega_p}. \tag{4.46}$$

In fact, this energy is a part of total energy of system. If minimum supports for this sub domain are provided the answers for stresses will be the same as those from solution of the whole domain. But according to Herrmann's theorem this is equivalent to minimization of error of energy with respect to $\bar{\mathbf{u}}$. Thus, we minimize

$$\Delta E = \int_{\Omega_p} \left[\mathbf{D}\left(\mathbf{u} - \mathbf{U}\bar{\mathbf{u}}^h\right)\right]^{\mathrm{T}} \mathbf{C} \left[\mathbf{D}\left(\mathbf{u} - \mathbf{U}\bar{\mathbf{u}}^h\right)\right] \mathrm{d}V, \tag{4.47}$$

where $\mathbf{u}$ is the exact displacement field. Now we have

$$\frac{\partial \Delta E}{\partial \bar{\mathbf{u}}^h} = \mathbf{0},$$

which gives

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{D}\mathbf{u}\mathrm{d}V = \int_{\Omega_p} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{B}\bar{\mathbf{u}}^h \mathrm{d}V. \tag{4.48}$$

From Equation (4.25) we recall that the new approach is based on

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{D}\mathbf{P}^*\mathbf{a}\mathrm{d}V = \int_{\Omega_p} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{B}\bar{\mathbf{u}}^h \mathrm{d}V. \tag{4.49}$$

Comparing (4.48) with (4.49),

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{D}\mathbf{u}\mathrm{d}V = \int_{\Omega_p} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{D}\mathbf{P}^*\mathbf{a}\mathrm{d}V, \tag{4.50}$$

or

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}}\mathbf{C}\left(\mathbf{D}\mathbf{u} - \mathbf{D}\mathbf{P}^*\mathbf{a}\right)\mathrm{d}V = \mathbf{0}. \tag{4.51}$$

The matrix $\mathbf{B}^{\mathrm{T}}\mathbf{C}$ can be considered as a discontinuous weight function over the domain. So Equation (4.51) implies that

$$\mathbf{D}\mathbf{u} \simeq \mathbf{D}\mathbf{P}^*\mathbf{a}.$$

Multiplying both sides of above equation by $\left(\mathbf{D}\mathbf{P}^*\right)^{\mathrm{T}}\mathbf{D}$ and taking integral over the patch gives

$$\int_{\Omega_p} \left(\mathbf{D}\mathbf{P}^*\right)^{\mathrm{T}}\mathbf{C}\mathbf{D}\bar{\mathbf{u}}^h \mathrm{d}V \simeq \int_{\Omega_p} \left(\mathbf{D}\mathbf{P}^*\right)^{\mathrm{T}}\mathbf{C}\mathbf{D}\mathbf{P}^*\mathbf{a}\mathrm{d}V. \tag{4.52}$$

Now we can compare the above result with that of minimization of ΔE with respect to $\mathbf{a}$:

$$\Delta E = \int_{\Omega_p} \left[\mathbf{D}\left(\mathbf{u} - \mathbf{P}^*\mathbf{a}\right)\right]^{\mathrm{T}} \mathbf{C} \left[\mathbf{D}\left(\mathbf{u} - \mathbf{P}^*\mathbf{a}\right)\right] \mathrm{d}V, \tag{4.53}$$

$$\frac{\partial \Delta E}{\partial \mathbf{a}} = \mathbf{0}, \tag{4.54}$$

which leads to

$$\int_{\Omega_p} (\mathbf{DP}^*)^{\mathrm{T}} \mathbf{CD}\bar{\mathbf{u}}^h \mathrm{d}V = \int_{\Omega_p} (\mathbf{DP}^*)^{\mathrm{T}} \mathbf{CDP}^* \mathbf{a} \mathrm{d}V. \tag{4.55}$$

From equations (4.52) and (4.55) it can be stated that REP gives the answers of minimization of least square of the stress errors in energy norm. In fact, we have changed the finite element bases from shape functions with discontinuous derivatives to a continuous polynomials with higher order.

Equality (4.52) is dependent on the form of weighting function $\mathbf{B}^{\mathrm{T}}\mathbf{C}$ in Equation (4.51). So the form of elements used in the mesh can affect the super-convergence and therefore it can be expected that for different kinds of meshes the rate of convergence will differ.

It is clear that the order of polynomial used in finite element procedure determines the order of convergence of answers. So obviously if we are able to change the order of finite element bases, we can increase the rate of convergence by using a higher order polynomial.

4.3 Recovery by Compatibility in Patches: RCP

In this section, a new stress recovery procedure is presented [Ubertini, 2002]–[Ubertini, 2004]–[Benedetti *et al.* , 2006]. The method proceeds in a way analogous to that used in both the SPR procedure and the REP procedure. The desired improved approximation for stresses over the whole domain, say $\boldsymbol{\sigma}^*$, is constructed by, first, recovering nodal values $\bar{\boldsymbol{\sigma}}^*$ and, then, interpolating these values by standard shape functions (see Equation (2.24)). Thus, the key step is to determine the nodal values $\bar{\boldsymbol{\sigma}}^*$. To accomplish this task, the problem domain is subdivided into local patches of elements. Each assembly node of the finite element discretization is associated to a patch. Such a patch is defined as a union of elements surrounding the particular assembly node considered. An illustration of typical patches associated to corner nodes internal to the domain is shown in Figure 4.5(a). Some additional details about patches associated to nodes in the interior or on the edge of an element as well as to nodes lying on the domain boundary are given in Section 4.3.4.

4.3.1 Formulation of the RCP procedure

Let Ω_p be the domain of the typical patch, $\partial\Omega_p$ the boundary of the patch and $\boldsymbol{\sigma}_p^*$ the local stress field in Ω_p which is to be determined aiming at improving the finite element solution. The current patch is considered as a separate system on which displacements are prescribed along the boundary $\partial\Omega_p$, as shown in Figure 4.5(b).

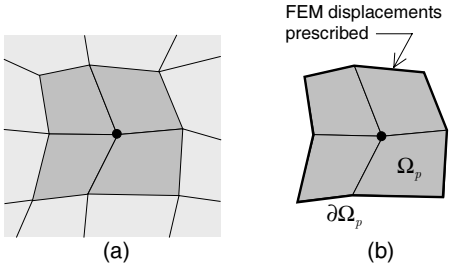


Figure 4.5: Example of pacth: (•) assembly node defining the patch

The prescribed displacements are those resulting from the finite element analysis. Then, the local stress field $\boldsymbol{\sigma}_p^*$ is determined by minimizing the convex functional of complementary energy associated to this separate system:

$$\Pi(\boldsymbol{\sigma}_p^*) = \frac{1}{2}\int_{\Omega_p} \boldsymbol{\sigma}_p^{*\,\mathrm{T}} \mathbf{C}^{-1} \boldsymbol{\sigma}_p^* \,\mathrm{d}V - \int_{\partial\Omega_p} \mathbf{u}^{h\,\mathrm{T}} \mathbf{N}^{\mathrm{T}} \boldsymbol{\sigma}_p^* \,\mathrm{d}S, \tag{4.56}$$

over a set of stress fields fulfilling equilibrium equations within the patch:

$$\mathbf{D}^* \boldsymbol{\sigma}_p^* = \mathbf{b} \quad \text{in } \Omega_p. \tag{4.57}$$

Since both $\mathbf{u}^h$ and $\boldsymbol{\sigma}_p^*$ are continuous over the patch, the above functional can be re-expressed in the equivalent form:

$$\Pi(\boldsymbol{\sigma}_p^*) = \sum_{j=1}^{nep} \left(\frac{1}{2}\int_{\Omega_j} \boldsymbol{\sigma}_p^{*\,\mathrm{T}} \mathbf{C}^{-1} \boldsymbol{\sigma}_p^* \,\mathrm{d}V - \int_{\partial\Omega_j} \mathbf{u}^{h\,\mathrm{T}} \mathbf{N}^{\mathrm{T}} \boldsymbol{\sigma}_p^* \,\mathrm{d}S \right), \tag{4.58}$$

where *nep* is the number of elements in the patch and Ω_j and $\partial\Omega_j$ are the domain and the boundary of the generic element of the patch, respectively. Notice that Equation (4.56) is the complementary energy of the separate patch system, which does not coincide in general with the exact complementary energy of the corresponding subdomain embedded into the original system, since $\mathbf{u}^h$ does not coincide in general with the exact displacement $\mathbf{u}$.

This simple procedure is based on the idea of improving stresses by enhancing equilibrium while relaxing compatibility. In fact, using the Gauss-Green formula (see Equation (2.16)) together with the following

$$\mathbf{u}^h = \mathbf{U}(\mathbf{x})\,\bar{\mathbf{u}}^h, \tag{4.59}$$

and with the Equation (4.57), functional (4.58) can be written as

$$\Pi(\boldsymbol{\sigma}_p^*) = \frac{1}{2}\int_{\Omega_p} \boldsymbol{\sigma}_p^{*\,\mathrm{T}} \mathbf{C}^{-1} \boldsymbol{\sigma}_p^* \,\mathrm{d}V - \int_{\Omega_p} \boldsymbol{\sigma}_p^{*\,\mathrm{T}} \mathbf{B}\bar{\mathbf{u}}^h \mathrm{d}V + \int_{\Omega_p} \mathbf{b}^{\mathrm{T}} \mathbf{U}\bar{\mathbf{u}}^h \mathrm{d}V. \tag{4.60}$$

Denoting by $\boldsymbol{\epsilon}^h$ the strains resulting from the finite element solution and by $\boldsymbol{\epsilon}_p^*$ the strains associated with the recovered stresses:

$$\boldsymbol{\epsilon}^h = \mathbf{B}\bar{\mathbf{u}}^h, \qquad \boldsymbol{\epsilon}_p^* = \mathbf{C}^{-1}\boldsymbol{\sigma}_p^*, \tag{4.61}$$

the first variation of the above functional takes the form

$$\delta\Pi = \int_{\Omega_p} \delta\boldsymbol{\sigma}_p^{*\,\mathrm{T}} \left(\boldsymbol{\epsilon}_p^* - \boldsymbol{\epsilon}^h\right) \mathrm{d}V. \tag{4.62}$$

This expression shows that the present procedure leads to local compatibility conditions. In particular, it enforces strains coming from equilibrated stresses to be equal to compatible finite element strains, by projecting the error into the new stress space over each patch. Hence, the procedure is a Recovery of stresses by Compatibility in Patches (RCP).

It is interesting to notice that the present procedure can be viewed as dual, in a certain sense, to the already mentioned procedure called Recovery by Equilibrium in Patches (REP). Indeed, dual should be intended only in the sense that the condition enforced by REP can be obtained from Equation (4.62) by exchanging the roles played by $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}$. This means to project the error between new and original stresses into the finite element strain space over the patch.

4.3.2 Implementation of the RCP procedure

In this section, a simple and low-cost implementation of the RCP procedure is presented. The first step is to select an approximation for local stresses over the patch. In order to obtain stresses which satisfy interior equilibrium, the following additive decomposition is introduced

$$\boldsymbol{\sigma}_p^* = \boldsymbol{\sigma}_{ph}^* + \boldsymbol{\sigma}_{pp}^*, \tag{4.63}$$

where $\boldsymbol{\sigma}_{ph}^*$ is the homogeneous solution of Equation (4.57) and $\boldsymbol{\sigma}_{pp}^*$ is a particular solution of the same equation. More details about $\boldsymbol{\sigma}_{ph}^*$ and $\boldsymbol{\sigma}_{pp}^*$ are given in Section 4.3.4. Notice that only $\boldsymbol{\sigma}_{ph}^*$ is indeterminate at this stage. Now, an approximation for $\boldsymbol{\sigma}_{ph}^*$ over the patch is assumed:

$$\boldsymbol{\sigma}_{ph}^* = \mathbf{P}^* \, \mathbf{a}, \tag{4.64}$$

where $\mathbf{P}^*$ is the matrix of basis functions and $\mathbf{a}$ is the vector of unknown parameters. For the above assumptions, matrix $\mathbf{P}^*$ should be selected such that

$$\mathbf{D}^* \mathbf{P}^* = \mathbf{0} \quad \text{in } \Omega_p, \tag{4.65}$$

i.e., $\mathbf{P}^*$ is a matrix of self-equilibrated stress modes.

The minimization condition results in the following system of linear algebraic equations

$$\mathbf{H}\,\mathbf{a} = \mathbf{g}, \tag{4.66}$$

with matrix $\mathbf{H}$ and vector $\mathbf{g}$ given by

$$\mathbf{H} = \sum_{j=1}^{nep} \mathbf{H}_j, \qquad \mathbf{g} = \sum_{j=1}^{nep} \left(\mathbf{g}_j^u - \mathbf{g}_j^p\right), \tag{4.67}$$

where

$$\mathbf{H}_j = \int_{\Omega_j} \mathbf{P}^{*\mathrm{T}} \mathbf{C}^{-1} \mathbf{P}^* \, \mathrm{d}V, \tag{68}$$

$$\mathbf{g}_j^u = \int_{\Omega_j} \mathbf{P}^{*\mathrm{T}} \mathbf{B} \bar{\mathbf{u}}^h \mathrm{d}V = \left(\int_{\Omega_j} \mathbf{P}^{*\mathrm{T}} \mathbf{B} \, \mathrm{d}V \right) \bar{\mathbf{u}}^h, \tag{69}$$

$$\mathbf{g}_j^p = \int_{\Omega_j} \mathbf{P}^{*\mathrm{T}} \mathbf{C}^{-1} \boldsymbol{\sigma}_{pp}^* \mathrm{d}V. \tag{70}$$

The outlined procedure can be easily implemented in existing finite element codes and possesses some features which make it simpler and cheaper than other patch-based recovery procedures. In fact, the number of unknown parameters $\mathbf{a}$ is equal to the number of equations of system (4.66) and the coefficient matrix $\mathbf{H}$ is symmetric and positive definite. Thus, system (4.66) can be solved directly using standard procedures. Notice that this is not the case of other recovery techniques, such as SPR and REP, which need to resort to least-squares scheme. Unlike the SPR procedure, RCP does not require any knowledge of super-convergent points. Moreover, matrix $\mathbf{H}$ and vector $\mathbf{g}$ are simply obtained by summing contributions from each element forming the patch, as shown by Equation (4.67). Finally, numerical stability is always guaranteed, so that the RCP procedure can be applied also to small patches (such as boundary ones) without any special modification to add extra constraints.

4.3.3 Consistency of the RCP procedure

A basic requirement for a recovery procedure, as well as for any numerical procedure, is that the exact solution should be reproduced under favorable circumstances. In the present context, such a condition can be formalized by requiring that if the finite element solution is exact, then also the local recovered stresses should be exact [Ainsworth & Oden, 2000]. It can be immediately realized that the RCP procedure satisfies this consistency condition if the local stress approximation is properly selected. In fact, let $\mathbf{u}^h$ be a polynomial of degree p which coincides with the exact solution $\mathbf{u}$. Then, Equation (4.56) actually represents the exact complementary energy for the subdomain corresponding to the current patch and the exact stresses are always recovered if the self-equilibrated field $\boldsymbol{\sigma}_{ph}^*$ belongs to the space of p-order polynomials which satisfy condition (4.65). Except for internally isostatic models such as rods or beams, this means that each component of $\boldsymbol{\sigma}_{ph}^*$ should be able to reproduce a complete polynomial of degree p. Thus, consistency is met if self-equilibrated stress approximations are selected among polynomials of degree not less than p, when finite element of order p are used.

4.3.4 Some remarks

In this section, some issues about the proposed procedure are addressed.

The first issue is the form of the typical patch. In principle, the number of elements forming the patch may be large at will. Also the whole domain could be considered as a patch. However, using large patches increases the computational cost and, as a consequence, the resultant procedure could become no more convenient in practice. Here, patches of conventional minimal form are used, as shown in Figure 4.5. In other words, each patch is formed using only one layer of elements surrounding the patch assembly node, so to minimize the computational cost.

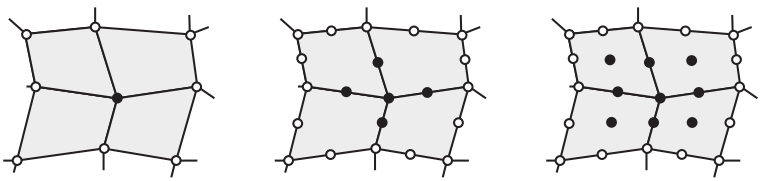


Figure 4.6: Computation of nodal values for four-, eight- and nine-node elements: (○) nodal points, (●) nodal values determined by recovery procedure

The second issue to be discussed is how to construct local stress approximations. Indeed, this can be easily done by employing the well-established techniques usually adopted to construct stress approximations for the original version of hybrid stress elements [Pian, 1973]. In particular, polynomial basis functions for self-equilibrated stresses $\boldsymbol{\sigma}^*_{ph}$ can be derived from assumed stress functions. It should be emphasized that these approximations should be derived having care to obtain complete polynomial expansions for each component of $\boldsymbol{\sigma}^*_{ph}$, so to retain invariance of the procedure with respect to any coordinate change and, in case, consistency, as stated in the previous section. A scheme to construct such approximations for some structural models is illustrated in Appendix A. Notice that the present procedure poses no upper limit to the order of polynomial approximation or to the number of unknown parameters $\mathbf{a}$ with any form of the patch. This does not occur with the SPR and REP procedures because of the limited number of independent equations they involve. Thus, for finite elements of order p any approximation of degree not less than p could be used in principle, provided it satisfies Equation (4.65). The numerical results obtained on some bi-dimensional tests using approximations of degree p and $p+1$ over patches of conventional form are compared and discussed in Section 4.6. Further increasing the order of approx-

imation seems to be not recommendable since the computational cost increases without any appreciable improvement in accuracy.

Another important issue regarding implementation of patch-based recovery procedures is how to compute smoothed stresses at nodes which are not located at the element vertices. In fact, when higher-order elements are used, improved stresses should be computed not only for corner nodes but also for edge nodes and/or nodes interior to the elements. Although the corresponding patches could be defined in different ways, the rules recommended for the SPR and REP procedures are adopted here. Only patches associated to corner nodes are formed and each patch is used to recover stresses at all the nodes inside the patch itself, as shown in Figure 4.6. By doing so, it is clear that edge and interior nodes generally belong to several patches and different recovered values are available from different overlapping patches. In order to obtain a single value for each node, a simple averaging of these values is taken. It should be remarked that the reason to adopt the above strategy for the present recovery procedure is not dictated by stability requirements. It has been adopted because provides good accuracy and allows computational savings.

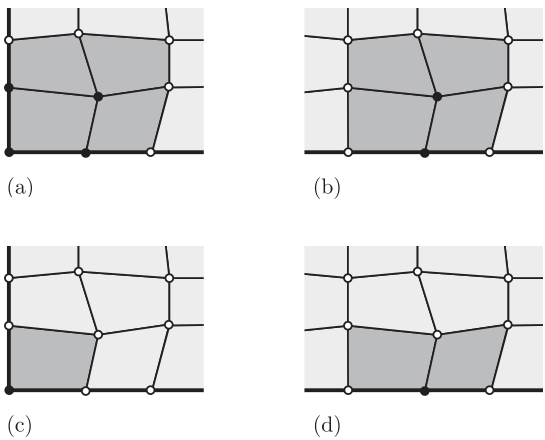


Figure 4.7: Boundary nodal recovery: (a) and (b) interior patches, (c) and (d) boundary patches; (○) nodal points, (●) nodal values determined by recovery procedure

Another particular situation arises for nodes lying on the domain boundaries. Also in this case, there are many ways to define patches for boundary nodes and compute the corresponding stress values. Here, two different techniques are

adopted. The first technique consists of forming and processing boundary patches in the same manner of interior patches. In particular, the smoothed stress field $\boldsymbol{\sigma}_p^*$ is taken as for interior patches. Notice that neither reducing the order of stress approximation nor adding any extra condition are needed for stability purpose. An illustration of boundary patches is reported in Figure 4.7. The second technique is the one recommended for the SPR procedure [Zienkiewicz & Zhu, 1995] and consists of evaluating stresses at boundary nodes by interior patches instead of boundary ones, as shown in Figure 4.7(a)-(b).

Finally, it should be added that locally based coordinate origins are used to avoid ill-conditioning.

4.3.5 An alternative version of RCP recovery

An alternative version of RCP recovery can be obtained by adopting a different strategy to form the patch, while retaining the same core. In the original form, the patch is defined as the union of the elements surrounding an assembly node of the finite element discretization (node patch, see Figure 4.7), while, in the new form, the patch is defined as the union of the elements surrounding an element (element patch, see Figure 4.8). In other words, in the original RCP each assembly node is associated to a patch, while in the new version each element is associated to a patch. Indeed, the above description of the node patch configuration is not exhaustive, since it refers only to internal vertex nodes. In fact, some additional details about node patches associated to nodes in the interior or on the edge of an element as well as to nodes lying on the domain boundary have been given in the previous Section, so leading to the need of special procedures in order to treat these cases. On the contrary, the above description completely defines the element patch configuration. For better clarity, Figure 4.9 shows an element patch associated to an element with edges on the domain boundary. As it can be noted, the same criterion as for internal elementsis used to form the patch.

Owing to the new patch configuration, parameters $\mathbf{a}$ can be directly used to recover stresses over the central element of the patch, without any additional interpolation and average process of values coming from adjacent patches.

The new form of the RCP procedure is simpler than the original one and easier to implement into existing finite element codes. In fact, no special data structures are needed to form the patch or to distinguish between vertex nodes, interior nodes and nodes lying on the domain boundary or other physical interfaces. Another advantage offered by the new form is that the larger support of the patch generally leads to higher accuracy without increasing the computational effort, which is substantially dictated by the assumed stress approximation over the patch. Indeed,

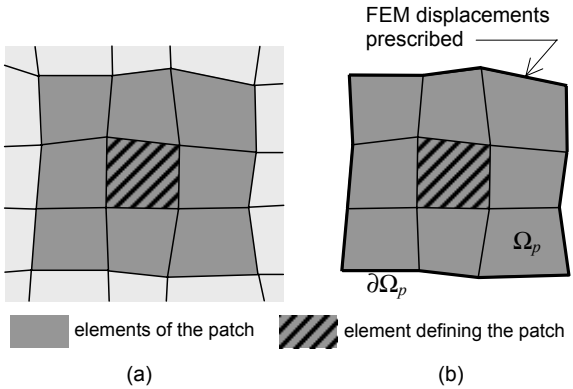


Figure 4.8: Example of element pacth: assembly element defining the patch

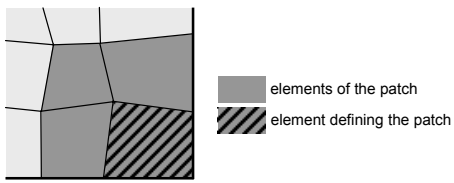


Figure 4.9: Boundary element recovery

unlike the original version, discontinuous solutions across element boundaries are obtained in this way, although discontinuities could be eliminated by an averaging process. However, in the spirit of the RCP recovery, the equilibrated nature of improved stresses within each element is retained. Finally, notice that the new version of the RCP recovery preserves the attractive feature of numerical stability.

4.4 Recovery on the boundary of the domain

It will be found that the accuracy of recovery at inner parts of the domain is greater than the accuracy on their boundary.

There are many ways to compute boundary stress values. One is computing these stresses from the smoothed stresses from interior patches without adding any extra condition for any imposed tractions.

In order to increase the accuracy at boundaries, we can enforce equilibrium with known traction boundary conditions, Figure 4.10(a) We can be sure that adding this form of constraint will not affect the convergence of the solution because the exact solution satisfies this conditions and we do nothing except to force the solution toward the exact solution. This constraint has been used in SPR lead-

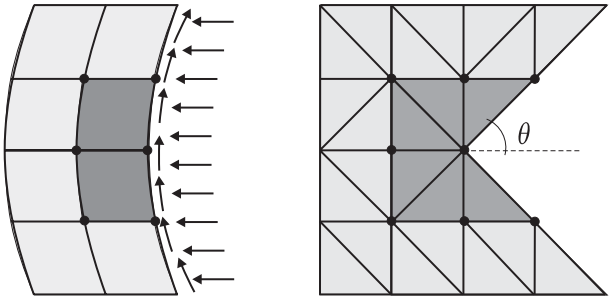


Figure 4.10: Typical boundary patches: (a) smooth boundary with traction; (b) geometrical singularity

ing to improvement of results elsewhere [Zienkiewicz & Zhu, 1995]–[Boroomand & Zienkiewicz, 1997b]. Considering the number of equations and unknowns, REP allows us to use such boundary patches with some extra traction conditions. For this, in the least squares procedure we add some terms which represent the difference between exact boundary condition and assumed one. Then we can write:

$$F(\mathbf{a}_i) = (\mathbf{H}_i\mathbf{a}_i - \mathbf{g}_i)^{\mathrm{T}}(\mathbf{H}_i\mathbf{a}_i - \mathbf{g}_i) + \sum_{i=1}^{NP}\left(\mathbf{N}^{\mathrm{T}}\boldsymbol{\sigma}_p^* - \bar{\mathbf{t}}\right)_i^{\mathrm{T}}\left(\mathbf{N}^{\mathrm{T}}\boldsymbol{\sigma}_p^* - \bar{\mathbf{t}}\right)_i . \qquad (4.71)$$

In the above equation, $\bar{\mathbf{t}}$ denotes the prescribed traction on the boundary, and NP is the number of integration points with prescribed tractions on the boundary of the patch. The experience shows that when a curved boundary exists and we are using low-order polynomial for smoothing, minimum number of integration points gives the best answer.

For geometrical singularities, like that shown in Figure 4.10(b), there will be singularities in strains and stresses. To overcome this difficulty, instead of using polynomial forms for σ_{11}, σ_{22} and τ_{12}, we could use polar expressions for smoothed values of σ_r, σ_θ and $\tau_{r\theta}$. A combination of polynomial for in radial direction (i.e. $1, r^{-1}, r^{-2}, \dots$) and periodic functions in rotational direction (i.e. 1, $cos(\theta)$, $sin(\theta), ...$) can be employed. This will help to have a discontinuous answer at $r = 0$. To evaluate $\mathbf{H}$, an appropriate rotational matrix must be used:

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}}\mathbf{T}\mathbf{P}^* \mathrm{d}V.$$

The above equation may need a larger number of integration points than conventionally used. Further it may be worthwhile to use large patches and add more unknown around singular points. To use higher order functions around these points, one can find all interior nodal stresses as a first step and use them, as the sec-

ond step, for the nodes belonging to the boundary patch as additional constraints. This can also be done for other boundary nodes when the user wishes to increase the number of polynomial terms.

4.5 The quality of an error estimator

The recovery procedures can be used to estimate the error resulting from a finite element analysis by replacing the exact solution with the recovered solution, as indicated by Equation (4.2). So, the estimated local error in the stresses is given by

$$\mathbf{e}^* = \boldsymbol{\sigma}^h - \boldsymbol{\sigma}^*, \tag{4.72}$$

where the symbol $\cdot^*$ is usual to distinguish the estimated error from the exact error (see Section 2.3 for more details on the error definitions). As for the exact error, if single-valued nodal errors are needed, because of interelement discontinuity the average of nodal errors can be computed in the same manner, element or global errors can be estimated. Adopting for instance the energy norm, the estimated element error is given by

$$\|\mathbf{e}^*\|_e = \left\|\boldsymbol{\sigma}^h - \boldsymbol{\sigma}^*\right\|_e = \left[\frac{1}{2}\int_{\Omega_e} \left(\boldsymbol{\sigma}^h - \boldsymbol{\sigma}^*\right) \mathbf{C}^{-1} \left(\boldsymbol{\sigma}^h - \boldsymbol{\sigma}^*\right) \mathrm{d}V\right]^{\frac{1}{2}}, \tag{4.73}$$

where Ω_e is the element domain under consideration. Finally, the estimated global error is

$$\|\mathbf{e}^*\| = \frac{\left(\sum_{e=1}^{n-elem} \|\mathbf{e}^*\|_e\right)^{\frac{1}{2}}}{E}, \tag{4.74}$$

where E is the exact energy norm of the problem

$$E = \left[\frac{1}{2}\int_{\Omega} \boldsymbol{\sigma}^{\mathrm{T}} \mathbf{C}^{-1} \boldsymbol{\sigma} \mathrm{d}V\right]^{\frac{1}{2}}. \tag{4.75}$$

Clearly the higher is the accuracy of the recovered solution, the higher is the quality of the error estimator.

To measure the accuracy of the recovered solution, the error definitions recalled above can be used again. A local measure is simply given by the pointwise error in the recovered stress

$$\mathbf{e}^{es} = \boldsymbol{\sigma}^* - \boldsymbol{\sigma}. \tag{4.76}$$

Notice that the superscript $\cdot^{es}$ is used to indicate this type of error, that is the exact error in the recovered solution, while symbol $\mathbf{e}$ is reserved for the exact error in the original finite element solution. In the case of interelement discontinuity the average of nodal errors can be considered also in this case.

The (exact) element and global errors in the recovered solution, assuming the

energy norm, read as

$$\|\mathbf{e}^{es}\|_e = \|\boldsymbol{\sigma}^* - \boldsymbol{\sigma}\|_e = \left[\frac{1}{2}\int_{\Omega_e} (\boldsymbol{\sigma}^* - \boldsymbol{\sigma})\, \mathbf{C}^{-1}\, (\boldsymbol{\sigma}^* - \boldsymbol{\sigma})\, \mathrm{d}V\right]^{\frac{1}{2}}, \tag{4.77}$$

$$\|\mathbf{e}^{es}\| = \frac{\left(\sum_{e=1}^{n-elem} \|\mathbf{e}^{es}\|_e^2\right)^{\frac{1}{2}}}{E}. \tag{4.78}$$

The effectiveness of the error-estimators is measured by the *effectivity index*, namely, the ratio between the estimated and the exact value of the error norm of interest:

$$\theta = \frac{\|\mathbf{e}^*\|}{\|\mathbf{e}\|}.$$

The error estimator is said to be asymptotically exact if θ approaches unity as the exact error $\|\mathbf{e}\|$ tends to zero (as $h \to 0$, or as $p \to \infty$). This means that the error estimator will always converge to the exact error while this decreases. Obviously, the reliability of $\|\mathbf{e}^*\|$ is dependent on the quality of $\boldsymbol{\sigma}^*$, and it can be proved (Zienkiewicz and Zhu, [Zienkiewicz & Zhu, 1992b]) that

$$1 - \frac{\|\mathbf{e}^{es}\|}{\|\mathbf{e}\|} \leq \theta \leq 1 + \frac{\|\mathbf{e}^{es}\|}{\|\mathbf{e}\|}. \tag{4.79}$$

In fact we can rewrite the error estimator $\|\mathbf{e}^*\|$ as

$$\|\mathbf{e}^*\| = \left\|\boldsymbol{\sigma}^h - \boldsymbol{\sigma}^*\right\| \equiv \left\|\left(\boldsymbol{\sigma}^h - \boldsymbol{\sigma}\right) - \left(\boldsymbol{\sigma}^* - \boldsymbol{\sigma}^h\right)\right\| = \|\mathbf{e} - \mathbf{e}^{es}\|, \tag{4.80}$$

and using the triangular inequality, we have

$$\|\mathbf{e}\| - \|\mathbf{e}^{es}\| \leq \|\mathbf{e}^*\| \leq \|\mathbf{e}\| + \|\mathbf{e}^{es}\|. \tag{4.81}$$

Dividing by $\|\mathbf{e}\|$, we obtain the (4.79).

Based on these result we have that the error estimator is asymptotically exact if

$$\frac{\|\mathbf{e}^{es}\|}{\|\mathbf{e}\|} \to 0 \qquad \text{as } \|\mathbf{e}\| \to 0. \tag{4.82}$$

This is achieved if $\|\mathbf{e}^{es}\|$ converges at a higher rate than $\|\mathbf{e}\|$. It follows that if $\|\mathbf{e}^{es}\|$ is superconvergent then asymptotic exactness of the error estimator is assured. The energy norm of the error of a finite element solution, with a p-order shape functions can be represented as follow

$$\|\mathbf{e}\| = O\left(h^p\right) \leq Ch^p, \tag{4.83}$$

where $p > 0$ is the rate of convergence and $C > 0$ is a constant independent of the element dimension h, but dependent on the patch (see Section 3.2). For the

recovered solution we can write

$$\|\mathbf{e}^{es}\| = O\left(h^{p+\alpha}\right), \tag{4.84}$$

with $\alpha > 0$, so the limits of the effectivity index, by substituting the equations (4.83) and (4.84) into (4.79), are :

$$1 - O(h^{\alpha}) \leq \theta \leq 1 + O(h^{\alpha}).$$

Naturally, one hopes that effectivity indices close to unity can be obtained, but global effectivity indices of $2.0 \div 3.0$ or even higher are often regarded as acceptable in many engineering applications.

4.5.1 The benchmarking approach

In engineering practice one often attempts to analyze the accuracy of the error estimators which are implemented in various finite element codes by checking the values of the effectivity index in a few example problems which are often called benchmarks. The benchmarking approach is a well accepted procedure for checking finite element codes [Babuška *et al.* , 1997a].

Can we classify the error estimator as reliable or unreliable based on these results? To answer to this question the following remarks should be considered [Babuška *et al.* , 1997b]:

1. The error in an element can be partitioned into the local error (the part of the error due to the residuals in the element and its neighbors) and the pollution error (the part of the error due to the residuals in the rest of the mesh and especially those in the neighborhood of the singular points and sharp fillets).
2. All the error indicators which employ only local computations (i.e. they use the input-data and the finite element solution in the element and its neighbors) cannot account for the pollution error.
3. The pollution error is the most significant component of the error in the majority of the elements for the types of meshes and domains employed in engineering computations. Note that when a globally adaptive mesh (which almost equidistributes the element error indicators) is employed, then the pollution error is practically controlled (see Reference [Zienkiewicz & Zhu, 1987]).

Therefore, it follow that

1. For uniform meshes (and singular problems), the global energy norm of the error practically coincides with the energy norm of the error over the patch of the elements which have a vertex at the singular point. Hence, for uniform meshes, the global effectivity index reflects only the accuracy of the error estimator for the elements adjacent to the singularity. The pollution error in

these elements is practically negligible and the element error indicators are close to the exact element norms of the error.

2. There is significant pollution practically everywhere, except in the elements with the biggest errors (which are the elements with a vertex at the singular point), and the element error indicators underestimate the element energy norms of the error in all the elements, except the elements with the biggest errors.

Hence, we conclude that if one is interested in the accuracy of the element error indicators in the elements adjacent to the singularity then a benchmark problem can be a proper choice. On the other hand, if one is interested in the accuracy of the element error indicators in the interior of the mesh or at the elements adjacent to a smooth boundary then benchmark tests could give no useful information unless we carefully avoid or properly control pollution.

Let us now summarize the conclusions which should be drawn from the above discussion (see also References [Babuška *et al.* , 1994b]-[Babuška *et al.* , 1994a]-[Babuška *et al.* , 1995b]).

1. The error in a patch of elements has a local and a global component and any error estimator based on local computations can only estimate the local component.
2. The element error indicators estimate only the local part of the error. The local error is practically the same as the error in the finite element approximation of the local $(p+1)$ Taylor series expansion of the exact solution.
3. The element effectivity indices depend on the local geometry of the mesh and should be reported for the worst $(p+1)$ degree Taylor series expansions corresponding to the particular mesh-geometry of interest.

Based on the above conclusions, Babuska and coworkers [Babuška *et al.* , 1994a]-[Babuška *et al.* , 1997a] prepared a numerical methodology to test the robustness of an error estimator in the asymptotic limit (when the mesh size tends to zero) for smooth problem (that is to check the capability of computing the local error). The methodology described in the next section provides an effective tool to evaluate and compare different procedures in a systematic and objective manner. It can be viewed as a patch test procedure which fixes minimal criteria to be met and avoids drawing erroneous conclusions based only on a limited number of test problems. However, it should be emphasized it is an asymptotic test and certain effects, such as those due to boundaries or singularities, have been isolated. In addiction no indications are given on the accuracy in the pre-asymptotic behaviour. Thus, this asymptotic testing should always be accompanied by some (carefully chosen)

benchmarking computation which give valuable information on the pre asymptotic behaviour.

4.5.2 The robustness test

The patch test idea was proposed by Bruce Irons in 1965 for to verify the quality of a model implemented in a finite element code, verifying, under regular condition, the capability of the element to reproduce, in a generic point, a constant state of strain when the dimension of the region surrounding that point tends to zero. So the elements surrounding this point must represent a constant strain state.

As shown before (see Section 3.2) the convergency is assured by the consistency of the finite element method: when the characteristic dimension of the elements h tends to zero, the approximated equations represent exactly the differential problem, at least in weak form. Based on the above conditions, the elements, with p degree polynomials, can reproduce exactly every problem with an exact solution based on that polynomial.

The *effectivity robustness test*, introduced by Babuška and coworkers, is based on the same idea of the traditional patch test, and is useful to compute the robustness of the error estimators. The method start from the fact that a smooth problem have a solution where the terms with order $(p+1)$ are predominant, so, as before, we can study a simple patch and its $(p+1)$ *exact* solution, that will give the asymptotic behaviour, its dimension tend to zero. We can also use this method to compute, over a specific patch, the maximum and minimum asymptotic value of θ, that reach the maximum and minimum values for the family problems with exact solution of order $(p+1)$.

These maximum and minimum bounds of the effectivity index are called θ_L and θ_U and for robustness in estimating of errors both values should be close to unity.

Babuška and coworkers defined a single robustness index as:

$$R = \max\left(|1-\theta_L| + |1-\theta_U|, \left|1-\frac{1}{\theta_L}\right| + \left|1-\frac{1}{\theta_U}\right|\right), \tag{4.85}$$

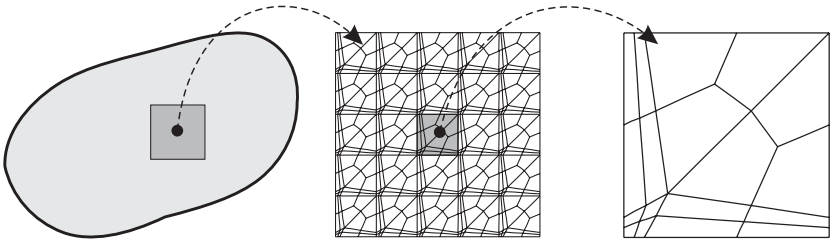
which obviously should be equal to zero in the ideal case when

$$\theta_L = \theta_U = 1, \tag{4.86}$$

and which measures how far both values depart from unity. The particular form of the index is chosen to give equal weight to the value of θ_L (or θ_U) below unity or above it, as clearly the coefficients must always be positive.

The test is based on the following assumptions:

Assumption 1. Since we are dealing with an asymptotic problem, it immediately follows that we should assume the element sizes to be near to



(a) problem domain (b) local (periodic) mesh (c) master-cell

Figure 4.11: Babuska patch test: (a) problem domain; (b) local (periodic) mesh; (c) master-cell

zero so the dimensions of any patch of elements, with limited number of elements, will be close to zero as well.

Assumption 2. We assume that the patch considered is far from singular points and that a smooth exact solution exists at the location of the patch.

Assumption 3. We assume that the mesh at the location of the patch has a repeatable pattern in all directions (we usually consider a single one of such repeatable patterns as a patch).

Summing up the test aims at assessing the asymptotic behaviour by considering that: (a) the mesh is locally uniform, (b) the exact solution is locally smooth and (c) the pollution error is negligible.

The idea is sketched in Figure 4.11 . Focusing the attention on the local periodic mesh (Figure 4.11(b)) and considering a finite element solution with p order shape function in all elements, we know thus that all reliable elements can reproduce an exact field of order p. Therefore the terms of order $p+1$ in the Taylor's series expansion of exact solution $\mathbf{u}$ have a predominant effect in the error (as $h \to 0$):

$$\|\mathbf{e}\| = \left\|\boldsymbol{\sigma}^h - \boldsymbol{\sigma}\right\| \simeq \left\|\boldsymbol{\sigma}^h - \boldsymbol{\sigma}_{asy}\right\|, \tag{4.87}$$

where

$$\boldsymbol{\sigma}_{asy} = \mathbf{C}\mathbf{D}\mathbf{u}^{p+1}, \tag{4.88}$$

is the stress corresponding to the $(p+1)$-order term $\mathbf{u}^{p+1}$ of the exact solution. Therefore for testing any recovery only such monomials need to be used as exact solutions:

$$\mathbf{u} \simeq \mathbf{u}^{p+1}. \tag{4.89}$$

The whole procedure of this test is based on estimating the asymptotic error and evaluating the effectivity indices, considering each one of these monomials as an exact solution over a local periodic mesh (Figure 4.11(b)). Having obtained these effectivity indices, the maximum and minimum bounds for the combination of the

monomials can be found through the solution of an eigenvalue problem.

The exact asymptotic error is given by

$$\|\mathbf{e}_{asy}\| = \left\|\boldsymbol{\sigma}^h_{asy} - \boldsymbol{\sigma}_{asy}\right\|, \tag{4.90}$$

where $\boldsymbol{\sigma}^h_{asy}$ is the asymptotic finite element solution on the periodic mesh. The estimated asymptotic error is

$$\left\|\mathbf{e}^*_{asy}\right\| = \left\|\boldsymbol{\sigma}^h_{asy} - \boldsymbol{\sigma}^*_{asy}\right\|, \tag{4.91}$$

where $\boldsymbol{\sigma}^*_{asy}$ is the asymptotic recovered solution.

The procedure consists of three steps.

1. Select the master cell (mesh pattern and aspect ratio), Figure 4.11(c), and take the exact solution as a general $(p+1)$ degree polynomial in the form
$$\mathbf{u} = \sum \alpha_i \mathbf{u}_i^{(p+1)}, \qquad \boldsymbol{\sigma} = \sum \alpha_i \boldsymbol{\sigma}_i, \tag{4.92}$$
where $\mathbf{u}_i^{(p+1)}$ are basis functions of degree $p+1$, α_i are unknown coefficients and
$$\boldsymbol{\sigma}_i = \mathbf{C}\mathbf{D}\mathbf{u}_i^{(p+1)}. \tag{4.93}$$
For example, for scalar-value problems and linear elements we can take
$$u = \alpha_1 x_1^2 + \alpha_2 x_1 x_2 + \alpha_3 x_2^2. \tag{4.94}$$
2. For each bases function $\mathbf{u}_i^{(p+1)}$, determine the (asymptotic) finite element solution $\boldsymbol{\sigma}^h_i$ and the (asymptotic) recovered solution $\boldsymbol{\sigma}^*_i$ over the periodic patch formed using the chosen master-cell.
3. Compute the extreme values of the (asymptotic) effectivity index, for all possible values of α_i, corresponding to the (asymptotic) solutions:
$$\boldsymbol{\sigma}^h_{asy} = \sum \alpha_i \boldsymbol{\sigma}^h_i, \qquad \boldsymbol{\sigma}^*_{asy} = \sum \alpha_i \boldsymbol{\sigma}^*_i. \tag{4.95}$$

To apply the above procedure we need to explain how to compute each finite element solution $\boldsymbol{\sigma}^h_i$. Indeed computing $\boldsymbol{\sigma}^*_i$ poses no particular difficulty as it can be done in the standard way using the recovery technique under investigation. In addition, we give some details on how to compute the extreme values of θ (step 3). The nature of this test is such that any recovery method which is truly superconvergent must produce the perfect answer of $\theta_L = \theta_U = 1$ or $R = 0$.

4.5.2.1 Computing asymptotic finite element solution

The relation between the finite element solution $\mathbf{u}^h_i$ and the exact solution $\mathbf{u}_i^{(p+1)}$ can always be written as

$$\int_\Omega \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{D}\left(\mathbf{u}^h_i - \mathbf{u}_i^{(p+1)}\right) \mathrm{d}V = 0. \tag{4.96}$$

Then, if $\bar{\mathbf{u}}^h_i$ are the nodal displacements of $\mathbf{u}^h_i$ and $\mathbf{u}^h_i = \mathbf{U}\bar{\mathbf{u}}^h_i$, the above relation

can be put in the form

$$\int_{\Omega} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{D}\mathbf{u}_i^h \mathrm{d}V = \int_{\Omega} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{D}\mathbf{U}\bar{\mathbf{u}}_i^h \mathrm{d}V, \tag{4.97}$$

By interpolating the exact $(p+1)$-degree solution $\mathbf{u}_i^{(p+1)}$ using the standard p-degree shape functions we have

$$\hat{\mathbf{u}}_i^{(p+1)} = \mathbf{U}\bar{\mathbf{u}}_i^{(p+1)} \tag{4.98}$$

where $\bar{\mathbf{u}}_i^{(p+1)}$ are the nodal values of $\mathbf{u}_i^{(p+1)}$. Then the exact solution can be rewritten as

$$\bar{\mathbf{u}}_i^{(p+1)} = \hat{\mathbf{u}}_i^{(p+1)} + \left(\bar{\mathbf{u}}_i^{(p+1)} - \hat{\mathbf{u}}_i^{(p+1)}\right) \tag{4.99}$$

and substituting in the Equation (4.96) yields

$$\int_{\Omega} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{D}\hat{\mathbf{e}}_i^h \mathrm{d}V = \int_{\Omega} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{D}\left(\mathbf{u}_i^{(p+1)} - \hat{\mathbf{u}}_i^{(p+1)}\right) \mathrm{d}V, \tag{4.100}$$

where $\hat{\mathbf{e}}_i$ is the difference between the finite element solution and the interpolant of the exact solution

$$\hat{\mathbf{e}}_i = \mathbf{u}_i^h - \hat{\mathbf{u}}_i^{(p+1)} = \mathbf{U}\left(\bar{\mathbf{u}}_i^h - \bar{\mathbf{u}}_i^{(p+1)}\right) = \tag{4.101}$$

$$= \mathbf{U}\bar{\mathbf{e}}_i. \tag{4.102}$$

Taking into account for this reason, Equation (4.97) can be rewritten as

$$\left(\int_{\Omega} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{B}\mathrm{d}V\right)\bar{\mathbf{e}}_i = \int_{\Omega} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{D}\left(\bar{\mathbf{u}}_i^{(p+1)} - \hat{\mathbf{u}}_i^{(p+1)}\right) \mathrm{d}V.$$

It can be shown that $\left(\mathbf{u}_i^{(p+1)} - \hat{\mathbf{u}}_i^{(p+1)}\right)$ is a periodic function with a period same as the patch length in all direction [Boroomand & Zienkiewicz, 1997a]. To prove this we consider the m-th monomial of the exact solution. With this important fact and with the knowledge of periodicity of the mesh we can write (4.97) for only the *master-cell* Ω_{mc}:

$$\left(\int_{\Omega_{mc}} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{B}\mathrm{d}V\right)\bar{\mathbf{e}}_i = \int_{\Omega_{mc}} \mathbf{B}^{\mathrm{T}}\mathbf{C}\mathbf{D}\left(\mathbf{u}_i^{(p+1)} - \hat{\mathbf{u}}_i^{(p+1)}\right) \mathrm{d}V. \tag{4.103}$$

This equation can be solved for $\bar{\mathbf{e}}_i$ once the appropriate boundary conditions are specified. In particular $\hat{\mathbf{e}}_i$ should be periodic over the master cell, therefore periodic boundary conditions on $\bar{\mathbf{e}}_i$ should be applied.

It can be easily verified that the above formulation has a unique solution a part for a constant which can be computed by subjecting the domain to a uniform "body force" and requiring that the absorbed energy from using the finite element method solution and the exact one should be the same. The resulting additional condition is

$$\int_{\Omega_{mc}} \bar{\mathbf{e}}_i^h \mathrm{d}V = \int_{\Omega_{mc}} \left(\mathbf{u}_i^{(p+1)} - \hat{\mathbf{u}}_i^{(p+1)}\right) \mathrm{d}V. \tag{4.104}$$

Now, the set boundary conditions is complete and Equation (4.103) can be solved

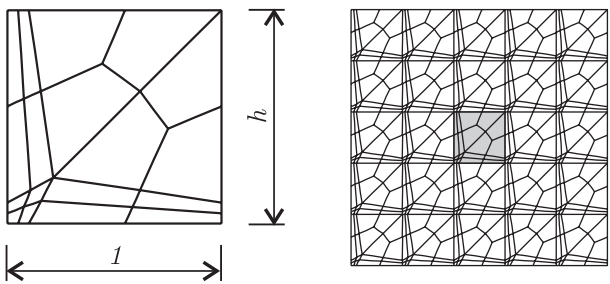
for $\bar{\mathbf{e}}_i$ and, subsequently, the finite element solution can be computed by

$$\mathbf{u}_i^h = \hat{\mathbf{u}}_i^{(p+1)} + \hat{\mathbf{e}}_i = \mathbf{U}\left(\bar{\mathbf{e}}_i + \bar{\mathbf{u}}_i^{(p+1)}\right). \tag{4.105}$$

4.5.2.2 Computing the effectivity index

For any procedure of error estimation we can independently determine the estimated error norm and, as the exact solution is known, also the exact error norm thus establishing the effectivity index θ corresponding to each bases function $\mathbf{u}_i^{(p+1)}$.

It would appear that a search amongst the values of indices so obtained would give the bounds of θ_L and θ_U or the possible effectivity. The determination of effectivity for each monomial separately does not, however, guarantee that absolute lowest and highest bounds have been determined.



(a) master-cell (b) periodic mesh

Figure 4.12: Computing the effectivity index: (a) master-cell; (b) periodic mesh

We give below a process by which all combination of the individual terms can be examined for the case of recovery based estimators. As a starting point we repeat the solution procedure for all $\mathbf{u}_i^{(p+1)}$ so the general finite element solution takes the form:

$$\bar{\mathbf{u}}^h = \sum_i \alpha_i \mathbf{u}_i^h = \sum_i \alpha_i \mathbf{U}\bar{\mathbf{u}}_i^h = \mathbf{U}\bar{\mathbf{U}}^h\boldsymbol{\alpha}, \tag{4.106}$$

where $\bar{\mathbf{U}}^h$ collects the nodal values of all the finite element solutions

$$\bar{\mathbf{U}}^h = \left[\begin{array}{c|c|c} \bar{\mathbf{u}}_1^h & \bar{\mathbf{u}}_2^h & \dots \end{array}\right], \tag{4.107}$$

and $\boldsymbol{\alpha}$ collects the coefficients α_i.

The general finite element solution in terms of stresses is given by

$$\boldsymbol{\sigma}^h = \mathbf{C}\mathbf{B}\bar{\mathbf{U}}^h\boldsymbol{\alpha}. \tag{4.108}$$

Analogously, the recovered solution in compact form reads as

$$\boldsymbol{\sigma}^* = \sum_i \alpha_i \boldsymbol{\sigma}_i^* = \mathbf{S}^* \boldsymbol{\alpha}, \tag{4.109}$$

where matrix $\mathbf{S}^*$ collects the solutions $\boldsymbol{\sigma}_i^*$ obtained by applying the recovery procedure under investigation to the finite element solution $\mathbf{u}_i^h$. Finally, also the exact solution can be put in a compact form as

$$\mathbf{u} = \sum_i \alpha_i \mathbf{u}_i^{(p+1)} = \mathbf{U}^{(p+1)} \boldsymbol{\alpha} \ ,$$

where matrix $\mathbf{U}^{(p+1)}$ collects the basis functions $\mathbf{u}_i^{(p+1)}$:

$$\mathbf{U}^{es} = \left[\begin{array}{c|c|c} & & \\ \mathbf{u}_1^{(p+1)} & \mathbf{u}_2^{(p+1)} & \dots \\ & & \end{array} \right]. \tag{4.110}$$

Consequently the exact solution in terms of stresses can be written as

$$\boldsymbol{\sigma} = \mathbf{C}\mathbf{B}\mathbf{U}^{(p+1)} \boldsymbol{\alpha}. \tag{4.111}$$

Using the above expressions the exact asymptotic error over the master-cell (in energy norm) results in

$$\begin{aligned} \|\mathbf{e}_{asy}\|^2 &= \frac{1}{2} \int_{\Omega_{mc}} \left(\mathbf{C}\mathbf{D}\mathbf{U}^{(p+1)} \boldsymbol{\alpha} - \mathbf{C}\mathbf{B}\bar{\mathbf{U}}^h \boldsymbol{\alpha} \right)^{\mathrm{T}} C^{-1} \left(\mathbf{C}\mathbf{D}\mathbf{U}^{(p+1)} \boldsymbol{\alpha} - \mathbf{C}\mathbf{B}\bar{\mathbf{U}}^h \boldsymbol{\alpha} \right) dV, \\ &= \boldsymbol{\alpha}^{\mathrm{T}} \left\{ \int_{\Omega_{mc}} \left(\mathbf{D}\mathbf{U}^{(p+1)} - \mathbf{B}\bar{\mathbf{U}}^h \right)^{\mathrm{T}} \mathbf{C}^{\mathrm{T}} \left(\mathbf{D}\mathbf{U}^{(p+1)} - \mathbf{B}\bar{\mathbf{U}}^h \right) \mathrm{d}V \right\} \boldsymbol{\alpha}. \end{aligned} \tag{4.112}$$

Posing

$$\mathbf{E} = \int_{\Omega_{mc}} \left(\mathbf{D}\mathbf{U}^{(p+1)} - \mathbf{B}\bar{\mathbf{U}}^h \right)^{\mathrm{T}} \mathbf{C}^{\mathrm{T}} \left(\mathbf{D}\mathbf{U}^{(p+1)} - \mathbf{B}\bar{\mathbf{U}}^h \right) \mathrm{d}V, \tag{4.113}$$

yields

$$\|\mathbf{e}_{asy}\|^2 = \boldsymbol{\alpha}^{\mathrm{T}} \mathbf{E} \boldsymbol{\alpha}. \tag{4.114}$$

On the other hand, the estimated asymptotic error over the master-cell can be expressed as

$$\begin{aligned} \left\|\mathbf{e}_{asy}^*\right\|^2 &= \frac{1}{2} \int_{\Omega_{mc}} \left(\mathbf{S}^* \boldsymbol{\alpha} - \mathbf{C}\mathbf{B}\bar{\mathbf{U}}^h \boldsymbol{\alpha} \right)^{\mathrm{T}} \mathbf{C}^{-1} \left(\mathbf{S}^* \boldsymbol{\alpha} - \mathbf{C}\mathbf{B}\bar{\mathbf{U}}^h \boldsymbol{\alpha} \right) \mathrm{d}V = \\ &= \boldsymbol{\alpha}^{\mathrm{T}} \left\{ \int_{\Omega_{mc}} \left(\mathbf{S}^* - \mathbf{C}\mathbf{B}\bar{\mathbf{U}}^h \right)^{\mathrm{T}} \mathbf{C}^{-1} \left(\mathbf{S}^* - \mathbf{C}\mathbf{B}\bar{\mathbf{U}}^h \right) \mathrm{d}V \right\} \boldsymbol{\alpha}. \end{aligned} \tag{4.115}$$

So, posing

$$\mathbf{E}^* = \int_{\Omega_{mc}} \left(\mathbf{S}^* - \mathbf{C}\mathbf{B}\bar{\mathbf{U}}^h \right)^{\mathrm{T}} \mathbf{C}^{-1} \left(\mathbf{S}^* - \mathbf{C}\mathbf{B}\bar{\mathbf{U}}^h \right) \mathrm{d}V, \tag{4.116}$$

we end with

$$\left\|\mathbf{e}_{asy}^*\right\|^2 = \boldsymbol{\alpha}^{\mathrm{T}} \mathbf{E}^* \boldsymbol{\alpha}. \tag{4.117}$$

Now we can write the (asymptotic) effectivity index as

$$\theta^2_{asy} = \frac{\left\|\mathbf{e}^*_{asy}\right\|^2}{\left\|\mathbf{e}_{asy}\right\|^2} = \frac{\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{E}^*\boldsymbol{\alpha}}{\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{E}\boldsymbol{\alpha}}. \tag{4.118}$$

The maximum and minimum values for θ^2_{asy} are the maximum and minimum eigenvalues of the following problem

$$\left(\mathbf{E}^* - \boldsymbol{\theta}^2\mathbf{E}\right)\mathbf{w} = \mathbf{0}, \tag{4.119}$$

where $\boldsymbol{\theta}^2$ is the diagonal matrix of the eigenvalues of the problem.

4.5.2.3 Self equilibrated solutions

In the test so far described we have used generic $(p+1)$ degree bases functions $\mathbf{u}_i^{(p+1)}$. Indeed it is convenient to use only self equilibrated, that is only these solutions of the homogeneous form of the governing equations of degree $(p+1)$. Such solutions can be found from the general form by enforcing a proper restrictions on coefficients $\boldsymbol{\alpha}$:

$$\boldsymbol{\alpha} = \mathbf{T}\tilde{\boldsymbol{\alpha}}. \tag{4.120}$$

Then Equations (4.114) and (4.117) can be written as

$$\left\|\mathbf{e}_{asy}\right\|^2 = \tilde{\boldsymbol{\alpha}}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{E}\mathbf{T}\tilde{\boldsymbol{\alpha}} = \tilde{\boldsymbol{\alpha}}^{\mathrm{T}}\tilde{\mathbf{E}}\tilde{\boldsymbol{\alpha}}, \tag{4.121}$$

and

$$\left\|\mathbf{e}^*_{asy}\right\|^2 = \tilde{\boldsymbol{\alpha}}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{E}^*\mathbf{T}\tilde{\boldsymbol{\alpha}} = \tilde{\boldsymbol{\alpha}}^{\mathrm{T}}\tilde{\mathbf{E}}^*\tilde{\boldsymbol{\alpha}}. \tag{4.122}$$

Therefore a modified procedure should be followed by finding the eigenvalues of

$$\left(\tilde{\mathbf{E}}^* - \tilde{\boldsymbol{\theta}}^2\tilde{\mathbf{E}}\right)\tilde{\mathbf{w}} = \mathbf{0}. \tag{4.123}$$

4.5.2.4 Summary of the procedure

In summary the test procedure is performed as described below:

1. select a mesh pattern for the master-cell;
2. select a proper basis function (monomial) one order higher than the order of shape-functions;
3. solve Equation (4.103) for $\bar{\mathbf{e}}_i$ with periodic boundary conditions and minimum support for the master-cell;
4. use Equation (4.104) to find rigid-body motions;
5. find the finite element solution through Equation (4.105);
6. perform the recovery procedure using the finite element solution from previous step. This step itself consists of three smaller steps:
   (a) produce a periodic mesh from the patch;

(b) apply the recovery procedure on the periodic mesh;

(c) consider the recovered answers only for one master-cell. (it should be noted that depending on the location of the patch the finite element solution may change but it can be shown that the errors will be periodic.)

7. repeat (2)-(6) for all possible basis functions (monomials);
8. evaluate $\mathbf{E}$ and $\mathbf{E}^*$ for exact and recovered errors. Special care should be taken here for rectangular elements since these elements can exactly reproduce some higher-order solutions, e.g. $x_1 x_2$ in bilinear elements, and the FEM solution, itself, will be exact. Therefore the corresponding terms in $\mathbf{E}$ and $\mathbf{E}^*$, which are close to zero, must be eliminated otherwise roundoff error will encounter the computation leading to some non-realistic answers.
9. find the eigenvalues of $\mathbf{E}^*$ (or $\mathbf{E}$) to determine the upper and lower bounds of the effectivity index and robustness index.

4.6 Numerical tests

In this section, the performance of the RCP recovery is validated numerically through a systematic testing based on the methodology proposed by Babuška et al. in [Babuška *et al.* , 1994a]-[Babuška *et al.* , 1997a], Section 4.6.1, and several benchmark tests proposed in the literature, Section 4.6.2. Comparisons are carried out with the improved version of the REP procedure [Boroomand & Zienkiewicz, 1997a], as quite natural considering the duality recognized in Section 4.3.1. For the sake of brevity, no further comparisons are included. However, the reader can refer to [Boroomand & Zienkiewicz, 1997b][Boroomand & Zienkiewicz, 1997a]-[Boroomand & Zienkiewicz, 1999] for a detailed comparison between the REP procedure and the widely used SPR procedure. As claimed by the authors of these papers, the two procedures yield very similar results.

All the numerical results presented in the following are obtained using patches of conventional minimal form. In the REP procedure, complete polynomial expansions of order p are selected for each stress component, being p the interpolation order for displacements in the finite element solution (i.e. the order of shape functions $\mathbf{U}$). In the RCP procedure, the influence of different choices for the order of stress approximation is investigated. It is worth to emphasize that the present procedure allows to enrich the approximation for smoothed stresses without any need to enlarge the form of the patch. As recommended in [Boroomand & Zienkiewicz, 1999], the REP procedure is applied without using boundary patches or any correction on boundaries, so that nodal boundary values are computed from interior

Table 4.1: Versions of RCP recovery tested and related orders of the stress approximation

Type	*order of the stress approximation*	*formation of the patch*	
RCP1n	p	node patch;	boundary patch are used;
RCP1nb	p	node patch;	boundary patch are not used;
RCP2n	$p+1$	node patch;	boundary patch are used;
RCP2nb	$p+1$	node patch;	boundary patch are not used;
RCP1e	p	element patch;	-
RCP2e	$p+1$	element patch;	-

patches. Since the present procedure allows to use boundary patches without introducing any additional condition or decreasing the order of approximation for smoothed stresses, this opportunity has been exploited. Thus, various versions of the RCP procedure have been implemented. Denoting by p theinterpolation order for displacements in the finite element solution, the various versions of the RCP procedure are summarized in Table 4.1.

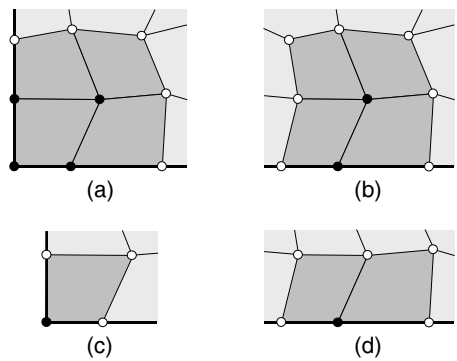


Figure 4.13: Boundary nodal recovery: (a) and (b) interior patches, (c) and (d) boundary patches; (○) nodal points, (●) nodal values determined by recovery procedure

4.6.1 Asymptotic behaviour and robustness

The analysis described in Section4.6.1 has been carried out for two different types of problems: equilibrium of thin membranes of uniform thickness or heat transfer (Class I problems, see Appendix A.1) and plane strain elasticity (Class II problems, see Appendix A.2). The Young's modulus and Poisson's ratio are taken as 1000 and 0.3, respectively. Various mesh patterns have been selected for the basic cell, using

both triangular and quadrangular elements, as illustrated by Figures 4.14 and 4.15. In all these tests the performance of the RCPn-based error estimator is compared with both the RCPe form and the REP-based estimator, in the improved version presented in [Boroomand & Zienkiewicz, 1997a]. The interested reader can refer to [Boroomand & Zienkiewicz, 1997a]-[Boroomand *et al.* , 2004] for a comparison with the SPR-based estimator. As suggested in [Boroomand & Zienkiewicz, 1997a], the REP procedure is used with complete polynomial expansions of degree p for each stress component. Stress approximations are selected according to Equation (A.3) for the thin membrane case and according to Equation (A.18) for the plane strain case.

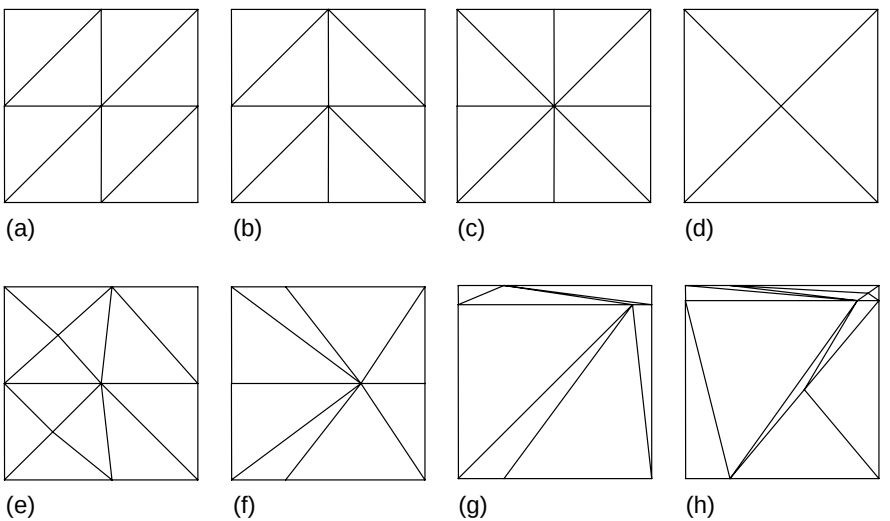


Figure 4.14: Mesh patterns used in robustness tests for triangular elements

4.6.1.1 Class I problems

Class I problems refer to thin membrane problems and are described in Appendix A.1. The results obtained for triangular and quadrangular elements in this case with the RCPn procedure are collected in Tables 4.2 and 4.3, respectively, while the corresponding results obtained with the RCPe procedure are collected in Tables 4.4 and 4.5. Moreover, Tables 4.6 and 4.7 collect some results obtained using regular mesh patterns of quadratic elements with various aspect ratios.

It is interesting to observe that the RCP procedure exhibits a super-convergent behavior for two simple patterns (Figures 4.14(a) and (c)) of linear triangular

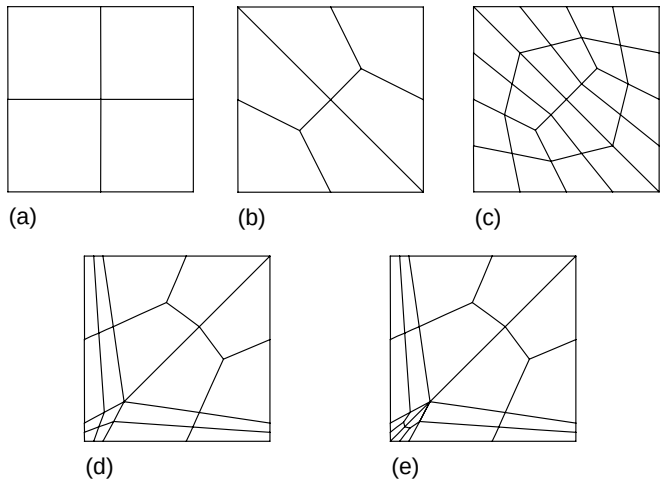


Figure 4.15: Mesh patterns used in robustness tests for quadrangular elements

elements. Super-convergence is observed also for different values of the aspect ratio of these mesh patterns. Indeed, super-convergence is lost by the RCPe on the *criss-cross* pattern (Figure 4.14(d)), where the RCPn and the REP are super-convergent, but the robustness index takes very small values: $R = 0.0243$ for RCP1e and $R = 0.0057$ for RCP2e (Table 4.2). On the other hand, full super-convergence is not achieved by none of the procedures considered for triangular elements of quadratic kind. Also in this case, however, the RCPe preserves small values of the robustness index, valuably smaller than those of the other procedures. In particular, $R < 0.07$ for RCP1e, $R < 0.27$ for RCP1n and $R < 0.33$ for REP (Table 4.4). This is confirmed also on regular meshes with different aspect ratios (see Table 4.7).

As regards quadrangular elements, the REP procedure is super-convergent for both four- and nine-node elements on regular meshes (Figure 4.15(a)), while the RCPn and the RCPe procedure only for four-node elements. However, although not exact, the RCP-results for nine-node elements are very good being $R < 0.08$ for RCPn (Table 4.6) and $R < 0.007$ for RCPe (Table 4.7). As for triangular elements, super-convergence is not observed for irregular patterns (Figures 4.15(b)-(e)), but excellent results are obtained.

In general, the robustness index obtained with the RCPe procedure takes, values smaller than those obtained with the RCPn version

The above results demonstrate the good quality of the RCPn and RCPe procedure. The RCP1 versions appear to be more robust than REP especially on irregular mesh patterns. No appreciable improvements of the asymptotic perfor-

Table 4.2: RCPn - Effectivity bounds and robustness indices for triangular elements applied to Class I problems (the mesh patterns are illustrated in Figure 4.14, the aspect ratio is 1/1).

	REP			RCP1n			RCP2n		
Mesh	θ_L	θ_U	R	θ_L	θ_U	R	θ_L	θ_U	R
	Three-node elements								
a	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
b	1.0000	1.0074	0.0074	0.9201	1.0039	0.0907	0.9298	1.0052	0.0807
c	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
d	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
e	0.9715	1.0190	0.0480	0.9636	1.0014	0.0392	0.9717	1.0020	0.0311
f	0.9828	1.0247	0.0420	0.9405	1.0035	0.0667	0.9479	1.0051	0.0600
g	0.9352	1.5400	0.6047	0.7121	1.0218	0.4256	0.7153	1.0316	0.4288
h	0.9939	1.8287	0.8347	0.5975	0.9806	0.6935	0.6311	0.9763	0.6089
	Six-node elements								
a	0.9516	1.0398	0.0891	0.9956	1.0114	0.0158	0.9956	1.0114	0.0158
b	0.9950	1.0222	0.0272	0.9839	1.0157	0.0318	0.9815	1.0070	0.0258
c	0.9320	0.9836	0.0897	0.9454	1.0136	0.0712	0.9454	1.0136	0.0712
d	0.9259	0.9594	0.1223	0.9454	1.0136	0.0712	0.9454	1.0136	0.0712
e	0.9481	1.0240	0.0782	0.9695	1.0095	0.0409	0.9582	1.0071	0.0507
f	0.9816	1.0045	0.0232	0.9929	1.0083	0.0154	0.9681	1.0058	0.0387
g	0.8972	1.1625	0.2653	0.8487	0.9950	0.1832	0.7767	1.0026	0.2900
h	0.8868	1.2142	0.3274	0.8033	0.9764	0.2691	0.7401	0.9613	0.3914

Table 4.3: RCPn - Effectivity bounds and robustness indices for quadrangular elements applied to Class I problems (the mesh patterns are illustrated in Figure 4.15, the aspect ratio is 1/1).

	REP			RCP1n			RCP2n		
Mesh	θ_L	θ_U	R	θ_L	θ_U	R	θ_L	θ_U	R
	Four-node elements								
a	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
b	0.9987	1.0174	0.0187	0.9987	1.0174	0.0187	0.9841	1.0008	0.0169
c	0.9979	1.0061	0.0082	0.9978	1.0037	0.0059	0.9889	1.0047	0.0159
d	0.9798	1.0680	0.0882	0.9394	1.0118	0.0762	0.9363	1.0000	0.0680
e	0.9833	1.0680	0.0847	0.9449	1.0077	0.0660	0.9334	1.0031	0.0745
	Nine-node elements								
a	1.0000	1.0000	0.0000	1.0107	1.0107	0.0214	1.0107	1.0107	0.0214
b	0.9816	1.0045	0.0232	0.9929	1.0083	0.0154	0.9681	1.0058	0.0387
c	0.9967	1.0131	0.0164	1.0016	1.0142	0.0158	0.9941	1.0074	0.0133
d	0.9911	1.0329	0.0418	0.9800	1.0018	0.0222	0.9467	0.9933	0.0630
e	0.9909	1.0334	0.0425	0.9730	1.0008	0.0285	0.9454	0.9934	0.0644

Table 4.4: RCPe - Effectivity bounds and robustness indices for triangular elements applied to Class I problems (the mesh patterns are illustrated in Figure 4.14, the aspect ratio is 1/1).

	REP			RCP1e			RCP2e		
Mesh	θ_L	θ_U	R	θ_L	θ_U	R	θ_L	θ_U	R
	Three-node elements								
a	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
b	1.0000	1.0074	0.0074	0.9705	1.0006	0.0310	0.9677	1.0008	0.0342
c	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
d	1.0000	1.0000	0.0000	0.9765	1.0002	0.0243	0.9946	1.0003	0.0057
e	0.9715	1.0190	0.0480	0.9812	1.0005	0.0196	0.9822	1.0006	0.0188
f	0.9828	1.0247	0.0420	0.9853	1.0023	0.0173	0.9888	1.0019	0.0133
g	0.9352	1.5400	0.6047	0.8504	1.0024	0.1783	0.8760	1.0011	0.1427
h	0.9939	1.8287	0.8347	0.7493	0.9942	0.3404	0.7701	0.9848	0.3140
	Six-node elements								
a	0.9516	1.0398	0.0891	0.9989	1.0007	0.0018	0.9979	1.0012	0.0033
b	0.9950	1.0222	0.0272	0.9986	1.0023	0.0038	0.9980	1.0105	0.0125
c	0.9320	0.9836	0.0897	0.9972	1.0044	0.0073	0.9931	1.0029	0.0098
d	0.9259	0.9594	0.1223	0.9972	1.0076	0.0104	0.9932	1.0015	0.0084
e	0.9481	1.0240	0.0782	0.9978	1.0051	0.0073	0.9960	1.0114	0.0154
f	0.9816	1.0045	0.0232	0.9978	1.0038	0.0060	0.9942	1.0054	0.0111
g	0.8972	1.1625	0.2653	0.9674	1.0125	0.0461	0.9419	1.0025	0.0642
h	0.8868	1.2142	0.3274	0.9465	0.9961	0.0605	0.9138	0.9879	0.1066

Table 4.5: RCPe - Effectivity bounds and robustness indices for quadrangular elements applied to Class I problems (the mesh patterns are illustrated in Figure 4.15, the aspect ratio is 1/1).

	REP			RCP1e			RCP2e		
Mesh	θ_L	θ_U	R	θ_L	θ_U	R	θ_L	θ_U	R
	Four-node elements								
a	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
b	0.9987	1.0174	0.0187	0.9999	1.0011	0.0011	0.9998	1.0079	0.0081
c	0.9979	1.0061	0.0082	0.9981	1.0016	0.0035	0.9926	0.9995	0.0079
d	0.9798	1.0680	0.0882	0.9364	1.0025	0.0704	0.9354	0.9924	0.0767
e	0.9833	1.0680	0.0847	0.9311	1.0032	0.0771	0.9280	0.9980	0.0797
	Nine-node elements								
a	1.0000	1.0000	0.0000	0.9981	0.9981	0.0038	0.9981	0.9981	0.0038
b	0.9816	1.0045	0.0232	0.9971	1.0022	0.0050	0.9930	1.0025	0.0095
c	0.9967	1.0131	0.0164	0.9978	1.0002	0.0024	0.9983	1.0028	0.0045
d	0.9911	1.0329	0.0418	0.9951	1.0021	0.0071	0.9936	1.0026	0.0091
e	0.9909	1.0334	0.0425	0.9946	1.0021	0.0075	0.9920	1.0009	0.0090

Table 4.6: RCPn - Effectivity bounds and robustness indices for regular meshes of triangular and quadrangular elements applied to Class I problems (the mesh patterns are illustrated in Figures 4.14(a) and 4.15(a)).

	REP			RCP1n & RCP2n		
Aspect ratio	θ_L	θ_U	R	θ_L	θ_U	R
	Six-node triangular elements					
1/1	0.9516	1.0398	0.0891	0.9956	1.0114	0.0158
1/2	0.9516	1.0425	0.0916	0.9968	1.0178	0.0210
1/4	0.9516	1.0450	0.0939	0.9988	1.0250	0.0262
1/8	0.9516	1.0459	0.0948	0.9997	1.0279	0.0283
1/16	0.9516	1.0462	0.0950	0.9999	1.0288	0.0288
1/32	0.9516	1.0462	0.0950	1.0000	1.0290	0.0290
1/64	0.9516	1.0463	0.0951	1.0000	1.0290	0.0290
	Nine-node quadrangular elements					
1/1	1.0000	1.0000	0.0000	1.0107	1.0107	0.0214
1/2	1.0000	1.0000	0.0000	1.0030	1.0297	0.0327
1/4	1.0000	1.0000	0.0000	1.0008	1.0530	0.0538
1/8	1.0000	1.0000	0.0000	1.0002	1.0659	0.0661
1/16	1.0000	1.0000	0.0000	1.0000	1.0701	0.0702
1/32	1.0000	1.0000	0.0000	1.0000	1.0713	0.0713
1/64	1.0000	1.0000	0.0000	1.0000	1.0716	0.0716

Table 4.7: RCPe - Effectivity bounds and robustness indices for regular meshes of triangular and quadrangular elements applied to Class I problems (the mesh patterns are illustrated in Figures 4.14(a) and 4.15(a)).

Aspect	REP			RCP1e			RCP2e		
ratio	θ_L	θ_U	R	θ_L	θ_U	R	θ_L	θ_U	R
	Six-node triangular elements								
1/1	0.9516	1.0398	0.0891	0.9989	1.0007	0.0018	0.9979	1.0012	0.0033
1/4	0.9516	1.0450	0.0939	0.9984	1.0021	0.0037	0.9981	1.0049	0.0068
1/16	0.9516	1.0462	0.0950	0.9983	1.0025	0.0042	0.9981	1.0080	0.0099
1/64	0.9516	1.0463	0.0951	0.9983	1.0025	0.0042	0.9981	1.0082	0.0102
	Nine-node quadrangular elements								
1/1	1.0000	1.0000	0.0000	0.9981	0.9981	0.0038	0.9981	0.9981	0.0038
1/4	1.0000	1.0000	0.0000	0.9942	0.9998	0.0060	0.9942	0.9998	0.0060
1/16	1.0000	1.0000	0.0000	0.9938	1.0000	0.0062	0.9938	1.0000	0.0062
1/64	1.0000	1.0000	0.0000	0.9938	1.0000	0.0062	0.9938	1.0000	0.0062

mance are observed using the RCP2 versions. It should be remarked that REP is not stable for linear triangular elements on the mesh patterns in Figures 4.14(c) and (f). Thus, the corresponding results have been obtained using the stabilization technique suggested in [Boroomand & Zienkiewicz, 1997b].

4.6.1.2 Class II problems

Class II problems refer to plane elasticity problems and are described in Appendix A.2. In particular, plane strain conditions and elastic isotropic properties are considered. The values of Young's modulus and Poisson's ratio are assumed as 1000 and 0.3, respectively. Tables 4.8 and 4.9 collect the results obtained for triangular and quadrangular elements using the RCPn procedure and the mesh patterns illustrated in Figures 4.14 and 4.15, respectively. Tables 4.12 and 4.10 collect some results obtained with the RCPe procedure. Finally, Tables 4.12 and 4.13 shows the influence of the aspect ratio for regular meshes of quadratic elements.

The above comments can be substantially repeated for triangular elements applied to the plane strain case. In particular, super-convergence is achieved for linear triangular elements on the simple patterns of Figures 4.14(a) and (c), also for aspect ratios different from 1/1. This ideal behavior is not observed using quadratic triangular elements, but the results reported in Tables 4.6, 4.7, 4.12 and 4.13 indicate a good performance of the RCP procedure.

In the case of quadrangular elements, both REP and RCP yield full superconvergence only for four-node elements on regular meshes. However, the robustness indices in Tables 4.9, 4.11, 4.12 and 4.13 reveal an excellent performance also for nine-node elements on regular meshes. The results obtained on distorted patterns are very good though not exact. These numerical tests confirm the good quality of the RCP procedure, already experienced on the previous class of problems. The asymptotic behaviour of the RCP recovery is similar to that of the REP recovery. Actually, the RCP1 version appears to be slightly superior to REP. Passing to the RCP2 versions no remarkable improvements of robustness are observed.

4.6.2 Benchmark tests

In the following, the numerical validation of the RCP recovery is completed by some benchmark computations, which provide valuable information on the pre-asymptotic behaviour.

Table 4.8: RCPn - Effectivity bounds and robustness indices for triangular elements applied to Class II problems (the mesh patterns are illustrated in Figure 4.14, the aspect ratio is 1/1).

	REP			RCP1n			RCP2n		
Mesh	θ_L	θ_U	R	θ_L	θ_U	R	θ_L	θ_U	R
	Three-node elements								
a	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
b	0.9763	1.0116	0.0358	0.8815	1.0063	0.1407	0.9207	1.0090	0.0950
c	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
d	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
e	0.9628	1.0165	0.0548	0.9447	1.0030	0.0615	0.9630	1.0062	0.0446
f	0.9590	1.0273	0.0693	0.8845	1.0038	0.1344	0.8826	1.0040	0.1369
g	0.8136	1.8299	1.0163	0.5533	1.0124	0.8196	0.5753	1.0308	0.7683
h	0.8270	1.8261	0.9991	0.4486	1.0021	1.2313	0.5005	0.9825	1.0157
	Six-node elements								
a	0.9558	1.1018	0.1461	0.9923	1.0361	0.0438	0.9923	1.0361	0.0438
b	0.9691	1.0757	0.1066	0.9811	1.0325	0.0513	0.9565	1.0126	0.0579
c	0.8587	0.9993	0.1653	0.8743	1.0216	0.1650	0.8743	1.0216	0.1650
d	0.8448	0.9871	0.1968	0.8743	1.0216	0.1650	0.8743	1.0216	0.1650
e	0.9174	1.0529	0.1403	0.9468	1.0216	0.0773	0.9184	1.0093	0.0980
f	0.8609	1.0300	0.1907	0.8747	1.0115	0.1546	0.8484	1.0040	0.1826
g	0.6944	1.3145	0.6794	0.7188	1.0034	0.3946	0.6617	1.0129	0.5240
h	0.7430	1.3508	0.6078	0.6740	0.9969	0.4867	0.6144	0.9801	0.6478

Table 4.9: RCPn - Effectivity bounds and robustness indices for quadrangular elements applied to Class II problems (the mesh patterns are illustrated in Figure 4.15, the aspect ratio is 1/1).

	REP			RCP1n			RCP2n		
Mesh	θ_L	θ_U	R	θ_L	θ_U	R	θ_L	θ_U	R
	Four-node elements								
a	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
b	0.9898	1.0231	0.0333	0.9895	1.0249	0.0354	0.9655	1.0014	0.0371
c	0.9924	1.0074	0.0150	0.9921	1.0066	0.0145	0.9827	1.0100	0.0274
d	0.9411	1.0975	0.1564	0.8968	1.0237	0.1383	0.8925	1.0319	0.1514
e	0.9517	1.0870	0.1353	0.9064	1.0191	0.1220	0.9002	1.0317	0.1416
	Nine-node elements								
a	1.0025	1.0102	0.0127	1.0000	1.0360	0.0360	1.0000	1.0360	0.0360
b	0.9834	1.0219	0.0386	0.9804	1.0203	0.0399	0.9420	1.0180	0.0792
c	0.9919	1.0170	0.0251	0.9974	1.0208	0.0234	0.9683	1.0127	0.0453
d	0.9675	1.0246	0.0576	0.9625	1.0028	0.0417	0.9049	0.9936	0.1116
e	0.9664	1.0235	0.0577	0.9562	1.0027	0.0485	0.9015	0.9937	0.1156

Table 4.10: RCPe - Effectivity bounds and robustness indices for triangular elements applied to Class II problems (the mesh patterns are illustrated in Figure 4.14, the aspect ratio is 1/1).

	REP			RCP1e			RCP2e		
Mesh	θ_L	θ_U	R	θ_L	θ_U	R	θ_L	θ_U	R
	Three-node elements								
a	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
b	0.9763	1.0116	0.0358	0.9642	1.0010	0.0381	0.9630	1.0011	0.0395
c	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
d	1.0000	1.0000	0.0000	0.9802	1.0005	0.0207	0.9964	1.0075	0.0112
e	0.9628	1.0165	0.0548	0.9761	1.0012	0.0257	0.9798	1.0014	0.0220
f	0.9590	1.0273	0.0693	0.9764	1.0026	0.0268	0.9847	1.0022	0.0177
g	0.8136	1.8299	1.0163	0.7612	1.0020	0.3157	0.7371	1.0039	0.3605
h	0.8270	1.8261	0.9991	0.6377	1.0016	0.5698	0.6336	1.0008	0.5791
	Six-node elements								
a	0.9558	1.1018	0.1461	0.9971	1.0018	0.0046	0.9973	1.0044	0.0070
b	0.9691	1.0757	0.1066	0.9957	1.0035	0.0078	0.9958	1.0161	0.0203
c	0.8587	0.9993	0.1653	0.9971	1.0345	0.0374	0.9894	1.0349	0.0455
d	0.8448	0.9871	0.1968	0.9970	1.0345	0.0376	0.9911	1.0349	0.0438
e	0.9174	1.0529	0.1403	0.9953	1.0149	0.0197	0.9937	1.0300	0.0363
f	0.8609	1.0300	0.1907	0.9961	1.0280	0.0319	0.9870	1.0272	0.0402
g	0.6944	1.3145	0.6794	0.9156	1.0157	0.1077	0.8676	1.0106	0.1631
h	0.7430	1.3508	0.6078	0.8626	1.0077	0.1669	0.8022	1.0008	0.2473

Table 4.11: RCPe - Effectivity bounds and robustness indices for quadrangular elements applied to Class II problems (the mesh patterns are illustrated in Figure 4.15, the aspect ratio is 1/1).

	REP			RCP1e			RCP2e		
Mesh	θ_L	θ_U	R	θ_L	θ_U	R	θ_L	θ_U	R
	Four-node elements								
a	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000	1.0000	1.0000	0.0000
b	0.9898	1.0231	0.0333	0.9991	1.0018	0.0027	0.9967	1.0098	0.0131
c	0.9924	1.0074	0.0150	0.9935	1.0038	0.0103	0.9886	1.0010	0.0125
d	0.9411	1.0975	0.1564	0.9137	1.0031	0.0976	0.9096	0.9952	0.1041
e	0.9517	1.0870	0.1353	0.9219	1.0056	0.0903	0.9148	0.9990	0.0941
	Nine-node elements								
a	1.0025	1.0102	0.0127	0.9965	1.0000	0.0035	0.9965	1.0000	0.0035
b	0.9834	1.0219	0.0386	0.9965	1.0079	0.0114	0.9918	1.0075	0.0157
c	0.9919	1.0170	0.0251	0.9976	1.0013	0.0037	0.9980	1.0079	0.0099
d	0.9675	1.0246	0.0576	0.9926	1.0041	0.0115	0.9903	1.0038	0.0136
e	0.9664	1.0235	0.0577	0.9931	1.0042	0.0111	0.9871	1.0028	0.0158

Table 4.12: RCPn - Effectivity bounds and robustness indices for regular meshes of triangular and quadrangular elements applied to Class II problems (the mesh patterns are illustrated in Figures 4.14(a) and 4.14(a)).

	REP			RCP1n & RCP2n		
Aspect ratio	θ_L	θ_U	R	θ_L	θ_U	R
	Six-node triangular elements					
1/1	0.9558	1.1018	0.1461	0.9923	1.036	0.0438
1/2	0.9548	1.0980	0.1432	0.9888	1.0392	0.0505
1/4	0.9527	1.0928	0.1401	0.9845	1.0437	0.0592
1/8	0.9519	1.0900	0.1381	0.9828	1.0457	0.0628
1/16	0.9517	1.0889	0.1372	0.9824	1.0462	0.0639
1/32	0.9516	1.0886	0.1370	0.9823	1.0464	0.0641
1/64	0.9516	1.0885	0.1369	0.9822	1.0464	0.0642
	Nine-node quadrangular elements					
1/1	1.0025	1.0102	0.0127	1.0000	1.0360	0.0360
1/2	1.0025	1.0181	0.0205	1.0000	1.0524	0.0524
1/4	1.0002	1.0136	0.0139	1.0000	1.0656	0.0656
1/8	0.9997	1.0061	0.0064	1.0000	1.0700	0.0700
1/16	0.9968	1.0061	0.0094	1.0000	1.0712	0.0712
1/32	0.9950	1.0061	0.0111	1.0000	1.0715	0.0715
1/64	0.9945	1.0061	0.0116	1.0000	1.0716	0.0716

Table 4.13: RCPe - Effectivity bounds and robustness indices for regular meshes of triangular and quadrangular elements applied to Class II problems (the mesh patterns are illustrated in Figures 4.14(a) and 4.15(a)).

Aspect	REP			RCP1e			RCP2e		
ratio	θ_L	θ_U	R	θ_L	θ_U	R	θ_L	θ_U	R
	Six-node triangular elements								
1/1	0.9558	1.1018	0.1461	0.9971	1.0018	0.0046	0.9973	1.0044	0.0070
1/4	0.9527	1.0928	0.1401	0.9947	1.0023	0.0076	0.9973	1.0048	0.0075
1/16	0.9517	1.0889	0.1372	0.9944	1.0027	0.0083	0.9973	1.0067	0.0094
1/64	0.9516	1.0885	0.1369	0.9944	1.0028	0.0084	0.9973	1.0075	0.0102
	Nine-node quadrangular elements								
1/1	1.0025	1.0102	0.0127	0.9965	1.0000	0.0035	0.9965	1.0000	0.0035
1/4	1.0002	1.0136	0.0139	0.9941	1.0000	0.0059	0.9941	1.0000	0.0059
1/16	0.9968	1.0061	0.0094	0.9938	1.0000	0.0062	0.9938	1.0000	0.0062
1/64	0.9945	1.0061	0.0116	0.9938	1.0000	0.0062	0.9938	1.0000	0.0062

4.6.2.1 Homogeneous elestic bar

Consider a one-dimensional sample problem governed by the following equations

$$k\frac{\mathrm{d}^2u}{\mathrm{d}x^2} = b, \qquad 0 < x < L,$$

$$u = 0, \qquad x = 0, L.$$

This problem can be viewed as the equilibrium problem of a clamped one-dimensional homogeneous elastic bar, of length L and axial rigidity k, subjected to a distributed axial force b. The values of the parameters are assumed as $L = 4$ and $k = 1$. The loading function is taken as

$$b = -\frac{\pi}{L}\cos\left(\frac{\pi x}{L}\right).$$

The exact solution in terms of stress is

$$\sigma = k\frac{\mathrm{d}u}{\mathrm{d}x} = \sin\left(\frac{\pi x}{L}\right).$$

This problem was used by Boroomand and Zienkiewicz [Boroomand & Zienkiewicz, 1997b] to illustrate the convergence properties of the REP procedure. The finite

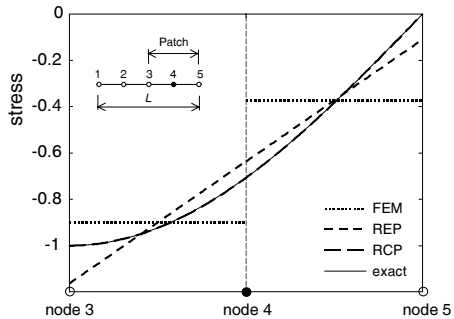


Figure 4.16: Recovered stress solutions for linear elements

element solution is obtained using uniform meshes of two-node (linear) elements. The undeterminate self-equilibrated stress in the RCP1n procedure is simply a constant term:

$$\sigma^*_{ph} = \beta.$$

Thus, the present procedure involves only one unknown parameter for each patch. The stress term in equilibrium with the prescribed load within the patch is taken as

$$\sigma^*_{pp} = \int_0^x b\,\mathrm{d}x.$$

In this problem, the stress fields recovered by the RCP1n procedure with and

without boundary patches are almost coincident. Thus, no distinction is made between RCP1n and RCP1nb in what follows. In the REP procedure, a linear stress expansion is assumed within the patch, so that two unknown parameters for each patch are involved.

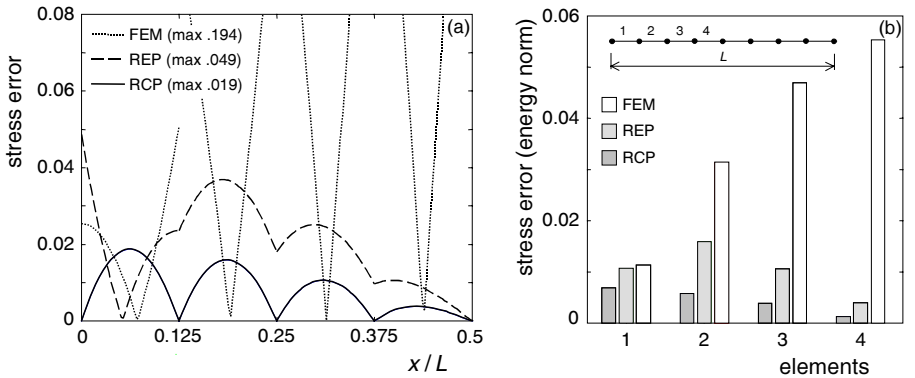


Figure 4.17: Error distributions for linear elements: (a) local error (maximum values are reported in brackets), (b) element error in energy norm

The local recovered solutions σ_p^* over a patch are shown in Figure 4.16 for the four-element mesh and compared with the finite element solution. Notice that the RCP-solution is indistinguishable from the exact one and much more accurate than the REP-solution. This is confirmed by the error distributions in the smoothed stresses σ^* reported in Figure 4.17 for the eight-element mesh. In particular, Figure 4.17(a) shows the local error and Figure 4.17(b) the element error. Due to symmetry, the distributions are plotted only on one half of the domain. The convergence properties of the recovery procedures are illustrated by Figure 4.18. Figure 4.18(a) shows the rate of convergence of the local error in the nodal stress value at point $x = (3/4)L$. Figure 4.18(b) shows the rate of convergence of the global error in energy norm. These graphs reveal that nodal stress values recovered by the RCP1n procedure are ultra-convergent, while those recovered by the REP procedure are super-convergent. Unfortunately, ultra-convergence is lost for the global error, because of the nodal interpolation by linear shape functions, and super-convergence is achieved by both the REP and RCP1n procedures.

4.6.2.2 Thin membrane subjected to transverse loading (Class I problem)

The second benchmark test is a typical problem of the class described in Appendix

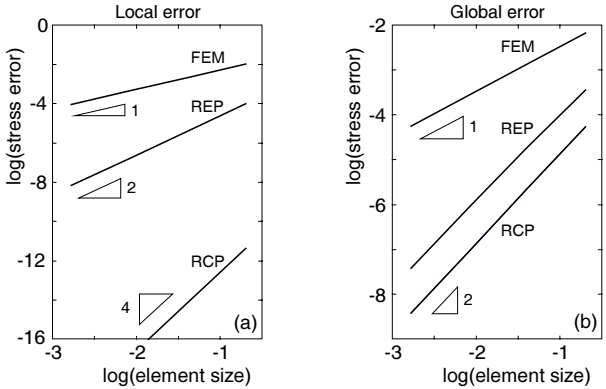


Figure 4.18: Rate of convergence for linear elements: (a) local error at point $x/L = 3/4$, (b) global error in energy norm (Problem No. 1)

A.1. In particular, the equilibrium problem of a thin square membrane of uniform thickness with fixed boundary is considered. Both the sidelength of the domain and the parameter k are taken of unity value. Two different load cases are considered. In load case No. 1, function b is selected such that the exact solution is given by

$$u(x, y) = xy \sin(\pi x) \sin(\pi y).$$

In load case No. 2, function b is selected such that the exact solution is given by

$$u(x, y) = x(1 - x)y(1 - y)(1 + 2x + 7y).$$

This problem has been extensively used in the literature to test recovery procedures on a bi-dimensional case. The finite element analysis is carried out on different meshes using quadrangular four- and nine node elements.

The error distributions for load case No. 1 on a regular mesh of four by four bi-linear elements are shown in Figures 4.19 and 4.20, in order to asses the accuracy properties of the RCP procedure. In particular, the distributions of the local error are shown in Figure 4.19 and the distributions of the element error are shown in Figure 4.20, together with the values of the relative percentage error η (Equation (2.37)). The local error is defined as the absolute value of the pointwise difference between the recovered stress $\boldsymbol{\sigma}^*$ and the exact stress $\boldsymbol{\sigma}$ for RCP and REP, while between the finite element stress $\boldsymbol{\sigma}^h$ and the exact stress $\boldsymbol{\sigma}$ for FEM. These results show a very good performance of the RCP procedure. In fact, RCP is much more accurate than the REP procedure almost everywhere, and the associated errors are more uniformly distributed over the domain. Moreover, RCP1e is, notwithstanding its simplicity, about as accurate as RCP1n. The figures reveal also that the RCPe seems to be slightly more sensitive than the RCPn to the increase of the stress

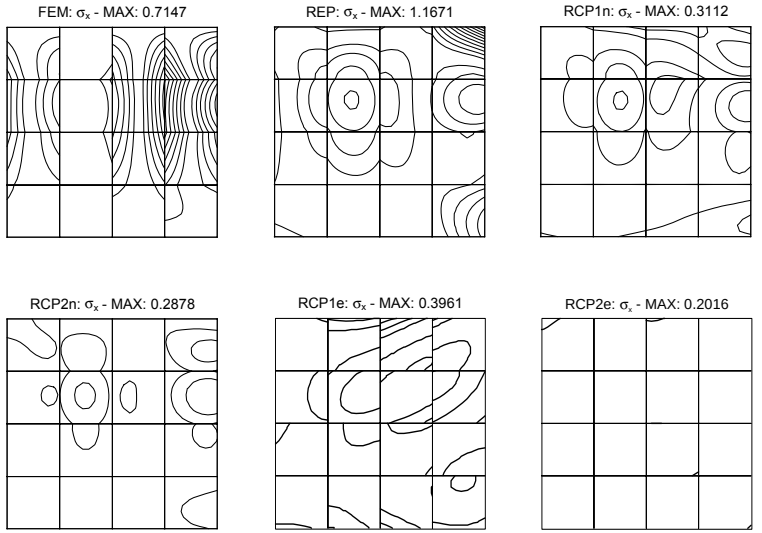


Figure 4.19: Thin membrane subjected to transverse loading, load case No. 1 - Local error distributions on a regular mesh of four-node elements (contour interval = 0.1)

approximation order. In particular, RCP2e remarkably improves accuracy all over the domain and the relative percentage error passes from 22% (RCP1e) to only 7%. Figure 4.21 shows the results of the convergence analysis of the RCP procedure, carried out on sequentially refined uniform meshes of bi-linear and bi-quadratic elements. In particular, the pointwise convergence of nodal stresses, evaluated at point $(x, y) = (0.5, 0.5)$ where $\sigma_x = \sigma_y$, is shown in Figures 4.21(a) and (c). Because of inter-element discontinuity, here and in the following, the single-valued RCPe and the FEM values are computed as average of nodal errors. Moreover, note that, since uniform meshes are used, also the recovered stress components turn out to be equal, so that only one graph is reported for each type of elements. The rate of convergence of the relative percentage error for bi-linear and bi-quadratic elements is shown in Figures 4.21(b) and (d). These results confirm that the RCP procedure possesses the desirable feature of superconvergence. Moreover, it can be observed that RCP2 is more accurate than RCP1, although it exhibits a similar convergence behavior, as expected from robustness tests. The effectiveness of the RCP procedure is also tested solving the present problem on an irregular unstructured mesh of bi-linear elements. The results are shown in Figures 4.22 and 4.23. In particular, Figure 4.22 shows the distributions of the local error in both the stress components, plotted as contours over the domain, and Figure 4.23 the distributions of the element error. The sensitivity of the RCP recovery to element geometry distortions is also investigated solving the some problem on a

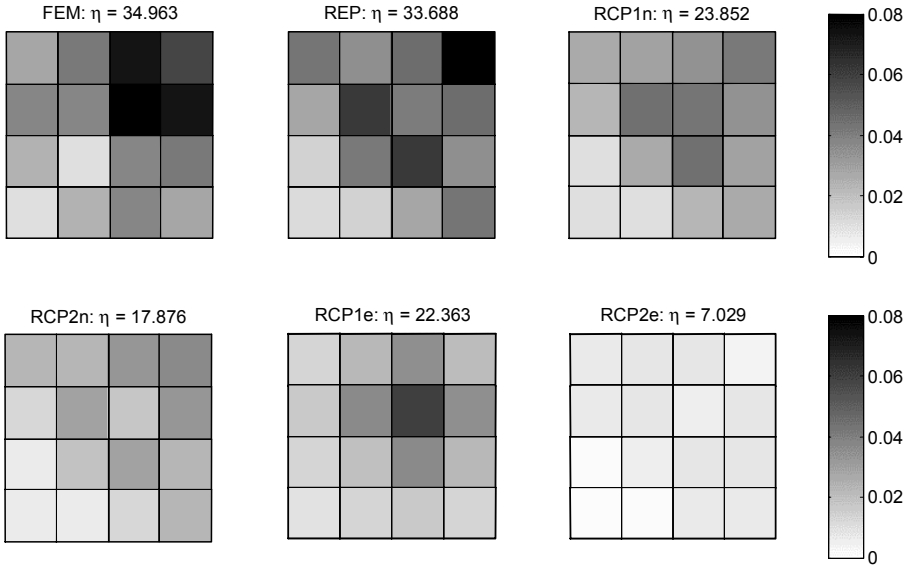


Figure 4.20: Thin membrane subjected to transverse loading, load case No. 1 - Distributions of element error in energy norm on a regular mesh of four-node elements (the relative percentage error is reported above each plot)

sequence of progressively distorted meshes. The irregular pattern used in this test is illustrated in Figure 4.26. The distortion is measured by the parameter d/L, which ranges from -0.375 to 0.375. The local error in the nodal stress value at point $x = 0.5$ and $y = 0.5$ is plotted against the distortion parameter in Figure 4.27 for both four- and nine-node elements. Figures 4.27(a) and (d) show the results obtained for the load case No. 1, while Figures 4.27(b), (c), (e) and (f) show the results obtained for the load case No. 2. In the first load case, only the results for one stress component are reported owing to symmetry considerations. It is interesting to notice the very good performance of the RCP procedure also for irregular meshes. The nodal stress values recovered by the present procedure are weakly sensitive to element geometry distortions and very accurate also for patches of highly distorted elements. In particular, the RCP-results are much less sensitive to geometry distortions than the REP-results. This is especially evident for nine-node elements and confirms the good quality of RCP also in the pre-asymptotic regime.

These experiments confirm what has been previously observed for uniform meshes, so demonstrating that RCPn is reliable also for arbitrary irregular meshes. This is a feature of importance in error estimation and adaptive analysis proce-

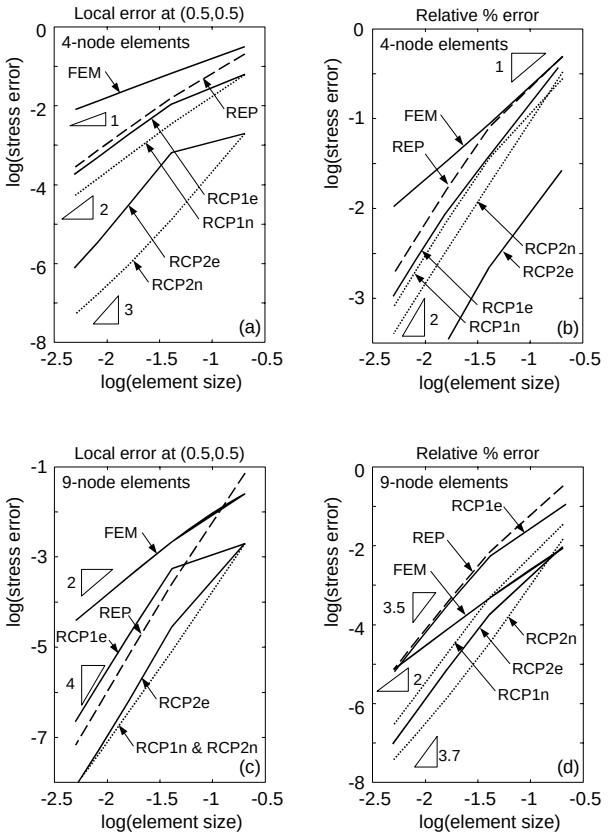


Figure 4.21: Thin membrane subjected to transverse loading, load case No. 1 - Rate of convergence: (a) local error at point $x = 0.5$, $y = 0.5$ ($\sigma_x = \sigma_y$) for four-node elements, (b) relative percentage error for four-node elements, (c) local error at point $x = 0.5$, $y = 0.5$ ($\sigma_x = \sigma_y$) for nine-node elements, (d) relative percentage error for nine-node elements

dures.

In order to evaluate the effect of using the boundary patch in the RCP procedure based on node patches, the above test have been conducted both with and without using boundary patches. The corresponding results, reported in Appendix B.1 show that using boundary patches actually enhances accuracy near boundaries.

Finally, using the unstructured mesh, the error distribution estimated basing on the present RCP procedure, according to Equations (4.1) and (4.2), is compared with the exact FEM error by Figure 4.28. As it can be observed, an excellent behavior of the RCP error estimator is experienced, for both bi-linear and bi-quadratic elements. Similar results are obtained using triangular elements. These are not included for the sake of brevity.

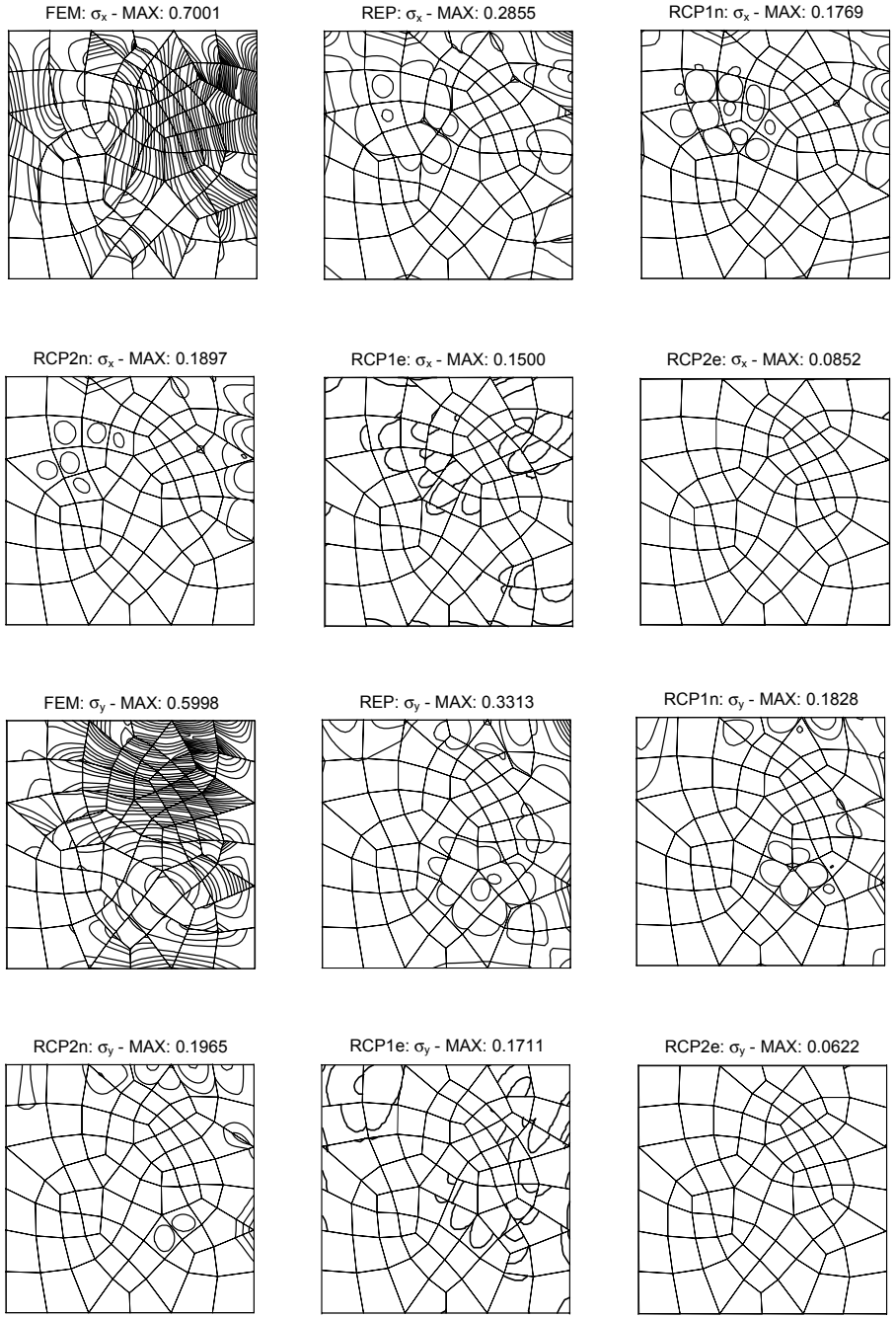


Figure 4.22: Thin membrane subjected to transverse loading, load case No. 1 - Local error distributions on an unstructured mesh of four-node elements (contour interval = 0.05)

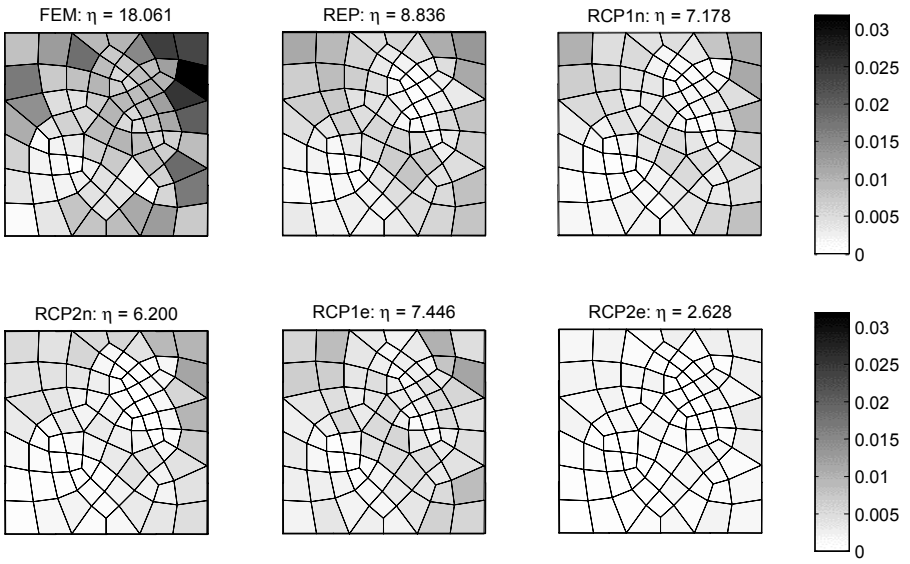


Figure 4.23: Thin membrane subjected to transverse loading, load case No. 1 - Distributions of element error in energy norm on an unstructured mesh of four-node elements (the relative percentage error is reported above each plot)

4.6.2.3 Stretched plate with circular hole (Class II problem)

The second class of problems considered refers to problem class described in the Appendix A.2. In particular, a portion of an infinite plate with a central circular hole of radius a subjected to a unidirectional tensile load $\sigma_0 = 1$ is considered. Plane stress conditions are assumed with elastic isotropic properties: Poisson's ratio is taken as 0.3 and Young's modulus as 1000. The geometry of the problem and the prescribed boundary conditions are shown in Figure 4.29. This is a popular benchmark test for which the exact solution reads as

$$\sigma_x = 1 - \frac{a^2}{r^2}\left(\frac{3}{2}\cos 2\theta + \cos 4\theta\right) + \frac{3}{2}\frac{a^4}{r^4}\cos 4\theta,$$
$$\sigma_y = -\frac{a^2}{r^2}\left(\frac{1}{2}\cos 2\theta - \cos 4\theta\right) - \frac{3}{2}\frac{a^4}{r^4}\cos 4\theta,$$
$$\tau_{xy} = -\frac{a^2}{r^2}\left(\frac{1}{2}\sin 2\theta + \sin 4\theta\right) + \frac{3}{2}\frac{a^4}{r^4}\sin 4\theta.$$

where (r, θ) are the usual polar coordinates (see Figure 4.29). The radius of the hole is taken as $a = 1$.

Three meshes are used in the finite element analysis and shown in Figure 4.29. The coarsest mesh consists of 72 elements. The subsequent meshes are created

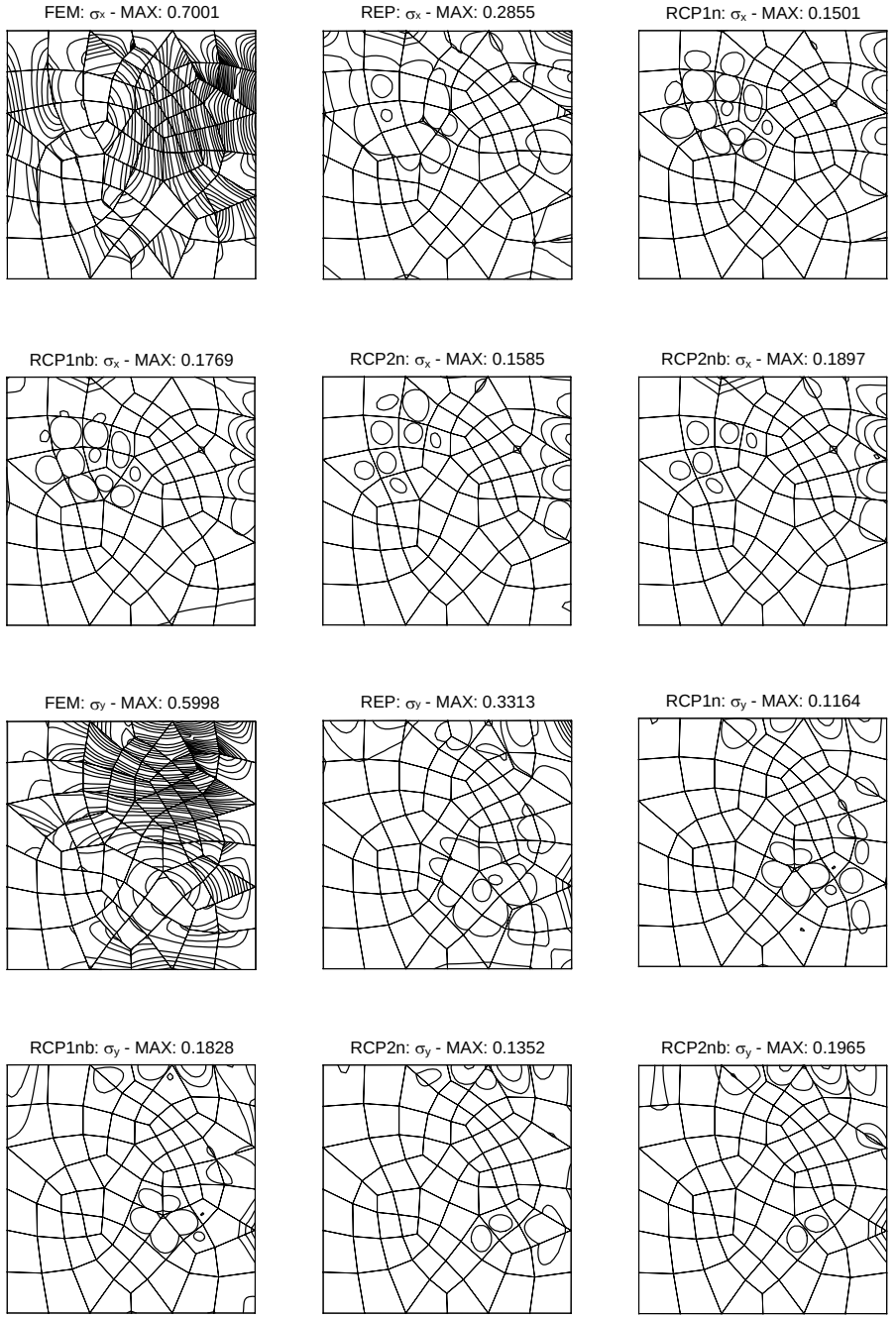


Figure 4.24: Thin membrane subjected to transverse loading - Error distributions on an unstructured mesh of four-node elements (contour interval = 0.05,load case No. 1)

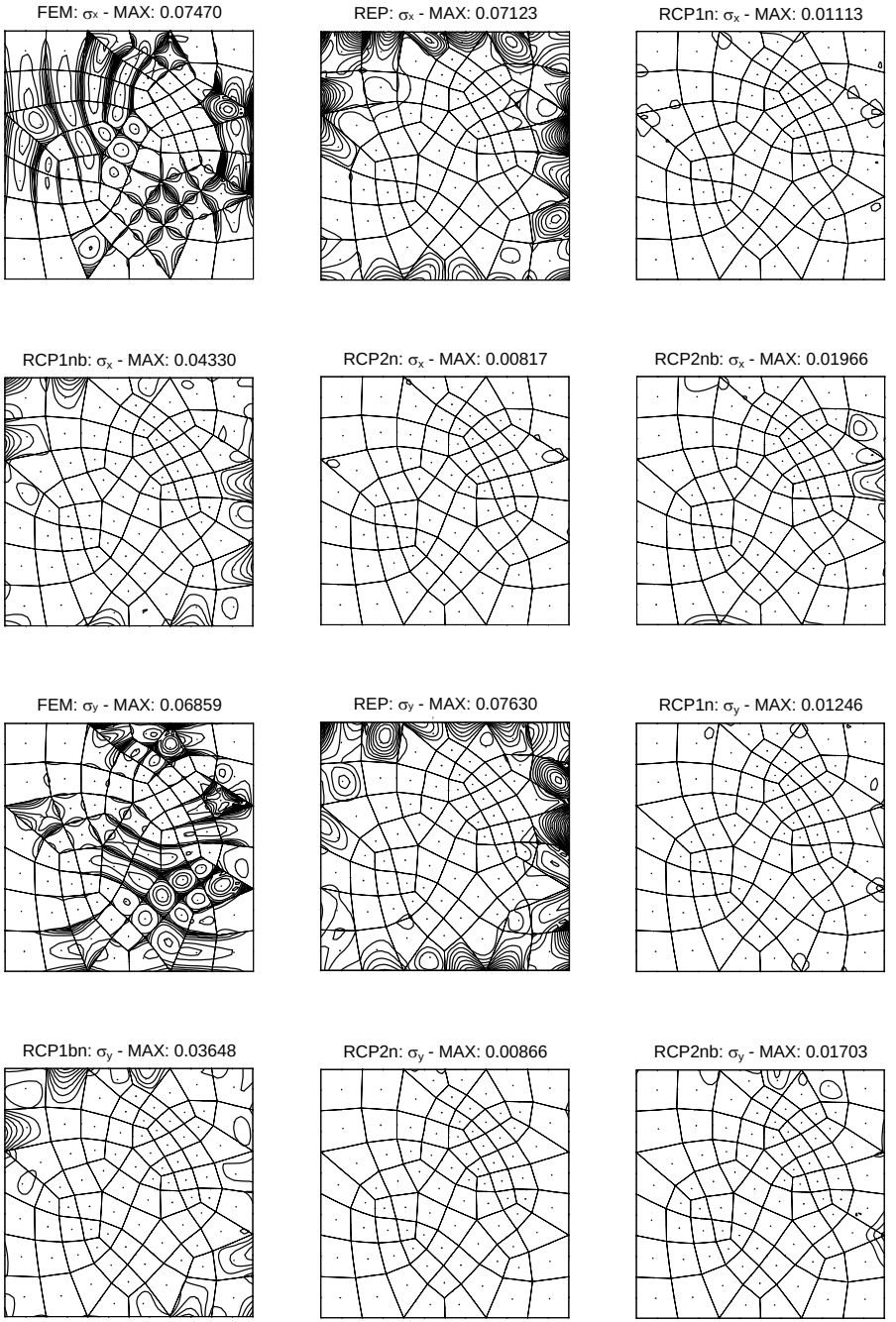


Figure 4.25: Thin membrane subjected to transverse loading - Error distributions on an unstructured mesh of nine-node elements (contour interval = 0.005, load case No. 1)

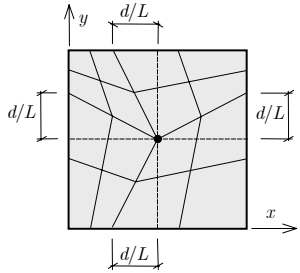


Figure 4.26: Thin membrane subjected to transverse loading - Irregular pattern used in sensitivity test to mesh distortion

by a uniform refining process: each element of the preceding mesh is divided into four new elements. The following quadrangular elements are employed: bi-linear four-node elements, quadratic serendipity eight-node elements and bi-quadratic nine-node elements. The stress approximations for the RCP and REP procedures are assumed as described in Section A.2.2. In particular, the same approximations are used for eight- and nine-node elements.

The distributions of the local error for four-node elements are displayed in Figures 4.30 and 4.31. The distributions of the element error for the same elements are displayed in Figure 4.32. Inspecting these figures reveals that conclusions analogous to those of the preceding benchmark problem can be drawn, so confirming the excellent performance of the RCP procedure also for plane elasticity problems. The convergence properties of the procedure are illustrated in Figure 4.33: the pointwise convergence of the recovered stress components, evaluated at the vertex node ($r = 1.5, \theta = \pi/4$) indicated on each mesh in Figure 4.29, is shown for four-node elements by Figures 4.33(a)-(c), the convergence of the relative percentage error is shown for four- and nine-node elements by Figures 4.33(d) and (f). As it can be noted, superconvergence of nodal stresses is experienced for the RCP procedure as well as for the REP procedure. On the contrary, the relative percentage error is not superconvergent with any procedure, as expected because the meshes are not fully regular. Anyway, the rate of convergence of the original finite element solution is remarkably improved. As in the previous test, the influence of using boundary patches in the RCPn procedure has also been investigated. The results, reported in Appendix B.2 demonstrate that boundary patches generally improve accuracy, especially on coarse meshes, but sometimes may deteriorate convergence.

Finally, the convergence behavior of the RCP-based error estimator is shown in Figure 4.34(a) and (b) for four- and nine-node elements, respectively, and compared to the exact one. In particular, the errors estimated by both RCP1e and RCP2e are reported. The excellent properties of the RCP1e estimator are confirmed.

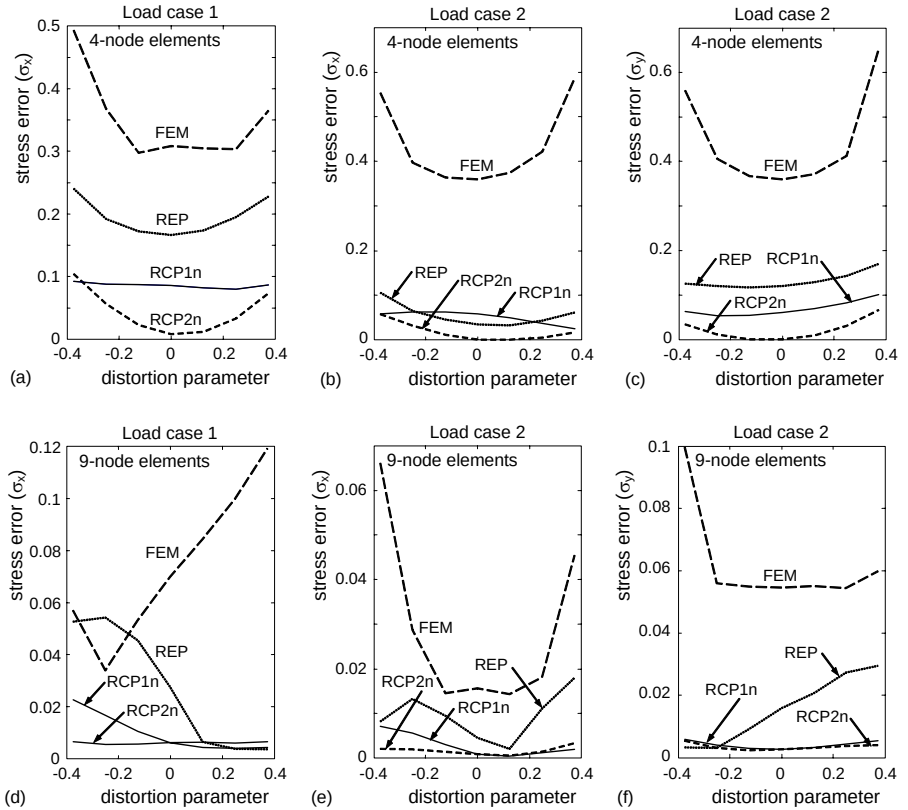


Figure 4.27: Thin membrane subjected to transverse loading - Sensitivity to mesh distortion of local error at point $x = 0.5$, $y = 0.5$: (a) four-node elements and load case No. 1, (b) and (c) four-node elements and load case No. 2, (d) nine-node elements and load case No. 1, (e) and (f) nine-node elements and load case No. 2

Moreover, it emerges that using the error estimator based on RCP2e could be convenient in conjunction with higher order elements. Similar results are obtained using triangular elements.

4.7 Recovery based error estimators can fail?

Recovery-based estimators possess a number of attractive features that have led to their popularity. In particular, their ease of implementation, generality, and ability to produce quite accurate estimators have led to their widespread adoption, especially in the engineering community. However, the estimators also have drawbacks.

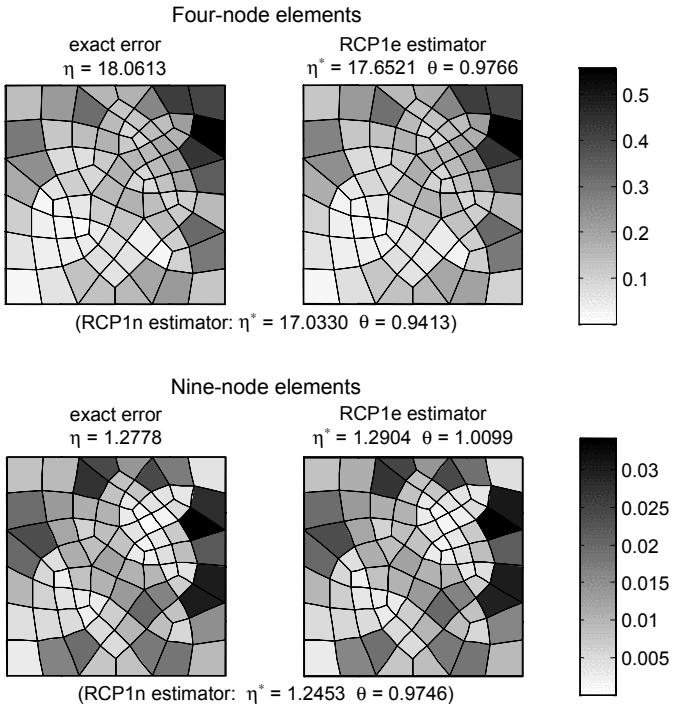


Figure 4.28: Thin membrane subjected to transverse loading, load case No. 1 - Exact and estimated distributions of the element error in energy norm on an unstructured mesh

To illustrate the dangers in the indiscriminate use of such techniques, we present a simple example by [Ainsworth & Oden, 2000]. The setting is chosen to be as simple as possible so that there is no question that the source of the problems is due to extraneous effects such as mesh distortions (the partition will be uniform) numerical quadrature (all integrals are evaluated exactly), nonsmooth coefficients (the data are smooth) or nonsmooth solution (the true solution is smooth). Furthermore, the example is even taken to be one-dimensional to avoid any question regarding mesh topology. Nevertheless, even in such an idealistic setting, it is possible to construct an example whereby the recovery based estimators produce an estimated error of zero, while the actual error can be arbitrarily large.

It would be quite wrong to dismiss such a situation as being purely academic and having no relevance to practical computation of much more complicated phenomena in multidimensions. The point is that if such anomalies can be present in such a simplified setting, then it is quite possible for similar effects to occur, possibly only locally, in more complex cases. The danger would be that the estimated error in the neighborhood of the local feature would be zero, with the net result

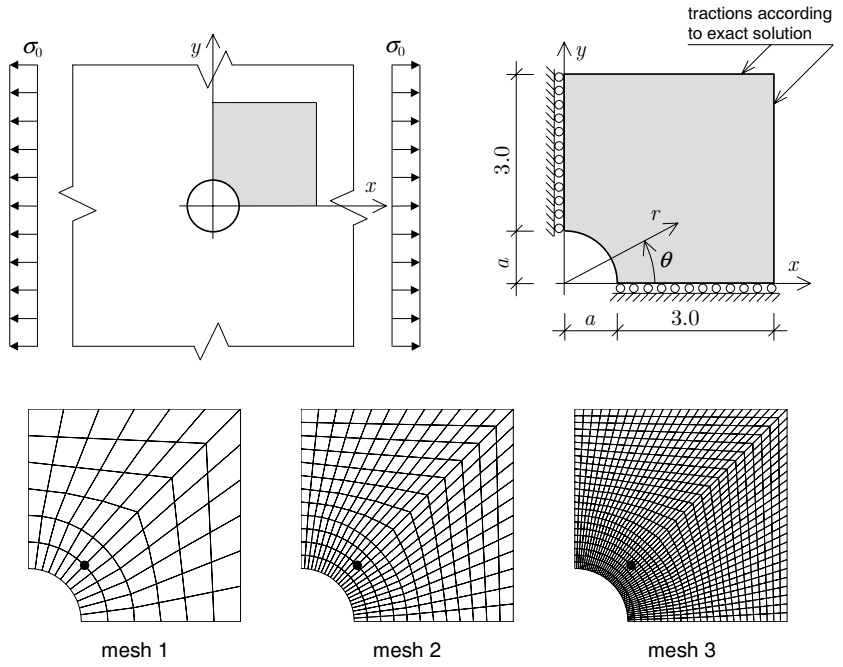


Figure 4.29: Stretched plate with circular hole

that the adaptive refinement procedure would miss the local feature altogether and the error estimator would not flag any difficulties. Obviously, this is a potentially disastrous situation. The example consists of approximating the problem

$$-u'' = f \qquad \text{in } \Omega = (0,1)\,, \tag{4.124}$$

$$u(0) = u(1) = 0. \tag{4.125}$$

The data f is chosen to be of the form

$$f(x) = \mu \sin\left(2^m \pi x\right), \tag{4.126}$$

where m is a fixed integer and μ is an arbitrary constant. The approximation consists of using piecewise linear finite elements on a uniform partition consisting of 2^n elements. Such a partition would arise beginning with a single element and performing $n-1$ successive subdivisions of the partition, corresponding to a sequence of uniform refinements. The element nodes are located at the points

$$x_k = \frac{k}{2^n}, \qquad k = 0, \ldots, 2^n. \tag{4.127}$$

It is well known that the finite element approximation u^h will actually be the piecewise linear interpolant to the exact solution u. Here, the exact solution is

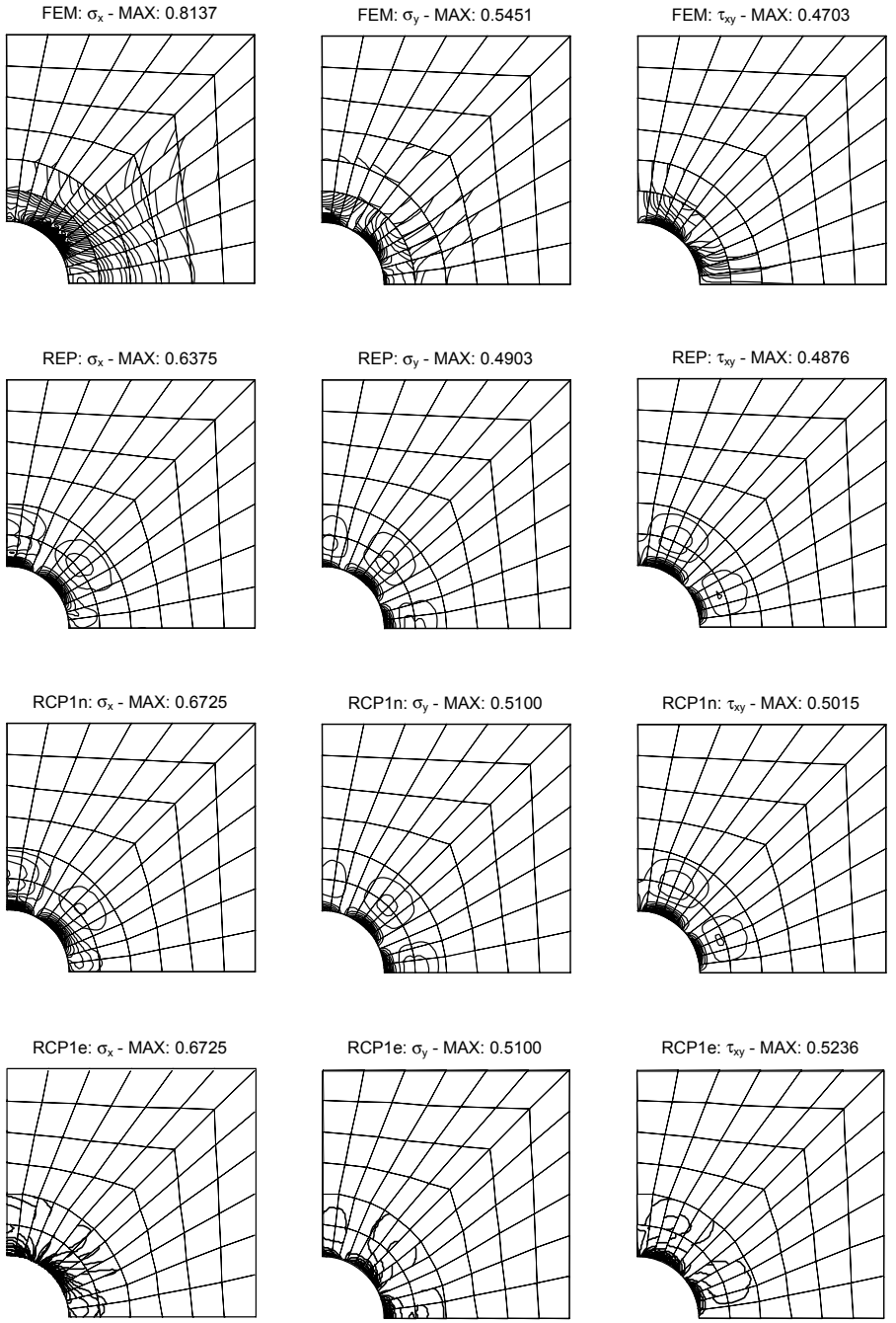


Figure 4.30: Stretched plate with circular hole - Local error distributions for four-node elements on mesh No. 1: FEM solution, REP, RCP1n and RCP1e (contour interval = 0.05)

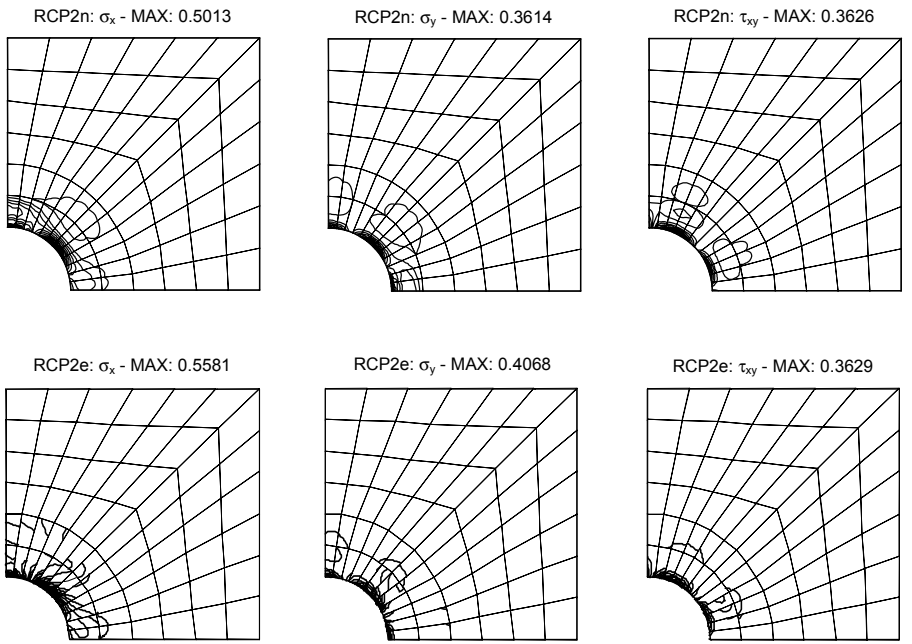


Figure 4.31: Stretched plate with circular hole - Local error distributions for four-node elements on mesh No. 1: RCP2n and RCP2e (contour interval = 0.05)

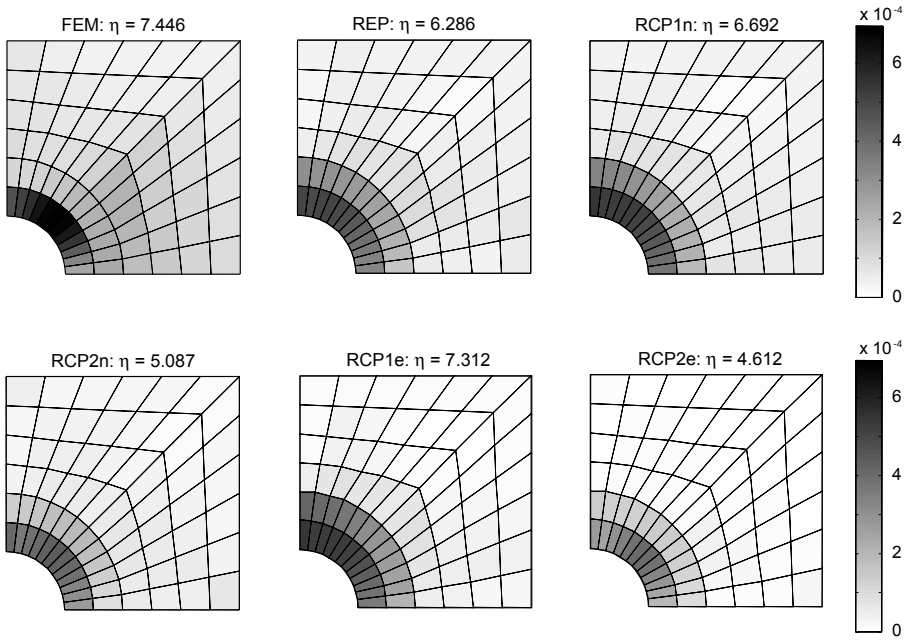


Figure 4.32: Stretched plate with circular hole - Distributions of element error in energy norm on mesh No. 1 of four-node elements (the relative percentage error is reported above each plot)

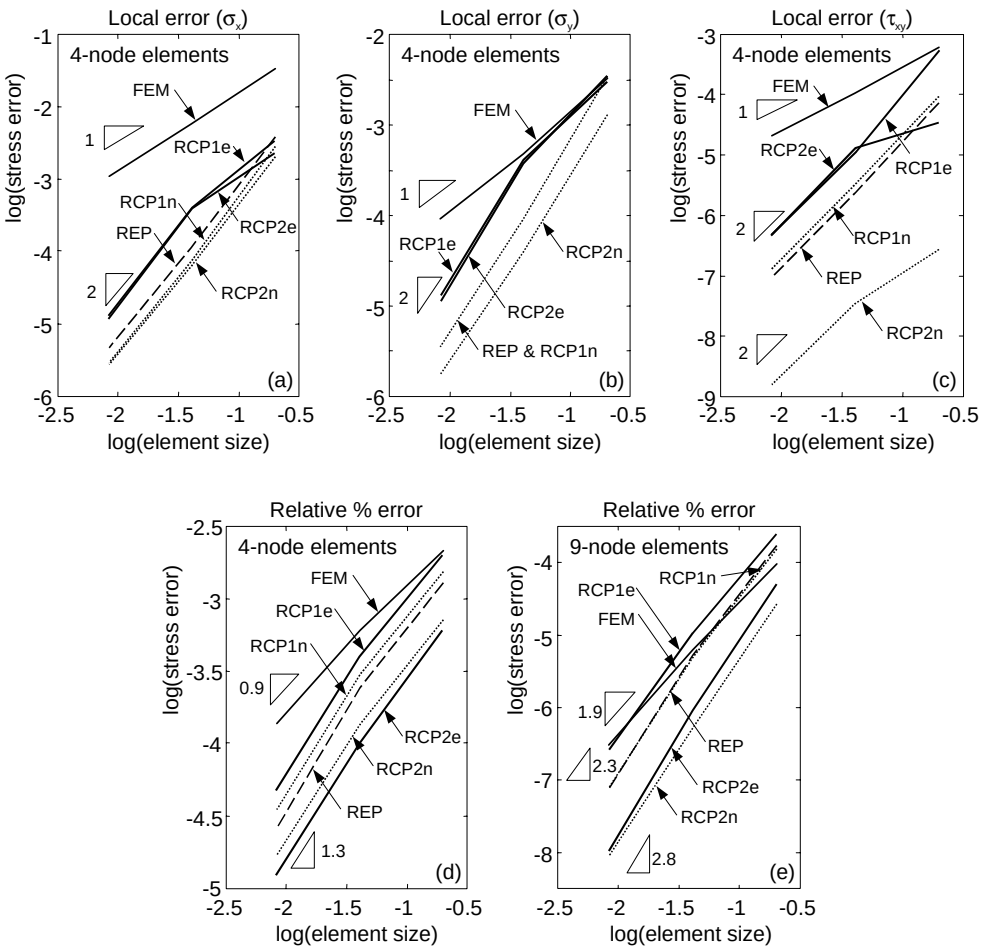


Figure 4.33: Stretched plate with circular hole - Rate of convergence: (a) local error in σ_x at point $r = 1.5$, $\theta = \pi/4$ for four-node elements, (b) local error in σ_y at point $r = 1.5$, $\theta = \pi/4$ for four-node elements, (c) local error in τ_{xy} at point $r = 1.5$, $\theta = \pi/4$ for four-node elements, (d) relative percentage error for four-node elements, (e) relative percentage error for nine-node elements

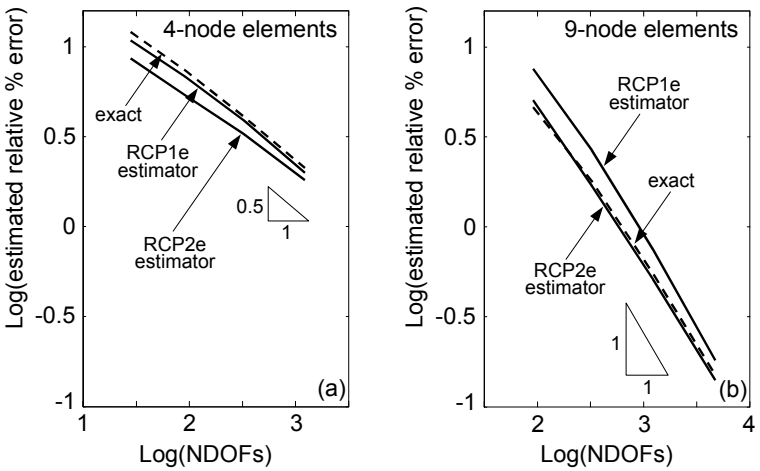


Figure 4.34: Stretched plate with circular hole - Convergence of the estimated relative percentage error: (a) four-node elements, (b) nine-node elements

given by

$$u(x) = \frac{1}{4^m \pi^2} \mu \sin(2^m \pi x) \tag{4.128}$$

and if $m \geq n$, then $u(x_k)$ vanishes identically. Consequently, the finite element approximation u^h is identically zero. Observe that, no matter how many times the partition is subdivided by increasing n, there always exists a set of data (corresponding to sufficiently large m) such that u^h will vanish identically. Obviously, the gradient of the finite element approximation also vanishes everywhere, and consequently the recovered gradient functions of u^h will also vanish everywhere. The majority of practical a posteriori error estimates are based on either the construction of an improved solution which is extracted locally from the finite element solution by some local averaging (recovery) or on the solutions of local boundary value problems set over the elements (or small subdomains associated with the elements) with input-data obtained from the local residuals. This means that the estimated error will be zero. However, the actual error is proportional to $|\mu|$ and could therefore be arbitrarily large.

Chapter 5

Adaptivity

In questo capitolo viene illustrato e discusso il ruolo dello stimatore d'errore presentato nel precedente capitolo quale parametro guida per la ricostruzione della discretizzazione finalizzata alla minimizzazione dell'errore stesso, ovvero quale parametro di controllo di analisi di tipo adattativo.

Dopo alcune considerazioni introduttive sulle varie strategie secondo cui si può essere condotta un'analisi adattativa, l'attenzione è focalizzata sull'impiego dello stimatore d'errore quale guida di analisi adattative condotte per completa rigenerazione della discretizzazione, mostrando l'efficacia dello stimatore RCP anche in questo ambito.

5.1 Mesh generation

Normal practice to solve engineering problems by means of the Finite Element Method involves increasing the number of discretization points in the computational domain and resolving the resulting system of equations to examine the relative change in the numerical solution. In general, this procedure is time consuming, it depends on the experience of the analyst, and it can be misleading if the solution has not entered an asymptotic range.

Ideally, with a robust and reliable self-adaptive scheme, one would be able to specify an initial discrete model which is sufficient to describe the geometry/topology of the domain and the boundary conditions, and to specify a desired error tolerance, according to an appropriate criterion. Then, the system would

automatically refine the model until the error measure falls below the prescribed tolerance. The process should be fully automatic and without any user intervention. This approach increases the overall reliability of the analysis procedure since it does not depend on the experience, or inexperience, of the analyst.

The need for developing better pre-processing techniques for the finite element method, for performing automated analysis, and for obtaining self-adaptive solutions (which is becoming a trend for commercial finite element software) have driven the development of automatic mesh generation algorithms, i.e. algorithms which are capable of discretizing an arbitrary geometry into a consistent finite element mesh without any user intervention.

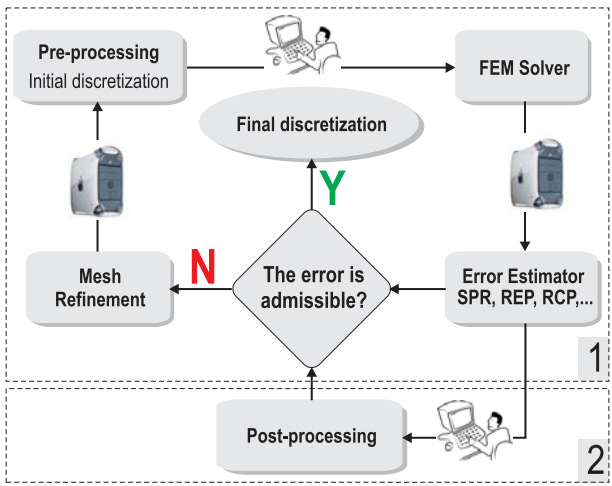


Figure 5.1: Simplified diagram of the iterative mesh design cycle: (1) self-adaptive tecnique; (1)+(2) self-adaptive tecnique with supervisor

Such software is nowadays accompanied by interactive graphics, which eases the mesh generation task considerably. Standardized data interfaces such as IGES and STEP allow CAD-generated data to be transferred between different software. The difficult step then follows of assigning the topological element descriptions to fill the required volume. However, the generation of meshes, particularly for complicated 3D structures, can still be a daunting task. Mesh generation techniques fall into two broad classes, which can be termed top-down and bottom-up.

Top-down takes for input an overall volume. This is defined to the program as a set of lines and curves in 2D, or equivalent surfaces in 3D. A mesh is automatically formed in the required volume. The problem here is that elements with badly distorted shapes can be generated to fill the space. When huge numbers of elements

are being generated in complicated 3D structures, such elements can be hard to spot by eye in the graphics. Some shape checks should be available in the software, and it is prudent to use this to check such distortions.

Bottom-up is based on designing the mesh element by element, building up using sets of commands in sequence. Hence a more continuous control on each element is available. However, distortions can occur in the later stages of the process when final parts of the overall shape have to be filled. The process is also time consuming for the user.

In both cases, the meshes produced should be checked for an accurate representation of the model, and that the elements are of suitable shape and size. Some elements will have to be somewhat distorted in the fitting process, particularly around detailed geometric features and curved boundaries [Zhu *et al.* , 1991b].

5.2 Adaptive finite element techniques

Adaptive procedures try to automatically refine, coarsen, or relocate a mesh and/or adjust the basis to achieve a solution having a specific accuracy in an optimal fashion. The computation typically begins with a trial solution generated on a coarse mesh with a low-order basis. The error of this solution is estimated. If it fails to satisfy the prescribed accuracy, adjustment are made with the goal of obtaining the desired solution with minimal effort. While adaptive finite element methods have been studied for nearly twenty years, surprising little is known about optimal strategies. Common procedures studied to date include

- local refinement and/or coarsening of a mesh
- relocating or moving a mesh
- locally varying the polynomial degree of the basis
- combinations of the precedent

We may guess that the relocating or moving a mesh method alone is generally not capable of finding a solution with a specified accuracy. If the mesh is too coarse, it might be impossible to achieve a high degree of precision without adding more elements or altering basis especially if geometry singularity occurs. This refinement is more useful with transient problems where elements move to follow an evolving phenomenon. By far the local refinement method is the most popular. It can increase the convergence rate, particularly when singularities are present. In some sense locally varying the polynomial degree as refinement is the most powerful. Exponential convergence rates are possible when solutions are smooth. When combined with a local refinement of the mesh size, these high rates are also

possible when singularities are present. The use of polynomial is most natural with a hierarchical basis, since portions of the stiffness and mass matrices and load vector will remain unchanged when increasing the polynomial degree of the basis.

In the previous chapter, we have discussed at some length various methods of recovery by which the finite element solution results could be made more accurate and this led us to devise various procedures for error estimation. In this chapter, we shall be concerned with methods which can be used to reduce the errors generally once a finite element solution has been obtained. As the process depends on previous results at all stages it is called adaptive. Such adaptive methods were first introduced to finite element calculations by Babuska and Rheinboldt in the late 1970s [Babuška & Rheinboldt, 1978]-[Babuška & Rheinboldt, 1979]. Before proceeding further it is necessary to clarify the objectives of refinement and specify *permissible error magnitudes,* and here the engineer or user must have very clear aims. For instance, the naive requirement that all displacements or all stresses should be given within a specified tolerance is not always acceptable. The reasons for this are obvious as at singularities, for example, stresses will always be infinite and therefore no finite tolerance could be specified. The same difficulty is true for displacements if point or knife edge loads are considered. The most common criterion in general engineering use is that of prescribing a total limit of the error computed in the energy norm. Often this error is required not to exceed a specified percentage of the total energy norm of the solution and in the many examples presented later we shall use this criterion. However, using a recovery type of error estimator it is possible to adaptively refine the mesh so that the accuracy of a certain quantity of interest, such as the root mean square error in displacement and/or root mean square error in stress, satisfy some user-specified criterion. Zienkiewicz and Zhu in [Zienkiewicz & Zhu, 1987] have used root mean square error in stress in the adaptive procedure to obtain more accurate stress solutions. We should recognize that adaptive procedure based on reducing the root mean square error in displacement is in effect reducing the average displacement error in each element; similarly adaptive procedure based on reducing the root mean square error in stress is the same as reducing the average stress error in each element. Here we could, for instance, specify directly the permissible error in stresses or displacements at any location. Others have used the requirement of constant energy norm density in the adaptive analysis, which is, in fact, equivalent to specifying a uniform distribution of root mean square error in stress in each element. We note that the recovery type of error estimators are particularly useful and convenient in designing adaptive analysis procedures for the quantities of interest.

5.2.1 Adapt elements or functions?

As we have already remarked in the previous chapter we will at all times consider the error in the actual finite element solution rather than the error in the recovered solution. It may indeed be possible in special problems for the error in the recovered solution to be zero, even if the error in the finite element solution itself is quite substantial. (Consider here for instance a problem with a linear stress distribution being solved by linear elements which result in constant element stresses. Obviously the element error will be quite large. But if recovered stresses are used, exact results can be obtained and no errors will exist). The problem of which of the errors to consider still needs to be answered. At the present time we shall consider the question of recovery as that of providing a very substantial margin of safety in the definition of errors.

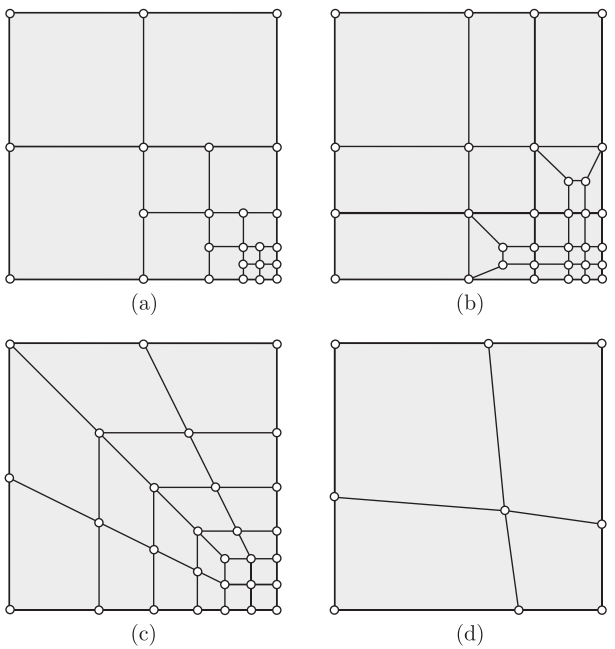


Figure 5.2: Various procedures by h-refinement: (a) mesh enrichment by transition elements; (b) mesh enrichment; (c) complete mesh regeneration; (d) repositioning node

Various procedures exist for the refinement of finite element solutions [Babuška & Miller, 1984]–[Carey, 2006]. Broadly these fall into two categories:

- The h-refinement, in which the same class of elements continue to be used but are changed in size, in some locations made larger and in others made smaller,

to provide maximum economy in reaching the desired solution.

- The p-refinement, in which we continue to use the same element size and simply increase, generally hierarchically, the order of the polynomial used in their definition.

It is occasionally useful to divide the above categories into subclasses, as the h-refinement can be applied and thought of in different ways. In Figure 5.2 we illustrate four typical methods of h-refinement.

1. The first of these h-refinement methods is element subdivision (enrichment) (Figure 5.2(a)). Here refinement can be conveniently implemented and existing elements, if they show too much error, are simply divided into smaller ones keeping the original element boundaries intact. Such a process is cumbersome as many hanging points are created where an element with mid-side nodes is joined to a linear element with no such nodes. On such occasions it is necessary to provide local constraints at the hanging points and the calculations become more involved. In addition, the implementation of de-refinement requires rather complex data management which may reduce the efficiency of the method. Nevertheless, the method of element subdivision is quite widely used.
2. The second method is the locally mesh enrichment by using transition zone (Figure 5.2(b)). Here, on the basis of a given solution, some element are subdivided by precise schemes and join together with other elements. Such a process needs sometime distorted element in the transition zone, and therefore is advise against use element sensitive to distortion.
3. The third method is that of a complete mesh regeneration or remeshing (Figure 5.2(c)). Here, on the basis of a given solution, a new element size is predicted in all the domain and a totally new mesh is generated. Thus a refinement and de-refinement are simultaneously allowed. This of course can be expensive, especially in three dimensions where mesh generation is difficult for certain types of elements, and it also presents a problem of transferring data from one mesh to another. However, the results are generally much superior and this method will be used in most of the examples shown in this chapter. For many practical engineering problems, particularly of those for which the element shape will be severely distorted during the analysis, adaptive mesh regeneration is a natural choice.
4. The final method, sometimes known as r-refinement [Figure 5.2(d)], keeps the total number of nodes constant and adjusts their position to obtain an optimal approximation. While this procedure is theoretically of interest it is difficult

to use in practice and there is little to recommend it. Further it is not a true refinement procedure as a prespecified accuracy cannot generally be reached.

We shall see that with energy norms specified as the criterion, it is a fairly simple matter to predict the element size required for a given degree of approximation. Thus very few re-solutions are generally necessary to reach the objective.

With p-refinement the situation is different. Here two subclasses exist:

1. One in which the polynomial order is increased uniformly throughout the whole domain;
2. One in which the polynomial order is increased locally using hierarchical refinement.

In neither of these cases a direct procedure, which allows the prediction of the best refinement to be used to obtain a given error, has been developed. Here the procedures generally require more resolutions and tend to be more costly. However, the convergence for a given number of variables is more rapid with p-refinement and it has much to recommend it.

On occasion it is possible to combine efficiently the h- and p-refinements and call it the hp-refinement. In this procedure both the size of elements h and their degree of polynomial p are altered. Much work has been reported in the literature by Babuska, Oden and others and it has been proved that an hp-adaptive system gives the optimal convergence (maximum accuracy for a given number of equations). However, its programming is difficult and requires careful planning of the data base structure [Babuška & Miller, 1984]–[Zienkiewicz *et al.* , 1989]–[Ainsworth & Oden, 1992]–[Ainsworth & Oden, 2000].

5.2.2 Structured or unstructured mesh?

The most basic form of mesh classification is based upon the connectivity of the mesh: structured or unstructured.

Structured mesh is one in which the elements have the topology of a regular grid. Structured meshes are typically easier to compute with (saving a constant factor in runtime) but may require more elements or worse-shaped elements.

Unstructured mesh is characterized by irregular connectivity. Unstructured meshes are often computed using quadtrees, or by Delaunay triangulation of point sets; however there are quite varied approaches for selecting the points to be triangulated. Compared to structured meshes, the storage requirements for an unstructured mesh can be substantially larger since the neighborhood connectivity must be explicitly stored.

A hybrid mesh is a mesh that contains structured portions and unstructured

portions. Note that this definition requires knowledge of how the mesh is stored (and used). There is disagreement as to the correct application of the terms "hybrid" and "mixed". The term "mixed" is usually applied to meshes that contain elements associated with structured meshes and elements associated with unstructured meshes (presumably stored in an unstructured fashion).

5.2.3 *h*-refinement on structured mesh

The quadrilateral type elements generally produce better solutions than the triangular-type, and have been frequently used in the adaptive $h-$refinement algorithm. However, when an area of complicated geometry is refined locally, the use of four node elements often only leads either to the meshes with distorted shapes or to the meshes with too many degrees of freedom not to result in an inefficient solution like those of Figure 5.3. The adaptive mesh refinements using variable-node elements have been reported to be effective [Gupta, 1978]–[Lo *et al.* , 2006].

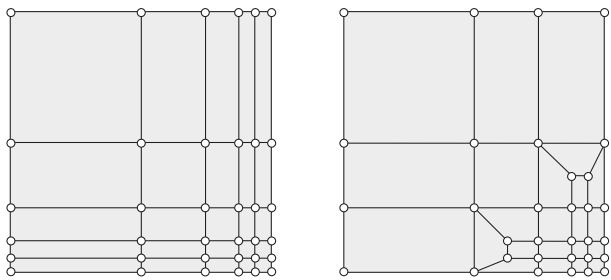


Figure 5.3: Examples of refinement without changing element type

The subdivided elements called *daughter elements* can maintain the same shapes as that of the original element, or *mother element*, when the variable-node elements are used for mesh refinement, as shown in Figure 5.4. If the well-composed initial mesh is used and the variable-node element is utilized in the transition zone, the mesh distortion will not occur or will be minimized. In evaluating the element mass and stiffness matrices of a variable-node element, a normal numerical integration may not be applied directly over the entire element domain; this is because the slope discontinuity of shape function derivative assumed in the element may cause a singular integral. Mesh refinement strategies for elliptic (steady) problems need not consider coarsening. We can refine an initially coarse mesh until the requested accuracy is obtained. This strategy might not be optimal and won't be, for example, if the coarse mesh is too fine in some regions.

5.2.3.1 Rules: quadtree balancing

The recursive binary refinement algorithm is simple. Each test point is located with respect to the current tree, as shown in Figure 5.5. The cell where the point is located is recursively divided by two until the size of the resulting cell is less than the characteristic size of the element associated with the test point. This process is repeated for each element edge segment along the curve. Each segment mid point is located in the tree resulting from the refinement due to the previous point.

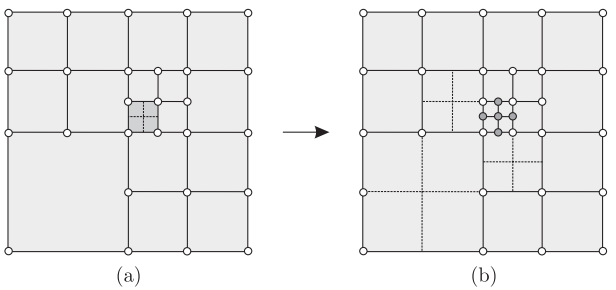


Figure 5.4: Quadtree balancing: (a) split of neighbors; (b) initial element split with new nodes and daughter-elements

In any spatial data structure, the domain is enclosed by unit squares (root) that are sub-divided into four equal elements (cells) which are the children of the root. This process can be repeated several times on each of the children until a stopping criteria is met. Two cells are adjacent if they have a common edge. Each child of the cell represents an element. A cell is called a leaf if it does not have any children. The level of a cell is the number of refinements needed to obtain that cell; the root is at level zero.

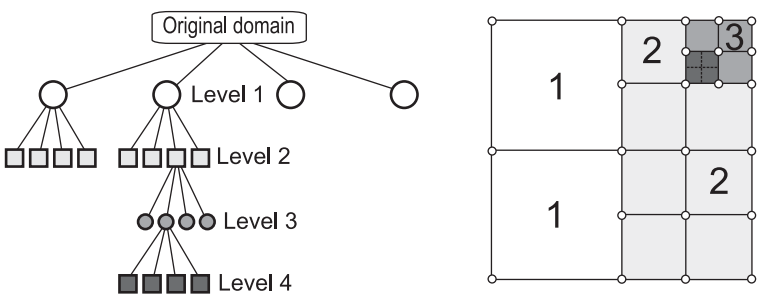


Figure 5.5: Quadtree connection: refinement by subsequent subdivision level

5.2.4 *h*-refinement on unstructured mesh

It is difficult to make general statements about unstructured mesh generation algorithms because the most prominent methods are very different in nature. The most popular family of algorithms are those based upon Delaunay triangulation, but other methods, such as quadtree/octree approaches are also used.

Many of the commonly used unstructured mesh generation techniques are based upon the properties of the Delaunay triangulation and its dual, the Voronoi diagram. Given a set of points in a plane, a Delaunay triangulation of these points is the set of triangles such that no point is inside the circumcircle of a triangle. The triangulation is unique if no three points are on the same line and no four points are on the same circle. A similar definition holds for higher dimensions, with tetrahedra replacing triangles in 3D.

5.3 Predicting the required element size in *h*-adaptivity

The error estimators discussed in the previous chapter allow the global energy (or similar) norm of the error to be determined and the errors occurring locally (at the element level) are usually also well represented. If these errors are within the limits prescribed by the analyst then clearly the work is completed. More frequently these limits are exceeded and refinement is necessary. The question which this section addresses is how best to effect this refinement [Zienkiewicz & Zhu, 1987]–[Zienkiewicz & Zhu, 1991]–[Zhu *et al.* , 1991a]. Here obviously many strategies are possible and much depends on the objectives to be achieved.

In the simplest case we shall seek, for instance, to make the relative energy norm percentage error η less than some specified value (say 5% in many engineering applications). Thus

$$\eta \leq \bar{\eta}, \tag{5.1}$$

is to be achieved.

In an *optimal mesh* it is desirable that the distribution of energy norm error (i.e., $\|\mathbf{e}\|$) should be equal for all elements. Thus if the total permissible error is determined (assuming that it is given by the result of the approximate analysis) as

$$\text{Permissible error} \equiv \bar{\eta}\,\|\mathbf{u}\| \approx \bar{\eta}\left(\left\|\mathbf{u}^h\right\|^2 + \|\mathbf{e}\|^2\right)^{1/2}, \tag{5.2}$$

here we have used

$$\|\mathbf{e}\|^2 = \|\mathbf{u}\|^2 - \left\|\mathbf{u}^h\right\|^2. \tag{5.3}$$

We could pose a requirement that the error in any element ($\cdot_e$) should be

$$\|\mathbf{e}\|_e \leq \bar{\eta} \left(\frac{\left\|\mathbf{u}^h\right\|^2 + \|\mathbf{e}\|^2}{ne} \right)^{1/2} = \bar{e}_{ne}, \tag{5.4}$$

where ne is the number of elements involved. Elements in which the above is not satisfied are obvious candidates for refinement. Thus if we define the ratio

$$\frac{\|\mathbf{e}\|_e}{\bar{e}_{ne}} = \xi_e, \tag{5.5}$$

we shall refine whenever

$$\xi_e > 1, \tag{5.6}$$

ξ_e can be approximated, of course, by replacing the true error in Equations (5.4) and (5.5) with the error estimates.

The refinement could be carried out progressively by refining only a certain number of elements in which ξ is higher than a certain limit and at each time of refining halve the size of such elements. This type of element subdivision process is also known as mesh enrichment. This process of refinement though ultimately leading to a satisfactory solution being obtained with a relatively small number of total degrees of freedom, is in general not economical as the total number of trial solutions may be excessive.

It is more efficient to try to design a completely new mesh which satisfies the requirement that

$$\xi_e \leq 1. \tag{5.7}$$

One possibility here is to invoke the asymptotic convergence rate criteria at the element level (although we have seen that these are not realistic in the presence of singularities) and to predict the element size distribution. For instance, if we assume

$$\|\mathbf{e}\|_e \propto h_e^p, \tag{5.8}$$

where h_e^p is the current element size and p the polynomial order of approximation, then to satisfy the requirement of Equation (5.4) the new generated element size should be no larger than

$$h_{new} = \xi_e^{-1/p} h_e. \tag{5.9}$$

The reason for the success of the mesh regeneration based on the simple assumption of asymptotic convergence rate implied in Equation (5.8) is the fact that with refinement the mesh tends to be *optimal* and the localized singularity influence no longer affects the overall convergence.

Of course the effects of singularity will still remain present in the elements adjacent to it and improved mesh subdivision can be obtained if in such elements

we use the appropriate convergence and write, if in Equations (5.8) and (5.9) p is replaced by λ,

$$h_{new} = \xi_e^{-1/\lambda} h_e, \tag{5.10}$$

in which λ is the singularity strength. A convenient number to use here is $\lambda = 0.5$ as most singularity parameters lie in the range $0.5 \div 1.0$. With this procedure, added to the refinement strategy, we frequently achieve accuracies better than the prescribed limit in one remeshing.

5.3.1 The $h-$refinement strategy with RCP

The refinement strategy depends on the nature of the accuracy criteria to be satisfied. A very common requirement is to specify the achievement of a certain value of the percentage error. Resorting to a recovery procedure, the percentage error in Equation (5.1) can be estimated as

$$\eta^* = \frac{\|e^*\|}{\left(\left\|\boldsymbol{\sigma}^h\right\|^2 + \|e^*\|^2\right)^{1/2}}.$$

Aiming at an equal distribution of error between all elements, requirement (5.1) can be reformulated by placing a limit on the error in each element. Following the strategy proposed in the previous section and denoting by ne the total number of elements in the current analysis, the following requirement on the permissible error in each element ($\cdot_e$) is posed:

$$\|\mathbf{e}\|_e \leq \bar{\eta} \left(\frac{\left\|\boldsymbol{\sigma}^h\right\|^2 + \|\mathbf{e}\|^2}{ne} \right)^{1/2} = \bar{e}_m. \tag{5.11}$$

Thus, the ratio

$$\xi_e = \frac{\|\mathbf{e}\|_e}{\bar{e}_m}, \tag{5.12}$$

defines the elements to be refined ($\xi_e > 1$) as well as the areas where a coarser subdivision is permissible ($\xi_e < 1$). In practice, an approximated value of ξ_e can be computed for each element by replacing the true error $\|e\|$ in Equations (5.11) and (5.12) with an error estimator $\|\mathbf{e}^*\|$. Then, the predicted values of ξ_e provide guidance for the refinement process. Here, the processes based on both mesh enrichment and complete mesh regeneration are considered. In the latter case the values ξ_e are used to predict the element size distribution. Invoking the asymptotic convergence rate criteria, the new generated element size should be no larger than

$$h_e^{new} = h_e \xi_e^{-1/p}, \tag{5.13}$$

where h_e is the current element size and p the assumed convergence rate of the error in the area covered by the element. This value is generally taken as equal

to the polynomial order of approximation, though of course it is not realistic near singularities.

5.3.2 The $hp-$refinement strategy

In an hp-adaptive solution one needs to pick which item to change first. Since p changes are relatively expensive and must be limited to integers it may be best to select p_{new}, first and to restrict the change in degree, say n to 0 or ± 1 Then, due to the integer choice on p some of the estimated refinement (or de-refinement) still needs to occur by also selecting a new mesh size. We can envision the refinement indicator as having two contributions, $\xi = \xi_p \cdot \xi_h$. If the new integer degree, $(p+n)$, was based on the current element size then the now known numerical value

$$\xi_p = \frac{h^p}{h^{(p+n)}}, \tag{5.14}$$

can be used to get the needed remaining spatial refinement indicator, ξ_h. Note that the product relation is

$$\xi = \frac{h^p}{h_{new}^{(p+n)}} = \frac{h^p}{h^{(p+n)}} \cdot \frac{h^{(p+n)}}{h_{new}^{(p+n)}} = \xi_p \cdot \xi_h, \tag{5.15}$$

which with ξ and ξ_h known simplifies to

$$\xi_h = \frac{\xi}{\xi_p} = \left(\frac{h}{h_{new}}\right)^{(p+n)}, \tag{5.16}$$

or finally

$$h_{new} = \frac{h}{\xi_h^{\frac{1}{(p+n)}}}. \tag{5.17}$$

Even with these rough estimates of desired changes one may need other rules to assure that the mesh size and local degree do not change rapidly from one solution iteration to the next, or oscillate between large and small values.

5.4 Numerical examples

To illustrate the efficiency of the RCP error estimator in adaptive analyses we present now some examples of adaptive refinement. The refinement strategies shown in the previous section that will be used are: h-refinement with transition element and h-refinement with full mesh regeneration.

5.4.1 Class I problems

The first class of problems considered refers to Class I problems (equilibrium problem of a thin membrane of uniform thickness, subjected to a distributed transverse force) described in Appendix A.1. The parameter k is set equal to 1, and the load

term b is.specified in the following.

5.4.1.1 Square domain: load case No.1

This example [Zienkiewicz & Taylor, 1989] is fairly straightforward and starts from a simple square domain in which the loading term b is selected such that the exact solution is given by

$$u(x,y) = x(1-x)y(1-y)\arctan\left[\alpha(\rho-\rho_0)\right], \tag{5.18}$$

where $\rho = (x+y)/\sqrt{2}$, $\rho_0 = 0.8$ and $\alpha = 20$. Figure 5.6 shows the exact solution in terms of stresses. As it can be observed, the exact stresses exhibit a relatively sharp concentration. In Figure 5.7 we show the first subdivision of this domain into regular linear and quadratic elements and the subsequent refinements.

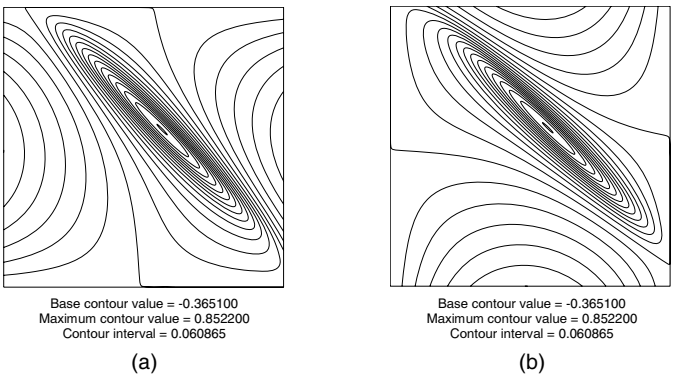
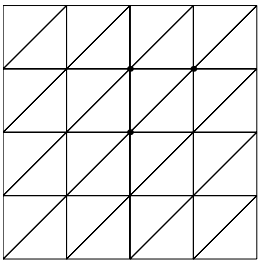


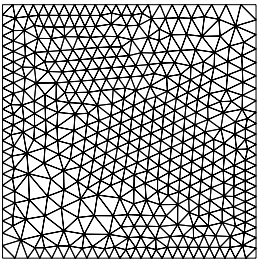
Figure 5.6: Exact solution: (a) contours of σ_x, (b) contours of σ_y

The elements are of both triangular and quadrilateral shape. For the linear elements the target error is 10% of total energy, while for quadratic elements the target error is 1% of total energy. In all cases, three refinements suffice to reach a very accurate solution satisfying the requirements despite the fact that the original mesh cannot capture in any way the high intensity region illustrated in the previous figure.

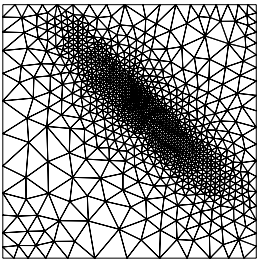
Notice that the prescribed accuracy has been achieved in all cases. Moreover, Figures 5.8-5.11 illustrate a comparison between the distributions of the exact and the estimated error in energy norm, showing a very good agreement. Finally, convergence rates of the exact and of the estimated error on the adaptive mesh refinement are shown in Figures. 5.12 and 5.13. The excellent behaviour of the RCP error estimator can be observed.



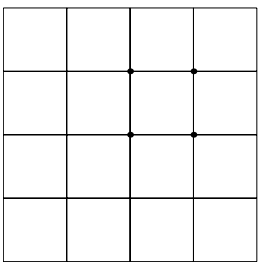
Mesh 1 – 32 elements (9 DOFs)
η = 78.71 % η^* = 65.21 %
θ = 0.778



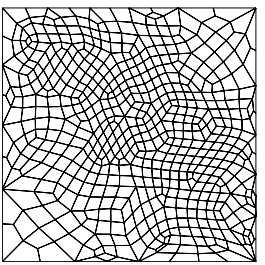
Mesh 2 – 828 elements (370 DOFs)
η = 16.85 % η^* = 17.02 %
θ = 0.010



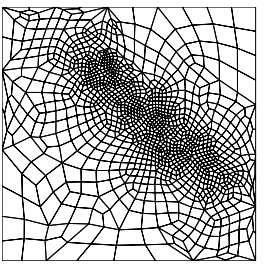
Mesh 1 – 1881 elements (911 DOFs)
η = 9.20 % η^* = 9.45 %
θ = 1.027



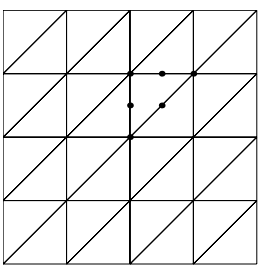
Mesh 1 – 32 elements (9 DOFs)
η = 57.48 % η^* = 55.04 %
θ = 0.938



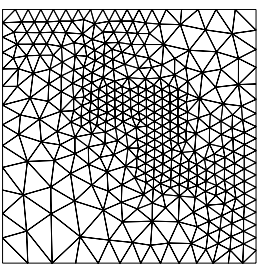
Mesh 2 – 580 elements (563 DOFs)
η = 13.52 % η^* = 13.61 %
θ = 1.007



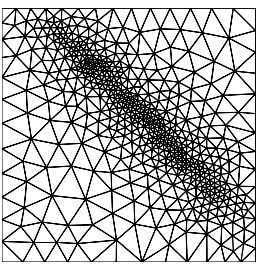
Mesh 3 – 1294 elements (1277 DOFs)
η = 7.98 % η^* = 8.11 %
θ = 1.017



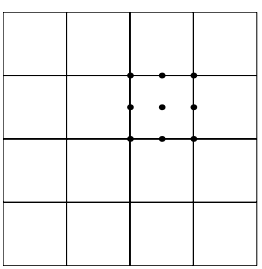
Mesh 1 – 32 elements (49 DOFs)
η = 41.11 % η^* = 43.53 %
θ = 1.124



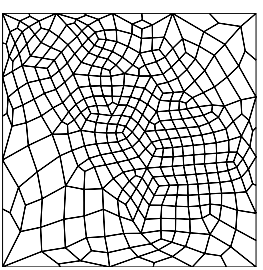
Mesh 2 – 690 elements (1319 DOFs)
η = 2.03 % η^* = 2.02 %
θ = 0.939



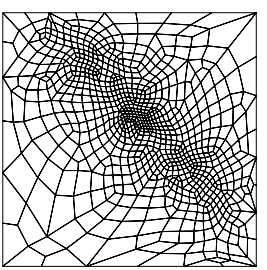
Mesh 3 – 1064 elements (2077 DOFs)
η = 0.84 % η^* = 0.85 %
θ = 1.018



Mesh 1 – 16 elements (49 DOFs)
η = 28.18 % η^* = 24.95 %
θ = 0.874



Mesh 2 – 382 elements (1501 DOFs)
η = 1.82 % η^* = 1.89 %
θ = 1.037



Mesh 3 – 876 elements (3477 DOFs)
η = 0.72 % η^* = 0.78 %
θ = 1.078

Figure 5.7: Adaptive solutions for: (a) and (b) linear elements with target error set to 10 %, (c) and (d) quadratic elements with target error set to 1 %

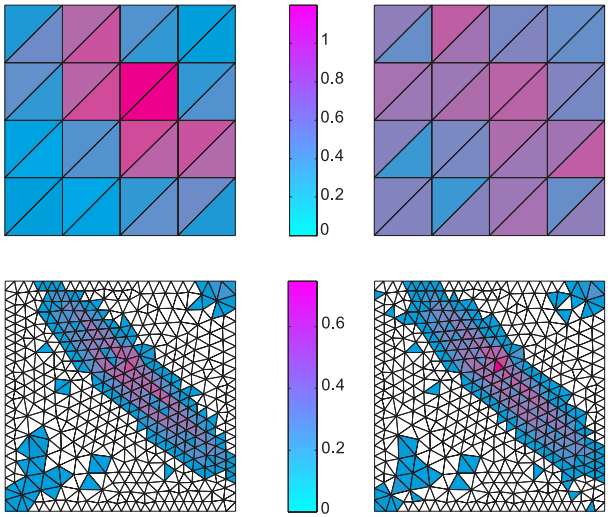


Figure 5.8: Distributions of element error in energy norm for linear triangular elements: (a) and (c) exact error, (b) and (d) error estimate

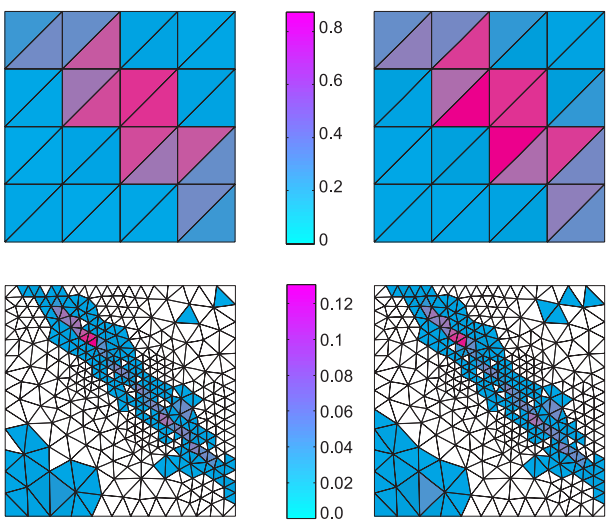


Figure 5.9: Distributions of element error in energy norm for quadratic triangular elements: (a) and (c) exact error, (b) and (d) error estimate

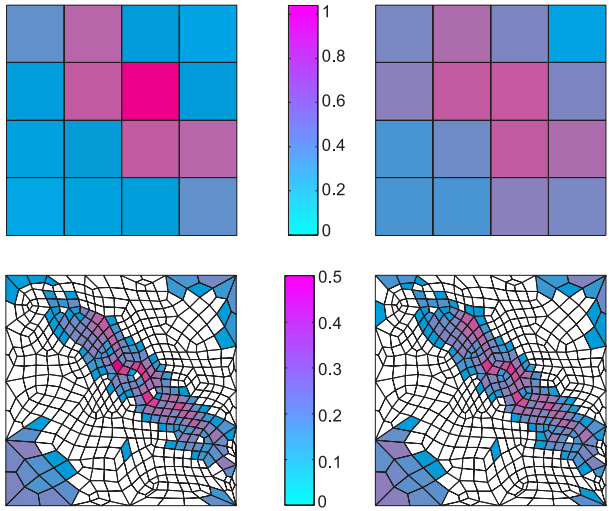


Figure 5.10: Distributions of element error in energy norm for linear quadrilateral elements: (a) and (c) exact error, (b) and (d) error estimate

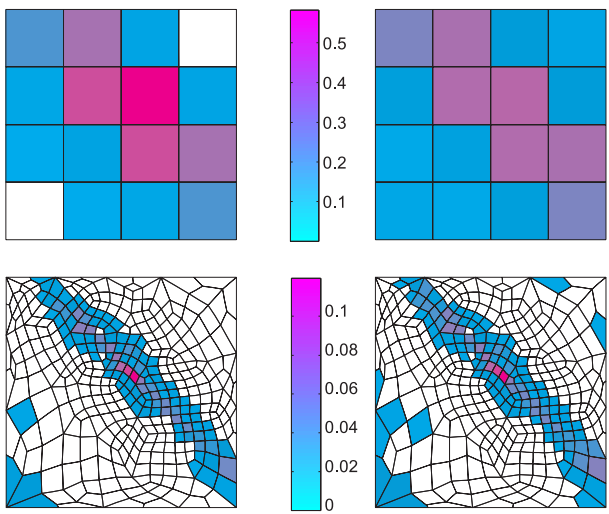


Figure 5.11: Distributions of element error in energy norm for quadratic quadrilateral elements: (a) and (c) exact error, (b) and (d) error estimate

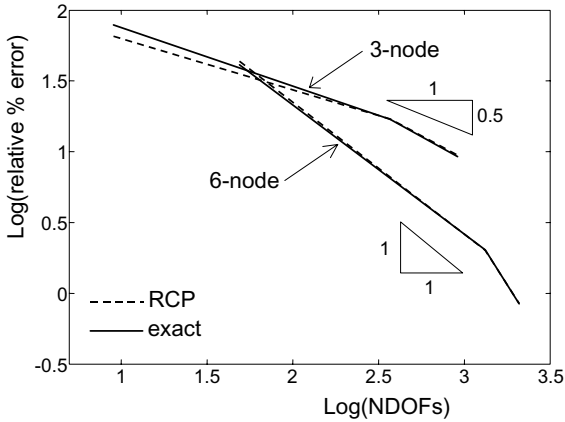


Figure 5.12: Convergence of adaptive refinement for triangular elements

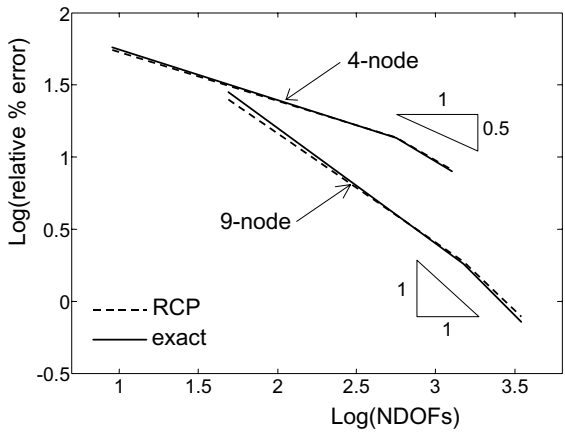


Figure 5.13: Convergence of adaptive refinement for quadrilateral elements

5.4.1.2 Square domain: load case No.2

A square domain is considered also in this example [Ainsworth & Oden, 2000], but the source term b is chosen so that the exact solution of the problem is

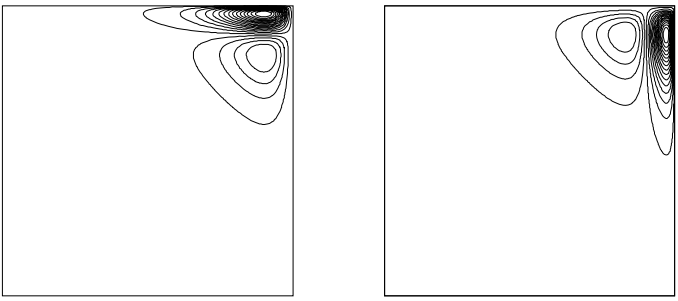


Figure 5.14: Exact solution: (a) contours of σ_x, (b) contours of σ_y

$$u(x,y) = 5x^2\left(1-x\right)^2\left(e^{10x^2}-1\right)y^2\left(1-y\right)^2\left(e^{10y^2}-1\right) \tag{5.19}$$

The solution is smooth, but possesses step gradients in the neighborhood of the boundaries as shown in Figure 5.14(a)-(b). For this reason, in this second load case the target error is higher than in the previous one: 15% of total energy for linear elements and, 5% of total energy for quadratic elements.

In Figure 5.15 we show the first subdivision of this domain into regular linear and quadratic elements and the subsequent refinements

Also in this example three refinements suffice to reach a very accurate solution satisfying the requirements, so confirming the good performance of the RCP estimator.

Moreover, Figures 5.16–5.19 compare the distribution of the exact error with that ao the estimated one, obtained both with the RCPn and the RCPe estimator.

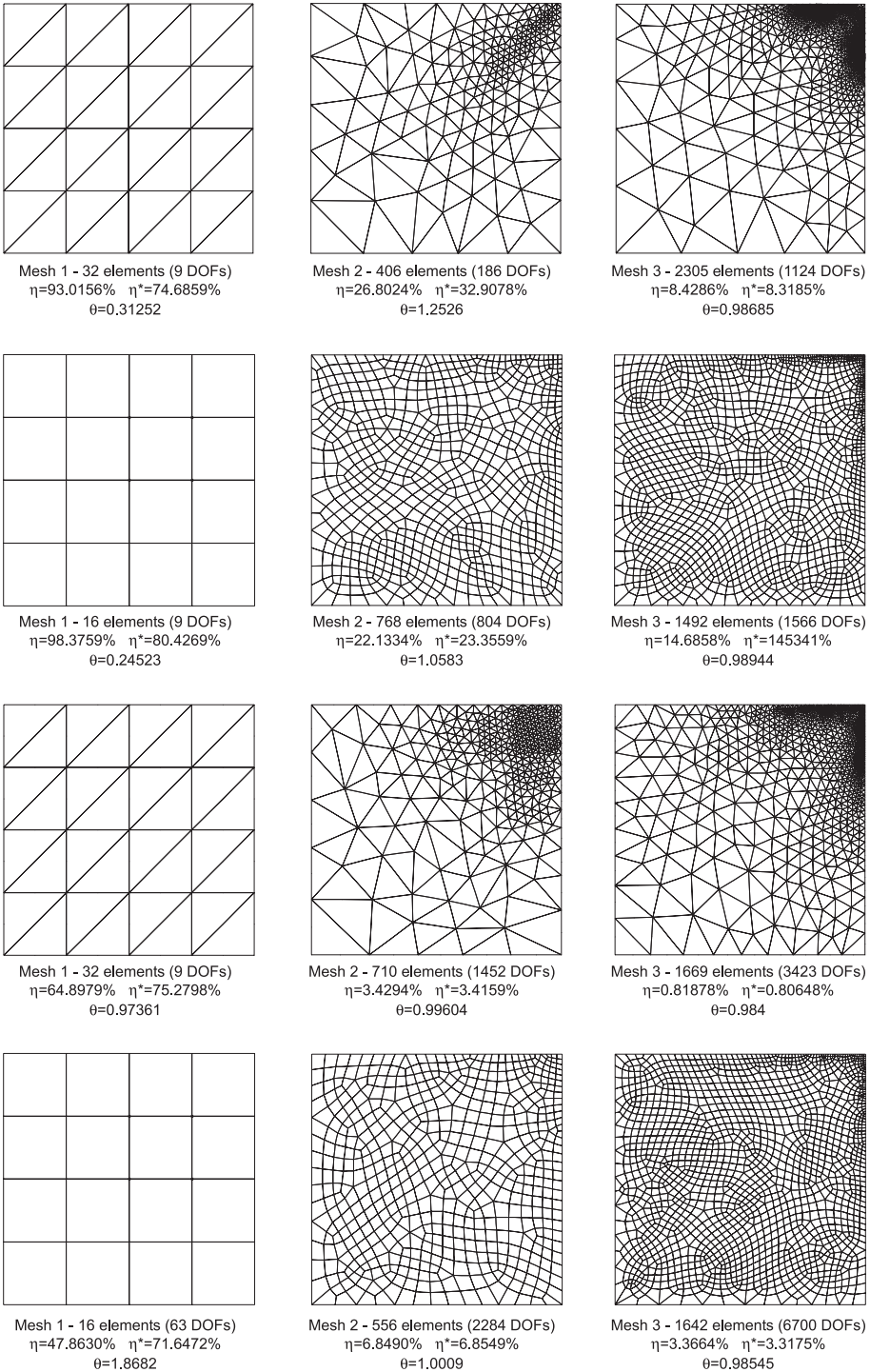


Figure 5.15: Adaptive solutions for: (a) and (b) linear elements with target error set to 15 %, (c) and (d) quadratic elements with target error set to 5 %

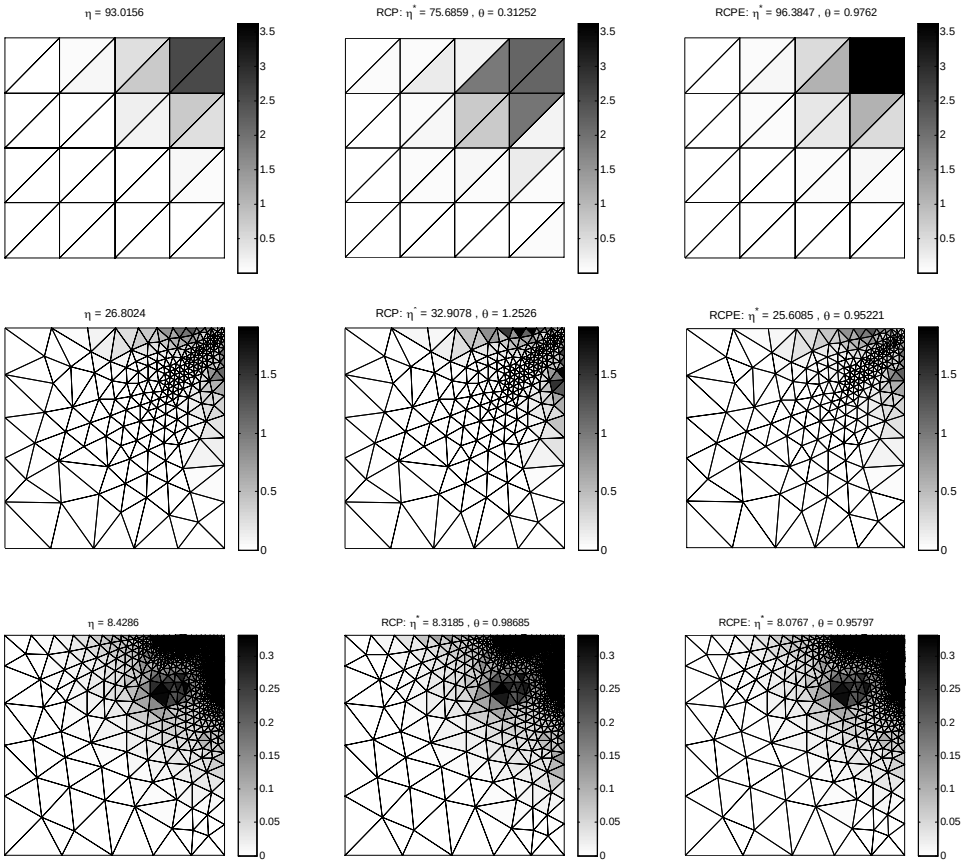


Figure 5.16: Distributions of element error in energy norm for linear triangular elements: (a), (c) and (e) exact error, (b), (d) and (f) error estimate

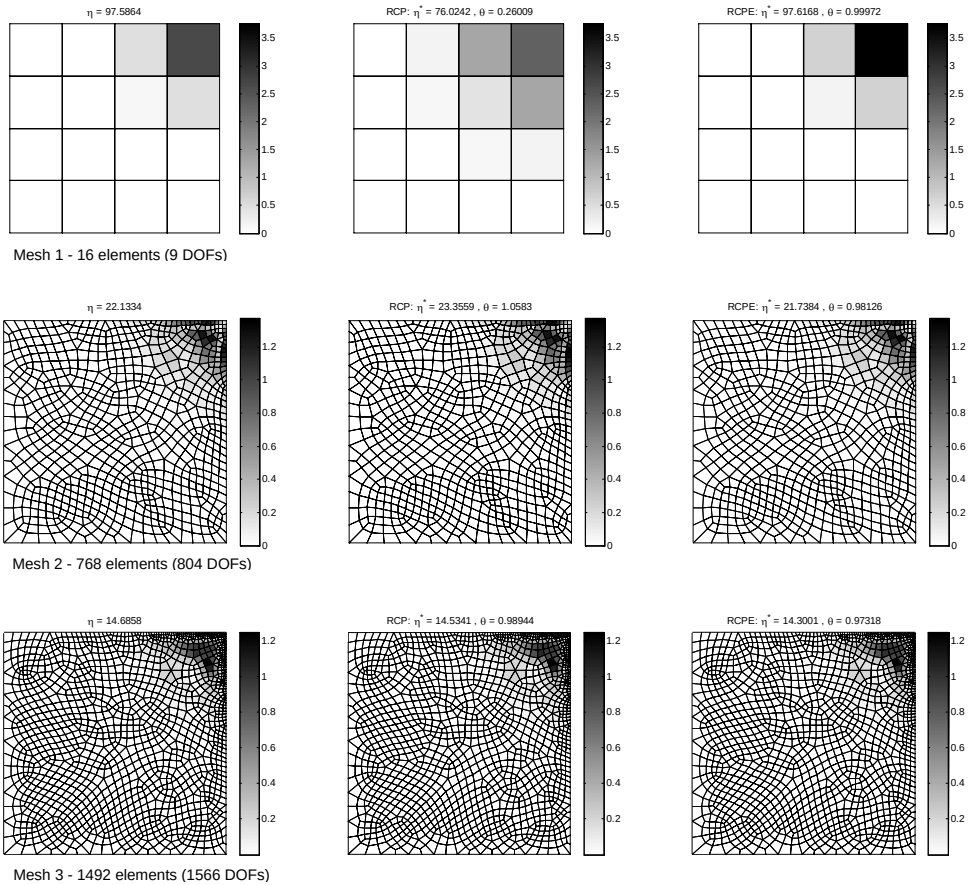


Figure 5.17: Distributions of element error in energy norm for linear quadrangular elements: (a), (c) and (e) exact error, (b), (d) and (f) error estimate

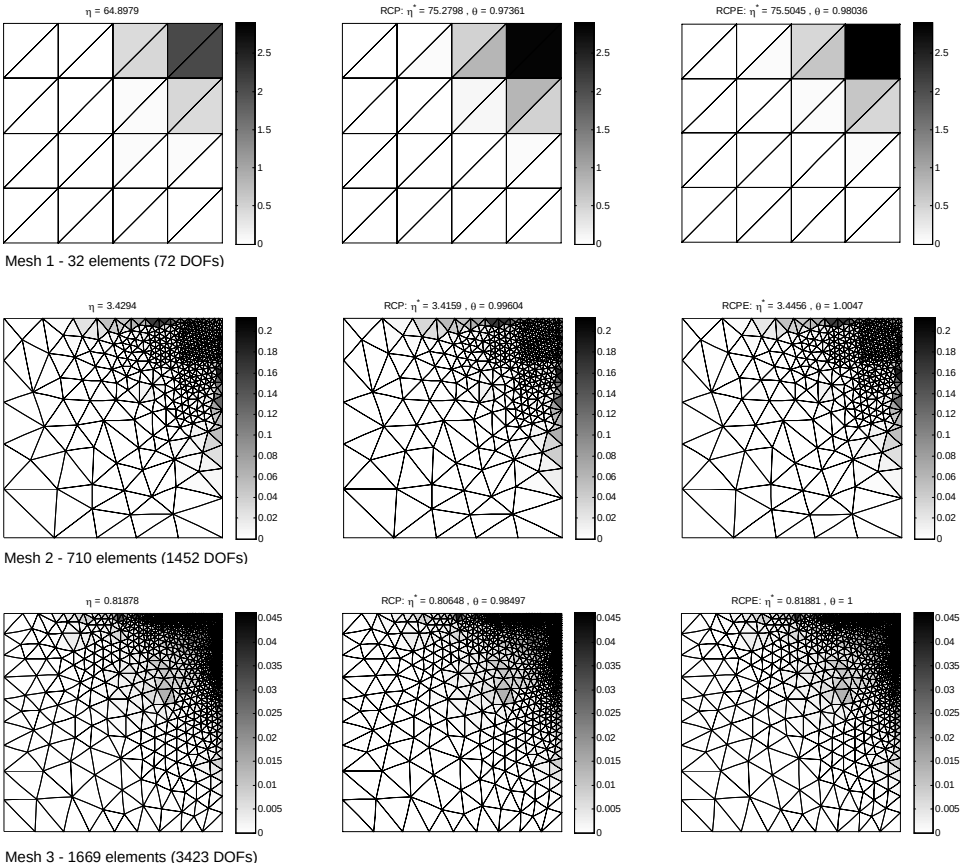


Figure 5.18: Distributions of element error in energy norm for quadratic triangular elements: (a), (c) and (e) exact error, (b), (d) and (f) error estimate

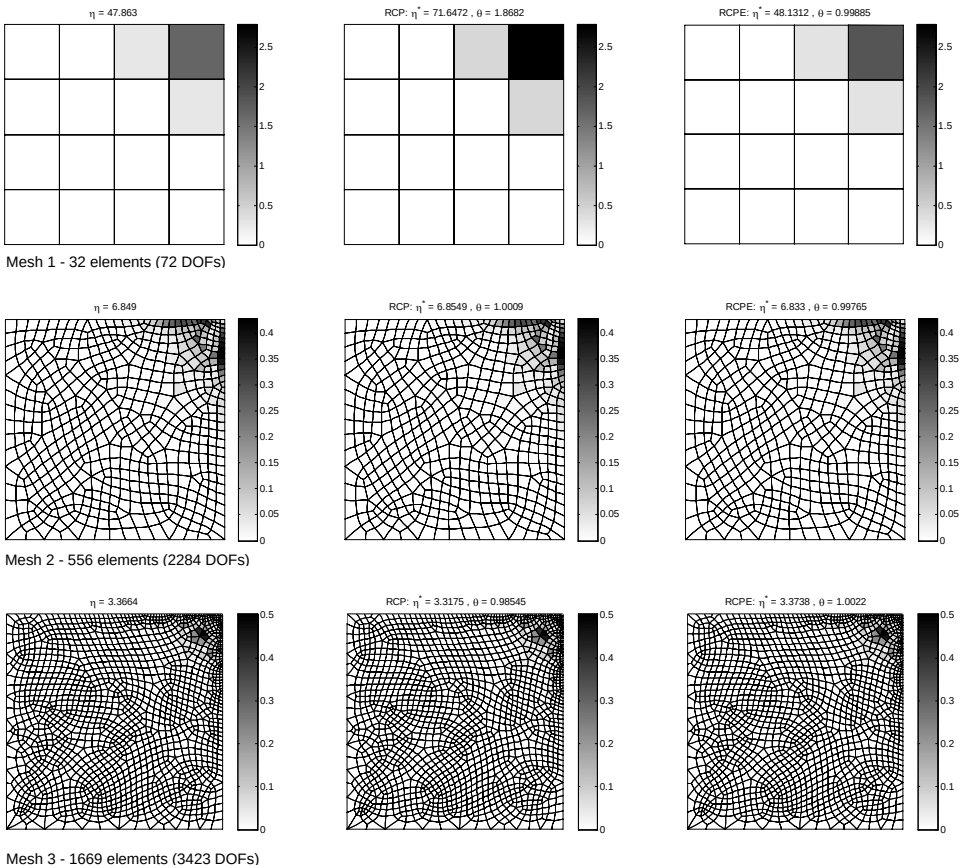


Figure 5.19: Distributions of element error in energy norm for quadratic quadrangular elements: (a), (c) and (e) exact error, (b), (d) and (f) error estimate

5.4.2 Class II problems

In this section some adaptivity examples concerning plane stress problems (Class II problems) are presented. Young modulus is set equal to 1 and Poissons ratio to 0.3.

5.4.2.1 Short cantilever with uniformly distributed load

A short cantilever subjected to uniformly distributed load is considered, Figure 5.20 [Zienkiewicz & Taylor, 1989]–[Zienkiewicz & Zhu, 1991].

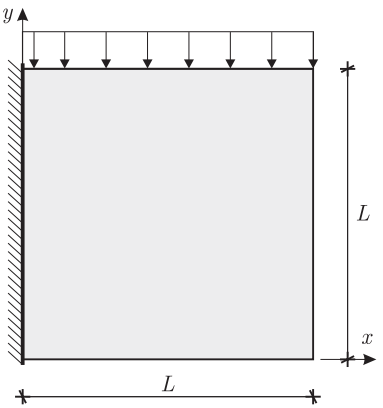


Figure 5.20: Short cantilever beam

The allowable error η is set equal to 5%. There are two singular stress points at the two corners of the edge clamped on the wall. The reference solution is obtained with a fine mesh of 64×64 elements. Adaptive analysis is performed employing mesh refinement with transition element. The details about the transition elements used in this example are given in Appendix D. As expected, in the refinement process, elements are significantly refined around the singular stress points.

5.4.2.2 L-shaped domain

This example considers a finite plate subjected to tensile loading along the orizontal direction, one quarter of which is shown in Figure 5.22. In the plate centre, there is a square cut out that introduces four singular points at the four corners of the cut out. As discussed in the previous section (see Equation (5.10)), the effects of the singularity will still be present in the elements adjacent to it and improved mesh subdivision can be obtained if, in such elements we use the appropriate singularity strength coefficient λ. As shown before most singularity parameters lie in the range $0.5 \div 1.0$. A convenient number to be use here is $\lambda = 0.5$ A complete

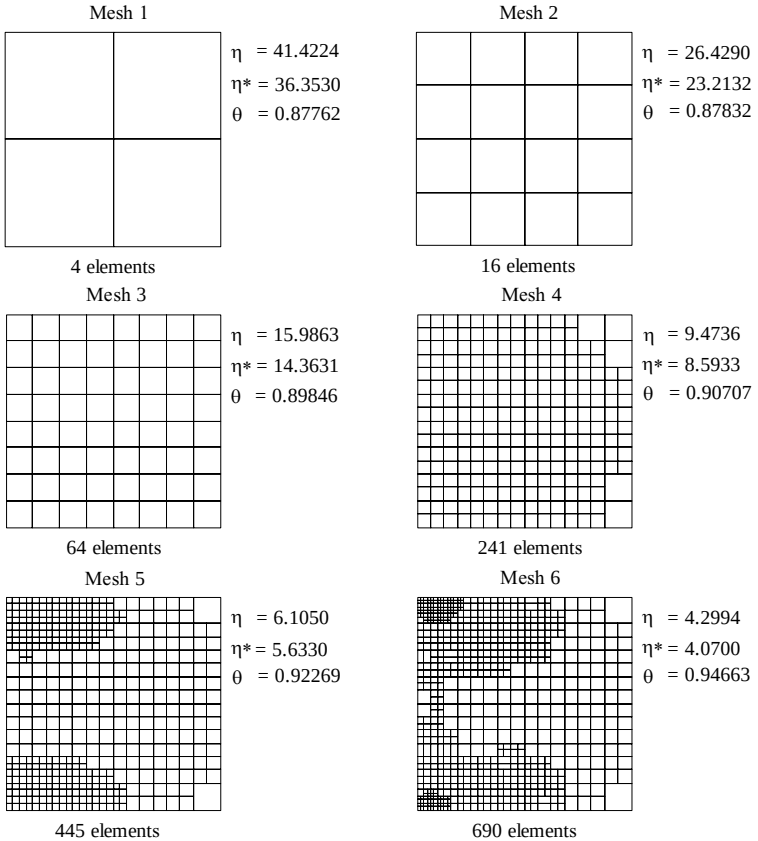


Figure 5.21: Short cantilever beam solved by mesh enrichment: linear quadrilateral elements

mesh regeneration with quadratic quadrilateral elements is used. The allowable error η is set equal to 1%. Figure 5.22 shows the progressive steps of adaptive refinement and Figure 5.23 shows the corresponding distributions of the estimated error with both RCPn and RCPe estimator. The convergence rate is given in Figure 5.24

5.4.2.3 Machine part

The machine part shown in Figure 5.25 is considered [Zienkiewicz & Taylor, 1989]–[Zienkiewicz & Zhu, 1991]. The allowable error η is set equal to 5%. Adaptive analysis is performed using complete mesh regeneration strategy. Figure 5.26 and 5.27 show the RCP-based adaptive mesh refinement with quadrangular and triangular elements, respectively. As it can be noted only two steps suffice to reach the desired accuracy.

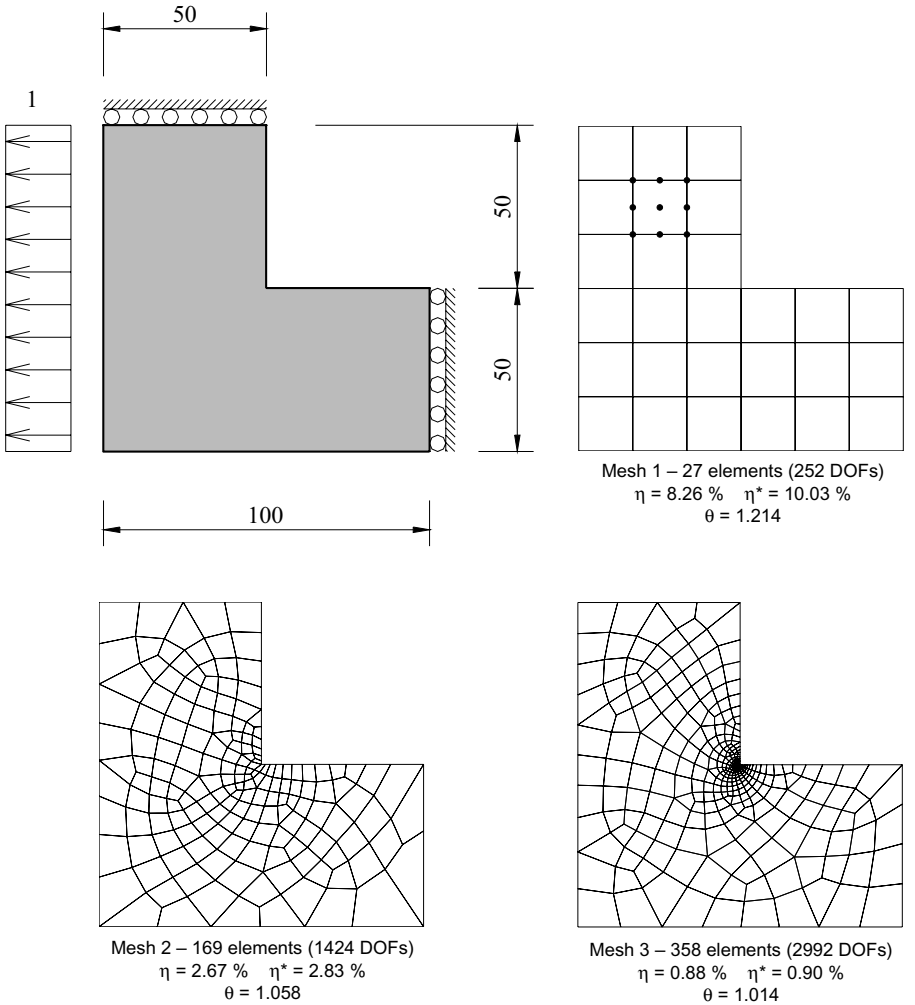


Figure 5.22: Adaptive refinement of an L-shaped domain in plane stress with prescribed error of 1%: nine-node elements

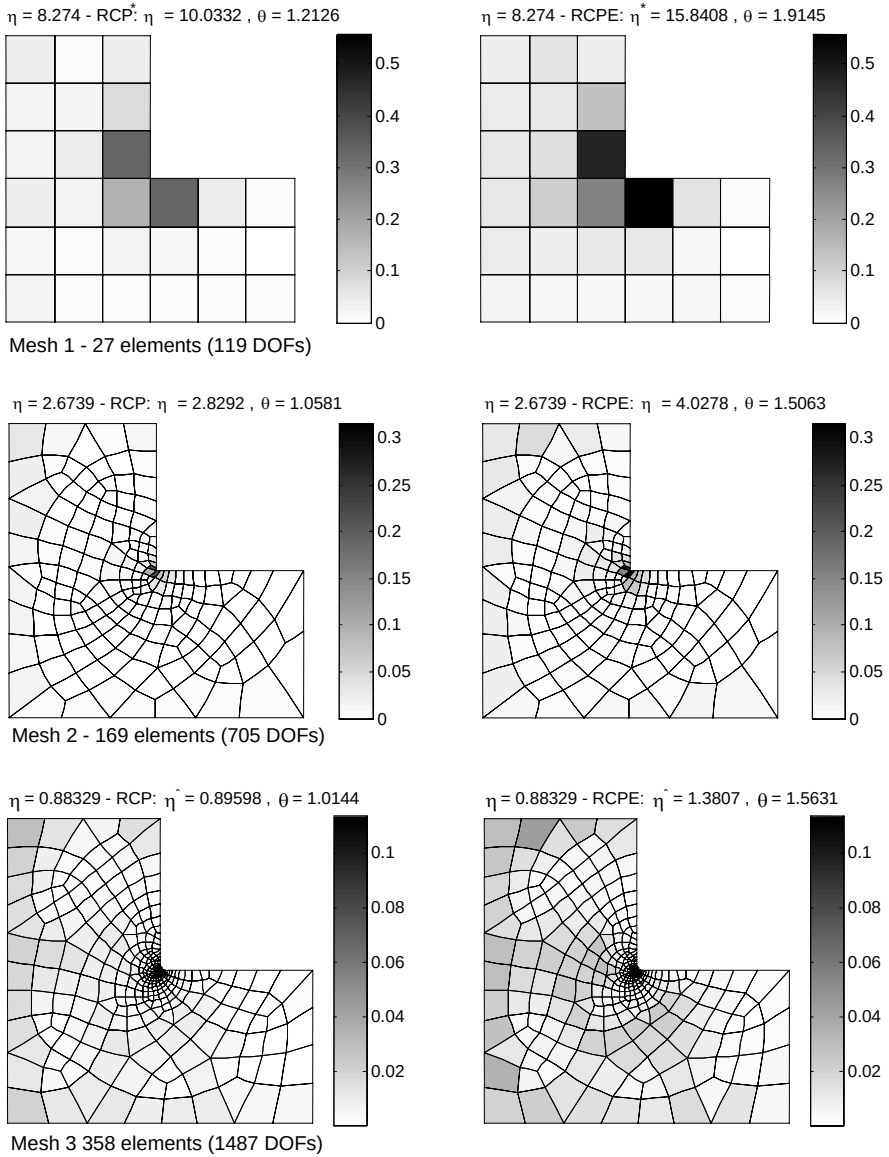


Figure 5.23: Adaptive solutions for quadratic elements with target error set equal to 2 %

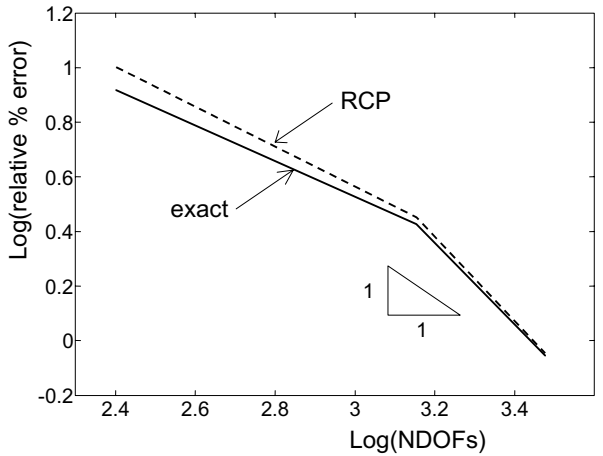


Figure 5.24: Convergence for L-shaped domain

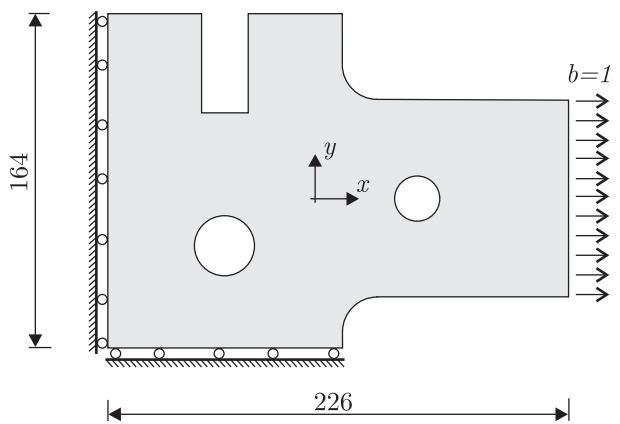
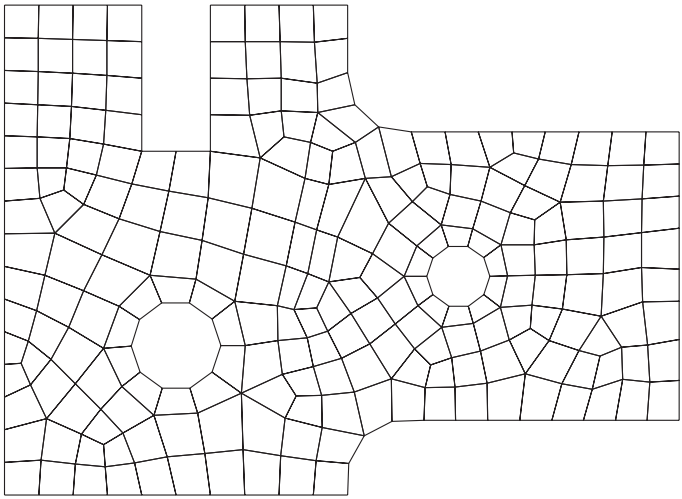
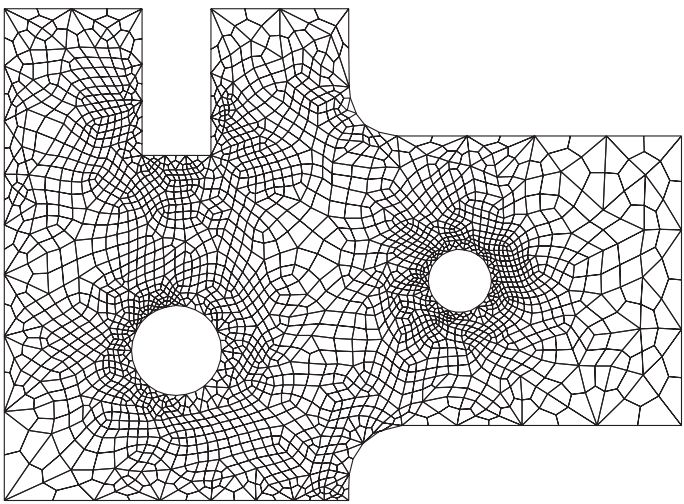


Figure 5.25: Machine part domain

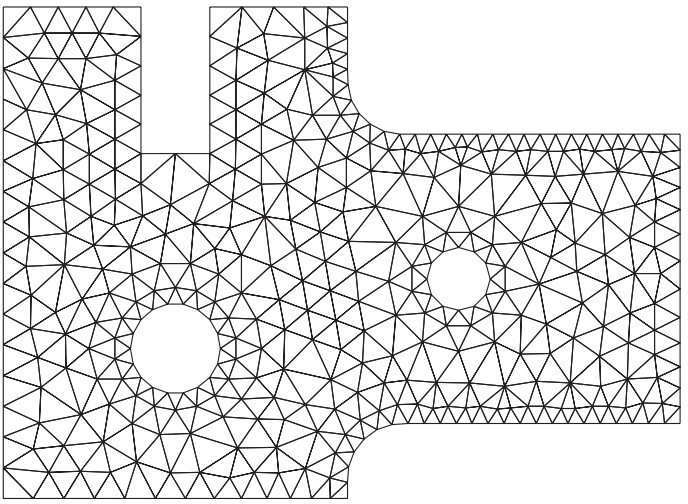


Mesh 1 - 215 elements (236 DOFs)



Mesh 2 - 2630 elements (2786 DOFs)

Figure 5.26: Adaptive refinement of machine part using linear quadrilateral elements.



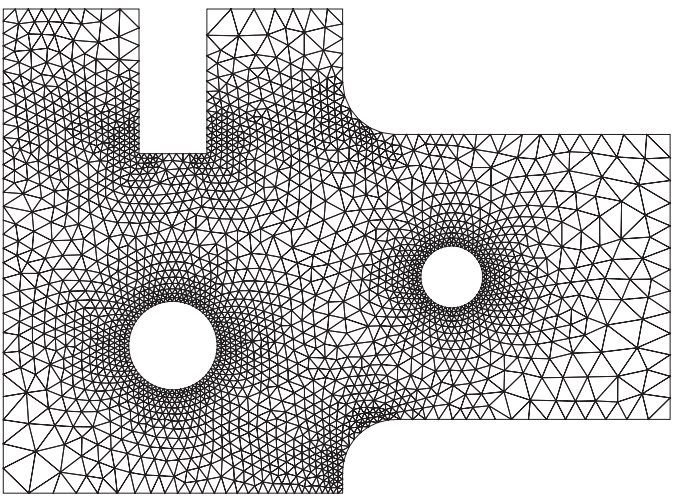


Figure 5.27: Adaptive refinement of machine part using linear triangular elements.

Chapter 6

Some advanced applications in structural mechanics

La formulazione dell'elemento piastra è un caso speciale della formulazione generale dell'elemento guscio ed è basata sulla teoria delle piastre con deformazioni trasversali taglianti incluse. Questi teoria, dovuta a E. Reissner e R. D. Mindlin, si basa sul presupposto che le particelle della piastra originalmente su una linea retta, normale alla superficie indeformata centrale rimangano su una linea retta durante il procedimento di deformazione, Tale linea non sarà necessariamente normale alla superficie centrale deformata.

Le strutture sottili come piastre e gusci costituiscono una categoria importante fra le analisi per elementi finiti a causa del loro grande campo di applicazione. Il presente capitolo punta l'attenzione sulla presentazione di un'estensione dello stimatore d'errore basato sulla procedura RCP all'analisi per elementi finiti piastre alla Reissner-Mindlin.

Come ben noto tali modelli strutturali soffrono di problemi legati alla loro formulazione quando il rapporto tra larghezza della piastra e spessore della piastra aumenta, manifestandosi in un abnorme aumento di rigidezza della piastra. Al fine di evitare tali problemi si ricorre abitualmente ad elementi di tipo misto agli sforzi, ove le variabili spostamento e tensione vengono coinvolti separatamente. Nel presente capitolo vengono inoltre presentate alcune applicazioni avanzate, quali l'impiego della tecnica proposta con elementi finiti di tipo misto agli sforzi, l'applicazione alle piastre e l'estensione alla ricostruzione del profilo delle tensioni tangenziali lungo

lo spessore di compositi laminati.

In this chapter we address some advanced problems in structural mechanics, such as structures having a "thin" dimension in one or more directions and constrained elastic problems. In particular we focus the attention on shear deformable plates structures, which are an important class of structures because of their large application fields. Indeed this application is one of the first to which finite elements were directed, but it still is a subject of continuing research [Alfano *et al.* , 2001]–[Auricchio *et al.* , 2006]. The reason lies on the well known pathology called shear locking, that is an abnormal stiffening exhibited by standard finite elements as the thickness/length ratio decreases. This locking performance is typical of constrained elastic problems such as nearly incompressible elasticity, and is more pronounced for lower order elements and in the presence of geometry distortion. A successful approach to cure this trouble is to develop non standard finite element based on mixed formulations. In the following the RCP based error estimation is extended to tackle with such advanced applications.

First the recovery procedure is extended to non standard finite element formulation????, such as assumed stress formulations. Then the recovery procedure is applied to plate structures based on Reissner-Mindlin theory [Castellazzi *et al.* , 2006a]–[Castellazzi *et al.* , 2006c]–[Castellazzi *et al.* , 2006b]. In particular, the RCP procedure is applied to a hybrid stress (locking-free) element recently proposed in [de Miranda & Ubertini, 2006] besides standard assumed displacement elements. Finally, laminated composite plates are considered, with the formulation of First order Shear Deformation Theory (FSDT), and the RCP procedure is applied to capture the transverse shear stresses along the thickness [Daghia *et al.* , 2007], which is crucial in order to predict critical phenomena such as delamination. Indeed, this last application can be also viewed as an attempt to pass from a "gross" model (FSDT plate model) to a "finer" model (3D elasticity) and, in this sense, it could be viewed as a first step towards model validation.

6.1 Recovery for assumed stress finite elements

Here, the procedures presented in the previous Chapter are applied to recover stresses from mixed and hybrid stress finite elements, based on the formulation early proposed by Pian [Yuan *et al.* , 1993]. Reference is made to the model problem discussed in Section 2.2 and the associated mixed and hybrid functionals (see Equations (2.20)–(2.21)).

Both displacements and stresses are involved as independent variables and are

independently approximated as

$$\mathbf{u}^h = \mathbf{U}\bar{\mathbf{u}}^h \tag{6.1}$$

$$\boldsymbol{\sigma}^h = \mathbf{P}\boldsymbol{\beta} \tag{6.2}$$

In particular a standard interpretation is assumed for displacements, based on standard shape functions $\mathbf{U}$ and nodal values $\mathbf{u}^h$. Identifying nodal values at common element boundaries by the usual assemblage procedure ensures the required displacement continuity. On the other hand, stress are approximated independently in each element, through a matrix of stress modes $\mathbf{P}$ and a vector of unknown parameters $\boldsymbol{\beta}$ which is local to each element.

In the hybrid version, stresses should satisfy the equilibrium equations pointwise within each element. This is obtained by the stress into two parts:

$$\boldsymbol{\sigma}^h = \boldsymbol{\sigma}_p^h + \boldsymbol{\sigma}_h^h. \tag{6.3}$$

The first is a known function equilibrating the external load within the element

$$\mathbf{D}^*\boldsymbol{\sigma}_p^h = \mathbf{b}. \tag{6.4}$$

The second is approximated according to Equation (6.2) by properly selecting self-equilibrating stress modes:

$$\mathbf{D}^*\mathbf{P} = \mathbf{0}. \tag{6.5}$$

Introducing the assumed displacements and stresses into mixed or hybrid functional referred to the single element and making it stationary yield the following discrete element equations:

$$\begin{bmatrix} -\mathbf{H} & \mathbf{G} \\ \mathbf{G}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \bar{\mathbf{u}}^h \end{bmatrix} - \begin{bmatrix} \mathbf{g}_\sigma \\ \mathbf{g}_u \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{h} \end{bmatrix}, \tag{6.6}$$

where

$$\mathbf{H} = \int_{\Omega_e} \mathbf{P}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{P}\mathrm{d}\mathbf{V}, \qquad \mathbf{G}^{\mathrm{T}} = \int_{\Omega_e} (\mathbf{D}\mathbf{U})^{\mathrm{T}}\,\mathbf{P}\mathrm{d}\mathbf{V}, \tag{6.7}$$

$\mathbf{h}$ are the element nodal forces, $\mathbf{g}_\sigma$ and $\mathbf{g}_u$ are terms due to the prescribed loads.

In the hybrid version matrix $\mathbf{G}$ reduces to

$$\mathbf{G}^{\mathrm{T}} = \int_{\Omega_e} \mathbf{U}^{\mathrm{T}}\mathbf{N}^{\mathrm{T}}\mathbf{P}\mathrm{d}\mathbf{V}. \tag{6.8}$$

At this level, inner parameters $\boldsymbol{\beta}$ can be condensed out and the following elemental equations involving only nodal displacements are obtained

$$\mathbf{h} = \mathbf{K}\bar{\mathbf{u}}^h - \mathbf{f}, \tag{6.9}$$

where $\mathbf{K}$ is the element stiffness matrix and $\mathbf{f}$ the equivalent nodal loads.

Notice that the above equations are in the standard format of assumed displacement finite elements.

All the recovery procedures presented in Section 4.3 can be applied to finite

element solutions obtained by mixed and hybrid stress elements. However, these elements offer new possibilities which deserve to be discussed. In particular, the finite element solution in terms of stresses is directly given by $\boldsymbol{\sigma}^h$ (see Equation (6.2)) or it can be computed by $\mathbf{u}^h$ as

$$\boldsymbol{\sigma}_u^h = \mathbf{C}\mathbf{D}\mathbf{u}^h = \mathbf{C}\mathbf{B}\bar{\mathbf{u}}^h. \tag{6.10}$$

As a consequence the recovery procedure can operate either on $\boldsymbol{\sigma}^h$ or on $\boldsymbol{\sigma}_u^h$.

Considering the RCP procedure and the compatibility condition it enforces (see Equation (4.62)), the following two alternatives can be derived

$$\int_{\Omega_p} \delta\boldsymbol{\sigma}^{*\mathrm{T}} \left(\epsilon_p^* - \epsilon^h\right) \mathrm{d}V = 0 \qquad \begin{cases} (1)\ \epsilon^h = \mathbf{B}\bar{\mathbf{u}}^h \\ \\ (2)\ \epsilon^h = \mathbf{C}^{-1}\boldsymbol{\sigma}^h \end{cases}, \tag{6.11}$$

depending on if we evaluate the finite element strains from displacement $\mathbf{u}^h$ or from stresses $\boldsymbol{\sigma}^h$. Both the choices can be used in principle. The first is coherent with the RCP formulation as described n Section 4.3, while the second can not be traced back to the minimization of functional (4.56), but should be formulated directly in the form of Equation (6.11), that is based on arguments related to the principle of virtual works. This does not appear as a limitation. Indeed this second choice is expected to perform better because stresses $\boldsymbol{\sigma}^h$ are more accurate than those computed from displacements. It is expected to be especially true in constrained problems. These two RCP versions are compared in Section 6.3 through some numerical tests on plate problems.

Based on the same arguments two versions of the other recovery procedures can be devised. In particular, the REP procedures can be implemented as a least square of the following conditions

$$\int_{\Omega_p} \mathbf{B}^{\mathrm{T}} \left(\boldsymbol{\sigma}_{pi}^* - \boldsymbol{\sigma}_i^*\right) \mathrm{d}V = 0 \qquad \begin{cases} (1)\ \boldsymbol{\sigma}^h \\ \\ (2)\ \boldsymbol{\sigma}^h = \mathbf{C}\mathbf{B}\bar{\mathbf{u}}^h \end{cases}. \tag{6.12}$$

These two possibilities are both tested in Section 6.3.

6.2 Plate structures

In this section the RCP recovery procedure is generalized to plate structures. In particular, shear deformable plates are considered. The corresponding generalization of the SPR and REP procedures can be found in [Boroomand *et al.* , 2004].

The governing equations for shear deformable plate structures are given in Appendix A.3. As well known standard finite elements for this class of problems suffers from shear locking in the thin plate element. In this work a four-node locking-free finite element based on hybrid stress formulation is employed for finite

element analysis [de Miranda & Ubertini, 2006].

The element is denoted by 9βQ4 and is briefly described in Appendix C. For completeness, the standard assumed displacement four node element (C4) is also included in the analysis.

The RCP recovery procedure proceeds as illustrated in Chapter 4. The two ways to form the patch are shown in Figure 6.1.

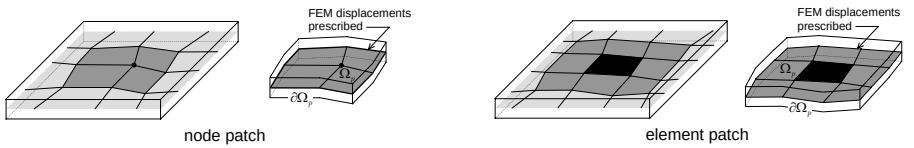


Figure 6.1: Example of node and element patch: (•) assembly node definig the patch, the dark element is the assembly element definig the patch.

The condition to be enforced on each patch specializes as

$$\int_{\Omega_p} \left[\delta\mathbf{N}^{*\mathrm{T}}\left(\boldsymbol{\mu}_p^* - \boldsymbol{\mu}^h\right) + \delta\mathbf{M}_p^{*\mathrm{T}}\left(\boldsymbol{\chi}_p^* - \boldsymbol{\chi}^h\right) + \delta\mathbf{S}_p^{*\mathrm{T}}\left(\boldsymbol{\gamma}_p^* - \boldsymbol{\gamma}^h\right)\right] \mathrm{d}V = 0, \quad \forall\ \delta\mathbf{N}^*,\ \delta\mathbf{M}^*,\ \delta\mathbf{S}^*, \tag{6.13}$$

where $(\delta\mathbf{N}^*,\ \delta\mathbf{M}^*,\ \delta\mathbf{S}^*)$ are the recovered stress resultants (over the patch) that should satisfy a priori the plate equilibrium equations, $\left(\boldsymbol{\mu}_p^*, \boldsymbol{\chi}_p^*, \boldsymbol{\gamma}_p^*\right)$ are the strain components corresponding to the recovered stress resultants and $\left(\boldsymbol{\mu}^h, \boldsymbol{\chi}^h, \boldsymbol{\gamma}^h\right)$ are the finite element strain components.

The recovered stress resultantes are expressed as

$$\begin{bmatrix} \mathbf{N}_p^* \\ \mathbf{M}_p^* \\ \mathbf{S}_p^* \end{bmatrix} = \begin{bmatrix} \mathbf{N}_{pp}^* \\ \mathbf{M}_{pp}^* \\ \mathbf{S}_{pp}^* \end{bmatrix} + \begin{bmatrix} \mathbf{N}_{ph}^* \\ \mathbf{M}_{ph}^* \\ \mathbf{S}_{ph}^* \end{bmatrix}, \tag{6.14}$$

where the first term is chosen to be in equilibrium with the first term chosen to be in equilibrium with the applied loads over the patch (i.e. as a particular solution as the plate equilibrium equations over the patch), whereas the second term is a self-equilibrated term, approximated in the form

$$\begin{bmatrix} \mathbf{N}_{ph}^* \\ \mathbf{M}_{ph}^* \\ \mathbf{S}_{ph}^* \end{bmatrix} = \mathbf{P}^*\mathbf{a}, \tag{6.15}$$

such that

$$\begin{bmatrix} \mathbf{D}_p^* & & \\ & \mathbf{D}_p^* & \mathbf{I} \\ & & \mathbf{D}_p^* \end{bmatrix} \mathbf{P}^* = \mathbf{0}. \tag{6.16}$$

The equilibrium operators in the above condition are defined in Appendix A.3.

The expressions adopted in this work for $\mathbf{P}^*$ as well as for the first term in Equation (6.14) are given in Appendix C.

The recovered strain components $\left(\boldsymbol{\mu}_p^*, \boldsymbol{\chi}_p^*, \boldsymbol{\gamma}_p^*\right)$ in Equation (6.13) are computed from the recovered stress resultants via the plate constitutive equations. Then, to apply the RCP condition (6.13) we need to specify how to compute the finite element strain components. When the recovery is used together with assumed stress elements, we have two possibilities as outlined in the previous section: they can be computed from the finite element kinematics via compatibility equations or from the finite element statics, that is from finite element stress resultants via the plate constitutive equations. Indeed, the hybrid stress element used in this work has a non standard displacement interpolation, or called linked interpolation. Thus the first version based on finite element kinematics can be implemented by computing $\left(\boldsymbol{\mu}^h, \boldsymbol{\chi}^h, \boldsymbol{\gamma}^h\right)$ either through a standard interpolation or through the true linked interpolation. All these possibilities are tested in the following. However in the thin plate limit, the version based on finite element statics is expected to be superior to the other since finite element stress resultants (that are directly interpolated) do not suffer from the appearance of internal constraints.

In the next section, the performance of RCP recovery is numerically verified. Since the membrane response has been already tested the attention is focused here on the bending response by considering homogeneous isotropic linearly elastic plates.

In Section 6.4 the coupled membrane bending response is investigated by considering laminated composite plates.

6.3 Some numerical tests on bending response

As shown in the previous chapters, two main approaches may be employed to investigate the performance of recovery methods. The first approach uses benchmark problems with exact solutions and attempts to generalize the conclusions to wider range of problems though no proof is given. Majority of studies on error estimators fall within this form of approach. In the second approach, the asymptotic behaviour is investigated. Although such a study does not play the role of a proof, the mathematical basis is much stronger and the range of application is much wider than the first approach. Unfortunately the second approach has not been yet generalized to plate bending problems, and such an extension is beyond the scope of this chapter. Hence the first approach is used in the following.

Two homogeneous isotropic linearly elastic square plates, are considered for

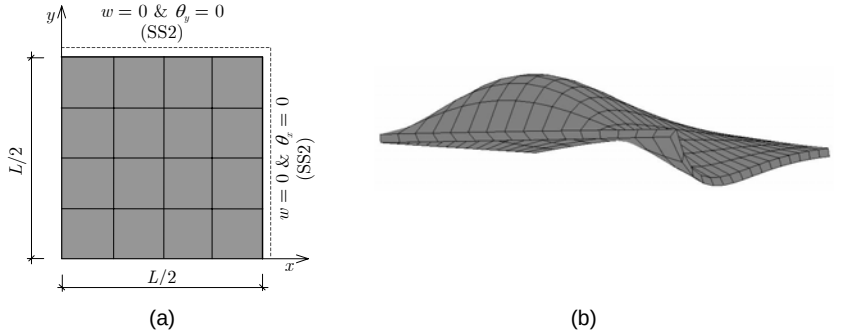


Figure 6.2: (a) Test No.1- SS2 square plate under uniform load; (b) test No.2 - square plate with prescribed displacement

various aspect ratios h/L from thick to thin cases (Figure 6.2).

The first square plate, of side length $L = 8$, is subjected to uniformly distributed transverse load $q = 1$, with simply supported boundary conditions of hard (SS2) type. This test case is referred to as Test No.1. Four meshes, created from a coarse mesh of 4×4 elements by a uniform refining process, are used in the finite element analysis, which is carried out for two values of the aspect ratio ($h/L = 0.2$ and $h/L = 0.002$) and assuming: $E = 10.92$ and $\nu = 0.3$. Conventional 4-node elements and hybrid stress 9βQ4 elements (see Appendix C) are employed. For this test some results on unstructured meshes are also discussed.

The second square plate, of side length $L = 4$, is subjected to prescribed transverse displacements along the boundary according to the following solution

$$w = \sin(x)\, e^{(-y)} y, \tag{6.17}$$

This test case is referred to as Test No.2. Four meshes, created from a coarse mesh of 4×4 elements by a uniform refining process, are used in the finite element analysis, which is carried out for an intermediate value of the aspect ratio ($h/L = 0.025$) and assuming: $E = 10.92$ and $\nu = 0.3$. Because of locking, only hybrid stress elements are employed for this test.

Depending on the thickness of the plate, the exact solution for the first benchmark test is given with assumption of thick or thin plate formulation. Fourier series can be used for both cases expressing the plate deflection as follows

$$w = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} C_{mn} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b}, \tag{6.18}$$

with a and b being the plate dimensions and C_{mn} representing a series of coefficients to be determined from the solution process. For thin plate solution, the deflection

Table 6.1: Versions of RCP recovery tested and related approximations

	REP	$\text{RCP}_{\text{linear}}$	$\text{RCP}_{\text{quadratic}}$
Q4 & 9βQ4	linear stress expansion $n_{\mathbf{a}} = 3+3$	quadratic stress expansion $n_{\mathbf{a}} = 9$	cubic stress expansion $n_{\mathbf{a}} = 17$

assumed above is directly substituted in to the governing differential equation derived from Kirchhoff assumptions. The solution is straightforward. For thick plate solution, however, some similar expressions for θ_x and θ_y are needed

$$\theta_x = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn}\cos\frac{m\pi x}{a}\sin\frac{n\pi y}{b}, \tag{6.19}$$

$$\theta_y = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} B_{mn}\sin\frac{m\pi x}{a}\cos\frac{n\pi y}{b}. \tag{6.20}$$

The solution may be found in Reference [Dobyns, 1981].

In the following the RCP recovery is tested by using different versions as summarized in Table 6.1 and compared with REP recovery.

The comparisons are carried out based on the following norms

Total energy norm:

$$\|\mathbf{u}\|_t = \left(\|\mathbf{u}\|_b^2 + \|\mathbf{u}\|_s^2\right)^{1/2}, \tag{6.21}$$

Energy norm of bending stresses

$$\|\mathbf{u}\|_b = \left[\frac{1}{2}\int_\Omega \mathbf{M}^{\mathrm{T}}\mathbf{C}_b^{-1}\mathbf{M}\,\mathrm{d}\Omega\right]^{1/2}, \tag{6.22}$$

Energy norm of shear stresses

$$\|\mathbf{u}\|_s = \left[\frac{1}{2}\int_\Omega \mathbf{S}^{\mathrm{T}}\mathbf{C}_s^{-1}\mathbf{S}\,\mathrm{d}\Omega\right]^{1/2}, \tag{6.23}$$

The associated error definitions read as

$$\|e^*\|_t = \left(\|e^*\|_b^2 + \|e^*\|_s^2\right)^{1/2}, \tag{6.24}$$

$$\|e^*\|_b = \left[\frac{1}{2}\int_\Omega \left(\mathbf{M}^h - \mathbf{M}^*\right)^{\mathrm{T}}\mathbf{C}_b^{-1}\left(\mathbf{M}^h - \mathbf{M}^*\right)\mathrm{d}\Omega\right]^{1/2}, \tag{6.25}$$

$$\|e^*\|_s = \left[\frac{1}{2}\int_\Omega \left(\mathbf{S}^h - \mathbf{S}^*\right)^{\mathrm{T}}\mathbf{C}_s^{-1}\left(\mathbf{S}^h - \mathbf{S}^*\right)\mathrm{d}\Omega\right]^{1/2}, \tag{6.26}$$

$$\eta_t^* = \frac{\|e^*\|_t}{\|\mathbf{u}\|_t}\times 100, \qquad \eta_b^* = \frac{\|e^*\|_b}{\|\mathbf{u}\|_b}\times 100, \qquad \eta_s^* = \frac{\|e^*\|_s}{\|\mathbf{u}\|_s}\times 100. \tag{6.27}$$

The results are presented in terms of convergence of both global error and point-

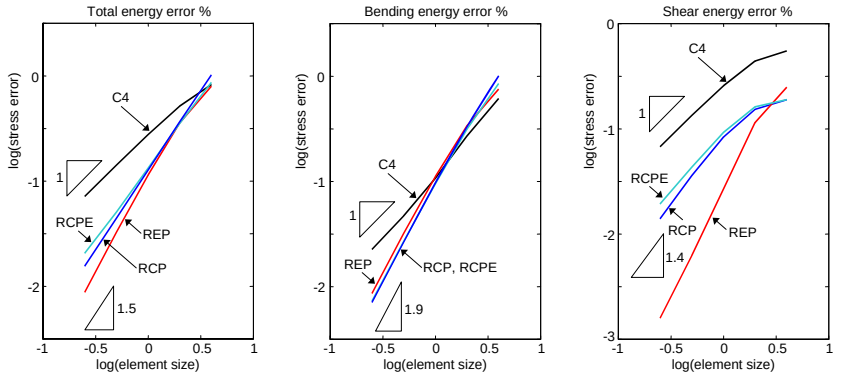


Figure 6.3: Test No.1 ($h/L = 0.2$) - convergence of global errors for recovery procedures based on linear stress approximations

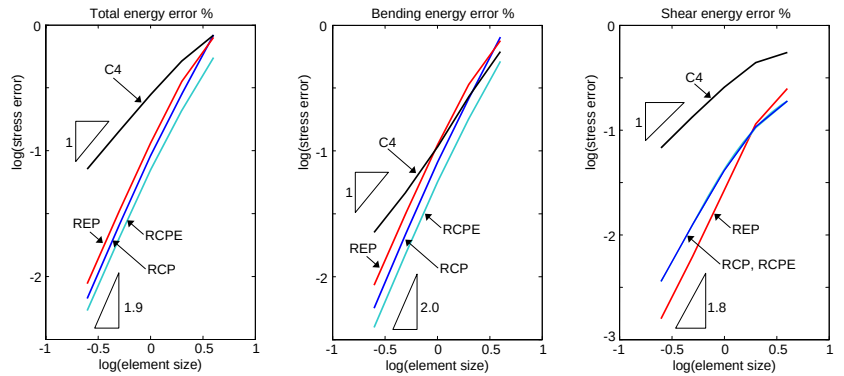


Figure 6.4: Test No.1 ($h/L = 0.2$) - convergence of global errors for recovery procedures based on quadratic stress approximations

wise error at point A located at the centre of the plate, from thick to thin cases. Convergence of the global errors in energy norm, for the thick case ($h/L = 0.2$), is shown in Figures 6.3 for the linear stress approximation and in Figure 6.4 for the quadratic stress approximation. These figures demonstrate the performance of RCP and RCPE when compatible C4 elements are used. As expected, the average rate of convergence of FEM solution is close to unity while those of the recovered solutions are higher than one, so evidencing a superconvergent behaviour similar

to that observed for two-dimensional elastic problems. Notice that the convergence rates for the various recovery methods are similar and tend to two when quadratic stress approximation is adopted.

The same case has been solved using hybrid stress elements instead of conventional C4 elements to obtain the finite element solution, and the corresponding results are shown in Figures 6.19–6.6 for linear and quadratic tress approximations, respectively. These results are similar to those observed in the previous analysis and no appreciable differences are noted among the various versions of the recovery procedures. In particular, RCP, $\mathrm{RCP_C}$ and $\mathrm{RCP_L}$ all give similar responses.

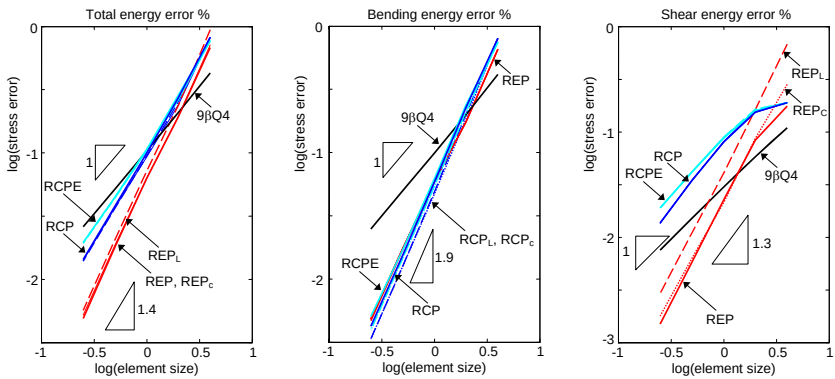


Figure 6.5: Test No.1 ($h/L = 0.2$) - convergence of global errors for recovery procedures based on Linear stress approximations

Passing to a thinner plate ($h/L = 0.025$) some curious responses are observed. The results obtained (using hybrid stress elements) in terms of both global and local errors are shown by Figures 6.7–6.10, for linear and quadratic stress approximations. The values of percentage point-wise errors in stress resultants in the plate centre are collected in Tables 6.2–6.6.

As it can be noted, the responses of recovery procedures based on FEM displacements (RCP, $\mathrm{RCPE_C}$, $\mathrm{RCPE_L}$ and $\mathrm{RCP_L}$) drastically deteriorates due to the poor performance on shear stresses. This becomes even more pronounced by increasing the order of stress approximation and can be interpreted as a sort of locking which destroyes the superconvergent properties of the recovery procedure in the thin plate limit. The reason lies on the fact that the recovery procedure is applied to the kinematic part of the finite element solution, which suffers from the internal constraint of thin plate limit. In fact the RCP procedures based on FEM stresses

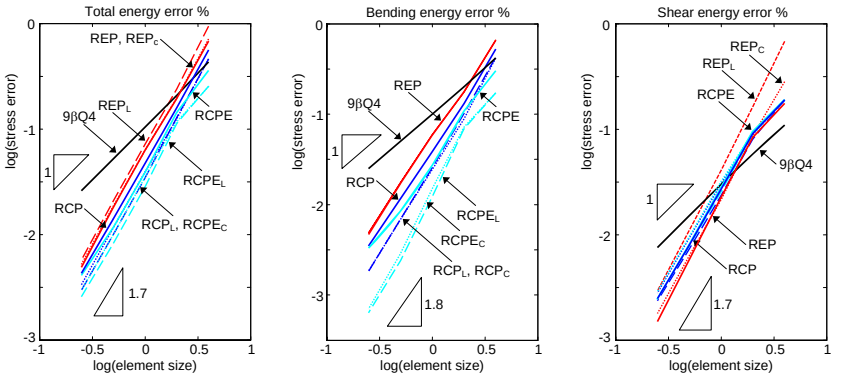


Figure 6.6: Test No.1 ($h/L = 0.2$) - convergence of global errors for recovery procedures based on quadratic stress approximations

(RCP and RCPE) do not exhibit such an analogous behaviour and maintain the superconvergent properties. This behaviour is confirmed by the predictions for the very thin plate, shown in Figure 6.11. As it can be observed, superconvergence is experienced using both RCP and RCPE.

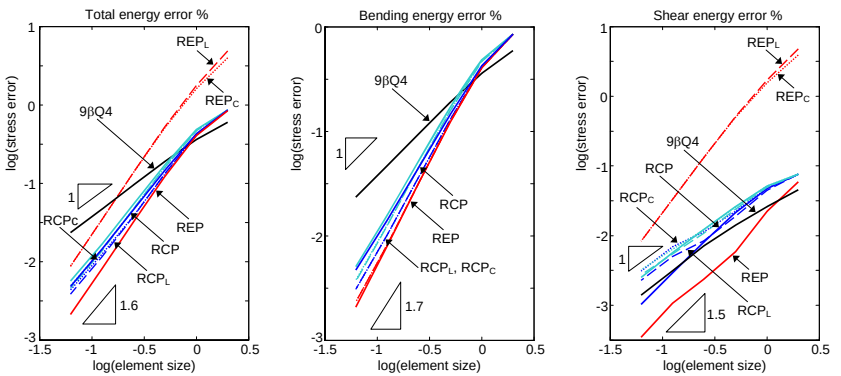


Figure 6.7: Test No.1 ($h/L = 0.025$) - convergence of global errors for recovery procedures based on linear stress approximations

For the very thin plate, that has been recognized as the most sever test problem, the local error distributions on the 8×8 uniform mesh are given in Figures **??**–**??**. In particular, Figure **??** refers to the element error in energy norm whereas Figures **??**–**??** to the pointwise errors in M_x, M_{xy} and S_x. As expected, the recovery procedures sensibility reduce the error also locally within the domain, while appear less effective near boundaries where small patches are used. Moreover,

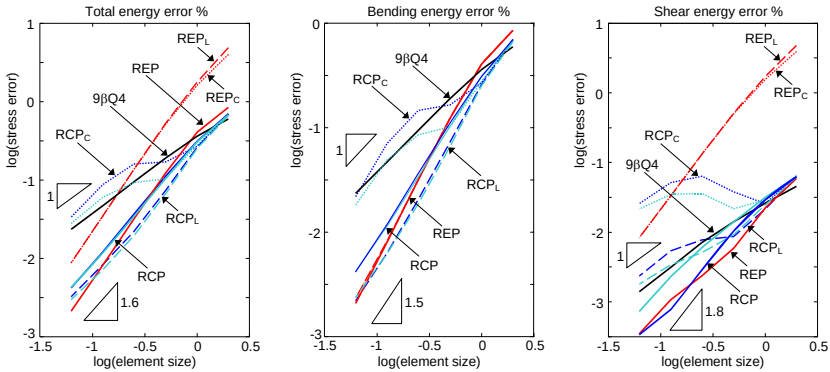


Figure 6.8: Test No.1 ($h/L = 0.025$) - convergence of global errors for recovery procedures based on quadratic stress approximations

no significant differences in accuracy emerge among the various procedures, so confirming the very good performance of the RCP procedure expecially in the element version.

Finally, the sensitivity to mesh distortion has been tested. For the sake of brevity, only the results obtained for the very thin plate (Test No.1) are reported using an unstructured mesh. In particular, Figures 6.16–6.18 show the pointwise errors in the stress resultants. These results confirm those experienced on regular geometry and demonstrates that the RCP procedure is reliable also for arbitrary irregulars meshes.

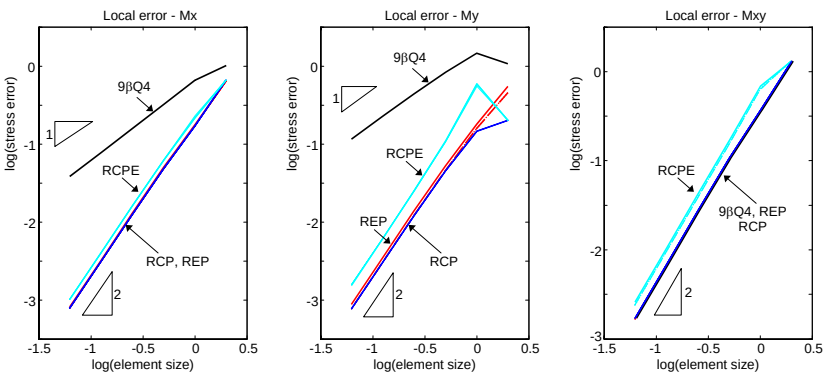


Figure 6.9: Test No.1 ($h/L = 0.025$) - convergence of local error at the plate centre for recovery procedures based on linear stress approximations

Summing up, the numerical experimentation evidences that the RCP procedure

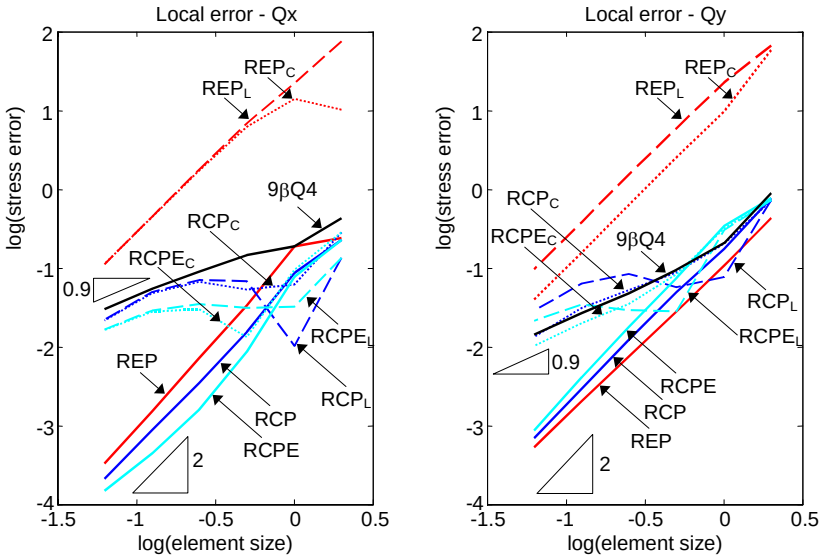


Figure 6.10: Test No.1 ($h/L = 0.025$) - convergence of local error at the plate centre for recovery procedures based on quadratic stress approximations

Table 6.2: Test No. 2 (h/L=0.025): Point-wise errors (per cent) of M_x at point A ,9βQ4.

mesh	FEM	REP	RCP	RCPE	REP_C	REP_L	RCP_C	RCP_L	RCPE_C	RCPE_L
4×4	102.2980	63.6449	65.8533	65.8533	65.4537	63.6449	65.8533	65.8533	65.8533	65.8533
8×8	65.9772	17.1101	16.9315	23.1429	17.8086	17.1101	16.9315	16.9315	21.6229	21.6269
16×16	32.2545	4.8034	4.8050	6.1959	5.0017	4.8034	4.8050	4.8050	6.0956	6.0984
32×32	15.7180	1.2325	1.2331	1.5951	1.2840	1.2325	1.2331	1.2331	1.5864	1.5889
64×64	7.7118	0.3106	0.3107	0.4036	0.3236	0.3106	0.3107	0.3107	0.4012	0.4032

is reliable and effective with either node or element patches, and that the linear stress approximation offers the best trade-off between accuracy and computational cost, especially for thin plates. Using quadratic stress approximations tends to increase accuracy in any case but improves also the rate os convergence only for thick plates.

Table 6.3: Test No. 2 (h/L=0.025): Point-wise errors (per cent) of M_y at point A, 9βQ4.

mesh	FEM	REP	RCP	RCPE	REP_C	REP_L	RCP_C	RCP_L	RCPE_C	RCPE_L
4×4	107.6726	45.5386	20.0198	20.0198	55.0367	45.5386	20.0198	20.0198	20.0198	20.0198
8×8	145.2619	15.8789	14.6587	54.7693	17.9765	15.8789	14.6587	14.6587	58.9709	58.9732
16×16	82.9392	4.5577	4.4799	10.5333	5.2186	4.5577	4.4799	4.4799	10.8004	10.8072
32×32	44.0788	1.2092	1.2031	2.4882	1.3800	1.2092	1.2031	1.2031	2.4986	2.5061
64×64	22.7161	0.3067	0.3061	0.6182	0.3497	0.3067	0.3061	0.3061	0.6130	0.6194

Table 6.4: Test No. 2 (h/L=0.025): Point-wise errors (per cent) of M_{xy} at point A, 9βQ4.

mesh	FEM	REP	RCP	RCPE	REP_C	REP_L	RCP_C	RCP_L	RCPE_C	RCPE_L
4×4	130.9319	128.3576	130.9319	130.9319	130.9319	128.3576	130.9319	130.9319	130.9319	130.9319
8×8	36.9052	35.6116	36.9052	64.2015	36.9052	35.6116	36.9052	36.9052	68.5926	61.9399
16×16	10.5748	10.2368	10.5748	15.3272	10.5748	10.2368	10.5748	10.5748	16.4785	14.6672
32×32	2.7141	2.6275	2.7141	3.8286	2.7141	2.6275	2.7141	2.7141	4.1263	3.6687
64×64	0.6830	0.6612	0.6830	0.9557	0.6830	0.6612	0.6830	0.6830	1.0330	0.9230

Table 6.5: Test No. 2 (h/L=0.025): Point-wise errors (per cent) of Q_x at point A, 9βQ4.

mesh	FEM	REP	RCP	RCPE	REP_C	REP_L	RCP_C	RCP_L	RCPE_C	RCPE_L
4×4	42.6506	1032.3483	28.5764	28.5764	23.9424	7585.5042	23.0386	13.5889	23.0386	13.5889
8×8	18.8570	1412.1877	6.2487	9.5095	19.0370	2225.5636	8.7652	1.0312	7.8487	3.2804
16×16	14.6230	623.2831	5.3557	1.3339	3.3935	692.7196	1.5386	6.8012	0.8631	3.1556
32×32	8.9459	174.4279	6.6641	3.0204	0.7210	179.6036	0.3516	7.0531	0.1598	3.4948
64×64	5.4583	44.8515	4.8443	2.7780	0.1520	45.2808	0.0880	4.9525	0.0447	2.8996

Table 6.6: Test No. 2 (h/L=0.025): Point-wise errors (per cent) of Q_y at point A, 9βQ4.

mesh	FEM	REP	RCP	RCPE	REP_C	REP_L	RCP_C	RCP_L	RCPE_C	RCPE_L
4×4	90.5749	5867.8725	72.6527	72.6527	43.4051	6692.6620	77.3863	73.7673	77.3863	73.7673
8×8	20.6936	971.0851	20.3957	31.6741	10.8521	2265.1508	17.7092	7.8062	34.1705	30.8678
16×16	9.5253	253.7822	8.8973	8.3888	2.8788	605.2156	4.9511	5.8076	7.5439	2.8718
32×32	4.8065	63.5138	5.2463	3.5083	0.7642	153.5090	1.2319	8.5986	1.7873	2.9443
64×64	2.7021	15.8915	3.1256	2.0107	0.2068	38.5098	0.2943	6.4601	0.4108	3.5097

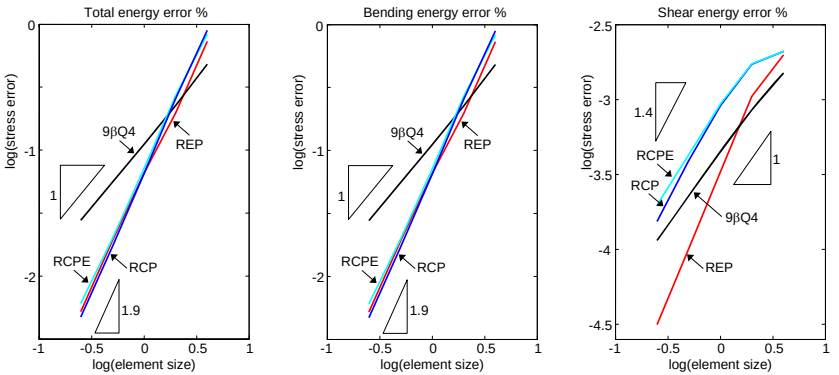


Figure 6.11: Test No.1 ($h/L = 0.002$) - convergence of global errors for recovery procedures based on linear stress approximations

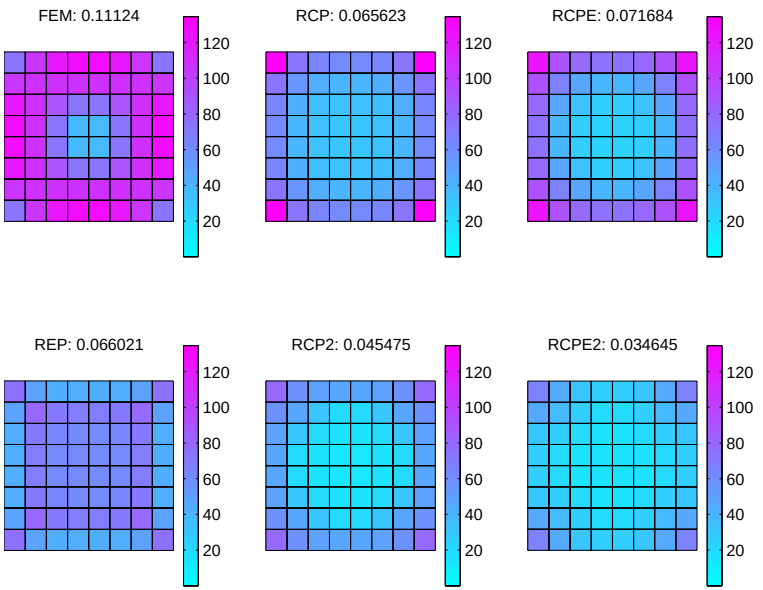


Figure 6.12: Test No.1 ($h/L = 0.002$) Distributions of element error in energy norm on uniform mesh wuth 9βQ4 elements (the relative percentage error is reported above each plot)

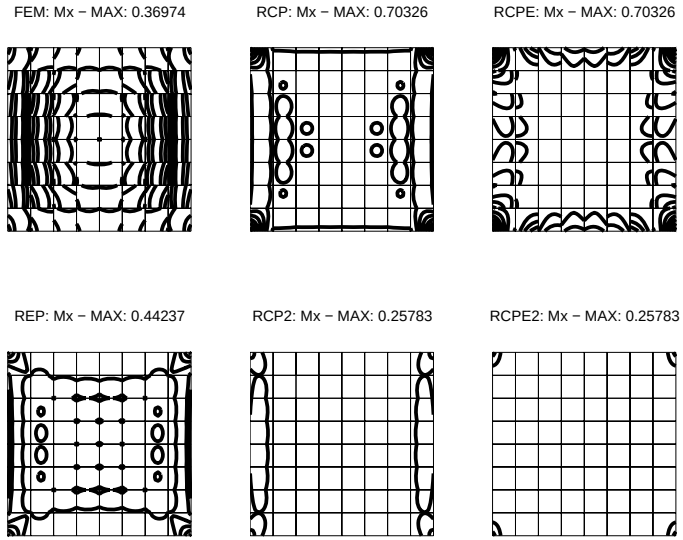


Figure 6.13: Test No.1 ($h/L = 0.002$) Distributions of poinwise errors in M_x on uniform mesh wuth 9βQ4 elements (contour interval =0.1)

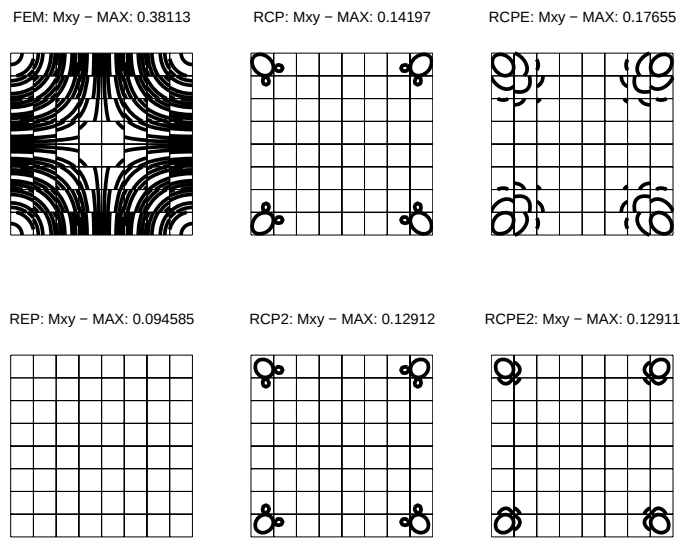


Figure 6.14: Test No.1 ($h/L = 0.002$) Distributions of poinwise errors in M_{xy} on uniform mesh wuth 9βQ4 elements (contour interval =0.1)

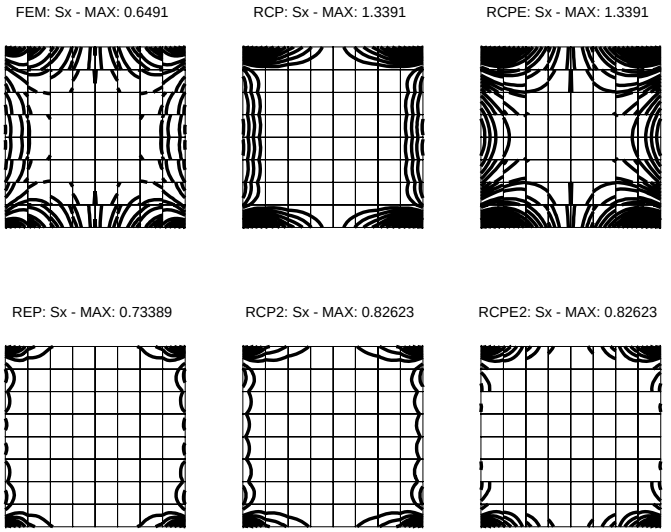


Figure 6.15: Test No.1 ($h/L = 0.002$) Distributions of poinwise errors in S_x on uniform mesh wuth 9βQ4 elements (contour interval =0.1)

Figure 6.16: Test No.1 ($h/L = 0.002$) Distributions of poinwise errors in M_x on uniform mesh wuth 9βQ4 elements (contour interval =0.2)

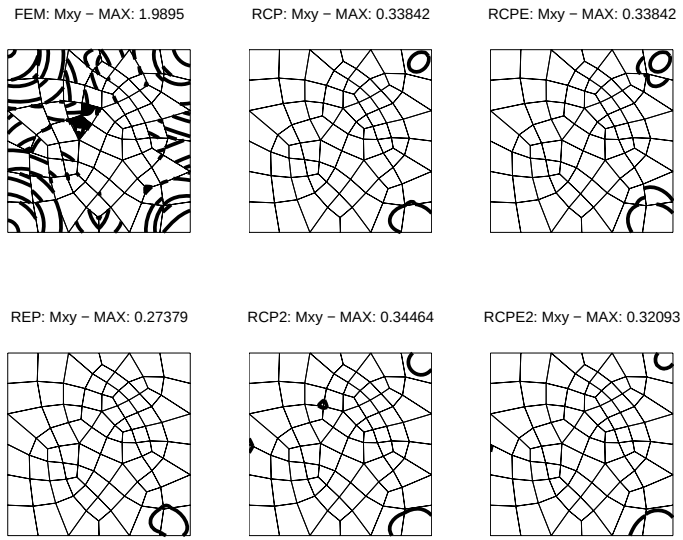


Figure 6.17: Test No.1 ($h/L = 0.002$) Distributions of poinwise errors in M_{xy} on uniform mesh wuth 9βQ4 elements (contour interval =0.2)

Figure 6.18: Test No.1 ($h/L = 0.002$) Distributions of poinwise errors in S_x on uniform mesh wuth 9βQ4 elements (contour interval =0.2)

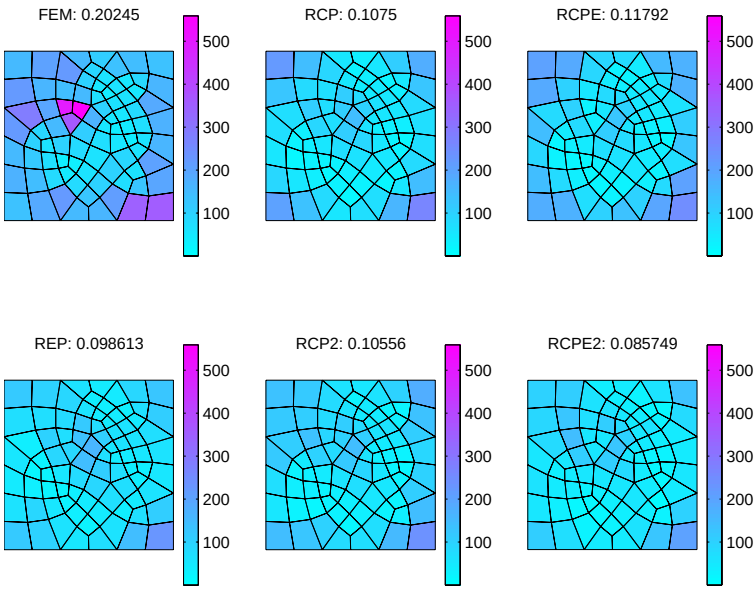


Figure 6.19: Test No.1 ($h/L = 0.002$) Distributions of element error in energy norm on uniform mesh wuth 9βQ4 elements (the relative percentage error is reported above each plot)

6.4 Laminated composite plates

This section deals with laminated composite plates based on First order Shear Deformation Theory, Figure 6.20.

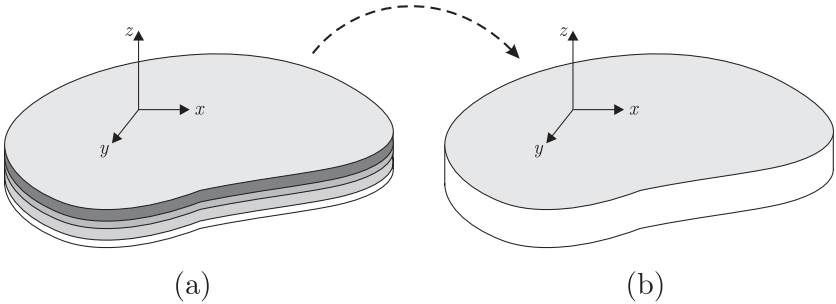


Figure 6.20: First order Shear Deformation Theory (FSDT) scheme for a laminated composite plate

The governing equations are briefly recalled in Appendix A, which can be derived from 3D elasticity as show in Appendix A.3.2. Using the same notation, the stress-strain relation for the kth layer can be written as

$$\mathbf{S}^{(k)} = \mathbf{C}_p^{(k)}\mathbf{e}, \qquad \boldsymbol{\tau}^{(k)} = \boldsymbol{\kappa} \odot \mathbf{C}_s^{(k)}\boldsymbol{\gamma}, \tag{6.28}$$

where $\mathbf{C}_p^{(k)}$ is the reduced in-plane elasticity matrix and $\mathbf{C}_s^{(k)}$ is the transverse shear elastic matrix. Thus, following Appendix A.3.2, the plate constitutive equations can be computed by integrating the local constitutive equations (6.28) through the thickness:

$$\mathbf{N} = \mathbf{C}_m\boldsymbol{\mu} + \mathbf{C}_{mb}\boldsymbol{\chi}, \tag{6.29}$$

$$\mathbf{M} = \mathbf{C}_{mb}\boldsymbol{\mu} + \mathbf{C}_b\boldsymbol{\chi}, \tag{6.30}$$

$$\mathbf{S} = \mathbf{C}_s\boldsymbol{\gamma}, \tag{6.31}$$

where the laminate stiffness matrices are given by

$$\mathbf{C}_m = \sum_{k}^{n.layers} \left(z_k - z_{k-1}\right)\mathbf{C}_p^{(k)}, \tag{6.32}$$

$$\mathbf{C}_{mb} = \frac{1}{2}\sum_{k}^{n.layers} \left(z_k^2 - z_{k-1}^2\right)\mathbf{C}_p^{(k)}, \tag{6.33}$$

$$\mathbf{C}_b = \frac{1}{3}\sum_{k}^{n.layers} \left(z_k^3 - z_{k-1}^3\right)\mathbf{C}_p^{(k)}, \tag{6.34}$$

$$\mathbf{C}_s = \kappa \odot \sum_{k}^{n.layers} (z_k - z_{k-1})\, \mathbf{C}_s^{(k)}. \tag{6.35}$$

The two vertical abscissas z_{k-1} and z_k define the bottom and top surfaces of the kth layer. Notice that bending and extension are generally coupled.

In this work, laminated composite plates are analyzed using the hybrid stress element called 9βQ4 (Appendix C) [de Miranda & Ubertini, 2006].

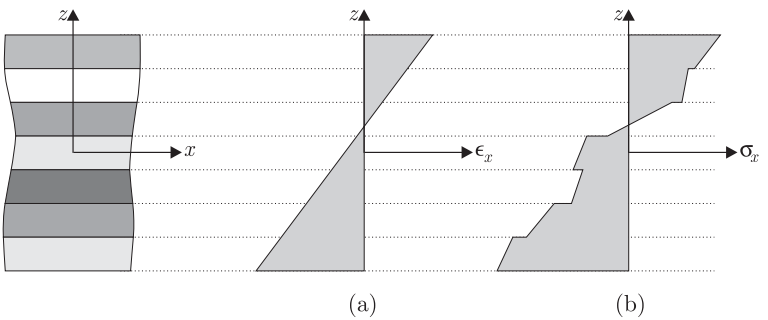


Figure 6.21: Variations of strains and stresses through layer and laminated thicknesses. (a) Variation of typical in-plane strain. (b) Variation of corresponding stress

The RCP recovery discussed in the previous section can be successfully applied to this class of problems, as shown by the numerical tests. Indeed, with this framework, the RCP recovery is found to be an effective tool in order to face a key issue for laminates: the evaluation of transverse shear stresses along the thickness of the laminate.

The kinematic hypothesis introduced in the first order shear deformation theory yields, obviously, inaccurate piecewise constant shear stresses which are discontinuous between the laminae. However, the specialistic literature suggests that accurate transverse shear stress profiles can be evaluated in the post-processing through three dimensional equilibrium equations (see Appendix A.3.2)

$$\mathbf{D}_p^* \mathbf{s}^{(k)} + \nabla_z \boldsymbol{\tau}^{(k)} + \mathbf{b}_x = \mathbf{0}. \tag{6.36}$$

From the FEM internal moments $\mathbf{M}^h$ and membrane forces $\mathbf{N}^h$ it is possible to evaluate the curvatures $\boldsymbol{\chi}^h$ and the membrane strains $\boldsymbol{\mu}^h$ through the laminate constitutive equation and, so, the in plane strains $\mathbf{e}^h$ (see Appendix A.3.2). The constitutive equations of each lamina allow, finally, to determine the in plane stresses $\mathbf{s}^{(k)}$, which can be introduced in the three dimensional equilibrium equations to

evaluate the transverse shear stresses:

$$\boldsymbol{\tau}_z^{(k)} = \boldsymbol{\tau}_0^{(k)} - \int_{z_{k-1}}^{z_k} \left(\mathbf{D}_p^* \mathbf{s}^{(k)} + \mathbf{b}_x \right) \mathrm{d}z, \tag{6.37}$$

where $\boldsymbol{\tau}_0^{(k)}$ represents the value of the shear stress vector at $z = z_{k-1}$ i.e.

$$\boldsymbol{\tau}_0^{(k)} = \boldsymbol{\tau}^{(k-1)} \left(z_{k-1} \right), \tag{6.38}$$

with

$$\boldsymbol{\tau}_0^{(1)} = -\mathbf{p}_x^{(-)} \tag{6.39}$$

The shear stresses so obtained should satisfy the equilibrium boundary conditions on the top surface of the laminate:

$$\boldsymbol{\tau}_0^{(n)} \left(z_n \right) = \mathbf{p}_x^{(+)}, \tag{6.40}$$

being n the number of layers. Moreover, in order to have consistency between the recovered shear stresses and the FSDT model, the shear stress resultants along the plate thickness should be equal to the transverse shear forces evaluated via finite element procedure:

$$\mathbf{S}^k = \int_f \boldsymbol{\tau} \mathrm{d}z = \mathbf{0} \tag{6.41}$$

It can be shown that both conditions are automatically satisfied if the internal moments and membrane forces satisfy the plate equilibrium equations.

Indeed, accuracy of the stress profiles depends upon the derivatives of the internal moments and membrane forces, therefore a good approximation not only of the internal forces but also of their derivatives is important. According to this, a two steps procedure is proposed to recover transverse shear stresses. First, RCP procedure is applied in order to improve the rate of convergence of the FEM internal forces. Then, using the recovered internal forces, the transverse shear stress profiles are determined as described above. In particular, RCP is applied using element patches so to preserve local equilibrium. In this way, transverse shear stresses can be actually recovered by simply appling Equation (6.37): conditions (6.40) and (6.41) were be automatically satisfied.

In the next section the effectiveness of RCP and the transverse shear stress recovery has been verified on various test problems.

6.5 Some numerical tests

In this section, some numerical tests are carried out in order to investigate the RCP performance and the transverse shear stress recovery procedure.

A simply supported square plate of side L (Figure 6.22) is considered under both uniform and sinusoidal load with maximum intensity q. The side to thickness ratio

of the laminate is $\eta = L/h = 10$. Two different stacking sequences are analyzed: a 0/90/0 symmetric laminate and a 0/90 antisymmetric one. Due to the double symmetry, only one quarter of the plate is studied. Some regular and distorted meshes used in the analyses are shown in Figure 6.23.

The lamina mechanical properties are the following:

$$E_1 = 25E_2, \qquad \nu_{21} = 0.25, \qquad G_{13} = G_{12} = 0.5E_2, \qquad G_{23} = 0.2E_2. \tag{6.42}$$

The shear correction factors are evaluated according to cylindrical bending as shown in [Vlachoutsis, 1992]–[Laitinen *et al.* , 1995]. The reference solution used for the convergence analysis and in the tables is obtained according to Reddy [Reddy, 1997].

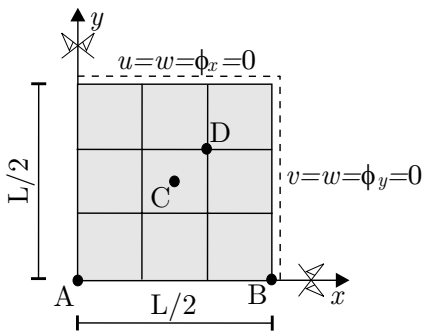


Figure 6.22: Simply supported square plate

Figure 6.24 shows the convergence in energy norm in the case of sinusoidal load for the two different stacking sequences.

Figures 6.25 and 6.26 show the convergence of the FEM stress resultants at $(L/4, L/4)$ (point C, see Figure 6.22) under uniform load for the symmetric and antisymmetric case, respectively. The values obtained are compared with the MITC laminated plate element implemented in the commercial software ADINA.

The next set of graphs highlight the importance of the RCP procedure in this context. 67

Figure 6.27 shows the convergence in energy norm of the shear stress profiles obtained using the internal moments and membrane forces before and after recovery, respectively. The error is defined as

$$e = \left(\frac{1}{2} \int_{\Omega} \left(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau} \right)^{\mathrm{T}} \mathbf{C}^{-1} \left(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau} \right) \mathrm{d}V \right)^{1/2}, \tag{6.43}$$

where $\hat{\boldsymbol{\tau}}$ is computed using directly thhe finite element stress resultants, in the

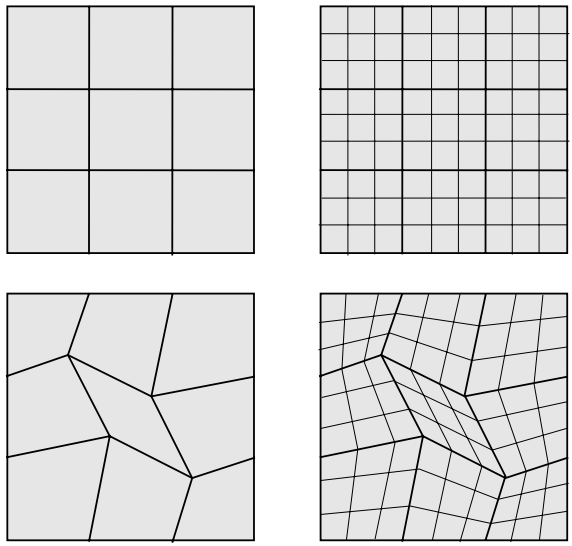


Figure 6.23: Some examples of regular and distorted meshes

first case, or the recovered stress resultants, in the second case. Notice the 9βQ4 element yelds stress resultants that satisfying the plate equilibrium equations and, thus, they can be used for shear stress recovery without validating conditions (6.40) and (6.41).

Both stacking sequences are reported, the load is sinusoidal. As can be clearly seen, the error does not decrease for the shear stresses obtained before the RCP, whereas for those obtained after the RCP the convergence rate is approximately 1.

This result can be explained by pointing out that the shear stress profiles obtained via three dimensional equilibrium depend upon the derivatives of the in plane stress components, thus the derivatives of the internal moments and membrane forces are involved. A Recovery by Compatibility in Patches procedure, while improving the convergence rate of the generalized stresses and thus of their derivatives, maintains the equilibrium within the element.

Some results concerning the effectiveness of the RCP procedure to improve the convergence rate are presented in Figures 6.28 and 6.29. Here, the convergence of the internal moments at point D are shown before and after the RCP procedure, respectively. Both stacking sequences are considered and the load is sinusoidal. Point D coincides with a node in all the considered regular meshes. Since the RCP procedure yields generalized stresses which are discontinuous across elements, in this case the value considered are those for the elements which have point D as the

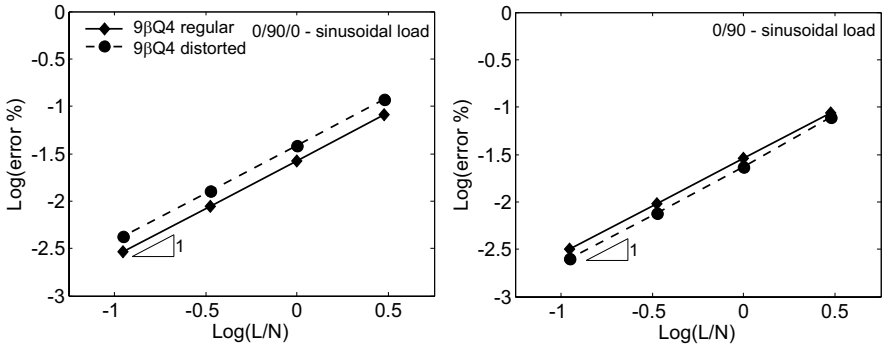


Figure 6.24: Convergence in energy norm (sinusoidal load and different staking sequences)

top right-hand node.

The stress values obtained at nodal points such as point D are usually more affected by error than the internal values. The improvement in convergence caused by the RCP is significant (the convergence rate changes from around 1 to around 2).

Figures 6.30 to 6.33 show the dimensionless transverse shear stress profiles at point C for sinusoidal load and the two different stacking sequences:

$$\bar{\tau}_{zs} = \tau_{zs} \frac{100}{q\eta} \tag{6.44}$$

In the first two figures, the profiles obtained after the Rep are shown for different meshes. The recovered profiles converge to the reference solution [Reddy, 1997]. In Figures 6.32 and 6.33, on the other hand, the profiles obtained before and after recovery for a 9×9 mesh are shown. Observing for example the result obtained for τ_{zy} using the distorted mesh in the 0/90/0 laminate (black dots in Figure 6.32), it is evident that the shape of the profile is very different from the reference solution, while the resultant force (represented by the area under the graph) is very similar. This means that, while the value of the generalized shear stress is near to the exact one, the profile does not converge unless a recovery procedure is implemented first. The importance of the RCP recovery is emphasized even more by observing Figures 6.34 and 6.35. They show the error density before and after recovery for two different meshes and the two stacking sequences. As can be noted, before recovery the order of magnitude of the error density does not change with mesh refinement, whereas a significant improvement can be seen in the error maps after recovery.

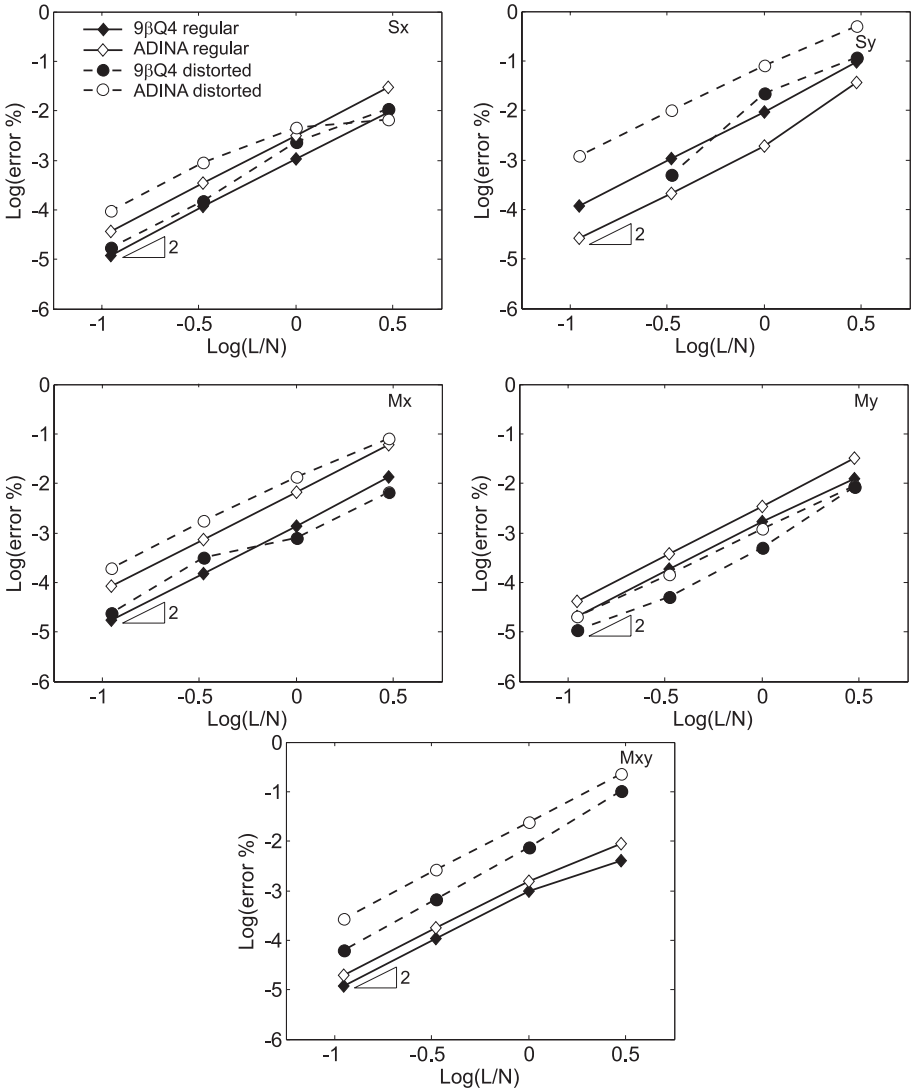


Figure 6.25: Convergence of stress resultant at point C (stacking sequence 0/90/0, uniform load)

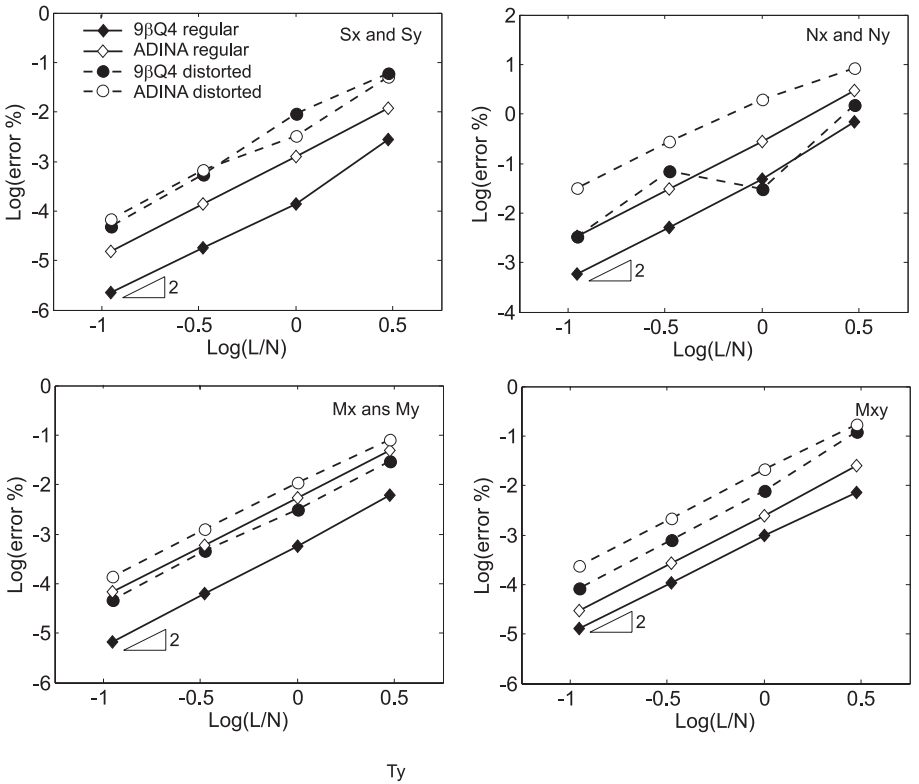


Figure 6.26: Convergence of stress resultant at point C (stacking sequence 0/90, uniform load)

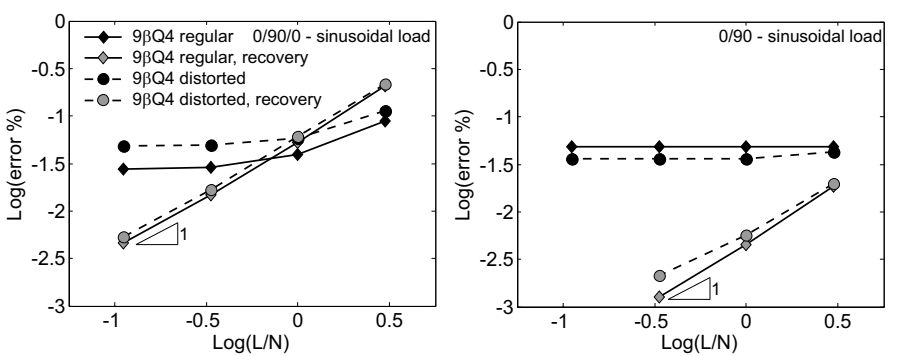


Figure 6.27: Convergence in energy norm of the shear stress profiles (sinusoidal load, different stacking sequences)

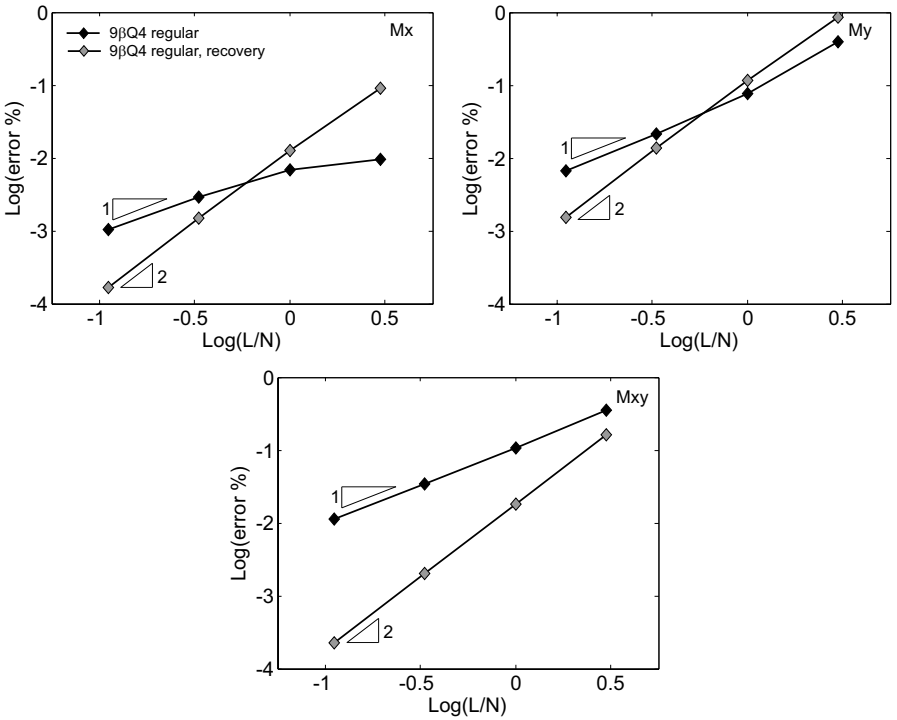


Figure 6.28: Convergence of the moments at point D before and after the recovery (stacking sequence 0/90/0, sinusoidal load)

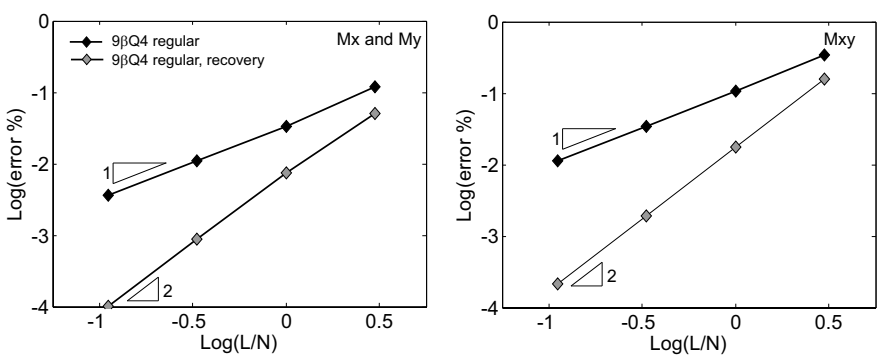


Figure 6.29: Convergenge of the moments at point D before and after the recovery (stacking sequence 0/90, sinusoidal load)

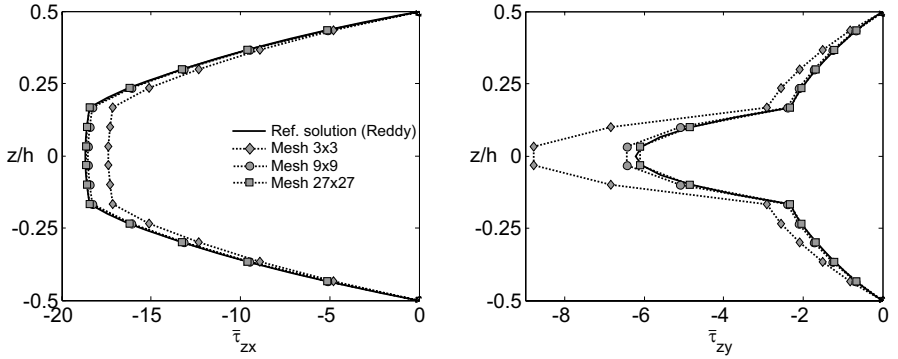


Figure 6.30: Shear stress profiles at point C after recovery for different regular meshes (stacking sequence 0/90/0, sinusoidal load)

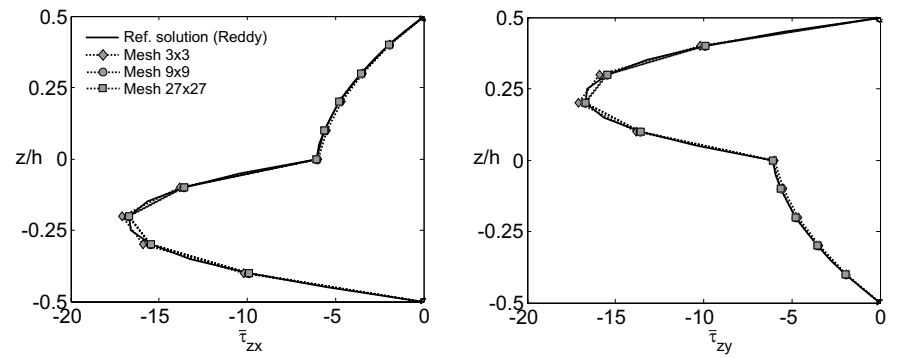


Figure 6.31: Shear stress profiles at point C after recovery for different regular meshes (stacking sequence 0/90, sinusoidal load)

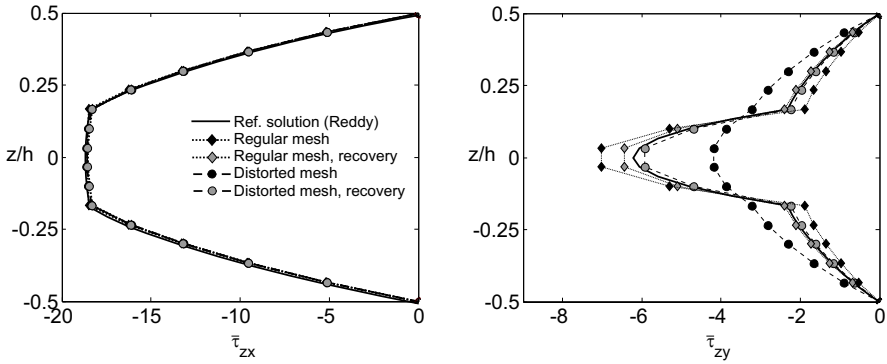


Figure 6.32: Shear stress profiles at point C after recovery for the mesh 9×9, regualr and distorted (stacking sequence 0/90/0, sinusoidal load)

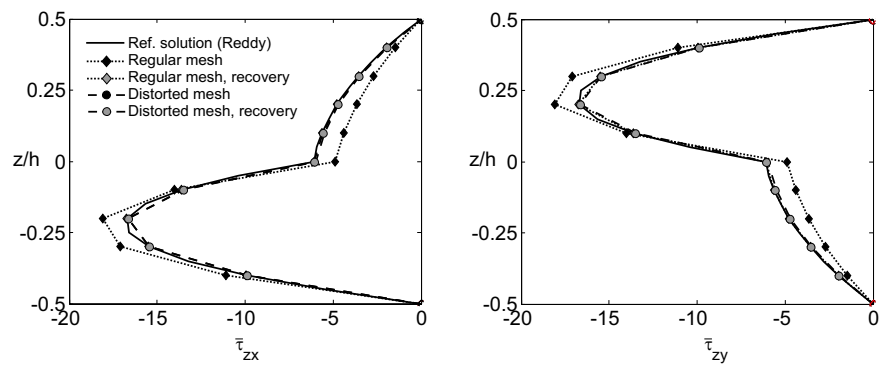


Figure 6.33: Shear stress profiles at point C after recovery for the mesh 9×9, regualr and distorted (stacking sequence 0/90, sinusoidal load)

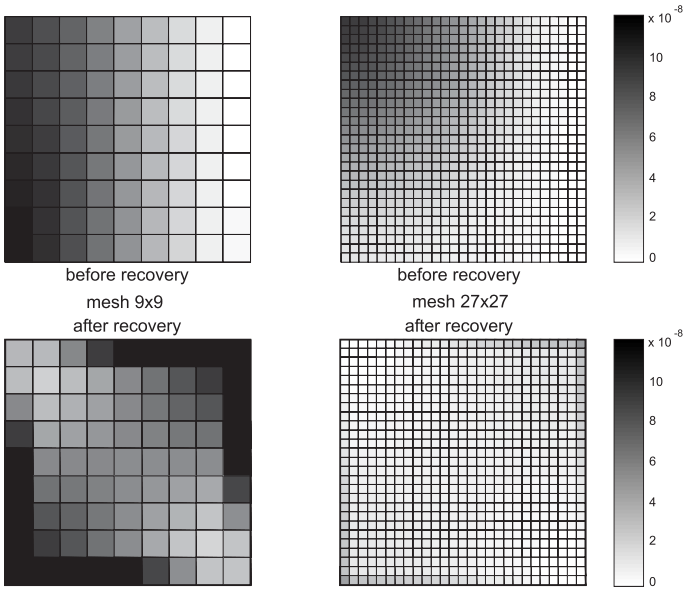


Figure 6.34: error density of the shear profiles for different meshes before and after recovery (stacking sequence 0/90/0, sinusoidal load)

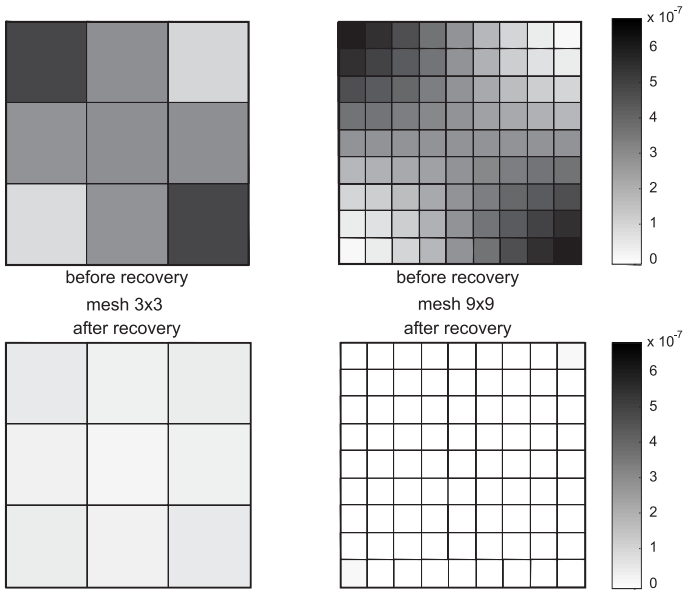


Figure 6.35: error density of the shear profiles for different meshes before and after recovery (stacking sequence 0/90, sinusoidal load)

Chapter 7

Concluding remarks: looking toward Validation

In questo capitolo, conclusivo, si intende richiamare alla mente l'intero processo di Verifica e Validazione per collocare il presente lavoro di tesi all'interno di questa ideale catena logica, anche alla luce della recente pubblicazione da parte dell 'ASME della Guida alla Verifica e Validazione nella Meccanica Computazionale dei Solidi. Di seguito sono riportati alcuni contenuti del documento, essenziali dunque alla suddetta collocazione, e vengono tratteggiati i possibili sviluppi del presente lavoro verso la validazione. A tal proposito mi piace ricordare quanto affermava il Professor A.A. Cannarozzi nel lontano 1991:

"*Gli aspetti di ideazione e di innovazione dei procedimenti di calcolo competono in linea di principio al Ricercatore. Il Produttore del codice, interagendo col Ricercatore (col quale talvolta si identifica) e col Produttore dell'apparecchiatura di calcolo, cura lo sviluppo, commercializza, aggiorna il codice e ne segue l'impiego presso l'utente. Quest'ultimo si avvale dei mezzi di calcolo come risorse per lo svolgimento della propria attività intellettuale, nell'ambito della quale i codici, per le opzioni finalizzate che offrono, diventano strumento assoluto di realizzazione di una consistente parte del processo logico. Infatti il ruolo attivo nell'impiego di un codice consiste essenzialmente nella comprensione e nella scelta della procedura logica di analisi, nella preparazione e nella consegna dei dati, nell'interpretazione e nel controllo dei risultati, mentre il coinvolgimento nell'impostazione concettuale e nelle modalità di esecuzione è per lo più impossibile. Il pericolo di impiego "a scatola*

chiusa", da molti lamentato, appare quindi effettivo. D'altra parte, la responsabilità dell'utente del codice per le conseguenze della propria attività, nei confronti del Committente o più in generale della collettività all'interno della quale egli opera, è assai spesso totale.

In considerazione di tutto questo appare comprensibile l'importanza di definire adeguate forme di regolamentazione e di controllo dell'operato delle figure interessate a vario titolo all'elaborazione automatica nella progettazione strutturale, dal Ricercatore all'utente" [Cannarozzi, 1991].

7.1 Assessing the numerical error in the finite element solution

During the last two decades, the National Agency for Finite Element Methods and Standards (NAFEMS, www.nafems.org) has developed some of the most widely known V&V benchmarks. Roughly 30 verification benchmarks have been constructed by NAFEMS. The majority of these benchmarks have targeted solid mechanics simulations, though some of the more recent benchmarks have been in fluid dynamics. Most of the NAFEMS verification benchmarks consist of an analytical solution or an accurate numerical solution to a simplified physical process described by a partial differential equation. The NAFEMS benchmark set is carefully defined, numerically demanding, and well documented. However, these benchmarks are currently very restricted in their coverage of various mathematical and/or numerical difficulties and in their coverage of physical phenomena. Further, the performance of a given code on the benchmark is subject to interpretation by the user of the code. It is also likely that the performance of a code on the benchmark is dependent on the experience and skill of the user.

Several large commercial code companies specializing in solid mechanics have developed an extensive set of well-documented verification benchmarks that can be exercised by licensed users of their codes. Such benchmarks are intended to be applied only to a particular code, and they describe how that code performed on the benchmark problems. The performance results of a code tested on the benchmark problems by a commercial company can be clearly compared with the results obtained by a user who tests the code with the same benchmark problems. These company and user-testing activities give the user a better understanding of the minimal performance that can be expected from a code. It should be noted here that information about a code's performance on a set of benchmark problems

prior to purchase of the code is often difficult to obtain, as this information is proprietary.

Moreover the most important step in any finite element analysis is clearly choosing the appropriate mathematical model to solve. This model must contain all the important ingredients regarding the geometry, loading, boundary conditions, and material data. The mathematical model must be reliable and effective and is solved using a finite element program with a mesh defined by the analyst.

The finite element solution of the mathematical model will contain some numerical error which the analyst clearly wants to be sufficiently small. This will of course be the case if a very fine mesh is used with appropriate finite elements. But, in practice, when the mesh used is not very fine, the assessment of the numerical error between the "exact" solution and the finite element solution can be important. Since the exact solution is unknown, lower and upper bounds, proven to bracket closely the exact solution, and obtainable with an acceptable computational effort, would be very valuable. Unfortunately, these bounds are not available for general analysis, not yet even in academic research. Hence for general analyses, the currently available error assessments are based on estimates without these bounds.

Since this error estimation is an important issue, and presents an excellent research challenge, a very large research effort by engineers and mathematicians has focused on the development of error estimators. The current state of art and hopefully some improvements on error estimation has been discussed in this work. But which is the role played by error estimation in the whole V&V process shown in Figure 7.1?

Recently the American Society of Mechanical Engineers (ASME) Standard Committee on Verification and Validation in Computational Solids Mechanics (PTC 60 / V&V 10) approved their first document (July 2006). This Guide has been submitted to ASME publications (http://catalog.asme.org/) and to American National Standards Institute (ANSI) for public review. In the following section some parts of this Guide are presented, underlying the role of error estimators.

7.2 Validate the model, not the codes

The notion of benchmarking a code refers to comparisons of results of the code to some standard "benchmark", which may be an analytical solution or, as many argue, a single physical experiment. Various uses of the notion of benchmarking are found in the literature. Among these are:

1. Code-to-Code Comparisons of Code Performance. Here execution speeds or

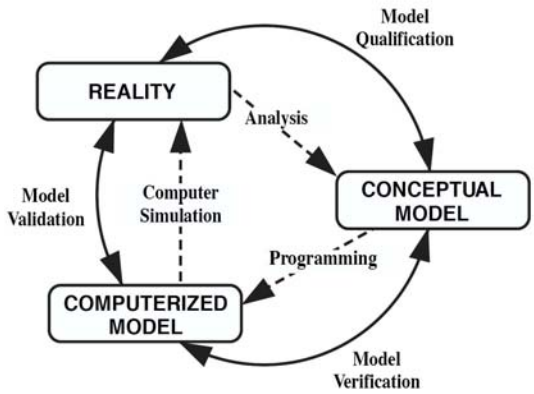


Figure 7.1: The scenario

numerical results, for example, of different codes solving one or more benchmark problems are compared on the same computer or on different computers running the same code.

2. Code-to-Code Comparisons Versus Verified Code. Here one assumes that a fully-verified code is in hand, and one compares results of different codes with those produced by the master code on a suite of benchmark problems.
3. Comparison of Code Predictions to Results of Physical Experiments. Here results of one or more codes are compared with results obtained by carefully executed physical experiments.
4. Comparison of Code Results to Analytical or Highly Accurate Solutions. Here a variety of features of basic algorithms in a code are tested by comparison of computed results with known analytic solutions or with "overkill" solutions obtained with very fine meshes using verified codes.

All of these approaches suffer from defects, some more serious than others.

Use 1 (*Code-to-Code for Performance*) has nothing to do with verification.

Use 2 (*Code-to-Verified Code Comparisons*) is not entirely useless, but is an exercise fraught with pitfalls. If the "verified" code is not as accurate as the tested code, how can useful conclusions be made by simply comparing results? Which is correct? Such comparisons may be meaningless, without reference to other information.

Use 3 (*Comparison with Experiments*), while a common benchmark scenario, is the worst of all. If discrepancies between computed and observed responses are observed, what is the source of error? Is the model bad or is the code inadequate or does the code have bugs? There is no systematic way to distinguish between

modeling error and approximation error by a simple comparison of computed predictions of a single experiment (which itself is subject to gross error). One must first "verify the code".

Use 4 (*Comparison with Accurate Solutions*) is somewhat better, but also imperfect. Exactly what should be compared? Here we come to the central question of benchmarking: what exactly are the comparisons with results of benchmark problems supposed to test? Going to the core of the issue: what do we want to know about a code to determine if it is "solving the model correctly"? The issue is often much more complicated than this: if we are given results of a "black box" commercial code where the precise model used to predict response is unknown (as it is propriety information, belonging to the code developer) what information can a suite of benchmark calculations provide?

We argue that: 1) benchmarking is primarily a verification exercise, not one of validation. Comparisons of computed predictions with experiments is the essence of validation, thus the subject code to be tested against experiments must have been already subject to verification test using, among other tools, benchmarking tests; 2) one must clearly set out tests designed to establish how well the subject code can solve the model used to depict natural phenomena or the behavior of engineering systems; 3) on the other hand, benchmarking need not necessarily be tied exclusively to the model used to develop the code, as the user has right to know where and when the model becomes invalid or inaccurate; 4) benchmarking, or the selection of a benchmark suite, is a delicate and complex proposition: tests on performance of a code on distorted meshes, on problems with nice smooth solutions, or on specific quantities of interest in a calculation must be made with a clear understanding of the capabilities of the model on which the code is based.

Some principles usually recommended for designing benchmark problems are:

1. **Failure** - Benchmark problems should be designed to establish the limits of the code under consideration. Limits means failure: when does the code fail to produce results in agreement with a suite of carefully-selected analytical solutions? Design the benchmarks so that bad results, not always glorious agreement, are obtained, by altering time steps, load steps, mesh size, approximation order, material mismatch, boundary-layers, point or line-loads, etc. Surely, good agreement on simple problems is necessary, but not sufficient to determine the real limitations of interest to the user.
2. **Convergence** - If a code produces pretty results but doesn't converge or doesn't converge at acceptable rates, then its use as a simulation tool is suspect at best. Also, the real meaning of the limit to which the code converges to is important. Benchmark problems should be designed to determine via

numerical experiments the rates of convergence of a model as the mesh size, the order of the approximation, time step, load step, etc. are varied, convergence being measured in a space of meaningful norms. Whenever possible these should be compared with a priori estimates if they are available for the problem at hand. Then a sequence of benchmarks can be designed to establish the rates-of-convergence for various regularities. Thus, if a solution of a problem is known, the error can be calculated for each choice of mesh size and order.

3. **Local Behavior, Diffusion, Propagation** - Users interested in simulating physical phenomena are concerned with how well the code in hand can depict so-called local phenomena. What are stresses at a point? How does the code handle convection of a signal propagating through a domain? Are bifurcations, localizations mesh dependent? Or do the predicted critical parameters agree with those depicted in appropriate analytical solutions? These issues must be considered in designing benchmark problems.
4. **Model Limitations** - Given a code (often a black box), design benchmark problems to fool or verify or, at least, determine the limitations of the model on which the code is based, this is not always easy. As a general rule, the design of effective benchmark problems (as a verification tool) should be based on models which supersede in complexity and sophistication those on which the code is based. Limitations of the subject code can be determined in a broader context. Then the limits of the model on which the code is based can also be roughly evaluated.

The conclusion: the results of benchmark problems must be viewed and evaluated within the context of what is predictable by the theory upon which the underlying model is based. Once again, the goal of verification is to determine if the model is solved correctly.

Thus, model verification depends upon what can and should be deliverable by the model. Such examples suggest that when the base model used in a code is unknown, the benchmarks should be based on models presumed to be of the highest sophistication and breadth appropriate for the class of problems under consideration.

For this and other reason in 1999 an ad hoc verification & validation specialty committee was formed under the auspices of the United States Association for Computational Mechanics (USACM). The purpose of this committee was to pursue the formation of a verification & validation standards committee under a professional engineering society approved to produce standards under the rules of the American National Standards Institute (ANSI). This goal was achieved in 2001 when the then Board on Performance Test Codes (PTC) of the American Society

of Mechanical Engineers (ASME) approved the committee's charter:

"To develop standards for assessing the correctness and credibility of modelling and simulation in computational solid mechanics."

and the committee was assigned the title and designation of the ASME Committee for Verification & Validation in Computational Solid Mechanics (PTC 601).

The membership is diverse with three major groups being industry, Government, and academia. The industry members include representatives from auto and aerospace industries and the Government members are primarily from the Departments of Defense and Energy. Particularly well represented are members from the three national laboratories under the National Nuclear Security Administration. This latter membership group is key to the committee as much of the recent progress in verification & validation has come from these laboratories and their efforts under the Advanced Simulation and Computing (ASC) Program, started in 1995.

7.3 An overview of the PTC 60 / V&V 10

In this section, we give an overview of the "Performance Test Code #60" (PTC60) recently appeared on the IACM bulletin (No. 20, January 2007) on the verification and validation topic. This article depicts the main steps done by the committee related to the necessity of standards and rules in the computational Solid Mechanics field, and starts with the following funny example:

question: "Are the sometimes lengthy and costly processes of verification & validation really necessary?"; answer: "Consider the following scenario that perhaps you can relate to first hand. A project review meeting is taking place and the project manager needs to make a critical decision to accept or reject a proposed design change. A relatively new employee, freshly minted from the nearby engineering university, makes an impressive presentation full of colorful slides of deformed meshes and skillfully crafted line plots indicating the results of many CPU and labour hours of non-linear numerical analyses, ending with a recommendation to accept the design change.

Hopefully, an astute project manager, aware of the vagaries of nonlinear numerical analyses, will not accept the analysis and its conclusion at face value, especially given the inexperience of the analyst. Rather, the project manager should seek some assurance that not only are the results reasonable, but a sound procedure was followed in developing the model and documenting the numerous physical and numerical parameters required for a typical analysis. The degree of assurance sought by the project manager is directly related to the criticality of the decision to

be made".

The processes of verification & validation are how evidence is collected, and documented, that help establish confidence in the results of complex numerical simulations.

The intended audience for this Guide is not the occasional computational mechanics user, e.g. a modern-day draftsman using an automated CAD/FEA package, rather it is computational analysts, experimentalists, code developers, and physics model developers, and their managers, who are prepared to read a technical document with a mixture of discussion concerning mathematics, numerics, experimentation, and engineering analysis processes.

7.3.1 The guide and the present work

The guidelines are based on the following key principles:

- Verification must precede validation.
- The need for validation experiments and the associated accuracy requirements for computational model predictions are based on the intended use of the model and should be established as part of V&V activities.
- Validation of a complex system should be pursued in a hierarchical fashion from the component level to the system level.
- Validation is specific to a particular computational model for a particular intended use.
- Validation must assess the predictive capability of the model in the physical realm of interest, and it must address uncertainties that arise from both simulation results and experimental data.

The Guide contains four major sections:

1. Introduction - the general concepts of verification and validation are introduced and the important role of a V&V Plan is described.
2. Model Development - from conceptual model, to mathematical model, and finally the computational model are the keys stages of model development.
3. Verification - is subdivided into two major components: code verification - seeking to remove programming and logic errors in the computer program, and calculation verification - to estimate the numerical errors due to discretization approximations.
4. Validation - experiments performed expressly for the purpose of model validation are the key to validation, but comparison of these results with model results depends on uncertainty quantification and accuracy assessment of the results.

In the next sections the contents of this four points will be sketched.

7.3.1.1 The Introduction Section

The processes of verification & validation start, and end, with modelling and models, and so we seek to verify & validate a computational model, for making predictions within the domain of intended use of the model. Three types of models, from the general to the specific, are described. The logic flow from the most general Conceptual, to Mathematical, to the most specific Computational Model, is illustrated in Figure 7.2.

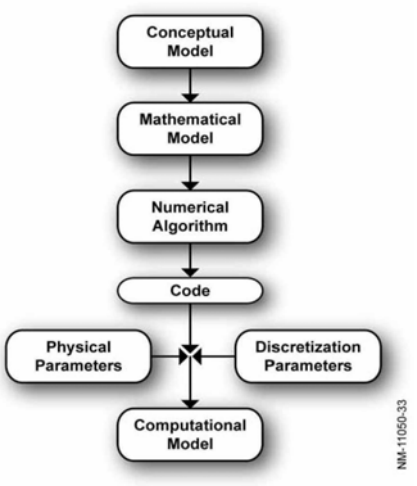


Figure 7.2: The path from Conceptual and Computational Model

Neither part of verification addresses the question of the adequacy of the selected Conceptual and Mathematical models for representing the reality of interest. Answering this question is the domain of validation, i.e. are the mechanics (physics) included in the Conceptual and Mathematical models sufficient for answering the questions in the problem statement.

The manner in which the Verification and Validation interact is illustrated in the flow chart shown in Figure 7.3. After the selection of the Conceptual model, the V&V process has two branches: the left branch contains the modelling elements and the right branch the experimental elements.

This figure is intentionally designed to illustrate the paramount importance of physical testing in the V&V process, as ultimately, it is only through experimentation that assessments about the adequacy of the selected Conceptual and

Mathematical models for representing the reality of interest can be made. Close cooperation among modelers and experimentalist is required during all stages of the V&V process, until the experimental outcomes are obtained. Close cooperation is required because the two groups will have quite different views of the Conceptual model, i.e. the mathematical and physical model will be different.

As an example consider the last example shown in the previous chapter about composite laminated plate with simply supported boundary conditions. Mathematically this boundary condition is quite easy to specify, but in the laboratory there is no such thing as a simply supported or 'clamped' boundary. In general, some parts of the Conceptual model will be relatively easy to include in either the mathematical or physical model, and others more difficult. A dialogue between the modelers and experimentalist is critical to resolve these differences. To aid in this dialogue, the 'cross-talk' activity labelled as "Preliminary Calculations" in Figure 7.3 is intended to emphasize the goal that both numerical modelers and experimentalist attempt to model the same Conceptual model.

Of equal importance is the idea that the experimental outcomes should not be revealed to the modelers until they have completed the simulation outcomes. The chief reason for segregation of the outcomes is to enhance the confidence in the model's predictive capability. When experimental outcomes are made available to modelers prior to establishing their simulation outcomes, the human tendency is to 'tune' the model to the experimental outcomes to produce a favorable comparison. This tendency decreases the level of confidence in the model's ability to predict, and moves the focus to the model's ability to mimic the provided experimental outcomes.

7.3.1.2 The Model Development Section

Before modelling begins, a reality of interest is identified, i.e. what is the physical system to be modelled. The reality of interest is typically described in the problem statement presented to the analyst, e.g. "We need to know the behaviour of a certain slab under a distributed load of x Newtons/meter". In this case the reality of interest is the composite laminated plate.

The most general form of the model addressed in the Guide is the Conceptual Model - "the collection of assumptions and descriptions of physical processes representing the solid mechanics behaviour of the reality of interest from which the mathematical model and validation experiments can be constructed". Continuing with the slab example, the conceptual model could be a simply supported plate of a laminated composite material, loaded uniformly.

With the Conceptual Model defined, the analyst next defines the Mathematical Model - "The mathematical equations, boundary values, initial conditions, and

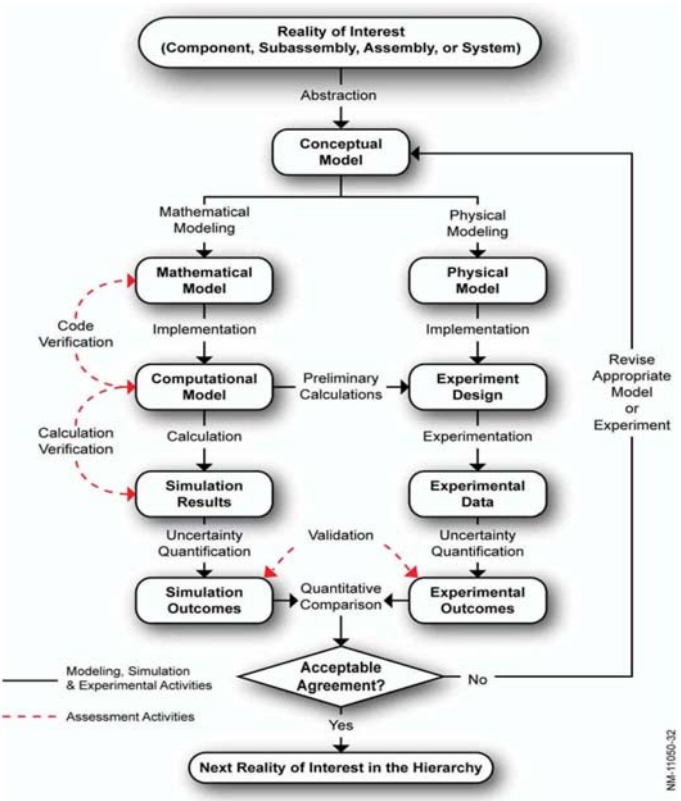


Figure 7.3: Verification and Validation activities and outcomes

modelling data needed to describe the conceptual model". Referring again to the slab example the mathematical model could be established based on the First order Shear Deformation Theory.

The final model in the sequence is the Computational Model - "The numerical implementation of the mathematical model, usually in the form of numerical discretization, solution algorithm, and convergence criteria". In our slab example, the computational model is derived within with the finite element framework using, for example, the 9βQ4 hybrid stress element.

At this point the computational model can be exercised (run) and the results compared to available experimental data for validation of the model. It is frequently the case that the results do not compare as favourably as requested in the original problem statement. Assuming a high degree of confidence in the experimental data, the analyst has two basic choices for revising the model: changing the model form or calibrating model parameters.

Changing the model form can apply to either the Mathematical or Conceptual

model. As an example, the first order shear theory assumption reveals to be too restrictive and needs to be replaced by a higher-order plate theory or the plate model reveals to be inadequate and should be replaced by a plate-like body.

7.3.1.3 The Verification Section

The Guide emphasizes that *Verification must precede Validation. The logic is that attempting to validate a model using a code that may still contain (serious) errors can lead to a false conclusion about the validity of the model.*

Two types of verification are generally recognized and defined in computational simulation:

1. *Code Verification - establish confidence, through the collection of evidence, that the mathematical model and solution algorithms are working correctly.*
2. *Calculation Verification - establish confidence, through the collection of evidence, that the discrete solution of the mathematical model is accurate.*

Code Verification *In general, Code Verification is the domain of software developers who hopefully use modem Software Quality Assurance techniques along with testing of each released version of the software. Users of software also share the responsibility for code verification, even though they typically do not have access to the software source. The large number of software users, typical of most commercial codes, provides a powerful potential code verification capability, if it is used wisely by the code developers.*

Among the code verification techniques, the most popular method is to compare code outputs with analytical solutions; this type of comparison is the mainstay of regression testing. Unfortunately, the complexity of most available analytical solutions pales compared to even rather routine applications of most commercial software. One code verification method with the potential to greatly expand the number and complexity of analytical solutions is what is termed in the V&V literature as manufactured solutions.

The basic concept of a manufactured solution is deceptively simple. Given a partial differential equation (PDE), and a code that provides general solutions of that PDE, an arbitrary solution to the PDE is manufactured, i.e. made up, then substituted into the PDE along with associated boundary and initial condition, also manufactured. The result is a forcing function (right-hand side) that is the exact forcing function to reproduce the originally selected (manufactured) solution. The code is then subjected to this forcing function and the numerical results compared with the manufactured solution. If the code is error free the two solutions should agree.

Calculation Verification *The above illustration of a manufactured solution used as part of code verification is only half of the verification effort. The other half is what is termed calculation verification, or estimating the errors in the numerical solution due to discretization. Calculation verification, of necessity, is performed after code verification, so that the two error types are not confounded.*

Any comparison of the numerical and analytical results will contain some error, as the discrete solution, by definition, is only an approximation of the analytical solution. So the goal of calculation verification is to estimate the amount of error in the comparison that can be attributed to the discretization.

The discretization error is most often estimated by comparing numerical solutions at two more discretizations (meshes) with increasing mesh resolution, i.e. decreasing element size. The objective of this mesh-to-mesh comparison is to determine the rate of convergence of the solution.

The main responsibility for Calculation Verification rests with the analyst, or user of the software. While it is clearly the responsibility of the software developers to assure their algorithms are implemented correctly, they cannot provide any assurance that a user-developed mesh is adequate to obtain the available algorithmic accuracy, i.e. large solution errors due to use of a coarse (unresolved) mesh are attributable to the software user.

7.3.1.4 The Validation Section

The validation process has the goal of assessing the predictive capability of the model. This assessment is made by comparing the predictive results of the model with validation experiments. If these comparisons are satisfactory, the model is deemed validated for its intended use, as stated in the V&V Plan. There is perhaps a subtle point here to be emphasized. The original reason for developing a model was to make predictions for applications of the model where no experimental data could, or would, be obtained. However, in the V&V Plan it was agreed that if the model could adequately predict some related, and typically simpler, instances of the intended use, where experimental data would be obtained, then the model would be validated to make predictions beyond the experimental data for the intended use. Simply put, if the model passes the tests in the V&V Plan, then it can be used to make the desired predictions with confidence. The V&V Plan is of paramount importance to the V&V process.

When it is said that the model is validated for the intended use, it is not just the Computational model, which likely will have to change for the predictions of interest, but the Mathematical and Conceptual models upon which the Computation model was built that have been validated. It is through the validation of the Conceptual model that confidence is gained that the correct physics (mechanics)

were included in the model development.

The key components of the validation process are the:

- *Validation Experiments - experiments performed expressly for the purpose of validating the model.*
- *Accuracy Assessment - quantifying how well the experimental and simulation outcomes compare.*

The goal of a validation experiment is to be a physical realization of an initial boundary value problem, since an initial boundary value problem is what the computational model was developed to solve. Most existing experiments do not meet the requirements of a validation experiment, as they were typically performed for purposes other than validation. Certainly appropriate existing experimental data should be used in the validation process, but the resulting confidence in the model's ability to make predictions, based on these experimental results, is diminished, relative to validation experiments. The reduced confidence arises from the necessity of an analyst needing to select physical and numerical parameters required for the model that were left undefined in the experiment.

The important qualities of a validation experiment include:

- *Redundancy of the Data - repeat experiments to establish experimental variation.*
- *Supporting Measurements - not only are measurements of the important system response quantities of interest recorded, but other supporting measurements are recorded. An example would be to record the curvature of a beam to support a strain gauge measurement.*
- *Uncertainty Quantification - errors are usually classified as being either random error (precision) or systematic error (bias).*

Once the experimental and simulation outcomes are obtained, the accuracy assessment phase of the validation process can begin. If possible, the comparison of the experimental and simulation outcomes should be made by an interested third party, as this helps to remove a bias that favors either the experimental or the simulation results. In addition to deciding what response quantities should be compared, the V&V Plan should state how the quantities are to be compared.

Validation metric is the term used to describe the comparison of validation experiment and simulation outcomes. These metrics can range from simple binary metrics, e.g. was the material's yield strength exceeded, to more complex comparisons involving magnitude and phase difference in wave forms, e.g. deceleration history in a vehicle crash. Whatever the form of the validation metric, the result should be a quantitative assessment of the agreement between the experiment and

simulation. Hopefully, this quantification will also include an estimate of the variability in the agreement and a confidence statement about the variability, e.g. the relative error between the experiment and simulations was 18% plus or minus 6% with a 85% confidence level. This three-part comparative statement is provided to the decision maker, along with all the supporting V&V documentation, to aide in their decision making process about the validity of the model for the intended use.

7.3.1.5 The Conclusion Section

Some of the remaining important V&V activities requiring guidance from the community:

- *Verification - this 'poor' sister of validation needs more attention from the V&V research community. Reliance on regression testing for code verification provides minimal confidence when using today's complex multi-physics and multi-scale software. Methods, and their implementation as tools, for verification of increasing software complexity are needed.*
- *Quantification of the Value of V&V - if program managers are asked to spend resources on V&V, they needed some measure of the value they are receiving for the resources expended.*
- *Incomplete V&V - if the V&V process is terminated before a successful conclusion, what is the best path forward for decision maker?*
- *Validation Experimentation - most experiments consume large amounts of resources, the value of these experiments to the V&V process needs to be quantified to enable decision makers to appropriately allocate resources for this important activity.*
- *Uncertainty Quantification - meaningful comparisons of simulations with experiments requires an estimate of the uncertainty in both sets of results, and a comparative assessment of these two uncertain outcomes.*
- *Predictive Confidence - when validated models are applied beyond the limited range of validation experiments, how can the confidence in these results be quantified?*

7.4 Concluding remarks

As a conclusion of the present work I would like to re interpret the last example shown in the previous chapter: the laminated composite plate, where our error verification tool (the RCP recovery procedure) is used to reconstruct the transverse shear stress along the laminate thickness. This is an attempt to improve simulation results by reviewing the underlying model, a sort of an ideal hierarchy of model

(from plates to plate-like bodies). Thus this could be viewed as an attempt to go beyond verification, looking at validation.

Looking back in the past, I would like to end this work with some advises coming from the far 1980 by the brilliant mind of Bruce Irons [Irons & Ahmadn, 1980] (that it still up-to-date), and with the encouraging willpower of J. Tinsley Oden.

- *Finite elements are no substitute for commonsense.*
- *Trouble-shooting will continue to involve engineers in the sort of calculations they have always done but have seldom been taught to do.*
- *A good engineer tries to understand every failure and learn from it.*
- *An element may succeed in one problem and fail in another. Finite elements are not infallible, but their strength lies in their versatility. This is the kind of strength valued by any engineer, who by choice would never buy a tool that does only one job.*
- *The matched solution involves thinking about computer results in an ordinary, engineering way and not merely accepting them as magic.*
- *When things go wrong, we want to try other approaches quickly, and when something quite new is demanded we do not expect to be kept waiting six months for a spare part.*
- *In seeking a suitable mesh, either by trial and error, or guessing from an intuitive impression of what the answers will be like, it is useful to regard the elements as interpolation formulae. If an element gives linear stresses, we judge the size the elements should be on the basis that they are representing the field by linear segments. The curvatures indicate, very roughly, what errors we should expect. Barlow points improve the estimates, and also make prediction more hazardous.*
- *Perhaps ultimately the most useful thing to emerge from the matched solution is the consciousness of, and a technique for identifying, the unwanted side-effects in the performance of an element. This will be important when we develop 'hybrid' elements to do a better job. For we can then choose the responses we want, in terms of stress fields, and exclude those we do not.*

As a last point of this list, Irons inserts a curious picture, Figure 7.4, that probably represent the little snares hidden behind the computer manipulations.

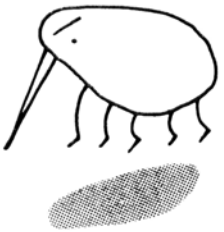


Figure 7.4: Cimex Fortranarius, magnification 10^9:1 (See [Irons & Ahmadn, 1980])

Professor J. Tinsley Oden in 1996, speaking about Computational mechanics:

"*Modeling and simulation represent catalysts capable of synthesizing formal theory and experimental results, breaking artificial barriers between these two main approaches to scientific discovery and engineering advances. Recently, computational simulation has become a third approach - along with theory and laboratory simulation - to studying and solving scientific and engineering problems. Computational simulation is based on the use of high-performance computers to model and simulate complex systems. In this approach, a computer equipped with problem-solving software tools may represent a virtual laboratory in which researchers can build a model for a given problem and run it under varying conditions.*

These increasingly complex computational methodologies require sophisticated models and numerical algorithms, and vice versa" (Excerpts from the Timoshenko Medal Acceptance Speech delivered by Professor J. Tinsley Oden in 1996. "The Revolution in Applied Mechanics from Timoshenko to Computation").

Appendix A
Structural models

A.1 Class I problems

The first class of problems considered refers to problems governed by the following differential equation

$$k\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}\right)=b, \qquad \text{in } \Omega. \tag{A.1}$$

As usual, stresses are defined by

$$\sigma_x = k\frac{\partial u}{\partial x}, \qquad \sigma_y = k\frac{\partial u}{\partial y}. \tag{A.2}$$

Within the context of linear elasticity, this problem may be viewed, for example, as the equilibrium problem of a thin membrane of uniform thickness, subjected to a distributed transverse force b. In the numerical tests, the parameter k is set to $k = 1$.

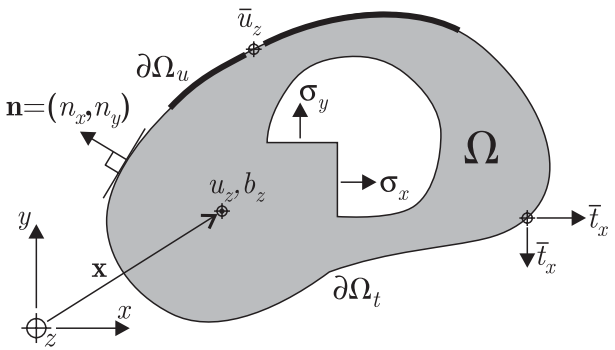


Figure A.1: Class I problems

A.1.1 The RCP stress approximation

The stress approximation for the RCP procedure can be derived from a stress function $\phi(x, y)$ as follows

$$\boldsymbol{\sigma}_{ph}^{*\,\mathrm{T}} = \left[\sigma^*_{h\,x}\ \sigma^*_{h\,y}\right], \qquad \sigma^*_{h\,x} = \frac{\partial \phi}{\partial y}, \qquad \sigma^*_{h\,y} = -\frac{\partial \phi}{\partial x}. \tag{A.3}$$

The particular solution $\boldsymbol{\sigma}^*_{pp}$ is taken as

$$\boldsymbol{\sigma}^{*\,\mathrm{T}}_{pp} = \left[\sigma^*_{p\,x}\ \sigma^*_{p\,y}\right], \qquad \sigma^*_{p\,x} = \frac{1}{2}\int_0^x b\,\mathrm{d}x, \qquad \sigma^*_{p\,y} = \frac{1}{2}\int_0^y b\,\mathrm{d}y. \tag{A.4}$$

Thus, selecting the stress approximation simply reduces to select the stress function. Here, ϕ is assumed as a complete polynomial expansion of degree $(q+1)$, so that each stress component is approximated by a complete polynomial of degree q. For finite elements of order p, two different values for q are tested: $q = p$ and $q = p+1$. The corresponding implementations of the present procedure are labelled by RCP1 and RCP2, respectively. Thus, local stresses are approximated by linear (RCP1) and quadratic (RCP2) polynomials, in the case of three- and four-node elements, and by quadratic (RCP1) and cubic (RCP2) polynomials, in the case of six- and nine-node elements. The expressions of these local approximations are reported in the following. The number of unknown parameters involved by linear, quadratic and cubic approximations are 5, 9 and 14, respectively. In the REP procedure, complete linear polynomials are used for three- and four-node elements ($3+3$ unknown parameters) and complete quadratic polynomials for six- and nine-node elements ($6+6$ unknown parameters).

The local stress approximations used in the RCP procedure for Class I problems are:

(a) linear 5-parameters approximation used in the RCP1 version for 4-node elements

$$\boldsymbol{\sigma}^*_{ph} = \begin{bmatrix} 1 & 0 & y & 0 & x \\ 0 & 1 & 0 & x & -y \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_5 \end{bmatrix}, \tag{A.5}$$

(b) quadratic 9-parameters approximation used in the RCP2 version for 4-node elements and in the RCP1 version for 9-node elements

$$\boldsymbol{\sigma}^*_{ph} = \begin{bmatrix} 1 & 0 & y & 0 & x & y^2 & 0 & x^2 & -2xy \\ 0 & 1 & 0 & x & -y & 0 & x^2 & -2xy & y^2 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_9 \end{bmatrix}, \tag{A.6}$$

(c) cubic 14-parameters approximation used in the RCP2 version for 9-node elements

$$\begin{aligned} \boldsymbol{\sigma}^*_{ph} \;=\; & \begin{bmatrix} 1 & 0 & y & 0 & x & y^2 & 0 & x^2 & 2xy & y^3 & 0 \\ 0 & 1 & 0 & x & -y & 0 & x^2 & -2xy & -y^2 & 0 & x^3 \end{bmatrix} \\ & \begin{bmatrix} x^3 & -3xy^2 & 2x^2y \\ -3x^2y & y^3 & -2xy^2 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_{14} \end{bmatrix}. \end{aligned} \tag{A.7}$$

A.2 Class II problems

The second class of problems is plane elasticity.

A.2.1 Basic equations

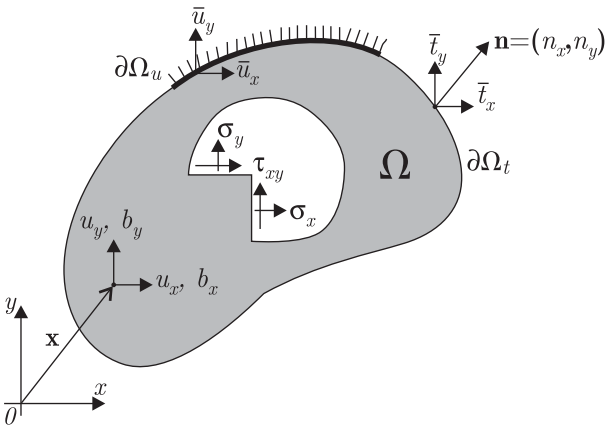


Figure A.2: Class II problems

Consider the elastic plane problem, shown in Figure A.2, governed by the following equations:

$$\mathbf{D}^*\boldsymbol{\sigma} = \mathbf{b} \quad \text{in } \Omega, \tag{A.8}$$

$$\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\epsilon} \quad \text{in } \Omega, \tag{A.9}$$

$$\boldsymbol{\epsilon} = \mathbf{D}\mathbf{u} \quad \text{in } \Omega, \tag{A.10}$$

where Ω is the domain, together with the following boundary conditions on $\partial\Omega$

$$\mathbf{u} = \bar{\mathbf{u}} \quad \text{in } \partial\Omega_u, \tag{A.11}$$

$$\mathbf{N}^{\mathrm{T}}\boldsymbol{\sigma} = \bar{\mathbf{t}} \quad \text{in } \partial\Omega_t, \tag{A.12}$$

where matrix $\mathbf{N}$ collects the components of the outward unit normal to $\partial\Omega$

$$\mathbf{N} = \begin{bmatrix} n_x & 0 & n_y \\ 0 & n_y & n_x \end{bmatrix}. \tag{A.13}$$

According to the figure, displacement $\mathbf{u}$, strains $\boldsymbol{\epsilon}$ and stress $\boldsymbol{\sigma}$ are given by

$$\mathbf{u}(\mathbf{x}) = \begin{bmatrix} u_x \\ u_y \end{bmatrix}, \qquad \boldsymbol{\epsilon}(\mathbf{x}) = \begin{bmatrix} \epsilon_x \\ \epsilon_y \\ \gamma_{xy} \end{bmatrix}, \qquad \boldsymbol{\sigma}(\mathbf{x}) = \begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix}. \tag{A.14}$$

Compatibility and equilibrium operators $\mathbf{D}$ and $\mathbf{D}^*$ take the form

$$\mathbf{D} = \begin{bmatrix} \partial/\partial x & 0 \\ 0 & \partial/\partial y \\ \partial/\partial y & 0\partial/\partial x \end{bmatrix}, \quad \mathbf{D}^* = \begin{bmatrix} -\partial/\partial x & 0 & -\partial/\partial y \\ 0 & -\partial/\partial y & -0\partial/\partial x \end{bmatrix}. \tag{A.15}$$

$\mathbf{b}$ is a prescribed source term,

$$\mathbf{b}(\mathbf{x}) = \begin{bmatrix} b_x \\ b_y \end{bmatrix}. \tag{A.16}$$

In the case of isotropic material in plane stress conditions, the elasticity matrix $\mathbf{C}$ takes the form

$$\mathbf{C} = \frac{Et}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{1-\nu}{2} \end{bmatrix}, \tag{A.17}$$

where E is the Young's modulus, ν is the Poisson's ratio and t is the thickness.

A.2.2 The RCP stress approximation

The stress approximation can be easily constructed employing the well-established techniques usually adopted for hybrid stress elements. In particular, for the equilibrium problem of plane elasticity, $\boldsymbol{\sigma}^*_{ph}$ and $\boldsymbol{\sigma}^*_{pp}$ can be selected according to the following relations:

$$\boldsymbol{\sigma}^{*\mathrm{T}}_{ph} = \left[\sigma^*_{h\,x}\ \sigma^*_{h\,y}\ \tau^*_{h\,xy}\right], \tag{A.18}$$

$$\sigma^*_{h\,x} = \frac{\partial^2\phi}{\partial y^2}, \quad \sigma^*_{h\,y} = \frac{\partial^2\phi}{\partial x^2}, \quad \tau^*_{h\,xy} = -\frac{\partial^2\phi}{\partial x \partial y}, \tag{A.19}$$

$$\boldsymbol{\sigma}^{*\mathrm{T}}_{pp} = \left[\sigma^*_{p\,x}\ \sigma^*_{p\,y}\ \tau^*_{p\,xy}\right], \tag{A.20}$$

$$\sigma^*_{p\,x} = -\int_0^x b_x\,\mathrm{d}x, \quad \sigma^*_{p\,y} = -\int_0^y b_y\,\mathrm{d}y, \quad \tau^*_{p\,xy} = 0. \tag{A.21}$$

The particular solutions given in Equation (A.20) are those commonly adopted for hybrid finite elements. Indeed, other choices are possible which would lead, obviously, to different final solutions. However, the various choices become all equivalent in practice when the mesh is refined, since all the recovered solutions have been proved to converge to the exact one [Ubertini, 2004].

Based on the above assumptions, choosing the stress approximation simply reduces to choose the stress function ϕ. Regarding this, it should be observed that the consistency of the procedure requires the self-equilibrated stress approximation to be selected among polynomials of degree not less than p, being p the interpolation order for displacements in the finite element solution. Moreover, the invariance of the procedure with respect to any coordinate change, requires that all the components of $\boldsymbol{\sigma}^*_{ph}$ are represented by complete polynomial expansions of the same degree. Therefore, the stress function should be chosen as a complete polynomial

of degree not less than $(p+2)$ in plane elasticity problems (see Equation (A.18)).

RCP1 and RCP2 denote the RCP procedure with p- and $(p+1)$-degree stress approximation, respectively. In particular RCP1 and RCP2 refer to the linear 7-parameters and quadratic 12-parameters approximations, when linear elements are used, while to the quadratic 12-parameters and cubic 18-parameters approximations, when quadratic elements are used. In the REP procedure, complete linear polynomials are used for linear elements $(3+3+3$ unknown parameters) and complete quadratic polynomials for quadratic elements $(6+6+6$ unknown parameters).

The local stress approximations used in the RCP procedure for Class II problems are:

(a) linear 7-parameters approximation used in the RCP1 version for 4-node elements

$$\boldsymbol{\sigma}^*_{ph} = \begin{bmatrix} 1 & 0 & 0 & y & 0 & x & 0 \\ 0 & 1 & 0 & 0 & x & 0 & y \\ 0 & 0 & 1 & 0 & 0 & -y & -x \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_7 \end{bmatrix}, \tag{A.22}$$

(b) quadratic 12-parameters approximation used in the RCP2 version for 4-node elements and in the RCP1 version for 8- and 9-node elements

$$\boldsymbol{\sigma}^*_{ph} = \begin{bmatrix} 1 & 0 & 0 & y & 0 & x & 0 & y^2 & 0 & x^2 & 2xy & 0 \\ 0 & 1 & 0 & 0 & x & 0 & y & 0 & x^2 & y^2 & 0 & 2xy \\ 0 & 0 & 1 & 0 & 0 & -y & -x & 0 & 0 & -2xy & -y^2 & -x^2 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_{12} \end{bmatrix}, \tag{A.23}$$

(c) cubic 18-parameters approximation used in the RCP2 version for 8- and 9-node elements

$$\boldsymbol{\sigma}^*_{ph} = \begin{bmatrix} 1 & 0 & 0 & y & 0 & x & 0 & y^2 & 0 & x^2 & 2xy & 0 & y^3 & 0 \\ 0 & 1 & 0 & 0 & x & 0 & y & 0 & x^2 & y^2 & 0 & 2xy & 0 & x^3 \\ 0 & 0 & 1 & 0 & 0 & -y & -x & 0 & 0 & -2xy & -y^2 & -x^2 & 0 & 0 \end{bmatrix}$$
$$\begin{bmatrix} x^3 & 3xy^2 & 0 & 3x^2y \\ 3xy^2 & 0 & 3x^2y & y^3 \\ -3x^2y & -y^3 & -x^3 & -3xy^2 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_{18} \end{bmatrix}. \tag{A.24}$$

A.3 Class III problems

The third class of problems is shear deformable plates.

A.3.1 Basic equations

Consider a plate referred to a Cartesian reference frame (O, x, y, z) with the origin O on the mid-surface Ω and the z-axis in the thickness direction, $-h/2 \leq z \leq h/2$

where h is the thickness.

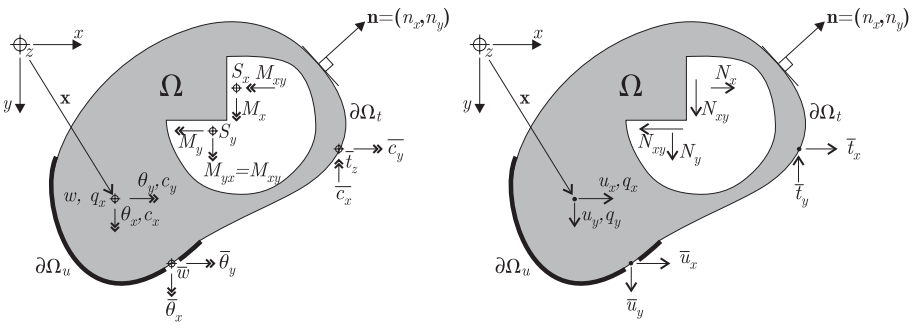


Figure A.3: Class III problems

Let $\partial\Omega$ be the boundary of Ω. The Reissner-Mindlin theory is employed. The compatibility equations are given by

$$\boldsymbol{\mu} = \mathbf{D}_p\mathbf{u}, \quad \boldsymbol{\chi} = \mathbf{D}_p\boldsymbol{\theta}, \quad \boldsymbol{\gamma} = \mathbf{D}_s w + \boldsymbol{\theta}, \tag{A.25}$$

where $\mathbf{u}^{\mathrm{T}} = [u_x, u_y]$ is the in plane displacement, w is the transverse displacement, $\boldsymbol{\theta}^{\mathrm{T}} = \left[\begin{array}{cc} \theta_x & \theta_y \end{array}\right]$ is the vector of rotations, $\boldsymbol{\mu}^{\mathrm{T}} = \left[\begin{array}{ccc} \epsilon_x & \epsilon_y & \gamma_{xy} \end{array}\right]$ is the vector of in-plane strains, $\boldsymbol{\chi}^{\mathrm{T}} = \left[\begin{array}{ccc} \chi_x & \chi_y & \chi_{xy} \end{array}\right]$ is the vector of curvatures, $\boldsymbol{\gamma}^{\mathrm{T}} = \left[\begin{array}{cc} \gamma_x & \gamma_y \end{array}\right]$ is the vector of transverse shear strains and

$$\mathbf{D}_b = \left[\begin{array}{cc} \partial/\partial x & 0 \\ 0 & \partial/\partial y \\ \partial/\partial y & \partial/\partial x \end{array}\right], \quad \mathbf{D}_s = \left[\begin{array}{c} \partial/\partial x \\ \partial/\partial y \end{array}\right]. \tag{A.26}$$

The equilibrium equations can be obtained *via* the principle of virtual work in the form

$$\mathbf{D}_p^*\mathbf{N} = \mathbf{q}_x, \quad \mathbf{D}_p^*\mathbf{M} + \mathbf{S} = \mathbf{c}, \quad \mathbf{D}_s^*\mathbf{S} = q_z, \tag{A.27}$$

where vectors $\mathbf{N}^{\mathrm{T}} = \left[\begin{array}{ccc} N_x & N_y & N_{xy} \end{array}\right]$, $\mathbf{M}^{\mathrm{T}} = \left[\begin{array}{ccc} M_x & M_y & M_{xy} \end{array}\right]$ and $\mathbf{S}^{\mathrm{T}} = \left[\begin{array}{cc} S_x & S_y \end{array}\right]$ collect the resultant stresses, $\mathbf{D}_p^*$ and $\mathbf{D}_s^*$ are differential operators adjoint to $\mathbf{D}_p$ and $\mathbf{D}_s$, respectively, and $\mathbf{q}_z$, $\mathbf{q}_x$, and $\mathbf{c}$ are the prescribed generalized loads.

The constitutive equations can be written in the following form typical of monoclinic laminated plates:

$$\mathbf{N} = \mathbf{C}_m\boldsymbol{\mu} + \mathbf{C}_{mb}\boldsymbol{\chi}, \tag{A.28}$$

$$\mathbf{M} = \mathbf{C}_{mb}\boldsymbol{\mu} + \mathbf{C}_b\boldsymbol{\chi}, \tag{A.29}$$

$$\mathbf{S} = \mathbf{C}_s\boldsymbol{\gamma}. \tag{A.30}$$

where $\mathbf{C}_m$, $\mathbf{C}_{mb}$, and $\mathbf{C}_b$ are respectively the membrane, membrane-bending coupling and bending elasticity matrices, whereas $\mathbf{C}_s$ represents the resultant elastic shear matrix.

For an isotropic linearly elastic material, the constitutive equations reduces to

$$\mathbf{N} = \mathbf{C}_m \boldsymbol{\mu}, \quad \mathbf{M} = \mathbf{C}_b \boldsymbol{\chi}, \quad \mathbf{S} = \mathbf{C}_s \boldsymbol{\gamma}, \tag{A.31}$$

$$\mathbf{C}_m = \frac{Eh}{(1-\nu^2)} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & (1-\nu)/2 \end{bmatrix}, \tag{A.32}$$

$$\mathbf{C}_b = \frac{Eh^3}{12(1-\nu^2)} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & (1-\nu)/2 \end{bmatrix}, \tag{A.33}$$

$$\mathbf{C}_s = \frac{\kappa Eh}{2(1+\nu)} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{A.34}$$

being E the Young's modulus and ν the Poisson's ratio and h is the plate thickness and κ the shear correction factor, ($\kappa = \frac{5}{6}$ if not otherwise specified).

Finally, the plate equilibrium problem can be stated once the boundary conditions on $\partial\Omega_u$ on $\partial\Omega$ are specified.

A.3.2 A plate-like body

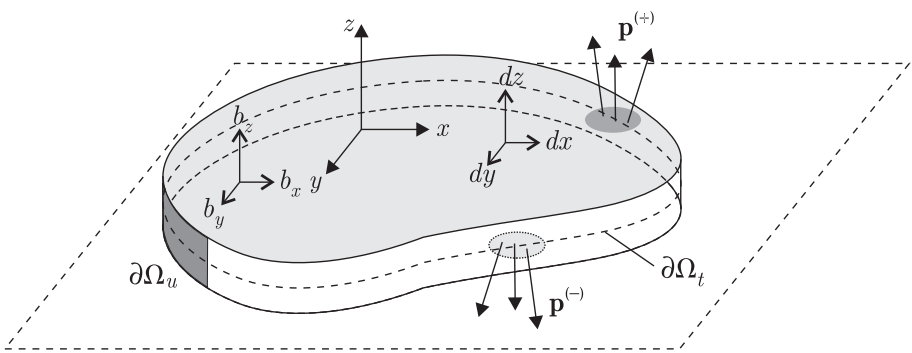


Figure A.4: A 3D plate-like body

Consider the 3D plate-like body in Figure A.4, where volume forces are indicated by $\mathbf{b}$ and surface tractions on top and bottom surfaces by $\mathbf{p}^{(+)}$ and $\mathbf{p}^{(-)}$.

In accordance with the Reissners-Mindlin plate theory, the displacement field

$\mathbf{d}$ is represented as a linear function of the thickness coordinate z in the form

$$\mathbf{d}\left(\mathbf{x},z\right)=\left[\begin{array}{c}\mathbf{u}\left(\mathbf{x}\right)+z\boldsymbol{\theta}\left(\mathbf{x}\right)\\ w\left(\mathbf{x}\right)\end{array}\right], \tag{A.35}$$

where the notation introduced in the previous section is used to indicate quantities defined on the plate mid-plane Ω.

The strain tensor $\boldsymbol{\epsilon}$ can be decomposed as follows

$$\boldsymbol{\varepsilon}=\boldsymbol{\nabla}^{(s)}\mathbf{d}=\left[\begin{array}{cc}\mathbf{e} & \frac{1}{2}\boldsymbol{\gamma}\\ \frac{1}{2}\boldsymbol{\gamma}^{\mathrm{T}} & \mathbf{0}\end{array}\right] \tag{A.36}$$

where $\mathbf{e}$ is the in-plane strain tensor and $\boldsymbol{\gamma}$ the transverse shear vector. Using the matrix notation, they can be written as:

$$\mathbf{e}=\left[\begin{array}{c}\epsilon_x\\ \epsilon_y\\ \gamma_{xy}\end{array}\right]=\boldsymbol{\mu}+z\boldsymbol{\chi},\quad \boldsymbol{\gamma}=\mathbf{D}_s\mathbf{w}+\boldsymbol{\theta}. \tag{A.37}$$

Analogously, the stress tensor $\boldsymbol{\sigma}$ can be separated into the in-plane stress tensor $\mathbf{S}$ and the transversal shear stress vector $\boldsymbol{\tau}$. As usual the out of plane normal stress is assumed to be null, trough it could be regarded as the "reactive" stress corresponding to the internal constant $\epsilon_z=0$ [Lembo & Podio-Guidugli, 2007].

The in-plane equilibrium equations are given by

$$\nabla_x\mathbf{s}+\boldsymbol{\nabla}_z\boldsymbol{\tau}+\mathbf{b}_x=\mathbf{0},\quad \mathbf{b}_x=\left[\begin{array}{c}b_x\\ b_y\end{array}\right], \tag{A.38}$$

or, in matrix notation, by

$$\mathbf{D}_p^*\mathbf{S}+\boldsymbol{\nabla}_z\boldsymbol{\tau}+\mathbf{b}_x=\mathbf{0},\quad \mathbf{S}=\left[\begin{array}{c}\boldsymbol{\sigma}_x\\ \boldsymbol{\sigma}_y\\ \boldsymbol{\tau}_{xy}\end{array}\right]. \tag{A.39}$$

Introducing the stress resultants

$$\mathbf{N}=\int_f\mathbf{s}\mathrm{d}z,\quad \mathbf{M}=\int_f z\mathbf{S}\mathrm{d}z,\quad \mathbf{S}-\int_f\boldsymbol{\tau}\mathrm{d}z, \tag{A.40}$$

the plate equilibrium equations can be derived in the form of Equation (A.31) where

$$\mathbf{q}_x=\int_f\mathbf{b}_x\mathrm{d}z+\left(\mathbf{p}_x^{(+)}+\mathbf{p}_x^{(-)}\right),$$

$$\mathbf{q}_z=\int_f\mathbf{b}_z\mathrm{d}z+\left(\mathbf{p}_z^{(+)}+\mathbf{p}_z^{(-)}\right),$$

$$\mathbf{c}=\int_f z\mathbf{b}_x\mathrm{d}z+\left(\mathbf{p}_x^{(+)}-\mathbf{p}_x^{(-)}\right)\frac{h}{2}.$$

Introducing the constitutive equations and taking into account for the above assumptions, the stress-strain relations can be written as

$$\mathbf{s}=\mathbf{C}_p\mathbf{e},\quad \boldsymbol{\tau}=\boldsymbol{\kappa}\odot\mathbf{C}_s\boldsymbol{\gamma}, \tag{A.41}$$

where $\mathbf{C}_p$ is the reduced in plane elastic matrix, $\mathbf{C}_s$ is the transverse shear elastic

matrix, matrix $\boldsymbol{\kappa}$ collects the shear correction factors

$$\boldsymbol{\kappa} = \begin{bmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{12} & \kappa_{22} \end{bmatrix}, \tag{A.42}$$

assumed to be constant along the plate thickness, symbol $\odot$ denotes a product component by component. In the case of homogeneous plates: $\kappa_{11} = \kappa_{22} = \frac{5}{6}$ and $\kappa_{12} = 0$.

By interpolating the stress-strain relations through the thickness, the plate constitutive equations are defined

$$\mathbf{N} = \mathbf{C}_m \boldsymbol{\mu} + \mathbf{C}_{mb} \boldsymbol{\chi}, \tag{A.43}$$

$$M = \mathbf{C}_{mb} \boldsymbol{\mu} + \mathbf{C}_b \boldsymbol{\chi}, \tag{A.44}$$

$$S = \mathbf{C}_s \boldsymbol{\gamma}. \tag{A.45}$$

A.3.3 The RCP stress approximation

The RCP stress approximation is selected as follows. The part which equilibrates external loads within the element is selected as

$$\mathbf{M}^*_{pp} = \begin{bmatrix} \int_0^x \left[c_x + \frac{1}{2} \int_0^x q_z \mathrm{d}x \right] \mathrm{d}x \\ \int_0^x \left[c_y + \frac{1}{2} \int_0^x q_z \mathrm{d}y \right] \mathrm{d}y \\ 0 \end{bmatrix} \tag{A.46}$$

$$\mathbf{S}^*_{pp} = -\frac{1}{2} \begin{bmatrix} \int_0^x q_z \mathrm{d}x \\ \int_0^y q_z \mathrm{d}y \end{bmatrix} \tag{A.47}$$

Two local approximations are tested for the sef-equilibrated part: a linear approximation ruled by 9 parameters

$$\begin{bmatrix} \mathbf{M}^*_{ph} \\ \mathbf{S}^*_{ph} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & y & 0 & x & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & x & 0 & y & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & y & x \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_9 \end{bmatrix}, \tag{A.48}$$

and a quadratic approximation ruled by 17 parameters

$$\begin{bmatrix} \mathbf{M}^*_{ph} \\ \mathbf{S}^*_{ph} \end{bmatrix} = \left[\begin{array}{ccccccccc|cccc} 1 & 0 & 0 & y & 0 & x & 0 & 0 & 0 & xy & 0 & y^2 & 0 \\ 0 & 1 & 0 & 0 & x & 0 & y & 0 & 0 & 0 & xy & 0 & x^2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & y & x & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & -y & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & -x & 0 & 0 \end{array}\right] \begin{bmatrix} x^2 & 0 & 0 & 0 \\ 0 & y^2 & 0 & 0 \\ -xy & -xy & x^2 & y^2 \\ -x & x & 0 & -2\,y \\ y & -y & -2\,x & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_{17} \end{bmatrix}. \quad \text{(A.49)}$$

Membrane forces can be approximated as described in Section A.2.2.

Appendix B
More benchmark tests

B.1 Class I problems

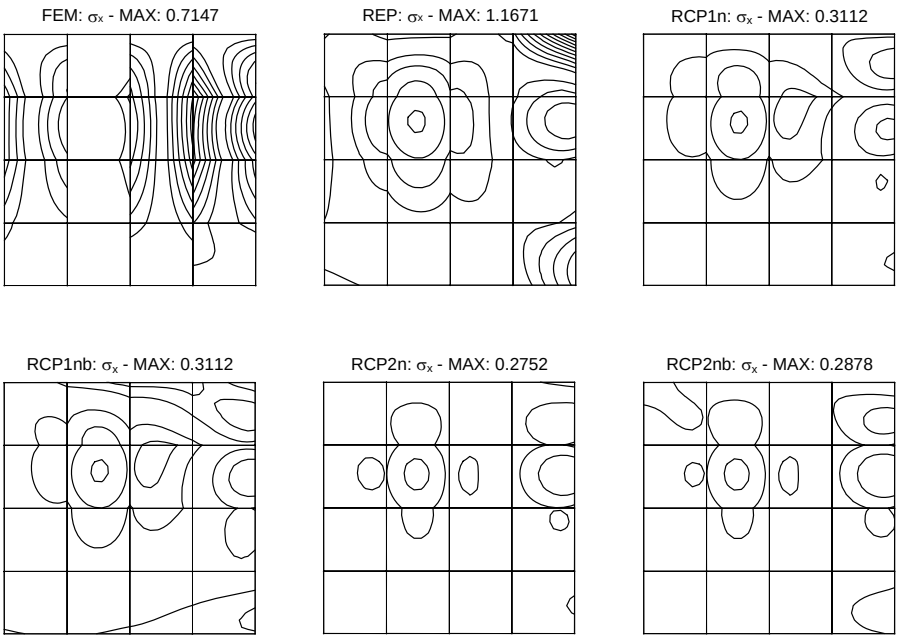


Figure B.1: Thin membrane subjected to transverse loading - Error distributions on a regular mesh of four-node elements (contour interval = 0.1, with load case No. 1)

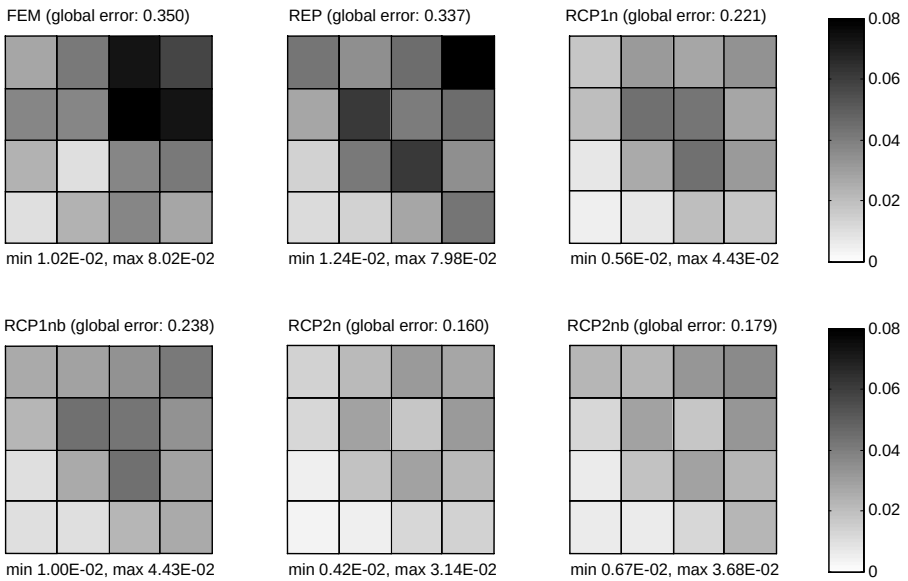


Figure B.2: Thin membrane subjected to transverse loading - Distributions of element error in energy norm on a regular mesh of four-node elements, load case No. 1 (minimum and maximum values are reported below each plot, the global error in energy norm is reported above each plot)

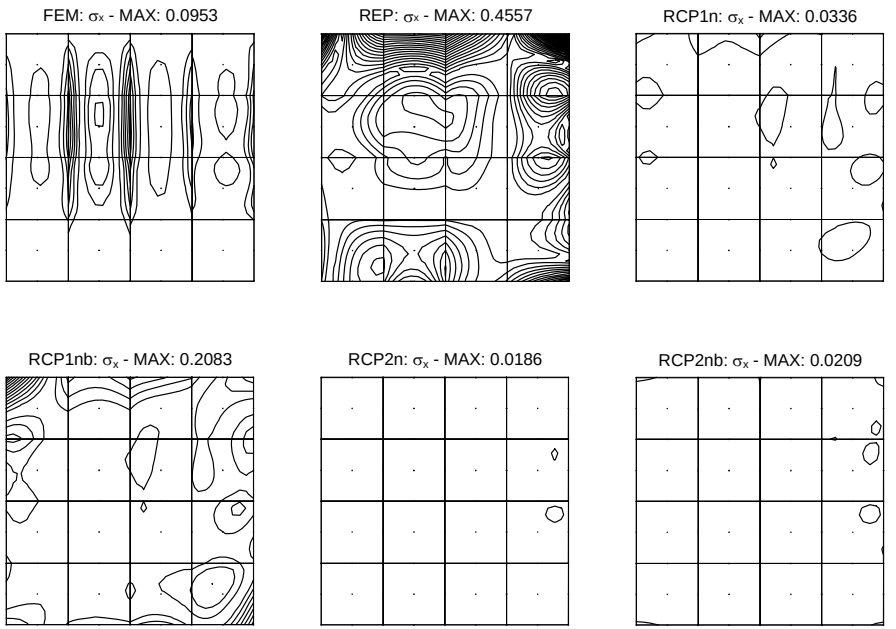


Figure B.3: Thin membrane subjected to transverse loading - Error distributions on a regular mesh of nine-node elements (contour interval = 0.015, load case No. 1)

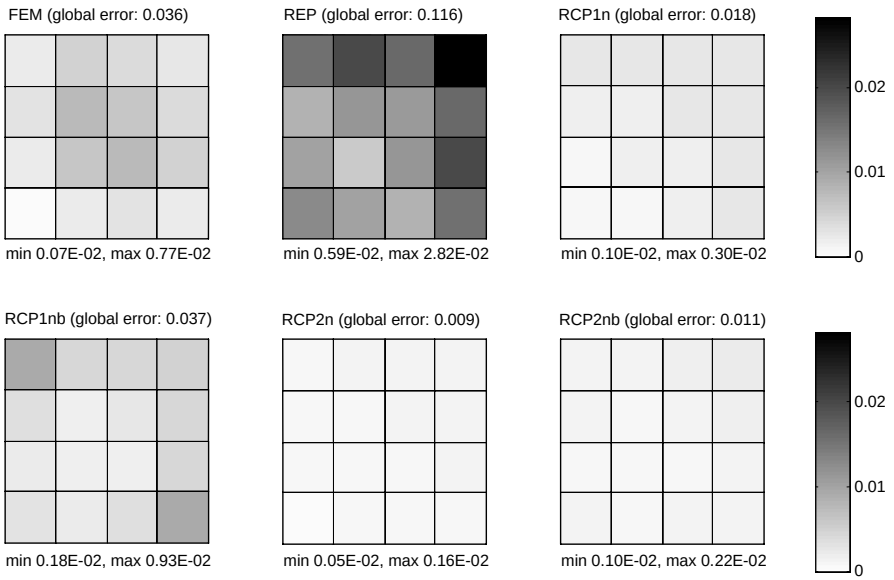


Figure B.4: Thin membrane subjected to transverse loading - Distributions of element error in energy norm on a regular mesh of nine-node elements, load case No. 1 (minimum and maximum values are reported below each plot, the global error in energy norm is reported above each plot)

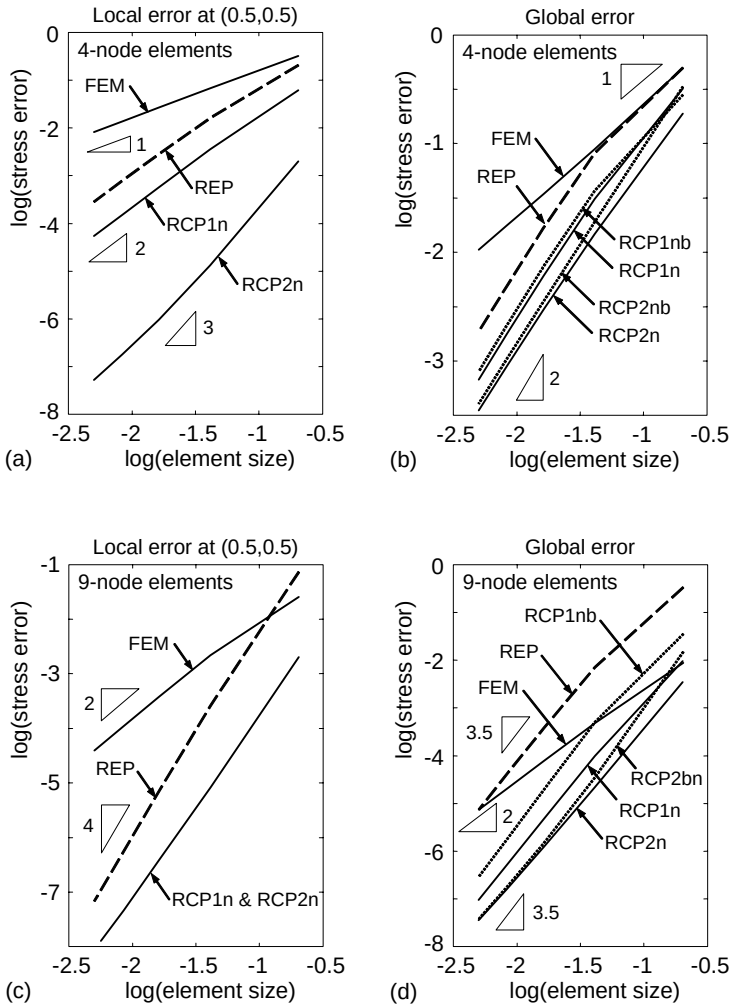


Figure B.5: Thin membrane subjected to transverse loading - Rate of convergence for load case No. 1: (a) local error at point $x = 0.5$, $y = 0.5$ ($\sigma_x = \sigma_y$) for four-node elements, (b) global error in energy norm for four-node elements, (c) local error at point $x = 0.5$, $y = 0.5$ ($\sigma_x = \sigma_y$) for nine-node elements, (d) global error in energy norm for nine-node elements

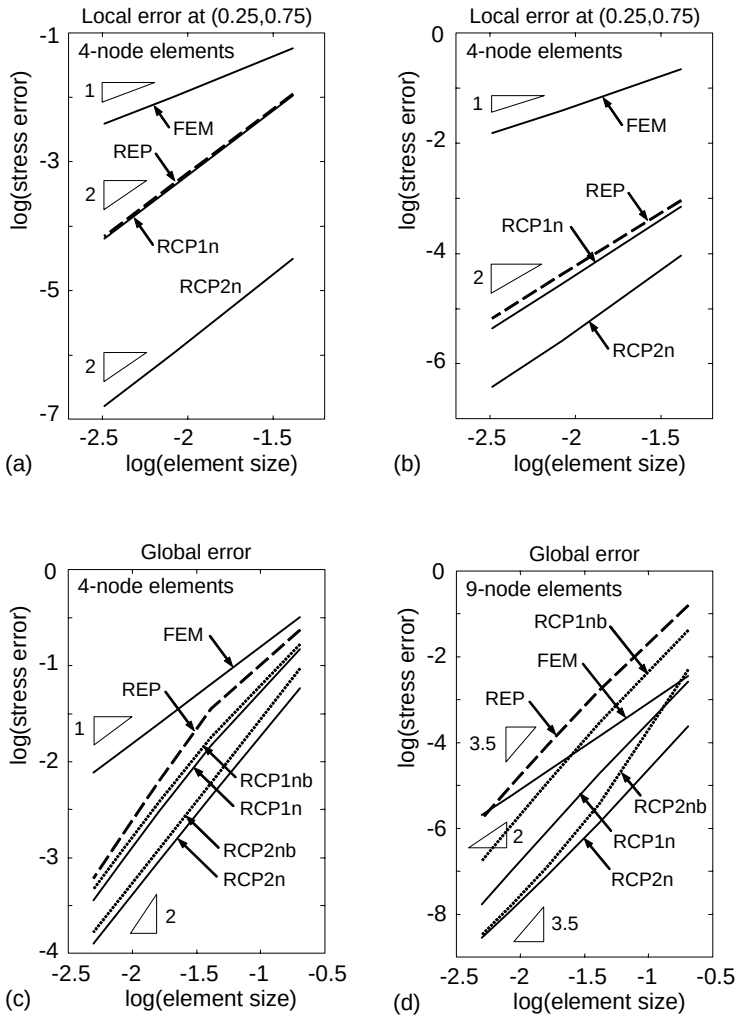


Figure B.6: Thin membrane subjected to transverse loading - Rate of convergence for load case No. 2: (a) local error in σ_x at point $x = 0.25$, $y = 0.75$ for four-node elements, (b) local error in σ_y at point $x = 0.25$, $y = 0.75$ for four-node elements, (c) global error in energy norm for four-node elements, (d) global error in energy norm for nine-node elements

B.2 Class II problems

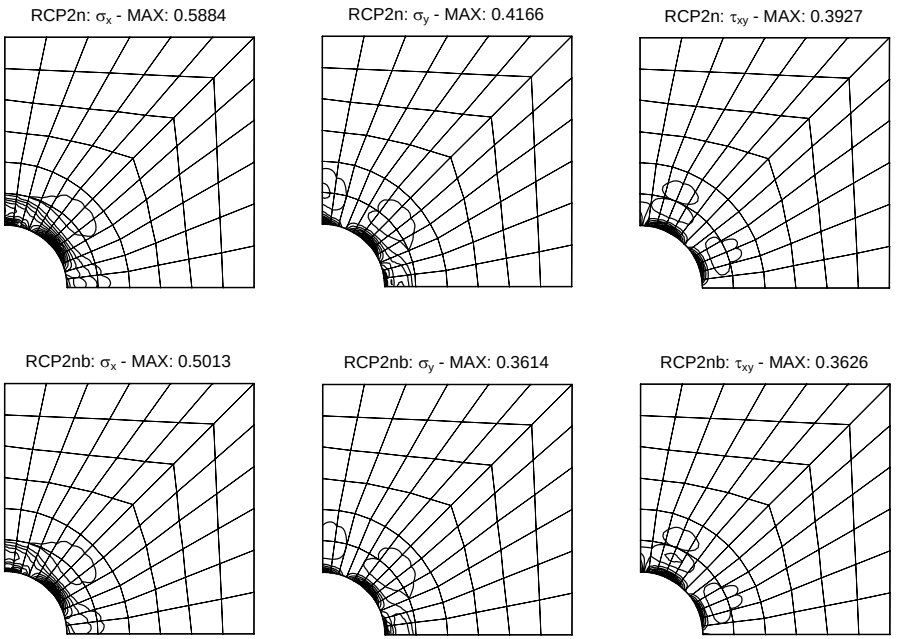


Figure B.7: Stretched plate with circular hole - Error distributions for four-node elements on mesh No. 1: RCP2 and RCP2b (contour interval = 0.05)

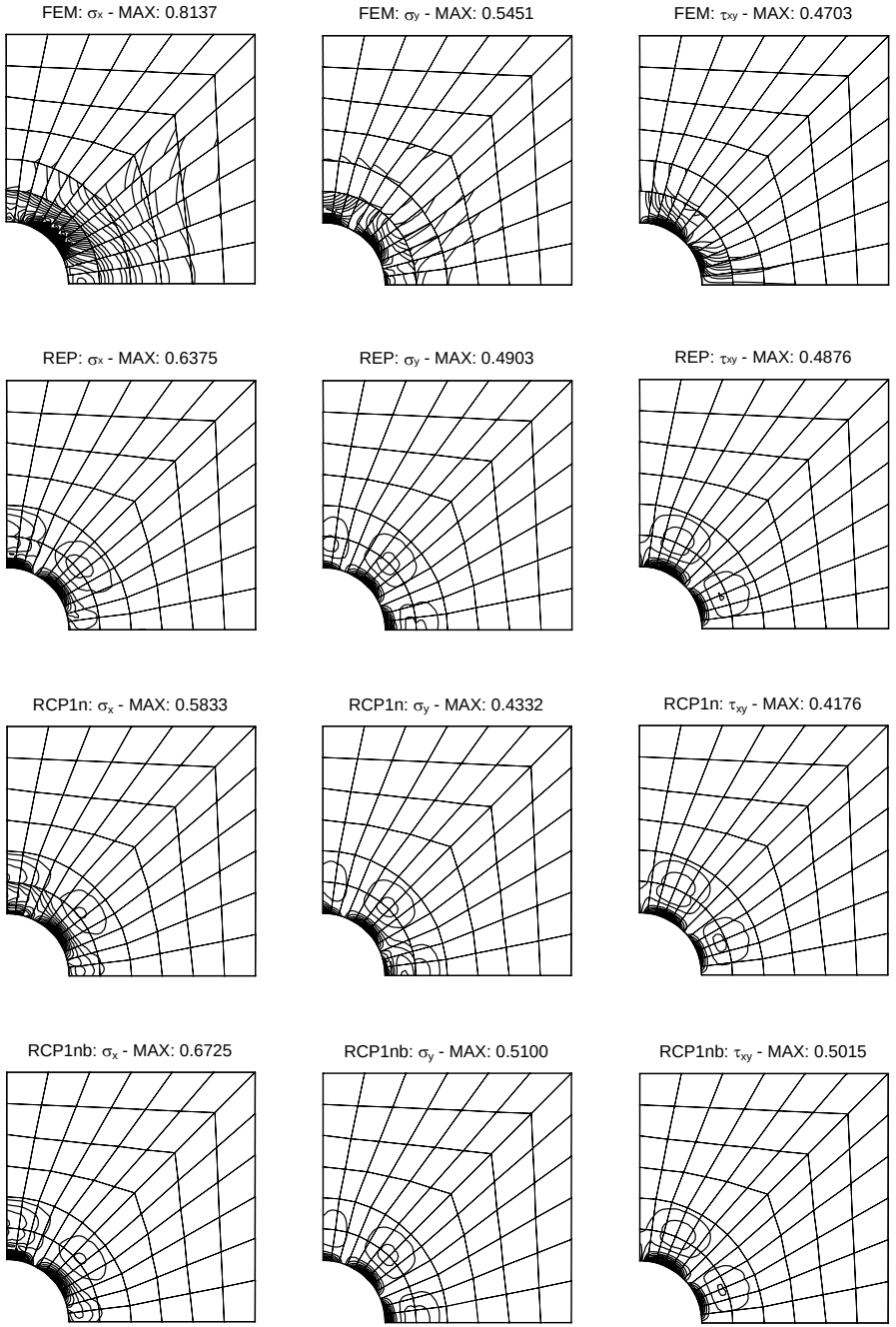


Figure B.8: Stretched plate with circular hole - Error distributions for four-node elements on mesh No. 1: FE solution, REP, RCP1 and RCP1b (contour interval = 0.05)

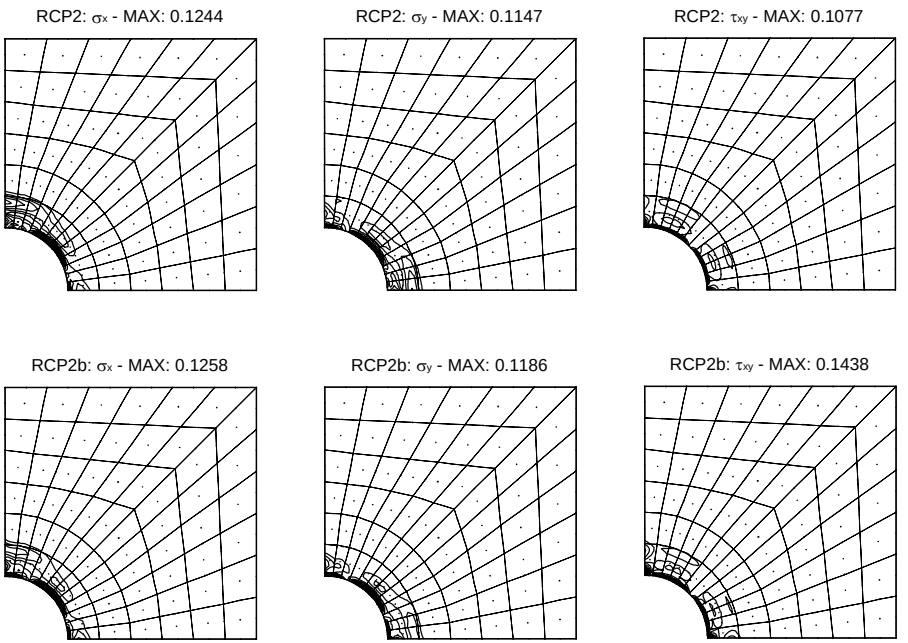


Figure B.9: Stretched plate with circular hole - Error distributions for nine-node elements on mesh No. 1: RCP2 and RCP2b (contour interval = 0.05)

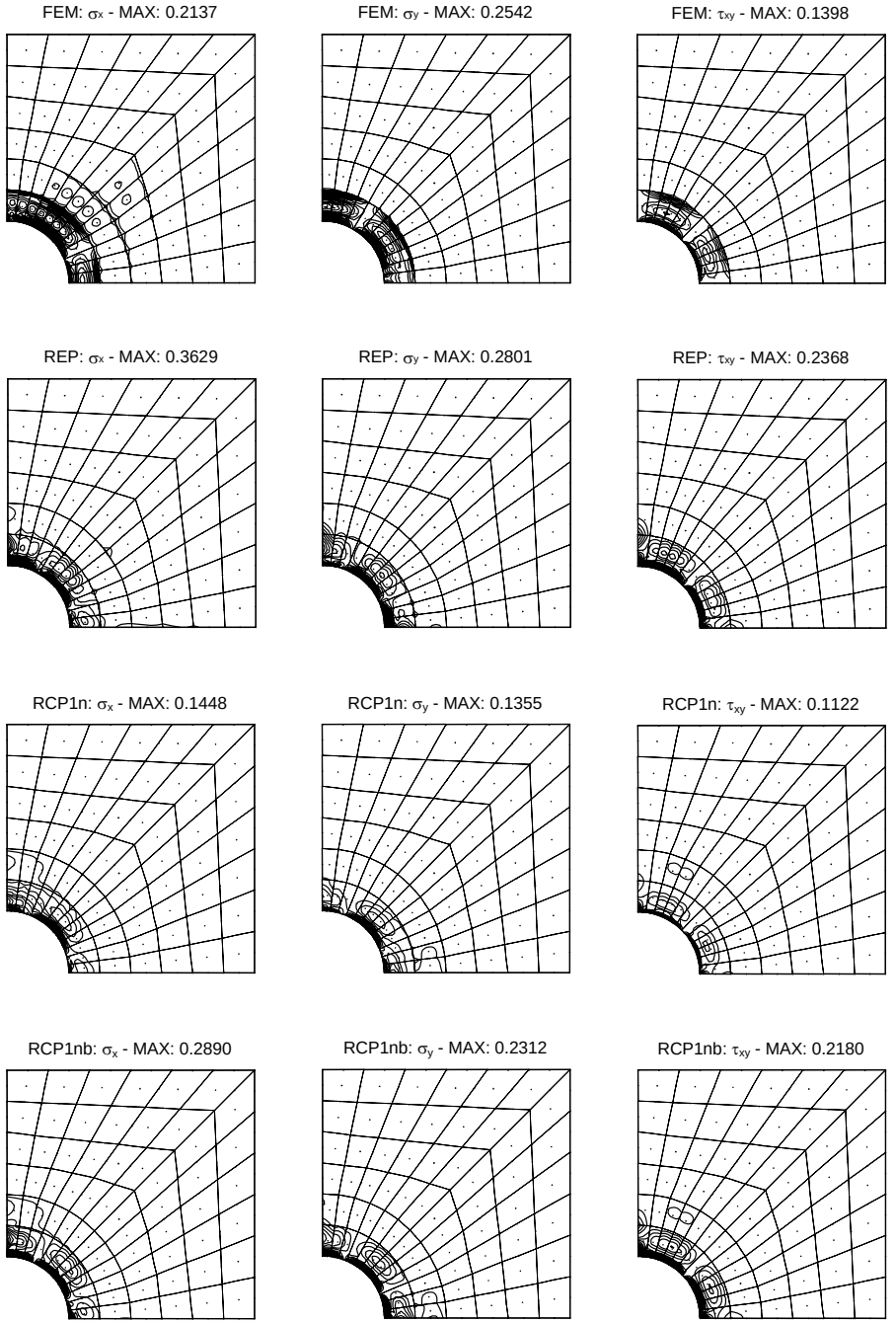


Figure B.10: Stretched plate with circular hole - Error distributions for nine-node elements on mesh No. 1: FE solution, REP, RCP1 and RCP1b (contour interval = 0.05)

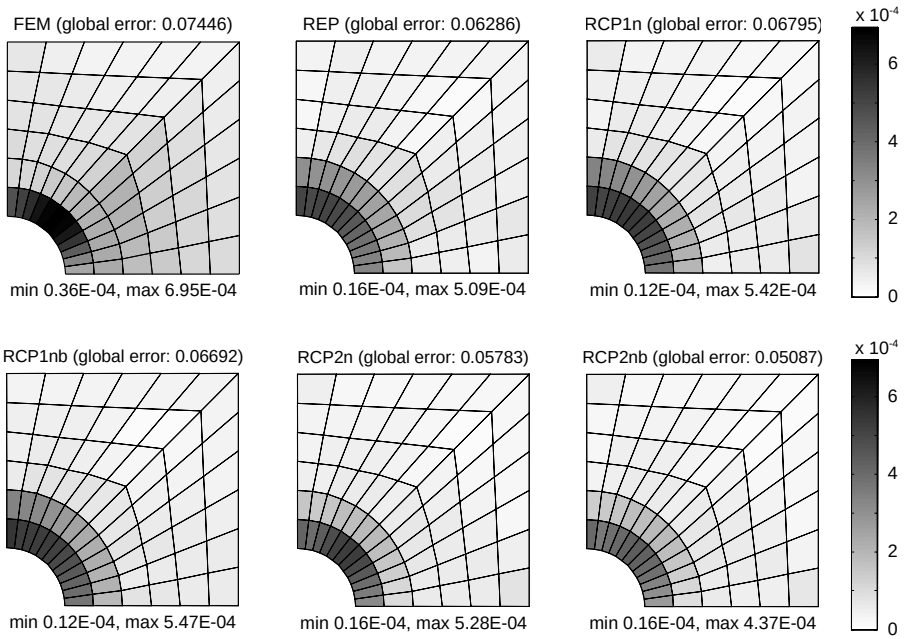


Figure B.11: Stretched plate with circular hole - Distributions of element error in energy norm on mesh No. 1: four-node elements (minimum and maximum values are reported below each plot, the global error in energy norm is reported above each plot)

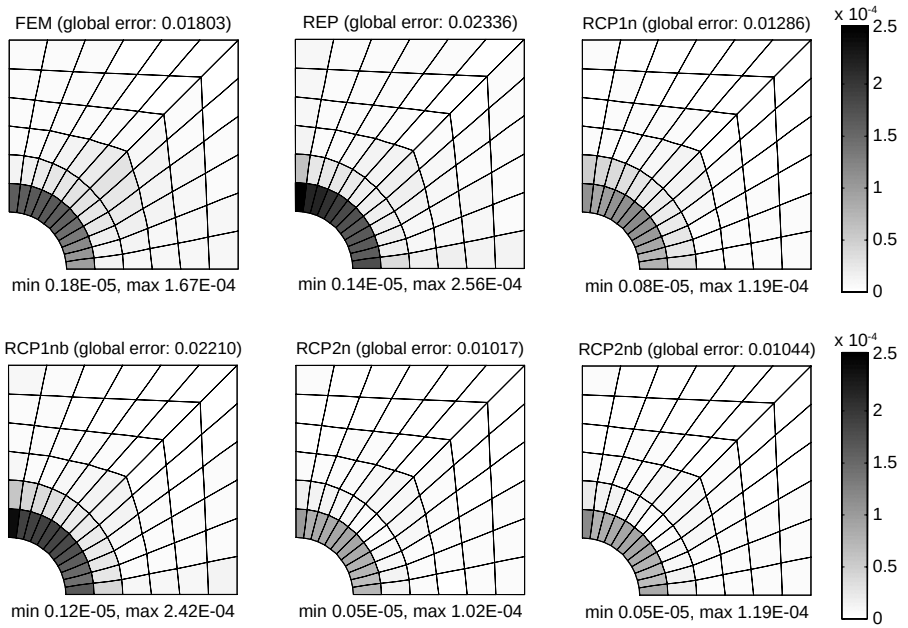


Figure B.12: Stretched plate with circular hole - Distributions of element error in energy norm on mesh No. 1: nine-node elements (minimum and maximum values are reported below each plot, the global error in energy norm is reported above each plot)

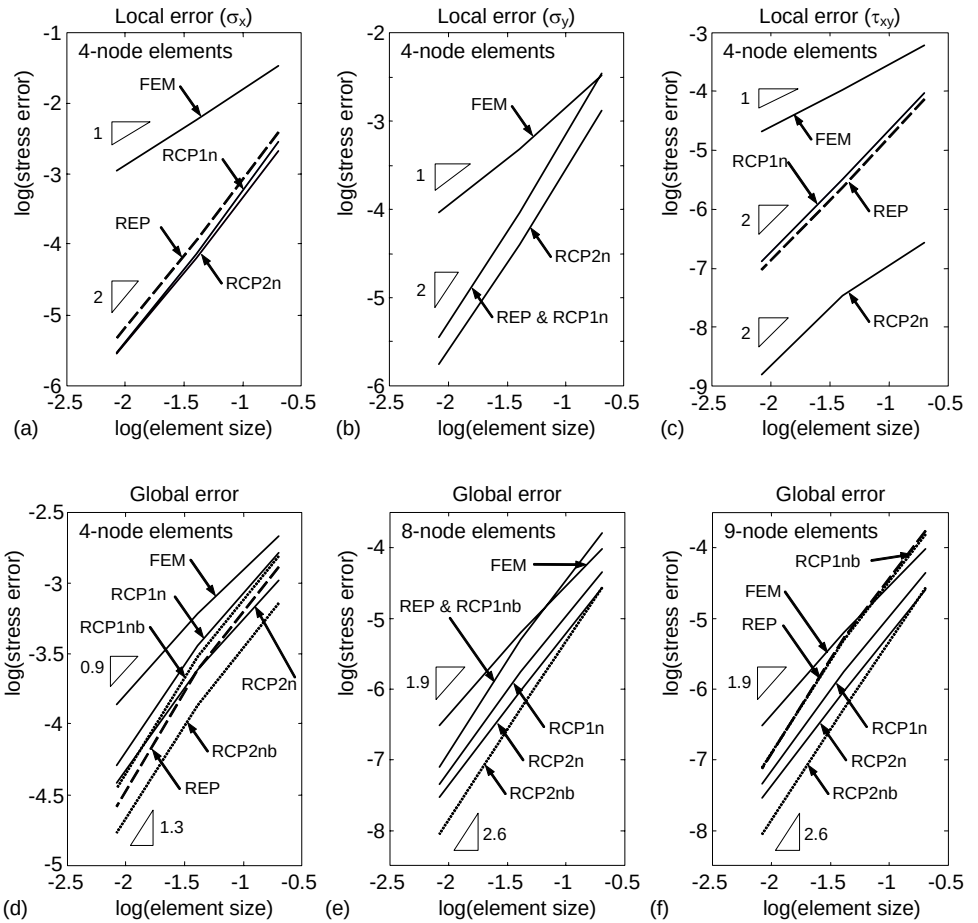


Figure B.13: Stretched plate with circular hole - Rate of convergence: (a) local error in σ_x at point $r = 1.5$, $\theta = \pi/4$ for four-node elements, (b) local error in σ_y at point $r = 1.5$, $\theta = \pi/4$ for four-node elements, (c) local error in τ_{xy} at point $r = 1.5$, $\theta = \pi/4$ for four-node elements, (d) global error in energy norm for four-node elements, (e) global error in energy norm for eight-node elements, (f) global error in energy norm for nine-node elements

Appendix C
The hybrid stress element for plate structures (9βQ4)

Recently, a new quadrilateral 4-node element for the analysis of shear deformable plates has been presented in [de Miranda & Ubertini, 2006]–[Daghia *et al.* , 2007] and labelled 9βQ4. The formulation is of hybrid stress type, involving equilibrating stress resultants within each element and compatible displacements along the interelement boundary. The assumed stress approximation, expressed in terms of skew coordinates, has the minimum number of stress modes ($9 + 5$ modes) and is coordinate invariant. Displacements are modelled using a linked interpolation ruled by the standard nodal degrees of freedom ($3 + 2$ dofs per node). The hybrid stress approach, where displacement acts on element boundary only, together with the linked interpolation allow to derive an element which is locking-free and passes all the patch tests. Moreover, the resulting element has been tested to be stable, accurate and relatively insensitive to geometry distortions.

The main features of the element are:

(1) it has 4-nodes ($12 + 8$ dofs) and involves only compatible displacement functions,

(2) it is locking-free,

(3) it passes all the patch tests,

(4) it is readily implementable into existing finite element codes.

The key ingredients of the element formulation are:

- a mixed variational approach involving stresses and displacements as independent variables, which allow to meet requirements (1) and (4);
- a linked interpolation for the transverse displacement field, which goes towards avoiding locking effects but it does not suffice to remove them;
- the stress resultants are assumed to satisfy a priori the plate equilibrium equations within each element.

It can be proved that the last ingredient together with the previous ones allows to meet also requirement (2) and (3).

The assumed approximation for stress resultants is

$$\begin{bmatrix} \mathbf{N}^h \\ \mathbf{M}^h_h \\ \mathbf{S}^h_h \end{bmatrix} = \begin{bmatrix} \mathbf{N}^h_p \\ \mathbf{M}^h_p \\ \mathbf{S}^h_p \end{bmatrix} + \begin{bmatrix} \mathbf{N}^{hh} \\ \mathbf{M}^h_h \\ \mathbf{S}^h_h \end{bmatrix} \tag{C.1}$$

where the first term is particular solution of the plate equilibrium equations, while the second term is the self equilibrated component which is approximated as follows

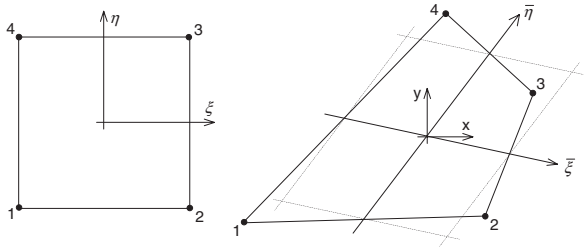


Figure C.1: Reference coordinate system and their coordinate axes

$$\left[\mathbf{N}^h\right] = \begin{bmatrix} 1 & 0 & 0 & a_1^2\bar{\eta} & a_3^2\bar{\xi} \\ 0 & 1 & 0 & b_1^2\bar{\eta} & b_3^2\bar{\xi} \\ 0 & 0 & 1 & a_1b_1\bar{\eta} & a_3b_3\bar{\xi} \end{bmatrix} \boldsymbol{\beta}_N, \tag{C.2}$$

$$\begin{bmatrix} \mathbf{M}_h^h \\ \mathbf{S}_h^h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & a_1^2\bar{\eta} & a_3^2\bar{\xi} & -a_1\bar{\xi} - a_3\bar{\eta} & 0 & -a_1^2\bar{\xi}\bar{\eta} & -a_3^2\bar{\xi}\bar{\eta} \\ 0 & 1 & 0 & b_1^2\bar{\eta} & b_3^2\bar{\xi} & 0 & -b_1\bar{\xi} - b_3\bar{\eta} & -b_1^2\bar{\xi}\bar{\eta} & -b_3^2\bar{\xi}\bar{\eta} \\ 0 & 0 & 1 & a_1b_1\bar{\eta} & a_3b_3\bar{\xi} & 0 & 0 & -a_1b_1\bar{\xi}\bar{\eta} & -a_3b_3\bar{\xi}\bar{\eta} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & a_1\bar{\eta} & a_3\bar{\xi} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & b_1\bar{\eta} & b_3\bar{\xi} \end{bmatrix} \tag{C.3}$$

Coefficients a_i and b_i in the above expressions depend on the element geometry:

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \end{bmatrix}. \tag{C.4}$$

Appendix D
Formulation of a transition element for plane problems

D.1 Shape functions

The typical configuration of a variable-node element is shown in Figure D.1. The midside nodes need to be generated when the refined elements are attached to the side of the unrefined element. In particular a quadrilateral element with a bilinear displacement field is considered but, the method is general and can be easily applied to other elements.

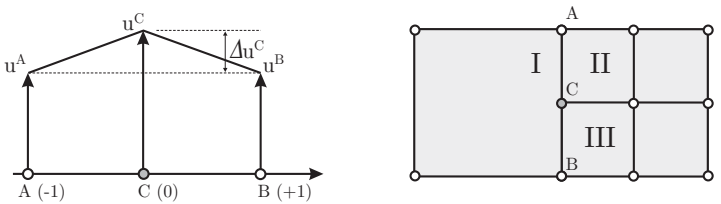


Figure D.1: The transition quadrilateral element and displacement field

In Figure D.1 edge AB of element I has discontinuity in the displacement slope at node C. The variation of any displacement u along edge AB can be expressed as

$$u = \frac{1}{2}(1-\xi)\,u^A + \frac{1}{2}(1+\xi)\,u^B + \frac{1}{2}(1-|\xi|)\,\Delta u^C, \tag{D.1}$$

where u^A and u^B are displacements at nodes A and B, respectively, and Δu^C is the departure of displacement at C from the straight variation defined by the displacements at A and B; ξ is the local coordinate along AB, $\xi = -1$ at A, $\xi = +1$ at B, $\xi = 0$ at C; $|\xi|$ denotes the absolute value of ξ. Equation D.1 gives

$$u^C = \frac{u^A + u^B}{2} + \Delta u^C, \tag{D.2}$$

or

$$\Delta u^C = u^C - \frac{u^A + u^B}{2}. \tag{D.3}$$

Substituting Equation (D.3) into Equation (D.1) one gets

$$u = \frac{1}{2}\left(|\xi| - \xi\right) u^A + \frac{1}{2}\left(|\xi| + \xi\right) u^B + \frac{1}{2}\left(1 - |\xi|\right) u^C,$$

or

$$u = N^A u^A + N^B u^B + N^C u^C, \tag{D.4}$$

where

$$N^A = \frac{1}{2}\left(|\xi| - \xi\right), \tag{D.5}$$

$$N^B = \frac{1}{2}\left(|\xi| + \xi\right), \tag{D.6}$$

$$N^C = \frac{1}{2}\left(1 - |\xi|\right). \tag{D.7}$$

Equations (D.5-D.7) give the shape functions along an edge with a transition node. The procedure to obtain these shape function, as shown in [Gupta, 1978] can be summarized as follows:

- write the shape functions for the corner nodes as if there were no intermediate transition node;
- write the shape function for the intermediate node treating the displacement at the intermediate node as a departure Δu^C from the linear displacement variation defined by the displacements at corner nodes;
- modify the shape function at the corner nodes such that the displacement at the intermediate node becomes the total displacement at the node, u^C. Note that the shape function at the intermediate node does not change;

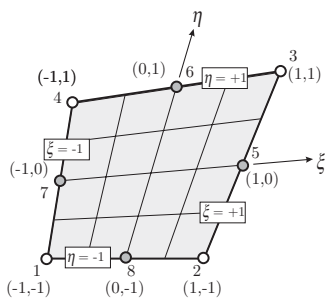


Figure D.2: The transition quadrilateral element

The above procedure was established for an edge, but the same procedure can be applied to the complete element. Figure D.2 shows a quadrilateral element with

corner nodes 1−2−3−4 and intermediate nodes 5−6−7−8. It is assumed that any intermediate node in this figure will introduce discontinuity in the displacement slope. However, it is not necessary that any or all of these intermediate nodes be present. Indeed, an element can be formulated with corner nodes only, or with one, two, three or four intermediate nodes. The procedure summarized above is applied as follows:

- write the shape functions for the corner nodes as if there were no intermediate nodes

$$\overline{N}^1 = \frac{1}{4}(1-\xi)(1-\eta), \tag{D.8}$$

$$\overline{N}^2 = \frac{1}{4}(1-\xi)(1+\eta), \tag{D.9}$$

$$\overline{N}^3 = \frac{1}{4}(1+\xi)(1-\eta), \tag{D.10}$$

$$\overline{N}^4 = \frac{1}{4}(1+\xi)(1+\eta), \tag{D.11}$$

- write the shape function for the intermediate node treating the displacement at the intermediate node as a departure from the linear displacement variation defined by the displacements at corner nodes

$$N^5 = \frac{1}{4}(1-|\xi|)(1-\eta), \tag{D.12}$$

$$N^6 = \frac{1}{4}(1-|\xi|)(1+\eta), \tag{D.13}$$

$$N^7 = \frac{1}{4}(1+\xi)(1-|\eta|), \tag{D.14}$$

$$N^8 = \frac{1}{4}(1+\xi)(1+|\eta|), \tag{D.15}$$

- if any intermediate node is not present, simply set its shape function equal to zero. For instance, if intermediate nodes 6 and 7 are not present, set $N^6 = 0$ and $N^7 = 0$. Modify the shape function at the corner nodes such that the displacement at the intermediate node becomes the total displacement at the respective nodes

$$N^1 = \overline{N}^1 - \frac{1}{2}\left(N^7 + N^8\right), \tag{D.16}$$

$$N^2 = \overline{N}^2 - \frac{1}{2}\left(N^8 + N^5\right), \tag{D.17}$$

$$N^3 = \overline{N}^3 - \frac{1}{2}\left(N^5 + N^6\right), \tag{D.18}$$

$$N^4 = \overline{N}^4 - \frac{1}{2}\left(N^6 + N^7\right), \tag{D.19}$$

Note, if nodes 6 and 7 were not present, their shape functions would be set to zero as explained above, which would automatically reflect in no change in

the shape function for node 2.

The shape functions for 5-node and 6-node elements are shown by Figures D.3 and D.4. Notice that some care should be posed on numerical integration when finite element equations are compute because of discontinuity of the shape functions. In particular a modified Gaussian quadrature rule is adopted. The patch test has been used to check the capability of the element above to represent constant strain states. All the elements pass the patch test.

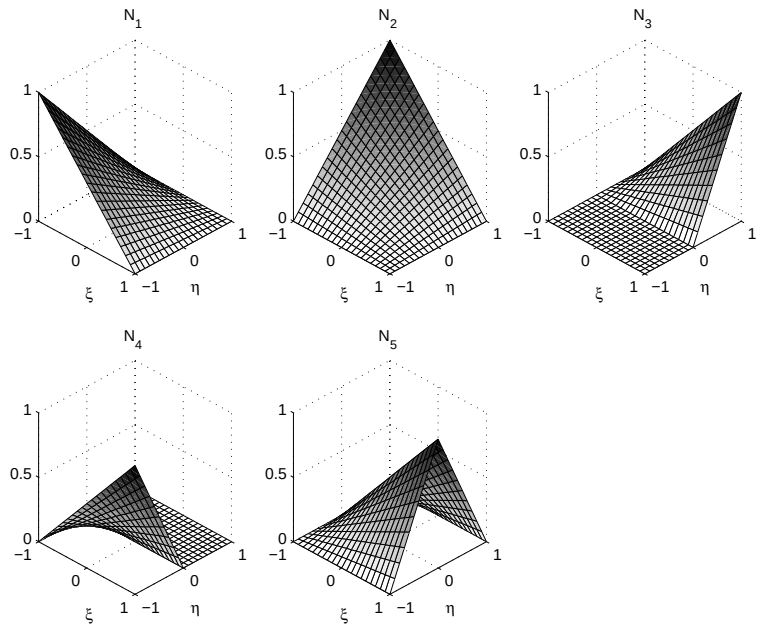


Figure D.3: The shape function for 5 mode element

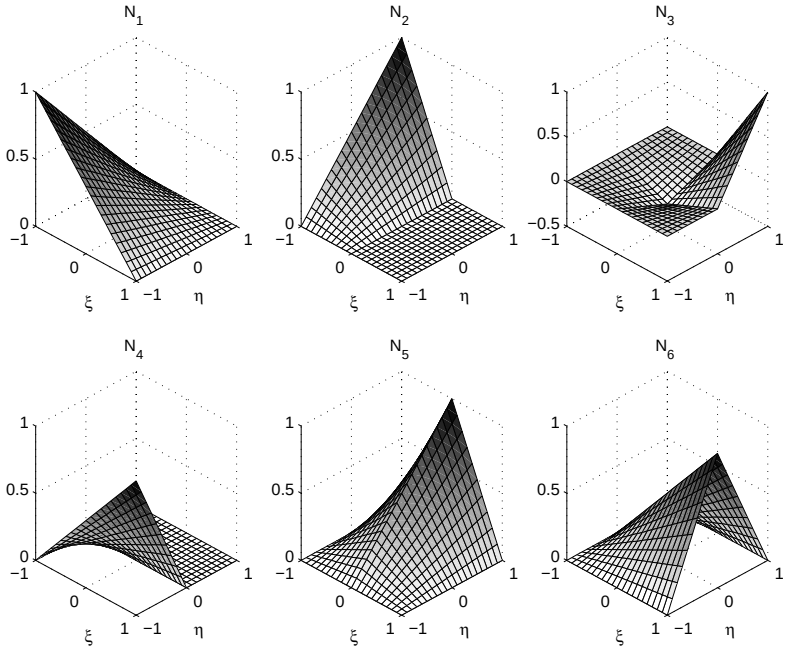


Figure D.4: The shape function for 6 mode element

Subject index

List of Tables

List of Figures

Bibliography

Ainsworth, M., & Oden, J.T. 1992. A Procedure for a Posteriori Error Estimation for $h-p$ Finite Element Methods. *Comput. Methods Appl. Mech. Engrg.*, **101**, 73–96.

Ainsworth, M., & Oden, J.T. 1993. A Unified Approach to a Posteriori Error Estimation Using Element Residual Methods. *Numerische Mathematik*, **65**, 23–50.

Ainsworth, M., & Oden, J.T. 1997. A Posteriori Error Estimation in Finite Element Analysis. *Comput. Methods Appl. Mech. Engrg.*, **142**, 1–88.

Ainsworth, M., & Oden, J.T. 2000. *A Posteriori Error Estimationn in Finite Element Analysis*. Wiley-Interscience, Jhon Wiley & Sons, Inc. New York.

Ainsworth, M., Zhu, J. Z., Craig, A. W., & Zienkiewicz, O. C. 1989. Analysis of the Zienkiewicz-Zhu a-Posteriori Error Estimator in the Finite Element Method. *Int. J. Numer. Meth. Engrg.*, **28**, 2161–2174.

Akin, J.E. 2006. *Finite Element Analysis with Error Estimators*. Elsevier.

Alfano, G, Auricchio, F, Rosati, L., & Sacco, E. 2001. MITC Finite Elements for Laminated Composite Plates. *Int. J. Numer. Meth. Engng.*, **50**, 707–738.

Almeida Pereira, O.J.B., Moitinho de Almeida, J.P., & Maunder, E.A.W. 1999. Adaptive Methods for Hybrid Equilibrium Finite Element Models. *Comput. Methods Appl. Mech. Engrg.*, **176**, 19–39.

Argyris, J.H. 1954. Energy Theorems and Structural Analysis. *Aircraft Engineering*, **26**. Part. 1 (Oct.- Nov.).

Argyris, J.H. 1955. Energy Theorems and Structural Analysis. *Aircraft Engineering*, **27**. Part 2 (Feb. - May).

Auricchio, F., Sacco, E., & Vairo, G. 2006. A Mixed FSDT Finite Element for Monoclinic Laminated Plates. *Computers and Structures*, **84**, 624–639.

Babuška, I., & Miller, A. 1984. The Post-Processing Approach in the Finite Element Method - Part 3: A-Posteriori Error Estimates and Adaptive Method Selection. *Int. J. Numer. Meth. Engng*, **20**, 2311–2325.

Babuška, I., & Miller, A. 1987. A Feedback Finite Element Method with a Posteriori Error Estiamtion: Part I. The Finite Element Method and some Basic Properties of the a Posteriori Error Estimator. *Comput. Methods Appl. Mech. Engrg.*, **61**, 1–40. Articolo cartaceo da trovare.

Babuška, I, & Oden, J.T. 2004. Verification and Validation in Computational Engineering and Science: Basic Concepts. *Comput. Methods Appl. Mech. Engrg.*, **193**, 4057–4066.

Babuška, I., & Oden, J.T. 2005. The Reliability of Computer Predictions: Can They Be Trusted? *International Journal of numerical analysis and modeling*, **1**, 1–18.

Babuška, I., & Rheinboldt, C. 1978. A-Posteriori Error Estimates for the Finite Element Method. *Int. J. Numer. Methods Engrg.*, **12**, 1597–1615.

Babuška, I., & Rheinboldt, C. 1979. Adaptive Approaches and Reliability Estimates in Finite Element Analysis. *Comput. Methods Appl. Mech. Engrg.*, **17–18**, 519–540.

Babuška, I., Strouboulis, T., & Upadhyay, C.S. 1994a. A Model Study of the Quality of a Posteriori Error Estimators for Linear Elliptic Problems. Error Estimation in the Interior of Patchwise Uniform Grids of Triangles. *Comput. Methods Appl. Mech. Engrg.*, **114**, 307–378.

Babuška, I., Strouboulis, T., Upadhyay, C.S., Gangaraj, S.K., & Copps, K. 1994b. Validation of a Posteriori Error Estimators by Numerical Approach. *Int. J. Numer. Meth. Engrg.*, **37**, 1073–1123.

Babuška, I., Strouboulis, T., & Upadhyay, C.S. 1995a. η%-Superconvergence of Finite Element Approximations in the Interior of General Meshes of Triangles. *Comput. Methods Appl. Mech. Engrg.*, **1995**, 273–305.

Babuška, I., Strouboulis, T., Upadhyay, C.S., & Gangaraj, S.K. 1995b. A Model Study of Element Residual Estimators for Linear Elliptic Problems: The Quality of the Estimators in the Interior of Meshes of Triangles and Quadrilaterals. *Computers & Structures*, **57**, 1009–1028.

Babuška, I., Strouboulis, T., & Upadhyay, C.S. 1997a. A Model Study of the Quality of a Posteriori Error Estimators for Finite Element Solutions of Linear Elliptic Problems, with Particular Reference to the Behavior Near the Boundary. *Int. J. Numer. Meth. Engng*, **40**, 2521–2577.

Babuška, I., Strouboulis, T., Gangarai, S.K., & Upadhyay, C.S. 1997b. Pollution Error in the H-Version of the Finite Element Method and the Local Quality of the Recovered Derivatives. *Comput. Methods Appl. Mech. Engrg.*, **140**, 1–37.

Babuška, I., Strouboulis, T., & Gangaraj, S.K. 1997c. A Posteriori Estimation of the Error in the Recovered Derivatives of the Finite Element Solution. *Comput. Methods Appl. Mech. Engrg.*, **150**, 369–396.

Babuška, I., Strouboulis, T., Gangaraj, S.K., Copps, K., & Datta, D.K. 1998. Practical Aspects of a-Posteriori Estimation for Reliable Finite Element Analysis. *Computers & Structures*, **66**, 627–664.

Backlund, J. 1978. On isoparametric elements. *Int. J. Numer. Meth. Engng.*, **195**, 731–732.

Barlow, J. 1976. Optimal Stress Locations in Finite Element Models. *Int. J. Num. Meth. Eng.*, **10**, 243–251.

Barlow, J. 1977. Comment on Optimal Stress Locations in Finite Element Models. *Int. J. Numer. Meth. Engrg.*, **11**, 604. Letters to the Editor.

Barlow, J. 1989. More on Optimal Stress Points-Reduced Integration, Element Distortions and Error Estimation. *Int. J. Numer. Methods Engrg.*, **28**, 1487–1504.

Bathe, K.J. 1996. *Finite Element Procedures.* Upper Saddle River, New Jersey, 07458: Prentice-Hall, Inc.

Benedetti, A., de Miranda, S., & Ubertini, F. 2006. A Posteriori Error Estimation Based on the Super Convergent Recovery by Compatibility in Patches. *Int. J. Numer. Meth. Engrg.*, **67**, 108–131.

Blacker, T., & Belytschko, T. 1994. Superconvergent Patch Recovery with Equilibrium and Conjoint Interpolation Enhancements. *Int. J. Numer. Methods Engrg.*, **37**, 517–536.

Boroomand, B., & Zienkiewicz, O.C. 1997a. An Improved REP Recovery and the Effectivity Robustness Test. *Int. J. Numer. Meth. Engrg.*, **40**, 3247–3277.

Boroomand, B., & Zienkiewicz, O.C. 1997b. Recovery by Equilibrium in Patches (REP). *Int. J. Numer. Meth. Engrg.*, **40**, 137–164.

Boroomand, B., & Zienkiewicz, O.C. 1999. Recovery Procedures in Error Estimation and Adaptivity. Part II: Adaptivity in Nonlinear Problems of Elasto-Plasticity Behaviour. *Comput. Methods Appl. Mech. Engrg.*, **176**, 127–146.

Boroomand, B., Zienkiewicz, O.C., & Zhu, J.Z. 1999. Recovery Procedures in Error Estimation and Adaptivity Part I: Adaptivity in Linear Problems. *Comput. Methods Appl. Mech. Engrg.*, **176**, 111–125.

Boroomand, B., Ghaffarian, M., & Zienkiewicz, O.C. 2004. On Application of Two Superconvergent Recovery Procedures to Plate Problems. *Int. J. Numer. Meth. Engrg.*, **61**, 1644–1673.

Brauchli, R.J., & Oden, J.T. 1971. On the Calculation of Consistent Stress Distributions in Finite Element Applications. *Int. J. Numer. Methods Engrg.*, **3**, 317–325.

Buchberger, B. 2001. Groebner Bases: A Short Introduction for Systems Theorists. *In:* Moreno-Diaz, R., Buchberger, B., & Freire, J.L. (eds), *EUROCAST 2001 (8th International Conference on Computer Aided Systems Theory - Formal Methods and Tools for Computer Science).* Springer - Verlag Berlin.

Bugeda, G. 2006. A New Adaptive Remeshing Scheme Based on the Sensitivity Analysis of the SPR Point Wise Error Estimation. *Comput. Methods Appl. Mech. Engrg.*, **195**, 462–478.

Burrows, D.J., & Enderby, L. 1993. Shape Measuring Criteria and the Establishment of Benchmark Tests for Single Membrane Elements (Summary). *NAFEMS*, **Order No.**

R0012.

Cannarozzi, A.A. 1991. *Illustrazione Della Normativa e Controlli Sui Codici.* Udine, Italy: CISM Press.

Cannarozzi, A.A., Momanyi, F.X., & Ubertini, F. 1997. Un Modello Ibrido Ai Flussi in Conduzione Del Calore. *Page Vol.3 of: Atti Del XIII Congresso Nazionale AIMETA.*

Carey, G.F. 2006. A Perspective on Adaptive Modeling and Meshing (AM&M). *Comput. Methods Appl. Mech. Engrg.*, **195**, 214–235.

Carey, G.F., & Oden, J.T. 1983. *Finite Elements.* Vol. II. Prentice-Hall, Inc.

Castellazzi, G. 2001. *Elementi Finiti Piani con Geometria Non Regolare.* M.Phil. thesis, University fo Bologna.

Castellazzi, G., & Ubertini, F. 2004. *Finite Elements with Non-Regular Geometry: Performance and Distortion Measures.* Tech. rept. 184. DISTART - University of Bologna.

Castellazzi, G., de Miranda, S., & Ubertini, F. 2006a (26–28 Giugno). Error Estimation by Compatibility in Patches for Plate Structures. *In: XVI Convegno Italiano di Meccanica Computazionale - GIMC.*

Castellazzi, G., de Miranda, S., & Ubertini, F. 2006b. A Posteriori Error Estimation in Finite Element Analysis of Plate Structures. *In: 7th World Congress on Computational Mechanics.*

Castellazzi, G., de Miranda, S., & Ubertini, F. 2006c (22-26 Maggio). Recovery Based Error Estimation for Plate Problems. *In: VIII Congresso SIMAI.*

Clough, R.W. 1960. The Finite Element Method in Plane Stress Analysis. *Proc. American Society of Civil Engineers (2nd Conference on Electronic Computation, Pitsburg, Pennsylvania)*, **23**, 345–378.

Courant, R. 1943. Variational Methods for the Solution of Problems of Equilibrium and Vibrations. *Bulletin of the American Mathematical Society*, **49**, 1–23.

Cox, D., Little, J., & O'Shea, D. 1992. *Ideals, Varieties and Algorithms.* New York, Inc.: Springer-Verlag.

Daghia, F., de Miranda, S., Ubertini, F., & E., Viola. 2007. A Hybrid Stress Approach for Laminated Composite Plates Within the First Order Shear Deformation Theory. *Int. J. Numer. Meth. Engrg.*, **submitted**.

de Miranda, S, & Ubertini, F. 2006. A Simple Hybrid Stress Element for Shear Deformable Plates. *Int. J. Numer. Meth. Engrg.*, **65**, 808–833.

de Veubeke, B.M. Fraeijs. 1965. Displacement and Equilibrium Models in the Finite Element Method. *Pages 145–197 of:* Zienkiewicz, O.C. (ed), *Stress Analysis.* London: Wiley.

Dobyns, A.L. 1981. Analysis of Simply-Supported Orthotropic Plates Subjected to Static and Dynamic Loads. *AIAA Journal*, **19**, 642–650.

Duarte, A.V.C., & Dutra Do Carmo, E.G. 2000. The Validity of the Superconvergent Patch Recovery in Discontinuous Finite Element Formulations. *Commun. Numer. Meth. Engng.*, **16**, 225–238.

El-Hamalawi, A. 2000. A Simple and Effective Element Distortion Factor. *Computer & Structures*, **75**, 507–513.

Felippa, C.A. 2004. *Introduction to Finite Element Methods.* Boulder, USA: University of Colorado.

Fenves, S.J., Logcher, R.D., Mauch, S.P., & Reinschmidt, K.P. 1964. *STRESS A Users's Manual - A Problem-Oriented Computer Language for Structural Engineering.* M.I.T. Press.

Frey, A.E., Hall, C.A., & Porsching, T.A. 1978. Some Results on the Global Inversion of Bilinear and Quadratic Isoparametric Finite Element Transformations. *Mathematics of Computation*, **32**, 725–749.

Gifford, L.N. 1979. More on distorted isoparametric elements. *Int. J. Numer. Meth. Engng*, **14**, 290–291.

Grätsch, T., & Bathe, K.J. 2005. A Posteriori Error Estimation Techniques in Practical Finite Element Analysis. *Computers & Structures*, **83**, 235–265.

Gu, H., Zong, Z., & Hung, K.C. 2004. A Modified Superconvergent Patch Recovery Method and its Application to Large Deformation Problems. *Finite Elements in Analysis and Design*, **40**, 665–687.

Gupta, A.K. 1978. A Finite Element for Transition from a Fine to a Coarse Grid. *Int. J. Numer. Methods Engrg.*, **12**, 34–45.

Herrmann, L.R. 1972. Interpretation of Finite Element Procedures in Stress Error Minimization. *Journal of the Engineering Mechanics Division*, **98**, 1330–1336.

Hinton, E., & Campbell, J.S. 1974. Local and Global Smoothing of Discontinuous Finite Element Functions Using a Least Squares Method. *Int. J. Numer. Meth. Engrg.*, **8**, 461–480.

Hughes, T.J.R. 1987. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis.* Englewood Cliffs: Prentice Hall.

Irons, B., & Ahmadn, S. 1980. *Techniques of Finite Elements.* Chichester: Ellis Horwood Limited.

Irons, B., & Razzaque, A. 1972. Experience with the Patch Test for Convergence of Finite Element Methods. *Pages 557–587 of:* Aziz, A.K. (ed), *Math Foundations of the Finite Element Method.* Academic Press.

Kelly, D.W. 1984. The Self-Equilibration of Residuals and Complementary a Posteriori Error Estimates in the Finite Element Method. *Int. J. Numer. Meth. Engng.*, **20**, 1491–1506.

Kirkpatrick, S.W., Simons, J.W., & Antoun, T.H. 1998. Development and Validation of High Fidelity Vehicle Crash Simulation Models. *Int. J. Crashworthiness*, 602–611.

Kvamsdal, T., & Okstad, K.M. 1998. Error Estimation Based on Superconvergent Patch Recovery Using Statically Admissible Stress Fields. *Int. J. Numer. Meth. Engrg.*, **42**, 443–472.

Ladeveze, P., & Leguillon, D. 1983. Error Estimate Procedure in the Finite Element Method and Applications. *SIAM J. Numer. Anal.*, **20**, 485–509.

Laitinen, M., Lahtonen, H., & Sjolind, S.-G. 1995. Transverse Shear Correction Factors for Laminated in Cylindrical Bending. *Int. J. Numer. Meth. Engng.*, **11**, 41–47.

Lautersztajn-S, N., & Samuelsson, A. 1997a. On Application of Differential Geometry to Computational Machanics. *Comput. Methods Appl. Mech. Engrg.*, **150**, 25–38.

Lautersztajn-S, N., & Samuelsson, A. 1997b. On Application of Differential Geometry to Computational Mechanics. *Comput. Methods Appl. Mech. Engrg.*, **150**, 25–38.

Lautersztajn-S, N., & Samuelsson, A. 1998a. Distortion Measures and Inverse Mapping for Isoparametric 8-Node Plane Finite Elements with Curved Boundaries. *Comm. Appl. Numer. Meth.*, **14**, 87–101.

Lautersztajn-S, N., & Samuelsson, A. 1998b. Distortion Measures and Inverse Mapping for Isoparametric 8-Node Plane Finite Elements with Curved Boundaries. *Comm. Appl. Numer. Meth.*, **14**, 87–101.

Lee, N.S., & Bathe, K.J. 1993. Effects of Element Distortions on Performance of Isoparametric Elements. *Int. J. Numer. Meth. Engng.*, **36**, 3553–3576.

Lee, T., Park, H.C., & Lee, S.W. 1997. A Superconvergent Stress Recovery Technique with Equilibrium Constraint. *Int. J. Numer. Meth. Engrg.*, **40**, 1139–1160.

Lembo, M., & Podio-Guidugli, P. 2007. How to Use Reactive Stresses to Improve Plate-Theory Approximations of the Stress Field in a Linearly Elastic Plate-Like Body. *International Journal of Solids and Structures*, **44**, 1337–1369.

Li, L.X., Kunimatsu, S., Han, X.P., & Xu, S.Q. 2004. The Analysis of Interpolation Precision of Quadrilateral Elements. *Finite elements in Analysis and Design*, **41**, 91–108.

Lo, S.H., & Lee, C.K. 1998. On Using Different Recovery Procedures for the Construction of Smoothed Stress in Finite Element Method. *Int. J. Numer. Meth. Engrg.*, **43**, 1223–1252.

Lo, S.H., Wan, K.H., & Sze, K.Y. 2006. Adaptive Refinement Analysis Using Hybrid-Stress Transition Elements. *Computers & Structures*, **84**, 2212–2230.

Macneal, R.H., & Harder, R.L. 1992. Eight Nodes or Nine? *Int. J. Numer. Meth. Engng.*, **33**, 1049–1058.

Oh, H.S., & Lim, J.K. 1996. A Simple Error Estimator for Size and Distortion of 2D Isoparametric Finite Elements. *Computers & Structures*, **59**, 989–999.

Ohnimus, S., Stein, E., & Walhorn, E. 2001. Local Error Estimates of FEM for Displacements and Stresses in Linear Elasticity by Solving Local Neumann Problems. *Int. J. Numer. Meth. Engng.*, **52**, 727–746.

Okstad, K.M., Kvamsdal, T., & Mathisen, K.M. 1999. Superconvergent Patch Recovery for Plate Problems Using Statically Admissible Stress Resultant Fields. *Int. J. Numer. Meth. Engrg.*, **44**, 697–727.

Pian, T.H.H. 1973. Hybrid Models. *Pages 50–80 of:* Fenves, S.J., Perrone, R., Robinson, R., & Schonbrich, W.C. (eds), *Numerical and Computer Methods in Structural Mechanics*. New York: Accademic Press.

Pian, T.H.H., & Tong, P. 1972. Finite Element Method in Continuum Mechanics. *In:* Yih, C.S. (ed), *Advances in Applied Mechanics*, vol. 12. New York: Academic Press.

Prathap, G. 1993. *The Finite Element Method in Structural Mechanics*. Dordrecht: Kluwer Academic Press.

Reddy, J.N. 1997. *Mechanics of Laminated Composite Plates - Theory and Analysis*. Boca Raton: CRC press.

Richardson, L.F. 1910. The Approximate Arithmetical Solution by Finite Differences of Physical Problems. *Trans. Roy. Soc.*, **A 210**, 307–357.

Ritz, W. 1909. Über eine Neue Metode Zur Lösung Gewisser Variationsprobleme der Matematischen Physik. *J. Reine Angew. Math.*, **135**, 1–61.

Robinson, J. 1978. Element Evaluation - A Set Assessment Points and Standard Tests. *In: Second World Congress of Finite Element Methods*.

Robinson, J. 1985. Basic and Shape Sensitivity Tests for Membrane and Plate Bending Finite Elements. *NAFEMS*, **Order No. P01**.

Robinson, J. 1987a. CRE Method of Element Testing and the Jacobian Shape Parameters. *Engng. Comps*, **4**, 113–118.

Robinson, J. 1987b. Some New Distortion Measures for Quadrilaterals. *Finite Elements in Analysis and Design*, **3**, 183–197.

Robinson, J. 1988. Distortion Measures for Quadrilaterals with Curved Boundaries. *Finite Elements in Analysis and Design*, **4**, 115–131.

Stein, E. 2003. *Error-Controlled Adaptive Finite Elements In Solid Mechanics*. Chichester: Wiley.

Strang, G., & Fix, G. 1973. *An Analysis of the Finiete Element Method*. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Striklin, L.A., Ho, W.S., Richardson, E.Q., & Haisler, W.E. 1977. On isoparametric vs linear strain triangular elements. *Int. J. Numer. Meth. Engng.*, **11**, 1041–1043.

Taylor, E.L., Beresford, P.J., & Wilson, E.L. 1976. A Non-Conforming Element for Stress Analysis. *Int. J. Num. Meth. Eng.*, **10**, 1211–1220.

Tran, Q.N. 2000. A fast algorithm for Gröbner basis conversion and its applications. *J. Symbolic Computation*, **30**, 451–467.

Turner, M.J., Clough, R.W., Martin, H.C., & Topp, L.J. 1956. Stiffness and Deflection Analysis of Complex Structures. *Journal of Aeronautical Science*, **23**, 805–824.

Ubertini, F. 2002 (July 7–12). Stress Recovey by Compatibility in Patches. *In: Fifth World Congress on Computational Mechanics*. Stress Recovery by RCP.

Ubertini, F. 2004. Patch Recovery Based on Complementary Energy. *Int. J. Numer. Meth. Engrg.*, **59**, 1501–1538.

Vlachoutsis, S. 1992. Shear Correction Factors for Plates and Shells. *Int. J. Numer. Meth. Engng.*, **33**, 1537–1552.

Wiberg, N.E. 1997. Superconvergent Patch Recovery - a Key to Quality Assessed FE Solutions. *Advances in Engineering Software*, **28**, 85–95.

Wiberg, N.E., & Abdulwahab, F. 1993. Patch Recovery Based on Superconvergent Derivatives and Equilibrium. *Int. J. Numer. Meth. Engrg.*, **36**, 2703–2724.

Wiberg, N.E., & Abdulwahab, F. 1997a. Error Estimation with Postprocessed Finite Element Solutions. *Computers & Structures*, **64**, 113–137.

Wiberg, N.E., & Abdulwahab, F. 1997b. Superconvergent Patch Recovery in Problems of Mixed Form. *Commun. Numer. Meth. Engng.*, **13**, 207–217.

Wiberg, N.E., & Li, X. 1999. Adaptive Finite Element Procedures for Linear and Non-Linear Dynamics. *Int. J. Numer. Meth. Engrg.*, **46**, 1781–1802.

Wilson, E.L., Taylor, R.L., Doherty, W., & Ghaboussi, J. 1973. Incompatible Displacement Models. *In:* Fenves, S.J., Perrone, R., Robinson, R., & Schonbrich, W.C. (eds), *Numerical and Computer Methods in Structural Mechanics*. New York: Academic Press Inc.

Yuan, K.Y., & Pian, T.H.H. 1994. The Reference Coordinates and Distortion Measures for Quadrilateral Hybrid Stress Membrane Element. *Computational Mechanics*, **14**, 311–322.

Yuan, K.Y., Huang, Y.S., & Pian, T.H.H. 1993. New Strategy for Assumed Stresses

for 4-Node Hybrid Stress Membrane Element. *Int. J. Numer. Meth. Engng.*, **36**, 1747–1763.

Zhu, J.Z., Hinton, E., & Zienkiewicz, O.C. 1991a. Adaptive Finite Element Analysis with Quadrilaterals. *Computers & Structures*, **40**, 1097–1104.

Zhu, J.Z., Zienkiewicz, O.C., Hinton, E., & Wu, J. 1991b. A new approach to the development of automatic quadrilateral mesh generation. *Int. J. Numer. Methods Engrg.*, **32**, 849–866.

Zienkiewicz, O.C., & Cheung, Y.K. 1967. *The Finite Element Method in Structural and Continuum Mechanics*. London: McGraw-Hill.

Zienkiewicz, O.C., & Taylor, R.L. 1989. *The Finite Element Method*. IV edn. McGraw-Hill Book Company.

Zienkiewicz, O.C., & Zhu, J.Z. 1987. A Simple Error Estimator and Adaptive Procedure for Practical Engineering Analysis. *Int. J. Numer. Methods Engrg.*, **24**, 337–357.

Zienkiewicz, O.C., & Zhu, J.Z. 1991. Adaptivity and Mesh Generation. *Int. J. Numer. Methods Engrg.*, **32**, 783–810.

Zienkiewicz, O.C., & Zhu, J.Z. 1992a. The Superconvergent Patch Recovery (SPR) and a Posteriori Error Estimates. Part I: The Recovery Technique. *Int. J. Numer. Methods Engrg.*, **33**, 1331–1364.

Zienkiewicz, O.C., & Zhu, J.Z. 1992b. The Superconvergent Patch Recovery (SPR) and a Posteriori Error Estimates. Part II: Error Estimates and Adaptivity. *Int. J. Numer. Methods Engrg.*, **33**, 1365–1382.

Zienkiewicz, O.C., & Zhu, J.Z. 1992c. The Superconvergent Patch Recovery (SPR) and Adaptive Finite Element Refinement. *Comput. Methods Appl. Mech. Engrg.*, **101**, 207–224.

Zienkiewicz, O.C., & Zhu, J.Z. 1995. Superconvergence and the Superconvergent Patch Recovery. *Finite Elements in Analysis and Design*, **19**, 11–23.

Zienkiewicz, O.C., Zhu, J.Z., & Gong, N.G. 1989. Effective and Pratical $h-p$-Version Adaptive Analysis Procedures for the Finite Element Method. *Int. J. Numer. Methods Engrg.*, **28**, 879–891.

Zienkiewicz, O.C., Zhu, J.Z., & Wu, J. 1993. Superconvergent Patch Recovery Techniques - some Further Tests. *Commun. Numer. Meth. Engng.*, **9**, 251–258.

Curriculum

Giovanni Castellazzi received the Laurea degree in Civil Engineering from the University of Bologna, Italy, in summer 2001. After receiving his Laurea degree, he spent time as Support Tech Specialist at Laboratory of Computational Mechanics (DISTART) Universityy of Bologna.
From January 2004 to January 2007 he attended at Ph.D. program in Structural Mechanics - University of Bologna.

This thesis is available in PDF format on LAMC web site: *www.lamc.ing.unibo.it*

Contacts

Laboratory of Computational Mechanics
DISTART - Structural Department
Faculty of Engineering
ALMA MATER STUDIORUM - University of Bologna
Viale del Risorgimento 2, 40136 Bologna - Italy
Phone: +39 051 2093503
Fax: +39 051 2093496
E-mail: giovanni.castellazzi@unibo.it
Web: www.lamc.ing.unibo.it