LEAST-SQUARES VARIATIONAL PRINCIPLES AND THE FINITE ELEMENT METHOD: THEORY, FORMULATIONS, AND MODELS FOR SOLID AND FLUID MECHANICS

A Dissertation

by

JUAN PABLO PONTAZA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2003

Major Subject: Mechanical Engineering

LEAST-SQUARES VARIATIONAL PRINCIPLES AND THE FINITE ELEMENT METHOD: THEORY, FORMULATIONS, AND MODELS FOR SOLID AND FLUID MECHANICS

A Dissertation

by

JUAN PABLO PONTAZA

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

J. N. Reddy
(Chair of Committee)

Ali Beşkök
(Member)

H. A. Hogan
(Member)

H. C. Chen
(Member)

Dennis L. O'Neal
(Head of Department)

December 2003

Major Subject: Mechanical Engineering

ABSTRACT

Least-squares Variational Principles and
the Finite Element Method: Theory, Formulations,
and Models for Solid and Fluid Mechanics. (December 2003)
Juan Pablo Pontaza, B.S., Texas A&M University;
M.S., Massachusetts Institute of Technology
Chair of Advisory Committee: Dr. J. N. Reddy

We consider the application of least-squares variational principles and the finite element method to the numerical solution of boundary value problems arising in the fields of solid and fluid mechanics. For many of these problems least-squares principles offer many theoretical and computational advantages in the implementation of the corresponding finite element model that are not present in the traditional weak form Galerkin finite element model. Most notably, the use of least-squares principles leads to a variational unconstrained minimization problem where stability conditions such as inf-sup conditions (typically arising in mixed methods using weak form Galerkin finite element formulations) never arise. In addition, the least-squares based finite element model always yields a discrete system of equations with a symmetric positive definite coefficient matrix. These attributes, amongst many others highlighted and detailed in this work, allow the development of robust and efficient finite element models for problems of practical importance. The research documented herein encompasses least-squares based formulations for incompressible and compressible viscous fluid flow, the bending of thin and thick plates, and for the analysis of shear-deformable shell structures.

To my Mother and Family, and

in loving memory of my Father.

ACKNOWLEDGMENTS

I would like to take this opportunity to thank my supervisor, Dr. J. N. Reddy, for the support that he has extended towards me during the course of this research. His reviews of my work and his dedication to teaching made my graduate studies at Texas A&M University an enjoyable and rewarding experience.

Special thanks to Dr. Ali Beşkök, for our many fruitful discussions regarding spectral/*hp* methods and well-posed open boundary conditions for compressible flow. The numerical simulations presented in this work were, to a large extent, performed using resources from the Texas A&M University Supercomputer Facility. Their support is greatly appreciated.

I would like to thank my mother, Dalila de Pontaza, for her support and encouragement throughout my academic life. Her vision and moral support made all this possible. I thank my wife, María José, for her love, care, and encouragement which allowed me to remain focused on my research. I also thank my daughters, Nicole and Cristina, for their support, patience, and understanding.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

TABLE Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

A. Background

It is well known that application of the weak form Galerkin procedure to problems whose solutions can be characterized as global minimizers results in global minimization of quadratic functionals, such is the case for linear elasticity problems; see [93]. In this case, given a conforming discretization, the finite element solution is an orthogonal projection of the exact solution onto the trial space, i.e., the finite element solution is a minimizer of an energy functional on the trial space so that it represents the best possible approximation in the energy norm. Such a setting, hereafter referred to as a *variational setting*, is the most favorable for a finite element approximation.

Failure to immediately recognize the link between global minimization of unconstrained convex functionals and the success of the method for linear elasticity problems led to early attempts to extend weak form Galerkin formulations beyond problems whose solutions could be characterized as unconstrained global minimizers. For example, problems of the saddle-point type, whose solution can be interpreted as a constrained minimization of a convex functional by requiring adherence of the discrete spaces to restrictive compatibility conditions; lack many of the attractive properties of the variational setting. Implementation of weak form Galerkin finite element models without accounting for the restrictive conditions leads to a non-optimal finite element approximation and un-reliable numerical results, in some cases revealing itself as spurious oscillations in the numerical solution. A typical example is given by the primitive variable formulation of the Stokes problem for which the velocity and

---

The journal model is *IEEE Transactions on Automatic Control*.

pressure approximation spaces cannot be chosen independently and must satisfy an inf-sup condition (see, e.g., [35]).

In the context of the Stokes and/or the Navier-Stokes equations, various finite element models attempting to fully or partially recover some of the properties of the variational setting have been proposed, and among them the Galerkin-Least-Squares and stabilized Galerkin methods have been extensively researched (see, e.g., [18, 32]). These approaches have failed to achieve widespread use and acceptance due to their explicit dependence on various mesh-dependent calibration parameters that need to be fine-tuned from application to application.

In the past few years finite element models based on least-squares variational principles have drawn considerable attention (see, e.g., [55, 13]). In particular, given a partial differential equation (PDE) or a set of partial differential equations, the least-squares method allows us to define an unconstrained minimization principle so that a finite element model can be developed in a variational setting. The idea is to define the least-squares functional as the sum of the squares of the equations residuals measured in suitable norms of Hilbert spaces. Assuming the governing equations (augmented with suitable boundary conditions) have a unique solution, the least-squares functional will have a unique minimizer. Thus, by construction, the least-squares functional is always positive and convex, ensuring coerciveness, symmetry, and positive definitiveness of the bilinear form in the corresponding variational problem. Moreover, if the induced energy norm is equivalent to a norm of a suitable Hilbert space, optimal properties of the resulting least-squares formulation can be established. However, as we shall elaborate further in Chapter II, an optimal least-squares formulation may result in an impractical finite element model. The reconciliation that must exist between practicality and optimality in least-squares based finite element models if of great importance and was first recognized by Bochev and

Gunzburger [11, 13]. The practicality of the resulting finite element model is, to a large extent, determined by the complexity of algorithm development and CPU solve time of the resulting discrete system of equations. Typically, the practicality is measured in terms of C^k continuity/regularity of the finite element spaces across inter-element boundaries. Ideally, a least-squares finite element model with "C^0 practicality" and full (mathematical) optimality is to be developed – unfortunately, this can seldom be achieved in a satisfactory manner.

The first rigorous mathematical analysis of least-squares finite element formulations can be traced back to the work of Bramble and Schatz [16] and Bramble and Nitsche [15]. Earlier work on least-squares is documented in the review of Easson [30]. Bramble and Schatz [16] analyzed least-squares models for $2m$th order scalar elliptic boundary value problems, where the least-squares functional was defined in terms of L_2 norms. Conforming discretizations require that the finite element space be spanned by functions that belong to the Hilbert space H^{2m}, in contrast to weak form Galerkin models which require only H^m regularity (due to the weakened differentiability requirements induced by the integration by parts). For the least-squares model, this implies a minimum of C^1 regularity of the finite element spaces across inter-element boundaries. Moreover, the least-squares discrete problem has associated with it a matrix conditioning of order h^{-4m}, in contrast to the h^{-2m} conditioning of the Galerkin models. As a result, least-squares formulations lost appeal and failed to gain popularity due to the higher regularity requirements (relative to the weak form Galerkin formulation).

To reduce the higher regularity requirements, the PDE or PDEs are first transformed into an equivalent lower order system by introducing additional independent variables, sometimes termed auxiliary variables, and then formulating the least-squares model based on the equivalent lower order system. The additional variables

imply an increase in cost, but can be argued to be beneficial as they may represent physically meaningful variables, e.g., fluxes, stresses or rotations, and will be directly approximated in the model. Such an approach, is believed to be first explored by Jesperson [52] and is the preferred approach in modern implementations of least-squares finite element models. For 2nd order PDEs, and equivalent first order system is introduced, and if the least-squares functional is defined in terms of L_2 norms only, the finite element model allows the use of nodal/modal expansions with merely C^0 regularity.

In this work, to retain the "C^0 practicality", we use L_2 norms to define the least-squares functional, which in turn is defined in terms of first-order PDEs only. Moreover, we can require the boundary conditions to be imposed either strongly or, alternatively, in a weak sense through the least-squares functional. Notable benefits from working in a variational setting include the fact that stability requirements such as inf-sup conditions will never arise and that the resulting algebraic problem will have a symmetric positive definite (SPD) coefficient matrix, which can be solved by using robust iterative methods, such as preconditioned conjugate gradient methods. In order to *fully* emulate the variational setting, one must define a L_2 least-squares functional that induces an energy norm that is equivalent to the H^1-norm. If this is achieved, the least-squares finite element solution can be interpreted as an orthogonal projection in the Hilbert space with respect to the induced energy inner product and thus results in an optimal approximation in the H^1-norm.

Identifying norm equivalent functionals associated with a partial differential equation is not a trivial task. We rely on the Agmon, Douglis, and Nirenberg (ADN) theory [1] for elliptic operators to identify appropriate spaces to define a least-squares functional that is norm-equivalent. First-order differential operators that are homogeneous elliptic in the ADN sense allow the construction of a L_2 least-squares functional

that is H^1-norm equivalent. For non-homogeneous elliptic first-order differential operators (in the ADN sense), a norm-equivalent least-squares functional may also be constructed. However, the least-squares functional will be defined in terms of computationally impractical norms (e.g., H^{-1} norms *or* norms that would require C^1 regularity of the finite element spaces across inter-element boundaries) which could be replaced by computable equivalents (e.g., weighted L_2 norms); perhaps at the expense of sacrificing properties of the variational setting.

First-order systems that allow the construction of a L_2 least-squares functional that is H^1-norm equivalent are commonly referred to as H^1-coercive formulations. Such systems yield optimal error estimates with respect to the H^1-norm for all variables.

B. Motivation of the research

In previous work concerned with least-squares finite element formulations, predominantly low order nodal expansions have been used to develop the discrete finite element model (see [55] and references therein). When the formulation is not H^1-norm equivalent (sometimes referred to as a *quasi-norm-equivalent* or a *non-equivalent* formulation), low order nodal expansions tend to lock and reduced integration techniques must be used to obtain acceptable numerical results. When enough redundant degrees of freedom are constrained the least-squares finite element solution using reduced integration yields a collocation least-squares finite element solution. However, the collocation solution may not always be reliable and the least-squares functional cannot be used to measure the quality of the solution (as it identically vanishes at the collocation points).

Least-squares finite element models that invoke the use of reduced integration techniques are thus not *bona fide* least-squares based models, but rather collocation least-squares based models. With this in mind, it is important to note that reduced integration techniques will only result in a collocation solution if a strict balance between the number of collocation points and total number of degrees of freedom is satisfied. It is thus safe to say that, in general, blind application of reduced integration techniques will not result in a collocation solution and should be avoided.

Even though a *quasi-norm-equivalent* or a *non-equivalent* formulation departs from the ideal mathematical setting, it does not lead to disastrous results; as a violation of the inf-sup condition would, in a mixed weak form Galerkin formulation. Motivation for this study arose from the resiliency of least-squares formulations to such departures from the mathematically ideal setting. Furthermore, we refrain from using reduced integration techniques and it is through our work that we wish to advance the practice of full integration and residual minimization through proper *hp*-refinement. In particular, we find that use of high p-levels (typically $p \geq 4$) are desirable for least-squares based finite element models and that such practice provides a desirable balance between practicality and optimality.

C. Scope of the research

The research started at Texas A&M University in the Fall of 2001 and encompasses the formulation of least-squares based finite element models for incompressible viscous fluid flow [87, 86], inviscid compressible flow [81], viscous compressible flow in the subsonic/transonic/supersonic flow regime [83], the bending of thin and thick plates [85], and shear-deformable shells [84]. In addition we develop a novel least-squares formulation for transient problems where the effects of space-time are fully coupled [86], resulting in an unconditionally stable time-stepping scheme where spectral accuracy

is achieved in space-time. Throughout this work we emphasize the advantages of using least-squares based finite element models, as opposed to the traditional weak form Galerkin based finite element models.

The dissertation is organized as follows. In Chapter II, we present the steps involved in developing a least-squares based finite model, we do so in the context of an abstract initial boundary value problem. We briefly present the mathematical theory of least-squares based formulations, specifically the notion of norm-equivalence of least-squares functionals and its consequences in the form of optimal a priori error estimates.

Chapters III and IV are concerned with incompressible and compressible fluid flows. In Chapter III, we develop least-squares based formulations and finite element models for viscous incompressible fluid flows governed by the Navier-Stokes equations. Least-squares finite element models based on equivalent first-order systems obtained by introducing vorticity, stresses, or velocity gradients as additional independent variables are presented and compared. Verification and validation of the models and associated computational algorithm is presented in the form of numerical results and comparisons with well established benchmark problems, including flow over a backward-facing step, lid-driven cavity flows, and flow past a circular cylinder. In Chapter IV, we present least-squares formulations for viscous compressible fluid flows. Such formulations are aimed at flow problems characterized by incompressible flow in parts of the domain with imbedded regions where compressibility effects are significant and cannot be neglected. The formulation is directly applicable to viscous flows in the subsonic, transonic, or supersonic regime. The ease by which characteristic based boundary conditions are imposed through the least-squares functional is emphasized. Numerical results for flow past a circular cylinder at different free-stream Mach numbers and different surface thermal loadings are presented.

Chapters V and VI are concerned with applications to solid mechanics, specifically plates and shells. In Chapter V, we present a least-squares formulation for the bending of thin and thick plates, i.e., plate models based on the Kirchhoff and Mindlin theories. The least-squares based plate elements are shown, through the solution of carefully chosen benchmark problems, to be insensitive to shear-locking and severe geometric distortions. In Chapter VI, we develop a least-squares formulation for the analysis of shear-deformable shell structures. Specifically we consider cylindrical shell structures and illustrate the performance of the formulation through the numerical solution of several well established benchmark problems, including the barrel-vault problem and the pinched cylinder problem. Finally in Chapter VII, we give concluding remarks and comment on the direction of ongoing and future research.

CHAPTER II

AN ABSTRACT LEAST-SQUARES FORMULATION

In this chapter we present the steps involved in developing and arriving at a least-squares based finite element model. We wish to present the procedure in a general setting, and to this end present the procedure in the context of an abstract initial boundary value problem.

First, we introduce notation that will be used throughout this chapter and in the remainder of this work. Given the abstract initial boundary value problem, our first task is to form the least-squares functional; which may be defined to yield a space-time coupled or decoupled formulation. The variational problem statement is obtained by defining the least-squares minimization principle in infinite dimensional spaces and the corresponding finite element model obtained by restricting the spaces to finite dimensional subspaces. These closed spaces are spanned at the element level by piecewise nodal/modal polynomial expansions.

Of paramount importance is to establish whether or not the resulting least-squares based formulation is within the ideal mathematical setting, i.e., whether or not the least-squares functional defines an equivalent norm in a suitable Hilbert space. Although norm-equivalence is always desirable, it sometimes needs to be sacrificed to yield practical finite element models. We discuss in detail the properties that are lost when this compromise between optimality and practicality is made, and whether or not the resulting formulation is able to yield optimal results.

We conclude the chapter by giving details on the nodal/modal expansions used in this work and discussing effective and efficient solution procedures designed to take advantage of the (always) symmetric positive definite structure of the resulting coefficient matrix associated with the discrete least-squares finite element model.

A. Notation

Let $\bar{\Omega}$ be the closure of an open bounded region Ω in $\mathbb{R}^d$, where $d = 2$ or 3 represents the number of space dimensions, and $\mathbf{x} = (x_1, \ldots, x_n) = (x, y, z)$ be a point in $\bar{\Omega} = \Omega \cup \partial\Omega$, where $\partial\Omega = \Gamma$ is the boundary of Ω.

For $s \geq 0$, we use the standard notation and definition for the Sobolev spaces $H^s(\Omega)$ and $H^s(\Gamma)$ with corresponding inner products denoted by $(\cdot,\cdot)_{s,\Omega}$ and $(\cdot,\cdot)_{s,\Gamma}$ and norms by $\|\cdot\|_{s,\Omega}$ and $\|\cdot\|_{s,\Gamma}$, respectively. Whenever there is no chance of ambiguity, the measures Ω and Γ will be omitted from inner product and norm designations. We denote the $L_2(\Omega)$ and $L_2(\Gamma)$ inner products by $(\cdot,\cdot)$ and $(\cdot,\cdot)_\Gamma$, respectively. By $\mathbf{H}^s(\Omega)$ we denote the product space $[H^s(\Omega)]^d$. We denote by $H_0^1(\Omega)$ the space consisting of $H^1(\Omega)$ functions that vanish on the boundary Γ and by $\bar{L}_2(\Omega)$ the space of all square integrable functions with zero mean with respect to Ω.

B. The abstract problem

Consider the following abstract initial boundary value problem:

$$\mathcal{L}_t(\mathbf{u}) + \mathcal{L}_\mathbf{x}(\mathbf{u}) = \mathbf{f} \qquad \text{in } \Omega \times (0, \tau] \tag{2.1}$$

$$\mathcal{G}(\mathbf{u}) = \mathbf{h} \qquad \text{on } \Gamma \times (0, \tau] \tag{2.2}$$

in which $\mathcal{L}_t$ and $\mathcal{L}_\mathbf{x}$ are first-order partial differential operators in time and space respectively, acting on the vector $\mathbf{u}$ of unknowns, $\mathbf{f}$ is a known vector valued forcing function, $\mathcal{G}$ is a trace operator acting on $\mathbf{u}$, and $\mathbf{h}$ represents a known vector valued function on the boundary. We assume initial conditions are given such that the problem is well posed and a unique solution exists.

C. L_2 least-squares formulation

The L_2 least-squares functional associated with the abstract initial boundary value problem is constructed by summing up the squares of the equations residuals in the L_2-norm and is given by

$$\mathcal{J}\left(\mathbf{u};\mathbf{f},\mathbf{h}\right) = \frac{1}{2}\left(\parallel \mathcal{L}_t(\mathbf{u}) + \mathcal{L}_{\mathbf{x}}(\mathbf{u}) - \mathbf{f} \parallel^2_{0,\,\Omega\times(0,\tau]} + \parallel \mathcal{G}(\mathbf{u}) - \mathbf{h} \parallel^2_{0,\,\Gamma\times(0,\tau]} \right). \qquad (2.3)$$

It is easy to see that the minimizer of (2.3) solves (2.1)-(2.2) and viceversa.

Note that in presenting the abstract initial boundary value problem and defining its associated least-squares functional we made two restrictions: (1) the temporal and spatial partial differential operators are of first-order and (2) the least-squares functional is defined exclusively in terms of L_2 norms. These restrictions are necessary in order to ensure a pre-determined level of practicality in the resulting least-squares based finite element model: specifically, the permission to use finite element spaces with merely C^0 regularity across inter-element boundaries. This is done with the understanding that the resulting finite element formulation may depart from the ideal mathematical setting and hence may not yield optimal a priori error estimates. Nevertheless, as we shall demonstrate with the aid of numerical examples, such departures from the ideal mathematical setting will not result in disaster (as a violation of an inf-sup condition would, in a weak form Galerkin formulation). In fact, least-squares formulations that depart from the ideal mathematical setting show remarkable robustness and are able to recover optimal properties provided the expansion order is high enough. For the purposes of this study we shall retain the aforementioned pre-determined level of practicality (i.e., C^0 practicality). However, ongoing research is pointing towards relaxing this level of practicality by considering finite element spaces with better regularity across inter-element boundaries: C^1 or C^2 regularity. This re-

laxes restrictions (1) and (2) and hence may improve on the compromise between optimality and practicality (see Chapter VII).

If the partial differential equations (PDEs) under consideration are not of first order, the "C^0 practicality" of the least-squares based finite element model comes at an extra cost, implied in restriction (1); which requires that the partial differential operators be of first order. This can always be achieved by introducing auxiliary variables until a first order system is attained. The added cost might be viewed as beneficial, in the sense that the auxiliary variables may have physical relevance to the problem under consideration, e.g., fluxes, stresses, or rotations.

1. Space-time coupled formulation

In addition, note that prior to defining functional (2.3) we did not replace the temporal operator with a discrete equivalent. This results in a fully space-time coupled formulation, implied in the definition of functional (2.3) where the L_2 norm is defined in space-time, i.e., $\|\cdot\|_{0,\,\Omega\times(0,\tau]}$ denotes the L_2 norm of the enclosed quantity in space-time:

$$\|u\|^2_{0,\,\Omega\times(0,\tau]} = \int_0^\tau \int_\Omega |u|^2 \, d\Omega \; dt \, .$$

This implies, for example, that a two-dimensional time-dependent problem will be treated as a three-dimensional problem in space-time domain. When dealing with the stationary form of the equations the integral over time domain is simply dropped.

In the space-time coupled approach, the effects of space and time are allowed to remained coupled. There is no approximation of the initial boundary value problem. Instead, a basis is introduced in time domain to represent the time evolution of the independent variables.

Invariably, we as analysts would like to simulate and study the time evolution of an initial value problem for large values of time. Taking into consideration modelling

issues, we realize that this would require a space-time mesh with a large number of elements in time domain. The size of the resulting set of assembled algebraic equations could be large and prohibitively expensive in terms of available computer memory and non-optimal in terms of CPU solve time. To alleviate the drawbacks, we adopt a time-stepping procedure in which the solution is obtained for space-time strips in a sequential manner. The initial conditions for the current space-time strip are obtained from the latest space plane from the previous space-time strip. Hence, for each space-time strip we solve a true initial boundary value problem, by minimizing the following functional in space-time domain:

$$\mathcal{J}\left(\mathbf{u};\mathbf{f},\mathbf{h}\right) = \frac{1}{2}\left(\left\| \mathcal{L}_t(\mathbf{u}) + \mathcal{L}_{\mathbf{x}}(\mathbf{u}) - \mathbf{f} \right\|^2_{0,\,\Omega\times[t_s,t_{s+1}]} + \left\| \mathcal{G}(\mathbf{u}) - \mathbf{h} \right\|^2_{0,\,\Gamma\times[t_s,t_{s+1}]} \right) \quad (2.4)$$

where the interval $[t_s, t_{s+1}]$ can be taken arbitrarily large, i.e., there are no restrictions on the size of the interval.

Since the initial boundary value problem in each space-time strip is represented and solved with predetermined accuracy of order p, the question of stability does not arise (see [86, 8, 9] and Chapters III and IV). The only issue that remains is accuracy; which we can control by *hp* refinements in time. Furthermore, the L_2 least squares functional can provide an error measure for adaptive h, p, or *hp* refinements in space-time.

2. Space-time decoupled formulation

In a space-time decoupled formulation, discretization in space and time are done independently. Traditionally, the temporal operators are represented by truncated Taylor series expansions in time domain. Such formulations result in an inherent approximation of the initial boundary value problem and thus the investigation of stability is essential. Representation of the temporal operator by high-order approximations,

such as multi-step schemes, are only conditionally stable; imposing severe limitations on the size of the allowable time increment. To further illustrate we present in the following a space-time decoupled formulation.

First, the temporal operator in Eq. (2.1) is replaced by a discrete equivalent:

$$\mathcal{L}_t(\mathbf{u}) \approx \mathcal{L}_{\Delta t}(\mathbf{u}^{s+1}, \mathbf{u}^{s-q}) \,,$$

where the time increment dependence of the discrete operator is explicit as well as its dependence on histories of previous time steps. For sufficiently small Δt, the modified problem is equivalent to the original problem. To march the problem in time using a least-squares spatial finite element model, we must minimize the following space functional at each time step:

$$\mathcal{J}_{\Delta t}\left(\mathbf{u}; \mathbf{f}, \mathbf{h}\right) = \frac{1}{2}\Big(\| \mathcal{L}_{\Delta t}(\mathbf{u}^{s+1}, \mathbf{u}^{s-q}) + \mathcal{L}_{\mathbf{x}}(\mathbf{u}^{s+1}) - \mathbf{f}^{s+1} \|^2_{0,\Omega} + \| \mathcal{G}(\mathbf{u}^{s+1}) - \mathbf{h}^{s+1} \|^2_{0,\Gamma} \Big) \quad (2.5)$$

where the dependence on the time increment $\Delta t = t_{s+1} - t_s$ is evident. Once the discrete model is obtained, the eigen-spectrum of the amplification matrix needs to be examined and the time-step restriction determined as a function of an upper bound for the maximum eigenvalue; see Reddy [94]. The upper bound for the maximum eigenvalue will undoubtedly be highly dependent on the spatial expansion order, resulting in severe restrictions in the allowable time increment for high-order spatial expansions. Note that, unlike the weak form Galerkin formulation of time-dependent problems, a semidiscrete system of ordinary differential equations in time for the element degrees of freedom does not arise because the temporal approximation was performed first.

D. The variational problem

Having defined the least-squares functional (2.3), the abstract least-squares minimization principle can be stated as:

$$\text{find } \mathbf{u} \in X \text{ such that } \mathcal{J}(\mathbf{u}; \mathbf{f}, \mathbf{h}) \leq \mathcal{J}(\mathbf{v}; \mathbf{f}, \mathbf{h}) \, \forall \, \mathbf{v} \in X \tag{2.6}$$

where X is a suitable vector space, e.g. $X = \mathbf{H}^1\left(\bar{\Omega} \times (0, \tau]\right)$, and we assume that the functions $\mathbf{f}, \mathbf{h}$ are sufficiently regular, e.g. $\mathbf{f} \in \mathbf{L}_2(\Omega \times (0, \tau])$ and $\mathbf{h} \in \mathbf{L}_2(\Gamma \times (0, \tau])$.

The Euler-Lagrange equation for this minimization problem is given by the following variational problem (also see [93]):

$$\text{find } \mathbf{u} \in X \text{ such that } \mathcal{B}(\mathbf{u}, \mathbf{v}) = \mathcal{F}(\mathbf{v}) \; \forall \, \mathbf{v} \in X \tag{2.7}$$

where $\mathcal{B}$ is a symmetric form given by

$$\mathcal{B}(\mathbf{u}, \mathbf{v}) = (\mathcal{L}(\mathbf{u}), \mathcal{L}(\mathbf{v}))_{0, \, \Omega \times (0,\tau]} + (\mathcal{G}(\mathbf{u}), \mathcal{G}(\mathbf{v}))_{0, \, \Gamma \times (0,\tau]}$$

and $\mathcal{F}$ is a functional given by

$$\mathcal{F}(\mathbf{v}) = (\mathbf{f}, \mathcal{L}(\mathbf{v}))_{0, \, \Omega \times (0,\tau]} + (\mathbf{h}, \mathcal{G}(\mathbf{v}))_{0, \, \Gamma \times (0,\tau]}$$

where $\mathcal{L} = \mathcal{L}_{\mathbf{x}} + \mathcal{L}_t$.

The inclusion of the boundary residual in (2.3) allows the use of spaces X that are not constrained to satisfy the boundary condition (2.2). In such a case, the boundary condition (2.2) is enforced in a weak sense through the least-squares functional. This is a tremendous advantage of least-squares based formulations, as it allows boundary conditions that are computationally difficult to impose to be efficiently included in the least-squares functional. An example where this property becomes extremely useful is for viscous or inviscid compressible flow (see [83] or Chapter IV), where characteristic-

based boundary conditions need to be prescribed at outflow/inflow boundaries. Of course, if the boundary condition (2.2) can be easily imposed and included in the space X, we omit the residual associated with the boundary term in (2.3).

E. The finite element model

The finite element model is obtained by either restricting (2.7) to the finite dimensional subspace X_{hp} of the infinite dimensional space X, or equivalently by minimizing (2.3) with respect to the chosen approximating spaces. This process leads to the discrete variational problem given by

$$\text{find } \mathbf{u}^{hp} \in X_{hp} \text{ such that } \mathcal{B}\left(\mathbf{u}^{hp}, \mathbf{v}^{hp}\right) = \mathcal{F}\left(\mathbf{v}^{hp}\right) \; \forall \, \mathbf{v}^{hp} \in X_{hp} \tag{2.8}$$

We proceed to define a discrete problem by choosing appropriate finite element subspaces for each of the components of the vector valued function $\mathbf{u}$. There are no restrictive compatibility conditions on the discrete spaces, so we choose the same finite element subspace for each of the primary variables. The only requirement on the approximating spaces is that we choose continuous piecewise polynomials that are at least bi-linear (in two dimensions) or tri-linear (in three dimensions).

F. Norm-equivalence and its implications

Our ultimate goal is to use (2.7) to compute approximate solutions to (2.1)-(2.2). Clearly, the least-squares functional is consistent in the sense that for sufficiently smooth data $\mathbf{f}$, $\mathbf{h}$ and smooth solutions $\mathbf{u}$ of (2.1)-(2.2), $\mathcal{J}\left(\mathbf{u}; \mathbf{f}, \mathbf{h}\right) = 0$. Furthermore, by construction, the least-squares functional is convex and positive. Which allows us to define an energy norm:

$$\|\cdot\|_{\mathrm{E}} = \mathcal{J}\left(\cdot; 0, 0\right)^{1/2} \; : \; X \to \mathbb{R}\,, \tag{2.9}$$

and an associated energy inner product:

$$(\,(\cdot,\cdot)\,)_{\mathrm{E}} \,:\, X \times X \to \mathbb{R}\,. \tag{2.10}$$

Then, if $X_{hp} \subset X$,

1. the variational problem (2.8) has a unique solution given by $\mathbf{u}^{hp} \in X^{hp}$, and
2. $\mathbf{u}^h$ is the orthogonal projection of $\mathbf{u}$ with respect to the energy inner product (2.10), and thus represents the best possible solution in the energy norm (2.9).

In addition, if $\{\varphi_i\}_{i=1}^n$ spans X_{hp}; the variational problem (2.8) is a linear system of algebraic equations whose coefficient matrix $\underline{\mathbf{K}}$ has entries given by

$$K_{ij} = ((\varphi_j, \varphi_i))_{\mathrm{E}}\,. \tag{2.11}$$

Thus, the coefficient matrix $\underline{\mathbf{K}}$ is a Gramm matrix with respect to the energy inner product (2.10) and thus is symmetric and positive definite. As a result, the system $\underline{\mathbf{K}}\mathbf{U} = \mathbf{F}$ has a unique solution.

Note that so far we have said nothing about norm-equivalence and already we have established that the least-squares based finite element model will yield a convex, unconstrained minimization problem with a unique minimizer that coincides with the best possible approximate solution to (2.1)-(2.2) in a well defined norm. In addition we have also established that the resulting discrete algebraic problem will have a symmetric positive definite coefficient matrix. This explains the robustness of least-squares based formulations, even when they depart from the ideal mathematical setting (which has not yet been established).

Establishing norm-equivalence (and hence, and ideal mathematical setting) will enable us to predict the asymptotic behavior of how $\mathbf{u}^{hp}$ approaches $\mathbf{u}$. For simplicity, consider the stationary version of (2.1)-(2.2) and let $X = \mathbf{H}^1(\bar{\Omega})$. If the coercivity

relation

$$C_2 \|\mathbf{u}\|_{1,\Omega} \leq \| \mathcal{L}(\mathbf{u}) \|_{0,\Omega} + \| \mathcal{G}(\mathbf{u}) \|_{0,\Gamma} \leq C_1 \|\mathbf{u}\|_{1,\Omega} \tag{2.12}$$

holds for all smooth solutions $\mathbf{u}$ of (2.1)-(2.2), then the L_2 least-squares functional defines an equivalent norm in H^1 in the sense

$$\frac{1}{2} C_2^2 \|\mathbf{u}\|_{1,\Omega}^2 \leq \mathcal{J}(\mathbf{u}) \leq \frac{1}{2} C_1^2 \|\mathbf{u}\|_{1,\Omega}^2 \tag{2.13}$$

or equivalently (and for the more general case),

$$\frac{1}{2} C_2^2 \|\mathbf{u}\|_X^2 \leq \|\mathbf{u}\|_{\mathrm{E}}^2 \leq \frac{1}{2} C_1^2 \|\mathbf{u}\|_X^2 \tag{2.14}$$

and optimal h-convergence rates should be attained as follows [13]:

$$\|\mathbf{u} - \mathbf{u}^{hp}\|_{r,\Omega} \leq C\, h^{p+1-r}, \quad r = 0, 1 \tag{2.15}$$

for smooth solutions $\mathbf{u}$, with C independent of h. If norm-equivalence cannot be established the constant C may be dependent on the mesh parameter h and/or not much can be said about the optimality of the convergence rates of the resulting formulation. This, however, does not imply that the resulting method is not optimal. It simply means that its optimality cannot be determined a priori using standard elliptic theory.

We illustrate the behavior of a non-equivalent least-squares functional with the solution of the stationary incompressible Navier-Stokes equations in its velocity-pressure-vorticity based first-order form with pure velocity boundary conditions. For the purposes of this demonstration, we solve the well-known two-dimensional lid-driven cavity problem in a series of meshes and for different p-levels such that the total number of degrees of freedom remains constant. We consider the three cases shown in Fig. 1 and flow conditions with $\mathrm{Re} = 10^3$. For each case we plot the u-

velocity profile along the geometric vertical mid-line of the cavity and the v-velocity profile along the geometric horizontal mid-line of the cavity. We take the target solution to be that reported and tabulated by Ghia et al. [34], frequently used and widely accepted as a verification benchmark.

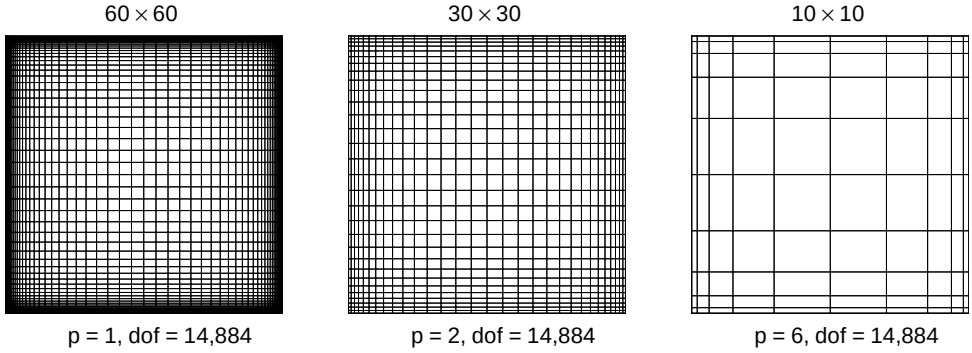

Fig. 1. Series of meshes used for the two-dimensional lid-driven cavity problem at flow conditions Re $= 10^3$. The meshes are chosen such that the total number of degrees of freedom remains constant for p-levels of 1,2, and 6, as shown.

Figures 2 and 3 show the u- and v- velocity profiles along the geometric vertical and horizontal mid-lines of the cavity. For the 60×60 finite element mesh with a p-level of 1 (i.e., bi-linear elements), the predicted velocity profiles are surprisingly of extremely poor quality. Initially one might be disappointed at the performance of the least-squares based formulation, as the 60×60 bi-linear finite element mesh will give considerably better results with a weak form Galerkin formulation. However, knowing that the least-squares functional we used to develop the finite element model does not define an equivalent norm in $X^{hp} \subset X$, we conjecture that the constant C in Eq. (2.15) depends on the mesh parameter h and thus expect a poor numerical solution. To keep the cost of the computation comparable and the total number of degrees of freedom constant we consider a 30×30 finite element mesh with a p-level of 2 (i.e., bi-quadratic elements). The predicted velocity profiles are significantly

improved, however not yet completely satisfactory. Still we are led to believe that the constant C depends on the mesh parameter h but with a weakened dependence at this p-level. Finally, we consider a 10×10 finite element mesh with a p-level of 6, where the total number of degrees of freedom is the same as for the previous to cases. The predicted velocity profiles are in excellent agreement with the benchmark solution and we are led to believe that at this p-level the dependence of the constant C on the mesh parameter h is negligible or nonexistent. Typically a p-level of 4 is sufficient to assure negligible or nonexistent dependence on the mesh parameter h.

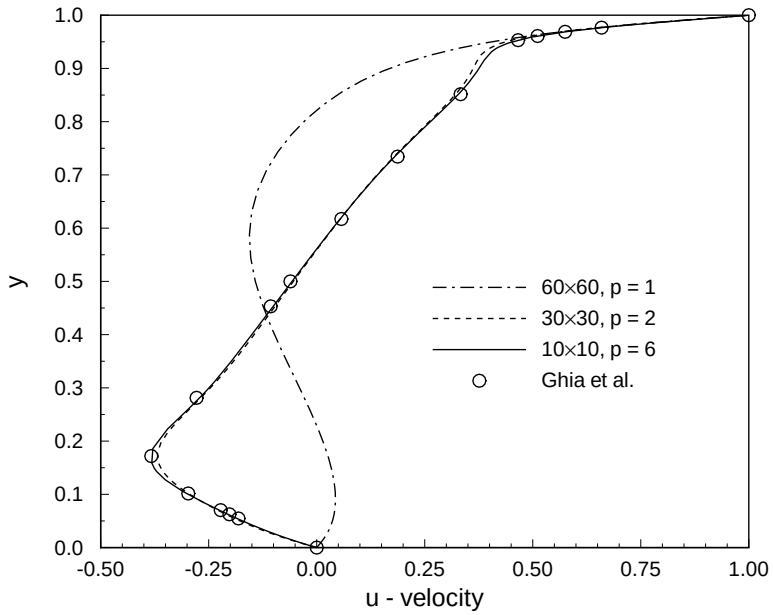

Fig. 2. u-velocity profiles along the vertical mid-line of the cavity at flow conditions $\mathrm{Re} = 10^3$.

The above illustrative example shows that if a non-equivalent least-squares functional is used to develop the finite element model, high-order expansions are desirable. In general, as we will show in subsequent chapters, high order expansions are always desirable when using *bona fide* least-squares based finite element models. If

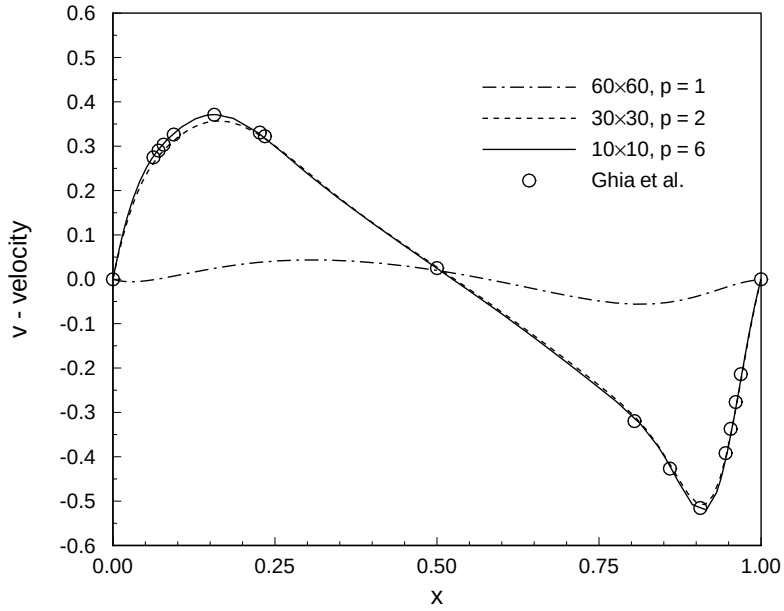

Fig. 3. v-velocity profiles along the horizontal mid-line of the cavity at flow conditions $\mathrm{Re} = 10^3$.

low-order expansions are to be used (i.e., p-levels of 1 or 2), it is best to use non-standard least-squares procedures such as collocation. This is the preferred procedure in the work presented by Jiang [55, 53, 54], Jiang et al. [58, 60, 59, 112], and Tang et al. [102, 101, 103], although they refer to the collocation solution as a reduced integration solution. However, in general, blind application of reduced integration techniques will not result in a collocation solution and should be avoided.

G. Nodal/modal expansions

Having motivated the use of high-order expansion, we present in this section some details on the high-order nodal/modal expansions used in this work. In accordance to the pre-determined level of practicality, the nodal/modal expansion considered herein have merely C^0 regularity across element boundaries.

Consider the two-dimensional case $\mathbf{u} = (u, v)$ and let $\mathcal{P}_h = \{\mathcal{Q}\}$ be a family of quadrilateral finite elements $\bar{\Omega}_e$ that make up the connected model $\bar{\Omega}^h$. We map $\bar{\Omega}_e$ to a bi-unit square $\hat{\Omega}_e = [-1, 1] \times [-1, 1]$, where $\boldsymbol{\xi} = (\xi_1, \xi_2) = (\xi, \eta)$ is a point in $\hat{\Omega}_e$. Over a typical element $\hat{\Omega}_e$, we approximate u by the expression

$$u(\xi, \eta) \approx u^{hp}(\xi, \eta) = \sum_{j=1}^{n} \Delta_j \, \varphi_j(\xi, \eta) \qquad \text{in } \hat{\Omega}_e \tag{2.16}$$

In a modal expansion, φ_j are tensor products of the one-dimensional C^0 p-type hierarchical basis

$$\psi_i(\xi) = \begin{cases} \frac{1-\xi}{2} & i = 1 \\ \left(\frac{1-\xi}{2}\right)\left(\frac{1+\xi}{2}\right) P_{p-2}^{\alpha,\beta} & 2 \le i \le p, \ p \ge 2 \\ \frac{1+\xi}{2} & i = p+1 \end{cases} \tag{2.17}$$

and Δ_j are coefficients associated with each of the modes of the hierarchical basis. In definition (2.17), $P_p^{\alpha,\beta}$ are the Jacobi polynomials of order p. We use ultraspheric polynomials corresponding to the choice $\alpha = \beta$ with $\alpha = \beta = 0$ or 1.

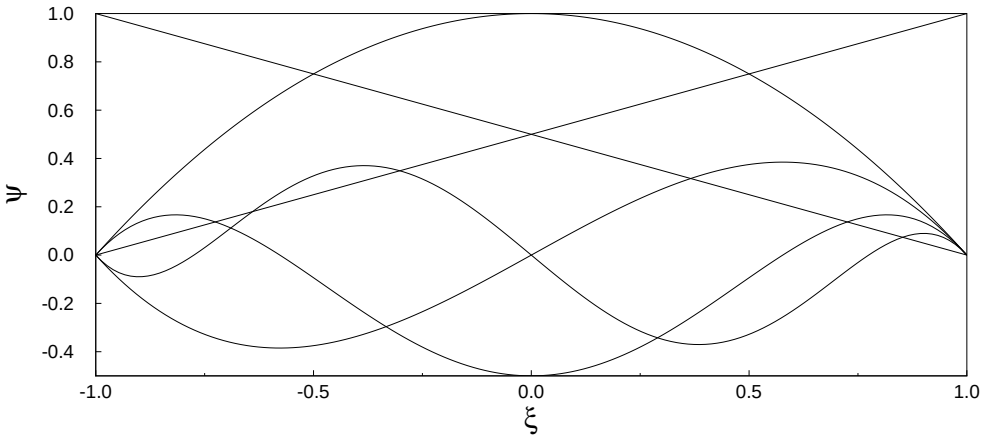

Fig. 4. C^0 p-type hierarchical modal basis. Shown is the case of $p = 5$. The p-bubbles are scaled by a factor of 4, for viewing ease.

Figure 4 shows the one-dimensional modal basis for the case $p = 5$. The linear basis or "hat-functions" ensure the C^0 continuity requirement across element boundaries and the p-bubbles hierarchically enrich the finite element space. Note that by construction the p-bubbles vanish at $\xi = -1$, $\xi = +1$ and have no nodes associated with them.

In a nodal expansion, φ_j are tensor products of the one-dimensional C^0 spectral nodal basis

$$h_i(\xi) = \frac{(\xi - 1)(\xi + 1)L_p'(\xi)}{p(p+1)L_p(\xi_i)(\xi - \xi_i)} \tag{2.18}$$

and Δ_j are nodal values due to the Kronecker delta property of the nodal basis. In Eq. (2.18), $L_p = P_p^{0,0}$ is the Legendre polynomial of order p and ξ_i denotes the location of the roots of $(\xi-1)(\xi+1)L_p'(\xi) = 0$ in the interval $[-1, 1]$. The set of points $\{\xi_i\}_{i=1}^{p+1}$ are commonly referred to as the Gauss-Lobatto-Legendre (GLL) points.

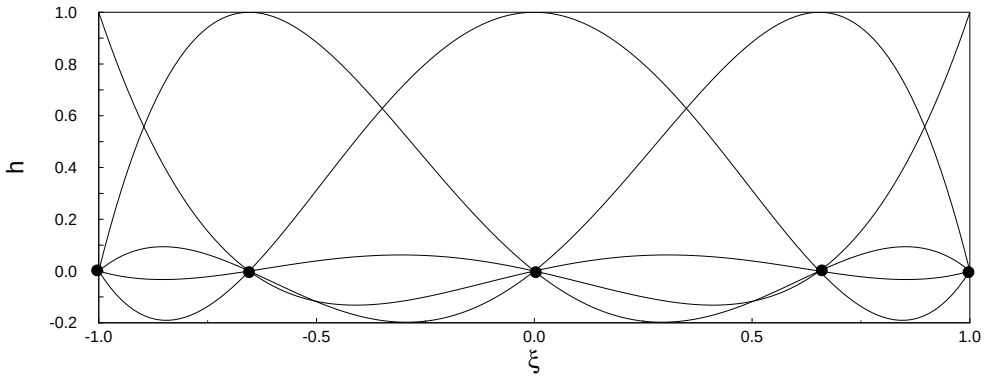

Fig. 5. C^0 p-type (spectral) nodal basis. Shown is the case of $p = 4$.

Figure 5 shows the one-dimensional nodal basis for the case $p = 4$. The location of the nodes coincides with the roots of the aforementioned Legendre polynomial and thus receives the name of a "spectral" basis. The Kronecker delta property is evident from the figure and is an attractive feature of this basis, as the coefficients Δ_j in (2.16) coincide with nodal values. The nodal basis is not hierarchical, i.e., all its "modes" are of polynomial order p.

In actual implementations the computationally more stable version of Eq. (2.18) is used to generate the nodal basis:

$$h_i(\xi) = \prod_{\substack{j=1 \\ j\neq i}}^{p+1} (\xi - \xi_j) \Big/ \prod_{\substack{j=1 \\ j\neq i}}^{p+1} (\xi_i - \xi_j) \ . \tag{2.19}$$

Details on the multidimensional construction of both the modal and nodal expansions can be found in Ref. [64].

We approximate the rest of the components of the vector valued function $\mathbf{u} = (u, v)$ in similar manner as we did for u in Eq. (2.16) and proceed to generate a system of linear algebraic equations at the element level using (2.8). The integrals in Eq. (2.8) are evaluated using Gauss quadrature rules. In our implementation the Gauss-Legendre rules are used for both the modal and nodal expansions, and *full integration* is used to evaluate the integrals.

The global system of equations is assembled from the element contributions using the direct stiffness summation assembly approach. The assembled system of equations can be written as

$$\begin{pmatrix} [K^{11}] & [K^{12}] \\ [K^{12}]^{\mathrm{T}} & [K^{22}] \end{pmatrix} \begin{pmatrix} \{\Delta^1\} \\ \{\Delta^2\} \end{pmatrix} = \begin{pmatrix} \{F^1\} \\ \{F^2\} \end{pmatrix} \tag{2.20}$$

where $\{\Delta^1\}, \{\Delta^2\}$ are the modal/nodal unknown coefficients associated with u and v. For details on standard finite element methods, such as mapping $\bar{\Omega}_e \rightleftarrows \hat{\Omega}_e$, numerical integration in $\hat{\Omega}_e$, and assembly using the direct stiffness summation approach, see Refs. [91, 94].

H. Solution procedures for SPD systems

The assembled system of linear equations resulting from the least-squares based finite element model will always have a symmetric positive definite (SPD) coefficient matrix.

It is thus appropriate to take full advantage of the symmetric positive definiteness by using solvers specially designed for such systems. In this section we briefly discuss the storage schemes and solution procedures used in the computational algorithms developed for the numerical solution of the least-squares finite element models.

1. Direct solvers

Given a SPD coefficient matrix $\underline{\mathbf{K}} \in \mathbb{R}^{N \times N}$ with bandwidth B, provided $N >> B$, and effective direct solution procedure is banded Cholesky factorization. The amount of work required for such an algorithm is approximately $N(B^2 + 3B)$ flops and N square roots [36]. This, of course, requires a suitable storage scheme, e.g., storing only the nonzero lower or upper triangular part in a $(B + 1)$ by N array. A more efficient storage scheme is the so called skyline storage, where B is allowed to vary from row to row and the data is stored in a one-dimensional array.

The constraint, $N >> B$, implies that a narrow band is always desirable. This is achieved, in the context of a finite element model, by numbering the local and global degrees of freedom in an optimal, problem dependent manner. This places severe restrictions on the size and geometry of the model. If we insist upon a banded direct solver, an alternate approach is to use graph-theory to minimize the bandwidth B of the matrix. For SPD matrices a popular choice is the Reverse Cuthill-McKee permutation (see Ref. [95] for details on permutations). Applying the permutation prior to the direct solve will always guarantee a minimum bandwidth, thus lifting the burden on the user to find an optimal degree of freedom numbering for a given problem.

2. Iterative solvers

Direct methods become impractical when N is large, e.g., large three-dimensional problems. In such cases, storage space may be limited in terms of available computer memory and solve times may become non-optimal in terms of CPU time.

In loose terms, iterative methods generate a sequence of approximate solutions $\{\mathbf{u}^k\}_{k=1}^{M}$ and essentially involve the matrix $\underline{\mathbf{K}}$ only for matrix-vector multiplications (`matvecs`). This implies that sparse-oriented storage schemes or element by element methods will prove useful in effectively computing the `matvecs`.

The performance of an iterative method is invariably measured on how quickly the iterates converge to within an acceptable tolerance, i.e., we want M small in achieving the prescribed tolerance. For a SPD coefficient matrix, an optimal choice are preconditioned conjugate gradient (PCG) methods, whose convergence rate is strongly dependent on the condition number of the (preconditioned) coefficient matrix [36, 95]. A suitable preconditioner will effectively lower the condition number of the coefficient matrix and result in fast convergence of the iterates. Ideally, the preconditiner would be a cheap, good approximation to the exact Cholesky factor of $\underline{\mathbf{K}}$. In this study we consider only a Jacobi preconditioner and a Symmetric Gauss-Seidel (SGS) preconditioner. Details of the PCG algorithms and construction of the Jacobi and SGS preconditioners can be found in Refs. [36, 95] among many others.

Application of the Gauss-Seidel preconditioner requires storage of the lower and upper parts of the assembled system of equations to perform the preconditioning step. The sparse storage scheme implemented for the computational algorithm is the compressed sparse row/column scheme, which is probably the most popular for storing general sparse matrices. The data structure uses three arrays with the following functions:

- a real one-dimensional array `GLK` containing real values K_{ij} stored row by row, from row 1 to N. The total size of `GLK` is `nnz` (number of non-zero entries in K_{ij}).

- an integer one-dimensional array `JA` containing the column addresses of the elements K_{ij} as stored in the array `GLK`. The total size of `JA` is also `nnz`.

- an integer one-dimensional array `IA` containing the address to the beginning of each row in the arrays `GLK` and `JA`. The total size of `IA` is $N + 1$.

Such data structure allows for fast `matvecs`, whose computational cost dominates each iterative step and thus its optimization is of paramount importance to the overall speed of the computations.

For sufficiently large N, even the sparse storage scheme may prove inconvenient. We therefore have to resort to storage-free techniques, also known as element-by-element solution algorithms, and implement a matrix-free version of the conjugate gradient method with a Jacobi preconditioner. The Jacobi preconditioner is easy and inexpensive to construct and apply, but is significantly of lower quality than the Gauss-Seidel preconditioner.

As implied earlier, the Gauss-Seidel preconditioner cannot be applied in a matrix-free setting, as it requires storage of the lower and upper parts of the assembled system of equations to perform the preconditioning step. It is necessary to emphasize that the matrix-free conjugate gradient algorithm with a Jacobi preconditioner does not require the assembly of a global matrix, not even an element matrix, which leads to tremendous savings in computer memory; and if implemented properly, considerable solve time speed-ups. Provided N is large enough, the matrix-free Jacobi CG solver will outperform the Gauss-Seidel CG solver in terms of CPU time – but never in terms of number of iterations required to converge within a specified tolerance.

CHAPTER III

VISCOUS INCOMPRESSIBLE FLUID FLOWS

In this chapter[†] we specialize the abstract formulation presented in Chapter II to the incompressible Navier-Stokes equations, governing the flow of viscous incompressible fluids – relevant to various engineering disciplines.

The numerical solution of the incompressible Navier-Stokes equations using least-squares based finite element models is among the most popular applications of least-squares methods. Least-squares formulations for incompressible flow circumvent the inf-sup condition, thus allowing equal-order interpolation of velocities and pressure, and result (after suitable linearization) in linear algebraic systems with a SPD coefficient matrix. This translates into easy algorithm development and leads to the use of robust and fast iterative solvers, resulting in substantial improvements over the weak form Galerkin finite element model – where the finite element spaces for velocities and pressure must satisfy the inf-sup compatibility condition and one must deal with an un-symmetric and indefinite coefficient matrix.

Recall from Chapter II, that to ensure the pre-determined "C^0 practicality" of the least-squares based finite element model the governing equations must be recast as an equivalent first-order system and the least-squares functional defined in terms of L_2 norms only. The incompressible Navier-Stokes equations, in particular the conservation of momentum equations, contain second-order derivatives implying $\mathbf{u} \in \mathbf{H}^2$ as a minimum requirement. A least-squares finite element model can indeed be

[†]Part of the numerical results reported in this chapter appear in the articles "Spectral/*hp* least-squares finite element formulation for the Navier-Stokes equations" by J. P. Pontaza and J. N. Reddy, *J. Comp. Phys.*, vol. 190, pp. 523–549, 2003 and "Space-time coupled spectral/*hp* least-squares finite element formulation for the incompressible Navier-Stokes equations" by J. P. Pontaza and J. N. Reddy, *J. Comp. Phys.*, accepted for publication. Copyright (2003) Elsevier Science.

developed by using the Navier-Stokes equations in their strong form, however such formulation would require C^1 regularity of the finite element spaces across inter-element boundaries as a minimum requirement (see Chapter VII for examples). In addition, if we are willing to upgrade to a "C^1 practicality" level, we are allowed to accommodate least-squares functionals defined in terms of H^1 norms for first-order operators.

Wishing to retain the "C^0 practicality" level we realize that auxiliary variables need to be introduced to recast the incompressible Navier-Stokes as an equivalent first order system. The most popular choice of an equivalent first-order system for the incompressible Navier-Stokes equations is the vorticity based first-order system. In two-dimensions the total number of variables is only increased by one and this formulation has the benefit of directly solving for a quantity of physical relevance, the vorticity. Formulations based on the velocity-pressure-vorticity first order system were first presented in the work of Jiang [55, 53, 54] and Jiang et al. [58, 60, 59]. However, Jiang's work has been mostly based on non-standard least-squares procedures: specifically, collocation (where the collocation points coincide with the Gauss-Legendre reduced integration points).

Yet another approach is to introduce the stresses as independent variables. This leads to a stress based first-order system. In two-dimensions the total number of variables is increased by three. Surana and co-workers prefer this first-order system in their work [110, 9].

A third option is to introduce all components of the gradient of the velocity vector field as independent variables. Such an approach was first suggested and studied by Cai et al. [20] and by Bochev et al. [12, 14]. In two dimensions the total number of variables is increased by four and this formulation has the added benefit to easily compute (in the post-processing stage) physical quantities of interest that are linear

combinations of the partial derivatives of the velocity vector field, e.g., vorticity and stresses.

Each of the three formulations has practical and mathematical advantages and disadvantages. For example, the L_2 least-squares functional for the vorticity based formulation fails to define an equivalent norm in H^1 for particular sets of boundary conditions (details are given in Refs. [13] and [11]) and thus the optimality of such formulation cannot be determined a priori (see Chapter II). Nevertheless, the formulation has had widespread acceptance, and in actual implementations performs exceptionally well. The L_2 least-squares functional for the stress based formulation is always non-equivalent, regardless of the choice of boundary conditions [13]. By adding additional constraints to weaken the dependencies between variables, the L_2 least-squares functional for the velocity gradient based first-order system can define an equivalent norm in H^1 [20, 12, 14]. However, the velocity gradient based system is the most expensive of the three choices considered here.

We start by presenting the governing equations and the aforementioned equivalent first-order systems where vorticity, stresses, or velocity gradients are introduced as auxiliary variables. Verification procedures are presented for our computational algorithm, showing exponentially fast decay (spectral convergence) of suitable error measures as the element expansion order (p-level) is increased. The analysis is carried out in geometrically un-distorted and distorted meshes, to reveal that the exponentially fast rate of decay is preserved. Even though the L_2 least-squares functional used for the vorticity and stress based formulations is non-equivalent, spectral convergence rates are achieved. Through the verification procedures we show that the vorticity based formulation is preferred, as it is associated with fewer degrees of freedom and yields same order of accuracy results as the other more expensive equivalent first-order formulations.

Results for verification benchmarks, such as flow over a backward-facing step and three-dimensional lid-driven cavity flow are presented. The chapter concludes with a validation benchmark: flow past a circular cylinder, where predictions of various flow metrics are compared against reliable experimental data and previously reported high-order numerical simulation results and found to be in excellent agreement.

A. The incompressible Navier-Stokes equations

We consider the solution of the Navier-Stokes equations governing incompressible flow, which in dimensionless form can be stated as follows:

Find the velocity $\mathbf{u}(\mathbf{x},t)$ and pressure $p(\mathbf{x},t)$ such that

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\,\mathbf{u} + \nabla p - \frac{1}{\mathrm{Re}}\nabla\cdot\left[(\nabla\mathbf{u}) + (\nabla\mathbf{u})^T\right] = \mathbf{f} \qquad \text{in } \Omega\times(0,\tau] \tag{3.1}$$

$$\nabla\cdot\mathbf{u} = 0 \qquad \text{in } \Omega\times(0,\tau] \tag{3.2}$$

$$\mathbf{u}(\mathbf{x},0) = {}^0\mathbf{u}(\mathbf{x}) \qquad \text{in } \Omega \tag{3.3}$$

$$\mathbf{u} = \mathbf{u}^s \qquad \text{on } \Gamma_u\times(0,\tau] \tag{3.4}$$

$$\hat{\mathbf{n}}\cdot\boldsymbol{\sigma} = \mathbf{f}^s \qquad \text{on } \Gamma_f\times(0,\tau] \tag{3.5}$$

where $\Gamma = \Gamma_u \cup \Gamma_f$ and $\Gamma_u \cap \Gamma_f = \emptyset$, τ is a real number (time) > 0, Re is the Reynolds number, $\nabla\cdot{}^0\mathbf{u} = 0$, $\boldsymbol{\sigma} = -p\,\mathbf{I} + 1/\mathrm{Re}\left[(\nabla\mathbf{u}) + (\nabla\mathbf{u})^T\right]$, $\mathbf{f}$ is a dimensionless force, $\hat{\mathbf{n}}$ is the outward unit normal on the boundary of Ω, $\mathbf{u}^s$ is the prescribed velocity on the boundary Γ_u, $\mathbf{f}^s$ are the prescribed tractions on the boundary Γ_f, and in Eq. (3.3) the initial conditions are given. We assume the problem is well posed.

In situations where outflow boundary conditions need to be modelled, the Navier-Stokes equations in the $\nabla^2\mathbf{u}$ form are preferred [96]. In such cases, (using the incompressibility constraint given in Eq. (3.2)) we would drop the $(\nabla\mathbf{u})^{\mathrm{T}}$ term in Eq. (3.1),

and the boundary conditions in Eq. (3.5) would then become

$$\hat{\mathbf{n}} \cdot \tilde{\boldsymbol{\sigma}} = \tilde{\mathbf{f}}^s \qquad \text{on } \Gamma_f \times (0, \tau] \tag{3.6}$$

where $\tilde{\boldsymbol{\sigma}}$ is a pseudo-stress, $\tilde{\boldsymbol{\sigma}} = -p\,\mathbf{I} + (1/\mathrm{Re})\,\nabla\mathbf{u}$, and $\tilde{\mathbf{f}}^s$ are the prescribed pseudo-tractions on the boundary Γ_f, prescribed as zero at an outflow boundary.

1. The vorticity based first-order system

Introducing the vorticity vector, $\boldsymbol{\omega} = \nabla \times \mathbf{u}$, then by making use of the vector identity

$$\nabla \times \nabla \times \mathbf{u} = -\nabla^2 \mathbf{u} + \nabla\,(\nabla \cdot \mathbf{u})$$

and in view of the incompressibility constraint given in Eq. (3.2), the non-stationary Navier-Stokes equations, Eqs. (3.1)-(3.5), can be replaced by their first-order system equivalent:

Find the velocity $\mathbf{u}\,(\mathbf{x}, t)$, pressure $p\,(\mathbf{x}, t)$, and vorticity $\boldsymbol{\omega}\,(\mathbf{x}, t)$ such that

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\,\mathbf{u} + \nabla p + \frac{1}{\mathrm{Re}}\,\nabla \times \boldsymbol{\omega} = \mathbf{f} \qquad \text{in } \Omega \times (0, \tau] \tag{3.7}$$

$$\boldsymbol{\omega} - \nabla \times \mathbf{u} = \mathbf{0} \qquad \text{in } \Omega \times (0, \tau] \tag{3.8}$$

$$\nabla \cdot \mathbf{u} = 0 \qquad \text{in } \Omega \times (0, \tau] \tag{3.9}$$

$$\nabla \cdot \boldsymbol{\omega} = 0 \qquad \text{in } \Omega \times (0, \tau] \tag{3.10}$$

$$\mathbf{u}\,(\mathbf{x}, 0) = {}^0\mathbf{u}\,(\mathbf{x}) \qquad \text{in } \Omega \tag{3.11}$$

$$\mathbf{u} = \mathbf{u}^s \qquad \text{on } \Gamma_u \times (0, \tau] \tag{3.12}$$

$$\boldsymbol{\omega} = \boldsymbol{\omega}^s \qquad \text{on } \Gamma_\omega \times (0, \tau] \tag{3.13}$$

The seemingly redundant equation (3.10) is needed in the three dimensional case to make the system of equations uniformly elliptic [13]. In addition, $\Gamma_u \cap \Gamma_\omega = \emptyset$, i.e., if

velocity is specified at a boundary, vorticity need not be specified there. This implies that no artificial boundary conditions for vorticity need to be devised at boundaries where the velocity is specified. Outflow boundary conditions are imposed in a weak sense through the least-squares functional as discussed in Chapter II.

2. The stress based first-order system

In the stress based first-order system we introduce the components of the stress tensor,

$$\underline{\mathbf{T}} = (\nabla \mathbf{u}) + (\nabla \mathbf{u})^T \ ,$$

as additional independent variables. Note that the stress tensor is symmetric, implying that $\mathrm{T}_{ij} = \mathrm{T}_{ji}$.

In terms of $\underline{\mathbf{T}}$, the non-stationary Navier-Stokes equations, Eqs. (3.1)-(3.5), can be replaced by their first-order system equivalent:

Find the velocity $\mathbf{u}(\mathbf{x}, t)$, pressure $p(\mathbf{x}, t)$, and stresses $\underline{\mathbf{T}}(\mathbf{x}, t)$ such that

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\,\mathbf{u} + \nabla p - \frac{1}{\mathrm{Re}} \nabla \cdot \underline{\mathbf{T}} = \mathbf{f} \qquad \text{in } \Omega \times (0, \tau] \tag{3.14}$$

$$\underline{\mathbf{T}} - \left[(\nabla \mathbf{u}) + (\nabla \mathbf{u})^T\right] = \underline{\mathbf{0}} \qquad \text{in } \Omega \times (0, \tau] \tag{3.15}$$

$$\nabla \cdot \mathbf{u} = 0 \qquad \text{in } \Omega \times (0, \tau] \tag{3.16}$$

$$\mathbf{u}(\mathbf{x}, 0) = {}^0\mathbf{u}(\mathbf{x}) \qquad \text{in } \Omega \tag{3.17}$$

$$\mathbf{u} = \mathbf{u}^s \qquad \text{on } \Gamma_u \times (0, \tau] \tag{3.18}$$

$$\hat{\mathbf{n}} \cdot \underline{\mathbf{T}} = \mathbf{T}^s \qquad \text{on } \Gamma_{\mathrm{T}} \times (0, \tau] \tag{3.19}$$

Outflow boundary conditions are imposed in a weak sense through the least-squares functional or can be modelled and imposed strongly by using the components of the stress tensor.

3. The velocity gradient based first-order system

To define the first-order velocity-pressure-velocity gradient system, we introduce all components of the gradient of the velocity vector field as independent variables. Viewing the d-vector functions as column vectors and the new d^2-vector functions as either block column vectors or matrices, using the notation of Ref. [20], we define the velocity gradients, $\underline{\mathbf{U}}$, as follows:

$$\underline{\mathbf{U}} = \nabla (\mathbf{u})^{\mathrm{T}}$$

In terms of $\underline{\mathbf{U}}$, the non-stationary Navier-Stokes equations, Eqs. (3.1)-(3.5), can be replaced by their first-order system equivalent:

Find the velocity $\mathbf{u}(\mathbf{x}, t)$, pressure $p(\mathbf{x}, t)$, and velocity gradients $\underline{\mathbf{U}}(\mathbf{x}, t)$ such that

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \underline{\mathbf{U}})^{\mathrm{T}} + \nabla p - \frac{1}{\mathrm{Re}} (\nabla \cdot \underline{\mathbf{U}})^{\mathrm{T}} = \mathbf{f} \qquad \text{in } \Omega \times (0, \tau] \tag{3.20}$$

$$\underline{\mathbf{U}} - \nabla (\mathbf{u})^{\mathrm{T}} = \underline{\mathbf{0}} \qquad \text{in } \Omega \times (0, \tau] \tag{3.21}$$

$$\nabla \cdot \mathbf{u} = 0 \qquad \text{in } \Omega \times (0, \tau] \tag{3.22}$$

$$\nabla \times \underline{\mathbf{U}} = \underline{\mathbf{0}} \qquad \text{in } \Omega \times (0, \tau] \tag{3.23}$$

$$\nabla (\operatorname{tr} \underline{\mathbf{U}}) = \mathbf{0} \qquad \text{in } \Omega \times (0, \tau] \tag{3.24}$$

$$\mathbf{u}(\mathbf{x}, 0) = {}^0\mathbf{u}(\mathbf{x}) \qquad \text{in } \Omega \tag{3.25}$$

$$\mathbf{u} = \mathbf{u}^s \qquad \text{on } \Gamma_u \times (0, \tau] \tag{3.26}$$

$$\hat{\mathbf{n}} \cdot \underline{\mathbf{U}} = \mathbf{U}^s \qquad \text{on } \Gamma_U \times (0, \tau] \tag{3.27}$$

where $\operatorname{tr} \underline{\mathbf{U}} = \sum_{i=1}^{d} U_{ii}$. For $d = 2$ (the two-dimensional case) we may also specify $\hat{\mathbf{t}} \cdot \underline{\mathbf{U}} = \hat{\mathbf{t}} \cdot \nabla (\mathbf{u}^s(\mathbf{x}))^{\mathrm{T}}$ on Γ_u , where $\hat{\mathbf{t}}$ is the unit tangent on the boundary of Ω.

Similarly, for $d = 3$, we may specify $\hat{\mathbf{t}}_1 \cdot \nabla (\mathbf{u}^s(\mathbf{x}))^{\mathrm{T}}$ and $\hat{\mathbf{t}}_2 \cdot \nabla (\mathbf{u}^s(\mathbf{x}))^{\mathrm{T}}$ on Γ_u, where $\hat{\mathbf{t}}_1$ and $\hat{\mathbf{t}}_2$ are unit tangent vectors on the boundary of Ω.

The seemingly redundant equations (3.23) and (3.24) are needed to make the system achieve H^1-coercivity [20, 12, 14]. Outflow boundary conditions are imposed in a weak sense through the least-squares functional or can be modelled and imposed strongly by using the components of the gradient of the velocity vector field.

B. Numerical examples: verification benchmarks

Having presented the three equivalent first-order systems for the incompressible Navier-Stokes, we use the procedures outlined in Chapter II to develop and arrive at the least-squares based finite element model. Depending on the size of the problem, the storage scheme and solution procedure are chosen as discussed in Chapter II.

In this section we are concerned with the *verification* of the spectral/*hp* least-squares based finite element computational algorithm, we follow the verification procedures suggested in Ref. [80]. The algorithm verification procedure must convincingly demonstrate that the governing equations, in this case the incompressible Navier-Stokes, are being solved consistently. The process of verification is thus a quantitative study of the accuracy of the numerical solution which deals with investigation of spatial and temporal convergence rates and independence of solutions to coordinate transformations. Theoretical a priori error estimates may prove useful in such a study, but are not considered to be part of the verification procedure: the code must be used to demonstrate that the optimal (or sub-optimal) spatial and temporal convergence rates are achieved.

Rigorous verification requires proof, in a practical sense, that the computational algorithm accurately represents the mathematical model (i.e., the governing equa-

tions) and its solution. With this is mind, it is of great importance to choose suitable error measures that convincingly show, through a convergence study, that the governing equations are being solved consistently.

A suitable error measure is, for example, the L_2 least-squares functional ($\mathcal{J}$). In fact, this error measure is of great value and importance in the verification procedure. Convergence of this measure to zero implies that the L_2 norm of the residuals of the governing equations are going to zero, i.e., conservation of mass and momentum are being satisfied. Equally important error measures are the L_2 norms of the difference between the numerical solution and the analytic solution. Convergence of this measure to zero implies that the numerical solution approaches the exact solution. The aforementioned measures will be used to demonstrate that the computational algorithm solves the governing equations consistently, i.e., verification.

In the following we present two analytic solutions to the incompressible Navier-Stokes equations and use them for the purpose of algorithm verification. The first benchmark problem is well suited for verification of the stationary (steady) equations and the second benchmark problem for the non-stationary (unsteady) case. A series of benchmark problems in the form of well-documented published numerical studies are then used to further verify the computational algorithm. Among these benchmarks we consider flow over a backward-facing step and two- and three-dimensional lid-driven cavity flows.

1. Kovasznay flow

The first benchmark problem for verification is an analytic solution to the two-dimensional, stationary incompressible Navier-Stokes due to Kovasznay [66]. The spatial domain in which Kovasznay's solution is defined is taken here as the bi-unit square $\bar{\Omega} = [-0.5, 1.5] \times [-0.5, 1.5]$. The solution is given by

$$\begin{aligned} u(x,y) &= 1 - e^{\lambda x}\,\cos(2\pi y) \\ v(x,y) &= \frac{\lambda}{2\pi}\,e^{\lambda x}\,\sin(2\pi y) \\ p(x,y) &= p_0 - \frac{1}{2}\,e^{2\lambda x} \end{aligned} \tag{3.28}$$

where $\lambda = \mathrm{Re}/2 - (\mathrm{Re}^2/4 + 4\pi^2)^{1/2}$ and p_0 is a reference pressure (an arbitrary constant). Figure 6a shows u-velocity contours of the exact solution for $\mathrm{Re} = 40$.

a. p-refinement study

First, we perform a p-refinement study. For such a study, we choose and fix a spatial discretization $\bar{\Omega}^h$, and systematically increase the p-level of the element approximation functions. Figure 6b shows the discretization of the domain, $\bar{\Omega}^h$, for the p-refinement study. The discretization is a non-uniform mesh of 8 quadrilateral finite elements $\bar{\Omega}_e$. Having chosen appropriate error measures, these measures should decay exponentially fast as the p-level is increased. In a logarithmic-linear scale the expected rate of convergence would appear as a straight line.

The incompressible Navier-Stokes equations in the vorticity based first-order form are used to develop the least-squares finite element model. Nodal expansions (spectral elements) are used to obtain the discrete model. The exact solution, given by Eq. (3.28), is used to prescribe Dirichlet velocity boundary conditions on Γ and pressure is specified at a point. No boundary conditions for the vorticity are necessary.

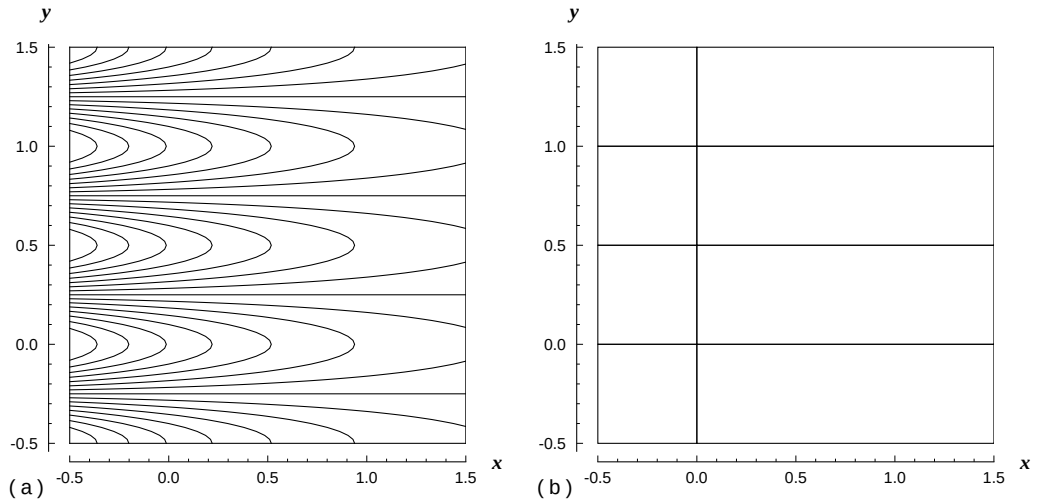

Fig. 6. Kovasznay flow: (a) u-velocity contours of the exact solution for Re = 40; (b) computational domain for p-refinement study.

The discrete system is linearized using Newton's method and the resulting linear algebraic system of equations with a symmetric positive definite coefficient matrix solved using Cholesky factorization at each Newton step. Nonlinear convergence was declared when the relative norm of the residual in velocities, $\|\Delta\mathbf{u}^{hp}\|/\|\mathbf{u}^{hp}\|$, was less than 10^{-4}, which typically required five Newton iterations.

Even though the resulting linear algebraic system of equations lends itself naturally to be solved using robust iterative methods, such as preconditioned conjugate gradient methods, we have chosen to use a direct solver. The reason for this choice, during the verification stages, is twofold. First, we prevent iterative convergence errors to pollute the error measures for verification of the algorithm. Second, the results obtained from the direct solver are to be used to verify results obtained from the iterative solver once the algorithm has been verified.

In Fig. 7 we plot the L_2 least-squares functional and L_2 error of the velocity, pressure, and vorticity fields as a function of the expansion order in a logarithmic-

linear scale. Exponentially fast decay (spectral convergence) of the L_2 least-squares functional and L_2 error is observed. For a p-level of 9, the L_2 norm of the error in velocities and pressure is already of machine order zero, meaning that the computed velocity and pressure fields are exactly those given by the analytic solution. Of course, for practical purposes, a satisfactory solution could already be claimed at p-levels of 5 or higher.

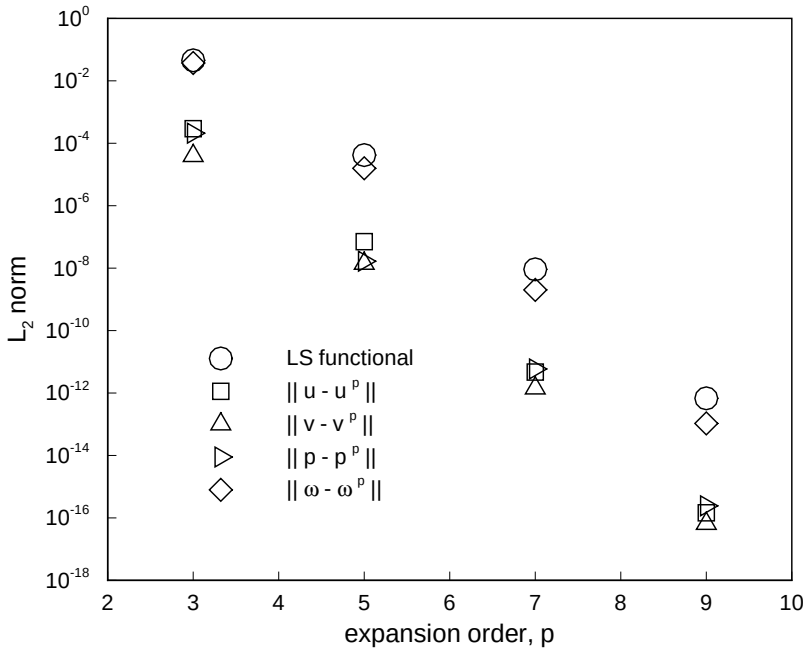

Fig. 7. Decay of the least-squares functional and convergence of the velocity, pressure, and vorticity fields to the Kovasznay solution in the L_2-norm. p-convergence.

b. h-refinement study

Next, we perform a h-refinement study. For such a study, we fix the p-level of the element approximation functions, and systematically refine the mesh. The error measures should decay at an algebraic rate as the mesh is refined. In a log-log scale the expected rate of convergence would appear as a straight line. For a fixed p-level of 4, standard elliptic theory predicts an optimal convergence rate of 5 in the L_2 norm for velocities.

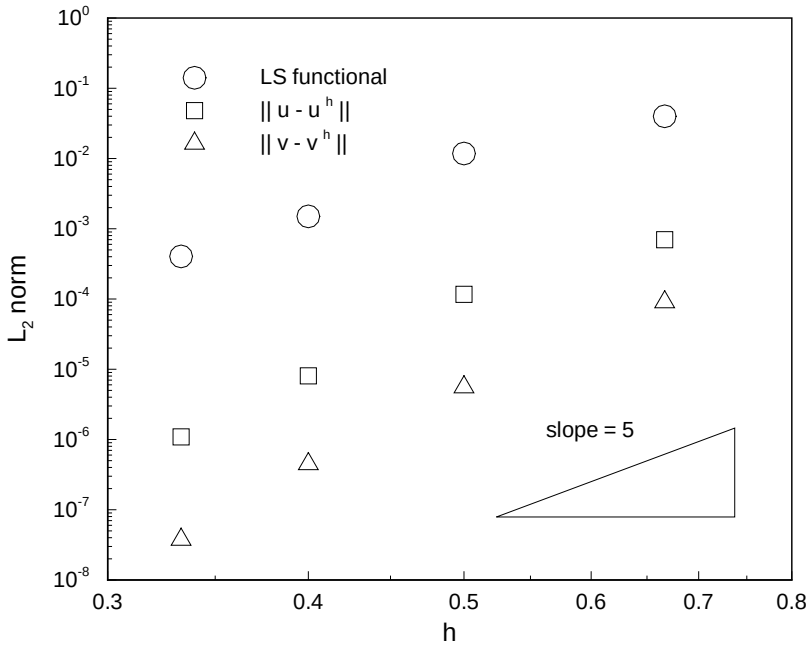

Fig. 8. Decay of the least-squares functional and convergence of the velocity field to the Kovasznay solution in the L_2-norm. h-convergence.

Five different uniform meshes are used to perform the h-refinement study. The meshes varied successively from 3×3 to 6×6 uniformly spaced elements in $\bar{\Omega}$. In Fig. 8 we plot the L_2 least-squares functional and L_2 error of the velocity field for successively refined meshes in a log-log scale. We see that an algebraic convergence rate slightly better than that predicted by standard elliptic theory is achieved.

c. Distorted meshes

Having demonstrated that the error measures tend to zero through h and/or p refinement studies using meshes aligned with an orthogonal coordinate system is generally not yet sufficient. To fully exercise all the terms in the finite element formulation, specifically cross-derivative terms, we still must perform a h or p refinement study using a distorted mesh. To illustrate the procedure, we perform a p-refinement study for the geometrically distorted mesh shown in Fig. 9.

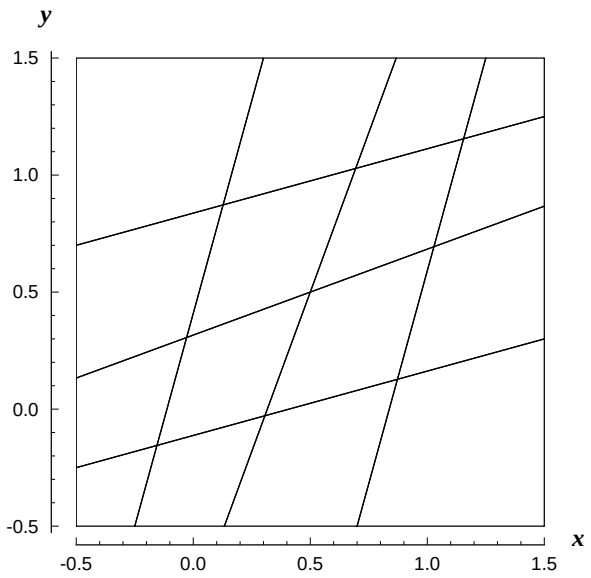

Fig. 9. Kovasznay flow: computational domain consisting of 4×4 geometrically distorted elements.

In Fig. 10 we plot the L_2 least-squares functional and L_2 error of the velocity, pressure, and vorticity fields as a function of the expansion order in a logarithmic-linear scale for the analysis in the geometrically distorted mesh. Exponentially fast decay (spectral convergence) of the L_2 least-squares functional and L_2 error is observed. As expected, for the distorted mesh, slightly higher p-levels are needed to obtain the same level of accuracy when compared to the geometrically undistorted mesh results.

Having performed a h- and/or p-refinement study in geometrically distorted and undistorted meshes, we have conclusively shown that the error measures tend to zero. Hence, the verification procedures have been completed successfully. Although the verification results shown here correspond to the vorticity based formulation, similar results are obtained for the stress and velocity gradient based formulations.

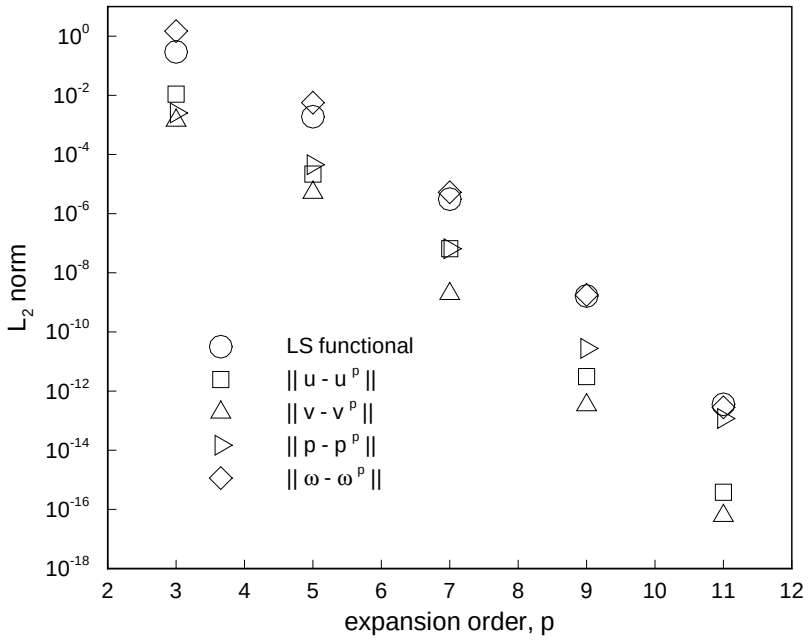

Fig. 10. Distorted mesh study: decay of the least-squares functional and convergence of the velocity, pressure, and vorticity fields to the Kovasznay solution in the L_2-norm. p-convergence for the geometrically distorted mesh.

d. Cost comparison

Also of interest is the cost and relative performance of the three equivalent first-order systems. A cost comparison between the vorticity, stress, and velocity gradient based formulations for the Kovasznay problem is illustrated in Fig. 11, by plotting the L_2 error of the velocity field as a function of the total number of degrees of freedom in a logarithmic-linear scale. Clearly, the vorticity based formulation is preferred because, for a given number of degrees of freedom, it allows for a more accurate solution at higher p-levels. Or equivalently, for a desired order of accuracy, a lower cost is incurred.

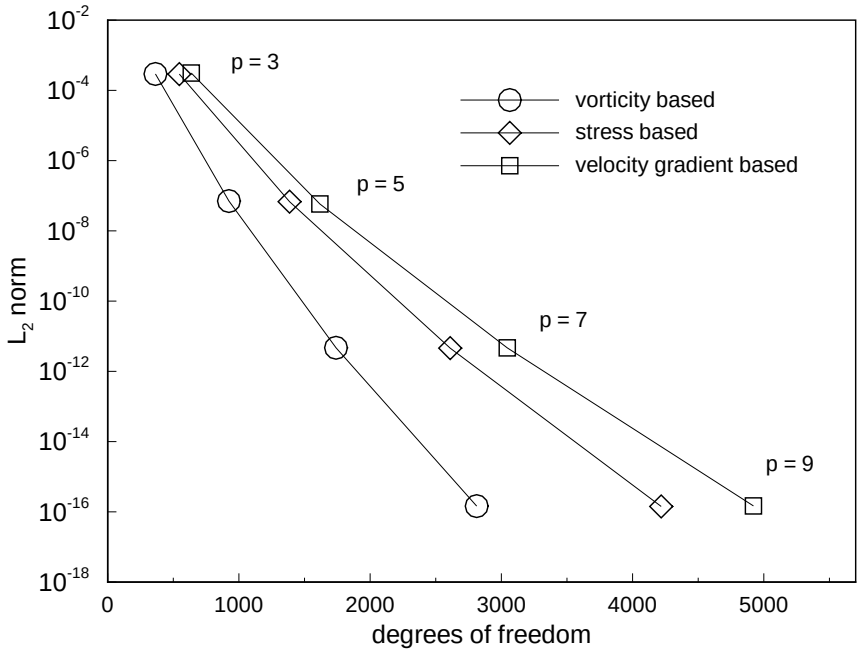

Fig. 11. Convergence of the velocity field to the Kovasznay solution in the L_2-norm for the vorticity, stress, and velocity gradient based first-order systems. Cost comparison.

The lower cost of the vorticiy based formulation is due to the fact that for $d = 2$ (the two-dimensional case) the vorticity first-order decomposition introduces only one component of the vorticity vector as an additional variable, whereas the stress first-order decomposition introduces three additional variables and the velocity gradient first-order decomposition introduces four additional variables. For $d = 3$ the vorticity first-order decomposition introduces the three components of the vorticity vector as additional variables, whereas the stress first-order decomposition introduces six additional variables and the velocity gradient first-order decomposition introduces nine additional variables.

In addition we note that for the p-convergence study, both the *non-equivalent* and *norm-equivalent* formulations achieve spectral convergence. Clearly, the fact that the L_2 least-squares functional for the vorticity and stress based formulations is not H^1-norm equivalent and hence defines a *non-equivalent* formulation; does not imply that the method is not optimal. It simply means that the optimality of the resulting method cannot be established a priori using standard elliptic theory. Similar observations were reported by Proot and Gerrtisma [88], for *non-equivalent* and *norm-equivalent* formulations, when verifying spectral convergence for the vorticity based first-order Stokes equations.

2. A manufactured solution

Having verified the computational algorithm with a stationary (steady) solution to the incompressible Navier-Stokes equations, we now wish to verify the algorithm with a non-stationary (unsteady) problem. The second benchmark problem for verification is based on a closed form manufactured solution. The basic idea behind manufactured solutions is to simply come up with an exact solution, preferably one that is infinitely differentiable (i.e., smooth), not trivially reproduced by the element approximation functions (i.e., no polynomials), and will exercise all the terms in the governing equation (or at least the ones we are interested in exercising).

We consider two-dimensional, unsteady flow in the bi-unit square, $\bar{\Omega} = [-1, 1] \times [-1, 1]$. We prescribe the closed form analytic solution to the non-stationary, incompressible Navier-Stokes equations to be of the form,

$$\begin{aligned} u(x, y, t) &= a_u + b_u \cos(\omega_u x) \sin(\omega_u y) \cos(\omega_{u,\tau} t) \\ v(x, y, t) &= a_v + b_v \sin(\omega_v x) \cos(\omega_{v,\tau} t) \\ p(x, y, t) &= a_p + b_p \sin(\omega_p y) \cos(\omega_{p,\tau} t) \end{aligned} \tag{3.29}$$

The prescribed analytic solution does not satisfy conservation of momentum or conservation of mass, i.e., if the above velocity and pressure fields are substituted into Eq. (3.1) (with $\mathbf{f} = \mathbf{0}$) and Eq. (3.2) a residual term for each of the equations will result. For the conservation of momentum, Eq. (3.1), a residual $\boldsymbol{\mathcal{R}}^{\text{mom}}$ will result. Likewise, for the conservation of mass given by Eq. (3.2), a residual $\mathcal{R}^{\text{con}}$ will result. These residuals are simply treated as source terms, belonging to their respective equations, that produce the desired (prescribed) solution. These source terms or residuals are commonly referred to as *the consistent forcing functions*.

The choice of constants in Eq. (3.29) are shown in Table I. Here we have chosen

$\omega_u = \omega_v = \omega_p = \omega = m\pi$ and $\omega_{u,\tau} = \omega_{v,\tau} = \omega_{p,\tau} = \omega_\tau = m\pi$, with $m = 2$. The higher the value of m and the lower the value of the constants b_u, b_v, b_p, the more challenge for the computational algorithm to reproduce the exact solution. As m is increased the solution becomes more oscillatory, the amplitude of the oscillations being controlled by the constants b_u, b_v, b_p. The smaller these constants are chosen, the harder it is to detect the oscillations.

Table I. Numerical values of the constants used in the manufactured solution for the incompressible Navier-Stokes equations.

	u	v	p
a	1.0	0.0	0.0
b	0.4	0.5	0.5
ω	2π	2π	2π
ω_τ	2π	2π	2π

For the purposes of verification, we need only consider a p-refinement study in space-time. The reason behind this is that when going from the stationary to the non-stationary formulation only minor changes are needed in the source code, mostly data management changes are made.

a. p-refinement study

The connected model in space-time for the bi-unit square, ${}_s^{s+1}\bar{\Lambda}^h = \bar{\Omega}^h \times [t_s, t_{s+1}]$, is shown in Figure 12. It consists of a 4×4 uniform finite element mesh in space and a single element layer in time. Having chosen the space-time discretization, we now systematically increase the p-levels of the element approximation functions (in space-time) and expect the error measures to decay exponentially fast as the p-level is increased. As before, in a logarithmic-linear scale the expected rate of convergence would appear as a straight line.

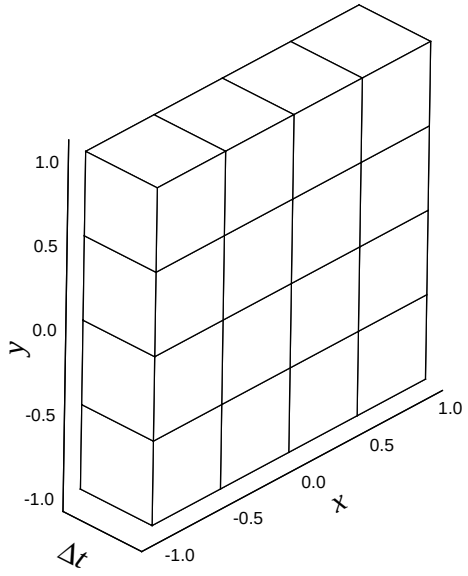

Fig. 12. Space-time computational domain and mesh for the bi-unit square on which the non-stationary incompressible Navier-Stokes problem is defined. The computational domain is a space-time strip, denoted by ${}_{s}^{s+1}\bar{\Lambda}^{h}$, with $\Delta t = t_{s+1} - t_s$.

The non-stationary incompressible Navier-Stokes equations in the vorticity first-order form were discretized using the space-time coupled least-squares finite element formulation with nodal expansions in space-time. We take $\Delta t = t_{s+1} - t_s = 0.5$ and compute the solution in a single space-time strip for increasing expansion orders, p, in space-time. The exact solution was used to compute the initial condition and Dirichlet boundary conditions for the velocity components. Pressure was prescribed only at a point, i.e., a space-time line. In Fig. 13 we plot the L_2 least-squares functional and L_2 error of the velocity, pressure, and vorticity fields in space-time as a function of the expansion order in a logarithmic-linear scale. Exponentially fast decay (spectral convergence) of the L_2 least-squares functional and L_2 error is observed in space-time. Note that only algebraic decay in time would be observed using a space-time decoupled formulation.

Implicitly, Fig. 13 also verifies the unconditional stability of the space-time cou-

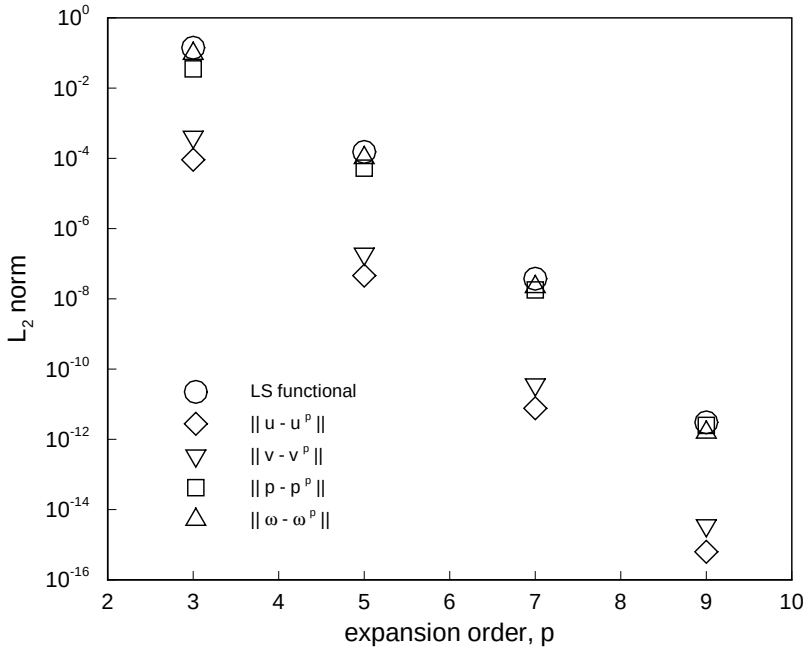

Fig. 13. Decay of the least-squares functional and convergence of the velocity, pressure, and vorticity space-time fields to the exact solution in the L_2 norm.

pled formulation. In a space-time decoupled formulation as the order of the time approximation is increased the stability region for the time approximation decreases and a smaller time increment must be used to remain stable. Here we have continuously increased the order of the time approximation and kept $\Delta t = t_{s+1} - t_s$ constant at 0.5. This completes the verification procedures (by means of convergence studies) for the incompressible Navier-Stokes equations.

3. Flow over a backward-facing step

We consider two-dimensional, steady flow over a backward-facing step at Re = 800. The geometry and boundary conditions are taken from the benchmark solution of Gartling [33] and are shown in Fig. 14. As shown in Fig. 14 the standard step geometry was simplified by excluding the channel portion upstream of the step. The boundary conditions for the step geometry include the no-slip condition at all solid surfaces and a parabolic inlet velocity profile given by $u(y) = 24y(0.5 - y)$ for $0 \leq y \leq 0.5$. The Reynolds number is based on the mean inlet velocity and height of the channel.

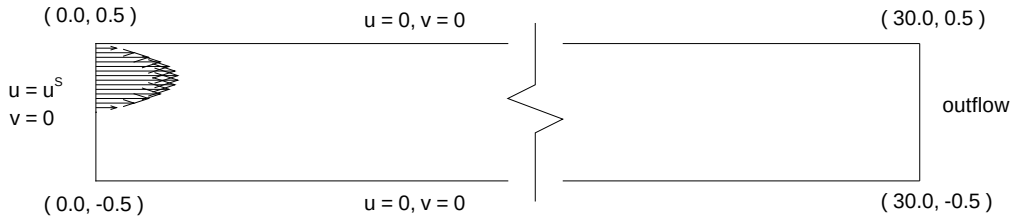

Fig. 14. Geometry and boundary conditions for flow over a backward-facing step.

The outflow boundary condition is of relevance in the context of the least-squares finite element formulation and is discussed in the following. Instead of imposing an outflow boundary condition in a strong sense we impose it in a weak sense through the least-squares functional. For example, if we use the vorticity based first-order system the L_2 least-squares functional is now

$$\mathcal{J}\left(\mathbf{u}, p, \boldsymbol{\omega}; \mathbf{f}, \tilde{\mathbf{f}}^s\right) = \frac{1}{2}\Big(\left\| \left(\mathbf{u}\cdot\nabla\right)\mathbf{u} + \nabla p + \frac{1}{\mathrm{Re}}\nabla\times\boldsymbol{\omega} - \mathbf{f} \right\|_0^2 + \left\| \boldsymbol{\omega} - \nabla\times\mathbf{u} \right\|_0^2 + \left\| \nabla\cdot\mathbf{u} \right\|_0^2 + \left\| \nabla\cdot\boldsymbol{\omega} \right\|_0^2 + \left\| \hat{\mathbf{n}}\cdot\tilde{\boldsymbol{\sigma}} - \tilde{\mathbf{f}}^s \right\|_{0,\Gamma_{\mathrm{outflow}}}^2 \Big) \quad (3.30)$$

where $\tilde{\boldsymbol{\sigma}}$ is a pseudo-stress (see Eq. (3.6)), $\tilde{\boldsymbol{\sigma}} = -p\mathbf{I} + (1/\mathrm{Re})\,\nabla\mathbf{u}$, and $\tilde{\mathbf{f}}^s$ are the prescribed pseudo-tractions, typically taken to be zero at an outflow boundary. The

strong outflow boundary condition $p = 0$ for the vorticity based first-order system or $p = 0$, $\partial v/\partial n = 0$ for the velocity gradient based first-order system also give good results. However, we prefer the weak imposition of outflow boundary conditions through the least-squares functional as it allows for more freedom in their modelling.

We discretize the domain, $\bar{\Omega} = [0, 30] \times [-0.5, 0.5]$, using 20 finite elements: two elements along the height of the channel and 10 uniformly spaced elements along the length of the channel. The numerical simulation is performed using the two-dimensional incompressible Navier-Stokes equations in the vorticity based first-order form. We use a 11th order modal expansion in each element and linearize the resulting discrete model (having a total of $10,212$ degrees of freedom) using Newton's method. At each Newton step, the linear system of equations with a SPD coefficient matrix is solved using the conjugate gradient method with a symmetric Gauss-Seidel preconditioner; see [95]. Convergence of the conjugate gradient method was declared when the norm of the residual was less than 10^{-6}. Nonlinear convergence was declared when the relative norm of the residual in velocities, $\|\Delta \mathbf{u}^{hp}\|/\|\mathbf{u}^{hp}\|$, was less than 10^{-4}, which typically required four Newton iterations. The analysis starts with $\mathrm{Re} = 100$ and steps to $\mathrm{Re} = 800$ using a solution continuation technique with increments of $\mathrm{Re} = 100$. Away from the corner of the step at $(x, y) = (0, 0)$, the L_2 least-squares functional remained below 10^{-5} through the Reynolds number stepping.

Figure 15 shows the streamlines, the vector velocity field, and pressure contours for $0 \leq x \leq 10$, where most of the interesting flow structures occur. The flow separates at the step corner and forms a large recirculation region with a reattachment point on the lower wall of the channel at approximately $x = 6$. A second recirculation region forms on the upper wall of the channel beginning near $x = 5$ with a reattachment point at approximately $x = 10.5$.

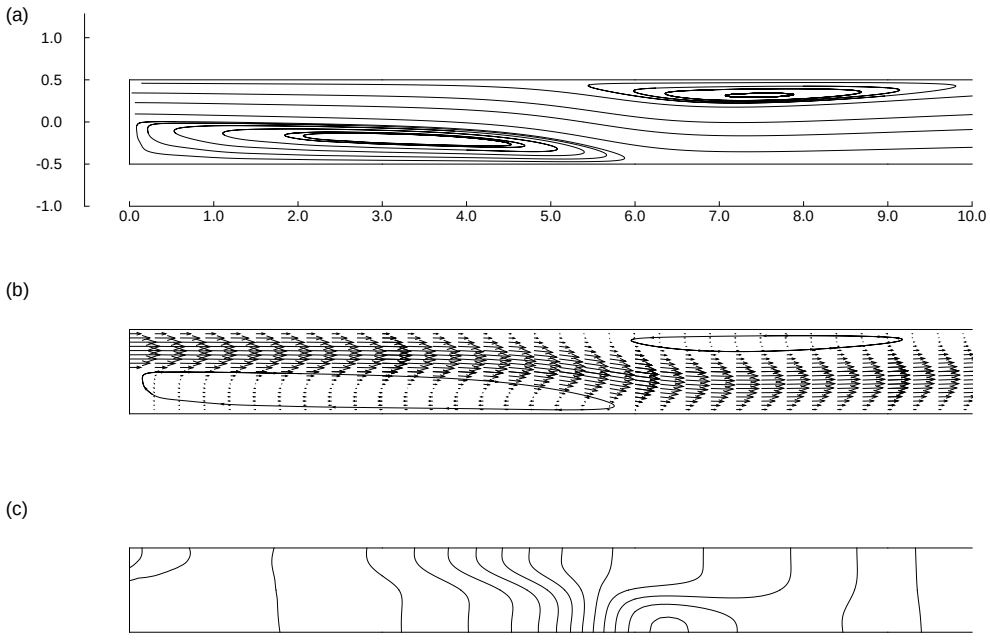

Fig. 15. Flow over a backward facing step at Re = 800: (a) streamlines, (b) vector velocity field, and (c) pressure field. Modal expansion of order 11 using 2 elements along the height and 10 uniformly spaced elements along the entire length of the channel.

Figure 16 shows u-velocity profiles along the channel height at $x = 7$ and $x = 15$. We compare with tabulated values from the benchmark solution of Gartling [33] and find excellent agreement. Gartling's benchmark solution is based on a mixed weak form Galerkin formulation using discretizations ranging from 6×120 to 40×800 bi-quadratic elements, corresponding to discrete systems of $8,426$ to $355,362$ degrees of freedom respectively. Figure 17 shows pressure profiles along the length of the channel walls. The slopes of the pressure profiles become constant near the exit plane, meaning that the flow has recovered to fully developed conditions at the exit.

In Fig. 18 we present the convergence history of the preconditioned conjugate gradient (PCG) method using a Jacobi preconditioner and a symmetric Gauss-Seidel preconditioner. The algebraic set of equations being solved corresponds to the 11th order modal expansion discrete model (having a total of $10,212$ degrees of freedom) and flow conditions of $\mathrm{Re} = 800$. We plot the convergence history corresponding to each of the four Newton iterations, after which the nonlinear tolerance is achieved. The number of PCG iterations reduces during the nonlinear iteration procedure because a continuation technique is used with respect to the nonlinear loop, i.e., the last converged solution is used as an initial guess for the PCG solver. As expected, the Gauss-Seidel preconditioner gives a higher convergence rate. It should be noted, however, that to apply the Gaus-Seidel preconditioner the assembled system of equations needs to be stored, this is not the case for the Jacobi preconditioner. For three-dimensional problems the cost of applying a Gauss-Seidel preconditioner might be high in terms of computer memory. Similar convergence rates and accuracy were observed for the stress and velocity gradient formulations, although at higher CPU solve times due to having additional degrees of freedom.

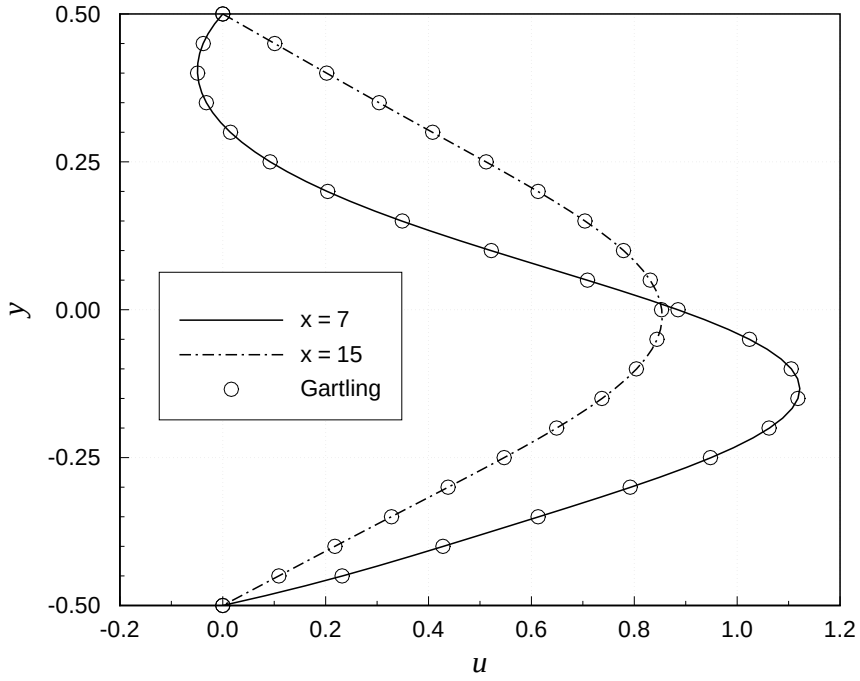

Fig. 16. Horizontal velocity profiles along the height of the channel at $x = 7$ and $x = 15$. Comparison with the benchmark solution of Gartling [33].

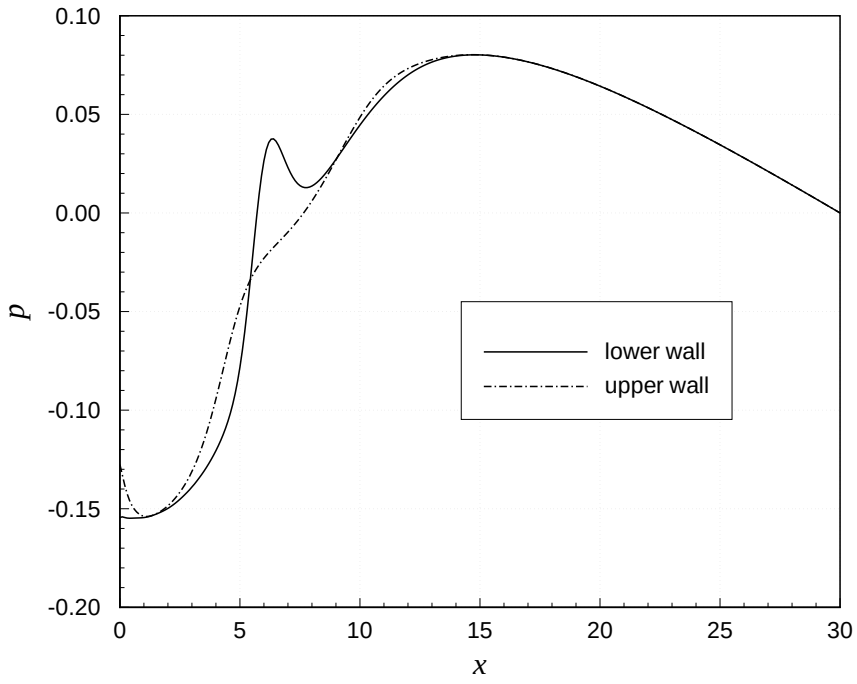

Fig. 17. Pressure profiles along lower and upper walls of the channel.

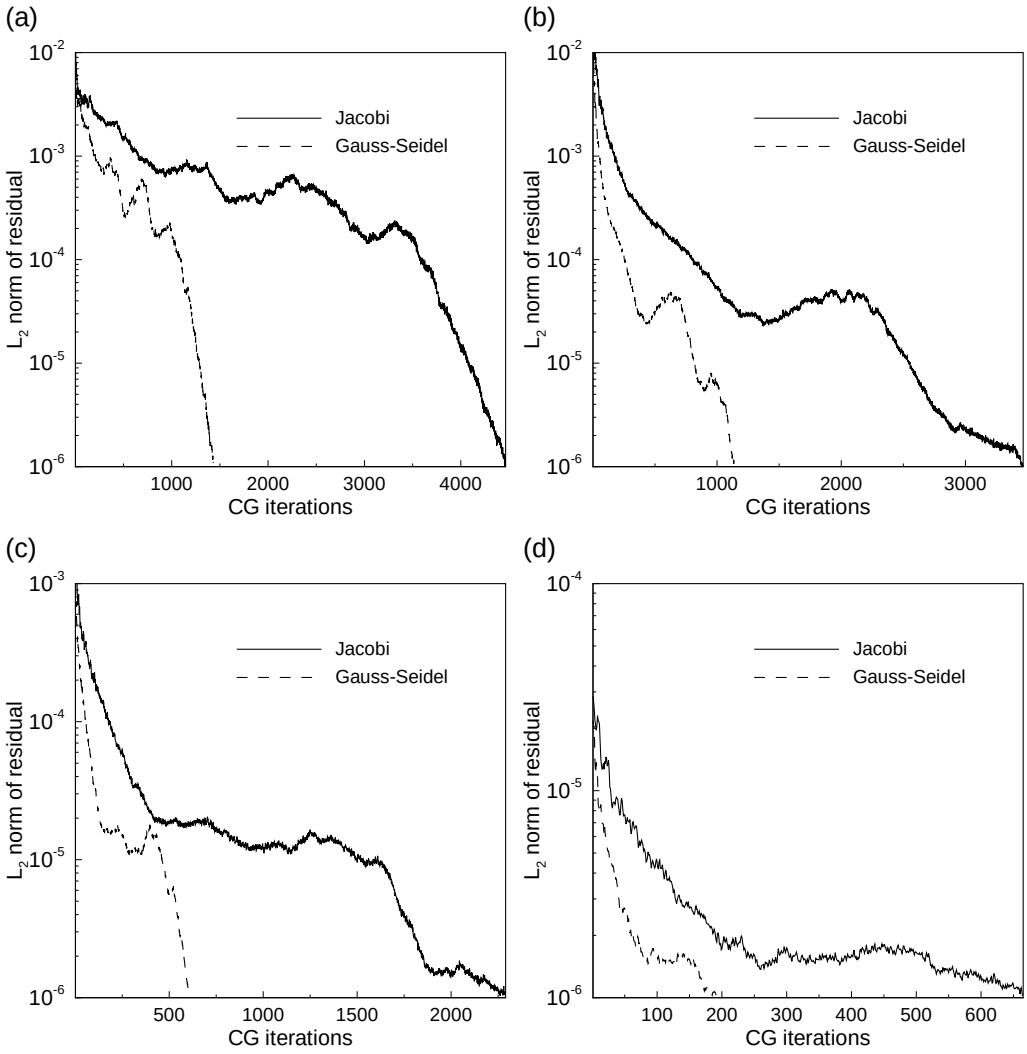

Fig. 18. Convergence history of the conjugate gradient (CG) solver using a Jacobi and a Gauss-Seidel preconditioner. (a)-(d) Correspond to Newton iterations 1-4, for an 11th order modal expansion and flow conditions of Re = 800 in the channel.

4. 3-D lid-driven cavity flow

Next, we consider the three-dimensional flow of an incompressible fluid bounded in a cubic enclosure, $\bar{\Omega} = [0,1] \times [0,1] \times [0,1]$, where the flow is driven by the translation of the top boundary. The connected model, $\bar{\Omega}^h = [0,1] \times [0,1] \times [0,0.5]$, using $6 \times 6 \times 3$ brick finite elements, $\bar{\Omega}_e$, is shown in Fig. 19. It has been established that, up to Re = 3200, the flow is symmetric about the plane $z = 0.5$ [59, 90]. We therefore model only half the domain.

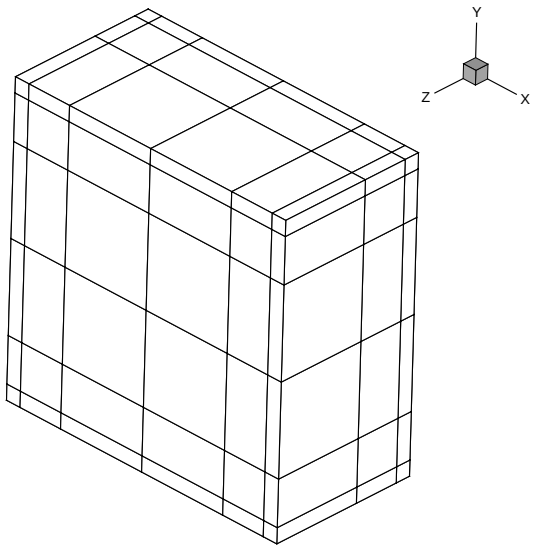

Fig. 19. Computational domain and mesh for the three-dimensional lid-driven cavity problem. Due to symmetry about the plane $z = 0.5$, only half the domain is modeled.

The Reynolds numbers considered here are 100 and 400, for which it has been well established that a steady-state solution exists. We perform the analysis using the three-dimensional incompressible Navier-Stokes equations in the vorticity based first-order form, for which there are a total of seven primary variables: three velocity components, pressure, and three vorticity components. The boundary conditions are as follows: $u = v = w = 0$ on all solid walls, $p = 0$ at a point, $w = \omega_x = \omega_y = 0$

on the symmetry plane ($z = 0.5$), and $u = u^s(x, z)$, $v = w = 0$ on the top driven surface (y=1.0). On the driven surface we specify a hyperbolic tangent u-velocity distribution: $u^s = g(x)\, g(z)$, where

$$g(s) = \begin{cases} \tanh(\beta\, s) & 0 \leq s \leq 0.5 \\ -\tanh(\beta\,(s-1)) & 0.5 < s \leq 1.0 \end{cases} \tag{3.31}$$

with $\beta > 0$. Here we take $\beta = 50$, which gives a smooth but at the same time sharp transition from $u = 0.0$ to $u = 1.0$ near the walls of the driven surface. This boundary condition results in a well-posed boundary condition, in the sense that singularities at the corners of the driven surface are eliminated.

We use a 5th order nodal expansion in each element, resulting in a discrete model with a total of $107,632$ degrees of freedom. The stress or velocity gradient based first-order system was not considered in this study for the solution of three-dimensional flow conditions, as it is significantly more expensive than the vorticity based first-order system. For the discretization considered here, the resulting velocity-gradient based discrete model would have a total of $199,888$ degrees of freedom. Which, based on results from the two-dimensional numerical examples, would yield equally accurate results when compared to the vorticity-based formulation.

At each Newton step the (SPD) linear system of equations is solved using the matrix-free conjugate gradient algorithm with a Jacobi preconditioner. Convergence of the conjugate gradient method was declared when the norm of the residual was less than 10^{-6}. Nonlinear convergence was declared when the relative norm of the residual in velocities was less than 10^{-4}, which typically required six Newton iterations. The analysis starts with Re = 100 and steps to Re = 400 using a solution continuation technique with increments of Re = 100. The L_2 least-squares functional remained below 10^{-3} for the flow conditions considered.

Figure 20 shows velocity vectors, vorticity contours , and pressure contours on the plane $z = 0.5$ for Re $= 400$. On this plane, we see the primary re-circulation region which looks similar to the well documented two-dimensional solution. Figure 21 show the solution on the planes $x = 0.5$ and $y = 0.5$. We see a pair of vortices at the bottom wall of the plane $x = 0.5$ and a weak pair of vortices near the top wall. On the plane $y = 0.5$ we see vortices on the side-wall of the cavity. Figure 22 shows u-velocity profiles along the vertical mid-line of the plane $z = 0.5$, our results are in good agreement with tabulated data from published results of Jiang et al. [59]. Jiang and co-workers used a least-squares finite element formulation with an almost uniform $50 \times 52 \times 25$ mesh with tri-linear elements and *reduced integration*, i.e., one collocation point per element.

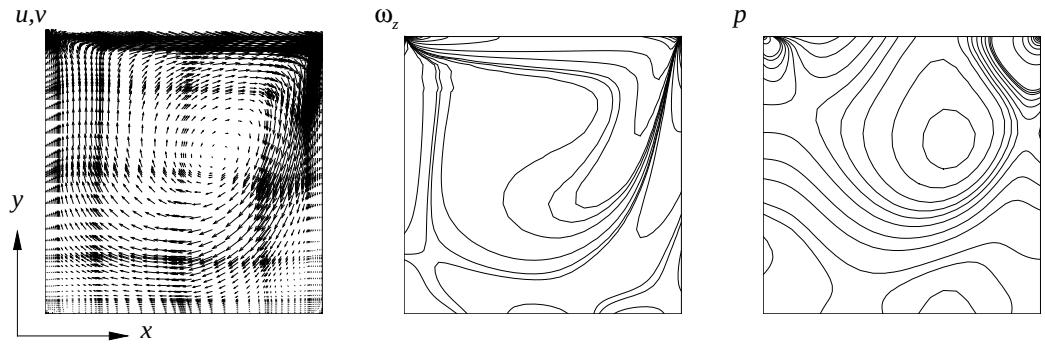

Fig. 20. Velocity vectors, vorticity, and pressure contours on plane $z = 0.5$ for Re $= 400$.

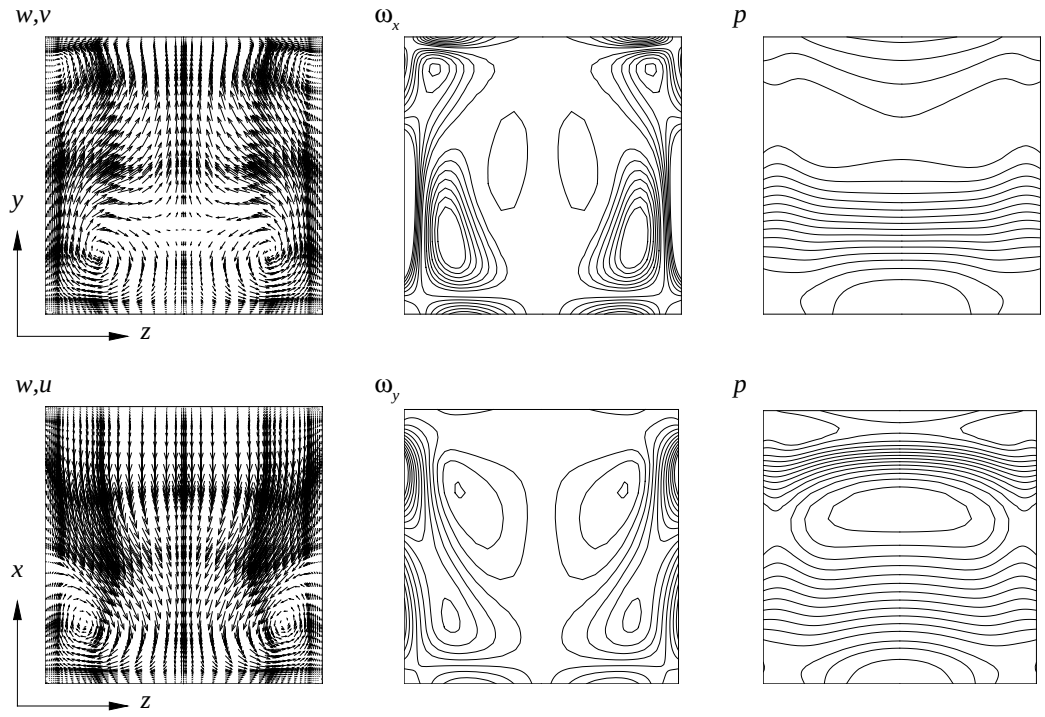

Fig. 21. Velocity vectors, vorticity, and pressure contours on planes $x = 0.5$ and $y = 0.5$ for $\mathrm{Re} = 400$.

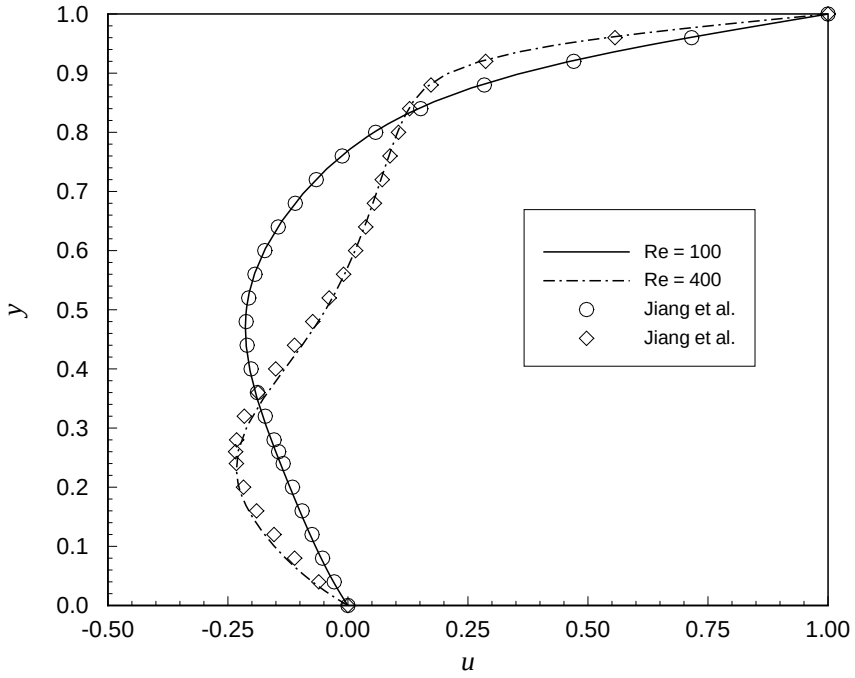

Fig. 22. Profiles of u-velocity along the vertical mid-line of the plane $z = 0.5$. Comparison with tabulated values from the published results of Jiang et al. [59].

5. Impulsively started lid-driven cavity flow

We consider the two-dimensional flow of an incompressible fluid bounded in a square enclosure, $\bar{\Omega} = [0,1] \times [0,1]$. The fluid is initially at rest and is put into motion by the sudden translation of the top boundary. The connected model in space-time, ${}_{s}^{s+1}\bar{\Lambda}^h = \bar{\Omega}^h \times [t_s, t_{s+1}]$, is shown in Fig. 23. It consists of 6×6 finite elements in space and a single element layer in time.

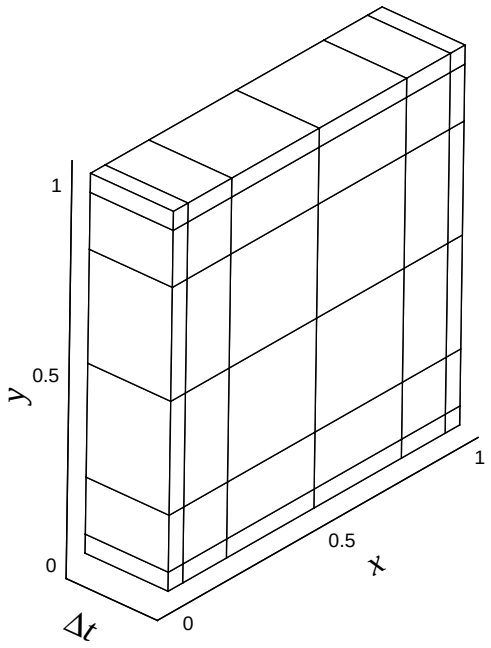

Fig. 23. Space-time computational domain and mesh for the lid-driven cavity problem. The computational domain is a space-time strip, denoted by ${}_{s}^{s+1}\bar{\Lambda}^h$, with $\Delta t = t_{s+1} - t_s$.

The boundary conditions are as follows: $u = v = 0$ on all solid walls, $p = 0$ at a point, and $u = u^s(x)$, $v = 0$ on the top driven surface ($y = 1.0$). On the driven surface we specify a hyperbolic tangent u-velocity distribution $u_{\text{lid}}(x) = g(x)$, with $g(s)$ given

in Eq. (3.31). Like before, we take $\beta = 50$ which gives a smooth but at the same time sharp transition from $u = 0.0$ to $u = 1.0$ near the walls of the driven surface. The u-velocity of the driven surface also varies in time according to a hyperbolic tangent distribution, $u_{\text{lid}}(t) = \tanh(t)$. This ensures a smooth and fast start-up translation of the driven surface.

Note that the boundary conditions are imposed on space-time surfaces, so that the boundary condition $p = 0$ at a point translates into $p = 0$ in a space-time line. Similarly, $u = v = 0$ on all solid walls implies $u = v = 0$ on all space-time surfaces associated with a solid wall. On the space-time surface corresponding to the "driven surface", spatial and temporal variations need to be accounted for when specifying the boundary condition.

The Reynolds number considered here is 400. We use the unsteady, two-dimensional incompressible Navier-Stokes equations in the vorticity based first-order form and a space-time coupled formulation with nodal expansions of $p_\xi = p_\eta = 5$ and $p_\gamma = 2$ in each element, i.e., resolution of order 5 in space and of order 2 in time. At each Newton step the linear system of algebraic equations is solved using the matrix-free conjugate gradient algorithm with a Jacobi preconditioner.

For the time marching procedure the size of the time step, $\Delta t = t_{s+1} - t_s$, was chosen as $\Delta t = 0.2$ for $t \in (0, 2]$, $\Delta t = 0.5$ for $t \in (2, 10]$, and $\Delta t = 1.0$ for $t > 10$. We march in time until a steady-state is reached.

Convergence of the conjugate gradient method was declared when the norm of the residual was less than 10^{-6}. Nonlinear convergence was declared when the relative norm of the residual in velocities between two consecutive iterations was less than 10^{-4}, which typically required three Newton iterations per space-time strip. Steady-state was declared when the relative norm of the residual in velocities be-

tween two consecutive space-time strips was less than 10^{-5}, which was reached at $t = 35$. The L_2 space-time least-squares functional remained below 10^{-3} throughout the time marching procedure.

Figure 24 shows the time history of the flow by streamline plots. Upon start-up a long narrow vortex forms close to lid. The vortex gradually moves to the right and begins to grow. Around $t = 6$ the recirculation region in the lower right corner of the cavity appears. The primary vortex continues to broaden and moves towards its steady-state position.

Figure 25 shows the time history of the u-velocity component at two locations along the vertical mid-line of the cavity, one 0.2 units away from the lid and the other 0.2 units away from the bottom surface. It is interesting to see that upon start-up the fluid close to the lid starts moving in the opposite direction. Figure 26 shows the steady-state velocity profile along the vertical mid-line of the cavity. The steady-state solution is in excellent agreement with the benchmark solution of Ghia et al. [34].

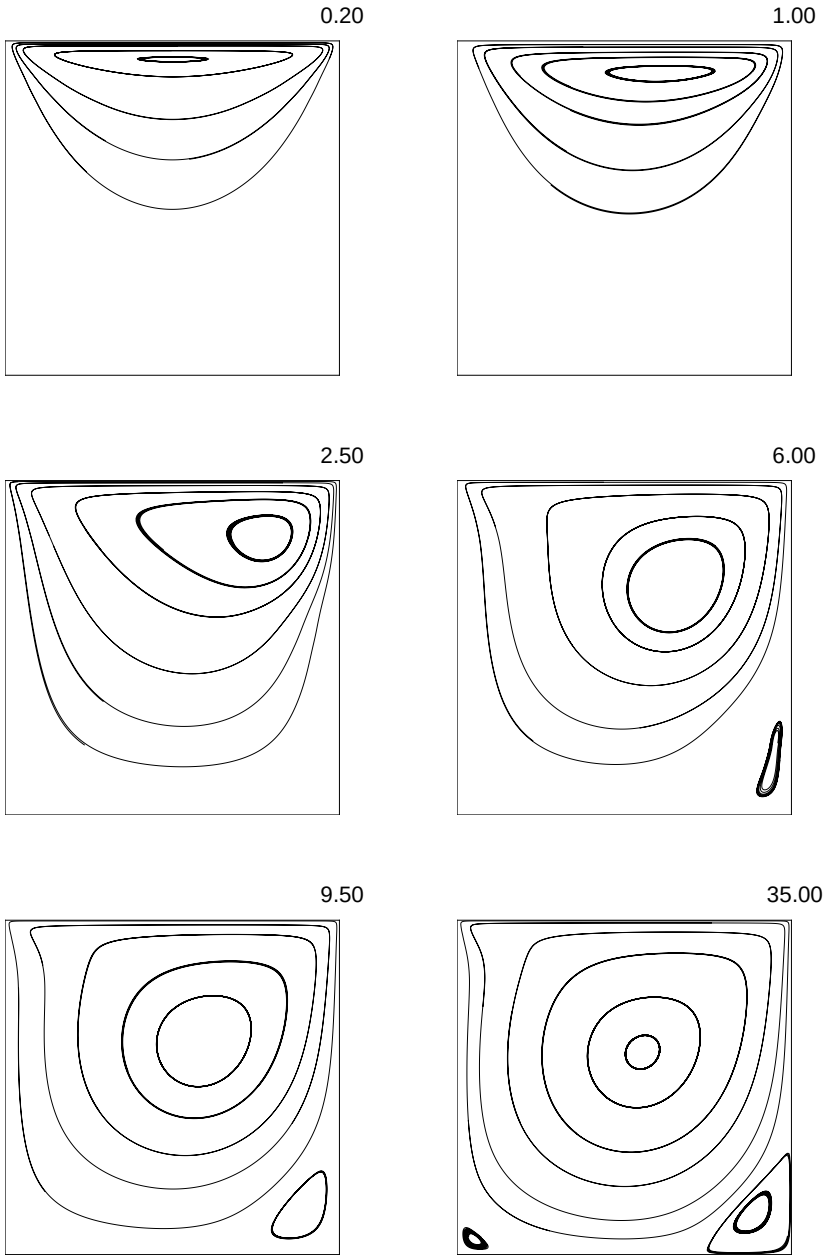

Fig. 24. Impulsively started lid-driven cavity flow: Time history streamline plots for $\mathrm{Re} = 400$.

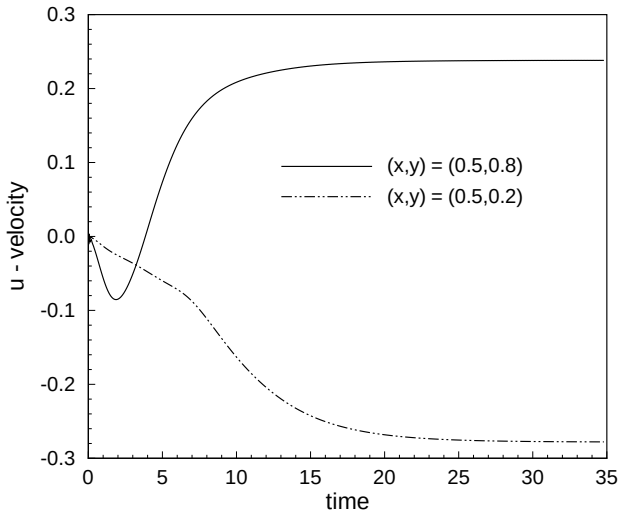

Fig. 25. Time history of the u-velocity component at two selected locations along the vertical mid-line of the cavity.

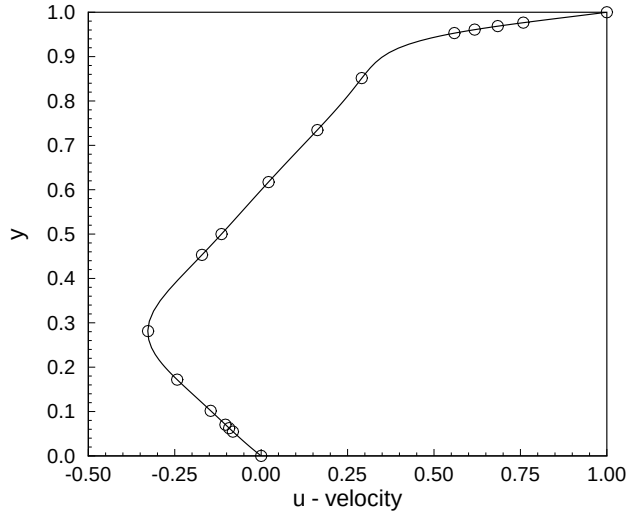

Fig. 26. Steady-state u-velocity profile along the vertical mid-line of the cavity. Re = 400: — computed, ∘ Ghia et al. [34].

6. Oscillatory lid-driven cavity flow

Again, we consider the two-dimensional flow of an incompressible fluid bounded in a square enclosure, $\bar{\Omega} = [0,1] \times [0,1]$. The fluid is initially in a steady-state motion brought about by the translation of the top boundary. The top boundary suddenly begins to oscillate in a periodic fashion. The connected model in space-time is the same as that used for the impulsively started lid-driven cavity flow, and is shown in Fig. 23. It consists of 6×6 finite elements in space and a single element layer in time.

The boundary conditions are the same as that used for the impulsively started lid-driven cavity flow, with the exception that now the u-velocity of the driven surface varies in time according to the cosine distribution: $u_{\text{lid}}(t) = \cos(t)$, with period $\mathcal{T} = 2\pi$.

The Reynolds number considered here is 400. We use the unsteady, two-dimensional incompressible Navier-Stokes equations in the vorticity based first-order form and a space-time coupled formulation with nodal expansions of $p_\xi = p_\eta = 5$ and $p_\gamma = 2$ in each element. At each Newton step the linear system of algebraic equations is solved using the matrix-free conjugate gradient algorithm with a Jacobi preconditioner. For the time marching procedure the size of the time step, $\Delta t = t_{s+1} - t_s$, was chosen as $\Delta t = 2\pi/20$; so that twenty time steps make one period. We march in time until a periodic steady-state is well established. Convergence of the conjugate gradient method was declared when the norm of the residual was less than 10^{-6}. Nonlinear convergence was declared when the relative norm of the residual in velocities between two consecutive iterations was less than 10^{-4}, which typically required three Newton iterations per space-time strip. The L_2 space-time least-squares functional remained below 10^{-3} throughout the time marching procedure.

Figure 27 shows streamline plots at time $t = \mathcal{T}$, $2\mathcal{T}$, $4\mathcal{T}$, and $8\mathcal{T}$. It takes about six periods to reach the periodic steady-state at which the solution is identical at time t and $t + \mathcal{T}$.

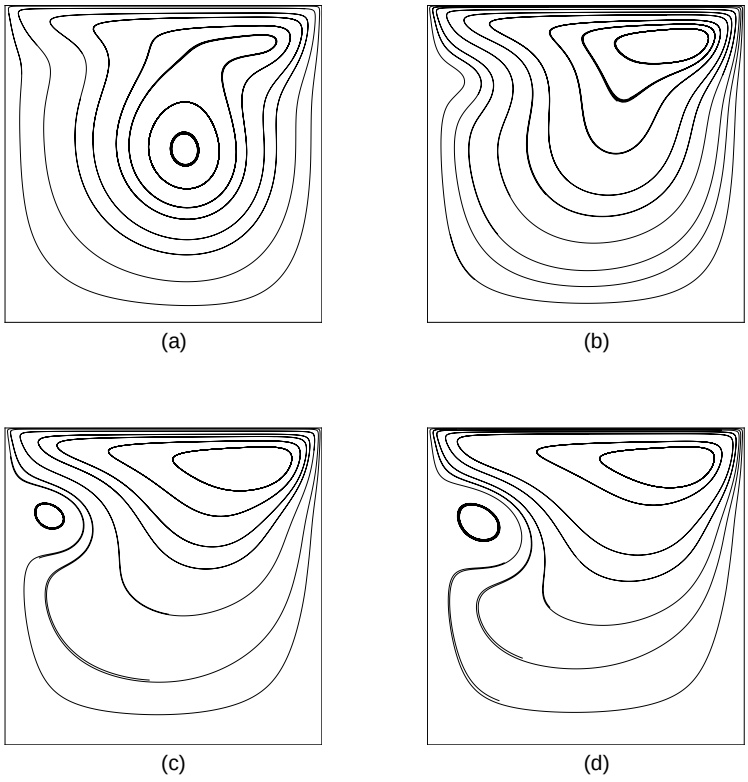

Fig. 27. Oscillatory lid-driven cavity flow: Time history streamline plots for Re = 400. Streamlines at (a) $t = \mathcal{T}$, (b) $t = 2\mathcal{T}$, (c) $t = 4\mathcal{T}$, and (d) $t = 8\mathcal{T}$.

Figure 28 shows the time history up to $t = 16\mathcal{T}$ of the u-velocity component at two locations along the vertical mid-line of the cavity, one 0.2 units away from the lid and the other 0.2 units away from the bottom surface. Streamline contour plots for the periodic steady state solution are presented in Fig. 29 at the 8th cycle for $8\mathcal{T} \leq t \leq 9\mathcal{T}$. Figures 29a-h correspond to the time sequence $t = 8\mathcal{T} + \beta\mathcal{T}$, where $\beta = 2/10$, $3/10$, $4/10$, $5/10$, $7/10$, $8/10$, $9/10$, $10/10$. The property of mirror images

is observed between Figs. 29a-d and Figs. 29e-h with respect to $x = 0.5$. Our results compare well with those by Iwatsu et al. [49] using the MAC method and Soh et al. [99] using an artificial compressibility method.

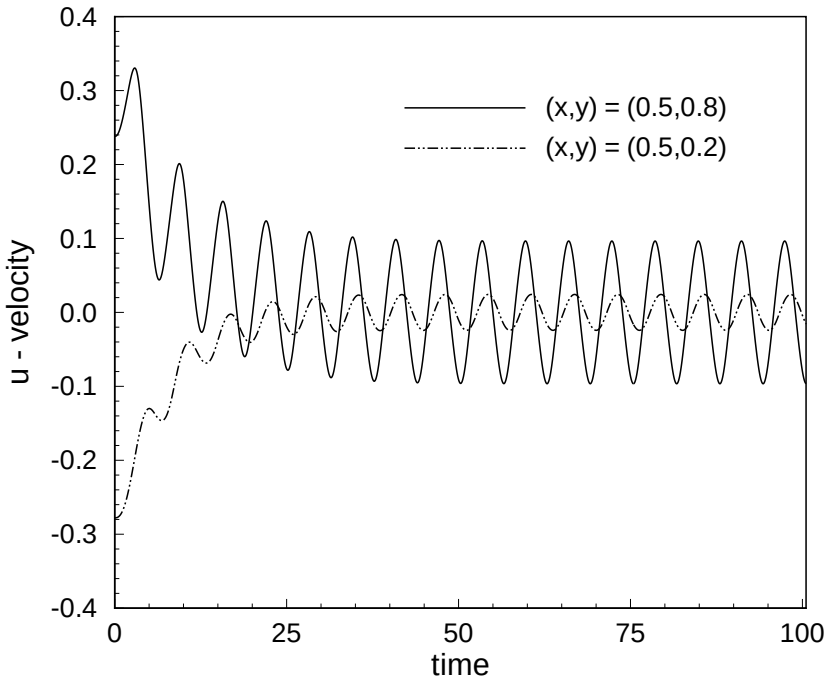

Fig. 28. Time history of the u-velocity component at two selected locations along the vertical mid-line of the cavity.

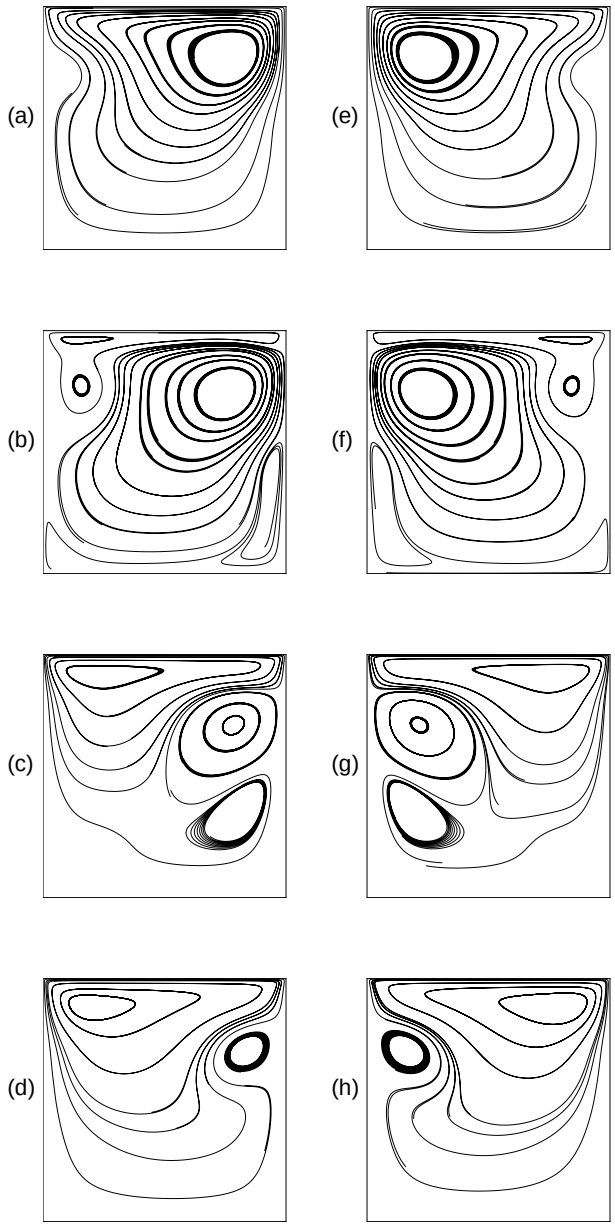

Fig. 29. Oscillatory lid-driven cavity flow: Periodic steady-state time history streamline plots for Re = 400. Streamlines at time $t = 8\mathcal{T} + \beta\mathcal{T}$; (a) $\beta = 2/10$, (b) $\beta = 3/10$, (c) $\beta = 4/10$, (d) $\beta = 5/10$, (e) $\beta = 7/10$, (f) $\beta = 8/10$, (g) $\beta = 9/10$, (h) $\beta = 10/10$.

7. Transient flow over a backward-facing step

We consider again the two-dimensional flow over a backward-facing step at Re = 800. In Section 3, the stationary incompressible Navier-Stokes equations were used directly to develop the finite element model (i.e., we assumed a priori that a steady-state solution existed) and our numerical results were compared and found to be in excellent agreement with the benchmark numerical solution of Gartling [33], who also used the stationary form of the incompressible Navier-Stokes equations directly.

Here we are concerned with the fundamental question of whether a steady-state solution exists for flow over a backward facing step at Re = 800. In the work of Kaiktsis et al. [61] it was concluded, by means of a (weak form Galerkin space-time decoupled) high-order splitting spectral/*hp* numerical simulation, that transition to turbulent flow had occurred by Re = 800 and that the flow was steady at Re = 500, time periodic at Re = 700, and chaotic at Re = 800. The time periodic and chaotic temporal behavior at Re = 700 and Re = 800, respectively, contradicted results reported in numerous previous published articles (summarized in Ref. [39]) which obtained steady flow results at those flow conditions. In response to the controversy concerning the stability at flow conditions of Re = 800, Gresho et al. [39] performed a detailed grid refinement study using four different numerical methods and conclusively showed that the flow over a backward-facing step at Re = 800 attains a stable (i.e., non-periodic) steady state. The numerical simulations of Gresho et al. [39] were performed on the simplified step geometry shown in Fig. 14 and among the four numerical methods considered, a (weak form Galerkin space-time decoupled) spectral element formulation similar to that of Kaiktsis et al. [61] was used.

The study of Gresho et al. [39] revealed that inadequate spatial resolution induces chaotic-like temporal behavior, whereas when sufficiently high spatial resolution

is used the flow evolves towards a steady state by a monotonic decay of the transient. This behavior was also observed in the work of Torczynski [104], who performed simulations using the standard step geometry (where the channel portion upstream of the step is included) and also used a (weak form Galerkin space-time decoupled) spectral element formulation. These observations were further confirmed in a subsequent study by Kaiktsis et al. [62]. In addition, a study by Fortin et al. [31] using dynamical systems theory, showed that the first Hopf bifurcation point (transition point) for two-dimensional flow over a backward-facing step occurs at least up to Re = 1600. Adding to the study of Torczynski [104], Yee et al. [111] showed that the initial condition (in addition to the spatial resolution) also strongly influences the temporal behavior of the flow.

In the present study, we wish to investigate the temporal behavior of the two-dimensional flow over a backward-facing step at Re = 800 using a least-squares based finite element formulation. We consider space-time decoupled and space-time coupled formulations. For comparison of the two formulations, we choose a second-order accurate in time discretization for the temporal terms in constructing the space-time decoupled formulation (see Chapter II) and restrict the space-time coupled formulation to a p-level of order 2 in time domain. The simulations are performed using the simplified step geometry shown in Fig. 14, with the length of the channel extending a distance $L/H = 15$ downstream of the step, which is sufficiently long to preclude undue influence of the finite channel length at this Reynolds number [33, 39]. Even though the geometric singularity of the corner step is excluded by ignoring the portion of the channel upstream of the step, the singularities in pressure and vorticity still exist due to the inlet boundary condition (see Fig. 14).

The boundary and initial conditions used here are those used in the work of Gresho et al. [39] for the spectral element numerical simulation: $u = v = 0$ on the

horizontal walls, $-p + \mu\,\partial u/\partial n = 0$ and $\partial v/\partial n = 0$ on the outflow boundary, and $u = [\tanh(t/4)]\,u_{\mathrm{B}}(y) + [1 - \tanh(t/4)]\,u_{\mathrm{P}}(y)$ and $v = 0$ on the inflow boundary and the step face. Here $u_{\mathrm{B}}(y) = \max[0, 24y(0.5 - y)]$ is the true inlet boundary condition and $u_{\mathrm{P}}(y) = 3(0.5 - y)(0.5 + y)$ is the Poiseuille flow observed infinitely far downstream at steady flow conditions. The initial velocity field is set to $u = u_{\mathrm{P}}(y)$ and $v = 0$ everywhere in the computational domain. Note that the inlet condition is varied fast but smoothly from Poiseuille flow to flow over a backward-facing step, thus inducing a transient wave strong enough to excite sustained unsteady behavior, if that is the correct asymptotic steady state behavior.

The transient wave will travel through the entire channel length. The main flow coming from the inlet will follow a sinuous path through the channel, forming a series of eddies along the upper and lower wall (see Fig. 35). As the flow evolves, the strength of the eddies should diminish and the two major separation zones near the inlet of the channel attain their steady-state position. Eventually the weaker eddies along the channel length die out and the flow reaches a steady state. Based on the observations made in the previous work of Gresho et al. [39] and Torczynski [104], high enough spatial resolution should be used to adequately resolve all spatial features of the flow. Otherwise, lack of spatial resolution will induce unrealistic temporal chaotic behavior resulting in an erroneous prediction of the long-term behavior of the flow. The rich physics of the flow and the danger of polluting the long-term behavior of the flow by inadequate spatial resolution render this benchmark problem a challenging one for high-order methods. In addition, as we subsequently show through numerical results, the problem is ideally suited to test and compare the performance of space-time coupled and decoupled formulations.

We discretize the domain, $\bar{\Omega} = [0, 15] \times [-0.5, 0.5]$, using 120 finite elements: four uniformly spaced elements along the height of the channel and 30 uniformly spaced

elements along the length of the channel. This mesh coincides with the high resolution mesh in the work of Gresho et al. [39] (although their computational domain extended a distance $L/H = 17$ downstream of the step). The numerical simulation is performed using the unsteady, two-dimensional incompressible Navier-Stokes equations in the vorticity based first-order form. The outflow boundary conditions are imposed in a weak sense through the least-squares functional.

First we use a spatial resolution of order 7, for which Gresho et al. [39] reported a steady monotonic decay of the transient. Recall that, for this simulation, a temporal resolution of order 2 is used for the space-time coupled formulation and a second-order accurate representation (the trapezoidal rule) is used for the temporal terms in the space-time decoupled formulation. For the time marching procedure the size of the time step, $\Delta t = t_{s+1} - t_s$, was chosen as $\Delta t = 0.20$ for the space-time coupled and decoupled formulations. We march in time until a steady-state is well established.

The resulting discrete models (having a total of $73,428$ dof for the space-time coupled finite element model and $24,476$ dof for the space-time decoupled finite element model) are linearized using Newton's method. At each Newton step, the linear system of equations with a SPD coefficient matrix are solved using the conjugate gradient method with a Jacobi preconditioner in matrix-free form. Convergence of the conjugate gradient method was declared when the norm of the residual was less than 10^{-6}. Nonlinear convergence was declared when the relative norm of the residual in velocities, $\|\Delta \mathbf{u}^{hp}\|/\|\mathbf{u}^{hp}\|$, was less than 10^{-4}, which typically required three Newton iterations per space-time strip/time-step.

Figure 30 shows a plot of the L_2 least-squares functional as a function of time, for the space-time coupled and decoupled formulations. The legend in Fig. 30 describes the resolution of the simulation in space-time and the size of the space-time strip/time-step. For example, $7/7/2$ $\Delta t = 0.20$ denotes a space-time coupled simulation with

isotropic resolution of order 7 in space and order 2 in time, with the the size of the space-time strip fixed at 0.20. On the other hand, 7/7/TR $\Delta t = 0.20$ denotes a space-time decoupled simulation with isotropic resolution of order 7 in space and a trapezoidal rule (TR) representation for the temporal terms, with the the size of the time-step fixed at 0.20.

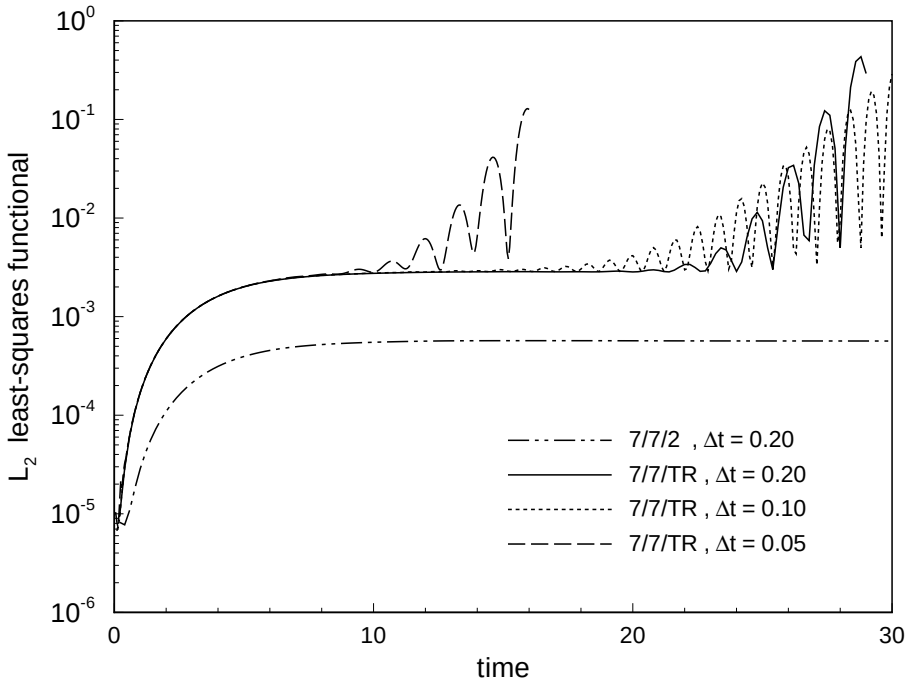

Fig. 30. Time history of the L_2 least-squares functional for space-time coupled and decoupled formulations. Fixed spatial resolution of order 7.

From Fig. 30 it is clear that the space-time decoupled formulation (for $\Delta t = 0.20$) becomes unstable, forcing an early termination of the simulation. On the other hand, the space-time coupled formulation remains stable throughout; at a lower value of the L_2 least-squares functional than that initially attained by the decoupled formulation. The decoupled formulation is stable at early times, during the smooth transition from Poiseuille flow to flow over a backward-facing step. However, by the time the transient

wave reaches the outflow boundary a severe instability had occurred and grown to such extent which forced the termination of the simulation.

Next, we investigate the effect of the time step size on the stability of the decoupled formulation. The results, also plotted in Fig. 30, show that the space-time decoupled formulation remains unstable, even for $\Delta t = 0.05$. In fact, when Δt is decreased the instability starts at earlier times in the simulation – perhaps suggesting a lack of numerical damping (inherent to the TR representation). With this in mind and in an attempt to stabilize the space-time decoupled simulation we implement the generalized α-method family of approximations [25, 51, 27], which retain second-order accuracy in time and allow for user controlled high frequency damping by the single free integration parameter, $0 \leq \rho_\infty^h \leq 1$. For $\rho_\infty^h = 1.0$ the method is identical to the trapezoidal rule and for choices of $0 \leq \rho_\infty^h < 1.0$ numerical damping is added with decreasing ρ_∞^h. Figure 31 shows the time history of the L_2 least-squares functional for the space-time decoupled formulation using the generalized α-method (GAM) for $\Delta t = 0.10$ and popular choices of ρ_∞^h (see Ref. [27]). Increasing the numerical dissipation clearly does not stabilize the simulation.

Realizing the underlying assumption in the decoupled formulation, namely that space and time are decoupled, we are led to believe that the spatial resolution is still not high enough and is inducing chaotic temporal instabilities that cause the space-time decoupled formulation to become unstable. Clearly this observation does not apply to the space-time coupled formulation, where no instability is observed.

Based on the above observation, we increase the spatial resolution of the space-time decoupled simulation to order 9 (having a total of $40,108$ dof for the space-time decoupled finite element model). For this spatial resolution the space-time decoupled formulation, with TR and $\Delta = 0.20$, is stable. Figure 32 shows the time history of the L_2 least-squares functional for the space-time coupled simulation $7/7/2$ $\Delta t = 0.20$ and

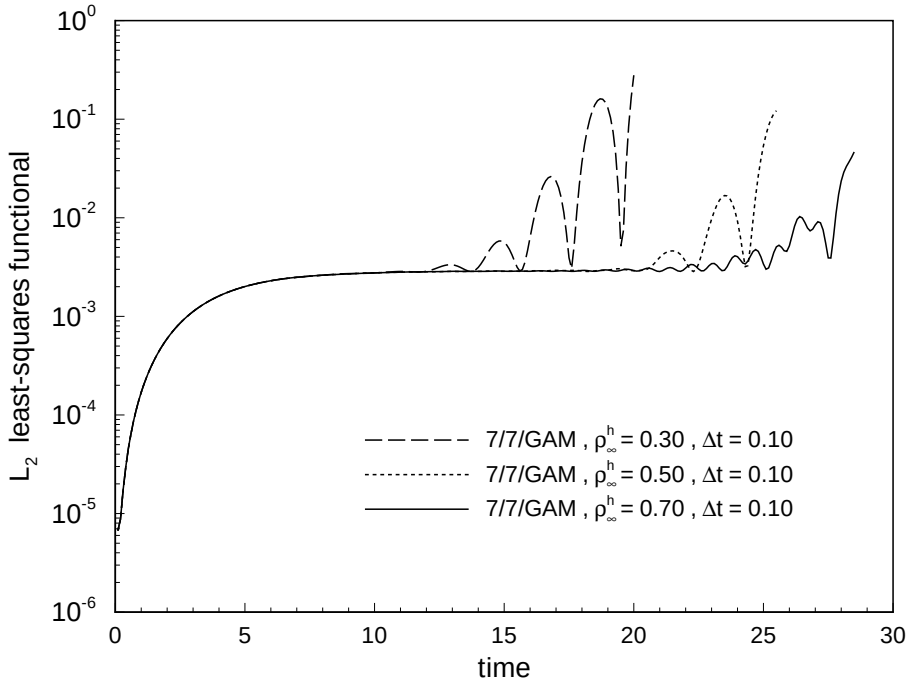

Fig. 31. Time history of the L_2 least-squares functional for space-time decoupled formulation using the generalized-α method. Fixed spatial resolution of order 7.

the space-time decoupled formulation 9/9/TR $\Delta t = 0.20$. Even though the space-time decoupled simulation has a higher spatial resolution, the space-time coupled formulation (with a lower spatial resolution) achieves a lower value for the L_2 least-squares functional.

At $t = 400.0$ the relative norm of the residual in velocities between two consecutive space-time strips was less than 10^{-5} for the space-time coupled formulation and less than 10^{-4} between two consecutive time-steps for the space-time decoupled formulation , indicating that a steady-state was achieved. Figure 33 shows the time history of the v-velocity component at two locations along the channel's mid-section for the space-time coupled (7/7/2 $\Delta t = 0.20$) and decoupled formulations (9/9/TR

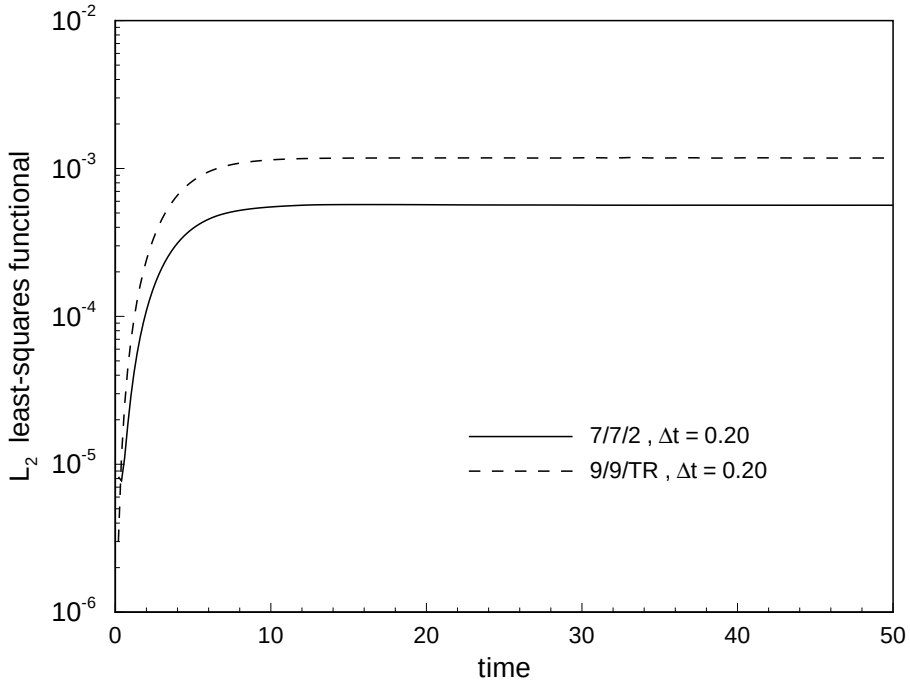

Fig. 32. Time history of the L_2 least-squares functional for space-time coupled simulation 7/7/2 $\Delta t = 0.20$ and decoupled simulation 9/9/TR $\Delta t = 0.20$.

$\Delta t = 0.20$). It is seen that both simulations do not exhibit a monotonic decay of the transient but rather an oscillatory decay of the transient, indicating that additional spatial resolution is needed. In addition, the space-time decoupled simulation exhibits a more oscillatory temporal behavior than the coupled simulation. Nevertheless, the desired steady-state flow condition is achieved using both approaches.

Of importance is also the computational cost associated with each of the simulations. Figure 34 shows the time histories of the number of PCG iterations at each space-time strip/time-step for the space-time coupled (7/7/2 $\Delta t = 0.20$) and decoupled formulations (9/9/TR $\Delta t = 0.20$). Each data point in the plot is the sum of PCG iterations at each Newton step per space-time strip/time-step (typically 3 Newton

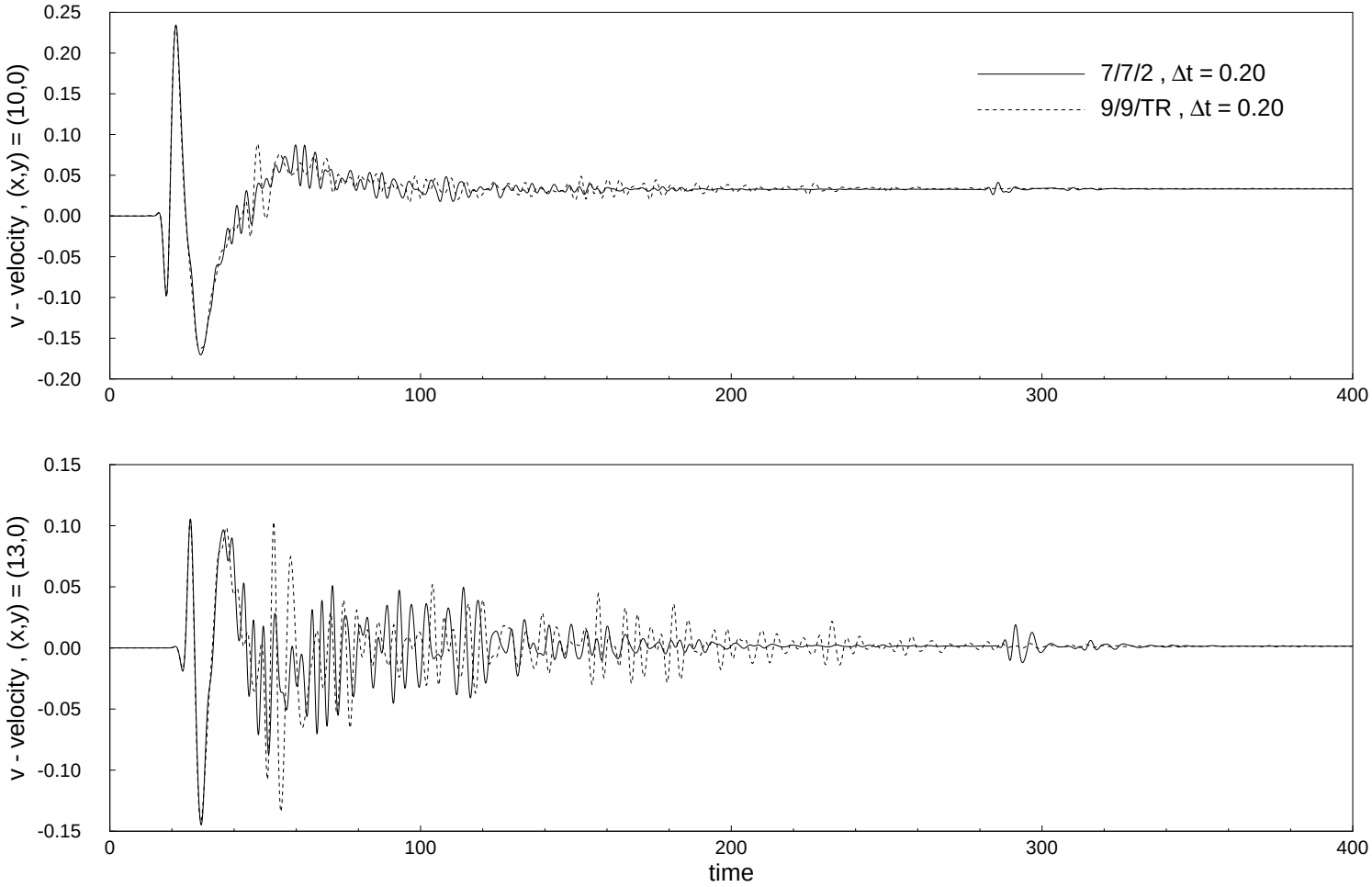

Fig. 33. Time history of the v-velocity component along the mid-section of the channel. Space-time coupled simulation 7/7/2 $\Delta t = 0.20$ and decoupled simulation 9/9/TR $\Delta t = 0.20$.

steps), thus representing the total cost per space-time strip/time-step. From Fig. 34 we see that initially both simulations have approximately the same cost, with the space-time coupled simulation eventually achieving an overall lower computational cost.

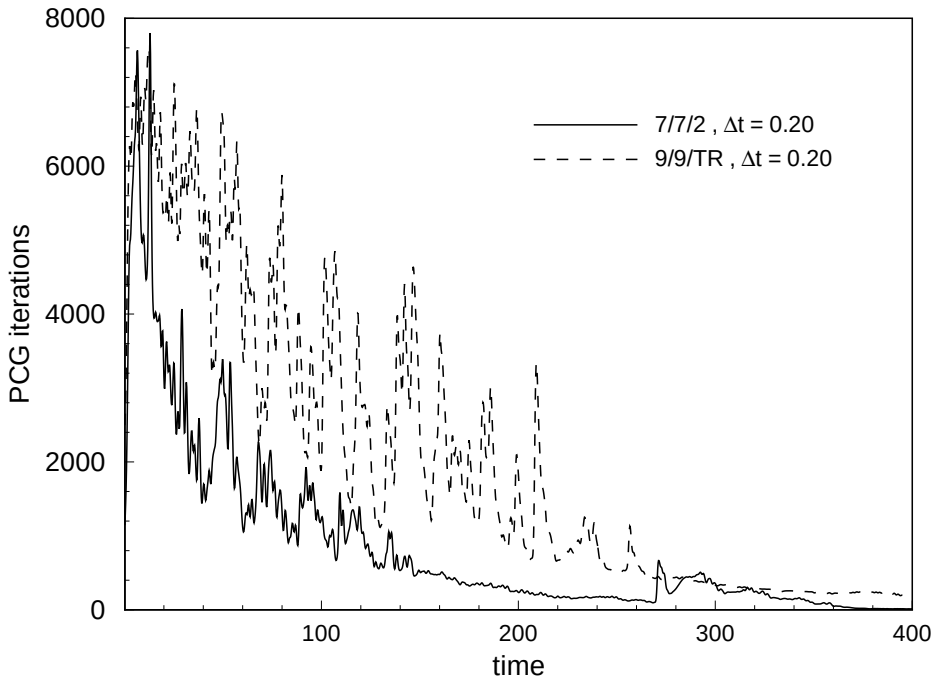

Fig. 34. Time history of total PCG iterations per space-time strip/time-step. Space-time coupled simulation 7/7/2 $\Delta t = 0.20$ and decoupled simulation 9/9/TR $\Delta t = 0.20$. Jacobi preconditioner with stopping criteria $\varepsilon_{PCG} = 10^{-6}$.

In an attempt to obtain an unstable simulation with the space-time coupled formulation, we decrease the spatial resolution to order 5, and run the simulation 5/5/2 $\Delta t = 0.20$. For such spatial resolution Gresho et al. [39] reported 'numerical errors', denoting termination of the numerical solution due un-physically large values in veloc-

ities. In contrast to the results reported using the high-order (space-time decoupled) weak-form Galerkin formulation in Gresho et al. [39], the least-squares space-time coupled simulation remained stable and predicted a steady-state. All the above results, indicate that the least-squares space-time coupled formulation (for unsteady two-dimensional simulations) is robust, computationally affordable, and superior to the space-time decoupled formulation.

This problem is clearly more demanding, in terms of space-time resolution, than the previously considered unsteady cavity flows. For the lid-driven cavity flows considered earlier, the space-time coupled and decoupled formulations give indistinguishable time histories (using time-domain p-resolution of order 2 for the space-time coupled simulation and a TR representation for the decoupled simulation). However, in general, we can never classify a problem a priori as needing low or high space-time resolution. In this sense, it is best to be conservative and use a space-time coupled formulation. For three-dimensional unsteady problems, a space-time coupled approach is currently prohibitively expensive and a decoupled approach must be used.

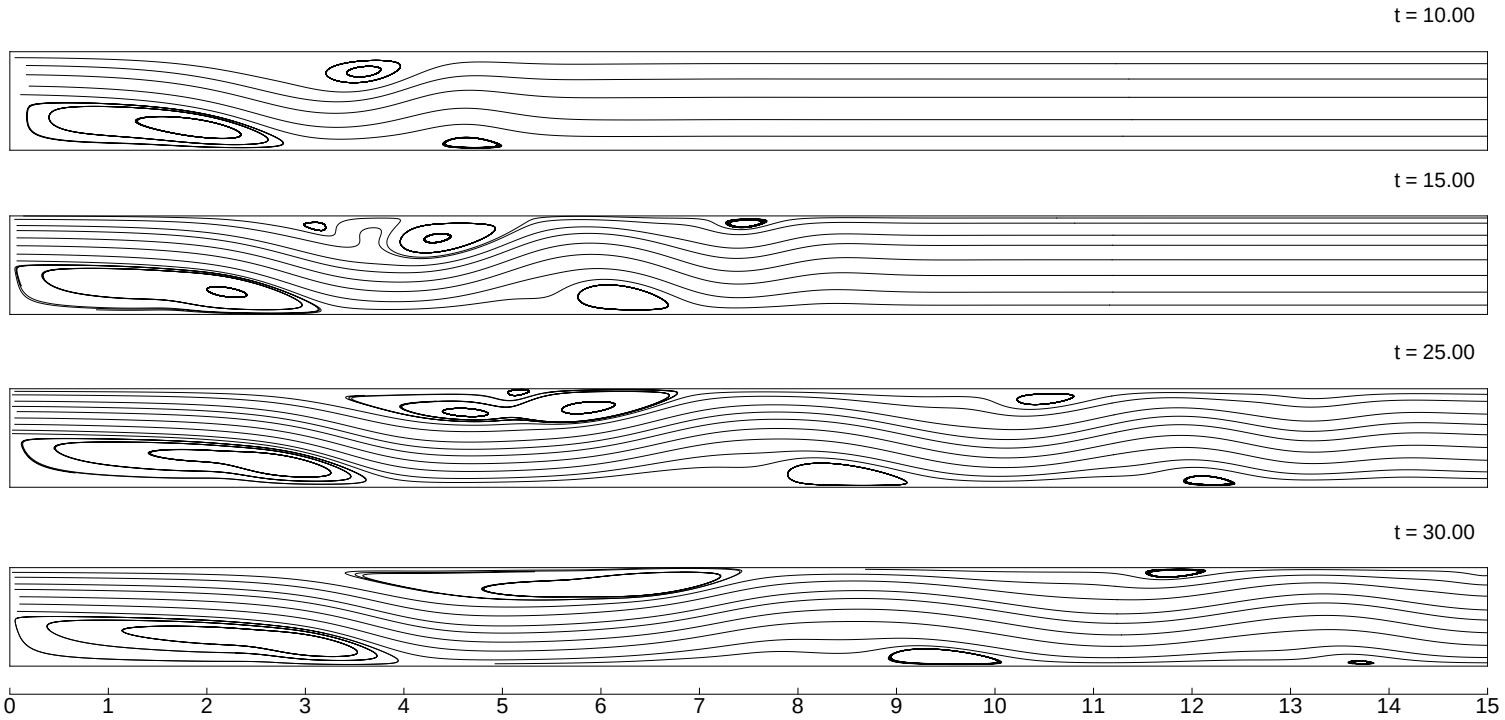

Fig. 35. Time history streamline plots of time-dependent simulation using Poiseuille flow as an initial condition. Space-time coupled simulation 7/7/2 $\Delta t = 0.20$.

C. Numerical examples: validation benchmarks

In this section we are concerned with the *validation* of the spectral/*hp* least-squares based finite element computational algorithm. Validation commonly involves comparison of the numerical results with reliable experimental data. In such a comparison, there is invariably error tolerance involved, and the levels of acceptable error are problem and purpose dependent.

The ongoing process of algorithm validation procedures must demonstrate, to within application dependent tolerances, that the computational algorithm reliably simulates nature through the mathematical model. Validation metrics are of great importance in validation procedures and are used to quantitatively compare numerical results with experimental measurements. Ideally, the validation metric should quantify errors and uncertainties in the comparison study. Examples of typical validation metrics are lift and drag coefficients for external flows, and surface pressure coefficient distributions for both external and internal flows.

In the following, we demonstrate the validation procedure of the (already verified) least-squares spectral/*hp* finite element algorithm by means of the benchmark problem of flow past a circular cylinder.

1. Flow past a circular cylinder

We consider the two-dimensional flow of an incompressible fluid past a circular cylinder. At low Reynolds numbers ($5 < \mathrm{Re} < 40.0$) the flow is stationary and characterized by a pair of standing vortices appearing behind the cylinder. The size of the separated flow region increases with increasing Reynolds number, until a limit in which the wake becomes unstable. At this critical Reynolds number, $\mathrm{Re}_c \approx 46.2$ [50, 29], the flow becomes unstable and vortices are shed from the aft of the circular cylin-

der, forming the well-known von Karman vortex street. The flow may be treated as two-dimensional until $\mathrm{Re}_c^{\text{3-D}} \approx 188.5$ [44], at which point the cylinder wake becomes three-dimensional due to a secondary instability of the vortex street.

From the numerical simulation point of view, the interest of this problem is in the accurate resolution of the vortex street and the modelling of the outflow boundary conditions, the latter which arises from the necessity of truncating the domain in a region where the vortex street is fully developed. This transient problem is frequently used to assess the accuracy of time-marching numerical procedures and open boundary conditions, as it is characterized by several important flow parameters, such as the non-dimensional frequency of the vortex shedding (the Strouhal number), the amplitude of the lift coefficient, and the mean drag coefficient. First, we perform a simulation for the cases $\mathrm{Re} = 20$ and 40, for which a steady-state solution exists and then consider the case $\mathrm{Re} = 100$ for which a transient simulation is necessary.

a. Simulation at $\mathrm{Re} = 20$ and 40

The issue of proper placement of the computational boundaries is of great importance, as it may significantly pollute the accuracy of the characteristic flow parameters. For the set of stationary simulations, the cylinder is of unit diameter and is placed in the finite region $\bar{\Omega} = [-15.5, 30.5] \times [-20.5, 20.5]$. The center of the cylinder lies at $(x, y) = (0, 0)$, so that the inflow boundary is located 15.5 cylinder diameters in front of the center of the cylinder and the outflow boundary 30.5 cylinder diameters downstream of the center of the cylinder. The top and bottom boundaries are located each 20.5 cylinder diameters above and below the center of the cylinder. The Reynolds number is based on the free-stream velocity and cylinder diameter.

Having considered a large computational domain allows us to impose free-stream boundary conditions at the top and bottom of the domain without noticeably affecting

the solution. The boundary conditions include a specified value of 1.0 for the x-component of velocity at the inflow, top, and bottom boundaries, i.e., the free-stream velocity u_∞ is specified to be unity. At these boundaries the y-component of velocity is set to zero. The outflow boundary conditions are imposed in a weak sense through the least-squares functional.

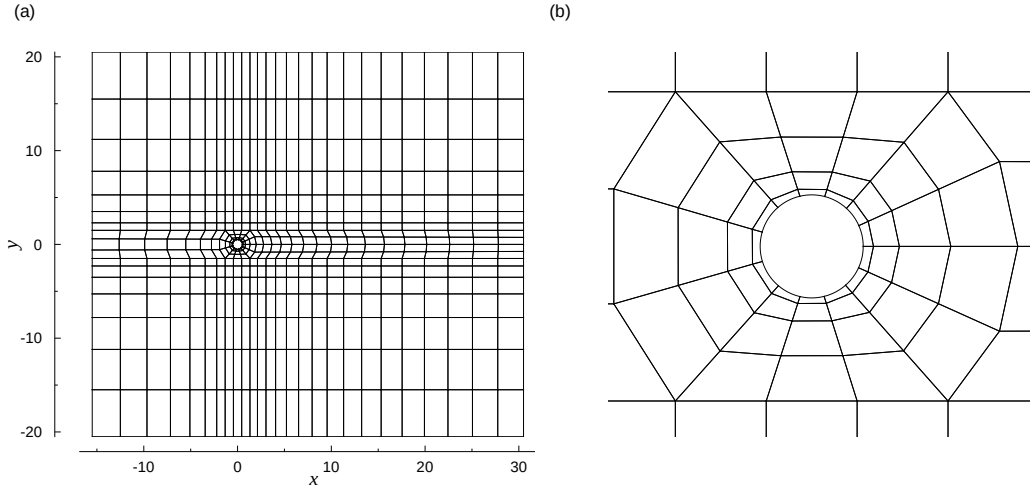

Fig. 36. Computational domain and mesh for flow past a circular cylinder. (a) Connected model, $\bar{\Omega}^h$. (b) Close-up view of the geometric discretization around the circular cylinder.

The connected model, $\bar{\Omega}^h$, consists of 501 finite elements and is shown in Fig. 36, where a close-up view of the geometric discretization around the circular cylinder is also shown. In the previous examples a subparametric formulation using a linear basis for the mapping, $\bar{\Omega}_e \rightleftarrows \hat{\Omega}_e$, was sufficient to exactly represent the straight-sided geometries. In order to accurately represent the circular surface, we implement an isoparametric formulation; i.e., we use the same expansion order for the element degrees of freedom and for the mapping, $\bar{\Omega}_e \rightleftarrows \hat{\Omega}_e$.

We use the two-dimensional incompressible Navier-Stokes equations in the vorticity based first-order form and a 6th order nodal expansions in each element, resulting

in a discrete model with a total of $73,344$ degrees of freedom. At each Newton step the (SPD) linear system of equations is solved using the matrix-free conjugate gradient algorithm with a Jacobi preconditioner. Convergence of the conjugate gradient method was declared when the norm of the residual was less than 10^{-6}. Nonlinear convergence was declared when the relative norm of the residual in velocities between two consecutive iterations was less than 10^{-4}, which required six Newton iterations. The L_2 least-squares functional remained below 10^{-6} for the flow conditions considered.

Figure 37 shows the computed surface pressure coefficient distributions along the cylinder surface for Re = 20 and 40, together with experimental measurements of Grove et al. [40] for the case Re = 40. We observe that the simulation result is in good agreement with the experimental measurements. The computed drag coefficients for flow conditions of Re = 20 and 40 were $C_D = 2.0862$ and 1.5537, respectively. Good agreement is found between the computed drag coefficients and the experimental mean curve of Tritton [105], where the corresponding values are $C_D = 2.05$ and 1.56, and with the high-order splitting spectral/hp numerical results of Henderson [44], reported as $\bar{C}_D = 2.06$ and 1.54.

Figure 38 shows computed pressure contours and streamlines in the wake region for Re = 20 and 40. The predicted wake extends 1.86 and 4.55 cylinder radii measured from the back of the cylinder, respectively. The values for the wake lengths are in good agreement with the numerical solution of Dennis and Chang [26], whose computed wake lengths for Re = 20 and 40 were reported as 1.88 and 4.69 cylinder radii, respectively. Better agreement for the case Re = 40 is found with the numerical solution of Kawaguti and Jain [65], who reported a computed wake length of 4.50 cylinder radii.

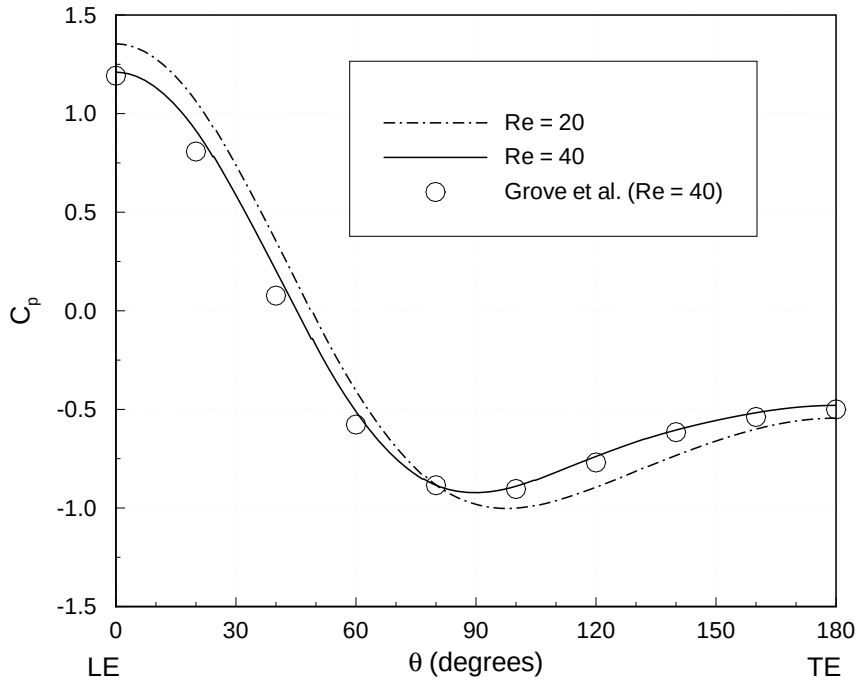

Fig. 37. Flow past a circular cylinder at Re = 20 and 40: Computed pressure coefficient distributions along the cylinder surface. Comparison with experimental measurements of Grove et al. [40] for the case Re = 40.

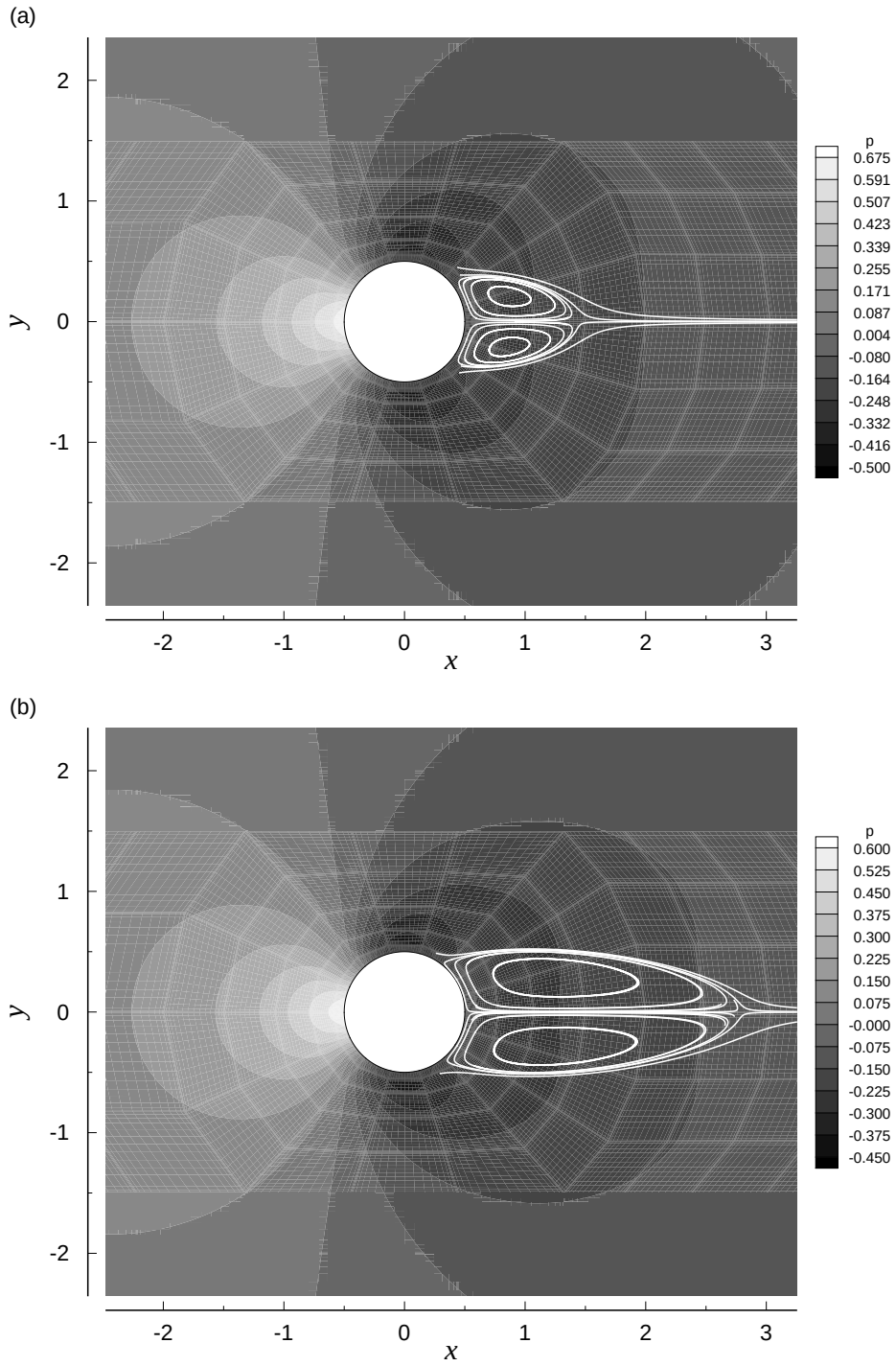

Fig. 38. Flow past a circular cylinder at (a) Re = 20 and (b) 40: Pressure contours and streamlines in the wake region.

b. Simulation at Re = 100

Next we consider the case Re = 100, for which a transient simulation is necessary. The spatial discretization used is similar to that shown in Fig. 36, except that the outflow boundary is placed 25.5 cylinder diameters downstream of the center of the cylinder. This was done in order to maintain the number of elements constant st 501 and redistribute the elements behind the cylinder as to provide better h-resolution. For the space-time coupled formulation, a single element layer is used to account for the temporal discretization. We use nodal expansions with $p_\xi = p_\eta = p_\gamma = 4$ in each element, i.e., uniform p-resolution of order 4 in space-time.

The fluid is initially at rest and the freestream velocity is gradually increased in time from 0.0 to 1.0 according to a hyperbolic tangent distribution, $u_\infty(t) = \tanh(t)$. An alternate initial condition is a steady-state solution (computed with the steady-state solver). The final periodic steady-state should be independent of the well-posed initial condition.

For the time marching procedure the size of the time step, $\Delta t = t_{s+1} - t_s$, was chosen as $\Delta t = 0.5$. Even though the space-time coupled formulation is stable for arbitrarily large time increments, a (somewhat) small time increment is desirable. When a small time increment is used the solution from the previous space-time strip serves as a very good initial guess for the solution at the current strip, thus the conjugate gradient method takes only a few iterations to converge which in turn significantly decreases the CPU solve time per space-time strip.

We march in time until a periodic steady-state is well established. Convergence of the conjugate gradient method was declared when the norm of the residual was less than 10^{-6}. Nonlinear convergence was declared when the relative norm of the residual in velocities between two consecutive iterations was less than 10^{-4}, which

typically required three Newton iterations. The L_2 space-time least-squares functional remained below 10^{-3} throughout the time marching procedure.

We choose two points in the near wake of the cylinder to trace the change of the velocity components and vorticity with time. Figures 39 and 40 show the time history of the v-velocity component and vorticity at the points $(x, y) = (1, 0)$ and $(x, y) = (2, 0)$, located 0.5 and 1.5 cylinder diameters behind the cylinder. From the figures we see that shedding starts around $t = 50$. No artificial perturbation is used to induce the vortex shedding. The flow reaches a periodic steady-state by $t = 100$. The shedding period for the v-velocity component is the same as that for the vorticity. The shedding period, obtained from Fig. 39 to within 0.05 time units, is found to be $\mathcal{T} = 6.05$; which gives a dimensionless shedding frequency of St $= 0.1653$. Our results are in good agreement with the experimental results of Williamson [109] and with the high-order splitting spectral/hp numerical results of Sherwin and Karniadakis [98], reported as St $= 0.1643$ and 0.1667 respectively.

The viscous and pressure forces acting on the cylinder are given by

$$\mathbf{F}_s = -\oint \mu \left[(\nabla \mathbf{u}) + (\nabla \mathbf{u})^T \right] \cdot \hat{\mathbf{n}}\, ds\,, \quad \mathbf{F}_p = -\oint p\, \hat{\mathbf{n}}\, ds$$

and the corresponding force coefficients obtained by normalizing the forces by the dynamic pressure, $\frac{1}{2}\rho u_\infty^2$, acting on a unit span of the circular cylinder. The predicted average drag coefficient is $\bar{C}_D = 1.345$, in good agreement with the high-order splitting spectral/hp numerical results of Henderson [44], reported as $\bar{C}_D = 1.35$. The amplitude of the lift coefficient is predicted as $C_L = \pm 0.332$. Figure 41 shows the time history of the lift coefficient with its pressure and viscous contributions.

Figure 42 depicts the vortex shedding cycle behind the circular cylinder by showing vorticity contours at four successive times during one period. Eddies are formed behind the cylinder and are washed away into the wake region. This flow pattern

is popularly known as the von Karman vortex street. Two eddies, alternatively of positive and negative vorticity, are shed within each period from the aft of the circular cylinder. Figure 43 shows instantaneous velocity and pressure contours at the reference dimensionless time $t_0 = 160$. From visual inspection of the contour plots in Figs. 42 and 43 it is clear that the outflow boundary condition allows the flow to exit the computational domain gracefully and does not disturb the upstream flow.

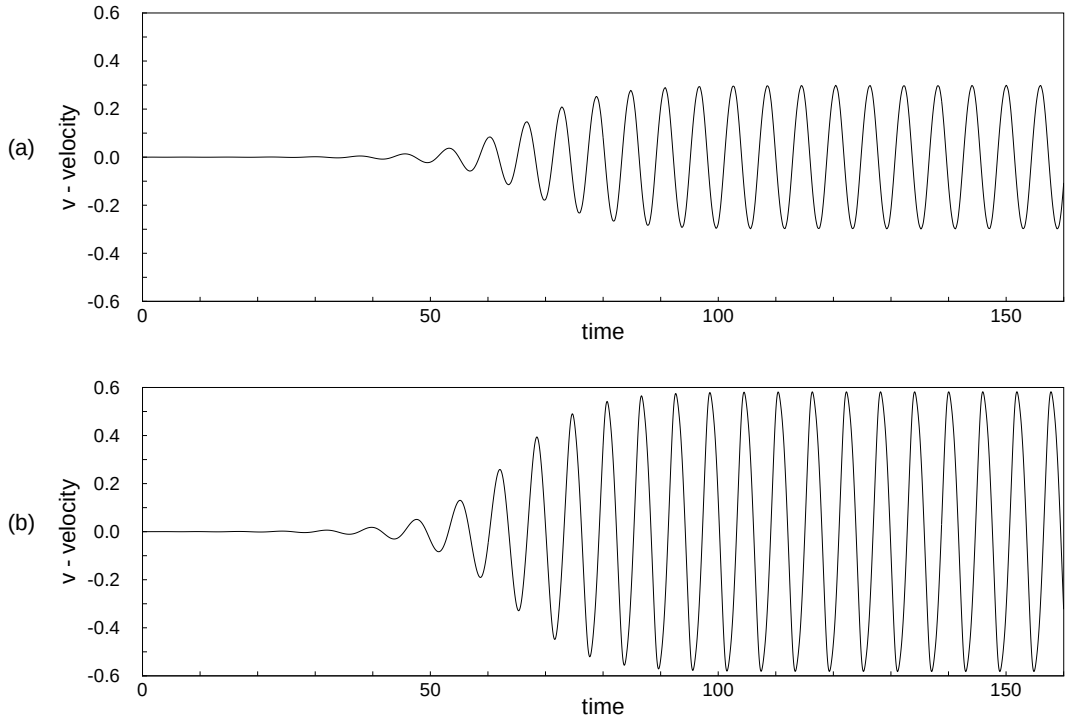

Fig. 39. Time history of v-velocity component behind the circular cylinder at points: (a) $(x, y) = (1, 0)$ and (b) $(x, y) = (2, 0)$.

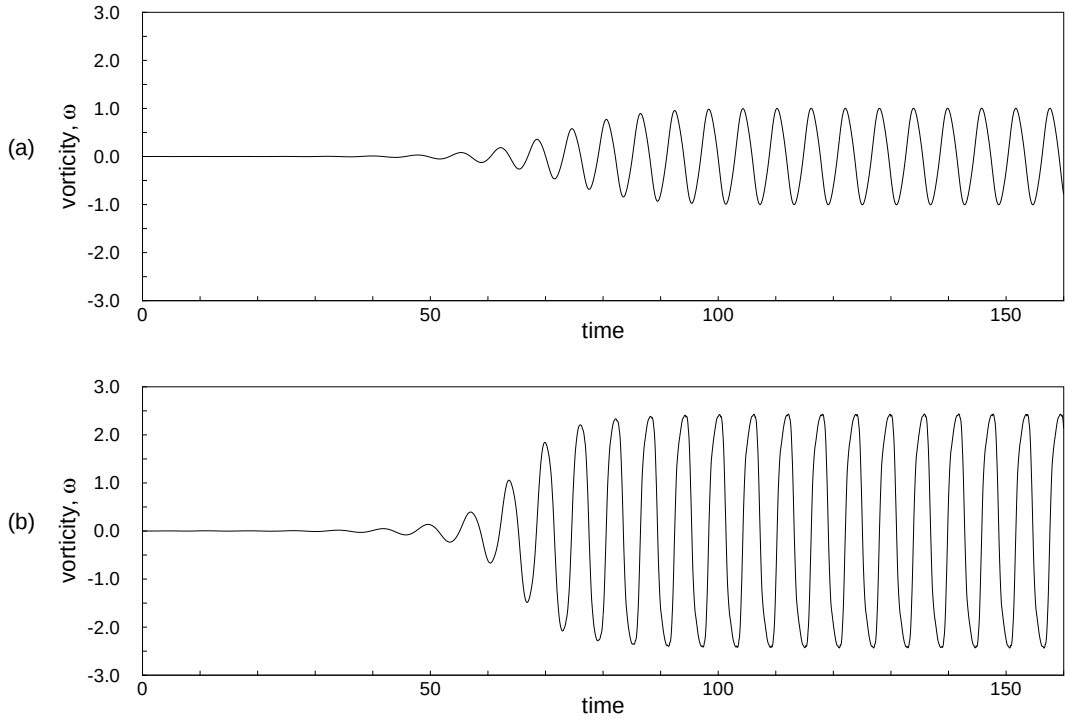

Fig. 40. Time history of vorticity behind the circular cylinder at points: (a) $(x, y) = (1, 0)$ and (b) $(x, y) = (2, 0)$.

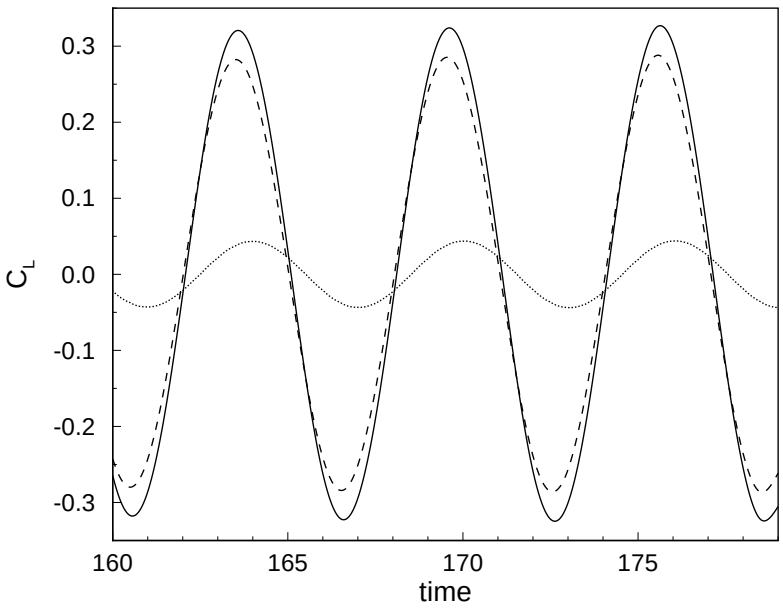

Fig. 41. Time history of lift coefficient (solid line). Shown is also the pressure contribution (dashed line) and the viscous contribution (dotted line).

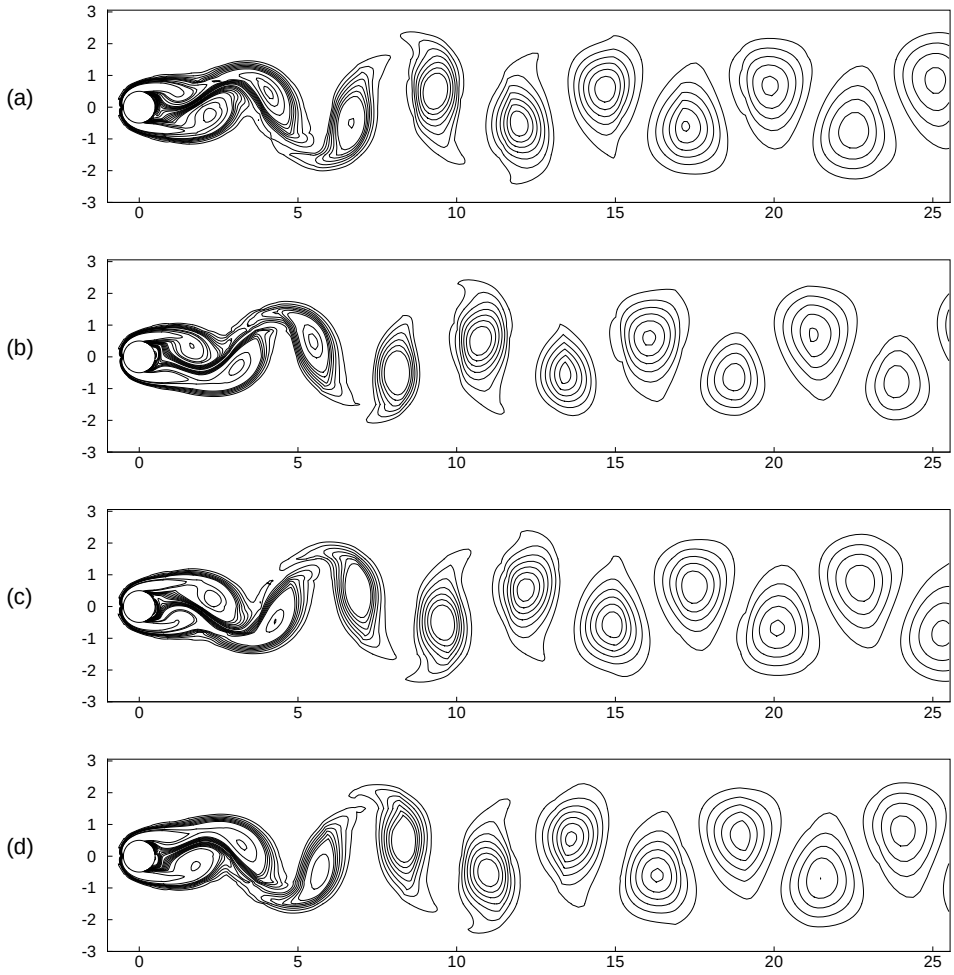

Fig. 42. Time history of vorticity contours behind the circular cylinder for four successive moments of time over a period. (a) t_0, (b) $t_0 + 3\Delta t$, (c) $t_0 + 6\Delta t$, (d) $t_0 + 9\Delta t$.

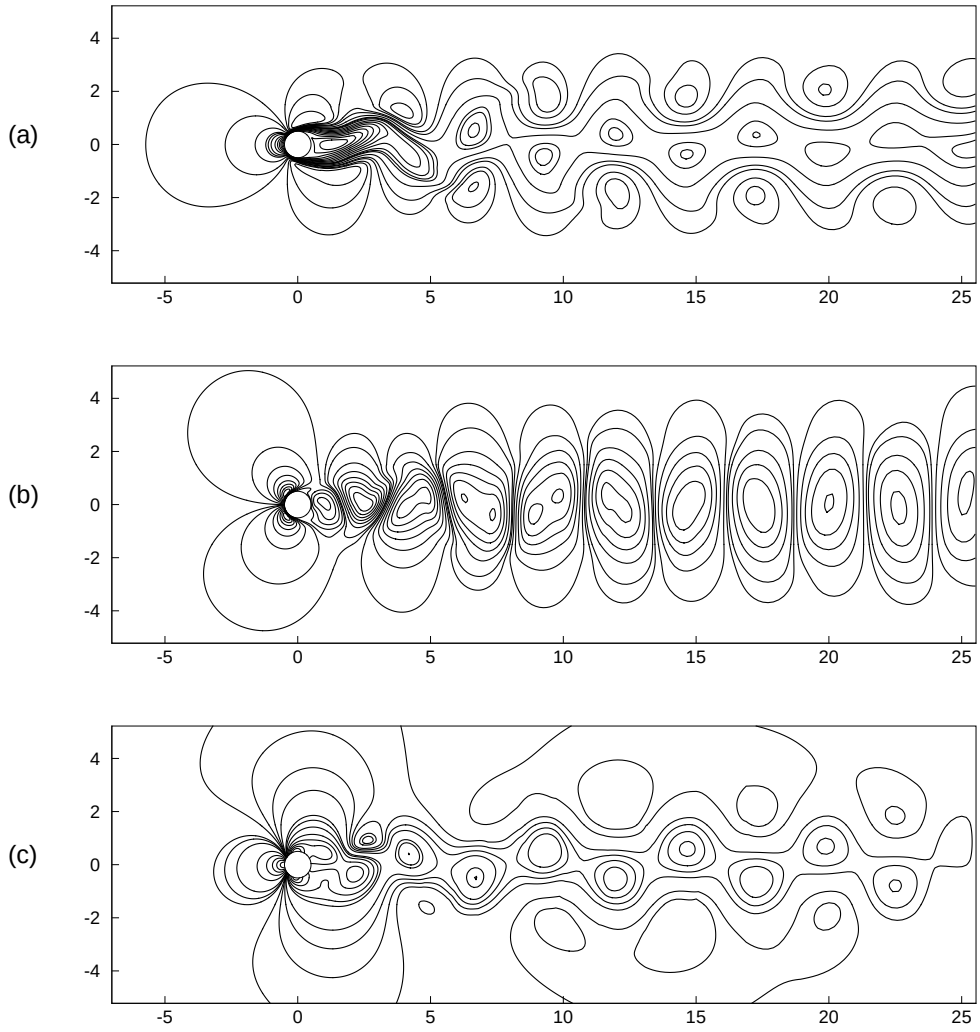

Fig. 43. Instantaneous (a) u-velocity, (b) v-velocity, and (c) pressure contours for the flow around a circular cylinder at $t = t_0$.

CHAPTER IV

VISCOUS COMPRESSIBLE FLUID FLOWS

In this chapter we present a least-squares formulation for the compressible Navier-Stokes equations. Numerical solutions for compressible flows using least-squares based finite element models have been limited to inviscid flows (i.e., the compressible Euler equations) [57, 69, 100, 81] and low-speed viscous compressible flows (i.e., anaelastic flows) [112, 87]. Here, for the first time, we present a least-squares formulation and accompanying numerical results for viscous compressible flows in the subsonic regime [83]. Such formulations are aimed at flow problems characterized by incompressible flow in parts of the domain with imbedded regions where compressibility effects are significant and cannot be neglected. For example, flow around an externally cooled turbine airfoil is characterized by an upstream Mach number of 0.3 (the upper end of the incompressible flow regime), but flow over the airfoil may reach local Mach numbers as high as 0.7 or 0.8. In addition, due to the external cooling mechanism, density variations induced by high temperature gradients cannot be neglected.

The formulation is not limited to simulate flows in the subsonic regime. It can be used directly to simulate transonic or supersonic flows. However, as we shall demonstrate later by a numerical example, such simulations would require an adaptive h-refinement scheme for adequate shock resolution, which we have not yet implemented. The formulation is a a density based formulation, developed using the non-conservative form of the compressible Navier-Stokes. A pressure based formulation is also possible and has been investigated, details of which will be reported in a near-future publication.

Traditional formulations for viscous compressible flows are based on treatment of the inviscid part of the equations by applying methods for the Euler equations and separate treatment of the viscous part of the equations. The underlying argument being that the viscous contributions can be thought of as corrections to the hyperbolic Euler equations, and thus a different numerical approach to treat the viscous part is allowed. For example, modern formulations based on the weak form Galerkin approach use a discontinuous formulation to treat the hyperbolic contributions and a mixed discontinuous (or continuous) formulation to treat the viscous contributions [72, 107]. Such an approach requires that interface patching schemes between elements be devised. Interface boundary conditions for the Euler part of the equations are imposed by specifying the incoming characteristic variables. In addition, when treating the viscous part with a mixed continuous formulation a reliable and optimal discretization requires that an inf-sup condition be satisfied. Another type of discontinuous weak form Galerkin formulation is that where a single variational statement is used to treat the conservation laws in flux vector form [7]. Such an approach does not require a mixed formulation for the viscous part of the equations.

In the proposed formulation no inviscid/viscous splitting is performed and no interface patching schemes are needed. The viscous compressible Navier-Stokes equations in non-conservative form are expressed as an equivalent set of first-order equations by introducing the velocity and temperature gradients as independent variables. The least-squares functional is constructed using the L_2 norms of the governing equations residuals and minimized with respect to the approximation spaces to develop the finite element model of the equations (see Chapter II). The formulation, based on least-squares variational principles, results in an unconstrained minimization problem through a variational framework of *residual* minimization. In such a variational setting the approximation spaces can be chosen independently, i.e. stability requirements such as the inf-sup condition never arise.

The issue of open boundary conditions of inflow and outflow type for the compressible Navier-Stokes equations is of paramount importance and is addressed herein. Artificial open boundaries are unavoidably introduced, for example, when addressing external flow problems where due to computational reasons one is forced to truncate the computational domain. Imposition of boundary conditions on the open boundaries becomes necessary and has been the subject of extensive research and numerical experimentation (see, for example, [41, 78, 37, 76, 28, 77, 46]).

For the purely hyperbolic problem, i.e. the Euler equations, it is well known that specifying the incoming characteristics results in a well posed problem and stable numerical approximation. The Euler characteristics have been used to impose boundary conditions at open boundaries for the viscous compressible Navier-Stokes equations. Although there are theoretical problems with this, it works well in practice (see, for example, [28]). The argument has been that sufficiently far away from the immersed body the flow is essentially inviscid (except in the wake region) and the Euler characterics can be used to impose the boundary conditions as a good approximation. In this study we account for the viscous effects and apply the energy method to the linearized, constant coefficient viscous compressible Navier-Stokes equations in non-conservative form to obtain energy inequalities which bound the temporal growth of solutions to the initial-boundary value problem. Boundary operators which ensure maximum energy dissipation are obtained from the inequalities. The resulting boundary conditions ensure a strongly well posed problem and in the limit of vanishing viscocity recover the Euler characteristics.

Discussion and derivation of well-posed boundary conditions for the viscous compressible Navier-Stokes equation in non-conservative form was considered by Gustafsson and Sundström [41], Oliger and Sundström [78], and Nordström [76]. They all used the energy method to derive well-posed, dissipative boundary conditions.

Gustafsson and Sundström's approach was to start with a set of boundary conditions that would give a well-posed problem for the corresponding hyperbolic system and modify them in such a way that the number of conditions was correct and the energy inequality remained satisfied. In the present study we follow the latest work of Nordström [77] and use energy inequalities in terms of characteristic variables to arrive at the desired boundary conditions. This results in *maximally* dissipative boundary conditions and a *strongly* well-posed problem. The same procedure was also used by Hesthaven and Gottlieb [46] in deriving boundary conditions for the viscous compressible Navier-Stokes equations in conservative form.

We also address numerical implementation regarding the imposition of the boundary conditions. We develop a fully coupled space-time least-squares finite element model for the viscous compressible Navier-Stokes equations and use high order nodal expansions (spectral basis) in space-time to discretize the model. Imposition of boundary conditions at the open boundaries is done through the least-squares functional. Weak imposition of the boundary conditions through the least-squares functional enforces the boundary conditions as well as considers the governing equations at the boundary. The use of least-squares principles leads to a symmetric and positive definite algebraic system of equations. The system of equations is linearized by Newton's method and solved by the preconditioned conjugate gradient method in matrix free form.

Numerical implementation regarding the imposition of the boundary conditions for the conservative form of the viscous compressible Navier-Stokes equations was addressed by Hesthaven and Gottlieb [46] in the context of a pseudo-spectral collocation method. Boundary conditions were enforced through an asymptotically stable penalty method. Appropriate bounds for the penalty parameter and its effect on the CFL condition when using explicit Runge-Kutta methods for time-stepping were

discussed. In the proposed formulation there is no need for a penalty parameter, as the boundary conditions are enforced in a least-squares sense through the functional. Furthermore, the issue of time-stepping stability never arises in a fully coupled space-time formulation [86].

First, we present the governing equations and derive maximally dissipative open boundary conditions for the two-dimensional viscous compressible Navier-Stokes equations in non-conservative form using primitive variables (ρ, u, v, T). Exponentially fast decay of the L_2 least-squares functional and L_2 error norms is verified using smooth solutions to the stationary and non-stationary viscous compressible Navier-Stokes equations. The performance of the formulation is tested by simulating subsonic compressible flow past a circular cylinder for a range of free-stream Mach numbers.

A. The compressible Navier-Stokes equations

We consider the solution of the Navier-Stokes equations governing viscous compressible flow of an ideal, Newtonian gas. The governing equations in dimensionless, non-conservative form can be written as:

Find the density $\rho(\mathbf{x}, t)$, velocity $\mathbf{u}(\mathbf{x}, t)$, pressure $p(\mathbf{x}, t)$, and temperature $T(\mathbf{x}, t)$ such that

$$\frac{\partial \rho}{\partial t} + (\mathbf{u} \cdot \nabla)\,\rho + \rho\,(\nabla \cdot \mathbf{u}) = 0 \qquad \text{in } \Omega \times (0, \tau] \tag{4.1}$$

$$\rho \frac{\partial \mathbf{u}}{\partial t} + \rho\,(\mathbf{u} \cdot \nabla)\,\mathbf{u} + \nabla p - \frac{1}{\mathrm{Re}}\,\nabla \cdot \mu\,\left[(\nabla \mathbf{u}) + (\nabla \mathbf{u})^T\right] - \frac{1}{\mathrm{Re}}\,\nabla\,\lambda\,(\nabla \cdot \mathbf{u}) = \mathbf{f} \qquad \text{in } \Omega \times (0, \tau] \tag{4.2}$$

$$\rho \frac{\partial T}{\partial t} + \rho\,(\mathbf{u} \cdot \nabla)\,T - \frac{1}{\mathrm{Pe}}\,\nabla\,(k\,\nabla T) - (\gamma - 1)\,\mathrm{M}^2\,\frac{Dp}{Dt} = \frac{(\gamma - 1)\,\mathrm{M}^2}{\mathrm{Re}}\,\Phi \qquad \text{in } \Omega \times (0, \tau] \tag{4.3}$$

$$1 + \gamma \, \mathrm{M}^2 \, p = \rho \, T \qquad \text{in } \Omega \times (0, \tau] \tag{4.4}$$

where τ is a real number (time) > 0, M is the Mach number, Pr is the Prandtl number, Re is the Reynolds number, $\mathrm{Pe} = \mathrm{Re}\,\mathrm{Pr}$ is the Peclet number, γ is the ratio of specific heats, μ is the dynamic viscosity, λ is the bulk viscosity, k is the coefficient of thermal conductivity, and Φ is the viscous dissipation function,

$$\Phi = \lambda \, (\nabla \cdot \mathbf{u})^2 + 2 \, \mu \, \underline{\mathbf{D}} : \underline{\mathbf{D}} \, ,$$

with $\underline{\mathbf{D}} = (1/2) \left[(\nabla \mathbf{u}) + (\nabla \mathbf{u})^T \right]$ denoting the deformation tensor. We assume initial and boundary conditions are given such that the problem is well posed.

1. Well-posed open boundary conditions

Next, we present in detail the derivation of boundary conditions that are maximally dissipative and lead to a strongly well posed problem. The derivation presented here is strongly based on previous work by Nordström [77] and Hesthaven [45].

We consider the two-dimensional viscous compressible Navier-Stokes equations in non-conservative form. For the sole purpose of deriving the open boundary conditions we neglect off-diagonal terms in the stress tensor, the full stress tensor will be considered in future work.

In dimensionless form the linearized, constant-coefficient Navier-Stokes equations can be written as

$$\frac{\partial \mathbf{W}}{\partial t} + \underline{\mathbf{A}}_1 \frac{\partial \mathbf{W}}{\partial x} + \underline{\mathbf{A}}_2 \frac{\partial \mathbf{W}}{\partial y} = \varepsilon \left(\underline{\mathbf{B}}_1 \frac{\partial^2 \mathbf{W}}{\partial x^2} + \underline{\mathbf{B}}_2 \frac{\partial^2 \mathbf{W}}{\partial y^2} \right) \tag{4.5}$$

where $\mathbf{W} = [\rho, u, T, v]^T$, $\varepsilon = 1/\mathrm{Re}$,

$$\underline{\mathbf{A}}_1 = \begin{bmatrix} \bar{u} & \bar{\rho} & 0 & 0 \\ \dfrac{\bar{c}^2}{\gamma\bar{\rho}} & \bar{u} & \dfrac{1}{\gamma \mathrm{M}^2} & 0 \\ 0 & (\gamma-1)\,\bar{c}^2 \mathrm{M}^2 & \bar{u} & 0 \\ 0 & 0 & 0 & \bar{u} \end{bmatrix}, \quad \underline{\mathbf{A}}_2 = \begin{bmatrix} \bar{v} & 0 & 0 & \bar{\rho} \\ 0 & \bar{v} & 0 & 0 \\ 0 & 0 & \bar{v} & (\gamma-1)\,\bar{c}^2 \mathrm{M}^2 \\ \dfrac{\bar{c}^2}{\gamma\bar{\rho}} & 0 & \dfrac{1}{\gamma \mathrm{M}^2} & \bar{v} \end{bmatrix},$$

$$\underline{\mathbf{B}}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \dfrac{(2\bar{\mu}+\bar{\lambda})}{\bar{\rho}} & 0 & 0 \\ 0 & 0 & \dfrac{\gamma\bar{k}}{\mathrm{Pr}\,\bar{\rho}} & 0 \\ 0 & 0 & 0 & \dfrac{\bar{\mu}}{\bar{\rho}} \end{bmatrix}, \qquad \underline{\mathbf{B}}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \dfrac{\bar{\mu}}{\bar{\rho}} & 0 & 0 \\ 0 & 0 & \dfrac{\gamma\bar{k}}{\mathrm{Pr}\,\bar{\rho}} & 0 \\ 0 & 0 & 0 & \dfrac{(2\bar{\mu}+\bar{\lambda})}{\bar{\rho}} \end{bmatrix},$$

c denotes the speed of sound, and the Navier-Stokes have been linearized around a constant state $(\bar{\rho}, \bar{u}, \bar{T}, \bar{v})$, which in general is different from the reference state (ρ_0, u_0, T_0, v_0) about which Re, Pr, and M have been defined and the Navier-Stokes non-dimensionalized. In Eq. (4.5) the pressure has been eliminated using the equations of state, Eq. (4.4), resulting in a density based system.

a. The Navier-Stokes equations in symmetric characteristic form

For ease of exposition, the boundary conditions will be imposed in the x direction and hence we work with the following form of the Navier-Stokes equations

$$\frac{\partial \mathbf{W}}{\partial t} + \underline{\mathbf{A}}_1 \frac{\partial \mathbf{W}}{\partial x} = \varepsilon\, \underline{\mathbf{B}}_1 \frac{\partial^2 \mathbf{W}}{\partial x^2} \tag{4.6}$$

and rewrite $\underline{\mathbf{A}}_1$ and $\underline{\mathbf{B}}_1$ as follows

$$\underline{\mathbf{A}}_1 = \begin{bmatrix} \underline{\mathbf{A}} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \bar{u} \end{bmatrix}, \quad \underline{\mathbf{B}}_1 = \begin{bmatrix} \underline{\mathbf{B}} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \bar{\mu}/\bar{\rho} \end{bmatrix}$$

where $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$ are 3×3 matrices containing the first three rows and columns of $\underline{\mathbf{A}}_1$ and $\underline{\mathbf{B}}_1$ respectively. In what follows we need only work with the 3×3 matrices since the last equation is decoupled from the first three and the last equation is already in the desired form. Hence for the time being we need only work with

$$\frac{\partial \mathbf{w}}{\partial t} + \underline{\mathbf{A}}\,\frac{\partial \mathbf{w}}{\partial x} = \varepsilon\,\underline{\mathbf{B}}\,\frac{\partial^2 \mathbf{w}}{\partial x^2} \tag{4.7}$$

where $\mathbf{w} = [\rho, u, T]^T$.

The energy method requires that we work with symmetric differential operators, to this end we symmetrize Eq. (4.7) with the symmetrizer $\underline{\mathbf{S}}$ as follows

$$\frac{\partial(\underline{\mathbf{S}}\mathbf{w})}{\partial t} + \underline{\mathbf{S}}\,\underline{\mathbf{A}}\,\underline{\mathbf{S}}^{-1}\frac{\partial(\underline{\mathbf{S}}\mathbf{w})}{\partial x} = \varepsilon\ \underline{\mathbf{S}}\,\underline{\mathbf{B}}\,\underline{\mathbf{S}}^{-1}\frac{\partial^2(\underline{\mathbf{S}}\mathbf{w})}{\partial x^2} \tag{4.8}$$

where

$$\underline{\mathbf{S}} = \sqrt{2}\begin{bmatrix} \bar{c}^{\,2}/\sqrt{\gamma} & 0 & 0 \\ 0 & \bar{\rho}\,\bar{c} & 0 \\ 0 & 0 & \bar{\rho}/\sqrt{\gamma\,(\gamma-1)\,\mathrm{M}^4} \end{bmatrix}$$

Defining $\underline{\mathbf{A}}^{\mathrm{S}} = \underline{\mathbf{S}}\,\underline{\mathbf{A}}\,\underline{\mathbf{S}}^{-1}$ and realizing that $\underline{\mathbf{B}}^{\mathrm{S}} = \underline{\mathbf{S}}\,\underline{\mathbf{B}}\,\underline{\mathbf{S}}^{-1} = \underline{\mathbf{B}}$ because $\underline{\mathbf{B}}$ is a diagonal matrix; Eq. (4.8) is written as

$$\frac{\partial(\underline{\mathbf{S}}\mathbf{w})}{\partial t} + \underline{\mathbf{A}}^{\mathrm{S}}\,\frac{\partial(\underline{\mathbf{S}}\mathbf{w})}{\partial x} = \varepsilon\,\underline{\mathbf{B}}\,\frac{\partial^2(\underline{\mathbf{S}}\mathbf{w})}{\partial x^2} \tag{4.9}$$

where

$$\underline{\mathbf{A}}^{\mathrm{S}} = \begin{bmatrix} \bar{u} & \bar{c}\,/\sqrt{\gamma} & 0 \\ \bar{c}\,/\sqrt{\gamma} & \bar{u} & \bar{c}\,\sqrt{(\gamma-1)/\gamma} \\ 0 & \bar{c}\,\sqrt{(\gamma-1)/\gamma} & \bar{u} \end{bmatrix}$$

The matrix $\underline{\mathbf{A}}^{\mathrm{S}}$ is symmetric with distinct eigenvalues and orthogonal eigenvectors. Thus $\underline{\mathbf{A}}^{\mathrm{S}}$ can be expressed as $\underline{\mathbf{A}}^{\mathrm{S}} = \underline{\mathbf{K}}\,\underline{\boldsymbol{\lambda}}\,\underline{\mathbf{K}}^{-1}$, in terms of a diagonal matrix $\underline{\boldsymbol{\lambda}}$ and a matrix $\underline{\mathbf{K}}$. The diagonal elements of $\underline{\boldsymbol{\lambda}}$ are the eigenvalues of $\underline{\mathbf{A}}^{\mathrm{S}}$ and the

columns of $\underline{\mathbf{K}}$ are the right eigenvectors of $\underline{\mathbf{A}}^{\mathrm{S}}$. Specifically,

$$\underline{\boldsymbol{\lambda}} = \begin{bmatrix} \bar{u}-\bar{c} & 0 & 0 \\ 0 & \bar{u} & 0 \\ 0 & 0 & \bar{u}+\bar{c} \end{bmatrix}, \quad \underline{\mathbf{K}}^{-1} = \begin{bmatrix} -1/\sqrt{2\gamma} & 1/\sqrt{2} & -\sqrt{(\gamma-1)/2\gamma} \\ \sqrt{(\gamma-1)/\gamma} & 0 & -1/\sqrt{\gamma} \\ 1/\sqrt{2\gamma} & 1/\sqrt{2} & \sqrt{(\gamma-1)/2\gamma} \end{bmatrix}$$

We define the symmetric characteristic variables as $\mathbf{c} = \underline{\mathbf{K}}^{-1}\underline{\mathbf{S}}\,\mathbf{w}$, so that an equation in terms of symmetric characteristic variables can be obtained from Eq. (4.9) as follows

$$\frac{\partial(\underline{\mathbf{K}}^{-1}\underline{\mathbf{S}}\,\mathbf{w})}{\partial t} + \underline{\mathbf{K}}^{-1}\underline{\mathbf{A}}^{\mathrm{S}}\,\underline{\mathbf{K}}\,\frac{\partial(\underline{\mathbf{K}}^{-1}\underline{\mathbf{S}}\,\mathbf{w})}{\partial x} = \varepsilon\,\underline{\mathbf{K}}^{-1}\underline{\mathbf{B}}\,\underline{\mathbf{K}}\,\frac{\partial^2(\underline{\mathbf{K}}^{-1}\underline{\mathbf{S}}\,\mathbf{w})}{\partial x^2}$$

or, equivalently

$$\frac{\partial \mathbf{c}}{\partial t} + \underline{\boldsymbol{\lambda}}\,\frac{\partial \mathbf{c}}{\partial x} = \varepsilon\,\underline{\mathbf{x}}\,\frac{\partial^2 \mathbf{c}}{\partial x^2} \tag{4.10}$$

where

$$\underline{\mathbf{x}} = \underline{\mathbf{K}}^{-1}\underline{\mathbf{B}}\,\underline{\mathbf{K}} = \frac{1}{2}\begin{bmatrix} \theta+\phi & \alpha\,\phi & \theta-\phi \\ \alpha\,\phi & \alpha^2\,\phi & -\alpha\,\phi \\ \theta-\phi & -\alpha\,\phi & \theta+\phi \end{bmatrix}$$

and

$$\theta = \frac{(2\bar{\mu}+\bar{\lambda})}{\bar{\rho}}, \quad \phi = \frac{(\gamma-1)\bar{k}}{\mathrm{Pr}\,\bar{\rho}}, \quad \alpha = \sqrt{\frac{2}{\gamma-1}}$$

The final equation based on symmetric characteristic form corresponding to Eq. (4.6) (the 4×4 system) is

$$\frac{\partial \mathbf{C}}{\partial t} + \underline{\boldsymbol{\Lambda}}\,\frac{\partial \mathbf{C}}{\partial x} = \varepsilon\,\underline{\mathbf{X}}\,\frac{\partial^2 \mathbf{C}}{\partial x^2} \tag{4.11}$$

where

$$\underline{\boldsymbol{\Lambda}} = \begin{bmatrix} \underline{\boldsymbol{\lambda}} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \bar{u} \end{bmatrix}, \quad \underline{\mathbf{X}} = \begin{bmatrix} \underline{\mathbf{x}} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \bar{\mu}/\bar{\rho} \end{bmatrix}$$

and $\mathbf{C} = [\mathbf{c}^T, v]^T = [\,\bar{\rho}\,\bar{c}\,u - p\,,\, \alpha\,(\rho\,\bar{c}^{\,2} - p)\,,\, \bar{\rho}\,\bar{c}\,u + p\,,\, v\,]^T$. Where p is a linearized

pressure defined as

$$p \equiv \frac{1}{\gamma \mathrm{M}^2} \left(\rho \bar{T} + \bar{\rho} T \right)$$

b. Maximally dissipative boundary conditions

Consider the following problem for the x-split Navier-Stokes equations in symmetric characteristic form:

$$\frac{\partial \mathbf{C}}{\partial t} + \underline{\mathbf{\Lambda}} \frac{\partial \mathbf{C}}{\partial x} = \varepsilon \underline{\mathbf{X}} \frac{\partial^2 \mathbf{C}}{\partial x^2} \qquad t \geq 0, \quad x \in [0,1] \tag{4.12}$$

$$\mathbf{C} = {}^0\mathbf{C}(x) \qquad t = 0, \quad x \in [0,1] \tag{4.13}$$

$$\underline{\mathbf{L}}_{-} \mathbf{C} = \mathbf{g}_{-}(t) \qquad t \geq 0, \quad x = 0 \tag{4.14}$$

$$\underline{\mathbf{L}}_{+} \mathbf{C} = \mathbf{g}_{+}(t) \qquad t \geq 0, \quad x = 1 \tag{4.15}$$

where $\underline{\mathbf{L}}_{-}$ and $\underline{\mathbf{L}}_{+}$ are boundary operators to be determined such that the temporal growth of energy of the initial-boundary value problem is bounded.

Applying the energy method, obtain the energy inequality by multiplying Eq. (4.12) with $\mathbf{C}^T$ and integrate over $x \in [0,1]$ to arrive at

$$\frac{1}{2} \frac{\partial ||\mathbf{C}||^2}{\partial t} = - \int_0^1 \mathbf{C}^T \underline{\mathbf{\Lambda}} \frac{\partial \mathbf{C}}{\partial x} \, dx + \varepsilon \int_0^1 \mathbf{C}^T \underline{\mathbf{X}} \frac{\partial^2 \mathbf{C}}{\partial x^2} \, dx$$

integration by parts yields

$$\frac{1}{2} \frac{\partial ||\mathbf{C}||^2}{\partial t} = -\frac{1}{2} \left[\mathbf{C}^T \underline{\mathbf{\Lambda}} \, \mathbf{C} - 2 \varepsilon \, \mathbf{C}^T \underline{\mathbf{X}} \frac{\partial \mathbf{C}}{\partial x} \right]_{x=0}^{x=1} - \varepsilon \int_0^1 \frac{\partial \mathbf{C}^T}{\partial x} \underline{\mathbf{X}} \frac{\partial \mathbf{C}}{\partial x} \, dx \leq 0$$

If the boundary conditions in the x-direction are constructed in such a way that

$$-\frac{1}{2} \left[\mathbf{C}^T \underline{\mathbf{\Lambda}} \, \mathbf{C} - 2 \varepsilon \, \mathbf{C}^T \underline{\mathbf{X}} \frac{\partial \mathbf{C}}{\partial x} \right]_{x=0}^{x=1} \leq 0$$

then it is evident that if we can show that

$$\varepsilon \int_0^1 \frac{\partial \mathbf{C}^T}{\partial x} \underline{\mathbf{X}} \frac{\partial \mathbf{C}}{\partial x}\, dx \;\geq 0$$

the initial-boundary value problem is strongly well-posed.

It is a matter of algebra to show that

$$\varepsilon \int_0^1 \frac{\partial \mathbf{C}^T}{\partial x} \underline{\mathbf{X}} \frac{\partial \mathbf{C}}{\partial x}\, dx = \int_0^1 (\, 2\,\theta\,\Delta_1^2 + 2\,\phi\,\Delta_2^2 + (\bar{\mu}/\bar{\rho})\,\Delta_4^2\,)\, dx \;\geq 0$$

if the following conditions are satisfied

$$\bar{\mu} \geq 0\,, \quad \bar{\lambda} + 2\bar{\mu} \geq 0\,, \quad \frac{\bar{k}}{\mathrm{Pr}} \geq 0\,, \quad \gamma \geq 1$$

These conditions are natural, in fact if not obeyed the Navier-Stokes equations violate the second law of thermodynamics.

We are now left with the task to construct an appropriate boundary operator. We rewrite the boundary operator inequality, aligned with the outward pointing unit normal, as follows

$$-\sum_{i=1}^{4} \frac{1}{\lambda_i^n} \left[\left(|\lambda_i^n| C_i^n - \varepsilon\, \frac{\lambda_i^n}{|\lambda_i^n|}\, G_i^n \right)^2 - (\,\varepsilon\, G_i^n\,)^2 \right] \tag{4.16}$$

where

$$\begin{pmatrix} \lambda_1^n \\ \lambda_2^n \\ \lambda_3^n \\ \lambda_4^n \end{pmatrix} = \begin{pmatrix} \bar{u}_n - \bar{c} \\ \bar{u}_n \\ \bar{u}_n + \bar{c} \\ \bar{u}_n \end{pmatrix}, \begin{pmatrix} C_1^n \\ C_2^n \\ C_3^n \\ C_4^n \end{pmatrix} = \begin{pmatrix} \bar{\rho}\,\bar{c}\,u_n - p \\ \alpha\,(\rho\,\bar{c}^{\,2} - p) \\ \bar{\rho}\,\bar{c}\,u_n + p \\ v \end{pmatrix}, \begin{pmatrix} G_1^n \\ G_2^n \\ G_3^n \\ G_4^n \end{pmatrix} = \begin{pmatrix} \theta\,\Delta_1^n + \phi\,\Delta_2^n \\ \alpha\,\phi\,\Delta_2^n \\ \theta\,\Delta_1^n - \phi\,\Delta_2^n \\ (\bar{\mu}/\bar{\rho})\,\Delta_4^n \end{pmatrix}$$

and

$$\Delta_1^n = \frac{1}{2}\frac{\partial}{\partial x}\left(C_1^n + C_3^n \right) = \bar{\rho}\,\bar{c}\,\frac{\partial u_n}{\partial x}$$

$$\Delta_2^n = \frac{1}{2}\frac{\partial}{\partial x}\left(C_1^n + \alpha\, C_2^n - C_3^n \right) = -\frac{\bar{\rho}}{(\gamma-1)\,\mathrm{M}^2}\,\frac{\partial T}{\partial x}$$

$$\Delta_4^n = \frac{\partial v}{\partial x}$$

It is now fairly obvious how to construct appropriate boundary operators for the continuous viscous compressible Navier-Stokes equations. Depending on the sign of the eigenvalues λ_i^n, the contribution of the boundary terms is made as negative as possible.

Without loss of generality, assume $\bar{u} > 0$, so that there is inflow at $x = 0$ and outflow at $x = 1$. For the moment we consider the case of homogeneous boundary data. Let us now consider all the possibilities at the open boundaries:

Inflow. At an inflow boundary $u_n < 0$ or $u_n = -u$

Subsonic inflow: $\lambda_1^n < 0, \quad \lambda_2^n < 0, \quad \lambda_3^n > 0, \quad \lambda_4^n < 0$

$$\begin{aligned} |\lambda_1^n| C_1^n + \varepsilon\, G_1^n &= 0 \\ |\lambda_2^n| C_2^n + \varepsilon\, G_2^n &= 0 \\ \varepsilon\, G_3^n &= 0 \\ |\lambda_4^n| C_4^n + \varepsilon\, G_4^n &= 0 \end{aligned} \tag{4.17}$$

Supersonic inflow: $\lambda_1^n < 0, \quad \lambda_2^n < 0, \quad \lambda_3^n < 0, \quad \lambda_4^n < 0$

$$\begin{aligned} |\lambda_1^n| C_1^n + \varepsilon\, G_1^n &= 0 \\ |\lambda_2^n| C_2^n + \varepsilon\, G_2^n &= 0 \\ |\lambda_3^n| C_3^n + \varepsilon\, G_3^n &= 0 \\ |\lambda_4^n| C_4^n + \varepsilon\, G_4^n &= 0 \end{aligned} \tag{4.18}$$

Outflow. At an outflow boundary $u_n > 0$ or $u_n = u$

Subsonic outflow: $\lambda_1^n < 0, \quad \lambda_2^n > 0, \quad \lambda_3^n > 0, \quad \lambda_4^n > 0$

$$\begin{aligned} |\lambda_1^n| C_1^n + \varepsilon\, G_1^n &= 0 \\ \varepsilon\, G_2^n &= 0 \\ \varepsilon\, G_4^n &= 0 \end{aligned} \tag{4.19}$$

Supersonic outflow: $\lambda_1^n > 0, \quad \lambda_2^n > 0, \quad \lambda_3^n > 0, \quad \lambda_4^n > 0$

$$\begin{aligned} \varepsilon\, G_1^n &= 0 \\ \varepsilon\, G_2^n &= 0 \\ \varepsilon\, G_4^n &= 0 \end{aligned} \tag{4.20}$$

Note that only three conditions need be specified at an outflow boundary. This is a consequence of $G_1^n = -G_3^n$ in the case when $G_2^n = 0$. Hence, for the two-dimensional viscous compressible Navier-Stokes equations, four boundary conditions are required at an inflow boundary and three at an outflow boundary; regardless if the flow is subsonic or supersonic. Furthermore, in the limit of the Reynolds number approaching infinity the characteristic boundary conditions for the compressible Euler equations are recovered.

2. The velocity/temperature gradient first-order system

Introducing the gradient of the velocity vector field and the gradient of the temperature field as independent variables, the compressible Navier-Stokes equivalent first-order system in dimensionless, non-conservative form can be written as:

Find the density $\rho(\mathbf{x},t)$, velocity $\mathbf{u}(\mathbf{x},t)$, velocity gradients $\underline{\mathbf{U}}(\mathbf{x},t)$, temperature $T(\mathbf{x},t)$, and temperature gradients $\mathbf{q}(\mathbf{x},t)$ such that

$$\frac{\partial \rho}{\partial t} + (\mathbf{u}\cdot\nabla)\,\rho + \rho\,(\nabla\cdot\mathbf{u}) = 0 \qquad \text{in } \Omega\times(0,\tau] \tag{4.21}$$

$$\rho\,\frac{\partial \mathbf{u}}{\partial t} + \rho\,(\mathbf{u}\cdot\nabla)\,\mathbf{u} + \frac{1}{\gamma\,\mathrm{M}^2}\,\nabla(\rho\,T) - \frac{1}{\mathrm{Re}}\,(\nabla\cdot\underline{\mathbf{U}})^T - \frac{1}{3}\,\frac{1}{\mathrm{Re}}\,\nabla(\mathrm{tr}\,\underline{\mathbf{U}}) = \mathbf{f} \qquad \text{in } \Omega\times(0,\tau] \tag{4.22}$$

$$\underline{\mathbf{U}} - \nabla(\mathbf{u})^T = \underline{\mathbf{0}} \qquad \text{in } \Omega\times(0,\tau] \tag{4.23}$$

$$\nabla\times\underline{\mathbf{U}} = \underline{\mathbf{0}} \qquad \text{in } \Omega\times(0,\tau] \tag{4.24}$$

$$\rho\,\frac{\partial T}{\partial t} + \rho\,(\mathbf{u}\cdot\nabla)\,T - \frac{\gamma}{\mathrm{Pe}}\,\nabla\cdot\mathbf{q} + (\gamma-1)\,\rho\,T\,(\mathrm{tr}\,\underline{\mathbf{U}}) = \frac{\gamma\,(\gamma-1)\,\mathrm{M}^2}{\mathrm{Re}}\,\Phi \qquad \text{in } \Omega\times(0,\tau] \tag{4.25}$$

$$\mathbf{q} - \nabla T = \mathbf{0} \qquad \text{in } \Omega\times(0,\tau] \tag{4.26}$$

$$\nabla\times\mathbf{q} = \mathbf{0} \qquad \text{in } \Omega\times(0,\tau] \tag{4.27}$$

where we've adopted Stokes's hypothesis to obtain $\lambda = -(2/3)\mu$, neglected temperature dependence effects on the dynamic viscosity and coefficient of thermal conductivity (i.e., $\mu = 1$, $k = 1$ in Eqs. (4.2) and (4.3)), and used the equation of state, Eq. (4.4), to eliminate pressure form the governing equations.

B. Numerical examples: verification benchmarks

Having presented the characteristic based, maximally dissipative boundary operators and the equivalent first-order system for the viscous compressible Navier-Stokes, we use the procedures outlined in Chapter II to develop and arrive at the least-squares based finite element model.

In this section we are concerned with the *verification* of the spectral/*hp* least-squares based finite element computational algorithm, we follow the verification procedures suggested in Ref. [80] and illustrated in Chapter III.

First, we present a (manufactured) analytic closed form solution to the compressible Navier-Stokes equations. We use the analytic solution for the purposes of algorithm verification, specifically to verify exponentially fast decay of errors for our stationary and non-stationary algorithms. We then use the benchmark of flow past a circular cylinder to further verify the computational algorithm. Keeping the free-stream Reynolds number fixed at 100, we consider a range of free-stream Mach numbers ($0.2 \leq \text{Mach} \leq 2.0$) and thermal loadings $T_{\text{cyl}}/T_{\infty}$ of 1.0 and 2.0. Where possible, we compare the predicted characteristic flow parameters with published numerical studies.

The limit problems of low-speed compressible flow and inviscid compressible flow are not documented here, since they are but special cases of this formulation. Instead, we refer the interested reader to our work in these flow regimes [87, 81].

1. Convergence

To verify spectral convergence, we use the method of manufactured solutions. We consider two-dimensional, unsteady flow in the bi-unit square, $\bar{\Omega} = [-1, 1] \times [-1, 1]$. We prescribe a solution to the viscous compressible Navier-Stokes equations, com-

pute the consistent forcing functions, and solve to verify spectral convergence of the numerical solution. We choose the exact solution to be of the form

$$\begin{aligned} \rho(x,y,t) &= a_\rho + b_\rho \sin(\omega_\rho x) \cos(\omega_{\rho,\tau} t) \\ u(x,y,t) &= a_u + b_u \cos(\omega_u x) \sin(\omega_u y) \cos(\omega_{u,\tau} t) \\ v(x,y,t) &= a_v + b_v \sin(\omega_v x) \cos(\omega_{v,\tau} t) \\ T(x,y,t) &= a_T + b_T \cos(\omega_T y) \cos(\omega_{T,\tau} t) \end{aligned} \tag{4.28}$$

First we perform a steady-state verification, i.e. we take $\omega_\tau = 0$. Table II shows the choice of constants in Eq. (4.28) for the verification procedure. The connected model of the bi-unit square is a 4×4 uniform finite element mesh. Having chosen the spatial discretization, we now systematically increase the p-levels of the element approximation functions and expect the error measures to decay exponentially fast as the p-level is increased. In a logarithmic-linear scale the expected rate of convergence would appear as a straight line.

Table II. Numerical values of the constants used in the manufactured solution for the compressible Navier-Stokes equations.

	ρ	u	v	T
a	1.0	1.0	0.0	1.0
b	0.2	0.4	0.5	0.5
ω	π	π	2π	2π
ω_τ	0.0	0.0	0.0	0.0

As noted in Chapter III, a suitable error measure is the L_2 least-squares functional ($\mathcal{J}$). Convergence of this measure to zero implies that the L_2 norm of the residuals of the governing equations are going to zero, i.e., conservation of mass, momentum, and energy are being satisfied. An equally important error measure is the L_2 norm of the

difference between the numerical solution and the analytic solution. Convergence of this measure to zero implies that the numerical solution approaches the exact solution.

The stationary viscous compressible Navier-Stokes equations in the velocity / temperature gradient first-order form were discretized using the least-squares finite element formulation with *nodal* expansions (spectral basis). The exact solution was used to compute Dirichlet velocity and temperature boundary conditions on Γ and density was specified only at the "inflow" (left) boundary. The resulting discrete system was linearized using Newton's method. The initial guess vector for the start-up of the non-linear iteration procedure was unity for the density and temperature fields and zero for all other degrees of freedom. Non-linear convergence was declared when the normalized norm of the residual in density, velocities, and temperature, $\|\Delta\mathcal{U}^{hp}\|/\|\mathcal{U}^{hp}\|$, was less than 10^{-4}; which for the aforementioned initial guess required eight Newton iterations. In Fig. 44 we plot the L_2 least-squares functional and L_2 error of the density, velocity, velocity gradient, temperature, and temperature gradient fields as a function of the expansion order in a logarithmic-linear scale. Exponentially fast decay (spectral convergence) of the L_2 least-squares functional and L_2 error is observed.

Next we consider the non-stationary case, i.e. $\omega_\tau \neq 0$. The choice of constants in Eq. (4.28) remain the same as shown in Table II, but now we take $\omega_\tau = 2\pi$. The connected model in space-time, consists of a 4×4 uniform finite element mesh in space and a single element layer in time.

The non-stationary viscous compressible Navier-Stokes equations in the velocity / temperature gradient first-order form were discretized using the space-time coupled least-squares finite element formulation with *nodal* expansions in space-time. We take $\Delta t = t_{s+1} - t_s = 0.5$ and compute the solution in a single space-time strip for increasing element expansion orders. The exact solution was used to compute

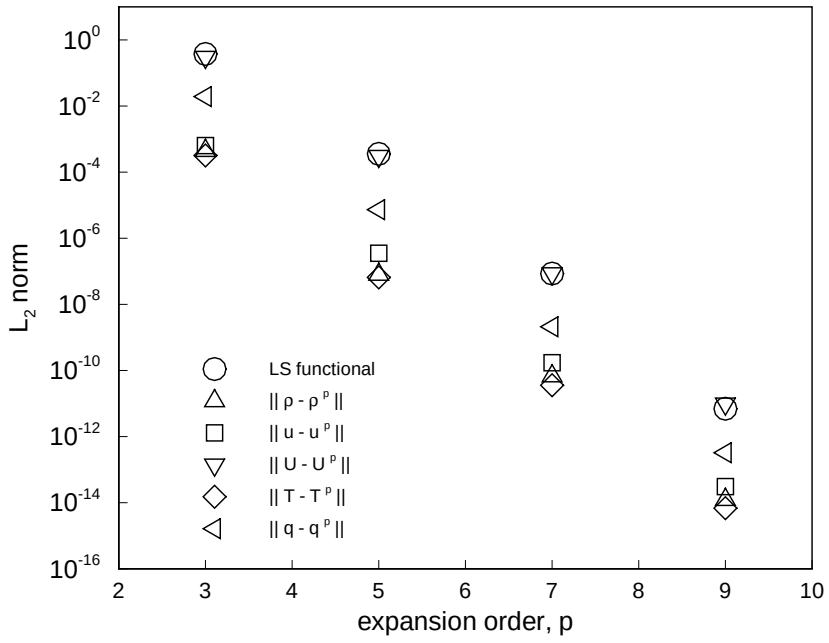

Fig. 44. Decay of the least-squares functional and convergence of the density, velocity, velocity gradient, temperature, and temperature gradient fields to the exact solution in the L_2-norm. Stationary case.

the initial condition and Dirichlet boundary conditions. In Figure 45 we plot the L_2 least-squares functional and L_2 error of the density, velocity, velocity gradient, temperature, and temperature gradient fields as a function of the expansion order in a logarithmic-linear scale. Like for the stationary case, exponentially fast decay of the L_2 least-squares functional and L_2 error is observed (for this case in space-time).

Implicitly, Fig. 45 also shows the unconditional stability of the space-time coupled formulation. In a space-time decoupled formulation as the order of the time approximation is increased the stability region for the time approximation decreases and a smaller time increment must be used to remain stable. Here we have continuously increased the order of the time approximation and kept $\Delta t = t_{s+1} - t_s$ constant at 0.5.

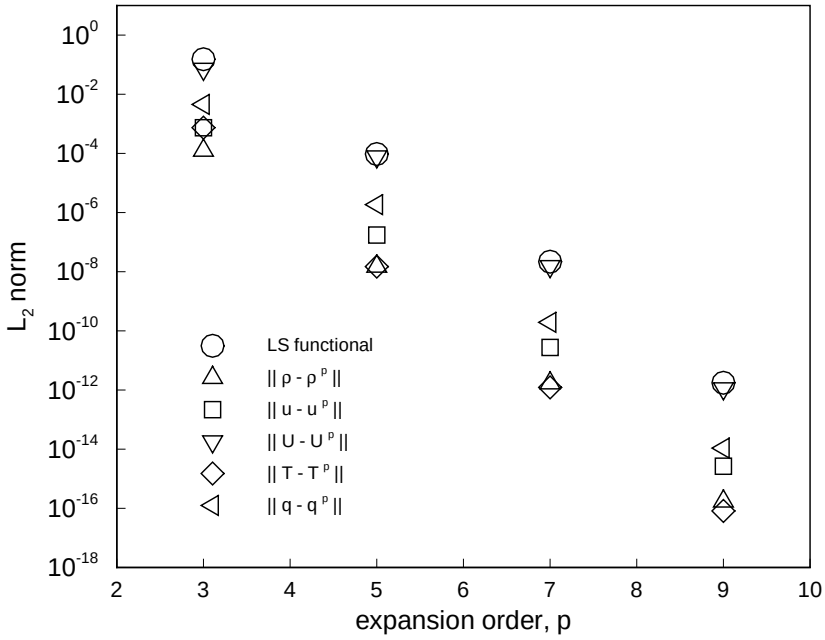

Fig. 45. Decay of the space-time least-squares functional and convergence of the density, velocity, velocity gradient, temperature, and temperature gradient space-time fields to the exact solution in the L_2-norm. Non-stationary case.

2. Flow past a circular cylinder

We consider the two-dimensional viscous flow of a compressible ideal gas past a circular cylinder. The interest of this problem is in the periodic flow pattern that develops when the flow remains subsonic and the free-stream Reynolds number is greater than a critical Reynolds number, e.g., $\mathrm{Re}_c \approx 46.2$ [50, 29] for incompressible flow. The problem is ideal to test the open boundary conditions as the inflow boundary may be well defined and the outflow boundary condition must be imposed in a region where the vortex street is fully developed.

In this study we consider two-dimensional, unsteady simulations for flow conditions with a free-stream Reynolds number of 100 only, for which the characteristic flow parameters should closely approximate values observed experimentally, i.e., three-dimensional effects are not yet dominant. Keeping the free-stream Reynolds number fixed, we consider a range of free-stream Mach numbers and thermal loadings ($T_{\mathrm{cyl}}/T_\infty$) of 1.0 and 2.0, with $\mathrm{Pr} = 0.7$ and $\gamma = 1.4$.

As noted in Chapter III for the incompressible flow simulation, the issue of proper placement of the computational boundaries is of great importance, as it may significantly pollute the accuracy of the characteristic flow parameters. For the set of simulations presented here, the cylinder is of unit diameter and is placed in the finite region $\bar{\Omega} = [-15.5, 30.5] \times [-20.5, 20.5]$.

For the numerical simulations we implement a space-time coupled least-squares formulation (see Chapter II). The spatial discretization used is similar to that shown in Fig. 36, except that a finer mesh consisting of 792 elements is used. A single element layer is used to account for the temporal discretization and use nodal expansions with $p_\xi = p_\eta = 4$, $p_\gamma = 3$ in each element, i.e., a spectral basis of isotropic order 4 in space-domain and a spectral basis of order 3 in time-domain. At each Newton step the

linear system of algebraic equations is solved using a matrix-free conjugate gradient algorithm with a Jacobi preconditioner. For the time marching procedure the size of the time step, $\Delta t = t_{s+1} - t_s$, was chosen as $\Delta t = 0.5$.

The boundary conditions at the outflow/inflow boundaries are of the maximally dissipative characteristic type (derived in Section A-1) and are imposed in a weak sense through the least-squares functional. At the lateral boundaries free-stream boundary conditions could be imposed, however we choose a weaker no-flux boundary condition: $\partial u/\partial n = 0$, $v = 0$, $\partial T/\partial n = 0$ imposed strongly. At the cylinder surface we prescribe the no slip boundary condition: $u = 0$, $v = 0$, and a prescribed cylinder temperature $T = T_{\text{cyl}}$. For a well-posed problem, the initial conditions should satisfy the compressible Navier-Stokes equations. To prescribe such a condition, we use as an initial condition a steady-state solution at the corresponding free-stream conditions.

a. Simulation at $M_\infty = 0.2$

We consider isothermal flow conditions ($T_{\text{cyl}}/T_\infty = 1.0$) and a free-stream Mach number of 0.2. At such free-stream Mach number we expect the flow to remain essentially incompressible, and thus the simulation results in a valuable verification/validation exercise as experimental data and extensive numerical simulation data is readily available. In other words, we expect the simulation to yield values for the characteristic flow parameters that are very close to the incompressible case.

To further emphasize the need for boundary conditions that lead to a strongly well-posed problem, we first perform a simulation with the so-called Neumann outflow-type boundary conditions, commonly used in low-order finite-difference/control-volume based simulations:

$$\frac{\partial u}{\partial n} = \frac{\partial v}{\partial n} = \frac{\partial T}{\partial n} = 0 \qquad \text{on } \Gamma_+ \times (0, \tau]\,. \tag{4.29}$$

Figure 46 shows the time history of density contours behind the circular cylinder when (4.29) is used as an outflow boundary condition.

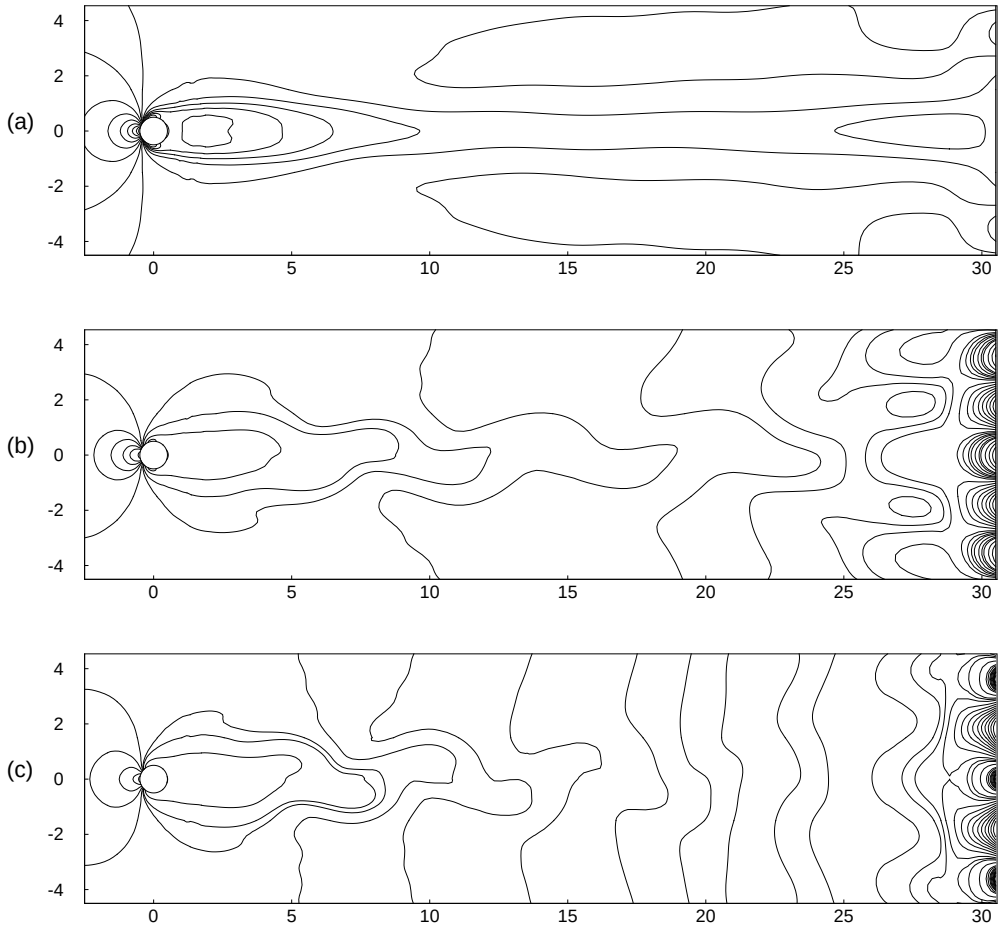

Fig. 46. Time history of density contours behind the circular cylinder for M = 0.2, Re = 100, Pr = 0.7. Outflow boundary conditions are *not* of the maximally dissipative characteristic type, resulting in an ill-posed problem and spurious density predictions which eventually lead to total loss of numerical stability. (a) $t = 50.0$, (b) $t = 70.0$, (c) $t = 72.0$.

From Fig. 46 we see that artificial boundary layers form at the outflow boundary and density waves start to propagate into the domain. Eventually the ill-posed outflow boundary condition causes the numerical simulation to become unstable due

to spurious predictions in density, temperature, and velocity fields. This illustrates the well know fact that high-order methods are less tolerant to ill-posed boundary conditions. Clearly, the same boundary conditions that produce stable (and perhaps erroneous) results in low-order finite-difference/control-volume based simulations are not tolerated in a high-order based simulation.

Having seen the catastrophic results resulting from using an open boundary condition that lead to an ill-posed problem, we now run a simulation with the maximally dissipative characteristic-based open boundary conditions derived in Section A-1 and imposed in a weak sense through the least-squares functional. Figure 47 shows the time history of density contours behind the circular cylinder when the maximally dissipative characteristic-based open boundary conditions are used. No boundary layers are present and no wave propagation into the domain occurs; the simulation carries on smoothly with no spurious predictions in density, temperature, or velocity fields.

As expected, the flow remains essentially incompressible: as shown in Fig. 47, where the density contours shown are in the range $[0.96, 1.03]$. Figure 48a shows instantaneous vorticity contours ar $t = 175.0$, at which time the vortex shedding cycle is already well established. The shedding period, to within 0.05 time units, is found to be $\mathcal{T} = 6.05$; which gives a dimensionless shedding frequency of $\mathrm{St} = 0.1653$. The predicted average drag coefficient is $\bar{C}_D = 1.375$ and the amplitude of the lift coefficient is $C_L = \pm 0.326$. The characteristic flow parameters are in good agreement with experimental measurements and with our incompressible numerical results (see Chapter III).

Figure 48b shows instantaneous local Mach number contours ar $t = 175.0$. Even though the flow may be regarded as essentially incompressible, we realize that compressibility effects are still present as the local Mach number reaches a maximum value of 0.27; a borderline value for the flow to still be regarded as incompressible.

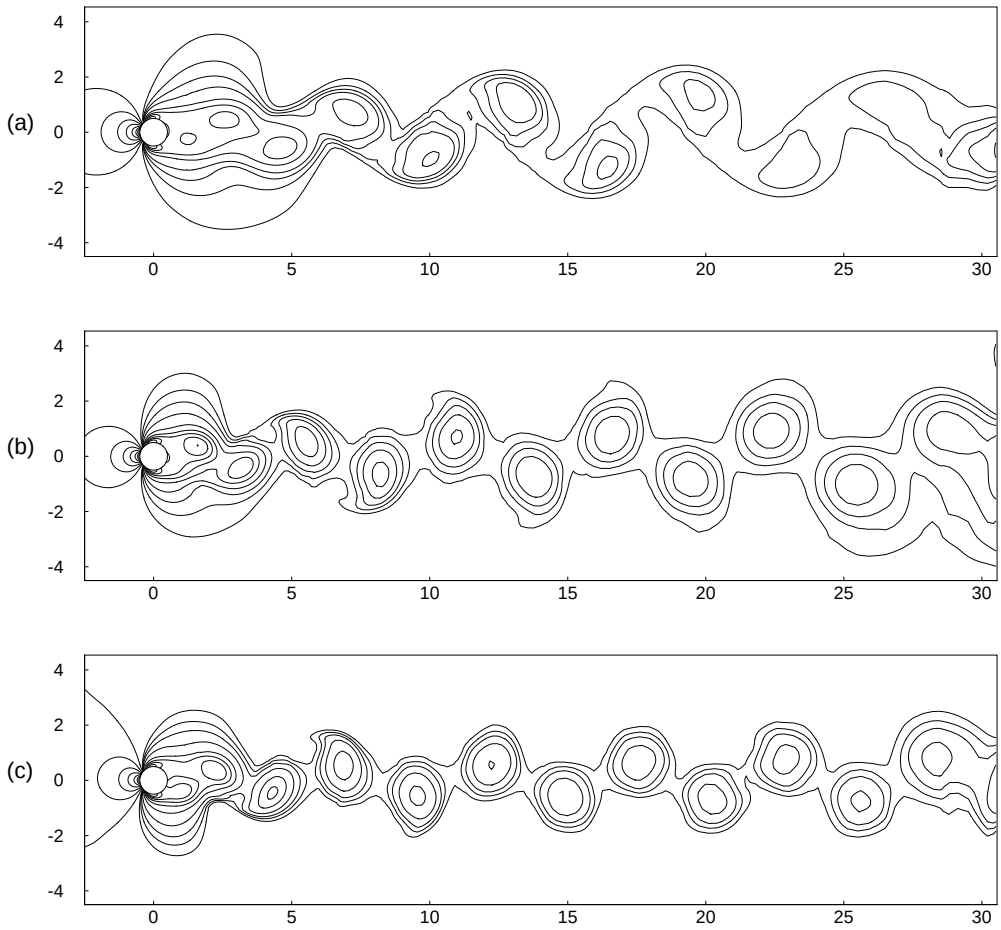

Fig. 47. Time history of density contours behind the circular cylinder for M = 0.2, Re = 100, Pr = 0.7. Outflow boundary conditions are of the maximally dissipative characteristic type, resulting in a strongly well-posed problem. Density contours are in the range $[0.96, 1.03]$. (a) $t = 75.0$, (b) $t = 100.0$, (c) $t = 150.0$.

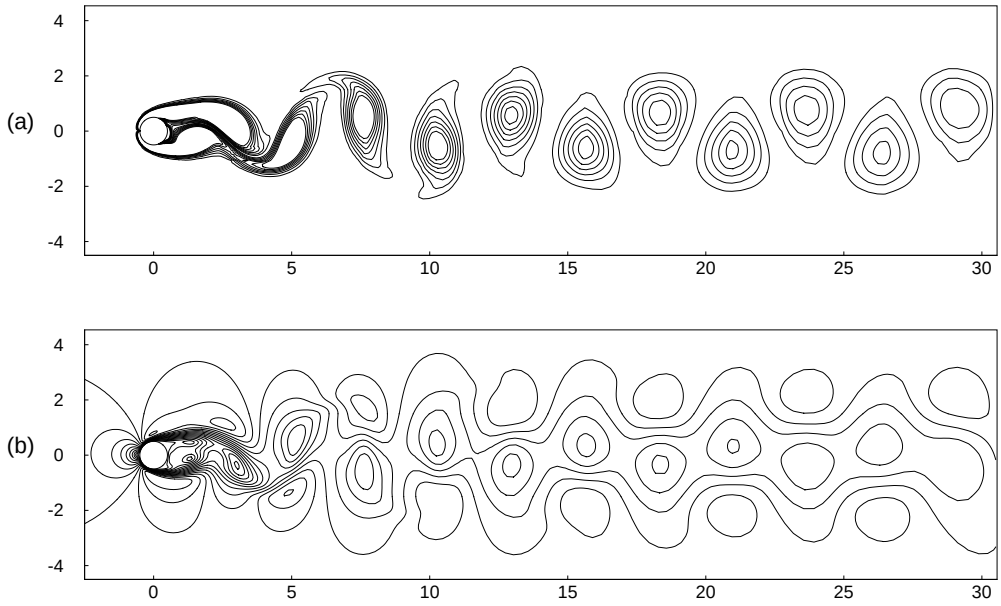

Fig. 48. Instantaneous (a) vorticity and (b) local Mach number contours behind the circular cylinder at $t = 175.0$ (M = 0.2, Re = 100, Pr = 0.7). Local Mach number contours are in the range $[0.0, 0.27]$.

b. Simulation at $M_\infty = 0.5$

We consider flow conditions with $T_{\text{cyl}}/T_\infty = 2.0$ and a free-stream Mach number of 0.5. The cylinder is now heated and the free-stream conditions are well into the subsonic regime. At this Mach number the vortex street still exists, as shown in Fig. 49 where at $t = 200$ the vortex street is already well established. However, unlike the (almost) incompressible flow conditions considered before, we see large density variations in the flow field. These can be attributed to compressibility effects and to high temperature gradients near the cylinder's surface: Fig. 49a shows density contours in the range $[0.40, 1.11]$. In addition, a thermal boundary layer now forms around the cylinder and high temperature gas is shed from the aft of the cylinder (see Fig. 49b).

The shedding period, to within 0.05 time units, is found to be $\mathcal{T} = 6.45$; which gives a dimensionless shedding frequency of St $= 0.1550$. The predicted average drag coefficient is $\bar{C}_D = 1.545$. The correct trends are observed, i.e., the shedding frequency decreases and the average drag coefficient increases (relative to the incompressible case) due to compressibility effects. We note, however, that for the thermal loading considered here the temperature dependence of the dynamic viscosity and thermal conductivity cannot be neglected. An appropriate constitutive relation like Sutherland's law [108] should be used in this simulation. In spite of this modelling issue, we see that the least-squares formulation is able to provide excellent resolution of the flow field and yield (qualitatively)† accurate predictions for the characteristic flow parameters.

†due to lack of experimental or numerical data at this flow conditions

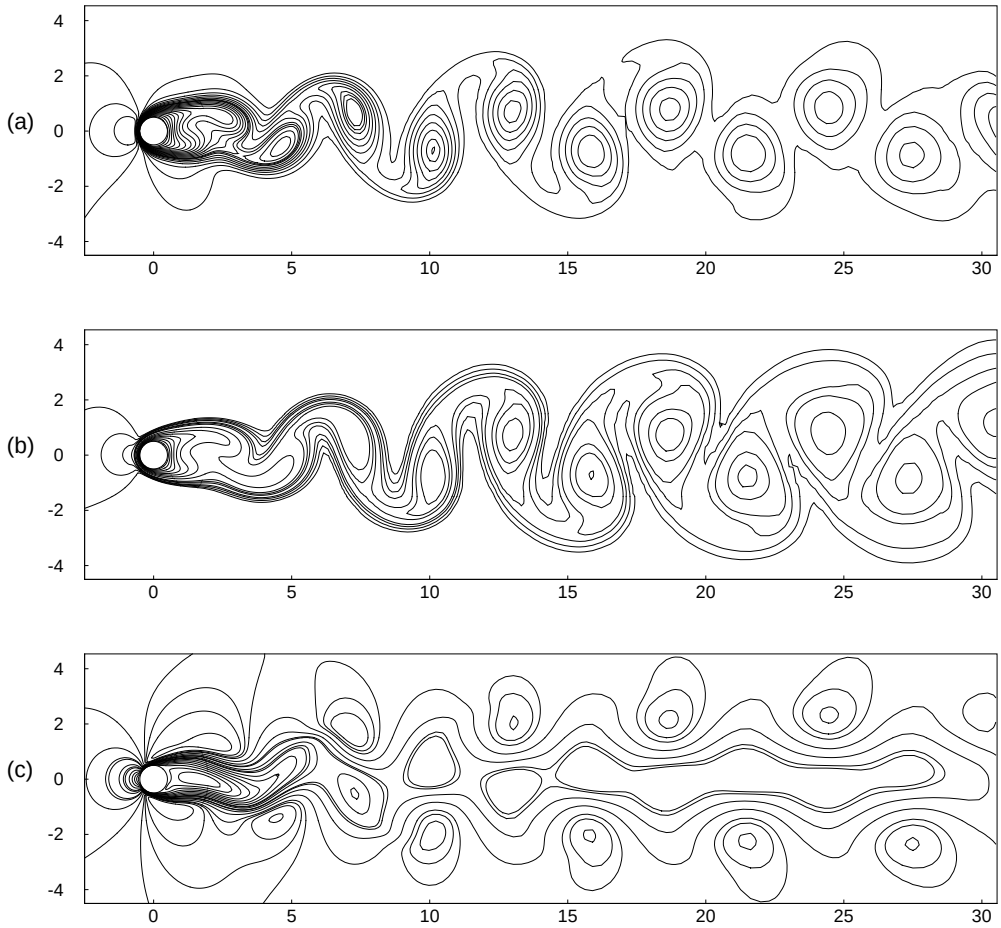

Fig. 49. Instantaneous (a) density, (b) temperature, and (c) u-velocity contours behind the heated circular cylinder at $t = 200.0$ (M $= 0.5$, Re $= 100$, Pr $= 0.7$). Density and temperature contours are in the range $[0.40, 1.11]$, $[0.97, 2.00]$ respectively.

c. Simulation at $M_\infty = 0.7$

We consider isothermal flow conditions ($T_{\text{cyl}}/T_\infty = 1.0$) and a free-stream Mach number of 0.7. As shown in Fig. 50, at this Mach number, the vortex street still exists and is well developed at time $t = 200$. Event though the cylinder is at the free-stream temperature (i.e., the cylinder is not heated) we see density variations in the flow field due to compressibility effects (see Fig. 50a). In addition, we also see temperature variations in the flow field. At this Mach number viscous dissipation is significant on the cylinder surface and acts like a source term in the energy equation, inducing a thermal boundary layer (see Fig. 50b).

The shedding period, to within 0.05 time units, is found to be $\mathcal{T} = 6.30$; which gives a dimensionless shedding frequency of St $= 0.1587$. Our results are in good agreement with the high-order splitting discontinuous Galerkin *hp* numerical results of Lomtev et al. [72], reported as St $= 0.158$. The predicted average drag coefficient is $\bar{C}_D = 1.843$, also in excellent agreement with the simulations of Lomtev et al. [72], who reported a value of $\bar{C}_D = 1.841$.

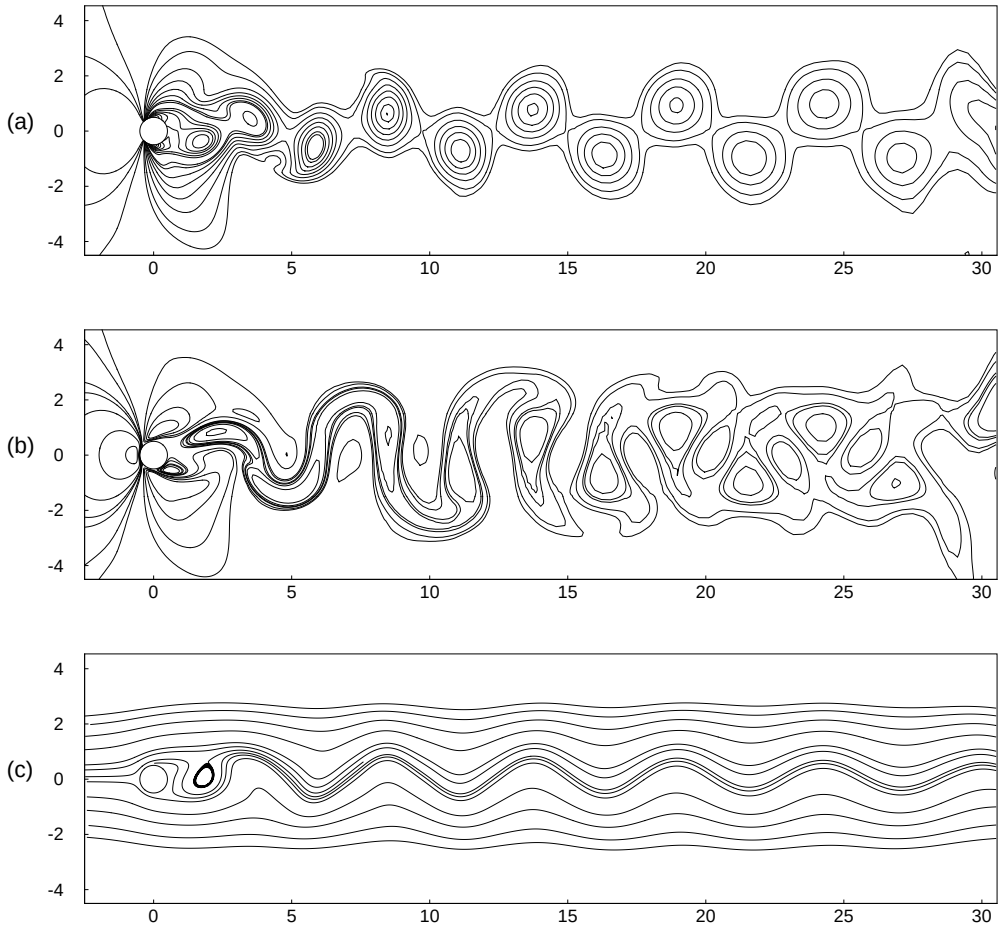

Fig. 50. Instantaneous (a) density contours, (b) temperature contours, and (c) streamlines behind the circular cylinder at $t = 200.0$ (M $= 0.7$, Re $= 100$, Pr $= 0.7$). Density and temperature contours are in the range $[0.54, 1.42]$, $[0.89, 1.09]$ respectively.

d. Simulation at $M_\infty = 2.0$

As a last example, we test the formulation in the supersonic flow regime. To this end, we consider isothermal flow conditions ($T_{\text{cyl}}/T_\infty = 1.0$) and a free-stream Mach number of 2.0. At this free-stream Mach number we expect a shock to appear in front of the cylinder.

At the time of this writing we have not yet developed the capability of adaptive mesh refinement, which could be used for accurate resolution of shocks. Nevertheless, we believe the simulation is valuable to verify whether the formulation is able to predict the formation of shocks.

At this Mach number the vortex street no longer exists and a steady state solution is obtained. Figure 51 shows the steady-state density contours, from which the formation of the shock in front of the cylinder is evident. Clearly, the shock is not well resolved. However, the least-squares formulation is able to predict the shock.

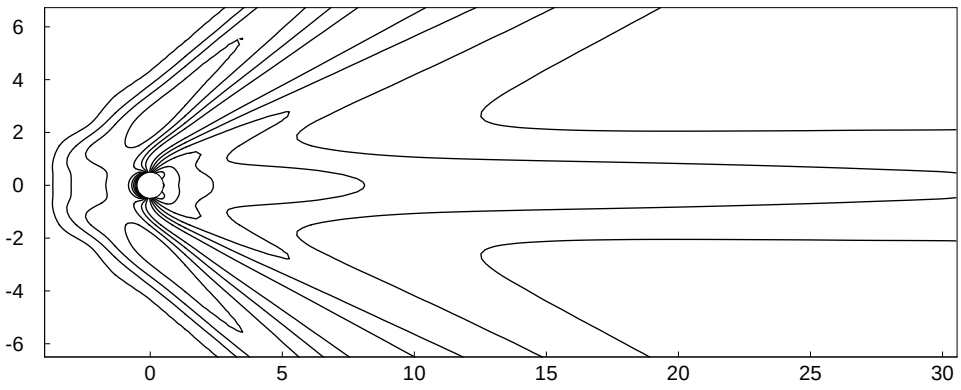

Fig. 51. Density contours for flow past a circular cylinder at supersonic free-stream conditions M = 2.0, Re = 100.

CHAPTER V

BENDING OF THIN AND THICK PLATES

This chapter and the following are concerned with the formulation and application of least-squares finite element models to problems in the field of solid mechanics. In this chapter[†] we consider the bending of thin and thick plates.

Finite element formulations for the bending of plates have been traditionally derived from the principle of virtual displacements or the principle of minimum total potential energy (see Reddy [92, 93, 91]). Equivalently, the traditional finite element formulations can be derived by applying the Ritz-Galerkin procedure (i.e., the weak form Galerkin procedure) to the plate bending equations. Displacement-based finite element models for the classical plate theory (CPT) require the use of C^1 interpolation functions because the essential boundary conditions involve specifying the transverse deflection as well as its derivatives. Displacement-based finite element models for the first-order shear deformation theory (FSDT) allow the use of C^0 interpolation functions for the transverse deflection and the two independent rotations. The displacement-based FSDT finite element model presents computational difficulties when the side-to-thickness ratio of the plate, a/h, is large (i.e., when modeling thin plates). A strong stiffening of the element matrices occurs, resulting in spurious shear stress predictions and erroneous results for the generalized displacements. This phenomenon is known as *shear-locking*. The problem is commonly treated by reduced integration techniques, although other remedies may be found (see, for example, [10, 89]).

[†]The numerical results reported in this chapter appear in the article "Mixed plate bending elements based on least-squares formulation" by J. P. Pontaza and J. N. Reddy, *Int. J. Numer. Meth. Engng*, accepted for publication. Copyright (2003) John Wiley & Sons.

Shear-locking is evident in Ritz-Galerkin displacement-based finite element models with equal order interpolation of all generalized displacements when full integration is used to analyze thin plates ($a/h \geq 20$), and is more pronounced when low-order elements are used. Higher-order elements show less sensitivity to shear-locking but with slower convergence. The Ritz-Galerkin plate elements behave uniformly well for thin and thick plates when reduced or selective reduced integration is used. Highly distorted meshes tend to aggravate the predictive capabilities of these elements, even for high-order elements and/or when reduced integration techniques are used. Shear-locking can also be alleviated by using mixed interpolation, i.e., un-equal order interpolation for the generalized displacements.

An alternate approach is to use mixed formulations, where in addition to the generalized displacements the stress resultants are introduced as independent variables [89]. Other types of mixed formulations are those where the shears strains are treated as independent variables [48, 5]. However, reliable and optimal Ritz-Galerkin mixed formulations require that the finite element approximation spaces satisfy an inf-sup condition [19]. In many cases it is very difficult to analytically show that the inf-sup condition is satisfied for a given combination of finite element spaces. Moreover, discrete models based on mixed formulations result in symmetric but indefinite systems of equations; although static condensation procedures may be invoked to remedy this shortcoming.

In the context of least-squares finite element formulations for plate bending, Jiang [56] presented a formulation based on the CPT where low order nodal expansions were used to develop the discrete finite element model. The resulting least-squares functional is non-equivalent and reduced integration techniques were used to obtain acceptable numerical results (i.e., to obtain a collocation solution). As recommended in Chapter II, blind application of reduced integration techniques will

not result in a collocation solution and should be avoided. Of course, in the work of Jiang [56] the conditions necessary to ensure that a reduced integration solution coincide with a collocation solution were considered and explained in detail.

The proposed *bona fide* formulation uses the generalized displacements and stress-resultants as independent variables for both the CPT and FSDT. The plate bending equations are used in their first-order form to allow the use of C^0 interpolation functions in the finite element model. The use of least-squares principles leads to a variational unconstrained minimization problem where the approximation spaces can be chosen independently, i.e. stability requirements such as the inf-sup condition never arise (see Chapter II). Equal order interpolation and full integration is used to develop the discrete finite element model, which results in a symmetric and positive-definite system of algebraic equations. We show by numerical examples that the proposed high-order plate bending elements are free of shear-locking, insensitive to geometric distortions, and predict accurate displacements as well as stress resultants along plate boundaries.

We present the governing equations for the bending of plates and use the procedures in Chapter II to develop and arrive at the finite element model. The performance of the plate bending elements is tested for circular, rectangular, and skew plates with various boundary conditions and plate thickness. Exponentially fast decay (spectral convergence) of the L_2 least-squares functional and L_2 error norms of the generalized displacements and stress-resultants is verified using a smooth solution to the FSDT plate bending equations.

A. Governing equations

Consider the bending of an isotropic plate of thickness h subjected to a transverse load $q(x,y)$. In this case, $\bar{\Omega}$ is the undeformed mid-plane of the plate. The problem can be stated as:

Find the generalized displacement fields $w(x,y)$, $\phi_x(x,y)$, $\phi_y(x,y)$, and the thickness-averaged stress resultants $M_{xx}(x,y)$, $M_{yy}(x,y)$, $M_{xy}(x,y)$, $Q_x(x,y)$, $Q_y(x,y)$, such that

$$\frac{\partial Q_x}{\partial x} + \frac{\partial Q_y}{\partial y} + q(x,y) = 0 \qquad \text{in } \Omega \tag{5.1}$$

$$\frac{\partial M_{xx}}{\partial x} + \frac{\partial M_{xy}}{\partial y} - Q_x = 0 \qquad \text{in } \Omega \tag{5.2}$$

$$\frac{\partial M_{xy}}{\partial x} + \frac{\partial M_{yy}}{\partial y} - Q_y = 0 \qquad \text{in } \Omega \tag{5.3}$$

$$M_{xx} - D\left(\frac{\partial \phi_x}{\partial x} + \nu \frac{\partial \phi_y}{\partial y}\right) = 0 \qquad \text{in } \Omega \tag{5.4}$$

$$M_{yy} - D\left(\nu \frac{\partial \phi_x}{\partial x} + \frac{\partial \phi_y}{\partial y}\right) = 0 \qquad \text{in } \Omega \tag{5.5}$$

$$M_{xy} - D\frac{(1-\nu)}{2}\left(\frac{\partial \phi_x}{\partial y} + \frac{\partial \phi_y}{\partial x}\right) = 0 \qquad \text{in } \Omega \tag{5.6}$$

$$\alpha\, Q_x - \beta\left(\frac{\partial w}{\partial x} + \phi_x\right) = 0 \qquad \text{in } \Omega \tag{5.7}$$

$$\alpha\, Q_y - \beta\left(\frac{\partial w}{\partial y} + \phi_y\right) = 0 \qquad \text{in } \Omega \tag{5.8}$$

Equations (5.1)-(5.3) are the equilibrium equations and Eqs. (5.4)-(5.8) are generalized displacements-stress resultants relationships (see Reddy [92] for further details). In Eqs. (5.4)-(5.6) D is the plate flexural rigidity

$$D = \frac{E\, h^3}{12\,(1-\nu^2)}\,,$$

E is the Young's modulus and ν the Poisson's ratio.

For $\alpha = 0$ and $\beta = 1$ we have the classical plate theory and Eqs. (5.7) and (5.8) simply become definitions for the rotations in terms of the transverse deflection w. For $\alpha = 1$ and $\beta = K_s\, G\, h$ we recover the first-order shear deformation plate theory and Eqs. (5.7) and (5.8) become relationships for the stress resultants in terms of the two independent rotations and the transverse deflection. For this definition of β, G is the shear modulus and $K_s = 5/6$ is the shear correction factor.

Appropriate boundary conditions for the support types considered here are given in the following

$$w = \phi_n = \phi_s = 0 \qquad \text{on } \Gamma_C \tag{5.9}$$

$$w = \phi_s = M_{nn} = 0 \qquad \text{on } \Gamma_{SS} \tag{5.10}$$

$$M_{nn} = M_{ns} = Q_n = 0 \qquad \text{on } \Gamma_F \tag{5.11}$$

where $\Gamma = \Gamma_C \cup \Gamma_{SS} \cup \Gamma_F$ and $\Gamma_C \cap \Gamma_{SS} \cap \Gamma_F = \emptyset$. In Eqs. (5.9)-(5.11) Γ_C is a clamped boundary, Γ_{SS} is a simply-supported boundary, and Γ_F is a free boundary.

Keeping in mind the "C^0 practicality" level, we realize that the equations governing the bending of plates, Eqs. (5.1)-(5.8), are already of first-order. Hence, we need only construct the L_2 least-squares functional and minimize it with respect to the chosen approximating spaces to obtain the finite element model.

Additional smoothness for the stress resultants may be requested for the CPT model ($\alpha = 0$ and $\beta = 1$) by including the following four seemingly redundant constraints

$$\frac{\partial \phi_y}{\partial x} - \frac{\partial \phi_x}{\partial y} = 0 \qquad \text{in } \Omega \tag{5.12}$$

$$\frac{\partial Q_y}{\partial x} - \frac{\partial Q_x}{\partial y} = 0 \qquad \text{in } \Omega \tag{5.13}$$

$$\frac{\partial}{\partial y}\left(M_{xx}-\nu\, M_{yy}\right)-\left(1+\nu\right)\frac{\partial M_{xy}}{\partial x}=0 \qquad \text{in } \Omega \tag{5.14}$$

$$\frac{\partial}{\partial x}\left(M_{yy}-\nu\, M_{xx}\right)-\left(1+\nu\right)\frac{\partial M_{xy}}{\partial y}=0 \qquad \text{in } \Omega \tag{5.15}$$

which implicitly enforce the following compatibility relations for the deflection w

$$\frac{\partial}{\partial y}\left(\frac{\partial w}{\partial x}\right)=\frac{\partial}{\partial x}\left(\frac{\partial w}{\partial y}\right)$$

$$\frac{\partial}{\partial y}\left(\frac{\partial^2 w}{\partial x \partial y}\right)=\frac{\partial}{\partial x}\left(\frac{\partial^2 w}{\partial y^2}\right)\,,\;\frac{\partial}{\partial y}\left(\frac{\partial^2 w}{\partial x^2}\right)=\frac{\partial}{\partial x}\left(\frac{\partial^2 w}{\partial x \partial y}\right)$$

$$\frac{\partial^2}{\partial y \partial x}\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}\right)=\frac{\partial^2}{\partial x \partial y}\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}\right)$$

Similarly, additional smoothness for the stress resultants may be requested for the FSDT model ($\alpha = 1$ and $\beta = K_s\, G\, h$) at the expense of introducing an additional degrees of freedom: the transverse twist (Λ) defined as

$$\Lambda - D\,\frac{(1-\nu)}{2}\left(\frac{\partial \phi_y}{\partial x}-\frac{\partial \phi_x}{\partial y}\right)=0 \qquad \text{in } \Omega \tag{5.16}$$

and including the following three seemingly redundant constraints

$$\frac{\partial Q_y}{\partial x}-\frac{\partial Q_x}{\partial y}-K_s\, G\, h\,\frac{2\,\Lambda}{D\,(1-\nu)}=0 \qquad \text{in } \Omega \tag{5.17}$$

$$\frac{\partial}{\partial y}\left(M_{xx}-\nu\, M_{yy}\right)-\left(1+\nu\right)\frac{\partial}{\partial x}\left(M_{xy}-\Lambda\right)=0 \qquad \text{in } \Omega \tag{5.18}$$

$$\frac{\partial}{\partial x}\left(M_{yy}-\nu\, M_{xx}\right)-\left(1+\nu\right)\frac{\partial}{\partial y}\left(M_{xy}+\Lambda\right)=0 \qquad \text{in } \Omega \tag{5.19}$$

which implicitly enforce the following compatibility relations for the deflection w and independent rotations ϕ_x and ϕ_y

$$\frac{\partial}{\partial y}\left(\frac{\partial w}{\partial x}\right)=\frac{\partial}{\partial x}\left(\frac{\partial w}{\partial y}\right)$$

$$\frac{\partial}{\partial y}\left(\frac{\partial \phi_x}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial \phi_x}{\partial y}\right), \frac{\partial}{\partial y}\left(\frac{\partial \phi_y}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial \phi_y}{\partial y}\right)$$

The transverse twist is a very important and physically meaningful quantity as it governs edge zone-behavior of rotations and stress resultants for the FSDT [42].

The additional compatibility constraints for the CPT and FSDT models make the dependencies between the stress resultants and generalized displacements sub-dominant, i.e., it helps the formulation "forget" that the stress resultants are linear combinations of the derivatives of the generalized displacements. The weakening of the interdependencies may help retain full-coercivity of the system, resulting in optimal convergence rates for the stress resultants, as given by error estimates based on standard elliptic theory (see [13] for further details). Neglecting the compatibility constraints may result in sub-optimal convergence rates for the stress resultants but will not lead to disastrous results. The effect of neglecting the compatibility constraints may be especially pronounced when low p-levels (e.g., p-levels of 1 or 2) are used to construct the finite element model, and the effect becomes less pronounced when using higher-order element expansions. Since the inclusion of the compatibility constraints does not represent a significant increase in cost, except perhaps for the FSDT model where an additional degree of freedom is introduced, we have included them in our implementation.

B. Numerical examples: verification benchmarks

In this section we present some numerical results obtained with the proposed formulation. The problems have been selected to assess the predictive capabilities of the plate bending elements.

1. Convergence

To illustrate the performance of the elements and verify exponentially fast decay of discretization errors for increasing p-levels we consider the following test problem. We prescribe a solution for the generalized displacement fields (for the FSDT) in the bi-unit square, $\bar{\Omega} = [-1,1] \times [-1,1]$, compute the consistent forcing functions and solve to verify convergence of the numerical solution to the exact solution. We choose the generalized displacement fields to be of the form

$$\begin{aligned} w(x,y) &= a_w + b_w \cos(\omega_w x) \sin(\omega_w y) \\ \phi_x(x,y) &= a_{\phi_x} + b_{\phi_x} \cos(\omega_{\phi_x} x) \cos(\omega_{\phi_x} y) \\ \phi_y(x,y) &= a_{\phi_y} + b_{\phi_y} \sin(\omega_{\phi_y} y) \sin(\omega_{\phi_y} y) \end{aligned} \tag{5.20}$$

The exact solution was used to prescribe values of the generalized displacements on the entire boundary. The connected model, $\bar{\Omega}^h$, consists of a 4×4 finite element mesh. We consider uniform and distorted meshes. The choice of constants in Eq. (5.20) were as follows: $a_w = a_{\phi_x} = a_{\phi_y} = 0$, $\omega_w = \omega_{\phi_x} = \omega_{\phi_y} = 2\pi$, $b_w = 0.5$, $b_{\phi_x} = 0.2$, and $b_{\phi_y} = 0.4$.

For the case considered here, namely $\omega_w = \omega_{\phi_x} = \omega_{\phi_y} = \omega$, the generalized displacement field in Eq. (5.20) is a prototypical response of a square plate subjected to a trigonometric type transverse load with spatial frequency ω.

In Figure 52 we plot the L_2 least-squares functional ($\mathcal{J}$) and L_2 error of the generalized displacements $(w, \boldsymbol{\phi})$ and stress resultants $(\mathbf{M}, \mathbf{Q})$ as a function of the expansion order in a logarithmic-linear scale. Exponentially fast decay (spectral convergence) of the L_2 least-squares functional and L_2 errors is observed for both uniform and distorted meshes.

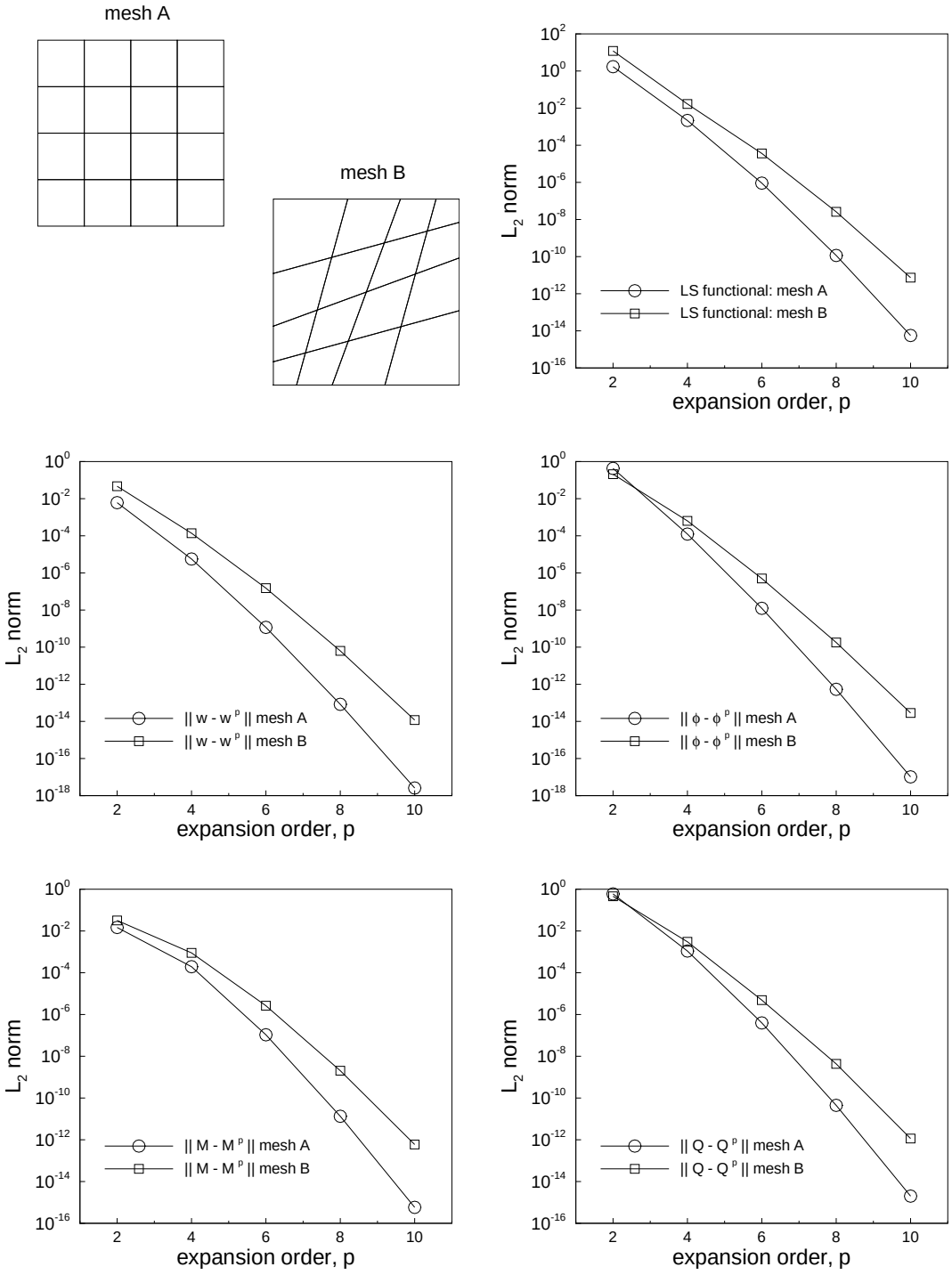

Fig. 52. Decay of the least-squares functional and convergence of the generalized displacements and stress resultants for uniform and distorted meshes.

2. Circular plates

We consider the bending of circular plates. The total domain of the plate is $\bar{\Omega} \times [-h/2, h/2]$, where h is the thickness of the plate. The undeformed mid-plane of the plate, $\bar{\Omega}$, has dimensions $[0, a] \times [0, 2\pi]$ and is shown in Fig. 53. The labeled open circles in Fig. 53 show locations at which deflections, stress resultants, and stresses will be tabulated for the different boundary conditions and used to compare against closed form analytic solutions for both the CPT and FSDT (see Reddy [92]).

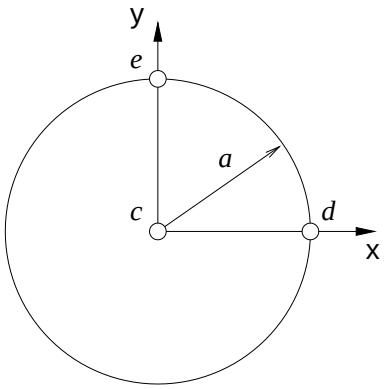

Fig. 53. Circular plate showing points and respectively assigned labels where displacement, stress resultants, and stresses are recorded and tabulated.

Unless otherwise stated stresses are reported at the top fibers of the plate, i.e. at $z = h/2$. The reported deflections (w), stress resultants ($\mathbf{M}, \mathbf{Q}$), and stresses ($\boldsymbol{\sigma}$) are normalized as follows

$$\bar{w} = w \, \frac{D}{q_0 \, a^4} \times 10^3, \quad \bar{\mathbf{M}} = \mathbf{M} \, \frac{100}{q_0 \, a^2}, \quad \bar{\mathbf{Q}} = \mathbf{Q} \, \frac{100}{q_0 \, a} \tag{5.21}$$

$$\begin{bmatrix} \bar{\sigma}_{xx} & \bar{\sigma}_{yy} & \bar{\sigma}_{xy} \end{bmatrix}^{\mathrm{T}} = \frac{10}{q_0 \, (a/h)^2} \begin{bmatrix} \sigma_{xx} & \sigma_{yy} & \sigma_{xy} \end{bmatrix}^{\mathrm{T}} \tag{5.22}$$

$$\begin{bmatrix} \bar{\sigma}_{yz} & \bar{\sigma}_{xz} \end{bmatrix}^{\mathrm{T}} = \frac{10}{q_0 \, (a/h)} \begin{bmatrix} \sigma_{yz} & \sigma_{xz} \end{bmatrix}^{\mathrm{T}} \tag{5.23}$$

a. Simply supported circular plate

Consider a simply supported, circular isotropic plate subjected to a uniformly distributed load of intensity q_0. Due to the symmetry only one quadrant of the plate need be modelled. The boundary conditions of the computational domain are specified as

$$\phi_x = 0\,, \quad M_{xy} = 0\,, \quad Q_x = 0 \qquad \text{on } x = 0\,,\ y = [0, a]$$

$$\phi_y = 0\,, \quad M_{xy} = 0\,, \quad Q_y = 0 \qquad \text{on } y = 0\,,\ x = [0, a]$$

$$w = 0\,, \quad \phi_s = 0\,, \quad M_{nn} = 0 \qquad \text{on } r = a\,,\ \theta = [0, \pi/2] \quad \text{CPT}$$

$$w = 0\,, \quad M_{nn} = 0\,, \quad M_{ns} = 0 \qquad \text{on } r = a\,,\ \theta = [0, \pi/2] \quad \text{FSDT}$$

The connected models of the quarter plate, $\bar{\Omega}^h$, used for the analysis are shown in Fig. 54. For the analysis of the simply supported plate we wish to fix the p-level at 4 and study the effect of h-refinement with meshes A and B in Fig. 54. The circular arc at $r = a$ is represented using the same nodal basis with a p-level of 4 in each element.

Table III shows results for normalized deflection, stress resultants, and stresses for the case $\nu = 0.30$. The left-superscript on each normalized variable in Table III denotes the location at which the parameter of interest was recorded (see Fig. 53). We observe excellent agreement between the numerical results and the closed form solution for the shear-deformable model at moderate and high diameter-to-thickness ratios and for the classical plate model. Mesh A gives sufficiently accurate results for moderate diameter-to-thickness ratios and for the CPT. For the case $2a/h = 100$ mesh B gives an improved solution; although the results obtained for mesh A are relatively good for such a coarse mesh and low p-level.

It is known that for the traditional (i.e., based on the Ritz-Galerkin procedure)

Table III. Normalized deflection, stress resultants, and stresses for a simply supported, isotropic ($\nu = 0.30$) circular plate under a uniformly distributed load. FSDT model ($K_s = 5/6$). p-level fixed at 4.

$2a/h$	mesh	$^{c}\bar{w}$	$^{c}\bar{M}_{xx}$	$^{d}\bar{M}_{yy}$	$^{e}\bar{Q}_{y}$	$^{c}\bar{\sigma}_{xx}$	$^{d}\bar{\sigma}_{yy}$	$^{e}\bar{\sigma}_{yz}$
5	A	75.1235	-20.6240	-8.7499	50.0179	-12.3744	-5.2499	6.0022
	B	75.1301	-20.6250	-8.7500	50.0000	-12.3750	-5.2500	6.0000
	Analytic	75.1305	-20.6250	-8.7500	50.0000	-12.3750	-5.2500	6.0000
20	A	64.3847	-20.6226	-8.7459	50.0224	-12.3735	-5.2475	6.0027
	B	64.4156	-20.6250	-8.7496	50.0000	-12.3750	-5.2498	6.0000
	Analytic	64.4162	-20.6250	-8.7500	50.0000	-12.3750	-5.2500	6.0000
100	A	61.9073	-20.5131	-8.6992	50.2649	-12.3079	-5.2195	6.0318
	B	63.6478	-20.6200	-8.7473	50.0567	-12.3719	-5.2484	6.0068
	Analytic	63.7305	-20.6250	-8.7500	50.0000	-12.3750	-5.2500	6.0000
CPT	A	63.6956	-20.6240	-8.7497	50.0005	-12.3744	-5.2498	0.0000
	B	63.7015	-20.6250	-8.7500	50.0000	-12.3750	-5.2500	0.0000
	Analytic	63.7019	-20.6250	-8.7500	50.0000	-12.3750	-5.2500	0.0000

displacement-based finite element formulations, even high-order elements still do not display a good predictive capability, particularly when the elements are geometrically distorted and used for stress predictions. Of particular interest is the transverse shear stress prediction, as it is often the most difficult stress component to predict for Ritz-Galerkin FSDT plate elements. The circular plate problem is ideal to asses the accuracy of the prediction as the analytical solution is readily available in closed form, no boundary layers are present, and the elements are naturally distorted to accommodate the plate geometry.

In Figs. 55 and 56 we present the predicted transverse shear stress distributions for a plate with diameter-to-thickness ratio of 100 using mesh A with a p-level of 4. Figure 55 shows the distributions along the radial line $\overline{ce}$ and Fig. 56 along the circular arc $\overline{de}$, with θ measured counterclockwise from the x-axis. We plot the stresses at the elements nodal mid- and end-points along radial line $\overline{ce}$ and along circular arc $\overline{de}$. For comparison, we also present results obtained using the displacement based

Ritz-Galerkin formulation using the same mesh and p-level. We do not employ stress smoothing to present the results, i.e. there may be different values for the shear stress where elements meet. Clearly, the stress predictions using the proposed least-squares formulation are superior.

In Fig. 57 we present contour plots of the predicted shear force $\bar{Q}_x$ for a plate with diameter-to-thickness ratio of 100 using mesh A with a p-level of 4. The Ritz-Galerkin FSDT plate elements give spurious shear force predictions, displaying a localized locking behavior near the curved boundary. On the other hand, the least-squares FSDT plate elements give a smooth and highly accurate shear force prediction throughout.

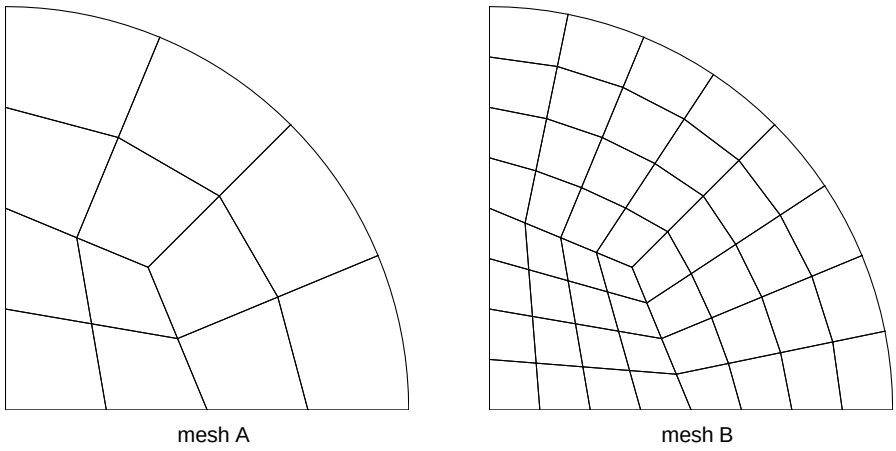

Fig. 54. Quarter plate computational domain for the analysis of the circular plate.

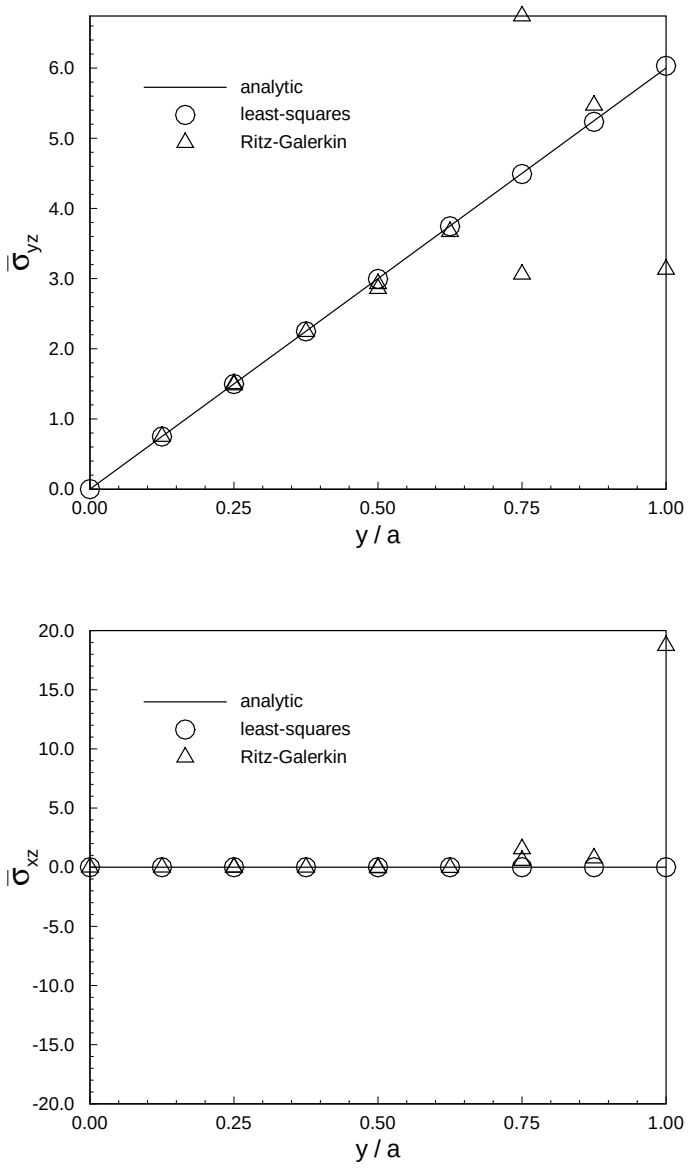

Fig. 55. Transverse shear stress predictions along radial line $\overline{ce}$. Simply supported circular plate with diameter-to-thickness ratio 100. Mesh A with p-level fixed at 4 for both least-squares and Ritz-Galerkin elements.

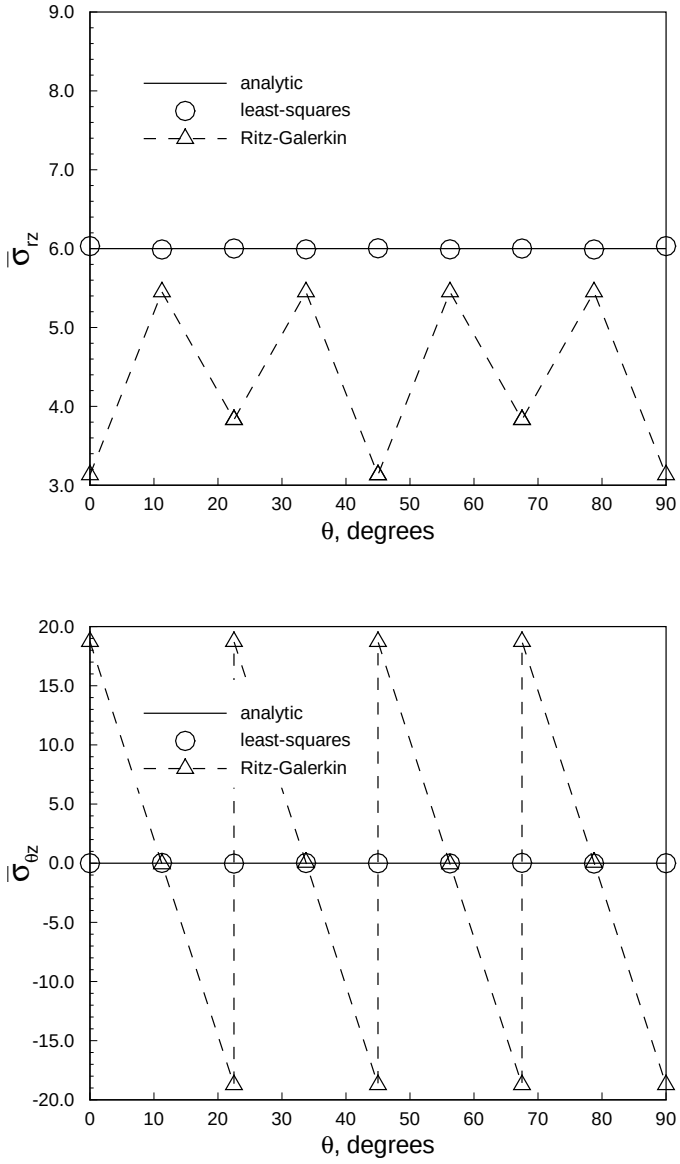

Fig. 56. Transverse shear stress predictions along circular arc $\overline{de}$. Simply supported circular plate with diameter-to-thickness ratio 100. Mesh A with p-level fixed at 4 for both least-squares and Ritz-Galerkin elements.

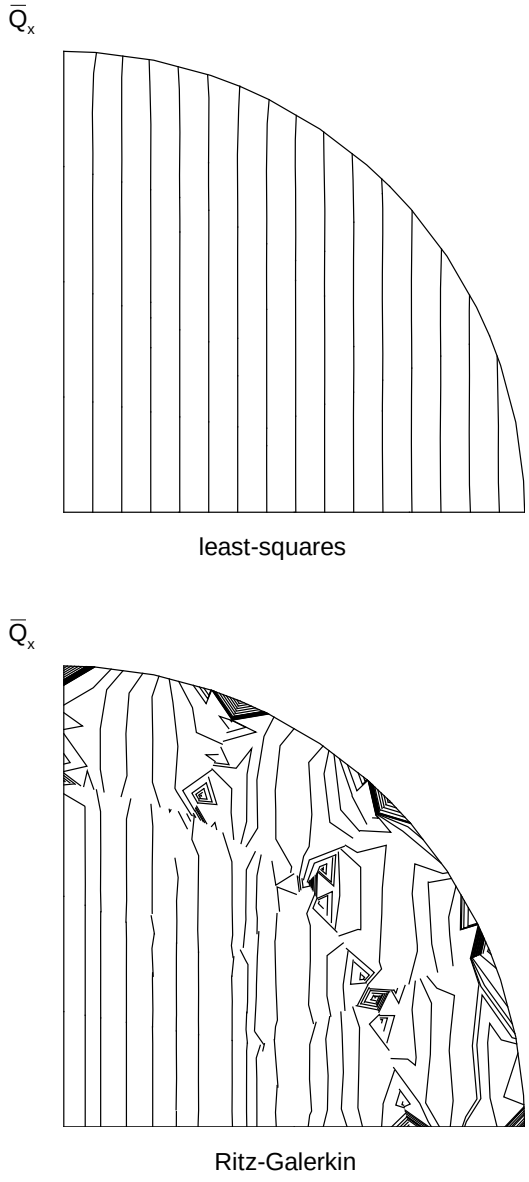

Fig. 57. Predicted shear force contours for a simply supported circular plate with diameter-to-thickness ratio 100. Mesh A with p-level fixed at 4 for both least-squares and Ritz-Galerkin elements.

b. Clamped circular plate

Consider a clamped, circular isotropic plate subjected to a uniformly distributed load of intensity q_0. Due to the symmetry only one quadrant of the plate need be modelled. The boundary conditions of the computational domain are specified as

$$\phi_x = 0\,, \quad M_{xy} = 0\,, \quad Q_x = 0 \qquad \text{on } x = 0\,,\ y = [0, a]$$

$$\phi_y = 0\,, \quad M_{xy} = 0\,, \quad Q_y = 0 \qquad \text{on } y = 0\,,\ x = [0, a]$$

$$w = 0\,, \quad \phi_n = 0\,, \quad \phi_s = 0 \qquad \text{on } r = a\,,\ \theta = [0, \pi/2] \quad \text{CPT}$$

$$w = 0\,, \quad \phi_n = 0\,, \quad M_{ns} = 0 \qquad \text{on } r = a\,,\ \theta = [0, \pi/2] \quad \text{FSDT}$$

Like for the simply supported case, we wish to fix the p-level at 4 and study the effect of h-refinement with meshes A and B in Fig. 54. Table IV shows results for normalized deflection, stress resultants, and stresses for the case $\nu = 0.30$. We observe excellent agreement between the numerical results and the closed form solution for the shear-deformable model and for the classical plate model.

In Fig. 58 we present the predicted transverse shear stress distributions along radial line $\overline{ce}$ for a plate with diameter-to-thickness ratio of 100 using mesh A with a p-level of 4. For the clamped case, the Ritz-Galerkin FSDT elements behave uniformly better when compared to the simply supported case. Nevertheless, the least-squares FSDT elements still give a more accurate prediction of the transverse shear stress distributions.

Table IV. Normalized deflection, stress resultants, and stresses for a clamped, isotropic ($\nu = 0.30$) circular plate under a uniformly distributed load. FSDT model ($K_s = 5/6$). p-level fixed at 4.

$2a/h$	mesh	${}^{c}\bar{w}$	${}^{c}\bar{M}_{xx}$	${}^{d}\bar{M}_{xx}$	${}^{d}\bar{M}_{yy}$	${}^{e}\bar{Q}_{y}$	${}^{d}\bar{\sigma}_{xx}$	${}^{d}\bar{\sigma}_{yy}$	${}^{e}\bar{\sigma}_{yz}$
5	A	27.0515	-8.1246	12.5026	3.7505	50.0121	7.5015	2.2503	6.0015
	B	27.0534	-8.1250	12.5000	3.7500	50.0006	7.5000	2.2500	6.0000
	Analytic	27.0536	-8.1250	12.5000	3.7500	50.0000	7.5000	2.2500	6.0000
20	A	16.3376	-8.1246	12.4977	3.7502	49.9253	7.4986	2.2501	5.9910
	B	16.3392	-8.1250	12.5000	3.7500	49.9956	7.5000	2.2500	6.0000
	Analytic	16.3393	-8.1250	12.5000	3.7500	50.0000	7.5000	2.2500	6.0000
100	A	15.6488	-8.1245	12.4554	3.7107	48.6536	7.4733	2.2264	5.8384
	B	15.6535	-8.1250	12.4967	3.7472	49.8244	7.4980	2.2483	5.9789
	Analytic	15.6536	-8.1250	12.5000	3.7500	50.0000	7.5000	2.2500	6.0000
CPT	A	15.6235	-8.1246	12.5008	3.7500	50.0001	7.5005	2.2500	0.0000
	B	15.6249	-8.1250	12.5000	3.7500	50.0000	7.5000	2.2500	0.0000
	Analytic	15.6250	-8.1250	12.5000	3.7500	50.0000	7.5000	2.2500	0.0000

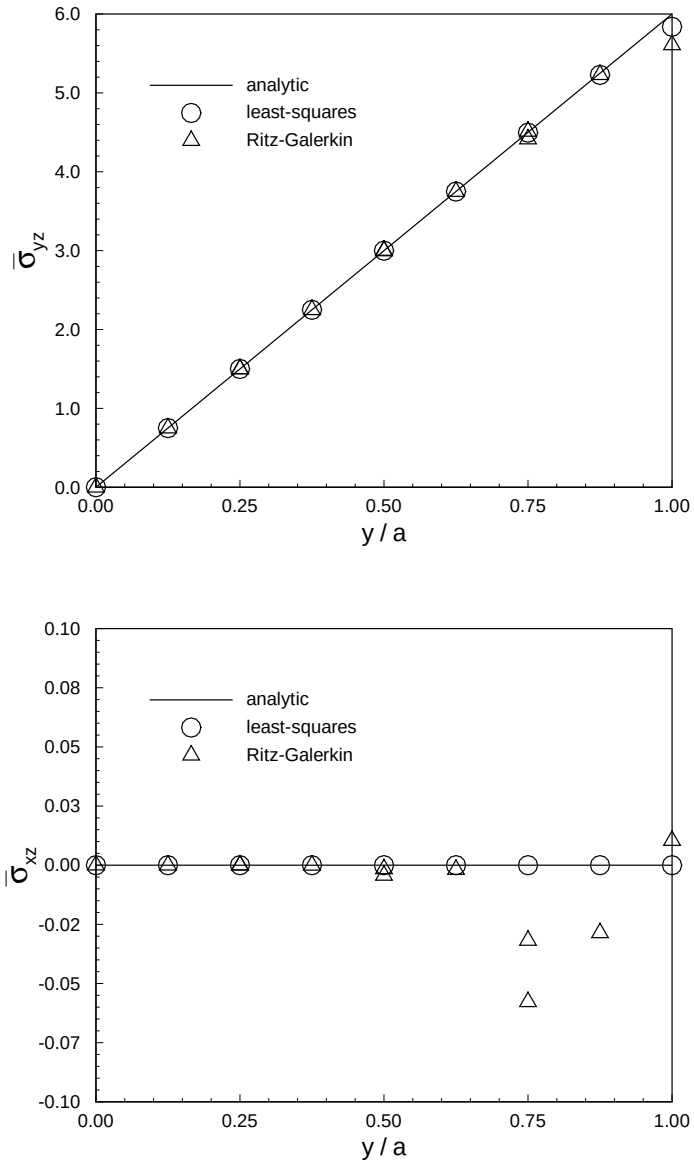

Fig. 58. Transverse shear stress predictions along radial line $\overline{ce}$. Clamped circular plate with diameter-to-thickness ratio 100. Mesh A with p-level fixed at 4 for both least-squares and Ritz-Galerkin elements.

3. Rectangular plates

We now consider the bending of rectangular plates with various boundary conditions. The total domain of the plate is $\bar{\Omega} \times [-h/2, h/2]$, where h is the thickness of the plate. The undeformed mid-plane of the plate, $\bar{\Omega}$, has dimensions $[-a/2, a/2] \times [-b/2, b/2]$ and is shown in Fig. 59. The labeled open circles in Fig. 59 show locations at which deflections, stress resultants, and stresses will be tabulated for the different boundary conditions and used to compare against available analytic solutions for both the CPT and FSDT. Results are reported in normalized form as defined in Eqs. (5.21) through (5.23).

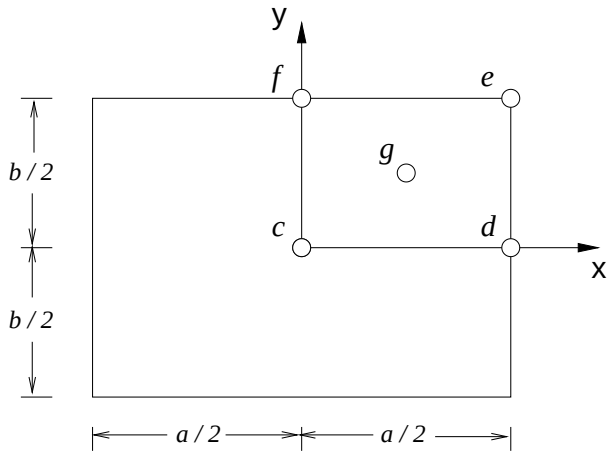

Fig. 59. Rectangular plate showing points and respectively assigned labels where displacement, stress resultants, and stresses are recorded and tabulated.

a. Clamped plate

Consider a square isotropic plate with clamped edges and subjected to a uniformly distributed load of intensity q_0. Due to the symmetry only one quadrant of the plate need be modeled. The boundary conditions of the computational domain are specified

as

$$\phi_x = 0\,, \quad M_{xy} = 0\,, \quad Q_x = 0 \qquad \text{on } x = 0\,,\ y = [0, b/2]$$

$$\phi_y = 0\,, \quad M_{xy} = 0\,, \quad Q_y = 0 \qquad \text{on } y = 0\,,\ x = [0, a/2]$$

$$w = 0\,, \quad \phi_x = 0\,, \quad \phi_y = 0 \qquad \text{on } x = a/2\,,\ y = [0, b/2]$$

$$w = 0\,, \quad \phi_x = 0\,, \quad \phi_y = 0 \qquad \text{on } y = b/2\,,\ x = [0, a/2]$$

The connected model of the quarter plate, $\bar{\Omega}^h$, consists of 2×2 quadrilateral finite elements, $\bar{\Omega}_e$, and is shown in Fig. 60. Table V shows result of a p-convergence study using the uniform mesh (mesh A in Fig. 60) for normalized deflection, bending moment, and in-plane normal stress for a square plate with $\nu = 0.30$. We consider p-levels in the range of 4 to 12, corresponding to discrete systems in the range of 648 to 5000 degrees of freedom respectively.

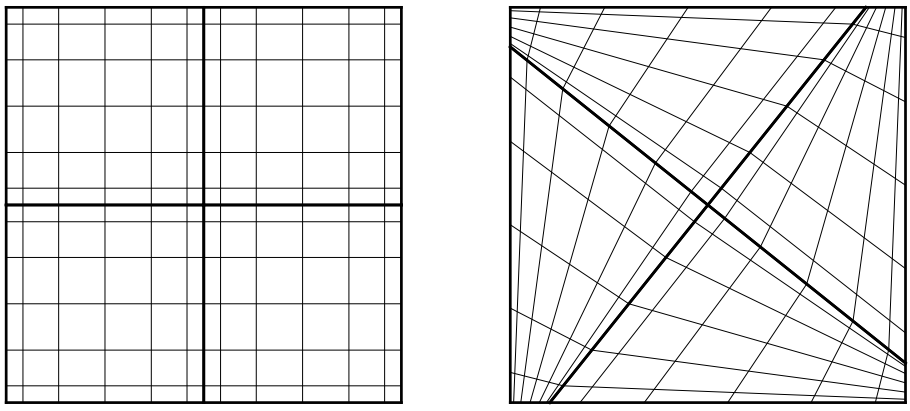

Fig. 60. Quarter plate computational domain using 4 quadrilateral finite elements. The element discretization is shown by the thick lines and the nodal points for sixth-order expansion are shown by the intersection of the thin lines. Mesh A: uniform mesh. Mesh B: distorted mesh.

The left-superscript on each normalized variable in Table V denotes the location at which it was recorded (see Fig. 59). We observe good convergence for the shear-deformable model at moderate and high side-to-thickness (a/h) ratios and for the classical plate model. We compare the computed values against the analytic *series solution* of Lim and Reddy [71]. The computed center deflection is exact for p-levels of 6 and higher. The slight discrepancy between the computed and analytic value for the bending moment is explained by realizing that a representative analytic value was taken by using only the first 20 terms in the series representation.

In Table VI we present results for normalized deflection, stress resultants, and stresses at various locations in the plate for a range of plate thickness. p-levels of 10 are used for side-to-thickness ratios up to 100, a p-level of 12 is used for the case $a/h = 200$. Such high p-levels were used to instill confidence in the results so that they may be used as a benchmark for comparison with low-order methods.

From Table V we see that no catastrophic loss of accuracy (shear-locking) is encountered for the least-squares formulation for such a course mesh and the range of p-levels considered. In the Ritz-Galerkin formulation, shear locking is more pronounced for low-order elements and for high side-to-thickness ratios (thin plates). Shear locking is alleviated by using selective reduced integration techniques or an appropriate mixed interpolation formulation [92, 93]. Again, these problems need not be treated here because they simply do not arise.

To further test the predictive capability and robustness of the plate bending elements using the least-squares formulation we consider a new connected model of the quarter plate, $\bar{\Omega}^h$, consisting of 2×2 geometrically distorted quadrilateral finite elements (mesh B in Fig. 60).

In Table VII we present the geometrically distorted mesh results for the same variables and their respective locations as in Table VI for p levels of 6 and 10. We

observe little or no change in the prediction of deflection, stress resultants, and stresses when compared to the geometrically undistorted mesh results of Table VI. Hence, the high-order plate bending elements based on the least-squares formulation not only display shear-locking-free behavior but are also insensitive to element geometric distortions.

Table V. p-convergence study showing normalized deflection, bending moment, and in-plane normal stress for a clamped, square, isotropic ($\nu = 0.30$) plate under a uniformly distributed load. FSDT model ($K_s = 5/6$). Quarter plate using a 2×2 uniform mesh.

a/h	p level	$^{c}\bar{w}$	$^{d}\bar{M}_{xx}$	$^{d}\bar{\sigma}_{xx}$
10	4	1.5037	4.9572	2.9239
	6	1.5046	4.9383	2.9608
	8	1.5046	4.9364	2.9610
	10	1.5046	4.9366	2.9616
	12	1.5046	4.9367	2.9619
	Analytic	1.5046	5.0738	
100	4	1.2693	5.1514	3.0432
	6	1.2680	5.1365	3.0829
	8	1.2679	5.1322	3.0799
	10	1.2679	5.1314	3.0791
	12	1.2679	5.1312	3.0788
	Analytic	1.2679	5.1327	
CPT	4	1.2654	5.1486	3.0822
	6	1.2653	5.1333	3.0801
	8	1.2653	5.1333	3.0802
	10	1.2653	5.1333	3.0800
	12	1.2653	5.1333	3.0800
	Analytic	1.2653	5.1334	

Table VI. Normalized deflection, stress resultants, and stresses for a clamped, square, isotropic ($\nu = 0.30$) plate under a uniformly distributed load. FSDT model ($K_s = 5/6$). Quarter plate using a 2×2 uniform mesh.

a/h	p level	${}^{c}\bar{w}$	${}^{d}\bar{M}_{xx}$	${}^{g}\bar{M}_{xy}$	${}^{f}\bar{Q}_{y}$	${}^{c}\bar{\sigma}_{xx}$	${}^{d}\bar{\sigma}_{xx}$	${}^{g}\bar{\sigma}_{xy}$	${}^{d}\bar{\sigma}_{xz}$
5	10	2.1722	4.6246	0.5728	38.2058	-1.4145	2.7747	0.3438	4.5848
	Analytic	2.1722	4.9797		38.1903				
10	10	1.5046	4.9366	0.6814	41.2186	-1.3920	2.9616	0.4089	4.9463
	Analytic	1.5046	5.0738		41.1982				
20	10	1.3273	5.0781	0.7287	42.8833	-1.3800	3.0466	0.4368	5.1460
	Analytic	1.3273	5.1166		42.8595				
100	10	1.2679	5.1314	0.7509	43.8450	-1.3746	3.0791	0.4471	5.2613
	Analytic	1.2679	5.1327		43.9145				
200	12	1.2660	5.1328	0.7500	43.9555	-1.3744	3.0800	0.4481	5.2746
	Analytic	1.2660	5.1332		44.0202				
CPT	10	1.2653	5.1333	0.7475	44.1244	-1.3743	3.0800	0.4485	0.0000
	Analytic	1.2653	5.1334		44.1193				

Table VII. Geometrically distorted mesh results: normalized deflection, stress resultants, and stresses for a clamped, square, isotropic ($\nu = 0.30$) plate under a uniformly distributed load. FSDT model ($K_s = 5/6$). Quarter plate using a 2×2 geometrically distorted mesh (mesh B in Fig. 60).

a/h	p level	${}^{c}\bar{w}$	${}^{d}\bar{M}_{xx}$	${}^{g}\bar{M}_{xy}$	${}^{f}\bar{Q}_{y}$	${}^{c}\bar{\sigma}_{xx}$	${}^{d}\bar{\sigma}_{xx}$	${}^{g}\bar{\sigma}_{xy}$	${}^{d}\bar{\sigma}_{xz}$
5	6	2.1722	4.6204	0.5720	38.2029	-1.4123	2.7691	0.3430	4.5836
	10	2.1722	4.6252	0.5730	38.2053	-1.4145	2.7750	0.3436	4.5847
	Analytic	2.1722	4.9797		38.1903				
10	6	1.5046	4.9359	0.6783	41.1520	-1.3895	2.9515	0.4074	4.9517
	10	1.5046	4.9369	0.6816	41.2184	-1.3920	2.9621	0.4087	4.9463
	Analytic	1.5046	5.0738		41.1982				
20	6	1.3272	5.0767	0.7243	42.3852	-1.3776	3.0458	0.4361	5.1520
	10	1.3273	5.0780	0.7284	42.8666	-1.3799	3.0468	0.4370	5.1462
	Analytic	1.3273	5.1166		42.8595				
100	6	1.2679	5.1159	0.7313	42.8093	-1.3719	3.0700	0.4486	5.1978
	10	1.2679	5.1297	0.7463	43.5255	-1.3743	3.0810	0.4482	5.2568
	Analytic	1.2679	5.1327		43.9145				
200	6	1.2660	5.1135	0.7324	43.5820	-1.3736	3.0734	0.4485	5.1758
	10	1.2660	5.1308	0.7461	43.5317	-1.3743	3.0799	0.4486	5.2520
	Analytic	1.2660	5.1332		44.0202				
CPT	6	1.2653	5.1348	0.7476	44.1323	-1.3741	3.0831	0.4484	0.0000
	10	1.2653	5.1333	0.7475	44.2010	-1.3743	3.0799	0.4485	0.0000
	Analytic	1.2653	5.1334		44.1193				

b. Orthotropic simply supported plate

Consider a square *orthotropic* plate with simply supported edges subjected to two different types of transverse loads: a sinusoidally distributed load with maximum intensity q_0 and a uniformly distributed load of intensity q_0. Due to the symmetry only one quadrant of the plate need be modelled. The boundary conditions of the computational domain are specified as

$$\phi_x = 0\,, \quad M_{xy} = 0\,, \quad Q_x = 0 \qquad \text{on } x = 0\,,\ y = [0, b/2]$$

$$\phi_y = 0\,, \quad M_{xy} = 0\,, \quad Q_y = 0 \qquad \text{on } y = 0\,,\ x = [0, a/2]$$

$$w = 0\,, \quad \phi_y = 0\,, \quad M_{xx} = 0 \qquad \text{on } x = a/2\,,\ y = [0, b/2]$$

$$w = 0\,, \quad \phi_x = 0\,, \quad M_{yy} = 0 \qquad \text{on } y = b/2\,,\ x = [0, a/2]$$

Like before we consider the 2×2 uniform and geometrically distorted meshes for the quarter plate, shown in Fig. 60.

For an orthotropic material the constitutive relationships are replaced by

$$\mathbf{M} = \underline{\mathbf{D}}\,\boldsymbol{\epsilon}^1 \qquad \text{in } \Omega \tag{5.24}$$

$$\alpha\,\mathbf{Q} = \boldsymbol{\beta}\,\boldsymbol{\epsilon}^0 \qquad \text{in } \Omega \tag{5.25}$$

with the following new definitions for $\underline{\mathbf{D}}$ and $\boldsymbol{\beta}$

$$\underline{\mathbf{D}} = \frac{h^3}{12}\begin{bmatrix} q_{11} & q_{12} & 0 \\ q_{12} & q_{22} & 0 \\ 0 & 0 & q_{66} \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} K_s\, q_{55}\, h & K_s\, q_{44}\, h \end{bmatrix}$$

where

$$q_{11} = \frac{E_1}{1 - \nu_{12}\,\nu_{21}}\,, \quad q_{12} = \nu_{21}\,q_{11}\,, \quad q_{22} = \frac{E_2}{1 - \nu_{12}\,\nu_{21}}$$

$$q_{44} = G_{23}\,, \quad q_{55} = G_{13}\,, \quad q_{66} = G_{12}$$

E_i are Young's moduli in the x_i^{m} material direction, ν_{ij} is the Poisson's ratio, and G_{ij} are shear moduli in the i-jth material planes. Here we take the material directions to coincide with the cartesian directions. Of course, for the CPT model $\boldsymbol{\beta} = [1 \quad 1]$.

For the uniform mesh (mesh A in Fig. 60), lines of constant ξ and η in $\hat{\Omega}_e$ coincide with the material directions. Of particular interest is to test the predictive capability of the elements when the geometrically distorted mesh is used (mesh B in Fig. 60). For this mesh, lines of constant ξ and η in $\hat{\Omega}_e$ no longer coincide with the material directions.

We consider the case $E_1 = 25\,E_2$, $G_{12} = G_{13} = 0.5\,E_2$, $G_{23} = 0.2\,E_2$, $\nu_{12} = 0.25$. Table VIII shows results for the sinusoidally distributed load case at a fixed p-level of 6, corresponding to a discrete system of 1352 degrees of freedom, for both uniform and geometrically distorted meshes. We present results for normalized deflection and stresses at various locations in the plate for a range of plate thickness and compare with the analytic *closed form solution* of Reddy [92]. The normalization for the deflection is defined as $\bar{w} = 10^3\,w\,D_{22}/q_0\,a^4$.

From results recorded in Table VIII we see that we predict exact deflection and stresses (including transverse shear stresses) for all the side-to-thickness ratios considered; using either the uniform or geometrically distorted mesh.

Table IX shows results for the uniformly distributed load case using the uniform mesh for p-levels of 6, 8, and 10; corresponding to discrete systems of 1352, 2312, and 3528 degrees of freedom respectively. We present results for normalized deflection and

stresses at various locations in the plate for a range of plate thickness and compare with the analytic *series solution* of Reddy [92].

We see from the table that the numerical solution at a p-level of 6 is already exact for the center deflection and well within an acceptable tolerance (when compared to higher p-level values) for the stresses for moderate and high side-to-thickness ratios. The slight discrepancy between the computed and analytic values for the stresses can be explained by realizing that only the first 19 terms of the series solution were used as a representative analytic value (see Reddy [92]). The analytic series representation associated with the shear stresses are typically characterized by slow asymptotic convergence, especially those associated with the transverse shear stresses.

Table X shows results using the geometrically distorted mesh for p-levels of 6, 8, and 10. We observe no deterioration in the numerical results and, like for the uniform mesh, a p-level of 6 is sufficient to obtain good accuracy.

Table VIII. Normalized deflection and stresses for a simply supported, square, orthotropic plate under a sinusoidally distributed load. FSDT model ($K_s = 5/6$). p-level fixed at 6.

a/h	mesh	${}^{c}\bar{w}$	${}^{c}\bar{\sigma}_{xx}$	${}^{c}\bar{\sigma}_{yy}$	${}^{e}\bar{\sigma}_{xy}$	${}^{d}\bar{\sigma}_{xz}$	${}^{f}\bar{\sigma}_{yz}$
10	A	0.5332	-5.2480	-0.3386	0.2463	3.4522	0.3675
	B	0.5332	-5.2480	-0.3385	0.2463	3.4522	0.3674
	Analytic	0.533	-5.248	-0.338	0.246	3.452	0.367
20	A	0.4040	-5.3499	-0.2859	0.2217	3.5008	0.3190
	B	0.4040	-5.3500	-0.2858	0.2218	3.5005	0.3185
	Analytic	0.404	-5.350	-0.286	0.222	3.501	0.319
50	A	0.3673	-5.3809	-0.2698	0.2143	3.5156	0.3042
	B	0.3673	-5.3809	-0.2697	0.2143	3.5143	0.3027
	Analytic	0.367	-5.380	-0.270	0.214	3.515	0.304
100	A	0.3620	-5.3854	-0.2675	0.2132	3.5177	0.3020
	B	0.3620	-5.3851	-0.2674	0.2131	3.5164	0.3014
	Analytic	0.362	-5.385	-0.267	0.213	3.517	0.302
CPT	A	0.3603	-5.3870	-0.2667	0.2128	0.0000	0.0000
	B	0.3603	-5.3869	-0.2667	0.2128	0.0000	0.0000
	Analytic	0.360	-5.387	-0.267	0.213	0.000	0.000

Table IX. Normalized deflection and stresses for a simply supported, square, orthotropic plate under a uniformly distributed load. FSDT model ($K_s = 5/6$). Quarter plate using a 2×2 uniform mesh.

a/h	p level	${}^{c}\bar{w}$	${}^{c}\bar{\sigma}_{xx}$	${}^{c}\bar{\sigma}_{yy}$	${}^{e}\bar{\sigma}_{xy}$	${}^{d}\bar{\sigma}_{xz}$	${}^{f}\bar{\sigma}_{yz}$
10	6	0.7953	-7.7065	-0.3523	0.5249	6.2739	1.6511
	8	0.7953	-7.7065	-0.3524	0.5321	6.2747	1.6517
	10	0.7953	-7.7065	-0.3524	0.5365	6.2748	1.6518
	Analytic	0.795	-7.706	-0.352	0.539	6.147	1.529
20	6	0.6067	-7.8288	-0.2727	0.4822	6.3192	1.5877
	8	0.6067	-7.8286	-0.2728	0.4830	6.3207	1.5886
	10	0.6067	-7.8286	-0.2728	0.4838	6.3209	1.5888
	Analytic	0.607	-7.828	-0.272	0.487	6.194	1.466
50	6	0.5531	-7.8611	-0.2491	0.4681	6.3277	1.5709
	8	0.5531	-7.8610	-0.2491	0.4687	6.3315	1.5726
	10	0.5531	-7.8610	-0.2491	0.4689	6.3321	1.5730
	Analytic	0.553	-7.860	-0.249	0.468	6.207	1.452
100	6	0.5454	-7.8654	-0.2456	0.4656	6.3265	1.5677
	8	0.5454	-7.8654	-0.2457	0.4659	6.3325	1.5703
	10	0.5454	-7.8654	-0.2457	0.4660	6.3335	1.5708
	Analytic	0.545	-7.865	-0.245	0.464	6.206	1.449
CPT	6	0.5428	-7.8670	-0.2446	0.4672	0.0000	0.0000
	8	0.5428	-7.8669	-0.2446	0.4658	0.0000	0.0000
	10	0.5428	-7.8669	-0.2446	0.4653	0.0000	0.0000
	Analytic	0.543	-7.866	-0.244	0.463	0.000	0.000

Table X. Geometrically distorted mesh results: normalized deflection and stresses for a simply supported, square, orthotropic plate under a uniformly distributed load. FSDT model ($K_s = 5/6$). Quarter plate using a 2×2 geometrically distorted mesh (mesh B in Fig. 60).

a/h	p level	${}^{c}\bar{w}$	${}^{c}\bar{\sigma}_{xx}$	${}^{c}\bar{\sigma}_{yy}$	${}^{e}\bar{\sigma}_{xy}$	${}^{d}\bar{\sigma}_{xz}$	${}^{f}\bar{\sigma}_{yz}$
10	6	0.7953	-7.7050	-0.3493	0.4611	6.2750	1.6491
	8	0.7953	-7.7062	-0.3526	0.4879	6.2749	1.6512
	10	0.7953	-7.7065	-0.3524	0.5084	6.2750	1.6518
	Analytic	0.795	-7.706	-0.352	0.539	6.147	1.529
20	6	0.6067	-7.8236	-0.2701	0.3968	6.3214	1.5855
	8	0.6067	-7.8289	-0.2732	0.4192	6.3208	1.5873
	10	0.6067	-7.8287	-0.2728	0.4389	6.3210	1.5884
	Analytic	0.607	-7.828	-0.272	0.487	6.194	1.466
50	6	0.5530	-7.8514	-0.2477	0.4428	6.3315	1.5728
	8	0.5531	-7.8616	-0.2495	0.4429	6.3326	1.5723
	10	0.5531	-7.8610	-0.2491	0.4444	6.3323	1.5727
	Analytic	0.553	-7.860	-0.249	0.468	6.207	1.452
100	6	0.5454	-7.8603	-0.2442	0.4647	6.3174	1.5575
	8	0.5454	-7.8660	-0.2459	0.4599	6.3344	1.5706
	10	0.5454	-7.8654	-0.2457	0.4592	6.3341	1.5709
	Analytic	0.545	-7.865	-0.245	0.464	6.206	1.449
CPT	6	0.5428	-7.8687	-0.2444	0.4683	0.0000	0.0000
	8	0.5428	-7.8666	-0.2446	0.4664	0.0000	0.0000
	10	0.5428	-7.8669	-0.2446	0.4653	0.0000	0.0000
	Analytic	0.543	-7.866	-0.244	0.463	0.000	0.000

c. Square plate with two opposite simply supported edges

Consider a square isotropic plate with two opposite simply supported edges (and two opposite free edges) subjected to a uniformly distributed load of intensity q_0. Due to the symmetry only one quadrant of the plate need be modelled. The boundary conditions of the computational domain are specified as

$$\phi_x = 0\,, \quad M_{xy} = 0\,, \quad Q_x = 0 \qquad \text{on } x = 0\,,\ y = [0, b/2]$$

$$\phi_y = 0\,, \quad M_{xy} = 0\,, \quad Q_y = 0 \qquad \text{on } y = 0\,,\ x = [0, a/2]$$

$$w = 0\,, \quad \phi_y = 0\,, \quad M_{xx} = 0 \qquad \text{on } x = a/2\,,\ y = [0, b/2]$$

$$M_{yy} = 0\,, \quad M_{xy} = 0\,, \quad Q_y = 0 \qquad \text{on } y = b/2\,,\ x = [0, a/2]$$

This case is of particular relevance, as it has been reported that when the p-Ritz method is used to analyze this problem (using the FSDT model) it is found that the shear forces and twisting moments are not accurate when thin plates are considered [63, 106]. Oscillations are observed in the computed stress-resultant distributions and the natural boundary conditions cannot be satisfied. A penalty function method was proposed to ensure the satisfaction of the natural boundary conditions and a post-processing curve fitting procedure to filter out the oscillations (see Wang et al. [106]).

Such "fixes" are not needed in the least-squares formulation as boundary conditions on shear forces and/or bending moments can be imposed strongly (pointwise). Figures 61 and 62 show the computed shear force and twisting moment distributions along the simply supported edge (with $\nu = 0.30$). No oscillations are observed as the plate thickness is decreased. A graded 4×4 mesh using a p-level of 6 was used to generate the distributions.

In Table XI we present results for normalized deflection and stress resultants at various locations in the plate for a range of plate thickness. We see that the computed maximum deflection is exact, when compared against the analytic *series solution* value of Lim and Reddy [71]. Analytic values for the stress resultants were not reported in Ref. [71].

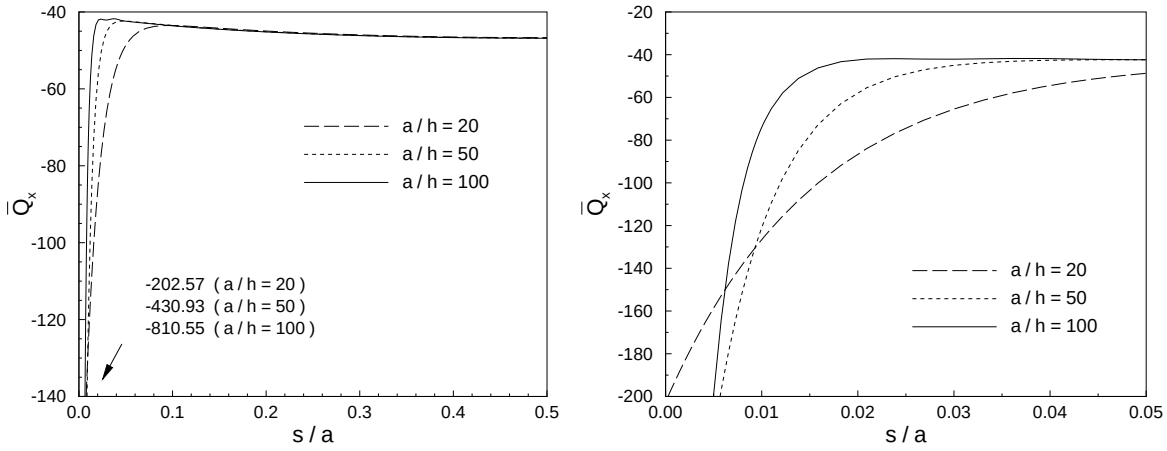

Fig. 61. Shear force $\bar{Q}_x$ distribution along simply supported edge for decreasing plate thickness: full-view ($0.0 \leq s/a \leq 0.5$) and close-up view ($0.0 \leq s/a \leq 0.05$). Uniformly loaded square plate with two opposite simply supported edges.

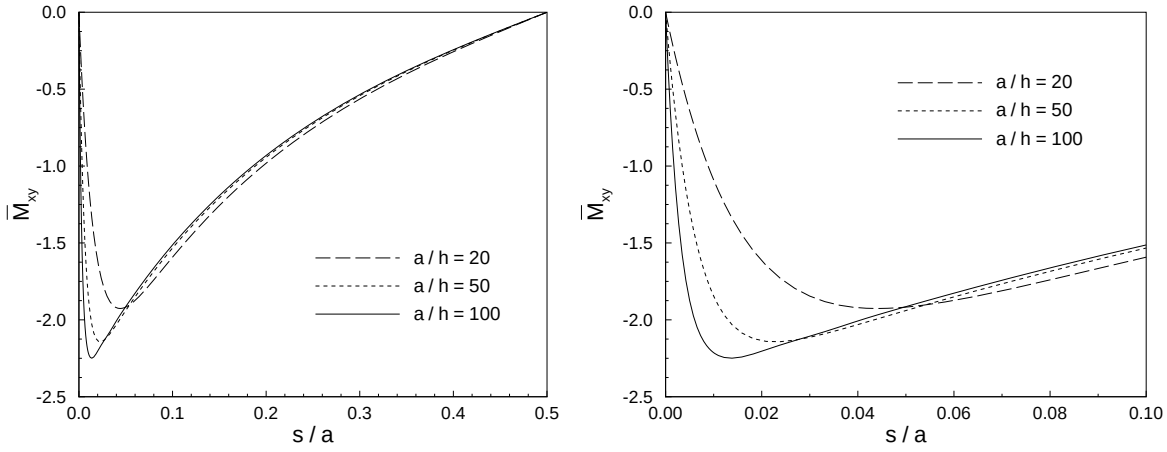

Fig. 62. Twisting moment $\bar{M}_{xy}$ distribution along simply supported edge for decreasing plate thickness: full-view ($0.0 \leq s/a \leq 0.5$) and close-up view ($0.0 \leq s/a \leq 0.1$). Uniformly loaded square plate with two opposite simply supported edges.

Table XI. Normalized deflection and stress resultants for a square, isotropic ($\nu = 0.30$) plate with two opposite simply supported edges under a uniformly distributed load. FSDT model ($K_s = 5/6$). Quarter plate using a 4×4 graded mesh.

a/h	p level	[f] $\bar{w}$	[f] $\bar{M}_{xx}$	[c] $\bar{M}_{yy}$	[d] $\bar{Q}_x$	[e] $\bar{Q}_x$
20	6	15.2165	12.9344	2.6415	46.6828	202.57
	Analytic	15.2165				
50	6	15.0706	13.0389	2.6826	46.7946	430.93
	Analytic	15.0706				
100	6	15.0380	13.0710	2.6954	46.8321	810.55
	Analytic	15.0380				

4. Skew plates

We now consider the bending of skew, simply supported, isotropic plates subjected to a uniformly distributed load of intensity q_0. We denote by θ the skew angle of the plate (see Fig. 63).

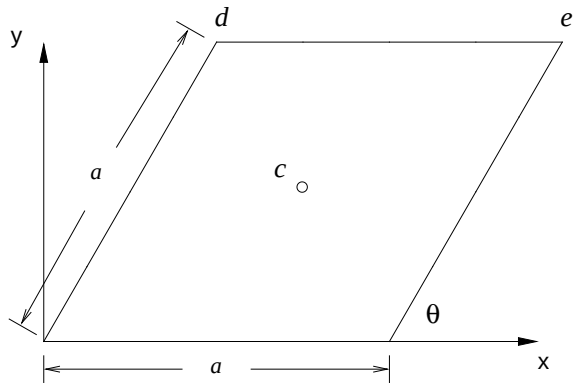

Fig. 63. Skew plate with side lengths a and skew angle θ.

We expect deterioration of accuracy in the numerical solution as θ becomes smaller. As pointed out by Babuška and Scapolla [4], the main cause in deterioration of the numerical solution is the rapid decay of smoothness of the exact solution with decreasing θ, and has little to do with the skewness of the elements which would worsen with decreasing θ.

For the CPT we find singular behavior of the bending moments at the obtuse angle corners of the plate. Babuška and Scapolla [4] estimated that for the analysis of a plate with skew angle $\theta = 30^\circ$ using a uniform mesh, 10^6 degrees of freedom for a displacement-based conforming C^1 element based on the Ritz-Galerkin formulation would be needed to achieve a five percent accuracy for the displacement at the center of the plate. Here we construct a connected model of the full plate, $\bar{\Omega}^h$, consisting of 4×4 uniformly spaced quadrilateral finite elements, $\bar{\Omega}_e$, shown in Fig. 64a.

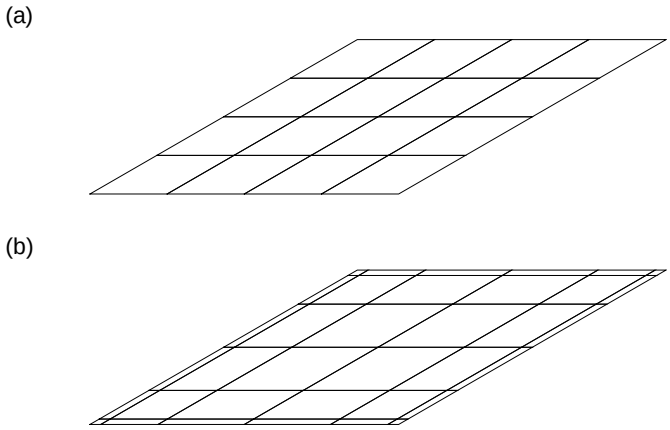

Fig. 64. Computational domain for $\theta = 30^\circ$ using (a) 4×4 uniform mesh and (b) 6×6 non-uniform mesh.

Table XII shows result of a p-convergence study for normalized deflection, principal bending moments, and principal stresses at the center of the plate for $\theta = 60^\circ$ and $\theta = 30^\circ$. Results are reported in normalized form as defined in Eqs. (5.21) through (5.23). We see excellent agreement with the analytic series solution of Morley [74]. As expected, we observe better convergence for the smoother case of $\theta = 60^\circ$, for which the solution at a p-level of 4 (2312 degrees of freedom) is already well within one percent accuracy for the displacement at the center of the plate. For the case of $\theta = 30^\circ$ a slower rate of convergence is observed; for a a p-level of 10 (13448 degrees of freedom) we are well within one percent accuracy for the displacement at the center of the plate.

For the FSDT simulation of the transversely loaded skew plate we specify simply supported boundary conditions of the soft type: $w = M_{nn} = M_{ns} = 0$; as opposed to the hard type: $w = \phi_s = M_{nn} = 0$. This allows the natural occurrence of boundary layers at the corners so as to neutralize the CPT singularity; a great improvement in the description of the actual physical behavior of the plate allowed by the FSDT model. Good resolution is desired near the obtuse corners to capture the boundary

Table XII. p-convergence study showing normalized deflection, principal bending moments, and principal stresses for a simply supported, isotropic ($\nu = 0.30$) skew plate under a uniformly distributed load. CPT model. Full plate using a 4×4 uniform mesh.

θ	p level	${}^c\bar{w}$	${}^c\bar{M}_{\min}$	${}^c\bar{M}_{\max}$	${}^c\bar{\sigma}_{\min}$	${}^c\bar{\sigma}_{\max}$
60°	4	2.5618	3.4146	4.1840	2.0208	2.5291
	6	2.5602	3.3433	4.2342	2.0119	2.5365
	8	2.5601	3.3304	4.2502	2.0010	2.5491
	10	2.5601	3.3285	4.2537	1.9977	2.5524
	Analytic	2.56	3.33	4.25		
30°	4	0.3823	1.1848	1.8333	0.6458	1.0428
	6	0.4000	1.1140	1.8852	0.6504	1.1185
	8	0.4053	1.0954	1.9009	0.6515	1.1371
	10	0.4069	1.0909	1.9045	0.6510	1.1415
	Analytic	0.408	1.08	1.91		

layers. We construct a connected model of the full plate, $\bar{\Omega}^h$, using a 6×6 non-uniform quadrilateral finite element mesh, shown in Fig. 64b. The spacing of the elements near the edges of the plate was chosen as $\delta/a = 0.035$.

Table XIII shows results for normalized deflection, principal bending moments, and principal stresses at the center of the plate for $\theta = 60°$, $\theta = 30°$, and decreasing plate thickness h. We present results for p-levels of 10. At these p-levels the change in normalized center deflection (between p-levels of p and $p-2$) was less than 0.05% and 1% for skew angles of $\theta = 60°$ and $\theta = 30°$ respectively. As expected, the effect of the transverse shear strains for thick plates is to increase the center deflection. We see nice convergence towards the CPT solution as the plate becomes thin. No shear-locking due to element skewness and/or decreasing plate thickness was observed.

Figure 65 shows the normalized transverse shear stress distribution along line $\overline{de}$ for a plate with skew angle $\theta = 60°$ and decreasing plate thickness. The normalization is taken here as $\bar{\sigma}_{xz} = 10\,\sigma_{xz}/q_0\,(a/h)^2$. We observe a mild boundary layer at the

obtuse corner for low values of a/h (thick plates). The boundary layer becomes more pronounced as the plate becomes thin.

Table XIII. Normalized deflection, principal bending moments, and principal stresses for a (softly) simply supported, isotropic ($\nu = 0.30$) skew plate under a uniformly distributed load. FSDT model ($K_s = 5/6$). Full plate using a 6×6 non-uniform mesh (see Fig. 64b).

θ	a/h	p level	${}^c\bar{w}$	${}^c\bar{M}_{\min}$	${}^c\bar{M}_{\max}$	${}^c\bar{\sigma}_{\min}$	${}^c\bar{\sigma}_{\max}$
60°	5	10	3.6449	3.7769	4.7776	2.2661	2.8665
	20	10	2.7505	3.5129	4.4174	2.1075	2.6505
	100	10	2.6001	3.3800	4.2930	2.0273	2.5758
30°	5	10	0.7306	1.2199	2.1414	0.7320	1.2848
	20	10	0.4553	1.1858	2.0123	0.7089	1.2060
	100	10	0.4073	1.1364	1.9346	0.6499	1.1433

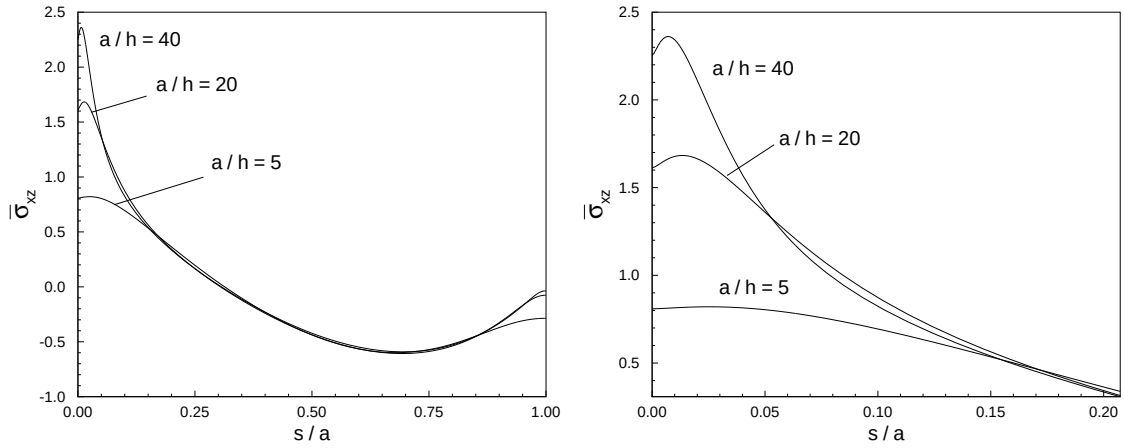

Fig. 65. Transverse shear stress distribution along line $\overline{de}$ for a plate with skew angle of 60° and decreasing plate thickness: full-view ($0.0 \leq s/a \leq 1.0$) and close-up view ($0.0 \leq s/a \leq 0.2$).

CHAPTER VI

SHEAR-DEFORMABLE SHELL STRUCTURES

In this chapter[†] we present a least-squares finite element model for the analysis of shear-deformable shells. Shell structures are encountered in many important engineering applications and a reliable method to numerically analyze such structures is highly desirable. Unlike (linear) plate bending, where the membrane and bending effects are uncoupled, shell structures naturally couple the membrane and bending effects to support externally applied loads. On the other hand, very much like plate bending, a shell model can be derived according to the *classical shell theory* or the *first-order shear deformation shell theory* (also known as the Kirchhoff/Koiter and Mindlin/Naghdi shell theories, respectively).

Finite element formulations for the analysis of shell structures are traditionally derived from the principle of virtual displacements or the principle of minimum total potential energy (see Reddy [93, 94]). Moreover, two distinct classes of shell finite element models are available: degenerate shell elements based on three-dimensional continuum theory and shell elements based on a (curvilinear) two-dimensional shell theory. The degenerate approach appeared first in the work of Ahmad et al. [2], it degenerates the shell element from a three-dimensional solid element via kinematical assumptions on the displacements. In this approach no explicit two-dimensional model is employed, although implicitly a first-order shear deformation shell theory model can be identified. Furthermore, shells of arbitrary shape can be modelled since the surface is constructed through isoparametric interpolation. In shell theory

[†]The numerical results reported in this chapter appear in the article "Hierarchical mixed least-squares shear-deformable shell elements" by J. P. Pontaza and J. N. Reddy, *Int. J. Numer. Meth. Engng*, submitted for publication. Copyright (2003) John Wiley & Sons.

based models the underlying kinematic assumptions associated with the description of the shell are identified explicitly and the strain measure, constitutive equations, and equilibrium equations derived in general curvilinear coordinates. Typically the analytic definition of the undeformed shell surface geometry is given, and restricted to relatively simple geometries. Shells of arbitrary shape can be represented by isoparametric interpolation in a local tangent surface using a two parameter representation typically provided by a CAD system. In the present study we work with a first-order shear deformation shell theory based finite element model, where an analytic definition of the undeformed shell surface geometry is given.

When considering the limiting behavior of a shell as the thickness becomes small, for a given shell geometry and boundary conditions, the shell problem will in general fall into either a membrane dominated or bending dominated state – depending on whether the membrane or bending energy component dominates the total energy. Displacement-based finite element models have no major difficulties in predicting the asymptotic behavior of the shell structure in the membrane dominated case. However, computational difficulties arise in the case when the deformation is bending dominated [70, 79, 43]. A strong stiffening of the element matrices occurs, resulting in spurious predictions for the membrane energy component. This phenomenon is known as *membrane-locking*, and occurs because the unconstrained minimization problem becomes a penalized form of an optimization problem under constraints that are not well represented in the numerical approximation of the mathematical model. Specifically, that the membrane and shear energy components should vanish and that the displacement field should belong to the space of inextensional displacement patterns. In shear-deformable shell models, yet another form of locking occurs and presents itself (again) in a strong stiffening of the element matrices, resulting in spurious predictions for the shear energy component. This form of locking is also

present in plate bending analysis when the side-to-thickness ratio of the plate is large (i.e., when modelling thin plates). This locking phenomenon is known as *shear-locking*.

Displacement-based finite element models for the first-order shear deformation theory (FSDT) allow the use of C^0 interpolation functions for the three displacements and the two independent rotations. Membrane- and shear-locking are evident in Ritz-Galerkin displacement-based finite element models with equal order interpolation of all generalized displacements when full integration is used to analyze thin shells, and is more pronounced when low-order elements are used. The performance of low-order elements can be considerably improved by the use of reduced integration techniques. However, highly distorted meshes tend to aggravate the predictive capabilities of these elements, even when reduced integration techniques are used. An alternate approach is to use mixed formulations, where in addition to the generalized displacements the shear and membrane strains are treated as independent variables [6, 47, 24, 3, 17]. These mixed approaches, also know as assumed strain methods, interpolate displacements and strains in such a way as to have no locking in bending-dominated states and aim to preserve good properties for membrane-dominated states. Such methods can efficiently relax the problems of membrane- and shear-locking to some extent for low-order elements [73].

Higher-order displacement based elements show less sensitivity to membrane- and shear-locking, and when sufficiently high p-levels are considered the problems associated with locking disappear [43, 23]. It is well known that to mitigate the locking effects it is far more effective to raise the element p-level than refining the finite element mesh at a fixed p-level. Furthermore, when a comparison between cost and accuracy is taken into account, high-order elements require a smaller computational effort to achieve a desired level of accuracy [23].

In this chapter we present a finite element formulation for the analysis of shell structures that is robust with regards to membrane- and shear-locking and yields highly accurate results for displacements as well as stresses (or stress resultants). The formulation is based on least-squares variational principles, which results in an unconstrained minimization problem through a variational framework of *residual* minimization.

The proposed formulation is a mixed formulation, in that it uses the generalized displacements and stress-resultants as independent variables. The equilibrium equations in general curvilinear coordinates are used in their first-order form to construct the least-squares functional which in turn allows the use of C^0 interpolation functions in the finite element model. The use of least-squares principles leads to a variational unconstrained minimization problem where the approximation spaces can be chosen independently, i.e., stability requirements such as the inf-sup condition never arise (see Chapter II). Equal order interpolation and full integration is used to develop the discrete finite element model, which results in a symmetric and positive-definite system of algebraic equations. We show by numerical examples that the proposed high-order shell elements are free of membrane- and shear-locking, insensitive to geometric distortions, and predict accurate displacements as well as stress resultants along the shell boundaries.

First, we outline the first-order shear deformation linear shell theory in tensor based form and arrive at the governing equations for the shear-deformable shell model. We use the procedures in Chapter II to develop and arrive at the finite element model, and verify exponentially fast decay (spectral convergence) of the L_2 least-squares functional and L_2 error norms of the generalized displacements and stress-resultants for a smooth solution to the FSDT shell equations. The performance of the shell elements is first tested for the membrane- and bending-dominated benchmark

problems suggested in Ref. [79], using uniform and distorted meshes. The accuracy and cost of the least-squares based elements is compared to that of displacement based Ritz-Galerkin elements for the same meshes and p-levels. The elements are further tested by considering the well-known Scordelis and Lo barrel vault problem and the pinched circular cylinder shell problem.

A. The shear-deformable shell model

1. Shell geometry

Let $\bar{\Omega}$ be the closure of an open bounded region Ω in $\mathbb{R}^2$ and (ξ^1, ξ^2) be a point in $\bar{\Omega} = \Omega \cup \partial\Omega$, where $\partial\Omega = \Gamma$ is the boundary of Ω. Suppose that the shell mid-surface $\mathcal{S}$ under consideration is parametrized by a single chart $\vec{\phi} = (\phi_1, \phi_2, \phi_3)$, $\vec{\phi} : \bar{\Omega} \subset \mathbb{R}^2 \longrightarrow \mathcal{S} \subset \mathbb{R}^3$, so that we can write $\mathcal{S} = \vec{\phi}\,(\bar{\Omega})$.

In the following we briefly recall the classical definitions and notation of differential geometry that we need for the purposes of presenting the shell model, see Ref. [38] for details. We employ the convention that Greek indices range over 1 and 2 and Latin indices over 1, 2, and 3. We start by defining a local curvilinear coordinate system for the mid-surface $\mathcal{S}$, and we do so by assuming that the one-to-one mapping $\vec{\phi}$ and the boundary of Ω, Γ, are sufficiently regular to define covariant base vectors of the tangential plane. At a point $\vec{\phi}(\xi^1, \xi^2)$, the covariant base vectors are given by

$$\vec{a}_\alpha = \frac{\partial \vec{\phi}(\xi^1, \xi^2)}{\partial \xi^\alpha} = \vec{\phi}_{,\alpha}\,, \tag{6.1}$$

so that the vectors $\vec{a}_\alpha$ define a basis for the tangent space of the mid-surface $\mathcal{S}$ at that point. The unit normal to the tangent plane is

$$\vec{a}_3 = \frac{\vec{a}_1 \times \vec{a}_2}{\|\vec{a}_1 \times \vec{a}_2\|}\,, \tag{6.2}$$

so that for each $(\xi^1, \xi^2) \in \Omega$, the vectors $\vec{a}_i$ define a basis for $\mathbb{R}^3$.

We define the first fundamental form of the mid-surface $\mathcal{S}$ by the relation

$$a_{\alpha\beta} = \vec{a}_\alpha \cdot \vec{a}_\beta \,, \tag{6.3}$$

or alternatively in contravariant form by

$$a^{\alpha\beta} = \vec{a}^{\,\alpha} \cdot \vec{a}^{\,\beta} \,, \tag{6.4}$$

where the contravariant base vectors are given by $\vec{a}_\alpha \cdot \vec{a}^{\,\beta} = \delta_\alpha^\beta$, and δ_α^β is the Kronecker delta operator.

The second fundamental form is defined by

$$b_{\alpha\beta} = \vec{a}_3 \cdot \vec{a}_{\alpha,\beta} = -\vec{a}_\alpha \cdot \vec{a}_{3,\beta} \,, \tag{6.5}$$

and the mixed and contravariant forms are easily derived as $b_\alpha^\beta = a^{\beta\gamma}\, b_{\gamma\alpha}$ and $b^{\alpha\beta} = a^{\alpha\gamma}\, a^{\beta\delta}\, b_{\gamma\delta}$.

Let a denote the determinant of the matrix $[a_{\alpha\beta}]$, given by

$$a = \|\vec{a}_1 \times \vec{a}_2\| = a_{11}\, a_{22} - (a_{12})^2 \,, \tag{6.6}$$

be a never-vanishing function on $\bar{\Omega}$, so that if $f : \mathcal{S} \longrightarrow \mathbb{R}$ is any function, then $\int_{\mathcal{S}} f \, d\mathcal{S} = \int_\Omega f \left(\vec{\phi}(\xi^1, \xi^2)\right) \sqrt{a}\, d\xi^1\, d\xi^2$.

Denoting by t the shell thickness and considering shells of constant thickness only, the shell occupies the region

$$\{\vec{\phi}(\xi^1, \xi^2) + \xi^3 \vec{a}_3(\xi^1, \xi^2) \,|\, (\xi^1, \xi^2) \in \Omega, -t/2 < \xi^3 < t/2\} \subset \mathbb{R}^3 \,, \tag{6.7}$$

so that the three-dimensional geometry of the shell can be parametrized by the chart $\vec{\Phi}$ defined by

$$\vec{\Phi}(\xi^1, \xi^2, \xi^3) = \vec{\phi}(\xi^1, \xi^2) + \xi^3 \vec{a}_3(\xi^1, \xi^2) \,, \tag{6.8}$$

2. Strain measures and strain energy

In Naghdi's shear-deformable shell model [75], the displacement vector is assumed to be of the form

$$\vec{U} = \left(u_\alpha + \xi^3\,\theta_\alpha\right)\,\vec{a}^{\,\alpha} + u_3\,\vec{a}^{\,3}\,, \tag{6.9}$$

where u_α are the displacements of the shell mid-surface, u_3 is the out-of-plane displacement, and θ_α are rotations of the transverse material fibers originally normal to the shell mid-surface. In addition, it is assumed that the transverse shear stress measure is constant through the thickness. From these assumptions, the linear membrane, bending, and shear strain measures in terms of displacements and mid-surface metrics are given by

$$\varepsilon_{\alpha\beta}\left(\vec{u}\right) = \frac{1}{2}\left(u_{\alpha|\beta} + u_{\beta|\alpha}\right) - b_{\alpha\beta}\,u_3\,, \tag{6.10}$$

$$\chi_{\alpha\beta}\left(\vec{u},\underline{\theta}\right) = \frac{1}{2}\left(\theta_{\alpha|\beta} + \theta_{\beta|\alpha} - b^\gamma_\alpha\,u_{\gamma|\beta} - b^\gamma_\beta\,u_{\gamma|\alpha}\right) + b^\gamma_\alpha\,b_{\gamma\beta}\,u_3\,, \tag{6.11}$$

$$\zeta_\alpha\left(\vec{u},\underline{\theta}\right) = u_{3,\alpha} + \theta_\alpha + b^\gamma_\alpha\,u_\gamma\,, \tag{6.12}$$

where covariant differentiation on the shell mid-surface is denoted by the vertical bar, e.g., $u_{\alpha|\beta} = u_{\alpha,\beta} - \Gamma^\delta_{\alpha\beta}\,u_\delta$ with $\Gamma^\delta_{\alpha\beta} = \vec{a}^{\,\delta}\cdot\vec{a}_{\alpha,\beta}$. Note that we have used over-arrows to indicate that the underlying vector has three components, and an under-line to indicate two components.

The strain energy of the shell can be written as the sum of the membrane, bending, and shear strain energies as $\mathcal{U} = \mathcal{U}_{\rm m} + \mathcal{U}_{\rm b} + \mathcal{U}_{\rm s}$, where

$$\mathcal{U}_{\rm m} = \frac{1}{2}\int\limits_\Omega N^{\alpha\beta}\,\varepsilon_{\alpha\beta}\,\sqrt{a}\,d\xi^1\,d\xi^2\,, \tag{6.13}$$

$$\mathcal{U}_{\rm b} = \frac{1}{2}\int\limits_\Omega M^{\alpha\beta}\,\chi_{\alpha\beta}\,\sqrt{a}\,d\xi^1\,d\xi^2\,, \tag{6.14}$$

$$\mathcal{U}_{\mathrm{s}} = \frac{1}{2} \int_{\Omega} Q^{\alpha}\, \zeta_{\alpha}\, \sqrt{a}\, d\xi^1\, d\xi^2\,, \tag{6.15}$$

and $N^{\alpha\beta}$, $M^{\alpha\beta}$, Q^{α} are the membrane, bending, and shear thickness-averaged stress-resultants. The stress-resultants are related to the strain measures through the following constitutive relations

$$N^{\alpha\beta} = t\, E^{\alpha\beta\gamma\mu}\, \varepsilon_{\gamma\mu}\,, \tag{6.16}$$

$$M^{\alpha\beta} = \frac{t^3}{12}\, E^{\alpha\beta\gamma\mu}\, \chi_{\gamma\mu}\,, \tag{6.17}$$

$$Q^{\alpha} = t\, J^{\alpha\beta}\, \zeta_{\beta}\,, \tag{6.18}$$

with the constitutive tensors defined as

$$E^{\alpha\beta\gamma\mu} = \frac{E}{2\,(1+\nu)} \left(a^{\alpha\gamma}\, a^{\beta\mu} + a^{\alpha\mu}\, a^{\beta\gamma} + \frac{2\,\nu}{1-\nu}\, a^{\alpha\beta}\, a^{\gamma\mu} \right) \tag{6.19}$$

$$J^{\alpha\beta} = \frac{E}{2\,(1+\nu)}\, K_s\, a^{\alpha\beta} \tag{6.20}$$

where E is the Young's modulus, ν is the Poisson's ratio, and K_s is the shear correction factor for the isotropic material, respectively.

Contributing to the total potential energy of the shell are the energy due to external distributed loads and applied resultant loads along the edge of the mid-surface:

$$W_{\Omega} = \int_{\Omega} \left(p^{\alpha}\, u_{\alpha} + p^3\, u_3 \right)\, \sqrt{a}\, d\xi^1\, d\xi^2\,, \tag{6.21}$$

$$W_{\Gamma} = \int_{\Gamma} \left(\bar{N}^{\alpha\beta}\, n_{\alpha}\, u_{\beta} + \bar{M}^{\alpha\beta}\, n_{\alpha}\, \theta_{\beta} + \bar{Q}^{\alpha}\, n_{\alpha}\, u_3 \right)\, d\,\Gamma\,, \tag{6.22}$$

where n_{α} are the components of the unit normal to Γ. The total potential energy of the shell is thus $\mathcal{I} = \mathcal{U} - W_{\Omega} - W_{\Gamma}$.

3. Variational formulation and asymptotics

According to the principle of minimum total potential energy, the shell assumes a state of deformation such that the total potential energy, $\mathcal{I}$, is minimized. This conveniently provides a variational setting for the solution of the shear-deformable shell model: find $(\vec{u}, \underline{\theta})$ in a suitable vector space, such that $\delta\mathcal{I} = 0$, or equivalently (and more formally)

Find $U = (\vec{u}, \underline{\theta}) \in X$ such that for all $V = \left(\vec{v}, \underline{\psi}\right) \in X$

$$t^3 \, A_b \left(U, V\right) + t \, A_{ms} \left(U, V\right) = F \left(V\right) \tag{6.23}$$

where

$$X = \left\{ \left(\vec{u}, \underline{\theta}\right) \in \left[H^1(\Omega)\right]^3 \times \left[H^1(\Omega)\right]^2 \right\} \tag{6.24}$$

and the space X satisfies the essential boundary conditions. The bilinear forms $A_b(\cdot,\cdot)$, $A_{ms}(\cdot,\cdot) = A_m(\cdot,\cdot) + A_s(\cdot,\cdot)$ represent the bending and membrane plus shear strain energies, respectively. Using Eqs. (6.13)-(6.15) and Eqs. (6.16)-(6.18), the bilinear forms are written as

$$A_{\mathrm{m}} \left(U, V\right) = \int\limits_{\Omega} \, E^{\alpha\beta\gamma\mu} \, \varepsilon_{\alpha\beta} \left(\vec{u}\right) \, \varepsilon_{\gamma\mu} \left(\vec{v}\right) \, \sqrt{a} \, d\xi^1 \, d\xi^2 \, , \tag{6.25}$$

$$A_{\mathrm{b}} \left(U, V\right) = \int\limits_{\Omega} \, \frac{E^{\alpha\beta\gamma\mu}}{12} \, \chi_{\alpha\beta} \left(\vec{u}, \underline{\theta}\right) \, \chi_{\gamma\mu} \left(\vec{v}, \underline{\psi}\right) \, \sqrt{a} \, d\xi^1 \, d\xi^2 \, , \tag{6.26}$$

$$A_{\mathrm{s}} \left(U, V\right) = \int\limits_{\Omega} \, J^{\alpha\beta} \, \zeta_{\alpha} \left(\vec{u}, \underline{\theta}\right) \, \zeta_{\beta} \left(\vec{v}, \underline{\psi}\right) \, \sqrt{a} \, d\xi^1 \, d\xi^2 \tag{6.27}$$

with the effect of external distributed loads represented by

$$F \left(V\right) = \int\limits_{\Omega} \, \left(p^{\alpha} \, v_{\alpha} + p^3 \, v_3\right) \, \sqrt{a} \, d\xi^1 \, d\xi^2 \, . \tag{6.28}$$

It is important to understand how the properties of the shell model are affected

when the thickness of the shell becomes small compared to other fixed length scales of the shell geometry. In such an asymptotic analysis the subspace X_0 of X plays an important role,

$$X_0 = \left\{ U \in X \mid A_{\text{ms}}(U, U) = 0 \right\}, \tag{6.29}$$

which corresponds to the space of pure bending displacements, sometimes referred to as the space of inextensional displacements. The case $X_0 = \{0\}$ corresponds to a membrane-dominated state, where the applied loading can be resisted by membrane and shear stresses only, while the case $X_0 \neq \{0\}$ corresponds to a bending-dominated state. A more in-depth asymptotic analysis of shells can be found in Ref. [22].

The Ritz-Galerkin displacement based finite element model is obtained by restricting (6.23) to the finite-dimensional subspace X_{hp} of the infinite dimensional space X. When the finite-dimensional subspace is spanned by low-order polynomials the displacement based finite element model is effective only when membrane-dominated problems are considered. If a bending-dominated problem is considered, numerical locking occurs, resulting in spurious predictions for the membrane and/or the shear strain energy components.

4. Equilibrium equations

To develop the least-squares finite element model we will use the equilibrium equations associated with the shear-deformable model to define the least-squares functional. The equilibrium equations are easily obtained as the Euler equations associated with the minimization of the total potential energy, $\delta\mathcal{I} = \delta\mathcal{U} - \delta W_\Omega - \delta W_\Gamma = 0$, where

$$\delta\mathcal{U} = \int_\Omega \left(N^{\alpha\beta}\, \delta\varepsilon_{\alpha\beta} + M^{\alpha\beta}\, \delta\chi_{\alpha\beta} + Q^\alpha\, \delta\zeta_\alpha \right) \sqrt{a}\, d\xi^1\, d\xi^2\,, \tag{6.30}$$

$$\delta W_\Omega = \int_\Omega \left(p^\alpha \, \delta u_\alpha + p^3 \, \delta u_3 \right) \, \sqrt{a} \, d\xi^1 \, d\xi^2 \, , \tag{6.31}$$

$$\delta W_\Gamma = \int_\Gamma \left(\bar{N}^{\alpha\beta} \, n_\alpha \, \delta u_\beta + \bar{M}^{\alpha\beta} \, n_\alpha \, \delta\theta_\beta + \bar{Q}^\alpha \, n_\alpha \, \delta u_3 \right) \, d\,\Gamma \, . \tag{6.32}$$

Using Eqs. (6.10)-(6.12), integrating by parts, and setting each of coefficients associated with the variations of the generalized displacements equal to zero yields the equilibrium equations,

$$\delta u_\alpha : \qquad N^{\alpha\beta}_{|\beta} - \left(M^{\gamma\beta} \, b^\alpha_\beta \right)_{|\gamma} - Q^\gamma \, b^\alpha_\gamma + p^\alpha = 0 \tag{6.33}$$

$$\delta u_3 : \qquad Q^\alpha_{,\alpha} + N^{\alpha\beta} \, b_{\alpha\beta} - M^{\alpha\beta} \, b^\gamma_\alpha \, b_{\gamma\beta} + p^3 = 0 \tag{6.34}$$

$$\delta\theta_\alpha : \qquad M^{\alpha\beta}_{|\beta} - Q^\alpha = 0 \tag{6.35}$$

and the variationally consistent boundary conditions

$$u_\alpha = \bar{u}_\alpha \quad \text{or} \quad N^{\alpha\beta} \, n_\beta - M^{\gamma\beta} \, b^\alpha_\beta \, n_\gamma = \bar{N}^{\alpha\beta} \, n_\beta \tag{6.36}$$

$$u_3 = \bar{u}_3 \quad \text{or} \quad Q^\alpha \, n_\alpha = \bar{Q}^\alpha \, n_\alpha \tag{6.37}$$

$$\theta_\alpha = \bar{\theta}_\alpha \quad \text{or} \quad M^{\alpha\beta} \, n_\beta = \bar{M}^{\alpha\beta} \, n_\beta \tag{6.38}$$

Keeping in mind the "C^0 practicality" level, we realize that the equations governing the shear-deformable shell model, Eqs. (6.33)-(6.35) and Eqs. (6.16)-(6.18), are already of first-order. Hence, we need only construct the L_2 least-squares functional and minimize it with respect to the chosen approximating spaces to obtain the finite element model.

We are now in a position to emulate a variational setting where the shell assumes a state of deformation such that the L_2 norms of the equilibrium and constitutive equations are minimized. This variational framework is one associated with residual

minimization, in contrast to that presented in Section 3 where the variational framework is one associated with energy minimization. We have thus constructed a new variational setting for the solution of the shear-deformable shell model.

B. Numerical examples: verification benchmarks

In this section we present some numerical results obtained with the proposed formulation. In the present study we restrict our analysis to problems where the undeformed shell surface geometry is given analytically. Specifically, we consider circular cylindrical shells, where the shell mid-surface $\mathcal{S}$ is given by

$$\mathcal{S} = \{-L < x_1 < L\,,\; x_2^2 + x_3^2 = R^2 \,|\, (x_1, x_2, x_3) \in \mathbb{R}^3\} \subset \mathbb{R}^3\,, \tag{6.39}$$

where $2L$ and R are the length and radius of the shell. The shell mid-surface $\mathcal{S}$, given by Eq. (6.39), can be parametrized by the single chart $\vec{\phi} = (\phi_1, \phi_2, \phi_3)$, $\vec{\phi} : \bar{\Omega} \subset \mathbb{R}^2 \longrightarrow \mathcal{S} \subset \mathbb{R}^3$,

$$\begin{aligned} \phi_1(\xi^1, \xi^2) =& \xi^1 \\ \phi_2(\xi^1, \xi^2) =& R\,\sin(\xi^2/R) \\ \phi_3(\xi^1, \xi^2) =& R\,\cos(\xi^2/R) \end{aligned} \tag{6.40}$$

so that Ω is the rectangle occupying the region

$$\{(\xi^1, \xi^2) \in \Omega \,|\, -L < \xi^1 < L, -R\pi < \xi^2 < R\pi\} \subset \mathbb{R}^2\,. \tag{6.41}$$

For the circular cylindrical shell, the membrane, bending, and shear strain measures, given in Eqs. (6.10)-(6.12), become

$$\varepsilon_{11} = u_{1,1}\,, \quad 2\,\varepsilon_{12} = u_{1,2} + u_{2,1}\,, \quad \varepsilon_{22} = u_{2,2} + \frac{u_3}{R} \tag{6.42}$$

$$\chi_{11} = \theta_{1,1}\,, \quad 2\,\chi_{12} = \theta_{1,2} + \theta_{2,1} + \frac{u_{2,1}}{R}\,, \quad \chi_{22} = \theta_{2,2} + \frac{1}{R}\left(u_{2,2} + \frac{u_3}{R}\right) \tag{6.43}$$

$$\zeta_1 = u_{3,1} + \theta_1 \,, \quad \zeta_2 = u_{3,2} + \theta_2 - \frac{u_2}{R} \tag{6.44}$$

and the equilibrium equations, given in Eqs. (6.33)-(6.35), take the form

$$\delta u_1 : \qquad N^{11}_{,1} + N^{12}_{,2} + p^1 = 0 \tag{6.45}$$

$$\delta u_2 : \qquad N^{12}_{,1} + N^{22}_{,2} + \frac{M^{12}_{,1}}{R} + \frac{M^{22}_{,2}}{R} + \frac{Q^2}{R} + p^2 = 0 \tag{6.46}$$

$$\delta u_3 : \qquad Q^1_{,1} + Q^2_{,2} - \frac{N^{22}}{R} - \frac{M^{22}}{R^2} + p^3 = 0 \tag{6.47}$$

$$\delta \theta_1 : \qquad M^{11}_{,1} + M^{12}_{,2} - Q^1 = 0 \tag{6.48}$$

$$\delta \theta_2 : \qquad M^{12}_{,2} + M^{22}_{,2} - Q^2 = 0 \tag{6.49}$$

The least-squares formulation and finite element model follow from the procedures in Chapter II. Note that if we let $R \to \infty$ we recover the (linear) shear-deformable plate bending strain measures and governing equations (see Chapter V), where membrane and bending effects are decoupled.

1. Convergence

To verify exponentially fast decay of appropriate error measures for increasing p-levels we consider the following test problem. We prescribe a solution for the generalized displacement fields in the bi-unit square, $\bar{\Omega} = [-1, 1] \times [-1, 1]$, compute the consistent forcing functions and solve to verify convergence of the numerical solution to the exact solution. We choose the generalized displacement fields to be of the form

$$\begin{aligned} u_i(\xi^1, \xi^2) &= a_{u_i} + b_{u_i} \cos(\omega\, \xi^1) \sin(\omega\, \xi^2) \,, \\ \theta_1(\xi^1, \xi^2) &= a_{\theta_1} + b_{\theta_1} \cos(\omega\, \xi^1) \cos(\omega\, \xi^2) \,, \\ \theta_2(\xi^1, \xi^2) &= a_{\theta_2} + b_{\theta_2} \sin(\omega\, \xi^1) \sin(\omega\, \xi^2) \,. \end{aligned} \tag{6.50}$$

The exact solution is used to prescribe values of the generalized displacements on the entire boundary. The connected model, $\bar{\Omega}^h$, consists of a 4×4 finite element mesh. We consider uniform and distorted meshes. The choice of constants in Eq. (6.50) were as follows: $a_{u_i} = a_{\theta_1} = a_{\theta_2} = 0$, $\omega = 2\pi$, $b_{u_1} = 0.5$, $b_{u_2} = 0.4$, $b_{u_3} = 0.7$, $b_{\theta_1} = 0.2$, and $b_{\theta_2} = 0.4$.

In Figs. 66 and 67 we plot the L_2 least-squares functional ($\mathcal{J}$) and L_2 error of the generalized displacements $(\vec{u}, \underline{\theta})$ and stress resultants $(\mathbf{N}, \mathbf{M}, \underline{Q})$ as a function of the expansion order in a logarithmic-linear scale. Exponentially fast decay (spectral convergence) of the L_2 least-squares functional and L_2 errors is observed for both uniform and distorted meshes.

This test problem does not illustrate or attempts to measure the predictive capability of the shell elements, it simply verifies the computational algorithm and shows that for smooth solutions exponentially fast decay of error measures can be achieved. In practice, the response of the shell structure may include strong boundary layers at several scales [79, 43], depending on the boundary conditions and shell geometry. Hence, we may not achieve exponentially fast decay of error measures, at least when using uniform meshes (i.e., non-graded meshes). This point is illustrated in the next two example problems, where we consider the membrane- and bending-dominated benchmark problems suggested in Ref. [79].

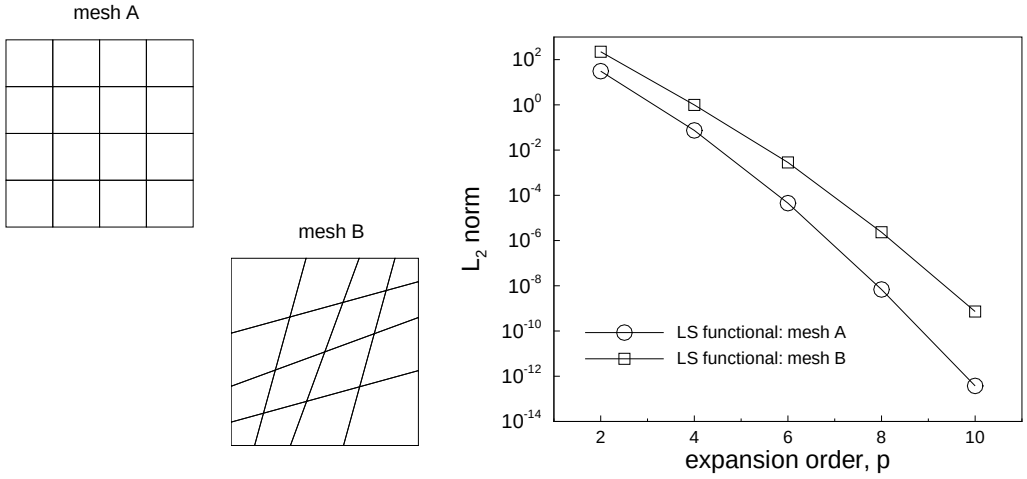

Fig. 66. Decay of the least-squares functional for uniform and distorted meshes.

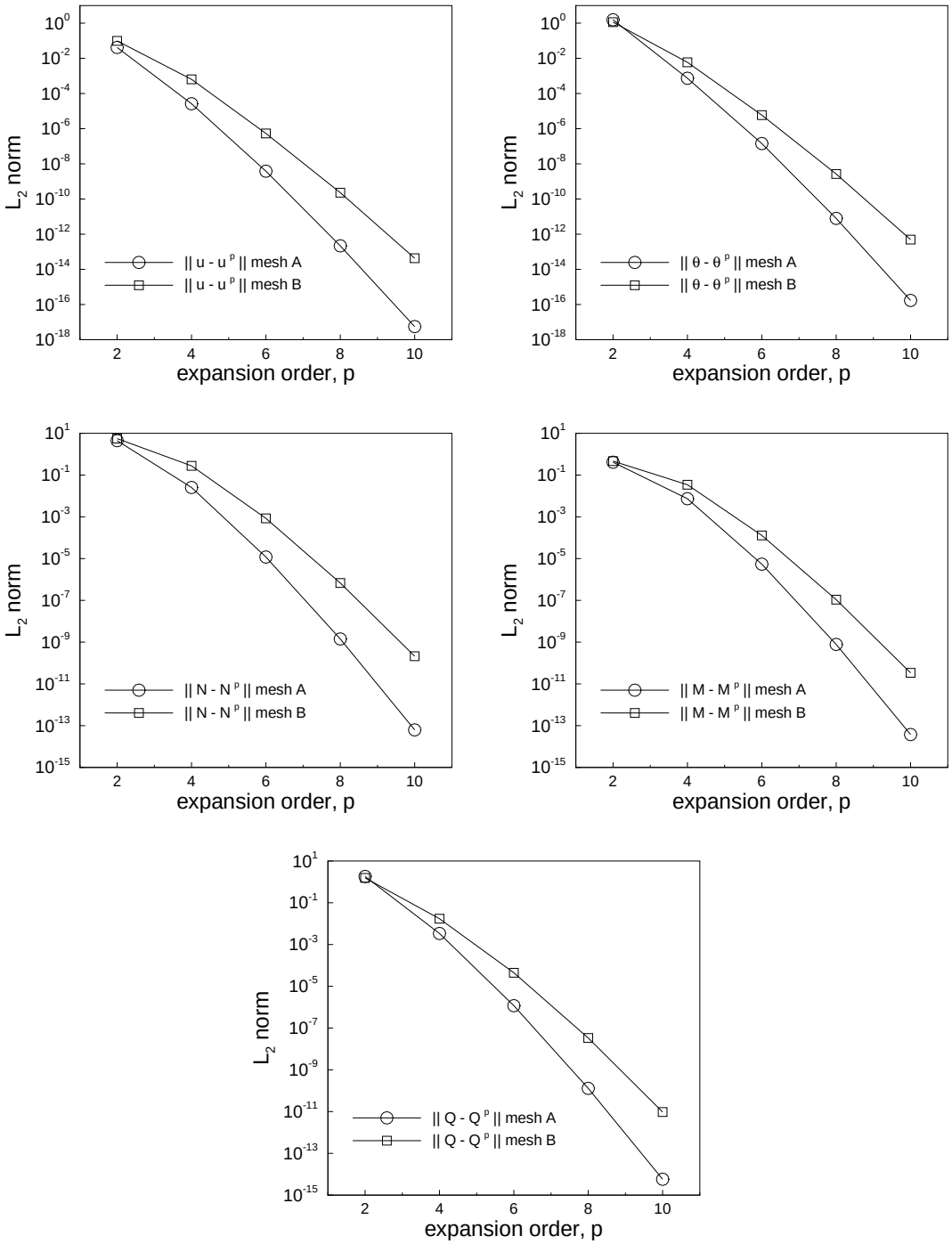

Fig. 67. Convergence of the generalized displacements and stress resultants for uniform and distorted meshes.

2. Clamped cylindrical shell

We consider a clamped circular cylindrical shell of uniform thickness t, length $2L$, and radius R with $L = R$. The shell is loaded by a self-balanced, axially-constant pressure distribution on its outer surface, given by

$$p^3(\xi^2) = p_0 \cos\left(2\,\xi^2/R\right) \,. \tag{6.51}$$

By symmetry considerations, the computational domain is limited to 1/16 of the total shell, so that

$$\Omega^h = \{(\xi^1, \xi^2) \,|\, 0 < \xi^1 < L \,,\, 0 < \xi^2 < \tfrac{\pi}{4}R\} \,. \tag{6.52}$$

The analytical solution for this problem, using the Naghdi shell theory, for any value of t is given in Ref. [79]. The shell is in a membrane-dominated state with the combined bending and shear energy components being less that 2% of the total energy for $R/t \geq 100$.

For the numerical simulation we use $R = L = 1$ m, $E = 2 \times 10^5$ MPa, $\nu = 1/3$, $K_s = 1$, and $p_0 = 1$ MPa. The connected model, $\bar{\Omega}^h$, consists of a 4×4 finite element mesh. We consider uniform and distorted meshes, shown in Fig. 68. By uniform mesh, we mean that the element sides are aligned with the principles directions of curvature. We systematically increase the p-level of the element approximation functions and monitor the decay of a suitable error measure for the case $R/t = 100$. For this problem boundary layers play a dominant role in a region of order $\sqrt{t}$ away from the clamped ends [79], thus we may not see a uniform convergence rate due to this non-smooth solution component.

The boundary conditions of the computational domain are specified as follows:

$$u_1 = 0 \,,\, \theta_1 = 0 \,,\, N^{12} = 0 \,,\, M^{12} = 0 \,,\, Q^1 = 0 \qquad \text{on } \xi^1 = 0 \,,\, \xi^2 = \left[0, \tfrac{\pi}{4}R\right] \quad \text{(sym.)} \,,$$

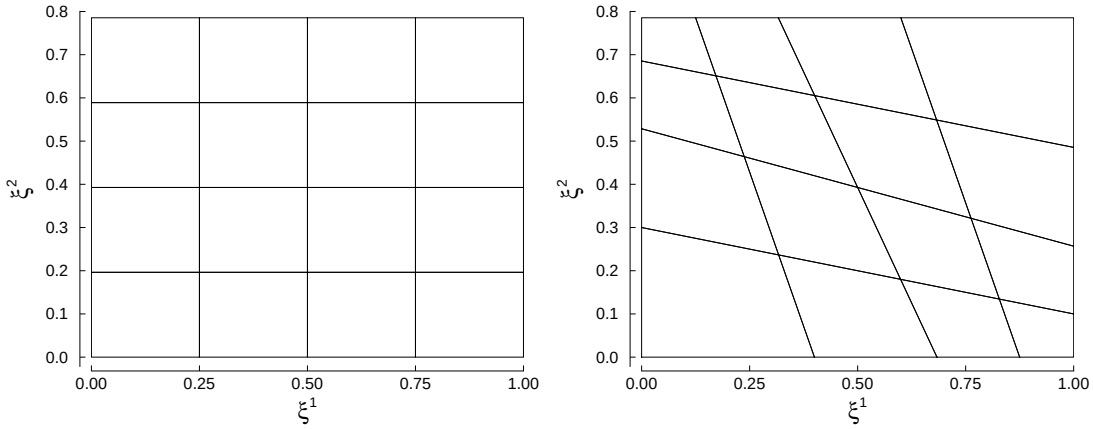

Fig. 68. Computational domain, using 4×4 uniform and distorted meshes, for the clamped and free cylindrical shell problems.

$$u_2 = 0\,,\ \theta_2 = 0\,,\ N^{12} = 0\,,\ M^{12} = 0\,,\ Q^2 = 0 \qquad \text{on } \xi^2 = 0\,,\ \xi^1 = [0, L] \qquad \text{(sym.)}\,,$$

$$u_1 = 0\,,\ u_2 = 0\,,\ u_3 = 0\,,\ \theta_1 = 0\,,\ \theta_2 = 0 \qquad \text{on } \xi^1 = L\,,\ \xi^2 = \left[0, \tfrac{\pi}{4}R\right] \qquad \text{(clamped)}\,,$$

$$u_2 = 0\,,\ u_3 = 0\,,\ N^{22} = 0\,,\ M^{22} = 0\,,\ M^{12} = 0 \qquad \text{on } \xi^2 = \tfrac{\pi}{4}R\,,\ \xi^1 = [0, L] \qquad \text{(anti-sym.)}\,.$$

Analytic values for the scaled strain energy of the shell are tabulated in Ref. [79], so we choose the error measure to be the relative error in the strain energy of the shell. Denoting by $\mathcal{U}$ the exact strain energy of the shell (see Section 2), the error measure is given by

$$E = \frac{|\mathcal{U} - \mathcal{U}^{hp}|}{\mathcal{U}} \tag{6.53}$$

We numerically solve the test problem using the proposed least-squares formulation and also using the displacement based Ritz-Galerkin formulation (see Section 3). Note that the Ritz-Galerkin formulation is based on a variational setting that naturally minimizes the chosen error measure, nevertheless we expect good performance from the least-squares formulation.

In Fig. 69 we plot the error measure E as a function of the expansion order in a logarithmic-linear scale, for both least-squares and Ritz-Galerkin formulations, using the uniform mesh. Despite the non-smooth solution component dominant near the clamped edge, good convergence is achieved by both formulations. No locking is observed and the Ritz-Galerkin formulation achieves a lower error in strain energy at any given p-level. In Fig. 70 we present the convergence results obtained using the distorted mesh. No dramatic changes in convergence rates are observed for either formulation.

Figure 71 shows the normalized radial displacement profiles at various p-levels, for both least-squares and Ritz-Galerkin formulations, using the uniform mesh and for the case $R/t = 1,000$. The normalized radial displacement is given by

$$\bar{u}_3 = u_3 \frac{D}{p_0 R t^2} , \tag{6.54}$$

with $D = E t^3/12(1-\nu^2)$. Slight oscillations appear near the clamped end due to the unresolved boundary layer. These oscillations disappear once the p-level is high enough to resolve the layer. The Ritz-Galerkin formulation is able to resolve the boundary layer at a lower p-level. Proper mesh grading near the clamped end would be effective to resolve the boundary layer at lower p-levels.

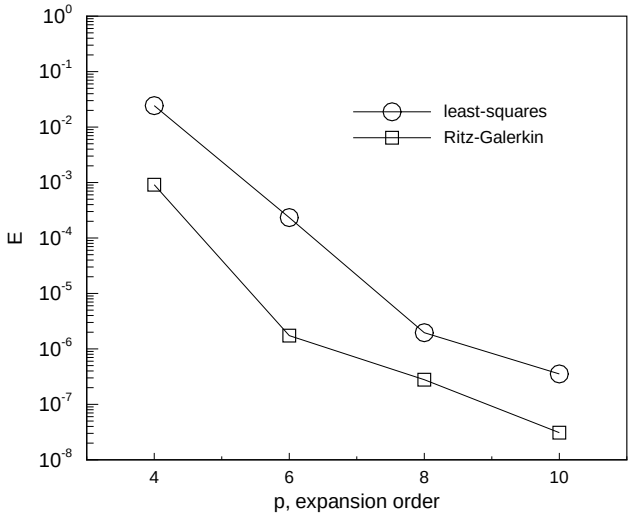

Fig. 69. Convergence of strain energy for the clamped cylindrical shell using the least-squares and Ritz-Galerkin formulations. 4×4 uniform mesh, $R/t = 100$.

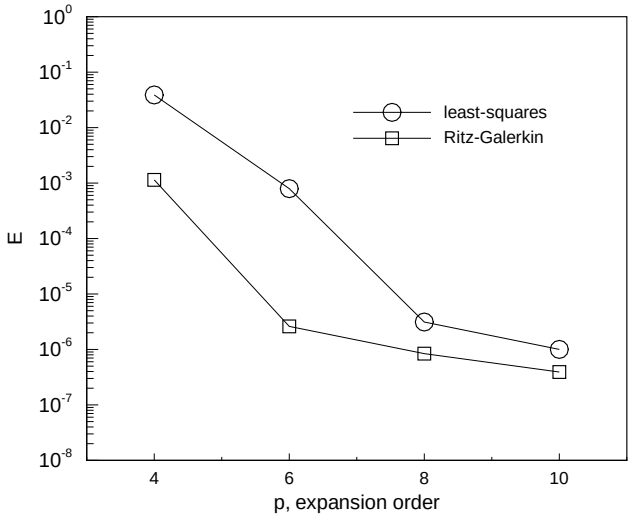

Fig. 70. Convergence of strain energy for the clamped cylindrical shell using the least-squares and Ritz-Galerkin formulations. 4×4 distorted mesh, $R/t = 100$.

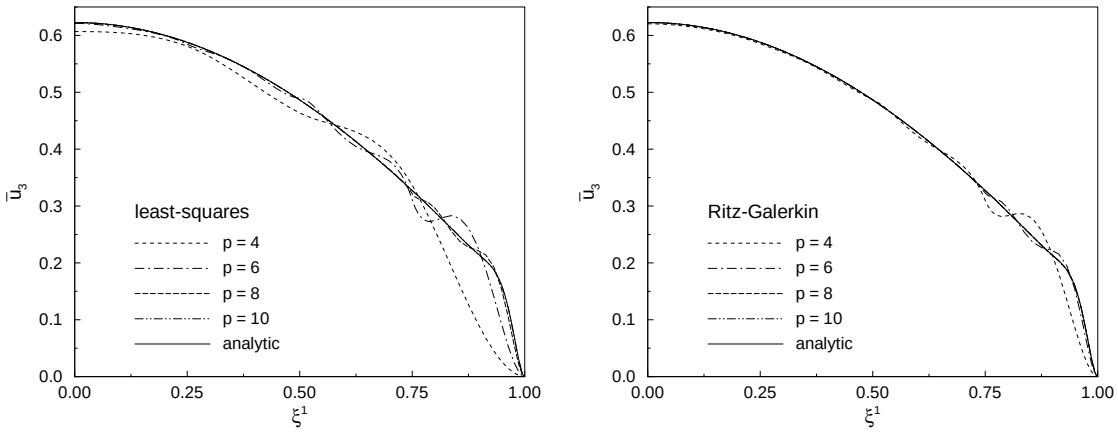

Fig. 71. Convergence of the normalized radial displacement for the clamped cylindrical shell using the least-squares and Ritz-Galerkin formulations. 4×4 uniform mesh, $R/t = 1,000$.

3. Free cylindrical shell

The geometry and loading for this test problem are the same as described for the clamped circular cylindrical shell. Like before, by symmetry considerations, the computational domain is limited to 1/16 of the total shell. The shell is in a bending-dominated state with the membrane and shear energy components being essentially negligible even for the moderately thick case of $R/t = 10$. The analytical solution for this problem, using the Naghdi shell theory, for any value of t is given in Ref. [79]. We systematically increase the p-level of the element approximation functions and monitor the convergence in strain energy for the case $R/t = 100$ using the uniform and distorted meshes, shown in Fig. 68.

The boundary conditions of the computational domain are the same as those specified for the previous problem, except that the clamped ends are now left free:

$$N^{11} = 0\,,\; N^{12} = 0\,,\; M^{11} = 0\,,\; M^{12} = 0\,,\; Q^{1} = 0 \quad \text{on } \xi^1 = L\,,\; \xi^2 = \left[0, \tfrac{\pi}{4}R\right] \quad \text{(free)}\,.$$

In Fig. 72 we present the convergence results obtained using the uniform mesh. For the purpose of comparison we also present results obtained with the displacement based Ritz-Galerkin formulation. The asymptotic convergence rate of the Ritz-Gakerkin formulation in the bending-dominated case is much lower than that seen for the membrane-dominated case, even though the initial error at a p-level of 4 is lower for the bending-dominated case. The least-squares formulation displays a good convergence rate in strain energy and asymptotically seems to achieve a slightly better convergence rate than the Ritz-Gakerkin formulation. In Fig. 73 we present the convergence results obtained using the distorted mesh. No dramatic changes in convergence rates or behavior are observed for the formulations.

Regarding the bending- and membrane-dominated shell problems presented in

this and the previous example we see from the convergence in strain energy curves that the least-squares solution, for the finite element meshes considered, is sufficiently accurate only for p-levels of 6 or higher. This was also the case for the smooth solution example presented in Section 1. Better accuracy at lower p-levels (e.g., $p = 1$ or 2) can be achieved by using reduced integration techniques to force a least-squares collocation solution. Or, an alternate least-squares functional that defines an equivalent norm in H^1 could be defined. Unfortunately, such an approach may yield a least-squares functional in terms of computationally impractical norms which at some point must be replaced by computable equivalents. Here we have chosen to define the least-squares functional in terms of L_2 norms only, which are readily and easily computable.

Having shown results using the displacement based Ritz-Galerkin formulation for this and the previous example, we note that when considering the computational cost associated with each of the formulations, for a given p-level and finite element mesh, the least-squares formulation is the most expensive; treating the generalized displacements $(\vec{u}, \underline{\theta})$ and stress resultants $(\mathbf{N}, \mathbf{M}, \underline{Q})$ as independent variables. A cost comparison between the least-squares and displacement based Ritz-Galerkin formulation is illustrated in Fig. 74, by plotting the convergence in strain energy for the free cylindrical shell as a function of the the total number of degrees of freedom in a logarithmic-linear scale. Clearly, the Ritz-Galerkin formulation is the most cost efficient allowing for more accurate solutions at higher p-levels at a lower computational cost. Nevertheless, the computational cost of the least-squares formulation could be reduced to that of the displacement based formulation by choosing an appropriate basis for the stress resultants so that these degrees of freedom can be condensed out at the element level. This approach is currently under investigation.

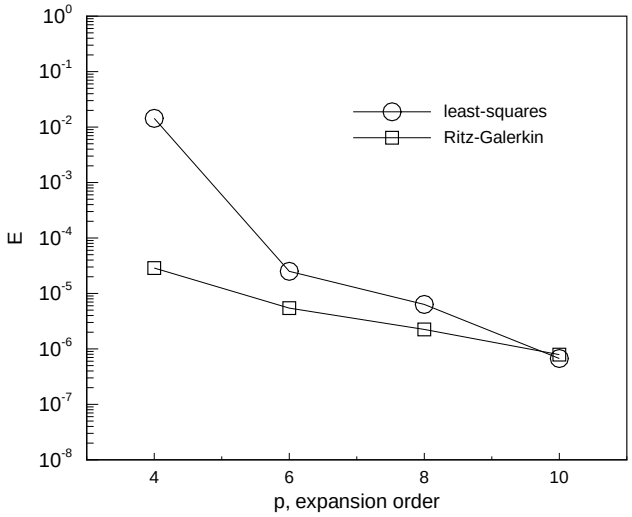

Fig. 72. Convergence of strain energy for the free cylindrical shell using the least-squares and Ritz-Galerkin formulations. 4×4 uniform mesh, $R/t = 100$.

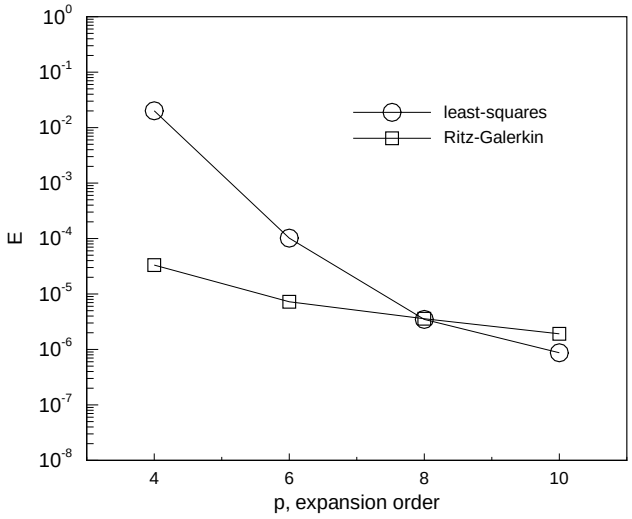

Fig. 73. Convergence of strain energy for the free cylindrical shell using the least-squares and Ritz-Galerkin formulations. 4×4 distorted mesh, $R/t = 100$.

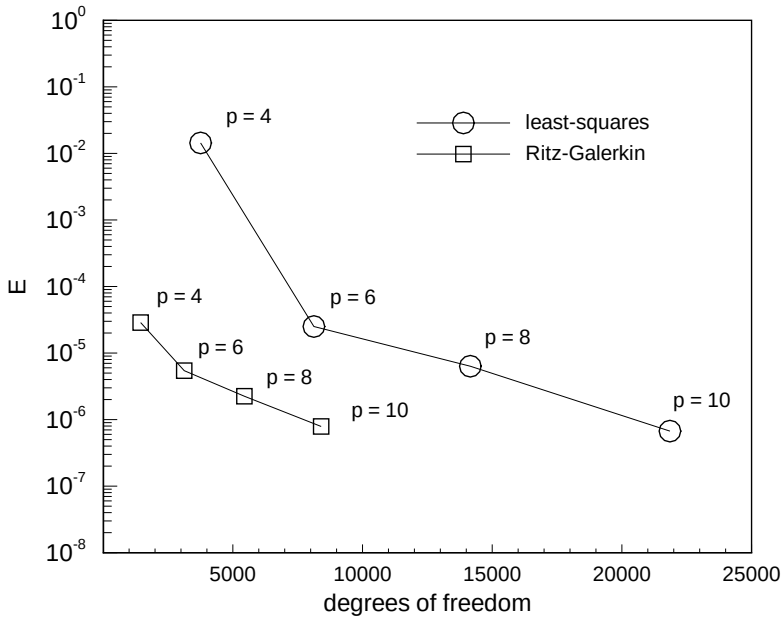

Fig. 74. A cost comparison between the least-squares and displacement based Ritz-Galerkin formulations: convergence of strain energy for the free cylindrical shell as a function of the the total number of degrees of freedom. 4×4 uniform mesh, $R/t = 100$.

4. Barrel vault

We consider a barrel vault loaded by its own weight. The barrel vault is a segment of a circular cylindrical shell whose mid-surface, after being parametrized by (6.40), is given by

$$\Omega = \{(\xi^1, \xi^2) \,|\, -L < \xi^1 < L \,,\, -\tfrac{2\pi}{9}R < \xi^2 < \tfrac{2\pi}{9}R\} \,. \tag{6.55}$$

The barrel vault is simply-supported on rigid diaphragms on opposite edges and is free on the other two edges. For the described loading, geometry, and boundary conditions, the problem is popularly known as the Scordelis-Lo roof.

The boundary conditions and geometry of the shell imply that, in the limit of vanishing thickness, this problem is a membrane-dominated problem [22]. However, the self-weight loading fails to meet certain regularity requirements that make the limit problem ill-posed; the limit problem being in a mixed state with the bending and membrane strain energy components asymptotically concentrated near the free edge [68]. Here we only consider the case $R/t = 100$, for which the limiting ill-posed behavior is not yet dominant.

By symmetry considerations, the computational domain is limited to 1/4 of the total shell, so that

$$\Omega^h = \{(\xi^1, \xi^2) \,|\, 0 < \xi^1 < L \,,\, 0 < \xi^2 < \tfrac{2\pi}{9}R\} \,. \tag{6.56}$$

The geometry of the barrel vault is specified as follows: $2L = 50$ ft, $R = 25$ ft, and $t = 3$ in.. The material is homogeneous and isotropic with $E = 3 \times 10^6$ psi and $\nu = 0$. The shear correction factor K_s is specified as 5/6 and the self-weight loading as $p_z = 90$ lb/ft^2 uniformly distributed over the surface area of the vault. This specific geometry, loading, and material data (but with $\nu = 0.3$) first appeared in the work of Cantin and Clough [21], who used Scordelis and Lo's [97] computer algorithm to verify their results.

The connected model, $\bar{\Omega}^h \subset \mathbb{R}^2$, consists of a 4×4 finite element mesh, shown in Fig. 75. For illustrative purposes we present in Fig. 76 the finite element mesh on the entire mid-surface of the barrel vault, $\mathcal{S} \subset \mathbb{R}^3$. The mesh is regular (i.e., not distorted) and graded. We expect strong boundary layers in the stress resultant profiles along the free and supported edges, so the mesh is graded towards those regions. In our experience with the mixed least-squares formulation, boundary layers in generalized displacements and stress resultants need to be adequately resolved to obtain an accurate prediction of the shell response. This justifies the use of the graded mesh.

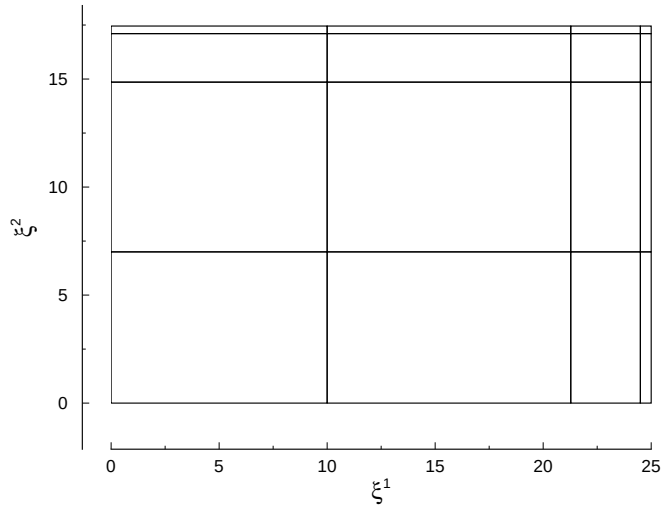

Fig. 75. Quarter shell computational domain for the analysis of the barrel vault. 4×4 graded mesh.

The boundary conditions of the computational domain are specified as

$$u_1 = 0\,,\ \theta_1 = 0\,,\ N^{12} = 0\,,\ M^{12} = 0\,,\ Q^1 = 0 \qquad \text{on } \xi^1 = 0\,,\ \xi^2 = \left[0, \tfrac{2\pi}{9} R\right] \quad \text{(sym.)}$$

$$u_2 = 0\,,\ \theta_2 = 0\,,\ N^{12} = 0\,,\ M^{12} = 0\,,\ Q^2 = 0 \qquad \text{on } \xi^2 = 0\,,\ \xi^1 = [0, L] \quad \text{(sym.)}$$

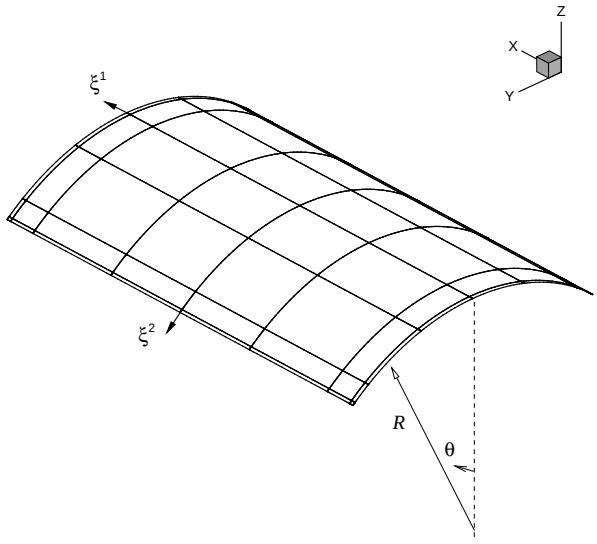

Fig. 76. Finite element mesh on the entire mid-surface of the barrel vault, $\mathcal{S} \subset \mathbb{R}^3$, showing the surface coordinate system $(\xi^1, \xi^2) \in \mathbb{R}^2$.

$$u_2 = 0\,,\; u_3 = 0\,,\; \theta_2 = 0\,,\; N^{11} = 0\,,\; M^{11} = 0 \qquad \text{on } \xi^1 = L\,,\; \xi^2 = \left[0, \tfrac{2\pi}{9}R\right] \quad \text{(supp.)}$$

$$N^{22} = 0\,,\; N^{12} = 0\,,\; M^{22} = 0\,,\; M^{12} = 0\,,\; Q^2 = 0 \quad \text{on } \xi^2 = \tfrac{2\pi}{9}R\,,\; \xi^1 = [0, L] \quad \text{(free)}$$

and the (vertical) pressure distribution is decomposed into its contravariant components, so that $p^2 = p_z\,\sin(\xi^2/R)$ and $p^3 = -p_z\,\cos(\xi^2/R)$.

First, we present a convergence study in strain energy for increasing p-levels of the element approximation functions. An analytic value for the strain energy is not available, so we use instead a reference value. The reference value was obtained with a p-level of 12 using the displacement based Ritz-Galerkin formulation in the graded finite element mesh (shown in Fig. 75). Denoting by $\mathcal{U}^{\text{ref}}$ the reference strain energy of the barrel vault, the error measure is given by

$$E = \frac{|\,\mathcal{U}^{\text{ref}} - \mathcal{U}^{hp}\,|}{\mathcal{U}^{\text{ref}}}\,. \tag{6.57}$$

In Fig. 77 we plot the error measure E for the least-squares formulation as a function of the expansion order in a logarithmic-linear scale. We see from the figure that the asymptotic exponential convergence rate is achieved for p-levels of 6 and higher. In the context of the least-squares formulation, this does not imply that the finite element solution for a p-level of 4 is membrane-locked. It simply means that, for this specific problem and the given finite element mesh, a p-level of 4 is not yet sufficient to yield a satisfactory least-squares solution. The point is further illustrated by Fig. 78, where we plot the vertical displacement and shear-stress resultant profiles along the mid-section of the vault for increasing p-levels. From Fig. 78 we see that the finite element solutions at p-levels of 6, 8, and 10 are essentially the same and fall on top of each other.

Table XIV shows results for the vertical displacement and stress resultants at the center of the free edge of the barrel vault, $(\xi^1, \xi^2) = (0, \frac{2\pi}{9}R)$, for p-levels of 4, 6, 8, and 10. Similarly, in Table XV we present results for the vertical displacement and stress resultants at the crown of the barrel vault, $(\xi^1, \xi^2) = (0, 0)$. We see from the tabulated data that a p-level of 4 is not sufficient to yield a satisfactory least-squares solution and that a p-level of 6 or 8 is already sufficient. The predicted vertical deflection at the center of the free edge (Table XIV) is in good agreement with the shallow shell analytical value of 3.7032 in. and the commonly used reference value for finite element analysis of 3.6288 in..

Figure 79 shows the vertical displacement and non-zero stress resultant profiles along the mid-section of the vault for a p-level of 10. Similarly, Fig. 80 shows the axial displacement and non-zero stress resultant profiles along the simply-supported section of the vault. Strong boundary layers are observed for the bending- and shear-stress resultants along this edge.

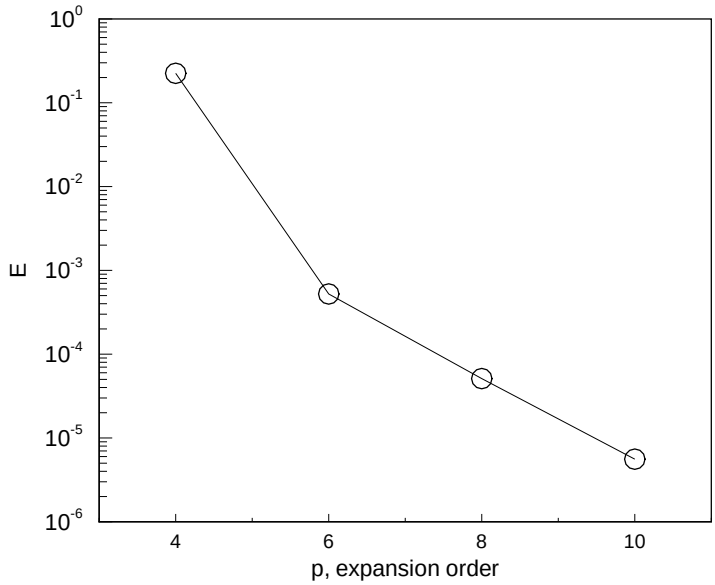

Fig. 77. Convergence of strain energy for the barrel vault problem. Least-squares formulation, 4×4 graded mesh.

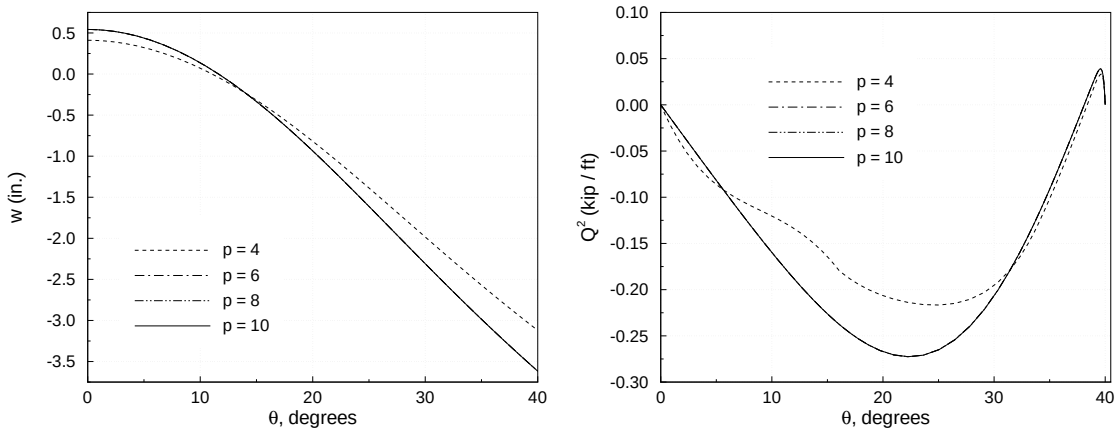

Fig. 78. Vertical displacement and shear stress-resultant profiles along the mid-section of the vault for increasing p-levels. Least-squares formulation, 4×4 graded mesh.

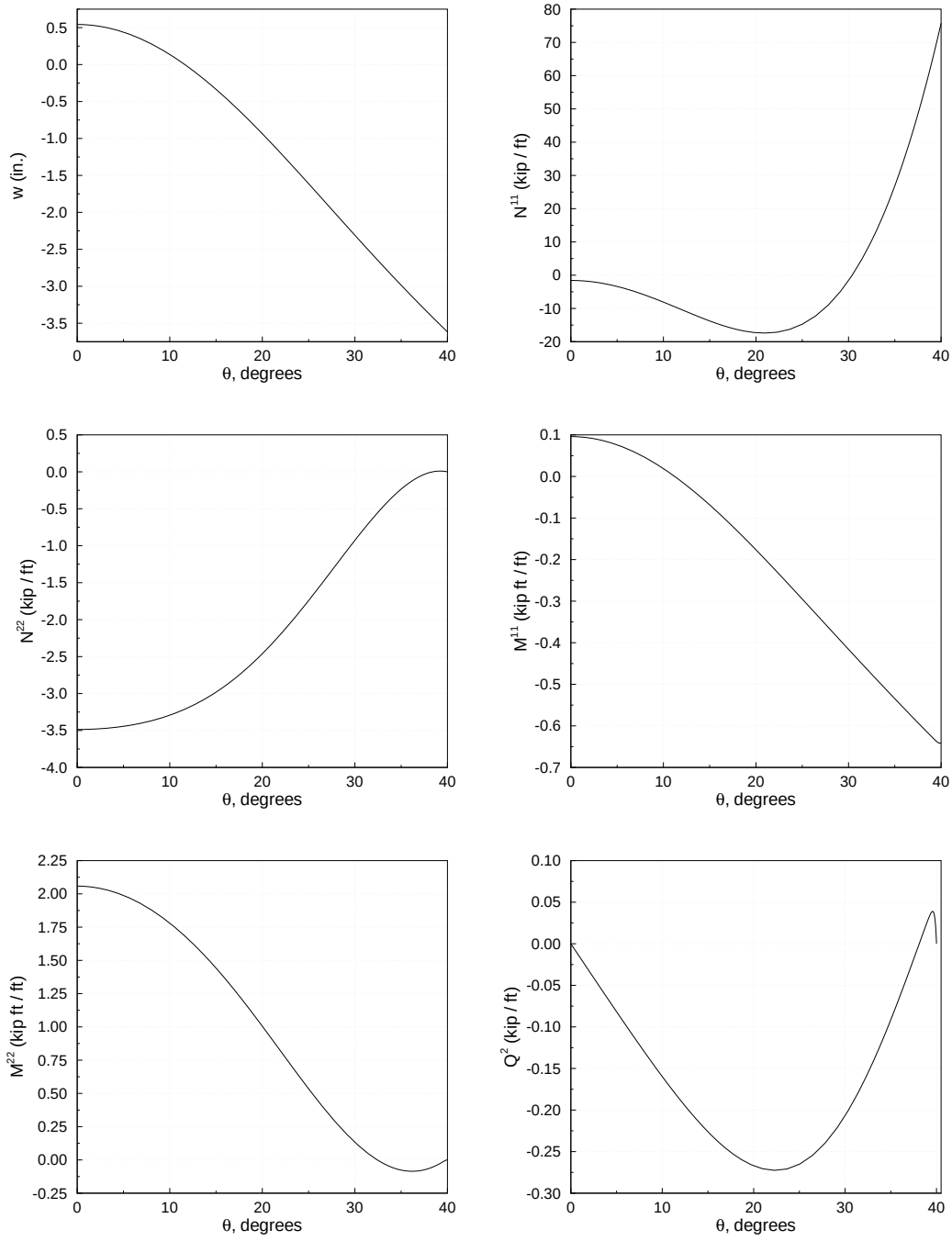

Fig. 79. Vertical displacement and stress resultant profiles along the central section of the vault. Least-squares formulation, p-level of 10, 4×4 graded mesh.

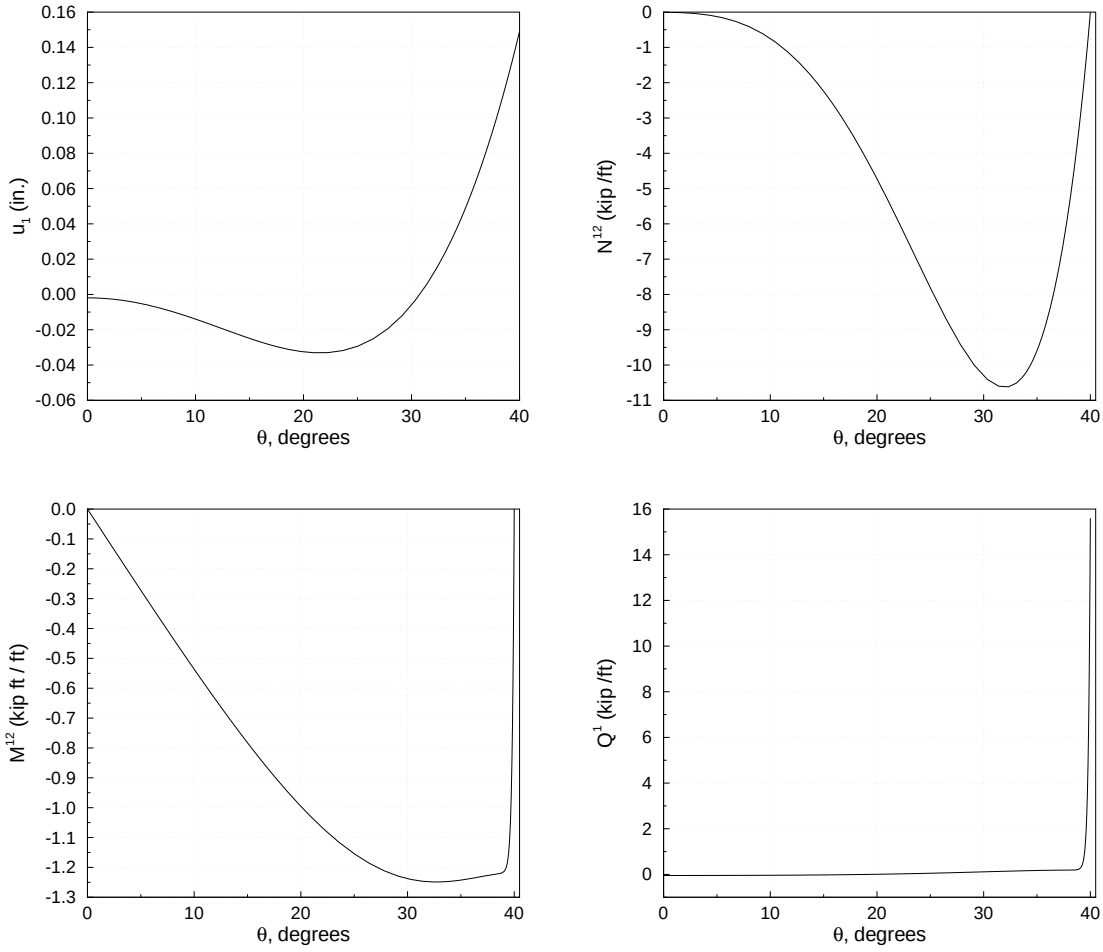

Fig. 80. Axial displacement and stress resultant profiles along the support section of the vault. Least-squares formulation, p-level of 10, 4×4 graded mesh.

Table XIV. p-convergence study showing vertical displacement and stress resultants at the center of the free edge of the barrel vault.

p level	w (in.)	N^{11} (kip/ft)	M^{11} (kip ft/ft)
4	-3.1208	68.3942	-0.5610
6	-3.6162	75.7476	-0.6400
8	-3.6173	75.7582	-0.6400
10	-3.6174	75.7593	-0.6400

Table XV. p-convergence study showing vertical displacement and stress resultants at the crown of the barrel vault.

p level	w (in.)	N^{11} (kip/ft)	N^{22} (kip/ft)	M^{11} (kip ft/ft)	M^{22} (kip ft/ft)
4	0.4109	-3.5870	-3.4148	0.0714	1.7597
6	0.5423	-1.5835	-3.4861	0.0959	2.0579
8	0.5425	-1.5805	-3.4862	0.0959	2.0583
10	0.5425	-1.5802	-3.4862	0.0959	2.0583

5. Pinched cylinder

We consider a circular cylindrical shell of uniform thickness t, length $2L$, and radius R with $L = R$. The shell is simply-supported on rigid diaphragms at each end and loaded by two opposed forces of equal magnitude acting on the shell's mid-section.

The boundary conditions and geometry of the shell imply that, in the limit of vanishing thickness, this problem is a membrane-dominated problem [22]. Typically the loading is specified as a transverse concentrated load (a point load), which fails to meet certain regularity requirements making the limit problem ill-posed. Furthermore, it has been established that the transverse displacement under the concentrated load grows unbounded when p-type refinements are performed to analyze the problem using a shear-deformable shell model [43], whereas low p-levels (and h-refinement) deceptively yield finite values. Even though we do not consider the limit problem in this study, we have modified the loading so that the applied load is not a point load, but a distributed pressure confined to a circle whose radius is of the order of the thickness of the shell. This eliminates all the singularities associated with the point load and closely emulates a point load.

By symmetry considerations, the computational domain is limited to 1/8 of the total shell, so that

$$\Omega^h = \{(\xi^1, \xi^2) \,|\, 0 < \xi^1 < L \,,\, 0 < \xi^2 < \tfrac{\pi}{2} R\} \,. \tag{6.58}$$

The geometry of the cylindrical shell is specified as follows: $2L = 600$ in., $R = 300$ in., and $t = 3$ in. so that $R/t = 100$. The material is homogeneous and isotropic with $E = 3 \times 10^6$ psi and $\nu = 0.3$. The shear correction factor K_s is specified as $5/6$ and the resultant of the distributed pressure as 1 lb (or $P = 0.25$ lb in the one-eight model).

The connected model, $\bar{\Omega}^h$, consists of a 6×6 finite element mesh, shown in Fig. 81. The mesh is regular (i.e., not distorted) and graded. The mesh is graded towards the region where the distributed pressure is applied, this is needed to appropriately resolve the imposed pressure distribution. For illustrative purposes we present in Fig. 82 the finite element mesh on one-eight of the mid-surface, $\mathcal{S} \subset \mathbb{R}^3$, of the cylinder.

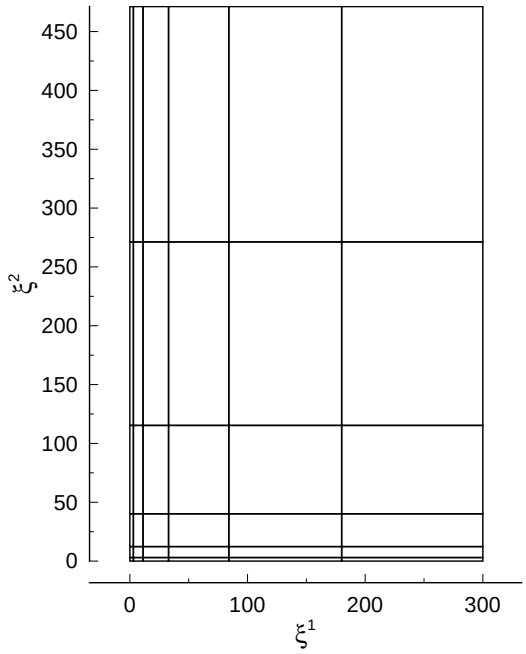

Fig. 81. One-eight shell computational domain for the analysis of the pinched cylinder. 6×6 graded mesh.

In the (ξ^1, ξ^2) space the (vertical) pressure distribution is taken to vary smoothly in the form of a sine curve, defined here as follows

$$p_z(r;\delta) = \begin{cases} \frac{C}{2}\left\{1 - \sin\left[\frac{\pi}{\delta}\left(r - \frac{\delta}{2}\right)\right]\right\} & \text{if } 0 \leq r < \delta, \\ 0 & \text{otherwise,} \end{cases} \tag{6.59}$$

where $r^2 = (\xi^1)^2 + (\xi^2)^2$ and δ defines the size of the domain of influence of the

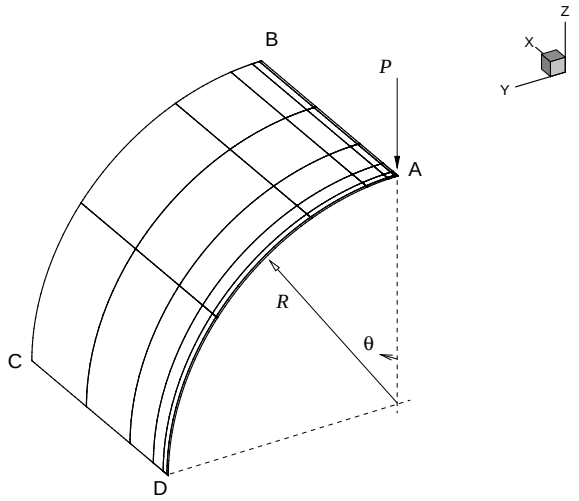

Fig. 82. Finite element mesh on one-eight of the mid-surface of the cylinder, $\mathcal{S} \subset \mathbb{R}^3$.

pressure distribution. To emulate the point load we take $\delta = t$. The constant C is chosen such that $P = \int_{\Omega^h} p_z \, d\Omega^h = 0.25$, which gives $C = 2\pi/\delta^2(\pi^2 - 4)$.

The boundary conditions of the computational domain are specified as

$$u_1 = 0\,,\ \theta_1 = 0\,,\ N^{12} = 0\,,\ M^{12} = 0\,,\ Q^1 = 0 \quad \text{on } \xi^1 = 0\,,\ \xi^2 = \left[0, \tfrac{\pi}{2}R\right] \quad \text{(sym.)}$$

$$u_2 = 0\,,\ \theta_2 = 0\,,\ N^{12} = 0\,,\ M^{12} = 0\,,\ Q^2 = 0 \quad \text{on } \xi^2 = 0\,,\ \xi^1 = [0, L] \quad \text{(sym.)}$$

$$u_2 = 0\,,\ u_3 = 0\,,\ \theta_2 = 0\,,\ N^{11} = 0\,,\ M^{11} = 0 \quad \text{on } \xi^1 = L\,,\ \xi^2 = \left[0, \tfrac{\pi}{2}R\right] \quad \text{(supp.)}$$

$$u_2 = 0\,,\ \theta_2 = 0\,,\ N^{12} = 0\,,\ M^{12} = 0\,,\ Q^2 = 0 \quad \text{on } \xi^2 = \tfrac{\pi}{2}R\,,\ \xi^1 = [0, L] \quad \text{(sym.)}$$

and the (vertical) pressure distribution is decomposed into its contravariant components, so that $p^2 = p_z \sin(\xi^2/R)$ and $p^3 = -p_z \cos(\xi^2/R)$.

In Fig. 83 we present a convergence study in strain energy for increasing p-levels of the element approximation functions. The error measure is given by Eq. (6.57), where the reference value for the strain energy was obtained using the displacement based

Ritz-Galerkin formulation in the graded finite element mesh (shown in Fig. 81) with a p-level of 12. Fig. 83 shows the error measure E for the least-squares formulation as a function of the expansion order in a logarithmic-linear scale. We infer from the convergence study that an acceptable least-squares solution is achieved for p-levels of 6 and higher.

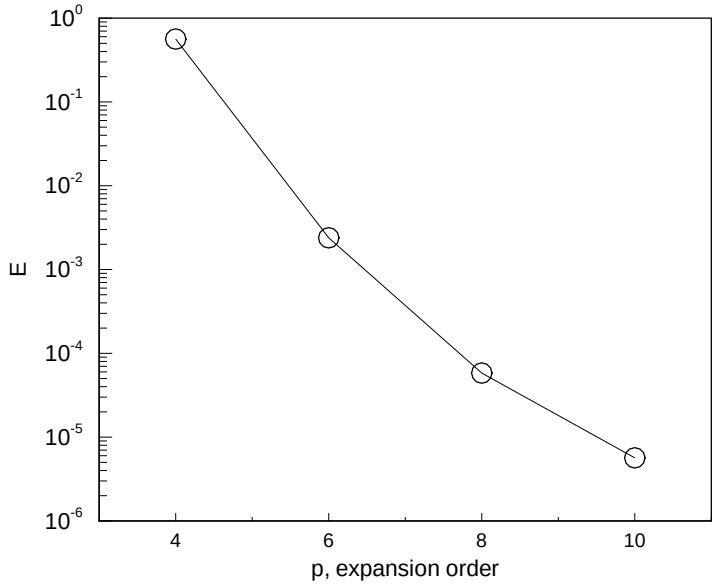

Fig. 83. Convergence of strain energy for the pinched cylinder problem. Least-squares formulation, 6×6 graded mesh.

As noted earlier, in the context of the least-squares formulation, this does not imply that the finite element solution for a p-level of 4 is membrane-locked. It simply means that, for this problem and the given finite element mesh, a p-level of 4 is not yet sufficient to yield a satisfactory least-squares solution.

Figure 78 shows the vertical displacement and bending-stress resultant profiles along the mid-section (arc AD) of the pinched cylinder for increasing p-levels. From

the figure we see that the finite element solutions at p-levels of 6, 8, and 10 are essentially indistinguishable from each other.

In Table XVI we present a p-convergence study for the vertical displacement and stress resultants at point A (see Fig. 82) of the pinched cylinder. The predicted vertical displacement is in good agreement with the Kirchhoff/Koiter model analytic value of 1.8248×10^{-5} in.. Table XVII shows a p-convergence study for the axial displacement at point B and radial displacement at point D of the pinched cylinder. The predicted displacements are in good agreement with the analytic Kirchhoff/Koiter model values of 4.5711×10^{-7} in. and 5.2222×10^{-8} in., respectively. By far the most challenging convergence study is that of the radial displacement at point D (see Ref. [67]), which is rarely reported in presentations of this example in the literature.

Figure 85 shows the vertical displacement and non-zero stress resultant profiles along the mid-section of the cylinder (arc AD) for a p-level of 10. Similarly, Fig. 86 shows the axial displacement and non-zero stress resultant profiles along the simply-supported section (arc BC) of the cylinder. Figure 87 shows the vertical displacement profile along the section AB and radial displacement profile along section DC of the cylinder. The deformed mid-surface of the pinched cylinder, magnified by a factor of 5×10^6, is shown in Fig. 88.

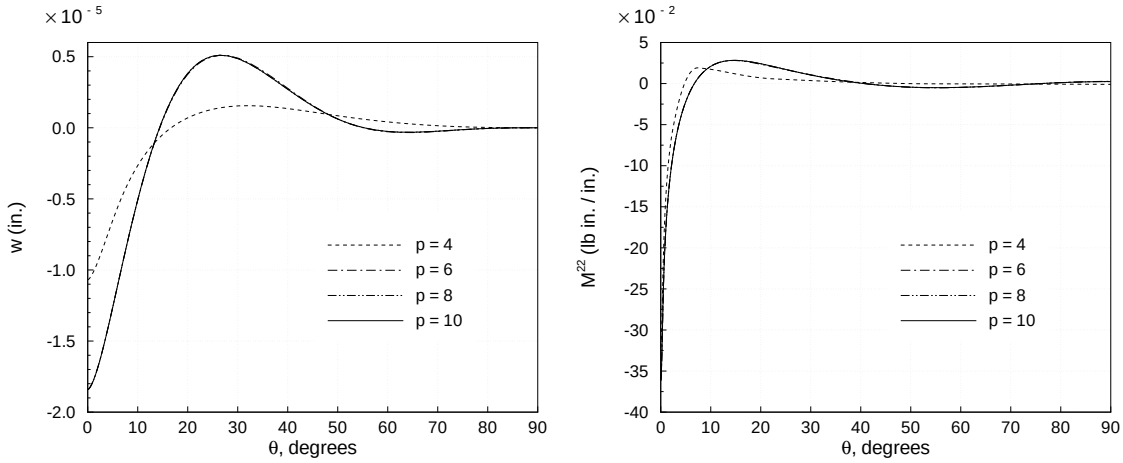

Fig. 84. Vertical displacement and bending stress-resultant profiles along the mid-section of the pinched cylinder for increasing p-levels. Least-squares formulation, 6×6 graded mesh.

Table XVI. p-convergence study showing vertical displacement and stress resultants at point A (see Fig. 82) of the pinched cylinder.

p level	w (in.) $\times 10^{-5}$	N^{11} (lb/in.) $\times 10^{-2}$	N^{22} (lb/in.) $\times 10^{-2}$	M^{11} (lb in./in.) $\times 10^{-2}$	M^{22} (lb in./in.) $\times 10^{-2}$
4	-1.0680	-4.1254	-6.6065	-30.7228	-32.1089
6	-1.8384	-5.7882	-6.7883	-32.3984	-36.1080
8	-1.8408	-5.7853	-6.7850	-32.4088	-36.1391
10	-1.8408	-5.7853	-6.7850	-32.4085	-36.1389

Table XVII. p-convergence study showing axial displacement at point B and radial displacement at point D (see Fig. 82) of the pinched cylinder.

p level	u_1 (in.) $\times 10^{-7}$	u_3 (in.) $\times 10^{-8}$
4	-2.1549	-91.2305
6	-4.5866	1.0132
8	-4.5787	-4.8651
10	-4.5787	-4.9439

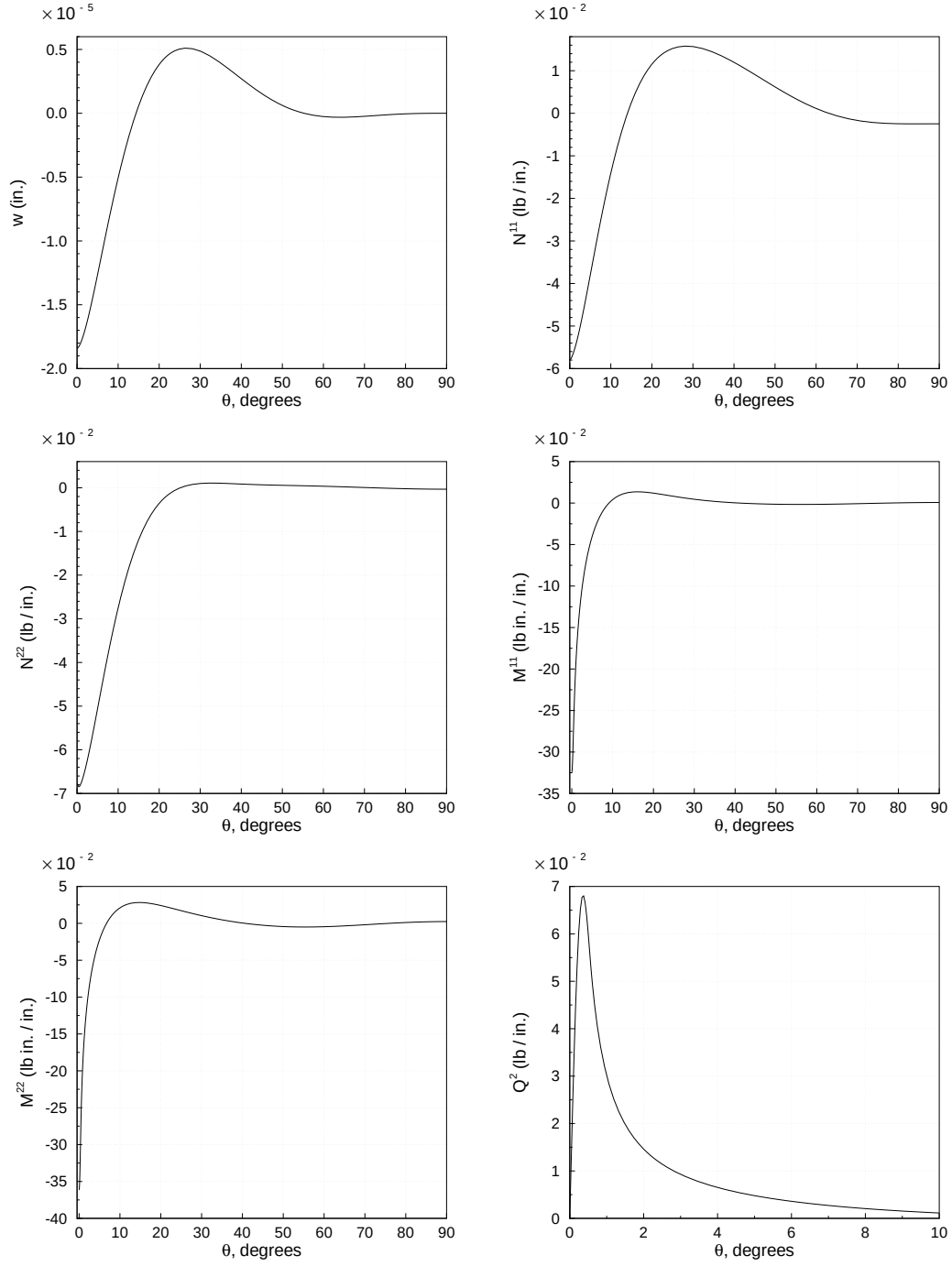

Fig. 85. Vertical displacement and stress resultant profiles along the mid-section (arc AD) of the pinched cylinder. Least-squares formulation, p-level of 10, 6×6 graded mesh.

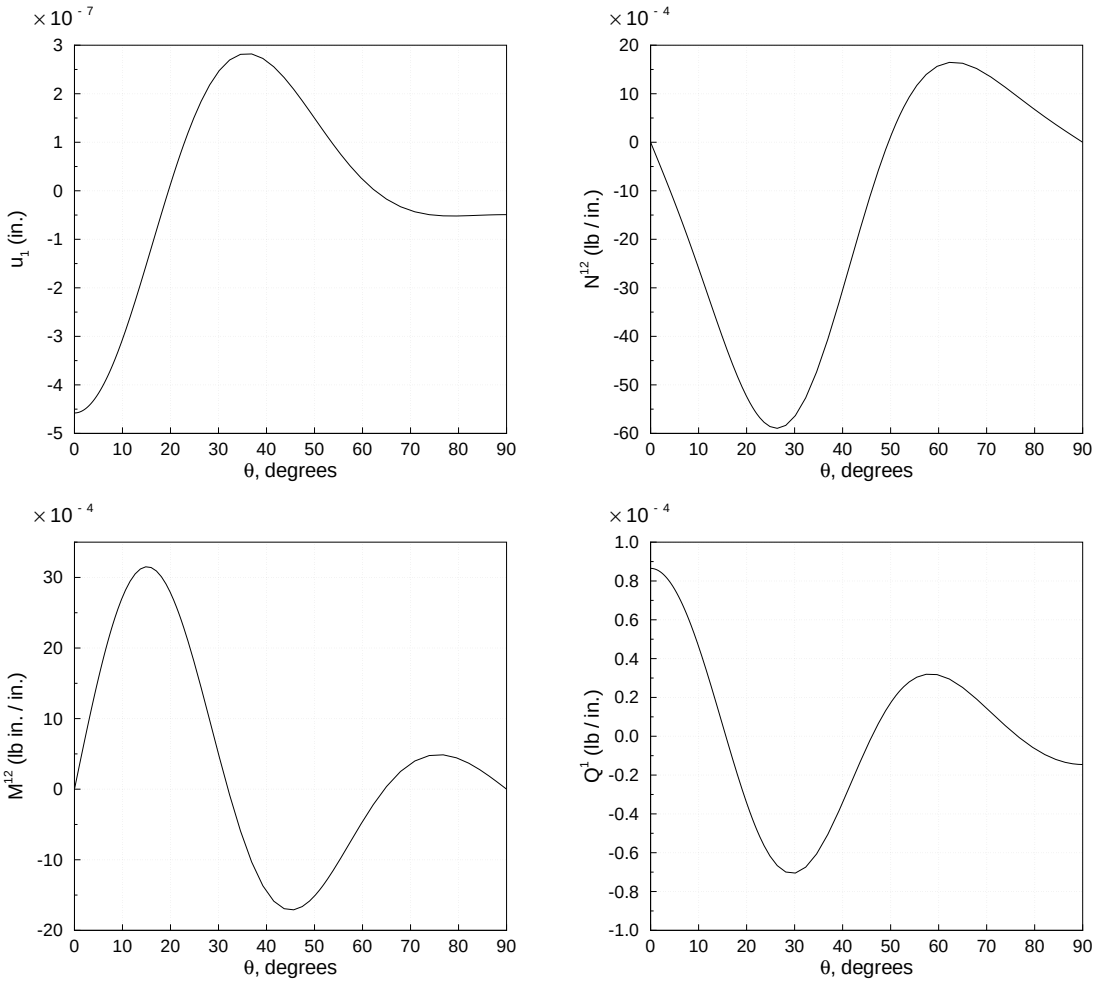

Fig. 86. Axial displacement and stress resultant profiles along the support section (arc BC) of the pinched cylinder. Least-squares formulation, p-level of 10, 6×6 graded mesh.

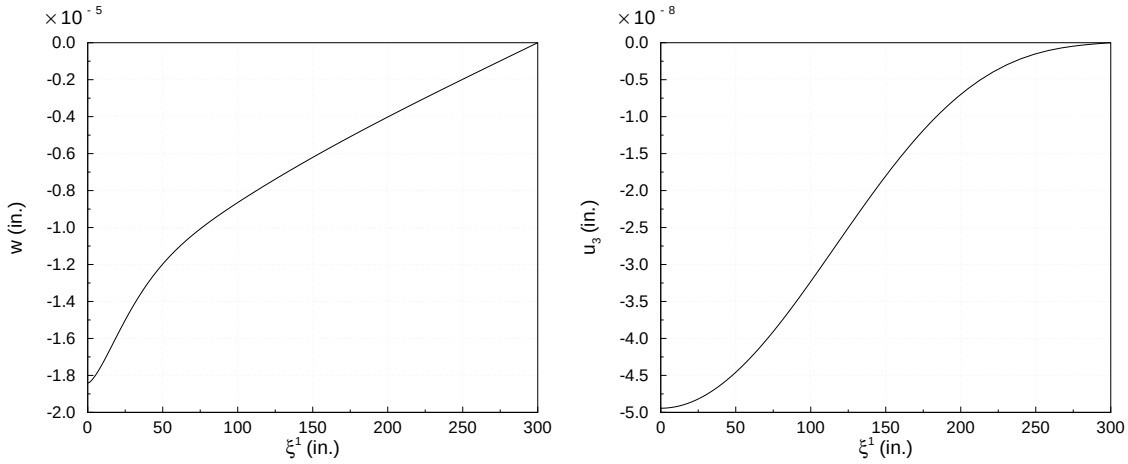

Fig. 87. Vertical displacement profile along the section AB and radial displacement profile along section DC of the pinched cylinder. Least-squares formulation, p-level of 10, 6×6 graded mesh.

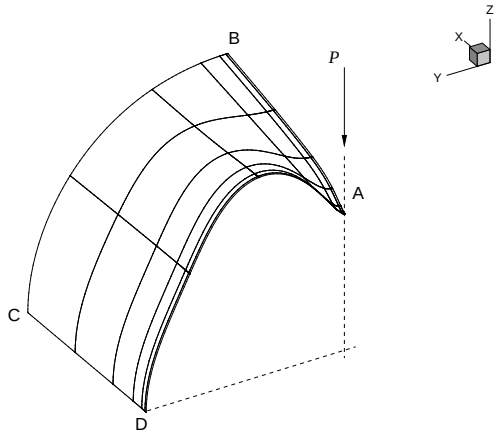

Fig. 88. Deformed mid-surface of the pinched cylinder. Deformation magnified by a factor of 5×10^6.

CHAPTER VII

CONCLUSIONS

A. Summary and concluding remarks

In this work we have presented least-squares based finite element formulations, as an alternate approach to the well-known weak form Galerkin finite element formulations. Formulations based on least-squares principles offer many theoretical and computational advantages in the implementation of the corresponding finite element model that are not present in the traditional weak form Galerkin finite element model. Most notably, the use of least-squares principles leads to a variational unconstrained minimization problem where stability conditions such as inf-sup conditions (typically arising in mixed methods using weak form Galerkin finite element models) never arise. In addition, the finite element model always yields a discrete system of equations with a symmetric positive definite coefficient matrix, allowing the use of robust and fast iterative methods for its solution.

The steps involved in developing a least-squares based finite element model were presented in Chapter II. The procedure was presented in a general setting by considering an abstract initial boundary value problem. The notion of norm equivalence and its implications on the resulting finite element model were discussed and assessed. Having developed an understanding of the compromise that must exist between the optimality and practicality of the finite element model, we chose to retain a predetermined level of "C^0 practicality" throughout this work. This level of practicality was achieved by assuring that the governing equations of the physical problem under consideration were recast as an equivalent first-order system *and* the least-squares functional defined by summing the squares of the equations residuals in the L_2 norm.

This level of practicality came at a price, which reflected upon the finite element formulation in two distinct forms:

1. the resulting formulation may depart from the ideal mathematical setting, in the sense that the resulting least-squares functional may not define an equivalent norm in a suitable Hilbert space. This prevents standard elliptic theory from establishing a priori estimates to predict asymptotic convergence behavior and thus the optimality of the formulation cannot be assured.

2. auxiliary variables need to be introduced to recast the governing equations as a first-order equivalent system of equations, thus increasing the size of the resulting finite element model.

Despite these drawbacks, *bona fide* least-squares formulations are still able to be competitive as the impact of these drawbacks are alleviated, respectively, in the following sense:

1. departure from the ideal mathematical setting does not imply that the resulting formulation is not optimal, it simply means that its optimality cannot be established a priori using standard elliptic theory. Through numerical examples we showed that the C^0 formulations recovered optimal properties if the finite element spaces were spanned by sufficiently high p-levels.

2. auxiliary variables may be viewed as beneficial, in the sense that they may be variables of physical interest. For example, in the context of fluid flow: vorticity or stresses, and in the context of solid mechanics: (bending, shear, or membrane) stress-resultants, rotations, and/or twists.

In Chapter III, we presented least-squares models for the incompressible Navier-Stokes equations. We considered equivalent first-order systems based on vorticity,

stresses, and velocity gradients. Due to the unconstrained minimization setting of the least-squares formulation, the inf-sup condition (also known as the LBB condition), ever present in the weak form Galerkin formulation, representing a strict compatibility condition between the velocity and pressure finite element spaces does not arise. In addition, the resulting discrete model yields a linear system of equations with a symmetric positive definite coefficient matrix, in contrast to the weak form Galerkin model where the resulting coefficient matrix is unsymmetric and indefinite. These properties inherent to the least-squares based model, resulted in the development of robust and efficient computational algorithms for the numerical solution of the incompressible Navier-Stokes equations. The performance of the formulations was demonstrated and exemplified through the solution of several verification and validation benchmarks, including flow over a backward-facing step, lid-driven cavity flows, and flow past a circular cylinder.

In Chapter IV, we considered the compressible Navier-Stokes equations and its equivalent first-order system based on velocity gradients and heat fluxes. Previous work on least-squares formulations for compressible flow dealt only with the limit problems of low-speed compressible flow or invscid compressible flow. Here, for the first time, we presented a formulation that is directly applicable to viscous flows in the subsonic, transonic, or supersonic regime. For such flow conditions the Navier-Stokes equations are of mixed hyperbolic and parabolic type and the the issue of well posed open boundary conditions is of relevance, as the problem naturally exhibits wave propagation. Boundary operators which ensure maximum energy dissipation and result in a strongly well posed problem were derived by applying the energy method to the characteristic based, linearized, constant coefficient viscous compressible Navier-Stokes equations. In the limit of vanishing viscocity, the boundary operators recover the Euler characteristics. The complex boundary conditions in terms

of characteristics are easily enforced in the finite element model through the least-squares functional. Numerical results for flow past a circular cylinder at different free-stream Mach numbers and different surface thermal loadings were presented to demonstrate the performance of the formulation.

In Chapters V and VI we considered least-squares formulations for applications in the field of solid mechanics, specifically we presented least-squares based plate and shell elements. The formulations were of the mixed type, having generalized displacements and stress-resultants as independent variables. For plate bending, finite element models for the classical plate theory (CPT) and first-order shear deformation theory (FSDT) were presented. We showed through several well chosen verification benchmarks that the proposed plate bending elements are free of shear-locking, insensitive to geometric distortions, and predict accurate displacements as well as stress resultants along plate boundaries. For the analysis of shell structures, we presented a finite element model for the Naghdi shear-deformable theory. The shell elements were tested for membrane- and bending-dominated benchmark problems, using uniform and distorted meshes, and found to be effective in both asymptotic states. Results for the well established benchmark problems of the Scordelis-Lo roof and the pinched cylinder were also presented and found to be in good agreement with reference solutions.

B. Topics of ongoing and future research

Throughout this work we have retained the pre-determined level of "C^0 practicality" in arriving at the least-squares based finite element models.

Ongoing research points towards relaxing the "C^0 practicality" level and allowing the finite element spaces to retain better regularity across inter-element boundaries.

This idea has spawned finite element spaces of class C^k, where regularity of order k is allowed across inter-element boundaries, i.e., the derivatives up to order k remain continuous across inter-element boundaries.

Such an approach may result in better least-squares finite element formulations, in the sense that a better compromise between optimality and practicality may be achieved. Retaining a pre-determined level of "C^k practicality" (for $k \geq 1$),

1. the least-squares functional may be defined in terms of L_2 and H^k norms, thus enlarging the class of least-squares functionals that define an equivalent norm in a suitable Hilbert space.

2. the strong form of the governing equations can be used to define the least-squares functional, thus eliminating the need to introduce auxiliary variables.

In the following we present some preliminary results for formulations of the class C^1. The results are presented in the context of viscous incompressible flow, details will be presented in [82].

With a pre-determined level of "C^1 practicality", the following least-squares functionals may be considered to develop a finite element model for the stationary incompressible Navier-Stokes:

$$\mathcal{J}(\mathbf{u}, p; \mathbf{f}) = \frac{1}{2}\left(\| (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p - \frac{1}{\mathrm{Re}}\nabla\cdot\left[(\nabla\mathbf{u}) + (\nabla\mathbf{u})^T\right] - \mathbf{f} \|_0^2 + \| \nabla\cdot\mathbf{u} \|_0^2 \right) \quad (7.1)$$

or

$$\mathcal{J}(\mathbf{u}, p; \mathbf{f}) = \frac{1}{2}\left(\| (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p - \frac{1}{\mathrm{Re}}\nabla\cdot\left[(\nabla\mathbf{u}) + (\nabla\mathbf{u})^T\right] - \mathbf{f} \|_0^2 + \| \nabla\cdot\mathbf{u} \|_1^2 \right) \quad (7.2)$$

Functional (7.1) is defined in terms of L_2 norms only, while functional 7.2 is defined in terms of L_2 norms and H^1 norms and defines an equivalent norm in H^1.

Figure 89 shows convergence of the velocity field in the L_2 norm as a function of total number of degrees of freedom for the Kovazsnay verification benchmark presented in Chapter III. The figure shows results for the three C^0 formulations presented in Chapter III and the C^1 formulation resulting from using functional (7.1) to develop a least-squares finite element model.

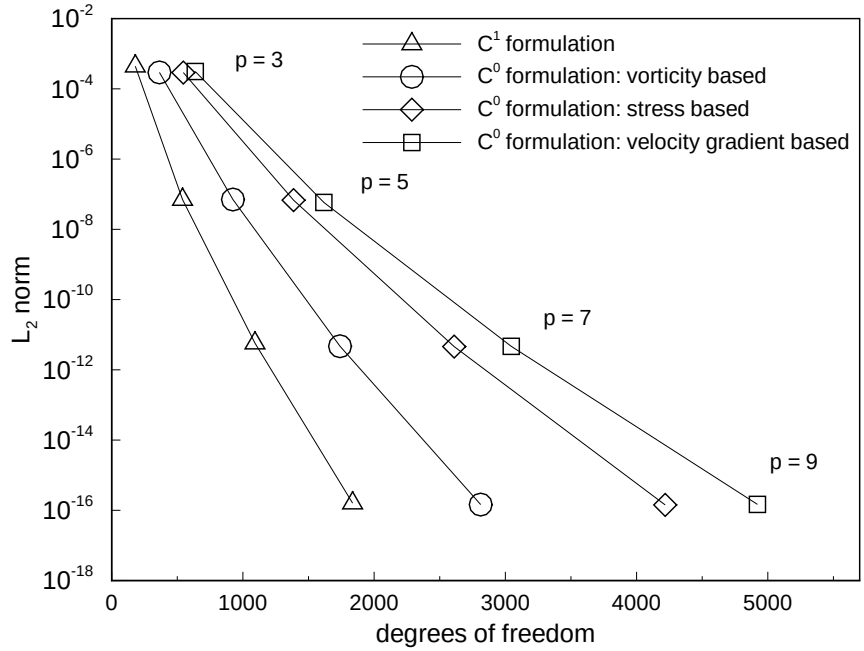

Fig. 89. Convergence of the velocity field to the Kovasznay solution in the L_2-norm for the C^1 formulation and C^0 vorticity, stress, and velocity gradient formulations. Cost comparison.

Clearly, the C^1 formulation outperforms the C^0 formulations, as it is able to achieve the same order of accuracy at significantly lower costs. This is due partly to the fact that the only independent variables are the velocities and pressure and also due to fact that more degrees of freedom are shared at the boundaries of the elements – to ensure the C^1 continuity of the velocity and pressure fields.

REFERENCES

[1] S. Agmon, A. Douglis, and L. Nirenberg, "Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions II," *Comm. Pure Appl. Math.*, vol. 17, pp. 35–92, 1964.

[2] S. Ahmad, B. M. Irons, and O. C. Zienkiewicz, "Analysis of thick and thin shell structures by curved finite elements," *Int. J. Numer. Meth. Engng*, vol. 2, pp. 419–451, 1970.

[3] D. N. Arnold and F. Brezzi, "Locking-free finite element methods for shells," *Math. Comp.*, vol. 66, pp. 1–14, 1997.

[4] I. Babuška and T. Scapolla, "Benchmark computation and performance evaluation for a rhombic plate bending problem," *Int. J. Numer. Meth. Engng*, vol. 28, pp. 155–179, 1989.

[5] K. J. Bathe and E. N. Dvorkin, "A four-node plate bending element based on Mindlin/Reissner plate theory and mixed interpolation," *Int. J. Numer. Meth. Engng*, vol. 21, pp. 367–383, 1985.

[6] ——, "A formulation of general shell elements – the use of mixed interpolation of tensorial components," *Int. J. Numer. Meth. Engng*, vol. 22, pp. 697–722, 1986.

[7] C. E. Baumann and J. T. Oden, "A discontinuous *hp* finite element method for the Euler and Navier-Stokes equations," *Int. J. Numer. Meth. Fluids*, vol. 31, pp. 79–95, 1999.

[8] B. C. Bell and K. S. Surana, "A space-time coupled p-version least-squares finite element formulation for unsteady fluid dynamics problems," *Int. J. Numer. Meth. Engng*, vol. 37, pp. 3545–3569, 1994.

[9] ——, "A space-time coupled p-version least-squares finite element formulation for unsteady two-dimensional Navier-Stokes equations," *Int. J. Numer. Meth. Engng*, vol. 39, pp. 2593–2618, 1996.

[10] T. Belytschko, C. S. Tsay, and W. K. Liu, "Stabilization matrix for the bilinear Mindlin plate element," *Comput. Methods Appl. Mech. Engrg.*, vol. 29, pp. 313–327, 1981.

[11] P. B. Bochev, "Analysis of least-squares finite element methods for the Navier-Stokes equations," *SIAM J. Numer. Anal.*, vol. 34, pp. 1817–1844, 1997.

[12] P. B. Bochev, Z. Cai, T. A. Manteuffel, and S. F. McCormick, "Analysis of velocity-flux first-order system least-squares principles for the Navier-Stokes equations: Part I," *SIAM J. Numer. Anal.*, vol. 35, pp. 990–1099, 1998.

[13] P. B. Bochev and M. D. Gunzburger, "Finite element methods of least-squares type," *SIAM Review*, vol. 40, pp. 789–837, 1998.

[14] P. B. Bochev, T. A. Manteuffel, and S. F. McCormick, "Analysis of velocity-flux least-squares principles for the Navier-Stokes equations: Part II," *SIAM J. Numer. Anal.*, vol. 36, pp. 1125–1144, 1999.

[15] J. H. Bramble and J. Nitsche, "A generalized Ritz-least-squares method for Dirichlet problems," *SIAM J. Numer. Anal.*, vol. 10, pp. 81–93, 1973.

[16] J. H. Bramble and A. H. Schatz, "Least-squares methods for $2m$th order elliptic boundary value problems," *Math. Comp.*, vol. 25, pp. 1–32, 1971.

[17] J. H. Bramble and T. Sun, "A locking-free finite element method for Naghdi shells," *J. Comp. App. Math.*, vol. 89, pp. 119–133, 1997.

[18] F. Brezzi and J. Douglas, "Stabilized mixed methods for the Stokes problem," *Numer. Math.*, vol. 53, pp. 225–235, 1988.

[19] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*. New York: Springer-Verlag, 1991.

[20] Z. Cai, T. A. Manteuffel, and S. F. McCormick, "First-order system least-squares for the Stokes equations, with application to linear elasticity," *SIAM J. Numer. Anal.*, vol. 34, pp. 1727–1741, 1997.

[21] G. Cantin and R. W. Clough, "A curved, cylindrical shell, finite element," *AIAA Journal*, vol. 6, pp. 1057–1062, 1968.

[22] D. Chappelle and K. J. Bathe, "Fundamental considerations for the finite element analysis of shell structures," *Computers & Structures*, vol. 66, pp. 19–36, 1998.

[23] C. Chinosi, L. D. Croce, and T. Scapolla, "Hierarchic finite elements for thin Naghdi shell model," *Int. J. of Solids and Structures*, vol. 35, pp. 1863–1880, 1998.

[24] M. Cho and H. Y. Roh, "Development of geometrically exact new shell elements based on general curvilinear co-ordinates," *Int. J. Numer. Meth. Engng*, vol. 56, pp. 81–115, 2003.

[25] J. Chung and G. M. Hulbert, "A time integration algorithm for structural dynamics with improved numerical dissipation: The generalized α-method," *J. Appl. Mech.*, vol. 60, pp. 371–375, 1993.

[26] S. C. R. Dennis and G. Z. Chang, "Numerical solutions for steady flow past a circular cylinder at Reynolds numbers up to 100," *J. Fluid Mech.*, vol. 42, pp. 471–489, 1970.

[27] W. Dettmer and D. Peric, "An analysis of the time integrator algorithms for the finite elelemt solutions of incompressible Navier-Stokes equations based on stabilised formulation," *Comput. Methods Appl. Mech. Engrg.*, vol. 192, pp. 1177–1226, 2003.

[28] W. S. Don and D. Gottlieb, "Spectral simulation of an unsteady compressible flow past a circular cylinder," *Comput. Methods Appl. Mech. Engrg.*, vol. 80, pp. 39–58, 1990.

[29] J. Dušek, P. L. Gal, and P. Fraunié, "A numerical and theoretical study of the first Hopf bifurcation in a cylinder wake," *J. Fluid Mech.*, vol. 264, pp. 59–80, 1994.

[30] E. D. Eason, "A review of least-squares methods for solving partial differential equations," *Int. J. Numer. Meth. Engng*, vol. 10, pp. 1021–1046, 1976.

[31] A. Fortin, M. Jardak, J. J. Gervais, and R. Pierre, "Localization of Hopf bifurcations in fluid flow problems," *Int. J. Numer. Meth. Fluids*, vol. 24, pp. 1185–1210, 1997.

[32] L. Franca and R. Stenberg, "Error analysis of some Galerkin least-squares methods for the elasticity equations," *SIAM J. Numer. Anal.*, vol. 28, pp. 1680–1697, 1991.

[33] D. K. Gartling, "A test problem for outflow boundary conditions - flow over a backward-facing step," *Int. J. Numer. Meth. Fluids*, vol. 11, pp. 953–967, 1990.

[34] U. Ghia, K. N. Ghia, and C. T. Shin, "High-Re solution for incompressible flow using the Navier-Stokes equations and the multigrid method," *J. Comput. Phys.*, vol. 48, pp. 387–411, 1982.

[35] V. Girault and P. Raviart, *Finite Element Methods for Navier-Stokes Equations*. New York: Springer-Verlag, 1986.

[36] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore: John Hopkins University Press, 1996.

[37] D. Gottlieb, M. Gunzburger, and E. Turkel, "On numerical boundary treatment of hyperbolic systems for finite difference and finite element methods," *SIAM J. Numer. Anal.*, vol. 19, pp. 671–682, 1982.

[38] A. E. Green and W. Zerna, *Theoretical Elasticity*, 2nd ed. Oxford: Oxford University Press, 1968.

[39] P. M. Gresho, D. K. Gartling, J. R. Torczynski, K. A. Cliffe, K. H. Winters, T. J. Garrat, A. Spence, and J. W. Goodrich, "Is the steady viscous incompressible two-dimensional flow over a backward-facing step at Re=800 stable?" *Int. J. Numer. Meth. Fluids*, vol. 17, pp. 501–541, 1993.

[40] A. S. Grove, F. H. Shair, E. E. Petersen, and A. Acrivos, "An experimental investigation of the steady separated flow past a circular cylinder," *J. Fluid Mech.*, vol. 19, pp. 60–80, 1964.

[41] B. Gustafsson and A. Sundström, "Incompletely parabolic problems in fluid dynamics," *SIAM J. Appl. Math.*, vol. 35, pp. 343–357, 1978.

[42] B. Häggblad and K. J. Bathe, "Specification of boundary conditions for Reissner/Mindlin plate bending finite elements," *Int. J. Numer. Meth. Engng*,

vol. 30, pp. 981–1011, 1990.

[43] H. Hakula, Y. Leino, and J. Pitkäranta, "Scale resolution, locking, and high-order finite element modelling of shells," *Comput. Methods Appl. Mech. Engrg.*, vol. 133, pp. 157–182, 1996.

[44] R. D. Henderson, "Details on the drag curve near the onset of vortex shedding," *Phys. Fluids*, vol. 7, pp. 2102–2104, 1995.

[45] J. S. Hestahven, "A stable penalty method for the compressible Navier-Stokes equations: III. Multidimensional domain decomposition schemes," *SIAM J. Sci. Comput.*, vol. 20, pp. 62–93, 1998.

[46] J. S. Hestahven and D. Gottlieb, "A stable penalty method for the compressible Navier-Stokes equations: I. Open boundary conditions," *SIAM J. Sci. Comput.*, vol. 17, pp. 579–612, 1996.

[47] H. C. Huang, "Membrane locking and assumed strain elements," *Computers & Structures*, vol. 27, pp. 671–677, 1987.

[48] H. C. Huang and E. Hinton, "A nine-node Lagrangian plate element with enhanced shear interpolation," *Engrg. Computations*, vol. 1, pp. 369–379, 1984.

[49] R. Iwatsu, J. M. Hyun, and K. Kuwahara, "Numerical simulation of flows driven by a torsionally oscillating lid in a square cavity," *J. Fluids Eng.*, vol. 114, pp. 143–151, 1992.

[50] C. P. Jackson, "A finite-element study of the onset of vortex shedding in flow past variously shaped bodies," *J. Fluid Mech.*, vol. 182, pp. 23–45, 1987.

[51] K. E. Jansen, C. H. Whitting, and G. M. Hulbert, "A generalized α-method for integrating the filtered Navier-Stokes equations with a stabilized finite element method," *Comput. Methods Appl. Mech. Engrg.*, vol. 190, pp. 305–319, 2000.

[52] D. Jesperson, "A least-squares decomposition method for solving elliptic equations," *Math. Comp.*, vol. 31, pp. 873–880, 1977.

[53] B. N. Jiang, "A least-squares finite element method for incompressible Navier-Stokes problems," *Int. J. Numer. Meth. Fluids*, vol. 14, pp. 843–859, 1992.

[54] ——, "On the least-squares method," *Comput. Methods Appl. Mech. Engrg.*, vol. 152, pp. 239–257, 1997.

[55] ——, *The Least-Squares Finite Element Method.* New York: Springer-Verlag, 1998.

[56] ——, "The least-squares finite element method for elasticity. Part II: Bending of thin plates," *Int. J. Numer. Meth. Engng*, vol. 54, pp. 1459–1475, 2002.

[57] B. N. Jiang and G. F. Carey, "Least-squares finite element methods for compressible Euler equations," *Int. J. Numer. Meth. Fluids*, vol. 10, pp. 557–568, 1990.

[58] B. N. Jiang and C. L. Chang, "Least-squares finite elements for the Stokes problem," *Comput. Methods Appl. Mech. Engrg.*, vol. 78, pp. 297–311, 1990.

[59] B. N. Jiang, T. L. Lin, and L. A. Povinelli, "Large-scale computation of incompressible viscous flow by least-squares finite element method," *Comput. Methods Appl. Mech. Engrg.*, vol. 114, pp. 213–231, 1994.

[60] B. N. Jiang and L. A. Povinelli, "Least-squares finite element method for fluid dynamics," *Comput. Methods Appl. Mech. Engrg.*, vol. 81, pp. 13–37, 1990.

[61] L. Kaiktsis, G. E. Karniadakis, and S. A. Orszag, "Onset of three-dimensionality, equilibria, and early transition in flow over a backward-facing step," *J. Fluid Mech.*, vol. 231, pp. 501–528, 1991.

[62] ——, "Unsteadiness and convective instabilities in two-dimensional flow over a backward-facing step," *J. Fluid Mech.*, vol. 321, pp. 157–187, 1996.

[63] T. Kant and E. Hinton, "Mindlin plate analysis by segmentation method," *J. Engrg. Mech.*, vol. 109, pp. 537–556, 1983.

[64] G. E. Karniadakis and S. J. Sherwin, *Spectral/hp Element Methods for CFD*. Oxford: Oxford University Press, 1999.

[65] M. Kawaguti and P. Jain, "Numerical study of a viscous fluid past a circular cylinder," *J. Phys. Soc. Japan*, vol. 21, pp. 2055–2062, 1966.

[66] L. I. G. Kovasznay, "Laminar flow behind a two-dimensional grid," *Proc. Camb. Phil. Soc.*, vol. 44, pp. 58–62, 1948.

[67] I. Kreja, R. Schmidt, and J. N. Reddy, "Finite elements based on a first-order shear deformation moderate rotation shell theory with applications to the analysis of composite structures," *Int. J. of Non-Linear Mechanics*, vol. 32, pp. 1123–1142, 1997.

[68] P. S. Lee and K. J. Bathe, "On the asymptotic behavior of shell structures and the evaluation in finite element solutions," *Computers & Structures*, vol. 80, pp. 235–255, 2002.

[69] D. Lefebvre, J. Peraire, and K. Morgan, "Finite element least-squares solution of the Euler equations using linear and quadratic approximations," *Int. J. Comput. Fluid Dynamics*, vol. 1, pp. 1–23, 1993.

[70] Y. Leino and J. Pitkäranta, "On the membrane locking of *h-p* finite elements in a cylindrical shell," *Int. J. Numer. Meth. Engng*, vol. 37, pp. 1053–1070, 1994.

[71] G. T. Lim and J. N. Reddy, "On canonical bending relationships for plates," *Int. J. of Solids and Structures*, vol. 40, pp. 3039–3067, 2003.

[72] I. Lomtev, C. B. Quillen, and G. E. Karniadakis, "Spectral/*hp* methods for viscous compressible flows on unstructured 2D meshes," *J. Comput. Phys.*, vol. 144, pp. 325–357, 1998.

[73] M. Malinem and J. Pitkäranta, "A benchmark study of reduced-strain shell finite elements: Quadratic schemes," *Int. J. Numer. Meth. Engng*, vol. 48, pp. 1637–1671, 2000.

[74] L. S. D. Morley, *Skew Plates and Structures*. Oxford: Pergamon Press, 1963.

[75] P. M. Naghdi, "Foundations of elastic shell theory," in *Progress in Solid Mechanics*, I. N. Sneddon and R. Hill, Eds. Amsterdam: North-Holland, 1963, vol. 4, pp. 1–90.

[76] J. Nordström, "The influence of open boundary conditions on the convergence to steady state for the Navier-Stokes equations," *J. Comput. Phys.*, vol. 85, pp. 210–244, 1989.

[77] ——, "The use of characteristic boundary conditions for the Navier-Stokes equations," *Computers and Fluids*, vol. 24, pp. 609–623, 1995.

[78] J. Oliger and A. Sundström, "Theoretical and practical aspects of some initial boundary value problems in fluid dynamics," *SIAM J. Appl. Math.*, vol. 35, pp. 419–446, 1978.

[79] J. Pitkäranta, Y. Leino, O. Ovakainen, and J. Piila, "Shell deformation states and the finite element method: A benchmark study of cylindrical shells," *Comput. Methods Appl. Mech. Engrg.*, vol. 128, pp. 81–121, 1995.

[80] J. P. Pontaza, "Verification, validation, and benchmarks for numerical fluid flow simulations," presented at the 7th U.S. National Congress on Computational Mechanics, Albuquerque, New Mexico, July 2003.

[81] J. P. Pontaza, X. Diao, J. N. Reddy, and K. S. Surana, "Least-squares finite element models of two-dimensional compressible flows," *Finite Elements in Analysis and Design*, 2003, in press.

[82] J. P. Pontaza and J. N. Reddy, "C^1-continuous least-squares finite element formulation for the incompressible Navier-Stokes equations," in preparation.

[83] ——, "A high-order least-squares finite element formulation for the viscous compressible Navier-Stokes equations," in preparation.

[84] ——, "Hierarchical mixed least-squares shear-deformable shell elements," *Int. J. Numer. Meth. Engng*, 2003, submitted.

[85] ——, "Mixed plate bending elements based on least-squares formulation," *Int. J. Numer. Meth. Engng*, 2003, in press.

[86] ——, "Space-time coupled spectral/*hp* least-squares finite element formulation for the incompressible Navier-Stokes equations," *J. Comput. Phys.*, 2003, submitted.

[87] ——, "Spectral/*hp* least-squares finite element formulation for the Navier-Stokes equations," *J. Comput. Phys.*, vol. 190, pp. 523–549, 2003.

[88] M. M. J. Proot and M. I. Gerrtisma, "Least-squares spectral elemets applied to the Stokes problem," *J. Comput. Phys.*, vol. 181, pp. 454–477, 2002.

[89] J. N. Reddy, "A penalty plate-bending element for the analysis of laminated anisotropic composite plates," *Int. J. Numer. Meth. Engng*, vol. 15, pp. 1187–1206, 1980.

[90] ——, "Penalty-finite-element analysis of 3-D Navier-Stokes equations," *Comput. Methods Appl. Mech. Engrg.*, vol. 35, pp. 87–97, 1982.

[91] ——, *An Introduction to the Finite Element Method*, 2nd ed. New York: McGraw-Hill, 1993.

[92] ——, *Theory and Analysis of Elastic Plates*, 2nd ed. New York: Taylor & Francis, 1999.

[93] ——, *Energy Principles and Variational Methods in Applied Mechanics*, 2nd ed. New York: John Wiley, 2002.

[94] ——, *An Introduction to Nonlinear Finite Element Analysis*. Oxford: Oxford University Press, 2003, in press.

[95] Y. Saad, *Iterative Methods for Sparse Linear Systems*. Boston: PWS Publishing Company, 1996.

[96] R. L. Sani and P. M. Gresho, "Resume and remarks on the open boundary condition minisymposium," *Int. J. Numer. Meth. Fluids*, vol. 18, pp. 983–1008, 1994.

[97] A. C. Scordelis and K. S. Lo, "Computer analysis of cylindrical shells," *J. American Concrete Institute*, vol. 61, pp. 539–561, 1964.

[98] S. J. Sherwin and G. E. Karniadakis, "A triangular spectral element method; applications to the incompressible Navier-Stokes equations," *Comput. Methods Appl. Mech. Engrg.*, vol. 123, pp. 189–229, 1995.

[99] W. Y. Soh and J. W. Goodrich, "Unsteady solution of incompressible Navier-Stokes equations," *J. Comput. Phys.*, vol. 79, pp. 113–134, 1988.

[100] F. Taghaddosi, W. G. Habashi, W. Guevremont, and D. Ait-Ali-Tahia, "An adaptive least-squares method for the compressible Euler equations," *Int. J. Numer. Meth. Fluids*, vol. 31, pp. 1121–1139, 1999.

[101] L. Q. Tang, T. Cheng, and T. T. H. Tsang, "Transient solutions for three-dimensional lid-driven cavity flows by a least-squares finite element method," *Int. J. Numer. Meth. Fluids*, vol. 21, pp. 413–432, 1995.

[102] L. Q. Tang and T. T. H. Tsang, "A least-squares finite element method for time dependent incompressible flows with thermal convection," *Int. J. Numer. Meth. Fluids*, vol. 17, pp. 271–289, 1993.

[103] ——, "Temporal, spatial and thermal features of 3-D Rayleigh-Benard convection by a least-squares finite element method," *Comput. Methods Appl. Mech. Engrg.*, vol. 140, pp. 201–219, 1997.

[104] J. R. Torczynski, "A grid refinement study of two-dimensional transient flow over a backward-facing step using a spectral element method," in *Separated Flows*, J. C. Dutton and L. P. Purtell, Eds. New York: ASME, 1993, vol. 149, pp. 44–62.

[105] D. J. Tritton, "Experiments on the flow past a circular cylinder at low Reynolds numbers," *J. Fluid Mech.*, vol. 6, pp. 547–567, 1959.

[106] C. M. Wang, Y. C. Wang, J. N. Reddy, and V. Thevendran, "Improved computation of stress resultants in the *p*-ritz method," *J. Structural Engrg.*, vol. 128, pp. 249–257, 2002.

[107] T. C. Warburton, I. Lomtev, Y. Du, S. J. Sherwin, and G. E. Karniadakis, "Galerkin and discontinuous Galerkin spectral/*hp* methods," *Comput. Methods Appl. Mech. Engrg.*, vol. 175, pp. 343–359, 1999.

[108] F. M. White, *Viscous Fluid Flow.* New York: McGraw-Hill, 1974.

[109] C. H. K. Williamson, "Oblique and parallel modes of vortex shedding in the wake of a circular cylinder at low Reynolds numbers," *J. Fluid Mech.*, vol. 206, pp. 579–627, 1989.

[110] D. Winterscheidt and K. S. Surana, "*p*-version least-squares finite element formulation for two-dimensional incompressible fluid flow," *Int. J. Numer. Meth. Fluids*, vol. 18, pp. 43–69, 1994.

[111] H. C. Yee, J. R. Torczynski, S. A. Morton, M. R. Visbal, and P. K. Sweby, "On spurious behavior of CFD simulations," *Int. J. Numer. Meth. Fluids*, vol. 30, pp. 675–711, 1999.

[112] S. T. Yu, B. N. Jiang, J. Wu, and N. S. Liu, "A div-curl-grad formulation for compressible buoyant flows solved by the least-squares finite element method," *Comput. Methods Appl. Mech. Engrg.*, vol. 137, pp. 59–88, 1996.

VITA

Juan Pablo Pontaza was born on July 31, 1976, in Guatemala City, Guatemala. He received his elementary education in Guatemala City, Guatemala and part of his secondary education in Altensteig, Germany. He enrolled at Texas A&M University in January 1995, majoring in mechanical engineering. While attending Texas A&M University he participated in the University Honors Program, the University Undergraduate Research Fellows, and the Engineering Scholars Program. He graduated with his B.S. degree, as Magna Cum Laude, and received the best thesis award as an Undergraduate Research Fellow in May 1999.

Juan Pablo then attended the Massachusetts Institute of Technology (MIT), where he worked on finite element formulations for high Reynolds number flow. The core idea of the research was to use local analytic solutions of the linear advection-diffusion problem to endow the finite element spaces with the ability to respond to local flow conditions, thus allowing under-resolved simulations to remain numerically stable at high Reynolds numbers. Juan Pablo graduated with his M.S. degree in mechanical engineering from MIT in June 2001.

Juan Pablo came back to Texas A&M University (TAMU) in September 2001 to continue his graduate studies under the supervision of Prof. J.N. Reddy, whom he had met during his undergraduate studies at TAMU. Juan Pablo's research at TAMU was concerned with least-squares finite element formulations and their application to solid and fluid mechanics. He graduated with his Ph.D. degree in mechanical engineering in December 2003.

Juan Pablo Pontaza may be contacted through Prof. J.N. Reddy at the Mechanical Engineering Department, Texas A&M University, College Station, TX 77843-3123.