UNIVERSITAT POLITÈCNICA DE CATALUNYA
E.T.S. D'ENGINYERS DE CAMINS, CANALS I PORTS
DEPARTAMENT DE MATEMÀTICA APLICADA III

ANALYSIS OF FRACTIONAL STEP, FINITE ELEMENT METHODS FOR THE INCOMPRESSIBLE NAVIER–STOKES EQUATIONS

by

Jorge Blasco Lorente

Advisors: Antonio Huerta Cerezuela,
Ramon Codina Rovira.

Program: Applied Mathematics.

Barcelona, December 1996.

Contents

Introduction **5**

1 Preliminaries **9**

1.1 Flow equations . . . 9
1.2 Function spaces, norms and forms . . . 12
1.3 Finite element approximation . . . 17
1.4 Mixed problems and the LBB condition . . . 18
1.4.1 Mixed problems . . . 19
1.4.2 Discrete approximations . . . 20
1.4.3 Some stabilizing techniques . . . 22
1.5 Description of fractional step methods . . . 24
1.5.1 Classical Projection Methods . . . 24
1.5.2 Higher Order Methods . . . 28
1.5.3 Viscosity splitting methods . . . 31
1.5.4 Other methods . . . 32
1.5.5 Further comments and conclusions . . . 33

2 A reformulated Stokes problem **35**

2.1 Stabilizing properties of projection methods . . . 36
2.2 Development of the method . . . 40
2.2.1 The continuous problem . . . 40
2.2.2 The discrete problem . . . 42
2.3 Stability and convergence of the method . . . 44
2.3.1 Stability . . . 44
2.3.2 Convergence in natural norms . . . 45
2.3.3 Convergence in L^2–norms . . . 50
2.4 A weakened *inf–sup* condition . . . 52
2.4.1 Macroelement technique . . . 53
2.4.2 Equal order interpolations . . . 60
2.5 Computational aspects . . . 62
2.5.1 Uncoupled block–Gauss–Seidel method . . . 63
2.5.2 Coupled block–Gauss–Seidel method. . . . 66
2.5.3 Acceleration of convergence . . . 67
2.5.4 Local stability parameter . . . 71
2.5.5 Summary of computational aspects . . . 73

2.6 Numerical results 74
2.6.1 Cavity flow problem 74
2.6.2 A test with an analytical solution 74
2.6.3 Behaviour of the pressure near the boundary 84

3 Reformulated Navier–Stokes equations 89
3.1 The steady, incompressible Navier Stokes equations 90
3.2 Development of the method 91
3.2.1 The continuous problem 91
3.2.2 The discrete problem 93
3.3 Stability and convergence of the method 98
3.3.1 Stability 98
3.3.2 Convergence in natural norms 99
3.3.3 Convergence in L^2-norm 100
3.4 Computational aspects 102
3.4.1 Nonlinear iterative solvers 103
3.4.2 Performance of the iterative schemes 107
3.4.3 Summary of computational aspects 110
3.5 Numerical results 110
3.5.1 Cavity flow problem 110
3.5.2 Kovasznay flow 111
3.5.3 Poiseuille flow through a junction of pipes 117

4 Viscosity splitting fractional step method 123
4.1 The unsteady, incompressible Navier Stokes equations 124
4.2 Viscosity splitting method 126
4.2.1 Development of the method 126
4.2.2 Convergence of the method 129
4.2.3 Error estimates 135
4.3 A pressure–correction method 143
4.3.1 Development of the method 143
4.3.2 Error estimates 145
4.3.3 Dependence of the steady state on the time step 151
4.4 Computational aspects 153
4.4.1 Finite element discretization 153
4.4.2 Numerical scheme 153
4.5 Numerical results 156
4.5.1 Numerical accuracy study 156
4.5.2 Backward facing step 158
4.5.3 Flow past a circular cylinder 159

5 A predictor–multicorrector algorithm 169
5.1 Semidiscrete form of the algorithm 170
5.2 Finite element discretization 175
5.3 Numerical results 176
5.3.1 Numerical accuracy study 176

5.3.2 Kovasznay flow 177
5.3.3 Cavity flow problem 186
5.3.4 Noflow problem 186
5.3.5 Plane jet simulation 190

Conclusions and future work **197**

A Improved error estimates **201**

A.1 Viscosity splitting method 201
A.2 Viscosity splitting, pressure correction method 206

List of Tables **211**

List of Figures **212**

References **215**

Introduction

The numerical solution of viscous incompressible flow problems is a hard and challenging subject, which has received much attention in the last decades. The main difficulties it poses are of three different kinds: first, the incompressibility condition, and consequently the pressure computation, establishes a coupling between the unknowns of the problem which, in standard formulations, restricts the freedom to choose discrete approximating spaces to those satisfying a certain compatibility condition; second, the advective–diffusive character of the equations may require of appropriate stabilizing techniques or extremely fine meshes, specially in convection dominated flow situations; finally, the nonlinearity of the equations increases the computational burder of the solution procedures.

Beyond those theoretical and computational challenges, the development of numerical methods for incompressible flow equations is of an undoubtable practical importance. These equations find numerous applications in different areas, both scientific and industrial, such as aeronautical sciences, metereology, ocean dynamics, environmental flows, oil industry, turbulent flows and many others. Besides, they are the basis for several extentions to more complex flow situations, such as thermal flows, free–surface flows, magnetohydrodynamics and others.

Many numerical schemes have been developed to approximate the solution of flow equations. For the space variables, discretizations range from finite differences, the simplest and most intuitive discretization method, to finite volume, finite element, boundary element and spectral element methods. Finite element methods have proved to be the most versatile, since they can cope with arbitrarily complex geometries and work on unstructured, automatically generated meshes, are based on more rigorous theoretical grounds and are liable to generalizations of arbitrary order of accuracy. As for the time integration, all these numerical schemes can be sorted, in a first approach, into single step, multistep and fractional step methods. These last methods are also sometimes known as operator splitting or projection methods.

The present work is devoted to the study of fractional step, finite element methods for the numerical solution of incompressible, viscous flow equations, and in particular of the Navier–Stokes equations. The main advantages of some of these methods over other time stepping strategies are the decoupling of the unknowns of the problem, thus reducing the size of the discrete problems to be solved, and the possibility of employing space interpolations which do not satisfy the compatibility condition, such as equal order ones. This last fact has been known for some time but, to our knowledge, not fully explained up to now; the first objective of this thesis is to provide a full explanation for it. On the other hand, projection methods are known to suffer from certain drawbacks, the main one being the need to impose some boundary conditions in one of the substeps of the method which are unphysical and may be a source of error; the second objective of this work is to develop a fractional step method which allows the imposition of the boundary conditions of the original problem in all substeps of the method.

In order to find the ultimate reason why fractional step projection methods allow the use of arbitrary space interpolations, a new method is developed for the simpler, linear, steady Stokes problem. This method retains the main features of projection methods for the full problem as far as space discretization is concerned; in particular, it allows the use of equal order interpolations, thus explaining why projection methods also do so. Optimal order convergence in the mesh size is proved for this method under a compatibility condition on the approximating spaces which is weaker than the standard one, and in particular satisfied by equal order interpolations. An extention of this method to the nonlinear, steady problem is also studied, and optimal order convergence, under the same compatibility condition as in the linear case and assuming a unique solution of the problem, is also proved.

As for the treatment of boundary conditions, a fractional step method is developed in which the viscous term is split into the two substeps of the scheme, which, unlike in standard projection methods, allows to enforce the boundary conditions conditions of the original problem in both substeps. Convergence in the time step both for the intermediate and end–of–step velocities of this method is proved in the spaces $L^2(\Omega)$ and $H_0^1(\Omega)$; in this last space, convergence of the end–of–step velocities does not hold for the classical projection method, due to the wrong boundary conditions they satisfy. Our fractional step method was also developed to explain the properties of a well known predictor multicorrector algorithm, which is here shown to be of a fractional step kind.

The primitive variable, velocity–pressure formulation of the equations has been considered throughout this work, and, although possible, no attempt has been made to extend it to other formulations. Moreover, simple boundary conditions have been used in the theoretical developments presented here, usually homogeneous Dirichlet conditions; in the numerical examples, however, 'natural' boundary conditions have also been considered sometimes. The extention of the theory to other kinds of boundary conditions is, again, possible, but not pursued here.

Since we have considered only low to moderate Reynolds' number flows, leaving aside highly convective flows, we have found no need to stabilize convection, and standard Galerkin formulations have been employed. In this sense, we have only dealt with laminar flow regimes.

We have also restricted ourselves to bounded domains and finite time problems, since extentions to other cases pose some additional theoretical difficulties into the formulations. Moreover, the numerical examples we have actually solved are all two–dimensional, although the theoretical developments are also valid for three–dimensional problems.

This work is structured into five Chapters. In the first one, a review of known results, which will be frequently referred to afterwards, is provided, where the notation and terminology used here is also introduced; in particular, a description of several fractional step methods is presented, classified according to different criteria. The second and third Chapters are devoted to the study of the new method for the steady Stokes and Navier–Stokes equations, respectively, which allow the use of equal order interpolations and explain why projection methods also do so. The structure of these two Chapters is similar, with some theoretical Sections first, where stability and optimal order convergence both in H^1 and L^2 norms are proved, followed by some computational aspects and the presentation of numerical results on some test problems.

In Chapter 4 the fractional step method that we consider is introduced and studied, for which first order convergence in the time step is proved both for the intermediate and end–of–step velocities. A convergence theorem which we originally proved for this method using more classical arguments and less restrictive assumptions on the solution and domain is also given. A variant of this method, using pressure correction, is also considered in this Chapter, and first order convergence for the velocities is also proved for this new scheme. An implementation of this pressure correction method as well as some numerical results obtained with it are also provided in this Chapter. Finally, in Chapter 5 a predictor–multicorrector algorithm is studied, showing in what sense it can be understood as a fractional step method and how it behaves in front of different space interpolations, both satisfying and not satisfying the standard compatibility condition. Numerical results obtained with this algorithm for two different space interpolations on several problems are also presented.

Chapter 1

Preliminaries

This first Chapter is devoted to the introduction of the basic mathematical concepts required for the development of the present work. In particular, we first recall the equations of motion of an incompressible fluid; then we review the basic function spaces, norms and forms needed for the study of those equations; later on we introduce the basic theory of finite element approximation and some standard results about mixed problems, and finally we give a comprehensive presentation of existing fractional step methods for the unsteady, incompressible Navier–Stokes equations, the study of which is the ultimate objective of this thesis.

1.1 Flow equations

Let us recall here the basic theory of fluid mechanics, which can be found in standard references such as [72]. The equations of fluid motion are obtained from principles of conservation of physical quantities, and simplified under various hypothesis. Several aspects of this theory are closely related to linear elasticity theory.

We consider a region $\Omega \subset \mathbb{R}^d$, where $d = 2$ or 3, filled with fluid material. The domain Ω is assumed to be open, bounded, connected and Lipschitz continuous, that is, its boundary Γ is a $(d-1)$-dimensional locally Lipschitz manifold. In particular, we will sometimes consider the case of Ω a convex polygon in $\mathbb{R}^2$ or a convex polyhedron in $\mathbb{R}^3$.

For a given $T > 0$, let $\rho(\mathbf{x}, t)$ and $\mathbf{u}(\mathbf{x}, t)$ denote the density and velocity of the fluid at a point $\mathbf{x} \in \Omega$ and time $t \in (0, T)$, respectively (boldface characters denote vector quantities). Conservation of mass leads to the continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad \text{in } \Omega \times (0, \mathrm{T}) \tag{1.1}$$

where $\nabla = (\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_d})$. If the fluid is subject to a volumetric force field $\mathbf{f}(\mathbf{x}, t)$ per unit density, conservation of momentum leads to the Cauchy equations:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} - \frac{1}{\rho}\nabla\cdot\boldsymbol{\sigma} = \mathbf{f} \quad \text{in } \Omega\times(0,T) \tag{1.2}$$

where $\boldsymbol{\sigma}$ is the stress tensor, representing the internal forces acting on the fluid. In the case of a viscous fluid, where internal frictions are taken into account, which is assumed to be Newtonian and isotropic, a constitutive equation of the form:

$$\boldsymbol{\sigma} = -\tilde{p}\mathbf{I} + 2\mu\epsilon(\mathbf{u}) + \lambda(\nabla\cdot\mathbf{u})\mathbf{I} \tag{1.3}$$

is obtained, where $\tilde{p}(\mathbf{x},t)$ is the fluid's pressure, $\mathbf{I}$ is the identity tensor, $\epsilon(\mathbf{u})$ is the deformation rate tensor with components $\epsilon_{ij} = \frac{1}{2}(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i})$, μ is the dynamic viscosity of the fluid and λ is its second viscosity. We will assume that these two scalar parameters remain constant if there are no temperature or density variations. The unsteady Navier–Stokes equations are then obtained under all these hypothesis:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p - 2\nu\nabla\cdot\epsilon(\mathbf{u}) - \frac{\lambda}{\rho}\nabla(\nabla\cdot\mathbf{u}) = \mathbf{f} \quad \text{in } \Omega\times(0,T) \tag{1.4}$$

where $\nu = \mu/\rho$ is the fluid's kynematic viscosity and $p(\mathbf{x},t)$ stands for the fluid's kynematic pressure (pressure divided by density).

We next consider the incompressibility condition, which establishes the limits of the scope of this work. Conservation of fluid volume leads to the condition:

$$\nabla\cdot\mathbf{u} = 0 \quad \text{in } \Omega\times(0,T) \tag{1.5}$$

which will be frequently referred to in what follows. Equation 1.5 replaces the continuity equation for an incompressible homogeneous fluid (that is, with constant density in space), since then ρ is constant at all times. No equation of state is then required to relate p and ρ. Equation 1.4 reduces in this case to:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p - 2\nu\nabla\epsilon(\mathbf{u}) = \mathbf{f} \quad \text{in } \Omega\times(0,T) \tag{1.6}$$

Under the incompressibility condition, 1.6 can further be rewritten as:

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p - \nu\Delta\mathbf{u} = \mathbf{f} \quad \text{in } \Omega\times(0,T) \tag{1.7}$$

where Δ is the Laplacian operator; equation 1.7 is the best known form of the unsteady, incompressible Navier–Stokes equations. A third possible formulation of the viscous term is obtained by making use of the vector identity $\Delta\mathbf{u} = \nabla(\nabla\cdot\mathbf{u}) - \nabla\times(\nabla\times\mathbf{u})$ and the incompressibility condition 1.5, resulting in the substitution of $\Delta\mathbf{u}$ in 1.7 by $-\nabla\times(\nabla\times\mathbf{u})$ (this is usually refered to as the *rot–rot* form of the viscous term, and is also employed in some numerical methods).

As for the convective term, some other formulations can also be considered. Thus, the j–th component of the *conservative* form is $\dfrac{\partial(u_i u_j)}{\partial x_i}$, where the summation convention is assumed on the i–th index. Under the incompressibility condition 1.5, this formulation is equivalent to that of 1.7. The *skew–symmetric* form $(\mathbf{u}\cdot\nabla)\mathbf{u}+\frac{1}{2}(\nabla\cdot\mathbf{u})\mathbf{u}$, also equivalent to that of 1.7 for an incompressible fluid, will also be frequently used.

Equation 1.7 is formally equivalent to its dimensionless form, provided $\nu = 1/\mathrm{Re}$, Re being the fluid's Reynolds number defined as $\mathrm{Re} = \rho\bar{u}L/\mu$. Here, $\bar{u}$ and L stand for a characteristic velocity and length of the fluid's motion, respectively. We will make this identification throughout this work.

The equation system 1.6–1.5 (or 1.7–1.5) has to be completed with suitable boundary and initial conditions to form a well-posed initial/boundary value problem. One generally assumes that the boundary Γ can be partitioned into two non–overlapping subsets Γ_D and Γ_N which accomodate given Dirichlet and Neumann boundary conditions, that is to say, prescribed velocitites and stresses, respectively:

$$\begin{aligned} \mathbf{u}(\mathbf{x},t) &= \tilde{\mathbf{u}}(\mathbf{x},t), \quad \mathbf{x}\in\Gamma_D, \;\; t\in(0,T) \\ \mathbf{n}\cdot\boldsymbol{\sigma}(\mathbf{x},t) &= \mathbf{h}(\mathbf{x},t), \quad \mathbf{x}\in\Gamma_N, \;\; t\in(0,T) \end{aligned} \qquad (1.8)$$

In equation 1.8, and throughout this work, $\mathbf{n}$ denotes the unit outward normal to Γ, $\boldsymbol{\sigma} = -p\mathbf{I} + 2\nu\epsilon(\mathbf{u})$ is the stress tensor (per unit density), $\tilde{\mathbf{u}}$ is the prescribed velocity and $\mathbf{h}$ the prescribed stress. In the formulation of equation 1.6 the condition $\mathbf{h} = \mathbf{0}$ in 1.8 comes up as a natural boundary condition, as is often employed in outflow boundaries, having the physical meaning of a no stress condition. On the contrary, when the formulation of equation 1.7 is employed, the natural condition for outflow boundaries does not have a physical meaning.

Purely Dirichlet type boundary conditions for the velocity, or equivalently $\Gamma_N = \emptyset$, are also considered sometimes. For consistency with the incompressibility condition 1.5, in that case it is required that the net flux of $\tilde{\mathbf{u}}$ through Γ be zero:

$$\int_\Gamma \mathbf{n}\cdot\tilde{\mathbf{u}}(\mathbf{x},t)\,d\Gamma = 0, \quad \forall t\in(0,T) \qquad (1.9)$$

In the theory to be developed in this work, homogeneous Dirichlet boundary conditions will be frequently assumed:

$$\mathbf{u}(\mathbf{x},t) = \mathbf{0}, \quad \mathbf{x}\in\Gamma, \;\; t\in(0,T) \qquad (1.10)$$

which is commonly referred to as the solid wall condition (no slip and no penetration). However, in some of the numerical examples presented, natural boundary conditions are also employed.

An initial condition is also required for the velocity:

$$\mathbf{u}(\mathbf{x},0) = \mathbf{u}_0(\mathbf{x}), \quad \mathbf{x} \in \Omega, \tag{1.11}$$

where $\mathbf{u}_0$ is assumed to be incompressible ($\nabla \cdot \mathbf{u}_0 = 0$). No initial or boundary conditions need be specified for the pressure, although this variable is subject to some a posteriori conditions (see Section 1.4).

The numerical approximation of the equation system 1.7–1.5–1.10–1.11 is the main concern of this work. This system is an unsteady, nonlinear problem coupled with the incompressibility constraint. Related but more simplified problems are also important to deal with. Thus, at steady state one gets the steady, incompressible Navier–Stokes equations:

$$(\mathbf{u} \cdot \nabla)\mathbf{u} + \nabla p - \nu \Delta \mathbf{u} = \mathbf{f} \quad \text{in } \Omega \tag{1.12}$$

which retain the nonlinear, convective/diffusive character of the full problem.

Furthermore, under the assumption of slow motion, and for low Reynolds number flows, the convective (quadratic) term can be neglected in 1.12, resulting in the Stokes problem, which consists of equation 1.5 together with:

$$-\nu \Delta \mathbf{u} + \nabla p = \mathbf{f} \quad \text{in } \Omega \tag{1.13}$$

This is the linear, steady counterpart of the original problem, still retaining the coupling with the incompressibility constraint.

1.2 Function spaces, norms and forms

We introduce here the basic mathematical theory of L^p and Sobolev spaces, where weak solutions of the preceeding equations belong. The results stated herein and the notation introduced will be of constant use in the following Chapters. The general theory presented in this Section is rather classical, and can be found in several texts such as [1] or [111]. However, some aspects are specific to incompressible flow equations, mainly those related to the divergence operator. These are treated in [19], [43], [71] and [105].

Let $C_0^\infty(\Omega)$ denote the set of infinitely differentiable real functions with compact support on Ω, and $D(\Omega)$ the space $C_0^\infty(\Omega)$ with a topology that makes derivation continuous (see [83], for instance); the dual space of $D(\Omega)$, denoted by $D'(\Omega)$, is the set of distributions on Ω. Distributions are infinitely differentiable in the sense of distributions.

Given $1 \leq p < \infty$, $L^p(\Omega)$ is the space of real functions u such that u^p is absolutely integrable in Ω with respect to the Lebesgue measure in $\mathbb{R}^d$. It is a Banach space for the norm $||u||_{L^p(\Omega)} = (\int_\Omega |u(x)|^p dx)^{1/p}$, and it is separable. For $1 < p < \infty$, $L^p(\Omega)$ is reflexive and its dual space is $L^q(\Omega)$, for q such that $1/p + 1/q = 1$; since we are assuming that Ω is bounded, one also has that for $1 \leq s < r < \infty$, $L^r(\Omega) \subset L^s(\Omega)$. The set $C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$ for $1 \leq p < \infty$.

The special case $p = 2$ is of main importance; $L^2(\Omega)$ is in fact a Hilbert space for the scalar product:

$$(u, v) \doteq \int_\Omega u(x)\, v(x)\, dx, \quad \forall u, v \in L^2(\Omega)$$

and norm:

$$|u| \doteq (u, u)^{1/2}$$

Here, and in what follows, the notation $\doteq$ is employed to denote equalities by definition. The space $L^2(\Omega)$ is usually identified with its dual space.

For $p = \infty$, the space $L^\infty(\Omega)$ consists of essentially bounded, real funtions on Ω, which is also a Banach space for the norm $||u||_\infty \doteq \operatorname{ess\,sup}_{x\in\Omega}\{|u(x)|\}$. Again, since Ω is bounded $L^\infty(\Omega) \subset L^p(\Omega)$ for all $p \in [1, \infty)$. One has that $(L^1(\Omega))' = L^\infty(\Omega)$, but $(L^\infty(\Omega))' \supset L^1(\Omega)$ with an strict inclusion.

Any function $f \in L^2(\Omega)$ can be understood as a distribution if one identifies $< f, u >= (f, u), \ \ \forall u \in D(\Omega)$ (we use the notation $<,>$ for duality pairings). The Sobolev space of order 1, $H^1(\Omega)$, consists of functions in $L^2(\Omega)$ such that their generalized first order derivatives (that is, derivatives in distribution sense) are also in $L^2(\Omega)$. It is also a Hilbert space with respect to the scalar product:

$$((u, v))_1 \doteq (u, v) + \sum_{i=1}^{d} (\frac{\partial u}{\partial x_i}, \frac{\partial v}{\partial x_i})$$

and norm:

$$||u||_1 \doteq ((u, u))_1^{1/2}$$

The inclusion $H^1(\Omega) \subset L^2(\Omega)$ is compact. The closure of $C_0^\infty(\Omega)$ in $H^1(\Omega)$ for the norm $||u||_1$ is denoted by $H_0^1(\Omega)$, and it is a proper subspace of $H^1(\Omega)$. To characterize the functions in $H_0^1(\Omega)$, we need to recall a classical trace theorem (see [43]): if the boundary Γ of Ω is Lipschitz continuous, then there exists a linear, continuous operator γ_0 mapping $H^1(\Omega)$ into $L^2(\Gamma)$ such that for any $u \in C^2(\bar{\Omega})$, $\gamma_0(u)$ is the restriction of u to Γ. The subspace $H_0^1(\Omega)$ can then be shown to be the kernel of γ_0, i.e., it consists of functions in $H^1(\Omega)$ which vanish at the boundary. The image space $\gamma_0(H^1(\Omega))$ is denoted by $H^{1/2}(\Gamma)$; its dual space is called $H^{-1/2}(\Gamma)$.

In the case Ω bounded, the classical Poincaré inequality holds; essentially, it says that there exists $C_\Omega > 0$ such that:

$$|u| \leq C_\Omega\, ||u||, \quad \forall u \in H_0^1(\Omega) \tag{1.14}$$

where:

$$||u||^2 \doteq \sum_{i=1}^{d} (\frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_i})$$

This shows that $||u||$ is a norm on $H_0^1(\Omega)$, equivalent to $||u||_1$; the associated scalar product is:

$$((u,v)) \doteq \sum_{i=1}^{d} (\frac{\partial u}{\partial x_i}, \frac{\partial v}{\partial x_i}), \ \forall u, v \in H_0^1(\Omega)$$

The dual space of $H_0^1(\Omega)$ is denoted by $H^{-1}(\Omega)$.

Sobolev spaces of order higher than one are also required sometimes. Thus, given $m \geq 1$, one considers the space $H^m(\Omega)$ made up with functions in $L^2(\Omega)$ whose generalized derivatives up to order m are in $L^2(\Omega)$. It is also a Hilbert space with respect to the scalar product:

$$((u,v))_m \doteq \sum_{|\eta| \leq m} (\partial^\eta u, \partial^\eta v), \ \forall u, v \in H^m(\Omega)$$

where $\eta = (\eta_1, \ldots, \eta_d) \in \mathbb{N}^d$ and $|\eta| \doteq \eta_1 + \cdots + \eta_d$. The norm in $H^m(\Omega)$ is denoted by $||u||_m$.

All the preceeding results have been stated for scalar functions u. The extentions to d-dimensional vector functions $\mathbf{u}$ are made in the usual way, with the help of product norms. Spaces like $\mathbf{D}(\Omega)$, $\mathbf{L}^2(\Omega)$, $\mathbf{H}^1(\Omega)$ or $\mathbf{H}_0^1(\Omega)$ will often be considered.

We now turn to the consideration of the subspaces needed for the treatment of the incompressibility condition 1.5. One usually defines the space:

$$H(\text{div}, \Omega) \doteq \{\mathbf{u} \in \mathbf{L}^2(\Omega) \ / \ \nabla \cdot \mathbf{u} \in L^2(\Omega)\}$$

which is a Hilbert space with respect to the norm $||\mathbf{u}||^2_{\text{div},\Omega} = |\mathbf{u}|^2 + |\nabla \cdot \mathbf{u}|^2$, $\forall \mathbf{u} \in H(\text{div}, \Omega)$. It is well known that there exists a normal trace operator for functions in $H(\text{div}, \Omega)$: if Γ is Lipschitz continuous, then there exists a linear, continuous operator γ_1 mapping $H(\text{div}, \Omega)$ into $H^{-1/2}(\Gamma)$ such that for every $\mathbf{u} \in \mathbf{D}(\bar{\Omega})$, $\gamma_1(u) = \mathbf{n} \cdot \mathbf{u}_{|\Gamma}$. The kernel of γ_1 is denoted by $H_0(\text{div}, \Omega)$.

The set of smooth, solenoidal vector fields is defined as:

$$\mathcal{V} \doteq \{\mathbf{u} \in \mathbf{D}(\Omega) \ / \ \nabla \cdot \mathbf{u} = 0\}$$

The closure of $\mathcal{V}$ in $\mathbf{L}^2(\Omega)$ is denoted by H, and plays a key role in theory of approximation of the Navier–Stokes equations . It can be shown that when Ω is bounded and Γ is Lipschitz continuous:

$$H = \{\mathbf{u} \in \mathbf{L}^2(\Omega) \ / \ \nabla \cdot \mathbf{u} = 0, \ \gamma_1(\mathbf{u}) = 0\}$$

that is, H consists of vector fields in $\mathbf{L}^2(\Omega)$ with zero divergence and zero normal trace at the boundary.

Since H is a closed subspace of $\mathbf{L}^2(\Omega)$, one has the decomposition $\mathbf{L}^2(\Omega) = H \oplus H^\perp$; the characterization of $H^\perp$ is a main concern in this context. It derives from a theorem due to Ladyzenskaya (see [71]), which essentially states that:

$$H^{\perp} = \{\mathbf{u} \in \mathbf{L}^2(\Omega) \;/\; \exists p \in H^1(\Omega),\ \mathbf{u} = \nabla p\}.$$

This is related to the classical Helmholtz decomposition of a vector field into the sum of a solenoidal field and the gradient of a scalar function, and ultimately to a powerful theorem proved by De Rham within the context of distributions (see [84]). This characterization implies, in particular, that for every $\mathbf{u} \in \mathbf{L}^2(\Omega)$, $(\mathbf{u}, \mathbf{v}) = 0 \;\; \forall \mathbf{v} \in \mathcal{V}$ if and only if $\mathbf{u} = \nabla p$ for some $p \in H^1(\Omega)$ defined up to an additive constant.

The projection of $\mathbf{L}^2(\Omega)$ onto H, denoted by P_H, is also of main importance, and actually gives name to a whole category of numerical methods (see Section 1.5). It is obviously continuous on $\mathbf{L}^2(\Omega)$, but it also maps $\mathbf{H}^1(\Omega)$ into itself and is continuous with respect to the norm of $\mathbf{H}^1(\Omega)$ (see [105] or [71]); that is to say, there exists a constant $C_1 > 0$ such that:

$$||P_H(\mathbf{u})||_1 \;\leq\; C_1||\mathbf{u}||_1, \quad \forall \mathbf{u} \in \mathbf{H}^1(\Omega)$$

One also considers the closure of $\mathcal{V}$ in $\mathbf{H}^1_0(\Omega)$. This space is classically denoted by V, but here we will refer to it as Y, keeping the notation V for other purposes. It can be shown (see [43]) that when Ω is bounded and Γ is Lipschitz continuous:

$$Y = \{\mathbf{u} \in \mathbf{H}^1_0(\Omega) \;/\; \nabla \cdot \mathbf{u} = 0\}$$

The decomposition $\mathbf{H}^1_0(\Omega) = Y \oplus Y^{\perp}$, analogous to the previous one, can be characterized in this case as follows (see [43]):

$$Y^{\perp} = \{\mathbf{u} \in \mathbf{H}^1_0(\Omega) \;/\; \exists p \in L^2(\Omega),\ \mathbf{u} = (-\Delta)^{-1}(\nabla p)\} \qquad (1.15)$$

where $(-\Delta)^{-1}$ is the inverse of the Riesz representation isomorphism, that is, $-\Delta\colon H^1_0(\Omega) \rightarrow H^{-1}(\Omega)$ defined by $< -\Delta u, v > \doteq ((u, v))$; 1.15 is to be understood in the following sense: for every $\mathbf{u} \in \mathbf{H}^1_0(\Omega)$, $((\mathbf{u}, \mathbf{v})) = 0 \; \forall \mathbf{v} \in \mathcal{V}$ if and only if $((\mathbf{u}, \mathbf{v})) =< \nabla p, \mathbf{v} >= -(p, \nabla \cdot \mathbf{v})$, $\forall \mathbf{v} \in \mathbf{H}^1_0(\Omega)$, for a certain $p \in L^2(\Omega)$ determined up to an additive constant. The indeterminacy of these functions p, as well as that of the pressure in some incompressible flow problems, leads to the introduction of the quotient space $L^2_0(\Omega) = L^2(\Omega)/\mathbb{R}$.

The strong form of the incompressible Navier-Stokes equations considered in the previuos Section is usually understood in distribution sense, resulting in a weak formulation. For this, we need to introduce some continuous forms, defined on appropriate function spaces, associated to each term of the equations. For the viscous term, two different forms will be considered, related respectively to the Laplacian formulation (as in equation 1.7) and to the rate of deformation tensor formulation (as in equation 1.6). For the former case, one defines:

$$a(\mathbf{u}, \mathbf{v}) \doteq \nu\,((\mathbf{u}, \mathbf{v})), \quad \forall \mathbf{u}, \mathbf{v} \in \mathbf{H}^1_0(\Omega) \qquad (1.16)$$

This is a bilinear, continuous form on $\mathbf{H}_0^1(\Omega)$ which is coercive with respect to the norm $||\mathbf{u}||$, since $a(\mathbf{u},\mathbf{u}) = \nu||\mathbf{u}||$. As for the latter case, one defines:

$$\tilde{a}(\mathbf{u},\mathbf{v}) = 2\nu\, \epsilon(\mathbf{u}) : \epsilon(\mathbf{v}) \doteq 2\nu \sum_{i,j=1}^{d} (\epsilon_{ij}(\mathbf{u}), \epsilon_{ij}(\mathbf{v})), \quad \forall \mathbf{u},\mathbf{v} \in \mathbf{H}_0^1(\Omega) \qquad (1.17)$$

This form is also bilinear, continuous and coercive on $\mathbf{H}_0^1(\Omega)$, due to the Körn's inequality (see [43], page 82).

On the other hand, both the pressure gradient term and the weak form of the incompressibility condition require of the bilinear form:

$$b(\mathbf{v},q) \doteq -(q, \nabla\cdot\mathbf{v}), \quad \forall q \in L^2(\Omega),\ \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega) \qquad (1.18)$$

which is also continuous with respect to the norms $|q|$ and $||\mathbf{v}||$. Finally, the standard formulation of the convective term gives rise to a trilinear form c defined by:

$$c(\mathbf{u},\mathbf{v},\mathbf{w}) \doteq \big((\mathbf{u}\cdot\nabla)\mathbf{v},\mathbf{w}\big), \quad \forall \mathbf{u} \in \mathbf{H}^1(\Omega), \mathbf{v} \in \mathbf{H}^1(\Omega), \mathbf{w} \in \mathbf{H}_0^1(\Omega) \qquad (1.19)$$

This form is well defined and continuous on these spaces (see [105]), and is skew–symmetric in its last two arguments if $\mathbf{u} \in H$, that is, if $\nabla\cdot\mathbf{u} = 0$ and $\mathbf{n}\cdot\mathbf{u} = 0$. Moreover, c posseses some other boundedness properties, such as (see [104]):

$$c(\mathbf{u},\mathbf{v},\mathbf{w}) \leq \begin{cases} C_{111}\ ||\mathbf{u}||\ ||\mathbf{v}||\ ||\mathbf{w}|| \\ C_{012}\ |\mathbf{u}|\ ||\mathbf{v}||\ ||\mathbf{w}||_2 \\ C_{021}\ |\mathbf{u}|\ ||\mathbf{v}||_2\ ||\mathbf{w}|| \\ C_{120}\ ||\mathbf{u}||\ ||\mathbf{v}||_2\ |\mathbf{w}| \\ C_{210}\ ||\mathbf{u}||_2\ ||\mathbf{v}||\ |\mathbf{w}| \end{cases}$$

The skew component of c is also used sometimes. Calling $\tilde{c}(\mathbf{u},\mathbf{v},\mathbf{w}) \doteq \frac{1}{2}(c(\mathbf{u},\mathbf{v},\mathbf{w}) - c(\mathbf{u},\mathbf{w},\mathbf{v})), \forall \mathbf{u} \in \mathbf{H}^1(\Omega), \mathbf{v} \in \mathbf{H}_0^1(\Omega), \mathbf{w} \in \mathbf{H}_0^1(\Omega)$, one has that $\tilde{c}$ is also trilinear continuous on these spaces (see [104]), and is the weak form of the *skew–symmetric* formulation of the convective term introduced in the previous Section. It satisfies $\tilde{c}(\mathbf{u},\mathbf{v},\mathbf{v}) = 0$ for all $\mathbf{u}$ and $\mathbf{v}$.

To end this Section, let us introduce the spaces requiered for the evolution problems. Given $T > 0$, $1 \leq p < \infty$ and a Banach space W with norm $||u||_W$, the space $L^p(0,T;W)$ consists of functions $u\colon (0,T) \to W$ such that: $||u||_{L^p(0,T;W)} \doteq (\int_0^T ||u(t)||_W^p)^{1/p} < \infty$. It is also a Banach space with respect to the norm $||u||_{L^p(0,T;W)}$. The space of essentially bounded functions on $(0,T)$ into W is denoted by $L^\infty(0,T;W)$, and is a Banach space with respect to the appropriate norm. The spaces $L^p(0,T;W)$ possess similar properties as far as separability and reflexiveness is concerned as $L^p(\Omega)$ for $1 \leq p \leq \infty$ whenever W also has those properties. The case $p = 2$ is, again, special: when W is a Hilbert space with scalar product $(u,v)_W$, the space $L^2(0,T;W)$ is likewise with respect to: $(u,v) \doteq \int_0^T (u(t),v(t))_W\, dt$. Spaces like $L^2(0,T;H_0^1(\Omega))$, $L^\infty(0,T;L^2(\Omega))$ and others will often be considered.

1.3 Finite element approximation

This Section is devoted to the introduction of the basic results concerning the approximation of the previous function spaces in finite elements. We first summarize the definitions of finite element function spaces and then state an approximation theorem and an inverse inequality in these spaces, under the usual regularity assumptions on the meshes. This general theory reviewed here can be found in standard references such as [25] or [83].

We consider a partition Θ_h of Ω into elements $\{K_e\}_{e = 1,\dots,n_e}$ (n_e is the number of elements). For each element K, the diameter of K is denoted by h_K, and its sphericity (diameter of the maximum sphere inscribed in K) by ϱ_K. We also call $h \doteq \max_{K\in\Theta_h}(h_K)$ and $\varrho \doteq \min_{K\in\Theta_h}(\varrho_K)$. We assume that each element is the image of a reference element $\hat{K}$ (bounded and connected) through transformations $F_K\colon \hat{K} \to K$, which are supposed to be diffeomorphisms. Functions $\hat{v}$ defined on $\hat{K}$ are transported to K by taking $v = \hat{v} \circ F_K^{-1}$.

A finite dimensional subspace $R_k(\hat{K})$ (indexed by $k \in \mathbb{N}$) of approximating functions is chosen on $\hat{K}$; polynomial functions are usually employed. Lagrange finite elements consider the degrees of freedom on $R_k(\hat{K})$ as values of the functions at a certain set of points $\hat{\Sigma} = \{\hat{a}_j\}_{j = 1,\dots,n_n}$ of $\hat{K}$, called (reference) nodes (n_n is the number of nodes per element). These points are chosen so that the set of linear restrictions $\{\hat{p}(\hat{a}_j)\}_{j = 1,\dots,n_n}$ on $\hat{p}$ is unisolvent in $R_k(\hat{K})$, that is, their values determine $\hat{p}$ in $R_k(\hat{K})$; in particular, $\dim R_k(\hat{K}) = \#\hat{\Sigma}$.

Two classes of isoparametric finite elements (that is, those in which the transformations F_K also belong to $R_k(\hat{K})$) will be considered. For simplicial finite elements, $\hat{K}$ is the standard simplex in $\mathbb{R}^d$. In this case, $R_k(\hat{K})$ is the set of polynomials in $\{x_1,\dots,x_d\}$ of degree less than or equal to k, called P_k. It is easy to see that $\dim P_k = \left(\begin{array}{c} d+k \\ k \end{array}\right)$.

On the other hand, for quadrilateral ($d = 2$) and hexahedral ($d = 3$) finite elements, $\hat{K}$ is the unit cube $[0,1]^d$; $R_k(\hat{K})$ then consists of polynomials in $\{x_1,\dots,x_d\}$ of degree less than or equal to k in each variable, space denoted by Q_k. One has that $\dim Q_k = (k+1)^d$.

With the help of these definitions, functions defined on Ω are approximated by other functions which, in each element, are the images of polynomials in $R_k(\hat{K})$. In other words, any function space V of those considered in the previous Section is approximated by a finite dimensional subspace V_h, whose degrees of freedom are the point values at the (mesh) nodes $\Sigma_h = \{a_j\}_{j=1,\dots,n_p}$ (n_p is the number of nodal points); these are the images of the reference nodes in $\hat{K}$. When the elements K have straight sides (or plane faces, for $d = 3$) for simplicial elements, or straight and parallel sides (or faces) for quadrilaterals (and hexahedra), the transformations F_K are affine (that is, they belong to $P_1(\hat{K})$); in this case, the functions $v = \hat{v} \circ F_K^{-1}$ belong to $P_k(K)$ and $Q_k(K)$, respectively, whenever $\hat{v}$ belongs to $P_k(\hat{K})$ or $Q_k(\hat{K})$.

Discrete finite element spaces like:

$$V_h = \left\{ v \in L^2(\Omega) \ / \ \forall e = 1, \ldots, n_e, \ v_{|K_e} = \hat{v}_e \circ F_K^{-1}, \hat{v}_e \in R_k \right\} \qquad (1.20)$$

are considered. If, moreover, one requires these finite element functions to be continuous on Ω, this approximating space is spanned by the functions $\{N_i\}_{i\,=\,1,\ldots,n_p} \subset V_h$ defined through the relations: $N_i(a_j) = \delta_{ij}$; these functions N_i are called the *standard shape functions.*

For continuous functions v, a classical interpolate can be defined by:

$$v \in C^0(\bar{\Omega}) \longrightarrow \Pi_h(v) \in V_h \ / \ \Pi_h(v)(x) \doteq \sum_{j=1}^{n_p} v(a_j) N_j(x) \quad \forall \mathrm{x} \in \bar{\Omega} \quad (1.21)$$

A projection operator Π_h can also be defined on more general spaces of (not necessarily continuous) functions onto V_h (see [98], for instance). To obtain approximating properties of the operator Π_h, some restrictions have to be enforced on the meshes. A family $\{\Theta_h\}_{h>0}$ of discretizations of Ω is called regular if there exists $\zeta_1 > 0$ independent of h such that $\frac{\varrho_K}{h_K} \geq \zeta_1 > 0$, for all $K \in \Theta_h$ and for all $h > 0$. Regularity of $\{\Theta_h\}_{h>0}$ means geometrically that the elements do not collapse into segments as h tends to zero. If $\{\Theta_h\}_{h>0}$ is regular, if $v \in H^r(\Omega)$ for $r \geq 2$ and if $R_k = P_k$ or Q_k, then the following approximation result holds (see, for instance, [98]):

$$\forall m = 0, \ldots, r, \ \ ||v - \Pi_h(v)||_m \leq Ch^s ||v||_r, \qquad (1.22)$$

where $s = \min\{k + 1 - m, r - m\}$. In 1.22, and throughout this work, C represents a generic constant independent of the mesh size h, possibly depending on Ω and other constants.

Moreover, $\{\Theta_h\}_{h>0}$ is called uniformly regular, or quasi-uniform, as h tends to zero if there exists $\zeta_2 > 0$ independent of h such that $\frac{\varrho}{h} \geq \zeta_2 > 0$ for all $h > 0$. Under this condition, the following inverse inequality can be proved by scaling arguments (see [13]):

$$||v_h||_1 \leq \frac{C}{h} |v_h|, \ \forall v_h \in V_h \qquad (1.23)$$

Both the approximation result 1.22 and the inverse inequality 1.23 will be used in what follows.

1.4 Mixed problems and the LBB condition

The understanding of the properties of discrete approximations of incompressible flow problems, as well as some other related mechanical problems (such as incompressible elasticity), led to the development of a general theory of *mixed* problems. From the variational viewpoint, these are understood as *saddle point* problems, the simplest of which is the optimization of a certain quadratic functional under a linear restriction on the appropriate function space. The main example of a *mixed* problem within the context of incompressible flow equations is the steady Stokes equations 1.13-1.5 with homogeneous boundary conditions 1.10; in this case, the linear restriction is the divergence–free condition 1.5, and the *Lagrange multiplier* associated with it is the pressure. This mixed character of the equations implies that the approximating (in our case, finite element) spaces for velocity and pressure should satisfy a compatibility condition in order to obtain optimal results, as will be seen in what follows.

1.4.1 Mixed problems

A complete exposition of the theory of mixed methods, which came about with the work of I. Babuška ([4]) and F. Brezzi ([14]), has been recently given in [19], reference which we mainly follow here. A different approach can also be found in [15].

If V and Q denote two real Hilbert spaces with norms $||v||_V$ and $||q||_Q$, respectively, $a: V \times V \rightarrow \mathrm{I\!R}$ and $b: V \times Q \rightarrow \mathrm{I\!R}$ are bilinear, continuous forms with norms $||a||$ and $||b||$, respectively, and $f \in V'$, $g \in Q'$ are given, a general mixed problem consists of finding $u \in V$ and $p \in Q$ such that:

$$\begin{array}{rcll} a(u,v) + b(v,p) & = & < f, v >, & \forall v \in V \\ b(u,q) & = & < g, q >, & \forall q \in Q \end{array} \qquad (1.24)$$

The study of this problem leads to introduce the forms:

$$B: V \rightarrow Q' \; / \; < B(v), q >_{Q' \times Q} = b(v,q) \quad \forall q \in Q, \; \forall v \in V$$

$$B^t: Q \rightarrow V' \; / \; < B^t(q), v >_{V' \times V} = b(v,q) \quad \forall v \in V, \; \forall q \in Q$$

Assuming that a is coercive on V, that is, $a(u,u) \geq \beta_a ||u||_V \;\; \forall u \in V$, the conditions for a solution of 1.24 to exist are that $g \in \operatorname{Im} B$ and that there exists a constant β_b such that:

$$\inf_{q \in Q} \Big(\sup_{v \in V} \frac{b(v,q)}{||v||_V \; ||q||_{Q/\mathrm{Ker}B^t}} \Big) \geq \beta_b > 0, \qquad (1.25)$$

in which case u is unique and p is determined up to an arbitrary element of $\mathrm{Ker}B^t$. Condition 1.25 is usually refered to as the *inf–sup* or LBB condition, after the work of O.A. Ladyzhenskaya ([71]), I. Babuŝka ([4]) and F. Brezzi ([14]).

The Stokes problem is cast into this framework by taking $V = \mathbf{H}_0^1(\Omega)$, $Q = L^2(\Omega)$, a and b defined by 1.16 and 1.18, respectively, $g = 0$ and

$f \in \mathbf{H}^{-1}(\Omega)$ given. In this case, the form a is coercive and symmetric; the forms B and B^t are defined by:

$$B\colon \mathbf{H}_0^1(\Omega) \to L^2(\Omega) \ / \ B(\mathbf{v}) = \nabla \cdot \mathbf{v}, \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega)$$

$$B^t\colon L^2(\Omega) \to \mathbf{H}^{-1}(\Omega) \ / \ B^t(q) = \nabla q, \quad \forall q \in L^2(\Omega)$$

One then has that KerB^t = $\{q \in L^2(\Omega) \ / \ q \text{ is constant on } \Omega\}$, space isomorphic to $\mathbb{R}$. Condition 1.25 was first proved for this case by O.A. Ladyzenskaya (see [71]); existence and uniqueness of the velocity solution $\mathbf{u}$ and existence of the pressure p defined up to an additive constant are thus established (we are considering the Dirichlet case). This indeterminacy in the pressure is usually surpassed by working on $Q = L_0^2(\Omega)$, where B^t is injective; however, in the discrete problem other linear restrictions may be used (such as fixing an arbitrary discrete value of the pressure to zero).

1.4.2 Discrete approximations

We now turn to the consideration of an approximate discrete solution of the mixed problem 1.24, where several difficulties may be encountered. Some of these were observed in practice in the early stages of Computational Fluid Dynamics, before this theory was even developed. We outline the basic results in what follows, inspired again in [19].

Let V_h and Q_h denote finite dimensional subspaces of V and Q, respectively, where the index h refers to a mesh size. The discrete version of problem 1.24 reads: find $u_h \in V_h$ and $p_h \in Q_h$ such that:

$$\begin{array}{rcll} a(u_h, v_h) \ + \ b(v_h, p_h) & = & < f, v_h >, & \forall v_h \in V_h \\ b(u_h, q_h) & = & < g, q_h >, & \forall q_h \in Q_h \end{array} \qquad (1.26)$$

Let B_h and B_h^t denote the discrete equivalent to the operators B and B^t on V_h and Q_h. For a given $g \in Q'$, let:

$$Z_h(g) \ \doteq \ \{v_h \in V_h \ / \ b(v_h, q_h) \ = < g, q_h >, \ \forall q_h \in Q_h\}$$

Then, if a is coercive on V, if $Z_h(g) \neq \emptyset$ and if the discrete equivalent of the LBB condition holds with a constant β_0 independent of h:

$$\inf_{q_h \in Q_h} \left(\sup_{v_h \in V_h} \frac{b(v_h, q_h)}{||v_h||_V \ ||q_h||_{Q/\mathrm{Ker}B^t}} \right) \geq \beta_h \geq \beta_0 > 0, \qquad (1.27)$$

then there exists a unique $u_h \in V_h$ and a $p_h \in Q_h$, defined up to an arbitrary element of KerB_h^t, solution of 1.26 which satisfy the following *optimal approximation* property:

$$||u-u_h||_V + ||p-p_h||_{Q/\mathrm{Ker}B^t} \leq C\Big(\inf_{v_h \in V_h} ||u-v_h||_V + \inf_{q_h \in Q_h} ||p-p_h||_{Q/\mathrm{Ker}B^t}\Big) \tag{1.28}$$

with constant C depending on $||a||$, $||b||$, β_a and β_0, but not on h.

Unfortunately, the discrete LBB condition 1.27 does not hold for simple combinations of finite element spaces for velocity and pressure, such as equal order ones. Those for which 1.27 holds are called *div-stable* in the terminology of [12]. Problems may develop in the following circumstances:

- The constant β_h in 1.27 may not be bounded away from zero uniformly in h; in this case, it is interesting to know the exact dependence of β_h with respect to h, so that weaker error estimates than 1.28 may be obtained.

- $\mathrm{Ker}B_h^t \not\subset \mathrm{Ker}B^t$; in this case, the discrete solution p_h may be *polluted* with unphysical (non constant) modes, called *spurious pressure modes.* Moreover, $Z_h(g)$ may be empty, in particular for some nonhomogeneous Dirichlet boundary conditions, leading to ill–posed discrete problems.

- $(\dim Q_h - 1) > \dim V_h$; this case leads to *locking* of the solution, since there are more restrictions on it than degrees of freedom. The only discrete divergence free vector field is the null one.

The problem of spurious pressure modes is, in essence, an algebraic problem. Calling K the matrix associated to the discretization of the form a, G the *discrete gradient* matrix and G^t the *discrete divergence* matrix, problem 1.26 can be written as:

$$\begin{pmatrix} K & G \\ G^t & 0 \end{pmatrix} \begin{pmatrix} U \\ P \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \tag{1.29}$$

where U and P are the vectors of discrete values of velocity and pressure, respectively, and F_1, F_2 come from external forces and (nonhomogeneous) boundary conditions. The system matrix of 1.29 will have a nontrivial kernel when $\mathrm{Ker}B_h^t \not\subset \mathrm{Ker}B^t$, since spurious modes satisfy $GP = 0,\ P \neq 0$.

Some of the most popular examples of mixed finite elements for incompressible flows are listed next, classified according to whether the discrete pressure is continuous on Ω or not; some of the terminology used applies only to the two–dimensional case:

1. Discontinuous pressure quadrilateral elements.

   (a) Q_1P_0 element: the bilinear–velocity, constant–pressure element does not satisfy the discrete LBB condition. For a regular mesh, the kernel of the discrete gradient matrix is two–dimensional, containing two independent spurious modes which are constant on the *red and white* cells of the mesh, viewed as a checkboard; they are called *checkboard modes*. The constant β_h can be shown to be $O(h)$ in this case (see [78]). A thorough study of the properties of this element was given in [85].

(b) Q_2P_1 element: the biquadratic–velocity, linear–pressure element is a *div–stable* element commonly used in practice. The pressure values on each element can be understood as the pressure and its first order spatial derivatives at the centroid of the element, in a hierarchical way. In general, the element Q_lP_{l-1} is *div–stable* for $l \geq 2$ (see [43]).

2. Continuous pressure elements.

(a) Equal order interpolations: elements where the velocity and pressure are interpolated by continuous functions on the same mesh points and to the same order of accuracy, such as Q_1Q_1 on quadrilaterals or P_1P_1 on triangles, are the simplest ones to implement; however, these elements also present spurious pressure modes, and yield unstable pressures which need to be filtered to get accurate results. A study of spurious modes for these elements was given in [86].

(b) Taylor–Hood elements: it was found experimentally that a P_2P_1 continuous pressure approximation on triangles and a Q_2Q_1 on quadrilaterals yielded stable and convergent results. Some analysis of these elements were given in [9] and [109], and an extension to P_3P_2 and Q_kQ_{k-1} (for $k \geq 2$) in [18].

1.4.3 Some stabilizing techniques

Several alternatives have been proposed to overcome the difficulties introduced by mixed methods. Efforts have been directed into three main directions:

- The development of *div–stable* finite element combinations, some of which we have just seen.

- The stabilization of known unstable elements, in particular the Q_1P_0 element, through the use of appropriate filtering techniques ([95], [94]), the use of macroelements ([97]) or by enriching the velocity space by bubble functions ([3]).

- Obtaining alternative formulations of the original equations, which *circumvent* the LBB restrictions, either by employing non primitive variables (vorticity, streamfunction or others) or by *augmented* or *stabilized* formulations. We summarize this last possibility for its relevance.

Stabilized formulations of the Stokes and incompressible Navier–Stokes equations developed from the original work on SUPG formulation for advective diffusive problems in [20] (which in turn was an extension of previous work on upwind finite differences), and the series of papers [57], [58] and [59]. The consistent Galerkin Least Squares (GLS) formulation then came about ([60], [61]). When applied to incompressible flow problems, it allows the use of arbitrary velocity–pressure elements. For a given mesh $\Theta_h = \{K_e\}_{e=1,\ldots,n_e}$ with element sizes h_K, the GLS method for the Stokes problem can be understood as a modification of the discrete problem 1.26 by adding to each equation a multiple of the strong form of the momentum equation:

$$\begin{aligned} a(\mathbf{u}_h,\mathbf{v}_h) + b(\mathbf{v}_h,p_h) &+ \sum_{K\in\Theta_h} \alpha_K(-\Delta\mathbf{u}_h + \nabla p_h - \mathbf{f}, -\Delta\mathbf{v}_h)_K \\ &= \; < f, \mathbf{v}_h >, \qquad \forall \mathbf{v}_h \in V_h \qquad (1.30) \\ b(\mathbf{u}_h,q_h) &+ \sum_{K\in\Theta_h} \alpha_K(-\Delta\mathbf{u}_h + \nabla p_h - \mathbf{f}, \nabla q_h)_K = 0, \;\; \forall q_h \in Q_h \end{aligned}$$

where $\alpha_K > 0 \; \forall K \in \Theta_h$ (we have neglected the term $\tau_2(\nabla\cdot\mathbf{u},\nabla\cdot\mathbf{v})$ from 1.30, which appears in the definition of the method, since it turns out to be unnecessary). These new integrals are evaluated on element interiors, where the approximating functions are sufficiently differentiable. Equations 1.30 are the Euler–Lagrange equations of a saddle point problem with a Lagrangian augmented by the addition, at element level, of a multiple of the square of the residual of equation 1.13, residual which is to be minimized; this is why this method is called Galerkin Least Squares.

Stability and optimal convergence for this and related methods were proved in [37] and [17], in mesh dependent norms such as:

$$|||(\mathbf{u},p)|||^2_{GLS} \doteq ||\mathbf{u}||^2 + \sum_{K\in\Theta_h} h_K^2 |\nabla p|_K^2$$

for any choice of V_h and Q_h; the coefficients α_K (called τ_1 in the usual terminology) are obtained as the diffusive limit of those of a similar method for the Navier–Stokes case, yielding, after some simplifications (see [37]):

$$\alpha_K = \alpha_0 \frac{h_K^2}{4\nu} \qquad (1.31)$$

For linear elements and small enough h_K, a value of $\alpha_0 = 1/3$ is optimal (see [37]). For quadratic elements, an optimal value of $\alpha_0 = 1/9$ was obtained in [27] for a related scheme.

The stabilization of the pressure in these residual methods is mainly due to the appearance of a nonzero diagonal term on the system matrix of the discrete problem 1.29 multiplying the pressure, which comes from the term $(\nabla p_h, \nabla q_h)$ in 1.30.

A great amount of work has been developed recently on stabilized methods (see [5], [16], [35], [36] and [107], for instance).

1.5 Description of fractional step methods

In this last Section we present several methods of fractional step type for the time integration of the unsteady, incompressible Navier–Stokes equations 1.7–1.5, with homogeneous Dirichlet boundary conditions 1.10 (for simplicity of exposition) and initial condition 1.11.

The common feature to these methods is the decomposition of each time advancement step into a sequence of two or more substeps. The way this decomposition is chosen in each method determines properties such as its stability, convergence, order of accuracy in the time step, steady state reached (if so), boundary conditions to be imposed in each substep, stabilization or not of the pressure and type of fully discrete problems actually solved, as will be explained in what follows.

We present the methods classified into four categories established according to different criteria, which may well overlap with one another. Given the great amount of fractional step methods developed nowadays, this presentation does not pretend to be exhaustive, but rather a wide view of the variety of existing methods, paying special attention to the most significant ones; besides, the classification could also respond to other criteria, but the ones chosen here emphasize some ideas to which we will come back later on.

Some representative methods of each category are explained in more detail. They are presented respecting as much as possible the structure and notation used in their original references. Some of them are introduced directly in fully discrete form, after some space discretization (finite differences, finite elements, finite volumes or spectral methods) has already been performed. However, we are mainly concerned with their semidiscrete formulations, which are more general, not depending on the particular form of space discretization used, and more suitable for the study of properties intrinsic to the time integration process, such as well–posedness of the intermediate problems or appropriate boundary conditions for them.

In what follows, we assume that a constant time step $\delta t > 0$ is given, and define the time levels $t_n = n\,\delta t$ for $n = 0, \ldots, [T/\delta t]$.

1.5.1 Classical Projection Methods

We group here the first fractional step methods to appear in the literature that gave rise to this kind of methods, as well as some of their closest variants and studies.

The original ideas of fractional step methods for general evolution equations go back to the work of Yanenko (see [110]). The concept of a splitting of the different operators appearing in the equations in succesive steps was first introduced there. In the early times, this splitting was usually associated to the different space dimensions. An interpretation of this general splitting in the case of the incompressible Navier–Stokes equations can be found in the *scheme with* $(n+1)$ *intermediate steps* of Temam (see III.7.2 in [105]).

But the actual origin of fractional step methods for Navier–Stokes equations is generally credited to the work of Chorin (see [22], [23] and [24]) and Temam (see [100], [101], [102] and [103]). The former is a 3–substep method, in which the first two substeps can be thought of as an ADI scheme (Alterning Directions Implicit) and the third one is a projection onto the subspace of solenoidal vector fields (in a sense to be explained in what follows); for the case of periodic boundary conditions in a unit cube and for a centered finite difference space approximation, it was shown in [24] that provided $\delta t = O(h^2)$, the convergence of this method was of first order in δt and second order in h. We present Temam's method in more detail, since it is the most popular of fractional step methods and one of the best studied. We name it the *classical method*, and follow the presentation of III.7.1 in [105].

Let us call $k = \delta t$, and assume that $\mathbf{f} \in \mathbf{L}^2(0, T; H)$ and $\mathbf{u}_0 \in H$. Given $\mathbf{u}^n \in H$, approximation of $\mathbf{u}$ at time t_n, the first step of the *classical method* consists of finding an *intermediate* velocity $\mathbf{u}^{n+1/2}$ such that:

$$\begin{aligned} \frac{\mathbf{u}^{n+1/2} - \mathbf{u}^n}{k} - \nu\Delta\mathbf{u}^{n+1/2} + (\mathbf{u}^{n+1/2} \cdot \nabla)\mathbf{u}^{n+1/2} &+ \\ \frac{1}{2}(\nabla \cdot \mathbf{u}^{n+1/2})\mathbf{u}^{n+1/2} &= \mathbf{f}^n \qquad (1.32) \\ \mathbf{u}^{n+1/2}_{|\Gamma} &= 0 \end{aligned}$$

An implicit backward Euler method is employed for the diffusive term, and the skew–symmetric form adopted for convection is also approximated implicitly. The force term $\mathbf{f}^n$ is the time average of $\mathbf{f}$ in $[t_n, t_{n+1}]$. On $\mathbf{u}^{n+1/2}$, the full Dirichlet boundary condition is imposed. The weak formulation of 1.32 consists of finding $\mathbf{u}^{n+1/2} \in \mathbf{H}^1_0(\Omega)$ such that for all $\mathbf{v} \in \mathbf{H}^1_0(\Omega)$:

$$\frac{1}{k}(\mathbf{u}^{n+1/2} - \mathbf{u}^n, \mathbf{v}) + a(\mathbf{u}^{n+1/2}, \mathbf{v}) + \tilde{c}(\mathbf{u}^{n+1/2}, \mathbf{u}^{n+1/2}, \mathbf{v}) = (\mathbf{f}^n, \mathbf{v}) \qquad (1.33)$$

Once $\mathbf{u}^{n+1/2}$ is determined, the second step consists of finding an *end–of–step* velocity $\mathbf{u}^{n+1}$ and a function p^{n+1} such that:

$$\begin{aligned} \frac{\mathbf{u}^{n+1} - \mathbf{u}^{n+1/2}}{k} + \nabla p^{n+1} &= 0 \qquad (1.34) \\ \nabla \cdot \mathbf{u}^{n+1} &= 0 \qquad (1.35) \\ \mathbf{n} \cdot \mathbf{u}^{n+1}_{|\Gamma} &= 0 \qquad (1.36) \end{aligned}$$

This is equivalent to saying that $\mathbf{u}^{n+1}$ is the projection of $\mathbf{u}^{n+1/2}$ onto the space H, so that equations 1.34–1.35–1.36 can also be written as $\mathbf{u}^{n+1} = P_H(\mathbf{u}^{n+1/2})$. This is the reason why this method, and other related schemes, is usually called the *projection method*.

As can be seen, the splitting of operators in this case consists of separating the effects of incompressibility from those of diffusion and convection, which are kept together.

This *classical projection method* posseses some advantages over standard single–step methods. From the computational standpoint, the main one is the decoupling of the computation of the 'pressure' p^{n+1} from that of the velocity; this is achieved with the help of a 'pressure Poisson equation' (PPE, from now on), obtained from equation 1.34. In fact, taking the divergence of 1.34 leads to:

$$\begin{aligned} \Delta p^{n+1} &= \frac{1}{k}\nabla \cdot \mathbf{u}^{n+1/2} && (1.37) \\ \mathbf{n} \cdot \nabla p^{n+1}_{|\Gamma} &= 0 && (1.38) \end{aligned}$$

Once this Neumann problem is solved for p^{n+1}, the final velocity $\mathbf{u}^{n+1}$ is obtained explicitly from 1.34. Another advantage of this scheme is that the space discretization used in combination with it is not restricted by the compatibility (*inf–sup*) condition encountered in the Stokes problem; this fact has been observed by some authors who have used this method together with different space discretizations, including some equal order finite elements. But, to the author's knowledge, the reason for this pressure stabilization has not yet been fully explained. We provide an explanation for it in Chapter 2.

The *classical projection scheme*, however, presents some drawbacks too. As can be seen in 1.36, the final velocity $\mathbf{u}^{n+1}$ does not satisfy the correct Dirichlet boundary condition, but only the normal component of it. This may result in the presence of a numerical boundary layer in the solution, whose size has been estimated to be $O(\sqrt{\nu\delta t})$ (see [45], [79] and [106]). Another side of the same problem is the need to impose the unphysical homogeneous Neumann boundary condition on p^{n+1}, while the exact pressure satisfies, for sufficiently smooth solutions (see [48]):

$$\begin{aligned} \Delta p(t) &= \nabla \cdot (\mathbf{f}(t) - (\mathbf{u}(t) \cdot \nabla)\mathbf{u}(t)) && (1.39) \\ \mathbf{n} \cdot \nabla p(t)_{|\Gamma} &= \mathbf{n} \cdot (\mathbf{f}(t) + \nu\Delta\mathbf{u}(t) - \frac{\partial \mathbf{u}}{\partial t} - (\mathbf{u} \cdot \nabla)\mathbf{u})_{|\Gamma} && (1.40) \end{aligned}$$

This has led several authors to believe that p^{n+1} is not an approximation of $p(t_{n+1})$, but a mere auxiliary mathematical variable needed to enforce the incompressibility condition on $\mathbf{u}^{n+1}$ (a Lagrange multiplier), although both $\mathbf{u}^{n+1/2}$ and $\mathbf{u}^{n+1}$ are legitimate approximations of $\mathbf{u}(t_{n+1})$ (see [106]).

Convergence of this scheme to a continuous solution $\mathbf{u}$ was proved by Temam (see [101] and [102]). He introduced the following approximating functions:

$\mathbf{u}^1_k: [0,T] \rightarrow \mathbf{L}^2(\Omega)$ / $\mathbf{u}^1_k(t) = \mathbf{u}^{n+1/2}$, $nk \leq t < (n+1)k$
$\mathbf{u}^2_k: [0,T] \rightarrow \mathbf{L}^2(\Omega)$ / $\mathbf{u}^2_k(t) = \mathbf{u}^{n+1}$, $nk \leq t < (n+1)k$

$\mathbf{u}_k: [0,T] \to \mathbf{L}^2(\Omega)$ / $\mathbf{u}_k$ is continuous, linear on t on each interval $[nk,(n+1)k]$ and $\mathbf{u}_k(t_n) = \mathbf{u}^n$, for $n = 0,\dots,[T/k]$.

He proved convergence of $\mathbf{u}_k^1$ (that is, of $\mathbf{u}^{n+1/2}$) to $\mathbf{u}$ in $\mathbf{H}_0^1(\Omega)$; but for $\mathbf{u}_k^2$ and $\mathbf{u}_k$, the convergence was only in $\mathbf{L}^2(\Omega)$, and this was due to the fact that $\mathbf{u}^{n+1}$ does not satisfy the correct boundary condition. In fact, he had:

- if $d=2$, $\mathbf{u}_k^i$ $(i=1,2)$ and $\mathbf{u}_k$ converge to $\mathbf{u}$ in $L^2(0,T;\mathbf{L}^2(\Omega))$ strongly as k tends to 0, and weak–star in $L^\infty(0,T;\mathbf{L}^2(\Omega))$; $\mathbf{u}_k^1$ converges to $\mathbf{u}$ in $L^2(0,T;\mathbf{H}_0^1(\Omega))$ strongly.

- if $d=3$, there exists a subsequence k' of k such that $\mathbf{u}_{k'}^i$ $(i=1,2)$ and $\mathbf{u}_{k'}$ converge to $\mathbf{u}$ in $L^2(0,T;\mathbf{L}^2(\Omega))$ strongly as k' tends to 0, and weak–star in $L^\infty(0,T;\mathbf{L}^2(\Omega))$; $\mathbf{u}_{k'}^1$ converges to $\mathbf{u}$ in $L^2(0,T;\mathbf{H}_0^1(\Omega))$ weakly.

Further studies of this method have been performed by other authors. The most relevant one is the work of J. Shen: in [90] he considered the *classical projection method* with a slightly different formulation of the convective term, namely, $(\mathbf{u}^n\cdot\nabla)\mathbf{u}^{n+1/2}$. This results in a skew–symmetric weak form $c(\mathbf{u}^n,\mathbf{u}^{n+1/2},\mathbf{v})$, since $\mathbf{u}^n \in H$. For this scheme, he proved first order error estimates in the time step for $\mathbf{u}^{n+1/2}$ and $\mathbf{u}^{n+1}$, and order 1/2 error estimates for p^{n+1} and $p^{n+1} - k\nu\Delta p^{n+1}$, in the appropriate sense. A mistake in the original proof pointed out by J.L. Guermond (see [50]) was corrected in [92]. The definitions of order of approximation employed in these proofs are as follows: given a Banach space X with norm $||\cdot||_X$, a continuous function $f: [0,T] \to X$ and a partition $\{t_n^k\}_{n\,=\,0,\dots,N}$ of $[0,T]$ whose maximum step tends to zero as k tends to zero, a function $f_k: [0,T] \to X$ is a weakly order α approximation of f in X if there exists $C>0$ independent of k such that:

$$k\sum_{n=0}^{N} ||f_k(t_n^k) - f(t_n^k)||_X^2 \le Ck^{2\alpha}$$

On the other hand, f_k is a strongly order α approximation of f in X if:

$$||f_k(t_n^k) - f(t_n^k)||_X^2 \le Ck^{2\alpha}, \ \forall n = 0,\dots,N$$

It was proved in [90] that both $\mathbf{u}^{n+1/2}$ and $\mathbf{u}^{n+1}$ are weakly first order approximations of $\mathbf{u}$ in $\mathbf{L}^2(\Omega)$, and that p^{n+1} and $p^{n+1} - k\nu\Delta p^{n+1}$ are weakly order 1/2 approximations of p in $L_0^2(\Omega)$. Once again, the incorrect boundary condition satisfied by $\mathbf{u}^{n+1}$ forbids to get satisfactory error estimates in $\mathbf{H}_0^1(\Omega)$.

In order to get improved error estimates, a modified scheme was also considered in [90]. It consists of adding the term ∇p^n to 1.32 and regarding the Lagrange multiplier of equation 1.34 as a *pressure correction*, rather than an end–of–step pressure, that is, $\phi\nabla(p^{n+1} - p^n)$ for some $\phi > 0$. In this case, both $\mathbf{u}^{n+1/2}$ and $\mathbf{u}^{n+1}$ are strongly first order approximations of $\mathbf{u}$ in $\mathbf{L}^2(\Omega)$, whereas p^{n+1} and $p^{n+1} - k\nu\Delta p^{n+1}$ are weakly first order approximations of p in $L^2_0(\Omega)$.

More recently, a general framework for these *classical projection methods* has been introduced in [96], where it is shown that the classical method, among others, with Shen's formulation of the convective term, is unconditionally stable (in the appropriate sense). A similar scheme was also considered in [10], but it was obtained by arguments of approximate matrix factorization, and proved to be equivalent to a form of fractional step method. The *classical projection method* was also used in [29] in conjunction with a finite volume space discretization on unstructured triangular meshes.

1.5.2 Higher Order Methods

We have seen how simple splittings of the incompressible Navier–Stokes equations generally lead to first order schemes in the time step. Several alternatives have been proposed to achieve higher order methods; we present some of the most outstanding ones.

To the author's knowledge, there are three main ways to develop higher order fractional step methods, which are the use of improved velocity boundary conditions, improved pressure boundary conditions and pressure correction, respectively, all of them developed in the mid–eighties.

As a representative of fractional step methods with improved velocity boundary conditions we consider the work of Kim and Moin (see [65]). It consists of the following fully discrete scheme, where a centered finite difference space approximation on a staggered grid is assumed (some staggered grid finite differences are equivalent to Q_1P_0 finite element discretizations with mass lumping):

$$\frac{u_i^{n+1/2} - u_i^n}{\delta t} = \frac{1}{2}\frac{1}{\mathrm{Re}}L(u_i^{n+1/2} + u_i^n) + \frac{1}{2}(3H_i^n - H_i^{n-1}) \qquad (1.41)$$

$$\frac{u_i^{n+1} - u_i^{n+1/2}}{\delta t} = -G\phi^{n+1}, \quad Du_i^{n+1} = 0 \qquad (1.42)$$

In 1.41 and 1.42 u_i represents the nodal vector containing the i–th component of velocity, H is a discretization of the conservative form of the convective operator, which is approximated by an explicit, second order multistep Adams–Bashforth method, and L is a centered discretization of the Laplacian operator (a second order, implicit Crank–Nicholson method is employed for diffusion, which enhances stability for low Reynolds number flows). In 1.42, G and D are the discrete gradient and divergence operators, respectively; this equation can be viewed as a projection step, and is actually solved by a discrete PPE. The main novelty of this scheme, however, is the boundary conditions imposed on the *intermediate* velocity: in the homogeneous Dirichlet case that we are considering, these are $u_i^{n+1/2} = \dfrac{\partial \phi^n}{\partial x_i}$, i.e., $\mathbf{u}^{n+1/2} = \nabla\phi^n$, obtained through a Taylor expansion of $\mathbf{u}^{n+1/2}$ (see [65]). An improvement of this method was presented in [74], where a three–step Runge–Kutta scheme was considered in which each step is decomposed into two fractional substeps in a similar manner to Kim and Moin's method. Improved velocity boundary conditions for fractional–step methods were also studied by M. Fortin and coworkers in [41].

Improved pressure boundary condition fractional–step methods stem from the work of Orzag, Israelli and Deville (see [79]). For a one dimensional linear model with no convection, a two–step method is devised in reversed order, that is, with a projection step first and a diffusion implicit Crank–Nicholson step second. The novelty, this time, is the second order boundary condition $\mathbf{n}\cdot\nabla p^{n+1} = -\nu\mathbf{n}\cdot(\nabla\times(\nabla\times\mathbf{u}^n))$ employed (the *rot–rot* form of the viscous term is used), which is closer to the continuous boundary condition 1.40 than the homogeneous Neumann condition 1.38. Several generalizations of this idea can be found in [63]; in this reference, a three–step method is considered consisting of an explicit Adams–Bashforth step for convection, followed by a projection step and an implicit Adams–Moulton step for diffusion. The projection step is solved via a continuous PPE with higher order pressure boundary conditions obtained from the continuous one. Stiffly stable schemes are also considered for the time derivative term, which enhance stability. All these methods are used in combination with a spectral element space discretization, and an extension to triangular spectral elements is provided in [93].

A second order pressure–correction fractional–step method was introduced by van Kan in [62]. It was developed for a system of ordinary differential equations with a linear constrain, representing a finite difference approximation of the Navier–Stokes equations on a uniform staggered-grid. Namely, he considered a system of the form $\dot{x} = f(x) + Gp$, $G^t x = g(t)$. In this context, the pressure correction methods reads:

$$\begin{aligned}\frac{x^{n+1/2}-x^n}{\delta t} &= \tfrac{1}{2}(f(x^{n+1/2})+f(x^n)) + Gp^n\\ \frac{x^{n+1}-x^{n+1/2}}{\delta t} &= \tfrac{1}{2}G(p^{n+1}-p^n), \quad G^t x^{n+1} = g^{n+1}\end{aligned}$$

The second step is actually solved by a discrete PPE:

$$\frac{1}{2}G^tG(p^{n+1}-p^n) = \frac{1}{\delta t}(g^{n+1} - G^t x^{n+1/2})$$

It is shown in [62] that the solution (x^n, p^n) of this split scheme differs from the solution of a coupled Crank–Nicholson method by $O(\delta t)^2$, so that this is also a second order method. A linearization of the convective term is used for the extension of the method to the Navier–Stokes equations. Van Kan's method was recently used in [81] with a spectral method for the space variables, in which the same mesh points were used for velocity and pressure.

Another second order projection method was introduced by Bell, Colella and Glaz in [8]. They considered an iterative scheme in each time step which converges to the solution of a coupled Crank–Nicholson scheme. Each iteration is decomposed into two substeps, the first one being a convective–diffusive step, which is explicit in convection and implicit in diffusion, and the second one an incompressibility step; thus, the k–th iteration of the scheme is split as follows (see [8]):

$$\begin{aligned} \frac{\mathbf{u}^{*,k}-\mathbf{u}^n}{\delta t} &= \frac{\nu}{2}\Delta(\mathbf{u}^{*,k}+\mathbf{u}^n) - [(\mathbf{u}\cdot\nabla)\mathbf{u}]^{n+1/2} \\ &\quad -\nabla p^{n+1/2,k} \\ \frac{\mathbf{u}^{n+1,k+1}-\mathbf{u}^n}{\delta t} + \nabla p^{n+1/2,k+1} &= \frac{\nu}{2}\Delta(\mathbf{u}^{*,k}+\mathbf{u}^n) - [(\mathbf{u}\cdot\nabla)\mathbf{u}]^{n+1/2} \\ \nabla\cdot\mathbf{u}^{n+1,k+1} &= 0 \end{aligned}$$

It is assumed that the convective term is computable from the velocity at time t_n and the current approximation of the pressure $p^{n+1/2,k}$ by an explicit, second order Godunov procedure (see [8]), and a standard finite difference approximation is used for the Laplacian term.

Several fractional step projection methods were studied by P. Gresho in [45]. They include *optimal* projection methods, with *optimal* boundary conditions for velocity and pressure, and simpler projection schemes, of which there is a first order (*Projection* 1, equivalent to the *classical projection method*), a second order (*Projection* 2, related to van Kan's, Kim and Moin's and Bell, Colella and Glaz's methods) and a third order version (*Projection* 3). These are presented in continuous, semidiscrete and fully discrete forms (in [46]), the latter with a Q_1P_0 finite element interpolation. Gresho's *Projection* 2 method also employs pressure correction.

More recently, error estimates of some of these and other higher order splitting methods were proved by J. Shen in [91], in a similar way to [90] and with the modifications of [92]. Roughly speaking, he showed that, under several hypothesis:

- A pressure correction method similar to van Kan's provided weakly order $(2-\epsilon)$ and strongly order $(3/2-\epsilon)$ approximations to the velocity in $\mathrm{L}^2(\Omega)$ and weakly order $(3/2-\epsilon)$ in $\mathrm{H}^1(\Omega)$ for any $\epsilon > 0$, both for the *intermediate* and the *end–of–step* velocities; he also proved weakly order $(3/2-\epsilon)$ error estimates for the pressure in $L^2_0(\Omega)$.

- A method similar to Kim and Moin's (for the linear unsteady Stokes problem) provided weakly order 3/2 approximations to the velocity in $\mathrm{L}^2(\Omega)$ both for the *intermediate* and the *end–of–step* velocities, and weakly order 1 estimates for a modified pressure in $L^2_0(\Omega)$.

- A penalty–projection scheme with pressure correction provided weakly order 2 approximations to the velocity in $\mathrm{L}^2(\Omega)$ for the *intermediate* velocity, strongly order 3/2 in $\mathbf{L}^2(\Omega)$ for the *end-of-step* velocity and weakly order 3/2 in $\mathbf{H}^1(\Omega)$ for both; the pressure was found to be weakly order 3/2 accurate.

To end this subsection, let us mention the work of Dukovicz and Dvinsky (see [32]), where some higher order splitting methods are developed by arguments of approximate matrix factorizations.

1.5.3 Viscosity splitting methods

We have seen that most fractional step methods employ a projection step at some point of the calculations, thus uncoupling the effects of incompressibility from all the other terms in the equations. Some other fractional step methods, however, do not fully uncouple incompressibility from diffusion, still splitting it from convection. We call them *viscosity splitting* methods.

As a clear example of this kind of methods, we consider the work of R. Natarajan (see [77]). Developed from a general splitting of operators for linear evolution equations, it consists of the following three-step procedure:

$$\begin{aligned}
\frac{\mathbf{u}^* - \mathbf{u}^n}{\lambda_1 \delta t} - \theta\nu\Delta\mathbf{u}^* + \nabla p^* &= \mathbf{f} + (1-\theta)\nu\Delta\mathbf{u}^n - (\mathbf{u}^n\cdot\nabla)\mathbf{u}^n \\
\nabla\cdot\mathbf{u}^* &= 0 \\
\mathbf{u}^*_{|\Gamma} &= 0
\end{aligned}$$

$$\begin{aligned}
\frac{\mathbf{u}^{**} - \mathbf{u}^*}{\lambda_2 \delta t} - (1-\theta)\nu\Delta\mathbf{u}^{**} + (\mathbf{u}^{**}\cdot\nabla)\mathbf{u}^{**} &= \mathbf{f} + \theta\nu\Delta\mathbf{u}^* - \nabla p^* \\
\mathbf{u}^{**}_{|\Gamma} &= 0
\end{aligned}$$

$$\begin{aligned}
\frac{\mathbf{u}^{n+1} - \mathbf{u}^{**}}{\lambda_1 \delta t} - \theta\nu\Delta\mathbf{u}^{n+1} + \nabla p^{n+1} &= \mathbf{f} + (1-\theta)\nu\Delta\mathbf{u}^{**} - (\mathbf{u}^{**}\cdot\nabla)\mathbf{u}^{**} \\
\nabla\cdot\mathbf{u}^{n+1} &= 0 \\
\mathbf{u}^{n+1}_{|\Gamma} &= 0
\end{aligned}$$

The parameters θ, λ_1 and λ_2 can be chosen to yield first and second order accurate methods. As can be seen, in the first and third substeps an implicit approximation of the viscous term is used together with the incompressibility condition, with the help of a Lagrange multiplier related to the pressure, and an explicit approximation of the convective term is considered. The second substep is a nonlinear problem, which is fully implicit both in convection and diffusion. This algorithm is discretized in space with a Q_2P_1 finite element interpolation.

The method just explained is similar to the well known *θ–method* of R. Glowinsky and others (see [44]). The convergence of two fully discrete *θ–schemes* to a continuous solution was first proved by E. Fernández–Cara and M. Marín (see [38]), where stability restrictions on the time step were also provided. Other stability and convergence results were proved in [67], assuming a *div–stable* mixed finite element interpolation. The *θ–scheme* was also considered in [108], named as $T3$ method; a $T6$ method was developed there, in which each substep of $T3$ was split into two, so as to apply efficient SUPG techniques to all convective terms appearing in the equations.

The three–step Runge–Kutta scheme of [74] mentioned in the previuos subsection can also be considered a *viscosity splitting* method. In it, each step is decomposed into two substeps: the first one is implicit in viscosity and explicit in convection; the second one is also implicit in viscosity and coupled with incompressibility. The implicitness parameters and boundary conditions are chosen so as to achieve second order accuracy.

Other viscosity splitting schemes were studied by L–a Ying in a series of papers in a continuous formulation (see [76] and the references therein). For one of them, and in the 2–dimensional case, he proved $O(\delta t)$ error estimates for both the *intermediate* and the *end–of–step* velocities in $L^\infty(0,T;\mathbf{H}_0^1(\Omega))$.

Finally, a linearized stability analysis for a fully discrete, staggered–grid finite difference two–step scheme was given in [70]; in this case, the second step is also implicit in viscosity and coupled with incompressibility.

1.5.4 Other methods

We briefly review here other fractional step methods also present in the literature.

A two–step projection scheme was considered by J. Donea *et al.* in [30], where the first step was explicit both in convection and diffusion, and the second was a projection step, solved by a discrete PPE; the method was approximated in space with a Q_1P_0 finite element interpolation. This scheme is related to the *velocity–correction* method, developed and extensively used by M. Kawahara and coworkers (see, for instance, [68]), employing a continuous PPE for the projection step and equal order Q_1Q_1 finite element space interpolation.

The *velocity correction* method was also studied in [113], among other schemes (such as a simple predictor–corrector method, Taylor–Galerkin and Runge–Kutta type schemes) for the linear, unsteady Stokes flow of a slightly compressible fluid. A fractional step method for both compressible and incompressible flow in a characteristic–Galerkin formulation was developed in [112]; in both references, the *inf–sup* restrictions on the discrete approximating spaces are shown to be bypassed by the appearance, in the steady state solution, of a nonzero diagonal term for the pressure in the system matrix.

A three–step explicit scheme was developed in [73], and more recently a two–step, fully implicit, pressure correction method was presented in [51].

An explicit and an implicit Taylor–Galerkin based three–step algorithms were introduced in [53]. Other studies on splitting methods can be found in [33], [99], [80] and [52], among several others.

1.5.5 Further comments and conclusions

We have now seen the great variety of fractional step methods for the unsteady, incompressible Navier–Stokes equations existing nowadays. The main advantages of many of these schemes, specially of projection methods using a PPE, are the uncoupling of the computation of the pressure from that of the velocity, thus reducing the order of the discrete systems of equations to be solved, and the possibility of using discrete approximations not restricted by the LBB compatibility conditions. The reasons for this latter fact appear not to be fully understood up to now; the ultimate objective of the next Chapter is to provide a full explanation of why these conditions do not apply to this kind of methods.

Moreover, the problem of which boundary conditions for velocity and pressure should be used in fractional step methods and the numerical boundary layer and tangential slip velocities introduced by some of them have also been met. Chapter 4 is mainly devoted to the study of a fractional step viscosity splitting method allowing the imposition of correct velocity boundary conditions in all substeps, while needing no boundary conditions at all for the pressure.

Chapter 2

A reformulated Stokes problem

As has already been said, understanding the properties of approximations of the Stokes problem 1.13–1.5 is crucial when trying to study the full unsteady Navier–Stokes equations, since it serves as a linear, steady model embracing the difficulties involved in the treatment of the incompressibility condition. Appart from some cases of steady creeping flow with large viscosity values, this problem is used as a physical model in incompressible elasticity problems.

The aim of this Chapter is to provide a *stabilized pressure* reformulated finite element method to solve the steady Stokes problem 1.13–1.5 numerically, which works with 'most' element pair velocity–pressure combinations. These are only restricted by a compatibility condition which is weaker than the standard *inf–sup* condition. In particular, the satisfaction of this weak condition is proved for most equal–order interpolations. Under this restriction, stability and optimal convergence both in H^1 and L^2–norms and both for the velocity and pressure variables are proved. The main idea behind the method consists of introducing a new variable which at the continuum level is the gradient of the pressure; a multiple of the residual of the equation defining this variable is then added to the continuity equation, yielding a consistent scheme.

But moreover, this method was ultimately studied to inherit the properties of classical fractional–step projection methods with a continuous PPE with respect to the stabilization of the pressure, so as to explain in particular why the compatibility conditions on the approximating spaces do not apply to these methods.

In Section 2.1 we study the stabilizing properties of projection methods for the unsteady, incompressible Navier–Stokes equations with respect to the pressure solution; in Section 2.2 we introduce the *reformulated* Stokes problem, which we analyse in Section 2.3. In 2.4 we study the weak compatibility condition required for the stability and convergence of this method, with the use of a *macroelement* technique. The next Section deals with the computational aspects of the method and the study of different iterative techniques of the *block Gauss–Seidel* type for the solution of the algebraic system of equations. Finally, we present some numerical results in Section 2.6.

2.1 Stabilizing properties of projection methods

We make here some considerations concerning the discretization of classical fractional step projection methods such as those of [22] and [101]. The basic idea for this analysis stems from previous work of [112] and [28], and for the incompressible case it can also be found in [26].

Since the analysis is linear in essence, we consider, for simplicity, the unsteady Stokes equations with homogeneous boundary conditions:

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} - \nu\Delta\mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in}\,\Omega\times(0,T) \\ \nabla\cdot\mathbf{u} &= 0 \quad \text{in}\,\Omega\times(0,T) \\ \mathbf{u} &= 0 \quad \text{on}\,\Gamma\times(0,T) \end{aligned} \tag{2.1}$$

The classical projection method for this problem reads:

$$\frac{\mathbf{u}^{n+1/2}-\mathbf{u}^n}{\delta t} - \nu\Delta\mathbf{u}^{n+1/2} = \mathbf{f}^n, \quad \mathbf{u}^{n+1/2}_{|\Gamma} = 0 \tag{2.2}$$

$$\frac{\mathbf{u}^{n+1}-\mathbf{u}^{n+1/2}}{\delta t} + \nabla p^{n+1} = 0, \quad \mathbf{n}\cdot\mathbf{u}^{n+1}_{|\Gamma} = 0 \tag{2.3}$$

$$\nabla\cdot\mathbf{u}^{n+1} = 0 \tag{2.4}$$

The projection step is usually solved via a PPE, which is deduced at the continuous level by taking the divergence of equation 2.3 and using 2.4:

$$\Delta p^{n+1} = \frac{1}{\delta t}\nabla\cdot\mathbf{u}^{n+1/2}, \qquad \mathbf{n}\cdot\nabla p^{n+1}_{|\Gamma} = 0 \tag{2.5}$$

A finite element discretization of each of these equations, not taking into account boundary conditions (see Remark 2.1), yields:

$$B\,U^{n+1/2} = \tilde{F}^n \tag{2.6}$$

$$M\,\frac{U^{n+1}-U^{n+1/2}}{\delta t} + G\,P^{n+1} = 0 \tag{2.7}$$

$$G^t\,U^{n+1} = 0 \tag{2.8}$$

$$L\,P^{n+1} = \frac{1}{\delta t}\,G^t U^{n+1/2} \tag{2.9}$$

from 2.2, 2.3, 2.4, and 2.5, respectively, where $B = M + \delta t\, K$, M is the *mass* matrix, K is the viscous stiffness matrix, $\tilde{F}^n = \delta t F^n + M U^n$, F^n comes from the force term $\mathbf{f}^n$, L is the scalar Laplacian matrix and the rest of the notation was introduced in Chapter 1.

At this point, two alternatives are possible in order to deduce a fractional step projection method, once a split–step time discretization as that of 2.2–2.3–2.4 has taken place, which lead to entirely different schemes. On the one hand, if a space discretization is introduced into the semidiscrete problem 2.2–2.3–2.4, it turns out that the linear equations to be finally solved are 2.6, 2.7 and 2.8. These two last equations have the form of the discretization of a mixed problem, and restrictions in the choice of discrete velocity and pressure spaces still apply. By isolating U^{n+1} from 2.7 and substituting it into 2.8, one finds:

$$\delta t\,(G^t M^{-1} G)\,P^{n+1} \;=\; G^t\,U^{n+1/2} \tag{2.10}$$

which, followed by 2.7, is the usual way to solve the projection step in this type of methods (see [30] and [46]). However, the consistent mass matrix which appears in 2.10 is too expensive to be inverted, and mass lumping is usually employed here, the effects of which are extensively discussed in [46].

On the other hand, if the segregation of the pressure from the velocity is done at a continuous level, by using equation 2.5, and then a space discretization is introduced, the resulting system of linear equations to be solved is, in this order, 2.6, 2.9 and 2.7. In this case, by eliminating $U^{n+1/2}$ from 2.7 and substituting it into 2.6 and 2.9, one gets:

$$\begin{aligned} BU^{n+1} \;+\; \delta t\,GP^{n+1} &= \tilde{F} + O(\delta t)^2 P^{n+1} && (2.11)\\ -\,G^t U^{n+1} \;+\; \delta t\,(L - G^t M^{-1} G)\,P^{n+1} &= 0 && (2.12) \end{aligned}$$

It is thus seen that although at the continuous level it is equivalent to use 2.3–2.4 or 2.3–2.5, at the discrete level it is quite different to use 2.7–2.8 than 2.7–2.9. In the latter case, the matrix $A = (L - G^t M^{-1} G)$ is introduced as a nonzero diagonal term which stabilizes the pressure, in a way that will be explained in what follows. This matrix, which appeared first in [87] and [64], can be understood as a difference between two discrete Laplacian operators.

The matrix A was recently proved to be positive semidefinite in [28], thus partially explaining its stabilizing properties, in a way which we outline next. The technique employed for this purpose sets the basic grounds of the theory to be developed in this Chapter. We are still considering no fixed boundary conditions.

Proposition 2.1: *for any combination of finite element spaces V_h and Q_h approximating the velocity and pressure variables with continuous functions, respectively, the matrix A is positive semidefinite.*

PROOF: one defines the space

$$\nabla Q_h = \{\mathbf{v}_h \in \mathbf{L}^2(\Omega) \;/\; \mathbf{v}_h = \nabla q_h,\ q_h \in Q_h\}$$

which is a finite dimensional subspace of $\mathbf{L}^2(\Omega)$. One defines, also:

$$E_h = V_h + \nabla Q_h,$$

another finite dimensional subspace of $\mathbf{L}^2(\Omega)$. Given a basis set $\{\mathbf{v}_1, \ldots, \mathbf{v}_{nc}\}$ of V_h, let us complete it with $\{\mathbf{v'}_1, \ldots, \mathbf{v'}_{nd}\} \subset V_h^\perp$, orthogonal subspace of V_h in E_h with respect to the L^2–product, to form a basis set of $E_h = V_h \oplus V_h^\perp$ (the indeces nc and nd refer to the continuous and discontinuous parts of ∇p_h, respectively). Given a vector P, let $p_h \in Q_h$ be the finite element function with nodal values given by the components of P, and let us express ∇p_h as:

$$\nabla p_h = \mathbf{z} + \mathbf{z}^\perp = \sum_{i=1}^{nc} y_i \mathbf{v}_i + \sum_{i=1}^{nd} y'_i \mathbf{v}'_i$$

We want to show that $P^tAP \geq 0$; one has:

$$P^tLP = (\nabla p_h, \nabla p_h) = (\mathbf{z}, \mathbf{z}) + (\mathbf{z}^\perp, \mathbf{z}^\perp) = |\mathbf{z}|^2 + |\mathbf{z}^\perp|^2$$

On the other hand, calling $M^{-1} = (M_{ij}^{-1})$, one has:

$$\begin{aligned}
P^tG^tM^{-1}GP &= \sum_{i,j=1}^{nc} (\nabla p_h, \mathbf{v}_i) M_{ij}^{-1} (\nabla p_h, \mathbf{v}_j) \\
&= \sum_{i,j=1}^{nc} \sum_{m,l=1}^{nc} y_m (\mathbf{v}_m, \mathbf{v}_i) M_{ij}^{-1} y_l (\mathbf{v}_l, \mathbf{v}_j) \\
&= \sum_{i,j,m,l=1}^{nc} y_m y_l (M_{mi} M_{ij}^{-1}) M_{lj} \\
&= \sum_{m,l=1}^{nc} y_m y_l (\sum_{j=1}^{nc} (\sum_{i=1}^{nc} M_{mi} M_{ij}^{-1}) M_{lj}) \\
&= \sum_{m,l=1}^{nc} y_m y_l M_{ml} = |\mathbf{z}|^2
\end{aligned}$$

One gets, thus, $P^tAP = |\mathbf{z}^\perp|^2 \geq 0$. □

The components of ∇p_h belonging to $V_h^\perp$, which we call *essentially discontinuous* pressure gradients, are stabilized by the matrix A; we will see in the next Sections how the other components are stabilized. We observe that for a given $p_h \in Q_h$ with associated nodal vector P, one has: $P^tAP = 0 \iff \nabla p_h \in V_h$, i.e., when ∇p_h is continuous. Defining the space $Q_h^c = \{q_h \in Q_h \;/\; \nabla q_h \in V_h\}$, we can determine the null space of A by studying Q_h^c, since $\dim Q_h^c = \dim(\mathrm{Ker} A)$. We present some results concerning the determination of this dimension for some common equal order interpolations.

We begin by the simplest one–dimensional case. For a mesh of linear elements, let $q_h \in Q_h$ have a continuous derivative. Since ∇q_h is constant on each element, it must be constant on Ω, and q_h globally linear. We find, this way, that $\dim Q_h^c = 2$. If we now consider a mesh of N quadratic elements, we have $2N+1$ degrees of freedom in Q_h and $N-1$ continuity conditions at element boundary nodes, yielding $\dim Q_h^c = N+2$. A mesh of N elements with polynomials of arbitrary degree k has $\dim Q_h^c = (k-1)N+2$.

In the two–dimensional case the situation is different. We consider each case separately:

- P_1 element: for a triangular mesh with linear polynomials, ∇q_h will be constant on each element; if it is continuous, it must be constant on Ω, and q_h globally linear. This gives $\dim Q_h^c = 3$.

- Q_1 element: for a mesh of quadrilaterals with bilinear polynomials, we have that $\dfrac{\partial q_h}{\partial x}$ is constant on each element with respect to x; if it is continuous, it will be constant on x on all the domain, $\dfrac{\partial q_h}{\partial x} = \eta_1(y)$ globally. By the same argument, $\dfrac{\partial q_h}{\partial y} = \eta_2(x)$ on Ω. Since q_h is a C^2–function on element interiors, the Schwarz theorem implies that $\eta_1'(y) = \eta_2'(x) = C$, that is, $\dfrac{\partial q_h}{\partial x} = Cx + D$, $\dfrac{\partial q_h}{\partial y} = Cy + E$. This leads to $q_h = Cxy + Dx + Ey + F$, so that q_h is globally a Q_1 polynomial. Thus, $\dim Q_h^c = 4$.

In higher d–dimensional cases, it is easy to see that $\dim Q_h^c = d+1$ for the P_1 element and $\dim Q_h^c = 2^d$ for the Q_1 element (functions of Q_h^c are globally linear and multilinear, respectively).

This study of the kernel of the matrix A will be useful in the theory to be developed in the next Sections.

REMARK 2.1: up to now we have deliverately omitted the imposition of boundary conditions in the discrete systems of equations. If we take them into account in the projection method 2.6–2.7–2.9, the matrix A should be modified to $\tilde{A} = L - G_0^t(M_\tau^{-1}G^\tau)^0$, where the subscript 0 indicates that the columns corresponding to all boundary components have been omitted, the superscript 0 likewise for rows, and the subscript and superscript τ refer to normal components on the boundary omitted but free tangential components. A similar analysis to the one performed for the matrix A, and in particular an appropriate decomposition of the space E_h, would explain the stabilizing properties of fractional step projection methods.

2.2 Development of the method

2.2.1 The continuous problem

We pretend to develop a finite element method for the steady Stokes equations with the same stabilizing properties as the fractional step methods just considered. In particular, we want the matrix $A = L - G^t M^{-1} G$ to be present in this method. We will restrict our attention to a class of Stokes problems with some additional regularity of the solution: we require the pressure gradient to be in $L^2(\Omega)$. Following [43] (page 126), we first define a regular Stokes problem as:

Definition 2.1: *let $\Omega \subset \mathbb{R}^d$ be an open, bounded, connected set; then the homogeneous Stokes problem:*

$$\begin{aligned} -\nu\Delta\mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in } \Omega \\ \nabla\cdot\mathbf{u} &= 0 \quad \text{in } \Omega \\ \mathbf{u} &= 0 \quad \text{on } \Gamma \end{aligned} \tag{2.13}$$

is called regular if $\mathbf{u} \in \mathbf{H}^2(\Omega) \cap Y$ and $p \in H^1(\Omega)$ whenever $\mathbf{f} \in \mathbf{L}^2(\Omega)$, and there exists a constant $C_r > 0$ such that:

$$||\mathbf{u}||_2 + ||p||_1 \leq C_r |\mathbf{f}|$$

According to P. Grisvard (see [49]), the Stokes problem is regular when Ω is of class C^2 in any dimension of space. Moreover, when $d = 2$ it is sufficient that Ω be a bounded, convex polygon.

This definition, however, is rather restrictive; for our purposes, it is sufficient that $p \in H^1(\Omega)$. We will call this case p-regular:

Definition 2.2: *let $\Omega \subset \mathbb{R}^d$ be an open, bounded, connected set; then the homogeneous Stokes problem* 2.13 *is called p-regular if $p \in H^1(\Omega)$ whenever $\mathbf{f} \in \mathbf{L}^2(\Omega)$.*

In any of these situations, we consider the spaces: $V_0 = \mathbf{H}^1_0(\Omega)$, $V = \mathbf{L}^2(\Omega)$ and $Q = H^1(\Omega)/\mathbb{R}$; this quotient space is isomorphic to the subspace $\{q \in H^1(\Omega) \ / \ \int_\Omega q \, d\Omega = 0\}$. We then define:

Definition 2.3: *given $\mathbf{f} \in \mathbf{L}^2(\Omega)$, the reformulated Stokes problem consists of finding $(\mathbf{u}, p, \mathbf{w}) \in V_0 \times Q \times V$ such that:*

$$\begin{array}{rcll} -\nu\Delta\mathbf{u} + \nabla p & = & \mathbf{f} & \text{in } \Omega \\ \nabla p - \mathbf{w} & = & 0 & \text{in } \Omega \\ \nabla\cdot\mathbf{u} + \alpha(-\Delta p + \nabla\cdot\mathbf{w}) & = & 0 & \text{in } \Omega \\ \mathbf{u} & = & 0 & \text{on } \Gamma \\ \mathbf{n}\cdot\nabla p - \mathbf{n}\cdot\mathbf{w} & = & 0 & \text{on } \Gamma \end{array} \tag{2.14}$$

where $\alpha > 0$ is a constant.

This problem is, at the continuous level, equivalent to the p–regular Stokes problem; its weak form is:

$$\begin{array}{rcll} \nu(\nabla\mathbf{u},\nabla\mathbf{v}) + (\nabla p,\mathbf{v}) & = & (\mathbf{f},\mathbf{v}), & \forall\mathbf{v}\in V_0 \quad (2.15) \\ (\nabla\cdot\mathbf{u},q) + \alpha(\nabla p,\nabla q) - \alpha(\mathbf{w},\nabla q) & = & 0, & \forall q\in Q \quad (2.16) \\ (\nabla p,\mathbf{y}) - (\mathbf{w},\mathbf{y}) & = & 0, & \forall\mathbf{y}\in V \quad (2.17) \end{array}$$

where the consistent boundary condition $\mathbf{n}\cdot\nabla p - \mathbf{n}\cdot\mathbf{w} = 0$ has been enforced weakly (see Subsection 2.6.3). For a p–regular Stokes problem, 2.14 has a unique solution:

Proposition 2.2: *if the Stokes problem* 2.13 *is p–regular and* $\mathbf{f}\in\mathbf{L}^2(\Omega)$, *then the reformulated Stokes problem* 2.14 *has a unique solution* $(\mathbf{u},p,\mathbf{w})$, *where* $(\mathbf{u},p)$ *is the solution of* 2.13 *and* $\mathbf{w}=\nabla p$ *in* Ω. *Moreover:*

$$||\mathbf{u}|| \le \frac{C_\Omega|\mathbf{f}|}{\nu} \tag{2.18}$$

where C_Ω was introduced in 1.14.

PROOF: existence is obtained from the properties of the solution $(\mathbf{u},p)$ of the p–regular Stokes problem; as for uniqueness, let us define the bilinear form D on $(V_0\times Q\times V)^2$ by:

$$\begin{array}{rcl} D(\mathbf{u},p,\mathbf{w};\mathbf{v},q,\mathbf{y}) & = & \nu(\nabla\mathbf{u},\nabla\mathbf{v}) + (\nabla p,\mathbf{v}) + (\nabla\cdot\mathbf{u},q) + \alpha(\nabla p,\nabla q) \\ & - & \alpha(\mathbf{w},\nabla q) - \alpha(\nabla p,\mathbf{y}) + \alpha(\mathbf{w},\mathbf{y}) \end{array} \tag{2.19}$$

and the linear form:

$$\mathcal{L}(\mathbf{v},q,\mathbf{y}) = (\mathbf{f},\mathbf{v})$$

Problem 2.15–2.16–2.17 can then be written as:

$$D(\mathbf{u},p,\mathbf{w};\mathbf{v},q,\mathbf{y}) = \mathcal{L}(\mathbf{v},q,\mathbf{y}), \qquad \forall(\mathbf{v},q,\mathbf{y})\in(V_0\times Q\times V) \tag{2.20}$$

The bilinear form D has the following coercivity property: for any $(\mathbf{u}, p, \mathbf{w})$, one has:

$$D(\mathbf{u}, p, \mathbf{w}; \mathbf{u}, p, \mathbf{w}) = \nu||\mathbf{u}||^2 \;+\; \alpha|\nabla p - \mathbf{w}|^2$$

Therefore, if $(\mathbf{u}, p, \mathbf{w})$ is a solution of 2.20:

$$\nu||\mathbf{u}||^2 \;\leq\; (\mathbf{f}, \mathbf{u}) \;\leq\; |\mathbf{f}|\,|\mathbf{u}| \;\leq\; C_\Omega\,|\mathbf{f}|\,||\mathbf{u}||,$$

so that the stability condition 2.18 holds. If now $(\mathbf{u}_1, p_1, \mathbf{w}_1)$ and $(\mathbf{u}_2, p_2, \mathbf{w}_2)$ are two solutions, the difference $(\mathbf{u}_1 - \mathbf{u}_2, p_1 - p_2, \mathbf{w}_1 - \mathbf{u}_2)$ satisfies the homogeneous problem, that is, 2.20 with $\mathbf{f} = 0$. Therefore, by 2.18, $\mathbf{u}_1 = \mathbf{u}_2$, and, by 2.15, $(\nabla(p_1 - p_2), \mathbf{v}) = 0$; $\forall \mathbf{v} \in \mathbf{H}_0^1(\Omega)$. The continuous LBB condition 1.25 ensures that $\nabla(p_1 - p_2) = 0$, so that p_1 and p_2 differ by a constant. Finally, equation 2.17 implies that $\mathbf{w}_1 = \mathbf{w}_2$. □

We have obtained, therefore, an equivalent formulation of the Stokes problem. Although at the continuous level not much is gained, we will see that this formulation allows 'almost' any combination of approximating spaces for the velocity and pressure, including equal order ones, at the expense of introducing a new variable which at the continuous level is the pressure gradient.

2.2.2 The discrete problem

We now consider a finite element discretization of the reformulated problem 2.15–2.16–2.17. With the notation of Section 1.3, the approximating spaces for each variable are:

$$\begin{aligned}
V_{h,0} &= \{\mathbf{v}_h \in C^0(\Omega) \cap \mathbf{H}_0^1(\Omega) \;/\; \forall K \in \Theta_h,\ (\mathbf{v}_h)_{|K} = \hat{\mathbf{v}}_K \circ F_K^{-1},\ \hat{\mathbf{v}}_K \in \mathbf{R}_{k_v}\} \\
Q_h &= \{q_h \in C^0(\Omega) \;/\; \forall K \in \Theta_h,\ (q_h)_{|K} = \hat{q}_K \circ F_K^{-1},\ \hat{q}_K \in R_{k_p}\} \\
V_h &= \{\mathbf{y}_h \in C^0(\Omega) \;/\; \forall K \in \Theta_h,\ (\mathbf{y}_h)_{|K} = \hat{\mathbf{y}}_K \circ F_K^{-1},\ \hat{\mathbf{y}}_K \in \mathbf{R}_{k_g}\}
\end{aligned}$$

Here, the indeces k_v, k_p and k_g refer to (possibly different) orders of approximation to the velocity, pressure and pressure gradient, respectively. The discrete equivalent to 2.15–2.16–2.17 is, then:

$$\begin{aligned}
\nu(\nabla \mathbf{u}_h, \nabla \mathbf{v}_h) + (\nabla p_h, \mathbf{v}_h) &= (\mathbf{f}, \mathbf{v}_h), \forall \mathbf{v}_h \in V_{h,0} & (2.21) \\
(\nabla \cdot \mathbf{u}_h, q_h) + \alpha(\nabla p_h, \nabla q_h) - \alpha(\mathbf{w}_h, \nabla q_h) &= 0,\ \forall q_h \in Q_h & (2.22) \\
(\nabla p_h, \mathbf{y}_h) - (\mathbf{w}_h, \mathbf{y}_h) &= 0,\ \forall \mathbf{y}_h \in V_h & (2.23)
\end{aligned}$$

Existence and uniqueness of a solution of 2.21–2.22–2.23 is established next, under a mild restriction on the approximating spaces. Let us first introduce the matrix form of this problem:

$$\begin{array}{rcl} KU + G_0P & = & F \qquad (2.24) \\ -G_0^tU + \alpha LP - \alpha G^tW & = & 0 \qquad (2.25) \\ GP - MW & = & 0 \qquad (2.26) \end{array}$$

where U, P and W represent the nodal vectors of $\mathbf{u}_h$, p_h and $\mathbf{w}_h$, respectively. By eliminating W from 2.26 and substituting it into 2.25 we get:

$$-G_0^tU + \alpha(L - G^tM^{-1}G)P = 0 \qquad (2.27)$$

The similarity with 2.12 is now clear; once again, the matrix $A = (L - G^tM^{-1}G)$ is introduced in the discrete continuity equation, and stabilizes the pressure.

Equation 2.23 essentially says that the discrete pressure gradient $\mathbf{w}_h$ is the L^2–projection of the gradient of the discrete pressure, ∇p_h, onto the space V_h. Recalling the space $E_h = V_h + \nabla Q_h$ introduced in Section 2.1, and the decomposition $E_h = V_h \oplus V_h^\perp$ of this space in the form $\nabla q_h = \mathbf{z} + \mathbf{z}^\perp$ for any $q_h \in Q_h$, equation 2.23 is also equivalent to $\mathbf{w}_h = \mathbf{z}$. We know by Subsection 2.1.1 that the component $\mathbf{z}^\perp$ of ∇q_h is stabilized by the matrix A. We now decompose $\mathbf{z}$ into a component vanishing on the boundary, that is, belonging to $V_{h,0}$, and an orthogonal component in V_h. The first one will be stabilized by equation 2.21; for the second one, we need to require a stability condition on the approximating spaces. Namely, we define:

$$\begin{array}{rcl} E_{h,1} & = & V_{h,0} \\ E_{h,2} & = & V_{h,0}^\perp \cap V_h \\ E_{h,3} & = & V_h^\perp \cap E_h \end{array}$$

so that $E_h = E_{h,1} \oplus E_{h,2} \oplus E_{h,3}$. For $i = 1, 2, 3$, we call $P_{h,i}$ the L^2–projection of E_h onto $E_{h,i}$, and for $i \neq j$, $P_{h,ij} = P_{h,i} + P_{h,j}$ and $E_{h,ij} = E_{h,i} \oplus E_{h,j}$. In this notation, $\mathbf{w}_h = P_{h,12}(\nabla p_h)$. We require the interpolating spaces to be such that the following stability condition holds: there exists $k_s' > 0$ such that for all $q_h \in Q_h$,

$$|\nabla q_h| \leq k_s' \, |P_{h,13}(\nabla q_h)| \qquad (2.28)$$

This inequality says, basically, that the second component of ∇q_h can be bounded in terms of the other two. As will be seen, condition 2.28 is weaker than the standard *inf–sup* condition 1.27, and in particular satisfied by equal order interpolations; it is a sufficient condition for existence, stability and convergence of the discrete solution of 2.15–2.16–2.17:

Proposition 2.3: *if* $(V_{h,0}, Q_h, V_h)$ *satisfy* 2.28, *then there exists a solution* $(\mathbf{u}_h, p_h, \mathbf{w}_h)$ *of* 2.21–2.22–2.23; $\mathbf{u}_h$ *and* $\mathbf{w}_h$ *are unique,* p_h *is determined up to* *an additive constant on Ω.*

PROOF: we will use both the matrix and function notation. Since the problem is finite dimensional, it suffices to consider the homogeneous case $\mathbf{f} = 0$. By multiplying 2.24 by U^t, 2.25 by P^t and adding them up, we get, using 2.26:

$$U^tKU \;+\; \alpha P^tAP \;=\; 0$$

This implies $U = 0$ and $P \in \operatorname{Ker} A$, that is, $\nabla p_h = \mathbf{w}_1 + \mathbf{w}_2 \in E_{h,12}$. By 2.24, we then have $G_0P = 0$; if we take $M = \begin{pmatrix} M_0 & 0 \\ 0 & M_0^\perp \end{pmatrix}$, associated to the decomposition $E_{h,12} = E_{h,1} \oplus E_{h,2}$, this implies, by 2.26, that $M_0W_1 = 0$, that is, $\mathbf{w}_1 = 0$. The inequality 2.28 then establishes that $\mathbf{w}_2 = 0$, so that $\mathbf{w}_h = 0$ and $\nabla p_h = 0$. □

2.3 Stability and convergence of the method

We present a numerical analysis of the reformulated method, from which we obtain optimal error estimates for the approximate solution, based on the satisfaction of condition 2.28. Most of these results can be found in [26].

2.3.1 Stability

We begin by the following stability result:

Proposition 2.4: *assume that the family of partitions Θ_h of Ω is such that the inverse inequality* 1.23 *holds, and that condition* 2.28 *also holds. Assume also that α satisfies:*

$$\alpha_- h^2 \;\le\; \alpha \tag{2.29}$$

for some $\alpha_- > 0$ independent of h. Then, the solution of 2.21–2.22–2.23 *satisfies the stability estimate:*

$$|||(\mathbf{u}_h, p_h, \mathbf{w}_h)||| \;\le\; C\,|\mathbf{f}|, \tag{2.30}$$

where we have used the mesh-dependent norm:

$$|||(\mathbf{v}_h, q_h, \mathbf{y}_h)||| \doteq \nu||\mathbf{v}_h|| \;+\; h|\nabla q_h| \;+\; h|\mathbf{y}_h|, \tag{2.31}$$

for all $(\mathbf{v}_h, q_h, \mathbf{y}_h) \in V_{h,0} \times Q_h \times V_h$.

PROOF: the solution $(\mathbf{u}_h, p_h, \mathbf{w}_h)$ satisfies:

$$D(\mathbf{u}_h, p_h, \mathbf{w}_h; \mathbf{u}_h, p_h, \mathbf{w}_h) = \nu||\mathbf{u}_h||^2 \;+\; \alpha|\nabla p_h - \mathbf{w}_h|^2 \;=\; (\mathbf{f}, \mathbf{u}_h) \tag{2.32}$$

where D is the bilinear form introduced in 2.19. The stability estimate for $\mathbf{u}_h$ is, therefore, the same as for the continuous problem:

$$||\mathbf{u}_h|| \;\leq\; \frac{C_\Omega|\mathbf{f}|}{\nu} \tag{2.33}$$

Since $\mathbf{w}_h = P_{h,12}(\nabla p_h)$, 2.32 also says that:

$$\alpha|P_{h,3}(\nabla p_h)|^2 \;=\alpha|\nabla p_h - \mathbf{w}_h|^2 \;\leq\; C_\Omega\,|\mathbf{f}|\,||\mathbf{u}_h||,$$

so that, from 2.33:

$$|P_{h,3}(\nabla p_h)| \;\leq\; \frac{C_\Omega|\mathbf{f}|}{\sqrt{\alpha\nu}} \tag{2.34}$$

On the other hand, since $P_{h,1}(\nabla p_h) \in V_{h,0}$, we have, from 1.23:

$$\begin{aligned}
|P_{h,1}(\nabla p_h)|^2 \;&=\; (\nabla p_h, P_{h,1}(\nabla p_h)) \qquad (2.35)\\
&=\; (\mathbf{f}, P_{h,1}(\nabla p_h)) \;-\; a(\mathbf{u}_h, P_{h,1}(\nabla p_h))\\
&\leq\; |\mathbf{f}|\,|P_{h,1}(\nabla p_h)| \;+\; ||a||\,||\mathbf{u}_h||\,||P_{h,1}(\nabla p_h)||\\
&\leq\; (|\mathbf{f}| \;+\; ||a||\frac{C_\Omega|\mathbf{f}|}{\nu}\frac{C}{h})\,|P_{h,1}(\nabla p_h)|,
\end{aligned}$$

Estimate 2.30 now follows from 2.28, 2.33, 2.34, 2.35 and 2.29, noticing that $|\mathbf{w}_h|^2 \;=\; |P_{h,1}(\nabla p_h)|^2 \;+\; |P_{h,2}(\nabla p_h)|^2$. □

2.3.2 Convergence in natural norms

We now provide a convergence analysis of the method in the natural norms for this problem, which are the H_0^1–norm for the velocity and the L^2–norm for the pressure gradient, as given by the stability estimate 2.30.

Theorem 2.1: *assume the same hypothesis as in Proposition* 2.4 *hold, but now suppose that* α *satisfies:*

$$\alpha_- h^2 \;\leq\; \alpha \;\leq\; \alpha_+ h^2 \tag{2.36}$$

with α_- *and* α_+ *independent of* h*. Then, the solution of* 2.21–2.22–2.23 *satisfies the error estimate:*

$$|||\,(\mathbf{u}-\mathbf{u}_h, p-p_h, \nabla p-\mathbf{w}_h)\,||| \;\leq\; C\,E(h) \tag{2.37}$$

with $C > 0$ *independent of* h *and:*

$$\begin{aligned}
E(h) \;=\;& \inf_{\mathbf{v}_h\in V_{h,0}} ||\mathbf{u}-\mathbf{v}_h|| \;+\; \frac{1}{h}\inf_{\mathbf{v}_h\in V_{h,0}} |\mathbf{u}-\mathbf{v}_h| \;+\; \inf_{q_h\in Q_h} |p-q_h|\\
&+\; h\inf_{q_h\in Q_h} |\nabla p - \nabla q_h| \;+\; h\inf_{\mathbf{y}_h\in V_h} |\nabla p - \mathbf{y}_h| \qquad (2.38)
\end{aligned}$$

PROOF: the discrete problem 2.21–2.22–2.23 can be written as:

$$D(\mathbf{u}_h, p_h, \mathbf{w}_h; \mathbf{v}_h, q_h, \mathbf{y}_h) = \mathcal{L}(\mathbf{v}_h, q_h, \mathbf{y}_h), \qquad \forall(\mathbf{v}_h, q_h, \mathbf{y}_h) \in (V_{h,0} \times Q_h \times V_h) \tag{2.39}$$

Substracting 2.39 from 2.20 and taking as test functions $(\mathbf{v}_h - \mathbf{u}_h, q_h - p_h, \mathbf{y}_h - \mathbf{w}_h) \in V_{h,0} \times Q_h \times V_h$ we obtain:

$$\begin{aligned} & D(\mathbf{u} - \mathbf{u}_h, p - p_h, \nabla p - \mathbf{w}_h; \mathbf{u} - \mathbf{u}_h, p - p_h, \nabla p - \mathbf{w}_h) \\ & = D(\mathbf{u} - \mathbf{u}_h, p - p_h, \nabla p - \mathbf{w}_h; \mathbf{u} - \mathbf{v}_h, p - q_h, \nabla p - \mathbf{y}_h), \end{aligned} \tag{2.40}$$

for all $(\mathbf{v}_h, q_h, \mathbf{y}_h) \in V_{h,0} \times Q_h \times V_h$. Using the definition of the form D, it is found from 2.40 that:

$$\begin{aligned} & a(\mathbf{u} - \mathbf{u}_h, \mathbf{u} - \mathbf{u}_h) + \alpha(\mathbf{w}_h - \nabla p_h, \mathbf{w}_h - \nabla p_h) = a(\mathbf{u} - \mathbf{u}_h, \mathbf{u} - \mathbf{v}_h) \\ & \quad +(\nabla p - \nabla p_h, \mathbf{u} - \mathbf{v}_h) + b(p - q_h, \mathbf{u} - \mathbf{u}_h) + \alpha(\mathbf{w}_h - \nabla p_h, \mathbf{y}_h - \nabla q_h). \end{aligned}$$

Using the coercivity of a, the continuity of a and b and Schwarz inequality, we get:

$$\begin{aligned} \|\mathbf{u} - \mathbf{u}_h\|_1^2 + \frac{\alpha}{\beta_a}|\mathbf{w}_h - \nabla p_h|^2 \quad \leq \quad & C\,[\|\mathbf{u} - \mathbf{u}_h\|_1 \|\mathbf{u} - \mathbf{v}_h\|_1 \\ + \quad & |\nabla p - \nabla p_h|\, \|\mathbf{u} - \mathbf{v}_h\| \\ + \quad & |p - q_h|\, \|\mathbf{u} - \mathbf{u}_h\|_1 \\ + \quad & \alpha|\mathbf{w}_h - \nabla p_h|\, |\mathbf{y}_h - \nabla q_h|] \end{aligned} \tag{2.41}$$

Let us denote by $E_m(\cdot)$ the error in the H^m norm of either $\mathbf{u}$, p or ∇p and $I_m(\mathbf{u}) \doteq \|\mathbf{u} - \mathbf{v}_h\|_m$, $I_0(p) \doteq |p - q_h|$, $I_0(\nabla p) \doteq |\nabla p - \nabla q_h|$ and $I_0(\mathrm{w}) \doteq |\nabla p - \mathbf{y}_h|$. Also, let $G \doteq |\mathbf{w}_h - \nabla p_h|$. We can thus write 2.41 as:

$$E_1^2(\mathbf{u}) + \frac{\alpha}{\beta_a}G^2 \leq C\,[E_1(\mathbf{u})I_1(\mathbf{u}) + E_0(\nabla p)I_0(\mathbf{u}) + I_0(p)E_1(\mathbf{u}) + \alpha G|\mathbf{y}_h - \nabla q_h|]\,. \tag{2.42}$$

Since:

$$|\mathbf{y}_h - \nabla q_h| \leq |\mathbf{y}_h - \nabla p| + |\nabla p - \nabla q_h| = I_0(\mathrm{w}) + I_0(\nabla p), \tag{2.43}$$

from 2.42 we obtain:

$$\begin{aligned} E_1^2(\mathbf{u}) + \frac{\alpha}{\beta_a}G^2 \quad \leq \quad & C\left[E_1(\mathbf{u}) + hE_0(\nabla p) + \alpha^{1/2}G\right] \\ \times \quad & \max\left\{I_1(\mathbf{u}), I_0(p), \frac{1}{h}I_0(\mathbf{u}), \alpha^{1/2}I_0(\mathrm{w}), \alpha^{1/2}I_0(\nabla p)\right\}. \end{aligned} \tag{2.44}$$

The problem now is to bound $E_0(\nabla p)$. We have that:

$$\begin{aligned} E_0(\nabla p) &\leq |\nabla p - P_{h,12}(\nabla q_h)| + |P_{h,12}(\nabla q_h) - \nabla p_h| \\ &\leq |\nabla p - P_{h,12}(\nabla q_h)| + |P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)| \\ &+ |P_{h,2}(\nabla q_h) - P_{h,2}(\nabla p_h)| + |P_{h,3}(\nabla p_h)|. \end{aligned} \tag{2.45}$$

Using now the stability condition 2.28, we obtain:

$$\begin{aligned} |P_{h,2}(\nabla q_h) - P_{h,2}(\nabla p_h)| &\leq C|P_{h,13}(\nabla q_h) - P_{h,13}(\nabla p_h)| \\ &\leq C|P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)| \\ &+ C|P_{h,3}(\nabla q_h)| + C|P_{h,3}(\nabla p_h)|. \end{aligned} \tag{2.46}$$

On the other hand:

$$\begin{aligned} |P_{h,3}(\nabla q_h)| &= |\nabla q_h - P_{h,12}(\nabla q_h)| \\ &\leq |\nabla q_h - \nabla p| + |\nabla p - P_{h,12}(\nabla q_h)|. \end{aligned}$$

Using this in 2.46 it is found that:

$$\begin{aligned} |P_{h,2}(\nabla q_h) - P_{h,2}(\nabla p_h)| &\leq C\,[|P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)| \\ &+ |\nabla q_h - \nabla p| + |\nabla p - P_{h,12}(\nabla q_h)| \\ &+ |P_{h,3}(\nabla p_h)|] \end{aligned}$$

Using this inequality in the estimate 2.45, we get:

$$\begin{aligned} E_0(\nabla p) &\leq (1+C)|\nabla p - P_{h,12}(\nabla q_h)| \\ &+ (1+C)|P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)| \\ &+ (1+C)|P_{h,3}(\nabla p_h)| + C|\nabla p - \nabla q_h|. \end{aligned} \tag{2.47}$$

Let us bound now the different terms in 2.47. If we still denote by $P_{h,12}$ the extension of the projection onto $E_{h,12} = V_h$ from the whole space $\mathbf{L}^2(\Omega)$, we have that:

$$|\nabla p - P_{h,12}(\nabla q_h)| \leq |\nabla p - P_{h,12}(\nabla p)| + |P_{h,12}(\nabla p) - P_{h,12}(\nabla q_h)|. \tag{2.48}$$

Since:

$$(\nabla p - P_{h,12}(\nabla p), \mathbf{y}_h) = 0 \qquad \forall \mathbf{y}_h \in V_h, \tag{2.49}$$

and $P_{h,12}(\nabla p) - \mathbf{y}_h \in V_h$ for $\mathbf{y}_h \in V_h$, we have that:

$$\begin{aligned}
|\nabla p - P_{h,12}(\nabla p)|^2 &= (\nabla p - P_{h,12}(\nabla p), \nabla p - P_{h,12}(\nabla p) + P_{h,12}(\nabla p) - \mathbf{y}_h) \\
&= (\nabla p - P_{h,12}(\nabla p), \nabla p - \mathbf{y}_h) \\
&\leq |\nabla p - P_{h,12}(\nabla p)|\, |\nabla p - \mathbf{y}_h|,
\end{aligned}$$

that is:

$$|\nabla p - P_{h,12}(\nabla p)| \leq I_0(\mathbf{w}). \tag{2.50}$$

If $||P_{h,12}||$ is the norm of $P_{h,12}$ as a linear operator from $\mathbf{L}^2(\Omega)$ to $E_{h,12}$, since this norm is less than or equal to 1, we have that:

$$\begin{aligned}
|P_{h,12}(\nabla p) - P_{h,12}(\nabla q_h)| &\leq ||P_{h,12}||\, |\nabla p - \nabla q_h| \\
&\leq I_0(\nabla p).
\end{aligned} \tag{2.51}$$

Using inequalities 2.50 and 2.51 in 2.48 we obtain:

$$|\nabla p - P_{h,12}(\nabla q_h)| \leq I_0(\mathbf{w}) + I_0(\nabla p). \tag{2.52}$$

The second term in 2.47 can be bounded using the first equation of the problem, that is, 2.21, and making use of the inverse estimate 1.23:

$$\begin{aligned}
|P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)|^2 &= (\nabla q_h - \nabla p_h, P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)) \\
&= (\nabla p - \nabla p_h, P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)) \\
&+ (\nabla q_h - \nabla p, P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)) \\
&= -a(\mathbf{u} - \mathbf{u}_h, P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)) \\
&+ (\nabla q_h - \nabla p, P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)) \\
&\leq \left(C\frac{||a||}{h}E_1(\mathbf{u}) + I_0(\nabla p)\right) \\
&\times |P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)|,
\end{aligned}$$

Therefore:

$$|P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)| \leq C\frac{||a||}{h}E_1(\mathbf{u}) + I_0(\nabla p). \tag{2.53}$$

The third term in 2.47 is $(1 + C)G$ and the last one is $CI_0(\nabla p)$. Thus, using bounds 2.52 and 2.53 in 2.47 we obtain:

$$E_0(\nabla p) \leq C\left[I_0(\mathbf{w}) + I_0(\nabla p) + \frac{||a||}{h}E_1(\mathbf{u}) + G\right], \tag{2.54}$$

and using this in 2.44, we get:

$$
\begin{aligned}
E_1^2(\mathbf{u}) + \frac{\alpha}{\beta_a} G^2 &\leq C \left[E_1(\mathbf{u}) + h I_0(\nabla p) + h I_0(\mathbf{w}) + (h + \alpha^{1/2}) G \right] \quad (2.55) \\
&\times \max \left\{ I_1(\mathbf{u}), I_0(p), \frac{1}{h} I_0(\mathbf{u}), \alpha^{1/2} I_0(\mathbf{w}), \alpha^{1/2} I_0(\nabla p) \right\}.
\end{aligned}
$$

From the behaviour assumed for the parameter α, 2.55 implies that there exist constants C_1 and C_2 such that:

$$
\begin{aligned}
E_1(\mathbf{u}) &\leq C_1 \max \left\{ I_1(\mathbf{u}), \frac{1}{h} I_0(\mathbf{u}), h I_0(\nabla p), I_0(p), h I_0(\mathbf{w}) \right\}, \quad (2.56) \\
G &\leq \frac{C_2}{h} \max \left\{ I_1(\mathbf{u}), \frac{1}{h} I_0(\mathbf{u}), h I_0(\nabla p), I_0(p), h I_0(\mathbf{w}) \right\}. \quad (2.57)
\end{aligned}
$$

Equation 2.56 is the error estimate for the velocity. Using 2.56 and 2.57 in 2.54, we obtain the error estimate for the pressure:

$$
h E_0(\nabla p) \leq C_3 \max \left\{ I_1(\mathbf{u}), \frac{1}{h} I_0(\mathbf{u}), h I_0(\nabla p), I_0(p), h I_0(\mathbf{w}) \right\}. \quad (2.58)
$$

On the other hand:

$$
|\mathbf{w}_h - \nabla p| = |\nabla p_h - \nabla p - P_{h,3}(\nabla p_h)| \leq E_0(\nabla p) + G. \quad (2.59)
$$

The theorem follows combining inequalities 2.56 to 2.59. □

We have obtained, therefore, 'optimal' error estimates for the velocity and pressure solutions. From the approximating properties 1.22 and the definitions of $V_{h,0}$, Q_h and V_h, it follows that:

Corolary 2.1: *if the solution* $(\mathbf{u}, p)$ *of* 2.13 *satisfies* $\mathbf{u} \in \mathbf{H}^r(\Omega) \cap Y$ *and* $p \in H^s(\Omega)$ *with* $r \geq 2$, $s \geq 1$, *then:*

$$
|||\,(\mathbf{u} - \mathbf{u}_h, p - p_h, \nabla p - \mathbf{w}_h)\,||| \leq C h^l \quad (2.60)
$$

with $l = \min\{r - 1, s, k_u, k_p + 1, k_g + 2\}$

It is thus possible to use discrete approximations in which the pressure is interpolated with polynomials of one degree less than the velocity, and the pressure gradient with two degrees less. Nevertheless, we will concentrate on the case of equal order interpolation.

We have seen that for stability it was necessary that $\alpha_- h^2 \leq \alpha$, whereas for convergence we needed $\alpha_- h^2 \leq \alpha \leq \alpha_+ h^2$. The behaviour of the coefficient α is mandated by this numerical analysis, and, as in the GLS method, we take it of the form:

$$\alpha = \alpha_0 \frac{h^2}{4\nu} \tag{2.61}$$

An extension of this theory to the case where α is defined elementwise by $\alpha_K = \alpha_0 \frac{h_K^2}{4\nu}$, $\forall K \in \Theta_h$, and the integrals it multiplies on 2.22 evaluated on each K, could be performed, which would allow the use of less uniform meshes, thus opening the door to selective mesh refinement. This local method would not admit a continuous interpretation, and in it local inverse inequalities like $||v_h||_{1,K} \leq \frac{C}{h_K}|v_h|_K$ should be used. We have considered this possibility in some of the numerical examples.

2.3.3 Convergence in L^2–norms

We use now the classical Aubin–Nitsche argument to obtain improved error estimates for the velocity and pressure in the space $L^2(\Omega)$, in a similar way to [17] for the GLS method. The shift used in these duality arguments requires of more regularity of the problem than was needed up to now.

Theorem 2.2: *Assume that* 1.23, 2.28 *and* 2.36 *hold, and that the Stokes problem* 2.13 *is regular. Then, the solution* $(\mathbf{u}_h, p_h, \mathbf{w}_h)$ *of* 2.21–2.22–2.23 *satisfies:*

$$|\mathbf{u} - \mathbf{u}_h| + h\,|p - p_h|_{L_0^2(\Omega)} \leq C\,h\,E(h) \tag{2.62}$$

where $E(h)$ *was defined in* 2.38

PROOF: we begin by the estimate for the velocity. Let $\mathbf{y} \in \mathbf{H}^2(\Omega) \cap Y$ and $\chi \in Q$ be the solution of the regular Stokes problem:

$$\begin{array}{rcl} -\Delta \mathbf{y} + \nabla \chi & = & \mathbf{u} - \mathbf{u}_h \quad \text{in } \Omega \\ \nabla \cdot \mathbf{y} & = & 0 \quad \text{in } \Omega \\ \mathbf{y} & = & 0 \quad \text{on } \Gamma \end{array} \tag{2.63}$$

so that:

$$\begin{array}{rcl} ||\mathbf{y}||_2 & \leq & C_r\,|\mathbf{u} - \mathbf{u}_h| \\ ||\chi||_1 & \leq & C_r\,|\mathbf{u} - \mathbf{u}_h| \end{array} \tag{2.64}$$

Let $\mathbf{y}_h \in V_{h,0}$ and $\chi_h \in Q_h$ be optimal order approximations to $\mathbf{y}$ and χ, respectively, satisfying:

$$\begin{array}{rcl} ||\mathbf{y} - \mathbf{y}_h||_m & \leq & C\,h^{2-m}\,||\mathbf{y}||_2 \\ ||\chi - \chi_h||_m & \leq & C\,h^{1-m}||\chi||_1 \end{array} \tag{2.65}$$

for $m = 0, 1$. We then have:

$$\begin{aligned} |\mathbf{u}-\mathbf{u}_h|^2 &= (\mathbf{u}-\mathbf{u}_h, \mathbf{u}-\mathbf{u}_h) = (\nabla\mathbf{y}, \nabla(\mathbf{u}-\mathbf{u}_h)) - (\chi, \nabla\cdot(\mathbf{u}-\mathbf{u}_h)) \\ &= \Big((\nabla(\mathbf{y}-\mathbf{y}_h), \nabla(\mathbf{u}-\mathbf{u}_h)) - (\chi-\chi_h, \nabla\cdot(\mathbf{u}-\mathbf{u}_h))\Big) \\ &+ (\nabla\mathbf{y}_h, \nabla(\mathbf{u}-\mathbf{u}_h)) - (\chi_h, \nabla\cdot(\mathbf{u}-\mathbf{u}_h)) = T_1 + T_2 + T_3 \end{aligned}$$

We bound each term separately:

$$\begin{aligned} T_1 &= (\nabla(\mathbf{y}-\mathbf{y}_h), \nabla(\mathbf{u}-\mathbf{u}_h)) - (\chi-\chi_h, \nabla\cdot(\mathbf{u}-\mathbf{u}_h)) \\ &\leq ||(\mathbf{y}-\mathbf{y}_h)||\,||(\mathbf{u}-\mathbf{u}_h)|| + C\,|\chi-\chi_h|\,||(\mathbf{u}-\mathbf{u}_h)|| \\ &\leq C\,||(\mathbf{u}-\mathbf{u}_h)||\,(h\,||\mathbf{y}||_2 + h\,||\chi||_1) \\ &\leq C\,h\,||(\mathbf{u}-\mathbf{u}_h)||\,|(\mathbf{u}-\mathbf{u}_h)| \end{aligned}$$

by 2.65, 2.63 and the continuity of the operator $\nabla\cdot$ on $\mathbf{H}_0^1(\Omega)$. Moreover:

$$\begin{aligned} T_2 &= (\nabla\mathbf{y}_h, \nabla(\mathbf{u}-\mathbf{u}_h)) \\ &= -\frac{1}{\nu}(\nabla(p-p_h), \mathbf{y}_h) = \frac{1}{\nu}(\nabla(p-p_h), \mathbf{y}-\mathbf{y}_h) \\ &\leq \frac{1}{\nu}\,|\nabla(p-p_h)|\,|\mathbf{y}-\mathbf{y}_h| \leq C\,h^2\,||\mathbf{y}||_2\,|\nabla(p-p_h)| \\ &\leq C\,h\,|(\mathbf{u}-\mathbf{u}_h)|\,(h\,|\nabla(p-p_h)|) \end{aligned}$$

by 2.21, 2.63 and 2.65. Finally:

$$\begin{aligned} T_3 &= -(\chi_h, \nabla\cdot(\mathbf{u}-\mathbf{u}_h)) \\ &= \alpha(\nabla(p-p_h), \nabla\chi_h) - \alpha(\nabla p - \mathbf{w}_h, \nabla\chi_h) \\ &= \alpha\Big[(\nabla(p-p_h), \nabla(\chi_h-\chi)) + (\nabla(p-p_h), \nabla\chi) \\ &- (\nabla p - \mathbf{w}_h, \nabla(\chi_h-\chi)) - (\nabla p - \mathbf{w}_h, \nabla\chi)\Big] \\ &\leq \alpha\Big[(|\nabla(p-p_h)| + |\nabla p - \mathbf{w}_h|)\,(|\nabla(\chi_h-\chi)| + |\nabla\chi|)\Big] \\ &\leq \alpha\,C\,|\nabla\chi|\,(|\nabla(p-p_h)| + |\nabla p - \mathbf{w}_h|) \\ &\leq C\,h^2\,|(\mathbf{u}-\mathbf{u}_h)|\,(|\nabla(p-p_h)| + |\nabla p - \mathbf{w}_h|) \end{aligned}$$

by 2.23, 2.63, 2.65 and 2.36. We obtain the error estimate for the velocity combining the above inequalities for T_1, T_2 and T_3. As for the pressure, we call $\mathbf{z} \in \mathbf{H}_0^1(\Omega)$ and $\xi \in L_0^2(\Omega)$ the solution of the Stokes problem:

$$\begin{aligned} -\Delta\mathbf{z} + \nabla\xi &= 0 \quad \text{in } \Omega \\ \nabla\cdot\mathbf{z} &= (p-p_h) \quad \text{in } \Omega \\ \mathbf{z} &= 0 \quad \text{on } \Gamma \end{aligned} \tag{2.66}$$

Standard results for this problem yield (see Remark 2.5, page 32 on [105]):

$$\begin{aligned} ||\mathbf{z}|| &\leq C\,|p-p_h| \\ |\xi| &\leq C\,|p-p_h| \end{aligned} \qquad (2.67)$$

If $\mathbf{z}_h \in V_{h,0}$ now satisfies:

$$||\mathbf{z}-\mathbf{z}_h||_m \leq C\,h^{1-m}\,||\mathbf{z}||_1 \qquad (2.68)$$

for $m = 0, 1$, we have:

$$\begin{aligned} |p-p_h|^2 &= (p-p_h, p-p_h) = (\nabla\cdot\mathbf{z}, p-p_h) \\ &= (\nabla\cdot(\mathbf{z}-\mathbf{z}_h), p-p_h) - (\mathbf{z}_h, \nabla(p-p_h)) \\ &= -(\mathbf{z}-\mathbf{z}_h, \nabla(p-p_h)) + \nu(\nabla(\mathbf{u}-\mathbf{u}_h), \nabla\mathbf{z}_h) \\ &= -(\mathbf{z}-\mathbf{z}_h, \nabla(p-p_h)) + \nu(\nabla(\mathbf{u}-\mathbf{u}_h), \nabla(\mathbf{z}_h-\mathbf{z})) \\ &+ \nu(\nabla(\mathbf{u}-\mathbf{u}_h), \nabla\mathbf{z}) \\ &\leq |\mathbf{z}-\mathbf{z}_h|\,|\nabla(p-p_h)| + C\,||\mathbf{u}-\mathbf{u}_h||\,(||\mathbf{z}-\mathbf{z}_h|| + ||\mathbf{z}||) \\ &\leq C\,h\,||\mathbf{z}||\,|\nabla(p-p_h)| + C\,||\mathbf{z}||\,||\mathbf{u}-\mathbf{u}_h|| \\ &\leq C\,(h\,|\nabla(p-p_h)| + ||\mathbf{u}-\mathbf{u}_h||)\,|p-p_h| \end{aligned}$$

and the estimate 2.62 is finally established. □

2.4 A weakened *inf–sup* condition

The stability and convergence results just proved rely on the satisfaction of condition 2.28. The very existence of a discrete solution is affected by this condition. We will see that this can be expressed in the form of an *inf–sup* condition, and then we present an analysis of sufficient conditions for 2.28 to hold, based on a macroelement technique. This will let us show that ours is weaker than the standard LBB condition 1.27, and prove that equal order, simplicial finite element interpolations of arbitrary order in two and three dimensions and first order quadrilateral interpolations satisfy condition 2.28, thus providing stable and convergent results of 2.15–2.16–2.17.

We first have the following previous result:

Lemma 2.1: *condition* 2.28 *is equivalent to the existence of a constant* $k_s > 0$ *such that:*

$$\inf_{q_h\in Q_h}\Big(\sup_{\mathbf{v}_h\in E_{h,13}} \frac{(\nabla q_h, \mathbf{v}_h)}{|\mathbf{v}_h|\,.\,|\nabla q_h|}\Big) \geq k_s > 0, \qquad (2.69)$$

PROOF: assume that 2.28 is satisfied; then, for all $q_h \in Q_h$:

$$\begin{aligned}(\nabla q_h, P_{h,13}(\nabla q_h)) &= (P_{h,13}(\nabla q_h), P_{h,13}(\nabla q_h)) = |P_{h,13}(\nabla q_h)|^2 \\ &\geq \frac{1}{k'_s}\, |P_{h,13}(\nabla q_h)|\, |\nabla q_h|,\end{aligned}$$

and 2.69 holds with $k_s = 1/k'_s$. On the other hand, if 2.69 is assumed, for all $q_h \in Q_h$:

$$k_s \leq \sup_{\mathbf{v}_h \in E_{h,13}} \frac{(\nabla q_h, \mathbf{v}_h)}{|\nabla q_h|\, |\mathbf{v}_h|} = \sup_{\mathbf{v}_h \in E_{h,13}} \frac{(P_{h,13}(\nabla q_h), \mathbf{v}_h)}{|\nabla q_h|\, |\mathbf{v}_h)|} \leq \frac{|P_{h,13}(\nabla q_h)|}{|\nabla q_h|},$$

so that 2.28 holds with $k'_s = 1/k_s$. □

We will use this equivalence between our stability condition 2.28 and 2.69 in what follows to obtain sufficient conditions for it to hold, which are simpler to check in practice than 2.28. Moreover, we will show that this condition is weaker than the standard LBB condition 1.27.

2.4.1 Macroelement technique

The ideas used here to obtain simple conditions for 2.69 to hold are an extension to our case of the theory of macroelement techniques, developed mainly by R. Stemberg (see [97]). We take part of our notation from this reference. A macroelement M is the union of one or more elements in Θ_h. For each $h > 0$, let $\mathcal{M}_h$ be a collection of macroelements covering Ω. One of these macroelements $M \in \mathcal{M}_h$ is said to be equivalent to another macroelement $M_0 \in \mathcal{M}_{h_0}$ if there exists an homeomorphism $\mathbf{G}_M : M_0 \longrightarrow M$ such that:

(i) $\mathbf{G}_M(M_0) = M$,

(ii) If $M_0 = \bigcup_{j=1}^m K_{0,j}$, then $M = \bigcup_{j=1}^m \mathbf{G}_M(K_{0,j})$, where $K_{0,j} \in \Theta_{h_0}, j = 1, ..., m$.

(iii) $\mathbf{G}_{M|K_0} = \mathbf{F}_K \circ \mathbf{F}_{K_0}^{-1}$, where $K = \mathbf{G}_M(K_0)$ and $\mathbf{F}_K$ and $\mathbf{F}_{K_0}$ are the mappings from the reference element $\hat{K}$ to $K \in \Theta_h$ and to $K_0 \in \Theta_{h_0}$, respectively, introduced earlier.

Notice that equivalent macroelements can be associated with the same or with a different finite element partition. Thus, with this definition, $\{\mathcal{M}_h\}_{h>0}$ is split into a finite number of equivalence classes $Ec_1, ..., Ec_{n_c}$.

Let us consider the spaces $V_{M,0}$, Q_M, V_M, E_M and $E_{M,i}$, $i = 1, 2, 3$, defined as their analogues $V_{h,0}$, Q_h, V_h, E_h and $E_{h,i}$, $i = 1, 2, 3$, but replacing the partition Θ_h by the partition of a macroelement $M \in \mathcal{M}_h$ (the zero mean restriction is not imposed on Q_M). Also, $P_{M,i}$ are the orthogonal projections from E_M to $E_{M,i}$, $i = 1, 2, 3$.

We first show that if a condition like 2.28 holds in a macroelement, then it also holds in Ω:

Lemma 2.2: *if there exists a constant $C > 0$ such that*

$$|\nabla q_h|_M \leq C|P_{M,13}(\nabla q_h)|_M \qquad \forall q_h \in Q_h, \tag{2.70}$$

for all $M \in \mathcal{M}_h$, then condition 2.28 *holds for a constant k_s' independent of h.*

PROOF: let $q_h \in Q_h$ and let $\mathbf{v}_{M,i}$ be the extension by zero of $P_{M,i}(\nabla q_h), i = 1, 3$, to the whole domain Ω. Consider also the vector field:

$$\mathbf{v}_h = \sum_M \mathbf{v}_M = \sum_M (\mathbf{v}_{M,1} + \mathbf{v}_{M,3}), \tag{2.71}$$

Clearly, $\mathbf{v}_{M,1} \in E_{M,1} \subset E_{h,1}$ $\forall M$ and thus $\sum_M \mathbf{v}_{M,1} \in E_{h,1}$. Let $\mathbf{v}_{h,12} \in E_{h,12}$. Since $\mathbf{v}_{h,12}|_M \in E_{M,12} = E_{M,3}^\perp$ (orthogonality in E_M) we have that:

$$\int_\Omega \mathbf{v}_{h,12} \cdot \left(\sum_M \mathbf{v}_{M,3}\right) \, d\Omega = \sum_M \int_M \mathbf{v}_{h,12}|_M \cdot \mathbf{v}_{M,3} \, d\Omega = 0, \tag{2.72}$$

that is, $\sum_M \mathbf{v}_{M,3} \in E_{h,12}^\perp = E_{h,3}$. Therefore, $\mathbf{v}_h$ in 2.71 belongs to $E_{h,13}$.

Let N_M be the maximum number of macroelements to which an element domain belongs, and N_K the maximum number of element domains per macroelement. Let us bound first $|\mathbf{v}_h|$:

$$\begin{aligned}
|\mathbf{v}_h|^2 &= \int_\Omega \left(\sum_M \mathbf{v}_M\right)^2 \, d\Omega \\
&= \int_\Omega \left(\sum_M (\mathbf{v}_M)^2 + 2 \sum_{M \neq M', M\cap M' \neq \emptyset} \mathbf{v}_M \cdot \mathbf{v}_{M'}\right) \, d\Omega \\
&\leq \sum_M |\mathbf{v}_M|^2 + 2 \sum_{M \neq M', M\cap M' \neq \emptyset} |\mathbf{v}_M||\mathbf{v}_{M'}| \\
&\leq \sum_M |\mathbf{v}_M|^2 + \sum_{M \neq M', M\cap M' \neq \emptyset} \left(|\mathbf{v}_M|^2 + |\mathbf{v}_{M'}|^2\right) \\
&\leq (1 + N_M N_K) \sum_M |\mathbf{v}_M|^2 \\
&\leq (1 + N_M N_K) \sum_M |\nabla q_h|_M^2 \\
&\leq (1 + N_M N_K) N_M |\nabla q_h|^2,
\end{aligned}$$

that is, there exists a constant $C_0 > 0$ such that:

$$|\mathbf{v}_h| \leq C_0 |\nabla q_h|. \tag{2.73}$$

On the other hand, from 2.70 it follows that:

$$\begin{aligned} \int_\Omega \nabla q_h \cdot \mathbf{v}_h \, d\Omega &= \sum_M \int_M \nabla q_h \cdot \mathbf{v}_M \, d\Omega \\ &= \sum_M |P_{M,13}(\nabla q_h)|^2 \\ &\geq \frac{1}{C^2} \sum_M |\nabla q_h|_M^2 \\ &\geq \frac{1}{C^2} |\nabla q_h|^2. \end{aligned} \tag{2.74}$$

But, using inequality 2.73:

$$\int_\Omega \nabla q_h \cdot \mathbf{v}_h \, d\Omega = \int_\Omega P_{h,13}(\nabla q_h) \cdot \mathbf{v}_h \, d\Omega \leq C_0 |P_{h,13}(\nabla q_h)| \, |\nabla q_h|. \tag{2.75}$$

The lemma follows combining inequalities 2.74 and 2.75 with $k_s' = C_0 C^2$. □

The next step is to give sufficient conditions for property 2.70 to hold. First we give a rather technical lemma:

Lemma 2.3: *Let M be a metric space with distance* dist, *X and Y two subsets of M and $\{Y_\mu\}_{\mu>0}$ a family of subsets of M such that:*

$$\lim_{\mu \to 0} \left[\sup_{y_\mu \in Y_\mu} \inf_{y \in Y} \operatorname{dist}(y_\mu, y) \right] = \lim_{\mu \to 0} \left[\sup_{y \in Y} \inf_{y_\mu \in Y_\mu} \operatorname{dist}(y_\mu, y) \right] = 0. \tag{2.76}$$

Let Z be another subset of M such that $Y \subset Z$ and $Y_\mu \subset Z$ for all $\mu > 0$. Consider a family of functions $\{f_\mu\}_{\mu>0}$ from $M \times M$ to $\mathbb{R}$ that converge uniformly in $X \times Z$ to a function f uniformly continuous in its second argument. Then:

$$\lim_{\mu \to 0} \left[\inf_{x \in X} \sup_{y_\mu \in Y_\mu} |f_\mu(x, y_\mu)| \right] = \inf_{x \in X} \sup_{y \in Y} |f(x, y)|. \tag{2.77}$$

PROOF: Let $\epsilon > 0$ be given. Since $\{f_\mu\}$ converges uniformly to f as $\mu \to 0$ in $X \times Z$:

$$\exists \mu_1 = \mu_1(\epsilon) \; : \; \forall \mu < \mu_1 \quad |f_\mu(x,y) - f(x,y)| < \frac{\epsilon}{4}, \quad \forall (x,y) \in X \times Z$$

Thus, if $\mu < \mu_1$:

$$\begin{aligned}\inf_{x\in X}\sup_{y\in Y_\mu}|f_\mu(x,y)| &\leq \inf_{x\in X}\left[\sup_{y\in Y_\mu}|f_\mu(x,y)-f(x,y)|+\sup_{y\in Y_\mu}|f(x,y)|\right]\\ &< \frac{\epsilon}{3}+\inf_{x\in X}\sup_{y\in Y_\mu\cup Y}|f(x,y)| \qquad (2.78)\end{aligned}$$

Given $x\in X$, let:

$$S_\mu(x) \doteq \sup_{y\in Y_\mu\cup Y}|f(x,y)| \geq S_0(x) \doteq \sup_{y\in Y}|f(x,y)|$$

Since f is uniformly continuous in its second argument in $X\times Z$ and $Y,\ Y_\mu\subset Z$:

$$\exists\delta=\delta(\epsilon)\ :\ \operatorname{dist}(y,y')<\delta\Rightarrow|f(x,y)-f(x,y')|<\frac{\epsilon}{3},\quad\forall x\in X,\ y,y'\in Z \qquad (2.79)$$

Condition 2.76 implies that:

$$\exists\mu_2=\mu_2(\delta(\epsilon))\ :\ \forall\mu<\mu_2,\ \forall y_\mu\in Y_\mu\cup Y,\ \exists y\in Y\ /\ \operatorname{dist}(y_\mu,y)<\delta \qquad (2.80)$$

On the other hand, we have that:

$$\begin{aligned}\exists y_\mu=y_\mu(\epsilon)\in Y_\mu\cup Y\quad/\quad |f(x,y_\mu)|\geq S_\mu(x)-\frac{\epsilon}{3} \qquad (2.81)\\ \forall y\in Y,\quad -|f(x,y)| \geq -S_0(x) \qquad (2.82)\end{aligned}$$

and therefore:

$$|f(x,y_\mu)-f(x,y)| \geq S_\mu(x)-S_0(x)-\frac{\epsilon}{3},\quad\forall y\in Y$$

If $\mu<\min\{\mu_1,\mu_2\}$ and we take y such that condition 2.80 holds for the y_μ that verifies condition 2.81, from condition 2.79 we have that:

$$\frac{\epsilon}{3} > S_\mu(x)-S_0(x)-\frac{\epsilon}{3}$$

that is:

$$\sup_{y\in Y_\mu\cup Y}|f(x,y)| < \frac{2}{3}\epsilon+\sup_{y\in Y}|f(x,y)|$$

Using this in inequality 2.78 it follows that:

$$\inf_{x\in X}\sup_{y\in Y_\mu}|f_\mu(x,y)|<\inf_{x\in X}\sup_{y\in Y}|f(x,y)|+\epsilon \qquad (2.83)$$

One can similarly show that if μ is small enough:

$$\inf_{x\in X}\sup_{y\in Y}|f(x,y)| < \inf_{x\in X}\sup_{y\in Y_\mu}|f_\mu(x,y)| + \epsilon \tag{2.84}$$

The lemma follows from inequalities 2.83 and 2.84. □

This result is used now to prove the following:

Lemma 2.4: *Let Ec_i be one of the equivalence classes introduced above, $i \in \{1, 2, ..., n_c\}$, and suppose that the following condition holds:*

$$\begin{aligned} &\exists M_0 \in Ec_i \text{ such that } \quad \forall q \in Q_{M_0}, \\ &\int_{M_0} \nabla q \cdot \mathbf{v} dM \;=\; 0, \quad \forall \mathbf{v} \in E_{M_0,13} \quad \Rightarrow \; \nabla q = 0. \end{aligned} \tag{2.85}$$

Then, there exists a constant $C_i > 0$ such that, for all $M \in Ec_i$:

$$|\nabla q|_M \le C_i |P_{M,13}(\nabla q)|_M \quad \forall q \in Q_M. \tag{2.86}$$

PROOF: Let us consider the following function defined on the class Ec_i:

$$\beta(M) \doteq \inf_{q\in Q_M}\sup_{\mathbf{v}\in E_{M,13}} \frac{(\nabla q, \mathbf{v})_M}{|\nabla q|_M |\mathbf{v}|_M}. \tag{2.87}$$

Inequality 2.86 is equivalent to saying that $\beta(M) \ge 1/C_i$ for all $M \in Ec_i$ (this can be proved as Lemma 2.1).

From assumption 2.85 it is easy to see that $\beta(M) > 0$ for all $M \in Ec_i$. Since M is defined by the coordinates of its nodes, β can be considered as a function of these coordinates. Due to the quasi-uniformity of the family $\{\Theta\}_{h>0}$ (or simply due to its non-degeneracy), all the nodes are isolated points of $\mathbb{R}^d$, and therefore they form a compact set. Thus, β can be considered as a function defined on a compact set. To prove that it is bounded below by a positive constant it is enough to prove that it is continuous.

Let M, $M' \in Ec_i$. We want to show that $\beta(M') \to \beta(M)$ as $M' \to M$. Let $\mathbf{G} : M \to M'$ be the homeomorphism that relates M and M'. We denote its Jacobian matrix (piecewise continuous) by $\mathbf{DG}$. Let also:

$$J' \doteq \max_{\mathbf{x}'\in M'} |\mathbf{DG}^{-1}|(\mathbf{x}'), \qquad j' \doteq \min_{\mathbf{x}'\in M'} |\mathbf{DG}^{-1}|(\mathbf{x}'), \tag{2.88}$$

where $|\cdot|$ stands now for the determinant of a matrix. Here and below, we use the symbol $'$ to refer to quantities associated with M'. The two functions in 2.88 depend on the macroelement M' and tend to 1 as $M' \to M$, that is, as $\mathbf{G} \to \mathbf{I}$.

Let us write the function β as:

$$\beta(M) = \inf_{q\in Q_{M,0}} \sup_{\mathbf{v}\in\mathbf{S}} f(\nabla q, \mathbf{v}), \qquad f(\nabla q, \mathbf{v}) \doteq \frac{(\nabla q, \mathbf{v})_M}{|\nabla q|_M |\mathbf{v}|_M}.$$

where $Q_{M,0} = \{q \in Q_M \mid \nabla q \neq \mathbf{0}\}$ and $\mathbf{S}$ is the the unit sphere of $E_{M,13}$.

Let $\mathbf{v}' \in E_{M',13}$, $q' \in Q_{M',0}$ and $\mathbf{v}$, q the pull-backs of $\mathbf{v}'$ and q' (that is, $\mathbf{v} = \mathbf{G}^*\mathbf{v}' = \mathbf{v}' \circ \mathbf{G}$, $q = \mathbf{G}^*q' = q' \circ \mathbf{G}$). It can be readily checked that:

$$\begin{aligned}
\int_{M'} \nabla' q' \cdot \mathbf{v}' dM' &= \int_M \nabla q \cdot \mathbf{DG}^{-1} \cdot \mathbf{v} |\mathbf{DG}| dM, \\
\int_{M'} \mathbf{v}' \cdot \mathbf{v}' dM' &= \int_M \mathbf{v} \cdot \mathbf{v} |\mathbf{DG}| dM, \\
\int_{M'} \nabla' q' \cdot \nabla' q' dM' &= \int_M \left(\nabla q \cdot \mathbf{DG}^{-1}\right) \cdot \left(\nabla q \cdot \mathbf{DG}^{-1}\right) |\mathbf{DG}| dM.
\end{aligned}$$

If we introduce the abbreviation $\nabla_G q \doteq \nabla q \cdot \mathbf{DG}^{-1}$ and denote by $(\cdot,\cdot)_{G,M}$ the L^2 scalar product in M with weight $|\mathbf{DG}|$, we have that:

$$f'(\nabla' q', \mathbf{v}') = \frac{(\nabla' q', \mathbf{v}')_{M'}}{|\nabla' q'|_{M'} |\mathbf{v}'|_{M'}} = \frac{(\nabla_G q, \mathbf{v})_{G,M}}{|\nabla_G q|_{G,M} |\mathbf{v}|_{G,M}} \doteq f_G(\nabla q, \mathbf{v}), \tag{2.89}$$

where $|\cdot|_{G,M}$ is the norm associated with $(\cdot,\cdot)_{G,M}$.

Since $\mathbf{DG}$ is nonsingular, if $\nabla' q' \neq 0$ then $\nabla q \neq 0$, that is, if $q' \in Q_{M',0}$ then $\mathbf{G}^* q' \in Q_{M,0}$. If $\mathbf{v}' \in \mathbf{S}'$, let us see where does $\mathbf{v} = \mathbf{G}^*\mathbf{v}'$ belong. Let $\mathbf{v}' = \mathbf{v}'_1 + \mathbf{v}'_3$, with $\mathbf{v}'_1 \in E_{M',1}$ and $\mathbf{v}'_3 \in E_{M',3}$. Since $\mathbf{v}'_1$ is continuous and vanishes on $\partial M'$ and $\mathbf{G}$ is continuous, $\mathbf{G}^*\mathbf{v}'_1 \in E_{M,1}$. In general, $\mathbf{G}^*\mathbf{v}'_{12} \in E_{M,12}$ for all $\mathbf{v}'_{12} \in E_{M',12}$. However, $\mathbf{G}^*\mathbf{v}'_3 \notin E_{M,3}$ if $\mathbf{v}'_3 \in E_{M',3}$. This is due to the fact that:

$$\forall \mathbf{v}_{12} \in E_{M,12}, \quad \int_M \mathbf{v}_{12} \cdot \mathbf{G}^*\mathbf{v}'_3 dM = \int_{M'} \left(\mathbf{v}_{12} \circ \mathbf{G}^{-1}\right) \cdot \mathbf{v}'_3 |\mathbf{DG}| dM', \tag{2.90}$$

which is in general not zero since $\mathbf{v}_{12} \circ \mathbf{G}^{-1} |\mathbf{DG}| \notin E_{M',12}$ if $|\mathbf{DG}|$ is not continuous. Therefore, if $\mathbf{S}_G = \mathbf{G}^*\mathbf{S}'$ then $\mathbf{S}_G \neq \mathbf{S}$.

Using the previous results, the function β evaluated at M' can be written as:

$$\beta(M') = \inf_{q\in Q_{M,0}} \sup_{\mathbf{v}\in\mathbf{S}_G} f_G(\nabla q, \mathbf{v}). \tag{2.91}$$

Now we use Lemma 2.3 to prove the continuity of β. Let:

$$Z = \left\{ \mathbf{v} \in E_M \;/\; \frac{1}{2} \leq |\mathbf{v}|_M \leq 2 \right\}. \tag{2.92}$$

We have that:

$$|\mathbf{G}^*\mathbf{v}|_M^2 = \int_{M'} \mathbf{v}' \cdot \mathbf{v}' |\mathbf{DG}^{-1}| dM', \tag{2.93}$$

and thus $\sqrt{j'} \leq |\mathbf{G}^*\mathbf{v}|_M \leq \sqrt{J'}$, with j' and J' defined in 2.88. If we take M' sufficiently close to M, $j' > 1/4$ and $J' < 4$, so that $\mathbf{S} \subset Z$ and $\mathbf{S}_G \subset Z$.

It is now easy to prove that $f(\nabla q, \mathbf{v})$ is uniformly continuous in the second argument in $Q_{M,0} \times Z$ and that $f_G(\nabla q, \mathbf{v})$ converges uniformly to $f(\nabla q, \mathbf{v})$ in $Q_{M,0} \times Z$. To apply Lemma 2.3 it remains to check condition 2.76 with $Y = \mathbf{S}$ and $Y_\mu = \mathbf{S}_G$, the parameter μ being now replaced by the function $\mathbf{G}$ and $\mu \to 0$ by $\mathbf{G} \to \mathbf{I}$.

Let $\bar{\mathbf{v}}_G \in \mathbf{S}_G \subset E_M$ and $\mathbf{v}' = \mathbf{v}_1' + \mathbf{v}_3' \in \mathbf{S}'$ such that $\bar{\mathbf{v}}_G = \mathbf{G}^*\mathbf{v}'$, with $\mathbf{v}_1' \in E_{M',1}$ and $\mathbf{v}_3' \in E_{M',3}$. Then $\bar{\mathbf{v}}_G = \mathbf{G}^*\mathbf{v}_1' + \mathbf{G}^*\mathbf{v}_3'$, with $\mathbf{G}^*\mathbf{v}_1' \in E_{M,1}$ but $\mathbf{G}^*\mathbf{v}_3' \notin E_{M,3}$ (in general). Let:

$$\mathbf{w} = \mathbf{G}^*\mathbf{v}_1' + \frac{\mathbf{G}^*\mathbf{v}_3'}{|\mathbf{DG}^{-1}|\circ\mathbf{G}}, \quad \bar{\mathbf{v}} = \frac{\mathbf{w}}{|\mathbf{w}|_M}. \tag{2.94}$$

It is easily verified that the second component in $\mathbf{w}$ belongs to $E_{M,3}$, and therefore $\bar{\mathbf{v}} \in \mathbf{S}$. A simple calculation shows that $\mathrm{dist}(\bar{\mathbf{v}}_G, \bar{\mathbf{v}}) \to 0$ as $\mathbf{G} \to \mathbf{I}$, that is, as j', $J' \to 1$. Hence:

$$\sup_{\mathbf{v}_G \in \mathbf{S}_G} \inf_{\mathbf{v} \in \mathbf{S}} \mathrm{dist}(\mathbf{v}_G, \mathbf{v}) \to 0 \quad \text{as} \quad \mathbf{G} \to \mathbf{I}. \tag{2.95}$$

Also, given $\bar{\mathbf{v}} = \mathbf{v}_1 + \mathbf{v}_3 \in \mathbf{S}$, with $\mathbf{v}_1 \in E_{M,1}$ and $\mathbf{v}_3 \in E_{M,3}$, let:

$$\mathbf{w}' = \mathbf{v}_1'\circ\mathbf{G}^{-1} + \frac{\mathbf{v}_3'\circ\mathbf{G}^{-1}}{|\mathbf{DG}|\circ\mathbf{G}^{-1}}, \quad \bar{\mathbf{v}}_G = \frac{\mathbf{G}^*\mathbf{w}'}{|\mathbf{w}'|_{M'}}. \tag{2.96}$$

It turns out that $\bar{\mathbf{v}}_G \in \mathbf{S}_G$ and that $\mathrm{dist}(\bar{\mathbf{v}}_G, \bar{\mathbf{v}}) \to 0$ as $\mathbf{G} \to \mathbf{I}$, thus proving that:

$$\sup_{\mathbf{v} \in \mathbf{S}} \inf_{\mathbf{v}_G \in \mathbf{S}_G} \mathrm{dist}(\mathbf{v}_G, \mathbf{v}) \to 0 \quad \text{as} \quad \mathbf{G} \to \mathbf{I}. \tag{2.97}$$

From 2.95 and 2.97 it may be concluded that hypothesis 2.76 holds in the present situation and ultimately that the function β defined in 2.87 is continuous, which is what had to be proved. □

Combining Lemmas 2.2 and 2.4 we obtain the following result:

Theorem 2.3: *Suppose that for all the equivalence classes Ec_i, $i = 1, ..., n_c$ of macroelements of $\{\Theta_h\}_{h>0}$ condition* 2.85 *holds. Then, there exists a constant $k_s > 0$, independent of h, for which the inf-sup condition* 2.69 *is verified.*

PROOF: let $C = \min\{C_1, ..., C_{n_c}\}$, where C_i is the constant for the equivalence class Ec_i established by Lemma 2.4. Since for all $h > 0$ functions $q_h \in Q_h$ restricted to a macroelement $M \in \mathcal{M}_h$ belong to Q_M, we are in the hypothesis of Lemma 2.2. The theorem follows from Lemma 2.1. □

We remark that condition 2.85 is the key to prove that a finite element interpolation is stable for our method; it is similar to the condition obtained in [97] for the standard LBB condition, but weaker than it: the space where v runs here ($E_{M_0,13}$) is larger than in the standard case ($E_{M_0,1}$).

2.4.2 Equal order interpolations

We prove now that condition 2.28 is fulfilled by simplicial equal order finite element interpolations with polynomials of arbitrary degree k on the simplex, both in two and three dimensions. They are only restricted by a weak condition on the meshes at the boundary that will be specified next. This result is achieved by applying the macroelement technique just considered, which can also be extended to the case of equal order interpolations with polynomials of Q_k on quadrilaterals and hexahedra.

Proposition 2.5: *let $k_u = k_p = k_g = k$ in the definition of $V_{h,0}$, Q_h and V_h, and $\hat{K}$ be the standard simplex in $\mathbb{R}^d$. Let Ec_i be a class of equivalent macroelements with reference macroelement $\hat{M}$ such that there is at least one interior vertex, and, for $d = 3$ and $k \geq 2$, no element $K \subset \hat{M}$ has three faces on $\partial\hat{M}$. Then, condition* 2.85 *is satisfied on $\hat{M}$.*

PROOF: we prove condition 2.85 by imposing continuity of ∇q_h on $\hat{M}$ rather than orthogonality to $E_{h,3}$, due to the difficulty of characterizing this space. Orthogonality to $E_{h,1}$ is enforced directly.

Let us consider the case of linear elements ($k = 1$) first, both for $d = 2$ and 3. For a given $q_h \in Q_h$, ∇q_h is constant on each element $K \subset \hat{M}$; if we assume ∇q_h is continuous, it must be constant on $\hat{M}$. Since we have assumed the existence of at least one node P interior to $\hat{M}$, orthogonality of ∇q_h with respect to velocity fields which take a value of one on P in each of the space dimensions and zero elsewhere, implies the vanishing of ∇q_h.

Let's now turn to the case of higher order elements ($k > 1$). Given $q_h \in Q_h$, for each $K \subset \hat{M}$ the components of $(\nabla q_h)_{|K}$ belong to $P_{k-1}(K)$. Thus, if these components are continuous, they can be determined by their nodal values on a discretization of $\hat{M}$ with the same elements K but with nodes corresponding to an interpolation with polynomials of degree $k - 1$. Let $n_{\rm int}$ be the number of nodes in the interior of $\hat{M}$, denoted by $\text{Int}(\hat{M})$, and n_{k-1} the number of nodes associated to an interpolation with polynomials of P_{k-1}. Since the orthogonality conditions with respect to all continuous vector functions that take arbitrary values at the nodes of $\text{Int}(\hat{M})$ are linearly independent restrictions on ∇q_h, it is enough to prove that $n_{\rm int} \geq n_{k-1}$.

Let's consider the two-dimensional case first; for any triangle $K \subset \hat{M}$, there are $(k-1)(k-2)/2$ nodes associated to P_k on $\text{Int}(K)$ and $(k+1)$ on each edge of K (including the vertices). Thus, there are $(k-2)(k-3)/2$ nodes associated to P_{k-1} on $\text{Int}(K)$ and k on each edge of K. If an element K lies on $\text{Int}(\hat{M})$, its contribution to $n_{\rm int}$ is clearly greater than to n_{k-1}. Thus, we restrict the analysis to the boundary. Suppose first that all the elements have at most one edge on the boundary. Let n_{ele} denote the number of elements in $\hat{M}$ with one edge on the boundary, and n_{edg} the number of edges with one node on the boundary; we study the various contributions to the difference $n_{\mathrm{int}} - n_{k-1}$:

- From element interiors: $n_{\mathrm{ele}} \times [(k-1)(k-2)/2 - (k-2)(k-3)/2] = n_{\mathrm{ele}} \times (k-2)$
- From edges with one boundary vertex (including it): $n_{\mathrm{edg}} \times [k-k] = 0$
- From boundary edges, of which there are n_{ele} due to the assumption on $\hat{M}$: $n_{\mathrm{ele}} \times [0-(k-2)] = -n_{\mathrm{ele}} \times (k-2)$

This proves that $n_{\mathrm{int}} - n_{k-1} \geq 0$ in this case. If we now include triangles with two edges on the boundary, the contribution to n_{int} is $(k-1)(k-2)/2 + (k-1)$, whereas the contribution to n_{k-1} is $(k-2)(k-3)/2 + 2(k-2) + 1$. These two quantities are equal, so that we still have $n_{\mathrm{int}} - n_{k-1} \geq 0$.

Finally, in the three dimensional case each tetrahedron $K \subset \hat{M}$ has $(k-1)(k-2)(k-3)/6$ nodes of $P_k(K)$ on $\mathrm{Int}(K)$, and $(k-2)(k-3)(k-4)/6$ of $P_{k-1}(K)$. As before, we first consider the case in which the elements have at most one face on $\partial\hat{M}$. If n_{ele} is the number of elements with one face on $\partial\hat{M}$, n_{fac} the number of faces with one edge on $\partial\hat{M}$, n_{edg} the number of edges with one node on $\partial\hat{M}$ and n_{bdr} the number of edges on $\partial\hat{M}$, contributions to $n_{\mathrm{int}} - n_{k-1}$ are:

- From element interiors: $n_{\mathrm{ele}} \times [(k-1)(k-2)(k-3)/6 - (k-2)(k-3)(k-4)/6] = n_{\mathrm{ele}} \times [(k-2)(k-3)/2]$
- From the interiors of faces with one edge on the boundary: $n_{\mathrm{fac}} \times [(k-1)(k-2)/2 - (k-2)(k-3)/2] = n_{\mathrm{fac}} \times (k-2)$
- From edges with one boundary vertex (including it): $n_{\mathrm{edg}} \times [k-k] = 0$
- From the interior of boundary faces, of which there are n_{ele} due to the assumption on $\hat{M}$: $n_{\mathrm{ele}} \times [0-(k-2)(k-3)/2] = -n_{\mathrm{ele}} \times (k-2)(k-3)/2$
- From boundary edges: $n_{\mathrm{bdr}} \times [0-(k-2)] = -n_{\mathrm{bdr}} \times (k-2)$

Since $n_{\mathrm{fac}} \geq n_{\mathrm{bdr}}$ in general, we find again that $n_{\mathrm{int}} - n_{k-1} \geq 0$. If we now consider elements with two faces on $\partial\hat{M}$, for each of them n_{int} increases by $(k-1)(k-2)(k-3)/6 + (k-1)(k-2) + (k-1)$, whereas the increase of n_{k-1} is only $(k-2)(k-3)(k-4)/6 + (k-2)(k-3) + (k-2)$. □

We have proved, in summary, that simplicial equal order interpolations of arbitrary order satisfy condition 2.28, thus yielding optimally convergent results. We now prove that this holds for equal order bilinear quadrilateral interpolations, under a mild nondegeneracy restriction on the mesh, to be specified next. The proof does not requiere of the macroelement technique, since it is given in the whole of the domain:

Proposition 2.6: *assume that the discretization Θ_h of the domain Ω is such that there are at least two nodes in the interior of Ω, if $d = 2$, or three nodes if $d = 3$. Consider a finite element interpolation such that $k_u = k_p = k_g = 1$ in the definition of $V_{h,0}$, Q_h and V_h, and $\hat{K}$ is the unit square ($d = 2$) or cube ($d = 3$). Then, condition* 2.28 *holds.*

PROOF: let us start by the two dimensional case first; we will show that condition 2.85 holds on all the domain Ω. We know by the study of the kernel of the matrix $A = L - G^t M^{-1} G$ of Section 2.1 that gradients of discrete pressures p_h which are orthogonal to $E_{h,3}$ are continuous, and that for the Q_1 element this can only hold if p_h is globally a Q_1 function, thus determined by 4 arbitrary constants. Orthogonality of ∇p_h, which depends on 3 arbitrary constants, with respect to velocity fields which vanish at the boundary of Ω and take arbitrary values at the two interior nodes of the mesh implies the vanishing of ∇p_h, since there are 4 of such fields that are linearly independent.

In the three dimensional case, discrete pressures with a continuous gradient are determined by $2^3 = 8$ constants, so that 7 linearly independent restrictions are enough to ensure the vanishing of ∇p_h. Since we are assuming that there are at least 3 interior nodes, orthogonality to velocity vectors defined from these nodes amounts to 9 independent restrictions, which imply that $\nabla p_h = 0$. □

We conjecture this result to be true also for equal order quadrilateral (and hexahedral) finite elements of higher order, but have not come up with a definite proof of this fact yet.

2.5 Computational aspects

We have studied several possibilities for the solution of the linear equation system 2.24–2.25–2.26, guided by some of the experience on the numerical solution of algebraic systems existing nowadays. Direct Gaussian-decomposition–based methods did not look appealing for solving 2.24–2.25–2.26, due to the large bandwidth of the system matrix for this problem, which is neither symmetric (although it can be symmetrized) nor positive definite. We propose iterative schemes which take advantage of the structure of problem 2.24–2.25–2.26; rather than standard Gauss–Seidel methods, we considered generalized *block Gauss–Seidel* schemes, in which each iteration is decomposed into a number of smaller linear problems with a symmetric, positive definite matrix, if possible.

We first present in 2.5.1 a simple scheme which we call *uncoupled block*

Gauss–Seidel method. In it, each equation from 2.24–2.25–2.26 is used to obtain updated values of one of the three variables, velocity, pressure and pressure gradient, from the others. This way, each of the smaller *uncoupled* subsystems has a symmetric, positive definite matrix, K, L and M respectively.

The slow convergence rates showed by this method led us to consider another scheme, which we call *coupled block Gauss–Seidel* method. In it, the velocity and pressure are solved together with an old pressure gradient, which is then updated using the new values just computed. In this case, the matrix for the velocity–pressure subsystem is either symmetric or positive definite, but not both at a time. The matrix for the pressure gradient is again the mass matrix. For it, the well known *lumping* technique was also considered in both the coupled and the uncoupled schemes, and comparison results with the consistent mass matrix case are provided.

Convergence results for the *coupled block Gauss–Seidel* scheme are much better than for the uncoupled one, but still not competitive. In 2.5.3, we present some techniques to accelerate this convergence, such as successive–over–relaxation methods or equation rescaling.

As mentioned earlier, a possible variant of the reformulated method 2.21–2.22–2.23 is the use of a local parameter α_K on each element K, specially suited for nonuniform meshes. We present some numerical experience concerning this possibility in 2.5.4.

In the implementation of the method we have studied 2–dimensional problems with interpolating polynomials of P_1 and P_2 on triangles and Q_1 and Q_2 on quadrilaterals. Where possible, the same mesh nodes have been used to define the elements for all four interpolations. In confined flow problems, where the velocity is prescribed on all the boundary, a pressure datum of 0 is enforced on the last node in the global numbering.

For a homogeneous external force $\mathbf{f} = \mathbf{0}$ and a nonhomogeneous boundary condition, the Stokes problem *scales* with the viscosity, in the sense that if $(\mathbf{u}_\nu, p_\nu)$ is the solution associated to a value of $\nu > 0$, one has that $\mathbf{u}_\nu = \mathbf{u}_1$ and $p_\nu = \nu p_1$. We have therefore considered unit viscosity throughout.

2.5.1 Uncoupled block–Gauss–Seidel method

We consider the following iterative scheme for the solution of 2.24–2.25–2.26, where $P^0 = 0$, $W^0 = 0$ and U^0 contains the prescribed velocity boundary conditions and is zero elsewhere:

$$\begin{aligned} KU^i &= F - G_0 P^{i-1} && (2.98)\\ \alpha L P^i &= \alpha G^t W^{i-1} + G_0^t U^i && (2.99)\\ MW^i &= GP^i && (2.100)\end{aligned}$$

Each subsystem has a symmetric, positive definite matrix; they are solved by the conjugate gradient method, to a given tolerance ϵ_{cg}. In the lumped mass matrix case, 2.100 is replaced by:

$$M^L W^i = GP^i \tag{2.101}$$

where M^L is the diagonal lumped matrix of M. In the P_2 element case, rather than by the standard *row–sum* technique, M^L is computed by a nodal quadrature rule obtained by splitting each P_2 element into 4 P_1 elements, in order to avoid null entries.

The scheme 2.98–2.99–2.100 was iterated to convergence at a given tolerance ϵ_{unc}. The following convergence criterion was chosen:

$$\begin{aligned} \text{Err}(U^i, P^i, W^i; U^{i-1}, P^{i-1}, W^{i-1}) &\doteq \\ \max\left(\frac{|U^i - U^{i-1}|_2}{|U^i|_2}, \frac{|P^i - P^{i-1}|_2}{|P^i|_2}, \frac{|W^i - W^{i-1}|_2}{|W^i|_2}\right) &\leq \epsilon_{\text{unc}} \end{aligned} \tag{2.102}$$

where $|X|_2$ is the Euclidean norm of a vector X.

Other permutations of the order in which the variables are updated in 2.24–2.25–2.26 were also considered. Two sets of three different permutations, even and odd respectively, are possible. Let us call I to 2.24, II to 2.25 and III to 2.26, equations which are used to update the velocity, pressure and pressure gradient, respectively, from the last updated values of the other variables; the initialization of the pressure as zero and the absence of boundary conditions for the pressure gradient imply that III-I-II is equivalent to I-II-III, and III-II-I equivalent to II-I-III. Four possibilities are therefore left. We performed some tests with them which showed that they all provide practically identical results, and chose I-II-III (that is, 2.98–2.99–2.100) throughout.

As a test problem, the standard cavity flow case was solved, which has become a compulsory benchmark problem for incompressible flow codes. We took the *ramp* case, in which the velocity is zero at the two upper corners and one in the horizontal direction at the rest of the upper lid. A uniform mesh of 21 × 21 nodes was used to discretize the unit square; it is shown in Figure 2.1 for the P_1 element. The pressure was set to zero at the top right corner.

We put off for the next Section the analysis of the results obtained for this problem, and concentrate here on the performance of the numerical scheme to reach a solution. We allowed a tolerance of $\epsilon_{\text{unc}} = 10^{-3}$. A study of the influence of the tolerance for the conjugate gradient method to solve each subsystem of equations on the convergence of the whole iterative scheme showed that a minimum number of iterations was needed for ϵ_{cg} in the order of 10^{-4}, which is the value that we selected.

We tried to use values of the coefficient α of the order given by equation 2.61, with a value of α_0 near unity. But we found that larger values of α_0 were needed for the iterative scheme to be stable; when a low value of this

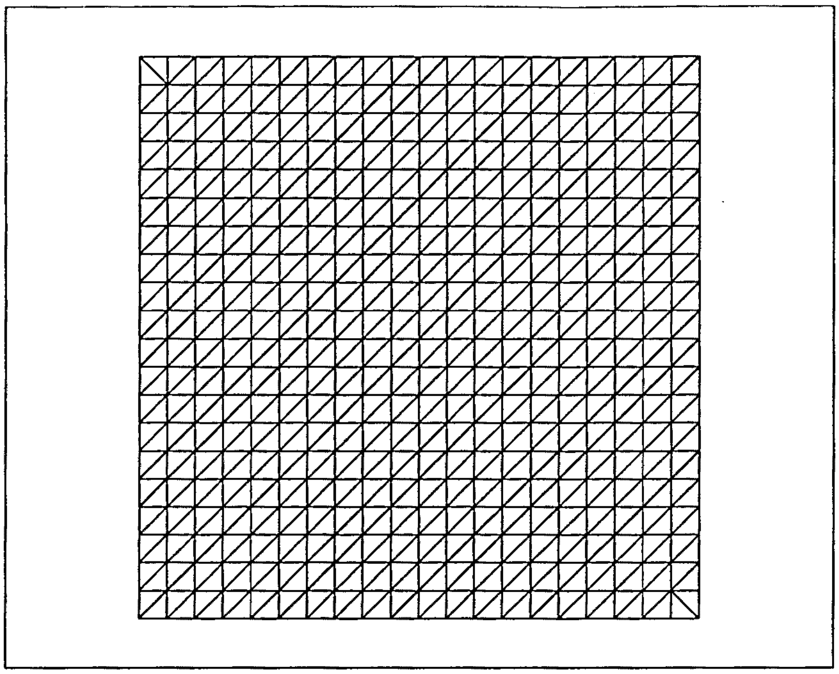


Figure 2.1: Cavity flow, uniform 21 × 21 mesh.

parameter was used, the scheme diverged. We will see in the next Section how this affects the accuracy of the solution. We finally selected $\alpha_0 = 40$ for the P_1 and Q_1 elements and $\alpha_0 = 5$ for the P_2 and Q_2.

The results obtained for this problem, in the form of number of iterations needed for convergence and CPU time spent, relative to the P_1 element case with a consistent mass matrix (in percentage of that case), are given in Table 2.1 for the four elements considered, both with consistent (C) and lumped (L) mass matrix. As can be seen, this scheme is too costly, due to the large numbers of iterations needed for convergence, specially in higher order elements. The P_2 element case with a lumped mass matrix did not converge at all.

Element	P_1-C	P_1-L	Q_1-C	Q_1-L	P_2-C	P_2-L	Q_2-C	Q_2-L
Iterations	106	66	126	59	212	-	571	195
Relative cost	100	65	61	30	122	-	316	99

Table 2.1: Convergence of the uncoupled block–Gauss–Seidel method.

2.5.2 Coupled block–Gauss–Seidel method.

We now introduce another iterative scheme for the solution of 2.24–2.25–2.26 with better convergence rates than the one just considered. The velocity and pressure are now coupled in a unique linear subsystem, in which the value of the pressure gradient at the previous iteration is used; this variable is then updated with the new pressure. With the same initializations as before, the scheme reads:

$$\begin{array}{rcl} KU^i + G_0P^i &=& F \\ -G_0^tU^i + \alpha LP^i &=& \alpha G^t W^{i-1} \\ MW^i &=& GP^i \end{array} \qquad \begin{array}{r} (2.103) \\ (2.104) \\ (2.105) \end{array}$$

Notice that the equation system 2.103–2.104 for (U^i, P^i) introduces a Laplacian term in the diagonal of the system matrix, in a similar way to stabilized methods of the GLS type. In fact, for linear elements the system matrix of this problem is the same as that of the GLS method 1.30; it is positive definite but non–symmetric. The solution (U^i, P^i) can be obtained in several different ways, such as a direct LU decomposition or the GMRES method. We tried these two possibilities and decided to use the first one, since the size of the problems that we deal with is small enough as to allow a direct method of solution.

Mass lumping was also considered for the solution of 2.105; when a consistent mass matrix was used, we tried solving 2.105 by a direct method and by the conjugate gradient algorithm. In the latter case, and for a tolerance of $\epsilon_{cg} = 10^{-5}$, it took about 10 iterations to find the solution in the first global iterations, but this reduced monotonically to 5 in the last iterations, as the initial approximation was closer to the solution. Nevertheless, we chose to use a direct method for 2.105.

Once again, due to the initialization of the pressure as zero and the absence of boundary conditions for the pressure gradient, it is inconsequencial to start the iterations by the pressure gradient equation 2.105 or by the velocity–pressure system 2.103–2.104.

The same convergence criterion 2.102 was used, and also the same test case and mesh. We took α as in 2.61, with $\alpha_0 = 1/3$ for P_1 and Q_1 and $\alpha_0 = 1/9$ for P_2 and Q_2. We present the convergence results for this case in Table 2.2. This time we show the number of iterations for convergence with a tolerance of $\epsilon_{cou} = 10^{-3}$, together with the CPU time spent in the first iteration (as a percentage of the total time in each case) and the total CPU time relative to the P_1-C case (in percentage of that case). We split the time of the first iteration to emphasize that when using direct methods to solve the linear subsystems of equations 2.103–2.104 and 2.105, it is the first iteration that requires most of the computing time, since it is then that the system matrices are assembled and factorized. The remaining iterations

Element	P_1-C	P_1-L	Q_1-C	Q_1-L	P_2-C	P_2-L	Q_2-C	Q_2-L
Iterations	16	12	35	13	41	45	118	42
1st. Iteration	76	74	66	71	67	64	50	64
Total CPU time	100	77	106	61	110	88	179	109

Table 2.2: Convergence of the coupled block–Gauss–Seidel method.

consist of the computation of the RHS vector and a backward and forward substitution only.

It can be seen, again, that linear and bilinear elements show better convergence properties than quadratic and biquadratic ones, and that mass lumping also accelerates the convergence. For the reference P_1-C case, the computing time for this coupled scheme was 34% that of the uncoupled scheme for the same case.

Finally, we present in Figures 2.2 and 2.3 the convergence history of each variable for all four interpolations, from which it can be deduced that it is the pressure and pressure gradient that dominate the convergence. It can also be observed that in the first two iterations there is a drastic reduction of the error.

2.5.3 Acceleration of convergence

The convergence results of the coupled block–Gauss–Seidel method are still not satisfactory. We employed simple techniques to accelerate this convergence, which yielded better results.

We first tried rescaling the different subsystems of linear equations in 2.24–2.25–2.26, to achieve a better conditioning of the global system matrix, hoping this way to accelerate the convergence of our block–Gauss-Seidel type schemes. Since the parameter α multiplying the Laplacian matrix in the pressure equations is of order h^2, too small for a diagonal term, we multiplied this equations by a parameter β, and replaced the pressure variable P by $R = \frac{1}{\beta}P$, so that 2.24–2.25–2.26 becomes:

$$\begin{array}{rcrcrcl} & & KU & + & \beta G_0 R & = & F \\ -\beta G_0^t U & + & \beta^2 \alpha L R & - & \beta \alpha G^t W & = & 0 \\ & & \beta G R & - & M W & = & 0 \end{array} \qquad (2.106)$$

A dimensional analysis (by comparison with the diagonal term of the first

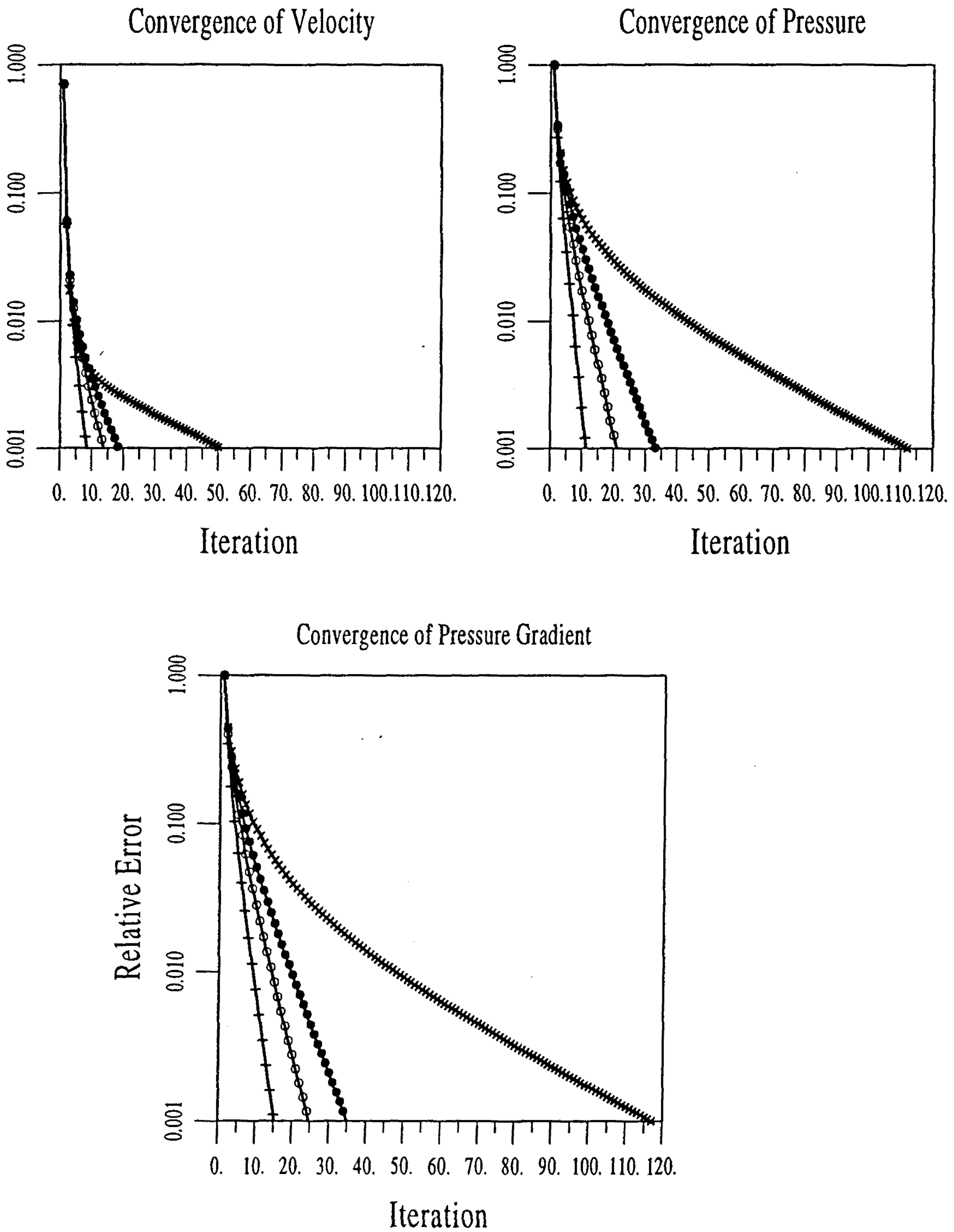


Figure 2.2: Convergence history, consistent mass matrix: + P_1 Element; • Q_1 Element; o P_2 Element; × Q_2 Element.

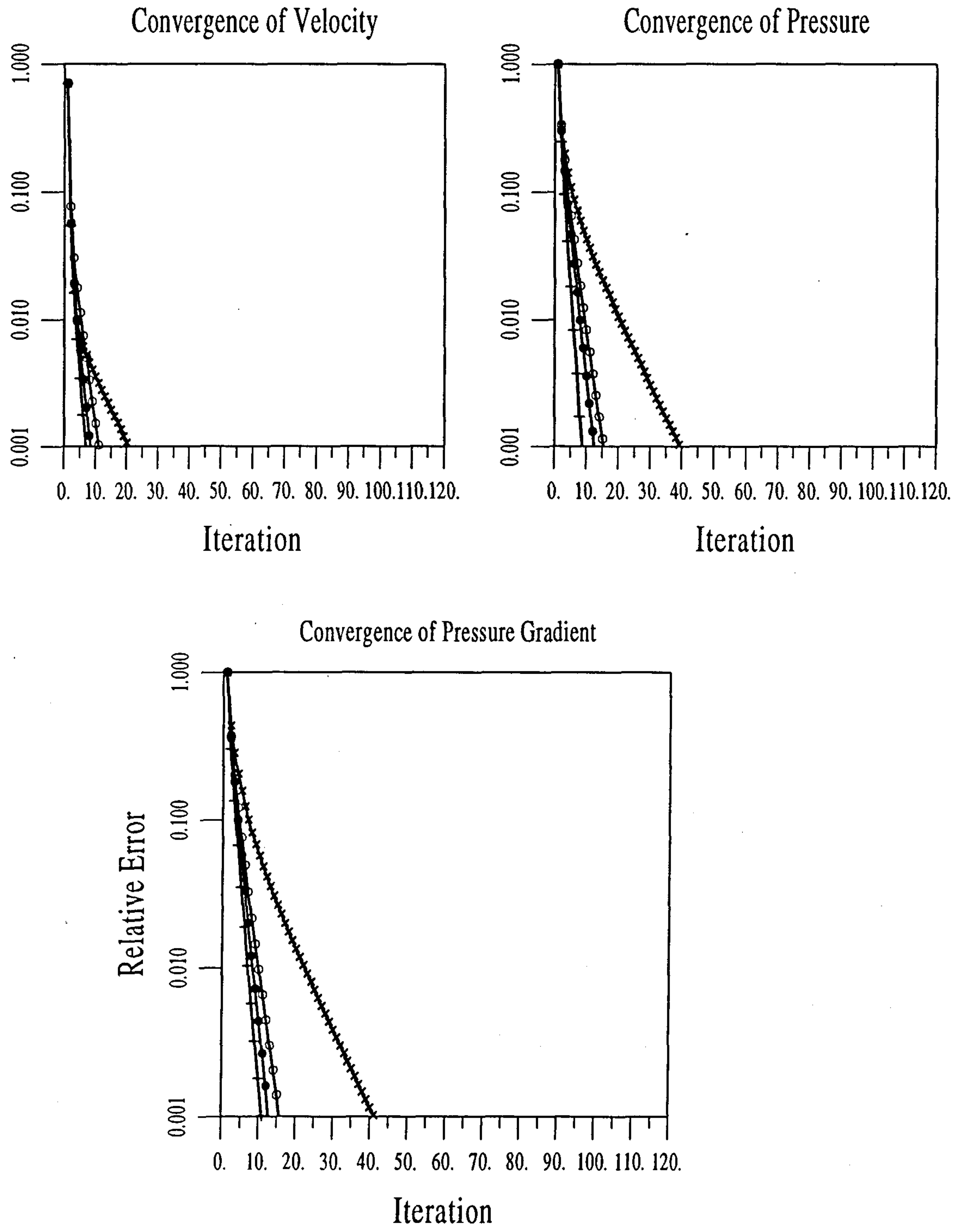


Figure 2.3: Convergence history, lumped mass matrix: + P_1 Element; • Q_1 Element; o P_2 Element; × Q_2 Element.

equation) suggested taking $\beta^2\alpha = \gamma^2\nu$, where γ is a free (dimensionless) parameter of order 1. The definition of α then gives:

$$\beta = \gamma\frac{2\nu}{\sqrt{\alpha_0 h}} \tag{2.107}$$

It can be shown that this scaling of equations is equivalent to a diagonal preconditioning strategy.

We performed several tests with the coupled–block–Gauss-Seidel scheme 2.103–2.104–2.105 applied to the rescaled equation system 2.106 and, after varying the value of γ by several orders of magnitude, we needed exactly the same number of iterations in each case as for the original problem 2.24–2.25–2.26, that is, those of Table 2.2. Nevertheless, scaling of equations improved the performance of the iterative GMRES method when it was used to solve the velocity–pressure system of equations.

We then considered standard successive–over–relaxation methods applied to 2.103–2.104–2.105, with the same relaxation parameter $\omega > 0$ for all three variables. Thus, this scheme was replaced by:

$$\begin{array}{rcrcl} K\tilde{U}^i & + & G_0\tilde{P}^i & = & F \\ -G_0^t\tilde{U}^i & + & \alpha L\tilde{P}^i & = & \alpha G^t W^{i-1} \\ & & M\tilde{W}^i & = & GP^i \end{array}$$

where the variables are updated by:

$$\begin{array}{rcrcl} U^i & = & \omega\,\tilde{U}^i & + & (1-\omega)\,U^{i-1} \\ P^i & = & \omega\,\tilde{P}^i & + & (1-\omega)\,P^{i-1} \\ W^i & = & \omega\,\tilde{W}^i & + & (1-\omega)\,W^{i-1} \end{array} \tag{2.108}$$

Different optimal relaxation parameters were found numerically for each element, all of them in the range $(1,2)$; they all lowered substantially the number of iterations needed for convergence. The results are summarized in Table 2.3, where we give the optimal relaxation values for each element together with the number of iterations needed for convergence and the total computing time, once again as a percentage of the reference P_1-C case of Table 2.2.

We then allowed the possibility of using a different relaxation parameter for each variable, ω_u, ω_p and ω_g, respectively, so that 2.108 was replaced by:

$$\begin{array}{rcrcl} U^i & = & \omega_u\,\tilde{U}^i & + & (1-\omega_u)\,U^{i-1} \\ P^i & = & \omega_p\,\tilde{P}^i & + & (1-\omega_p)\,P^{i-1} \\ W^i & = & \omega_g\,\tilde{W}^i & + & (1-\omega_g)\,W^{i-1} \end{array}$$

Element	P_1-C	P_1-L	Q_1-C	Q_1-L	P_2-C	P_2-L	Q_2-C	Q_2-L
ω	1.3	1.3	1.5	1.5	1.4	1.4	1.79	1.79
Iterations	7	7	12	12	10	10	30	30
CPU time	90	76	90	76	93	86	122	100
Iterations with $\omega_u = 1$	7	7	10	10	9	9	25	25

Table 2.3: Convergence of the relaxed coupled block–Gauss–Seidel method.

We found slightly improved convergence results only when we set $\omega_u = 1$, i.e., no relaxation for the velocity, and ω_p and ω_g equal to the optimal value for each element. We expected this to be the 'best' choice according to the convergence histories shown in Figures 2.2 and 2.3, where it is observed that it is the pressure and pressure gradient that slow the convergence. We also show in Table 2.3 the number of iterations for convergence in this case, where it can be observed that there is an improvement in some cases.

2.5.4 Local stability parameter

The uniform mesh used up to now is impractical in many situations, such as when convection is present. We solved the same test case on a non-uniform 39×39 noded mesh, with increasing density of elements near the boundary, made up with triangular linear elements (it is shown in Figure 2.4).

A question arises in this case about what value of the coefficient α is to be taken when the mesh size is not constant for all elements. We first considered the simple possibility of taking a unique value of α for all elements, as defined in 2.61 for a value of h equal to the maximum element diameter (just as it is defined). Then we adopted the idea of using a different value of α on each element, evaluating the integrals it multiplies on 2.22 elementwise. The local discrete reformulated problem reads:

$$\begin{aligned}
\nu(\nabla \mathbf{u}_h, \nabla \mathbf{v}_h) + (\nabla p_h, \mathbf{v}_h) &= (\mathbf{f}, \mathbf{v}_h), \forall \mathbf{v}_h \in W_{h,0} \\
(\nabla \cdot \mathbf{u}_h, q_h) + \sum_{K \in \Theta_h} \alpha_K((\nabla p_h, \nabla q_h)_K - (\mathbf{w}_h, \nabla q_h)_K) &= 0,\ \forall q_h \in Q_h \\
(\nabla p_h, \mathbf{x}_h) - (\mathbf{w}_h, \mathbf{x}_h) &= 0,\ \forall \mathbf{x}_h \in W_h
\end{aligned}$$

The local coefficients are, again, defined as $\alpha_K = \alpha_0 \frac{h_K^2}{4\nu}$, with $\alpha_0 = 1/3$ for the P_1 element. Although it is not computationally practical, the third equa-

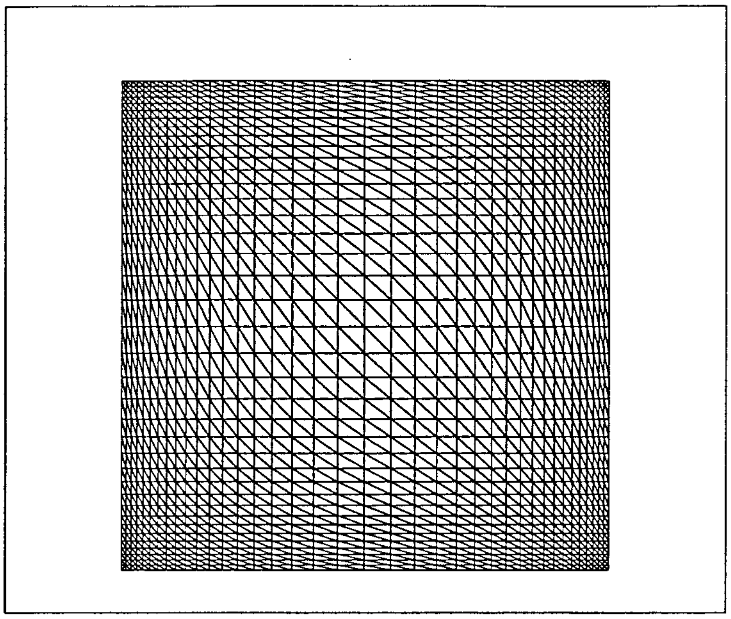


Figure 2.4: Cavity flow, nonuniform 39 × 39 mesh.

tion should actually be replaced by $\sum_{K\in\Theta_h} \alpha_K\Big((\nabla p_h, \mathrm{x}_h)_K - (\mathrm{w}_h, \mathrm{x}_h)_K\Big) = 0$, so as to be consistent with the second equation.

A comparison of the two methods shows that the local one provides faster convergence rates (as can be observed in Table 2.4) and more accurate results (see the next Section). In Table 2.4 we present the number of iterations for convergence in each case, the CPU time spent in the first iteration (as a percentage of the total time in each case) and the total computing time relative to the global method with a consistent mass matrix. A unique relaxation parameter $\omega = 1.3$ was used in this problem for all the variables.

Method	Global-C	Global-L	Local-C	Local-L
Iterations	13	8	7	7
1st. Iteration	80	82	84	82
CPU time	100	75	93	76

Table 2.4: Comparison of global and local parameters.

2.5.5 Summary of computational aspects

We have considered several possibilities for the solution of the linear equation system 2.24–2.25–2.26, resulting from a finite element discretization of the reformulated Stokes problem 2.14. They include two iterative schemes of the *block–Gauss–Seidel* type, which we call uncoupled and coupled, respectively, and some other ideas such as rescaling or diagonally preconditioning the equations and the use of successive–over–relaxation methods. For the pressure gradient mass matrix, we also considered lumping methods.

After several tests on a uniform mesh for the lid–driven cavity flow problem with the P_1, Q_1, P_2 and Q_2 elements, we have found that the most efficient method is the coupled block–Gauss-Seidel scheme applied to the original equation system (without rescaling) with selective successive–over–relaxation, which acts on the pressure and pressure gradient variables with an optimal relaxation parameter ω_{opt}, but leaves the velocity unrelaxed. Different optimal relaxation parameters were found experimentally for each element: 1.3 for the P_1, 1.5 for the Q_1, 1.4 for the P_2 and 1.79 for the Q_2.

For the P_1 element, the system matrix of the velocity–pressure subsystem to be solved at each iteration of that method is the same as that of the GLS method; this is also the case for the Q_1 element, if the Laplacian term is omitted on the GLS method from element interiors (for a mesh of parallelograms, the transformations from the reference element are affine, and this term vanishes identically for this element). The GLS formulation is one of the most widely used methods nowadays for incompressible flow problems; like our method, it is formulated in terms of primitive velocity–pressure variables, and it also allows the use of equal order interpolations. When direct solution methods are used to solve the velocity–pressure subsystems, the extra cost of our scheme with respect to the GLS method is that of computing and factorizing the mass matrix for the pressure gradient (if it is not lumped), forming the right–hand–side vectors and performing forward and backward substitutions at a few extra iterations, since the system matrices are computed and factorized only once. This represents, in average, about 25% of extra cost.

When nonuniform meshes are used, it seems more efficient to use local stability parameters α_K defined elementwise, rather than a unique global parameter α.

Finally, mass lumping also accelerates the convergence of the unrelaxed coupled block–Gauss-Seidel scheme, at the expense of a loss of accuracy (see the next Section). However, it does not affect the convergence of the faster over–relaxed coupled block–Gauss-Seidel schemes. It is, therefore, not recommended for general use with this method.

2.6 Numerical results

We now present the numerical results obtained with the reformulated method 2.21–2.22–2.23 for three test cases: the cavity flow problem considered in the previous Section, a problem with an analytical solution and a channel flow problem problem on a trapezoidal domain. These problems highlight other features of the method than those proved up to now, and confirm some of these.

2.6.1 Cavity flow problem

We present some selected numerical solutions obtained for the cavity flow problem in the convergence studies of the previous Section. The main flow features which we looked at in this problem are flow symmetry about the vertical centerline of the cavity and the pressure singularity at the two top corners. We will show that the method provides an excellent capturing of this singularity. Since the zero prescription for the pressure is enforced at the top right corner (the 'last node'), the minimum pressure value corresponds to the top left corner, and stablishes the (negative) pressure singularity.

We first present results obtained with the uniform 21 × 21 mesh, both with and without mass lumping. Pressure singularity capturing degrades with mass lumping. These results can be observed in Figure 2.5 for triangular elements and Figure 2.6 for quadrilateral elements.

Better results were obtained with the finer 39 × 39 nonuniform mesh of Figure 2.4. We compare the results obtained with a global stability parameter α and with local parameters α_K, both with and without mass lumping. The solution was symmetric in all cases. We present the pressure results in Figure 2.7, where it is observed that the best results are achieved with local stability parameters and a consistent mass matrix.

We conclude that the best results obtained for this problem with the nonuniform 39 × 39 mesh were for local stability parameters and a consistent mass matrix, iterating the scheme 2.103–2.104–2.105 to convergence: a pressure minimum of -1070 was achieved in this case. For the uniform 21 × 21 mesh, the best results correspond to the Q_2 element with a consistent mass matrix: the pressure minimum was of -1160 in that case.

2.6.2 A test with an analytical solution

We next consider a test problem with an analytical polynomial solution on the unit square, with homogeneous Dirichlet boundary conditions, which was introduced by J.T. Oden and coworkers (see [78]). We study this case in order to check numerically the optimal error estimates proved in Section 2.3. Setting $\nu = 1$, a polynomial force is selected so that the solution of 1.13 is $\mathbf{u} = (u_x, u_y)$ with:

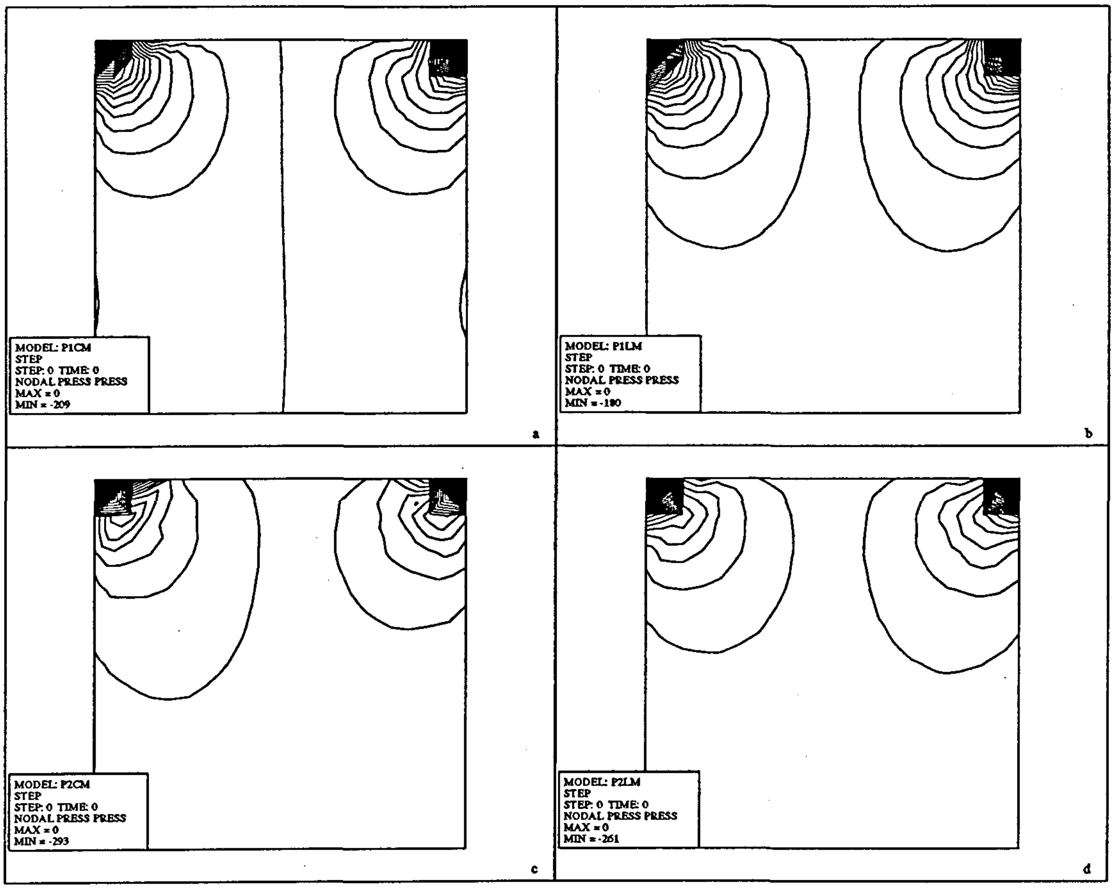


Figure 2.5: Cavity flow, uniform 21 × 21 mesh, triangular elements, pressure contours: a) $P_1 - C$; b) $P_1 - L$; c) $P_2 - C$; d) $P_2 - L$.

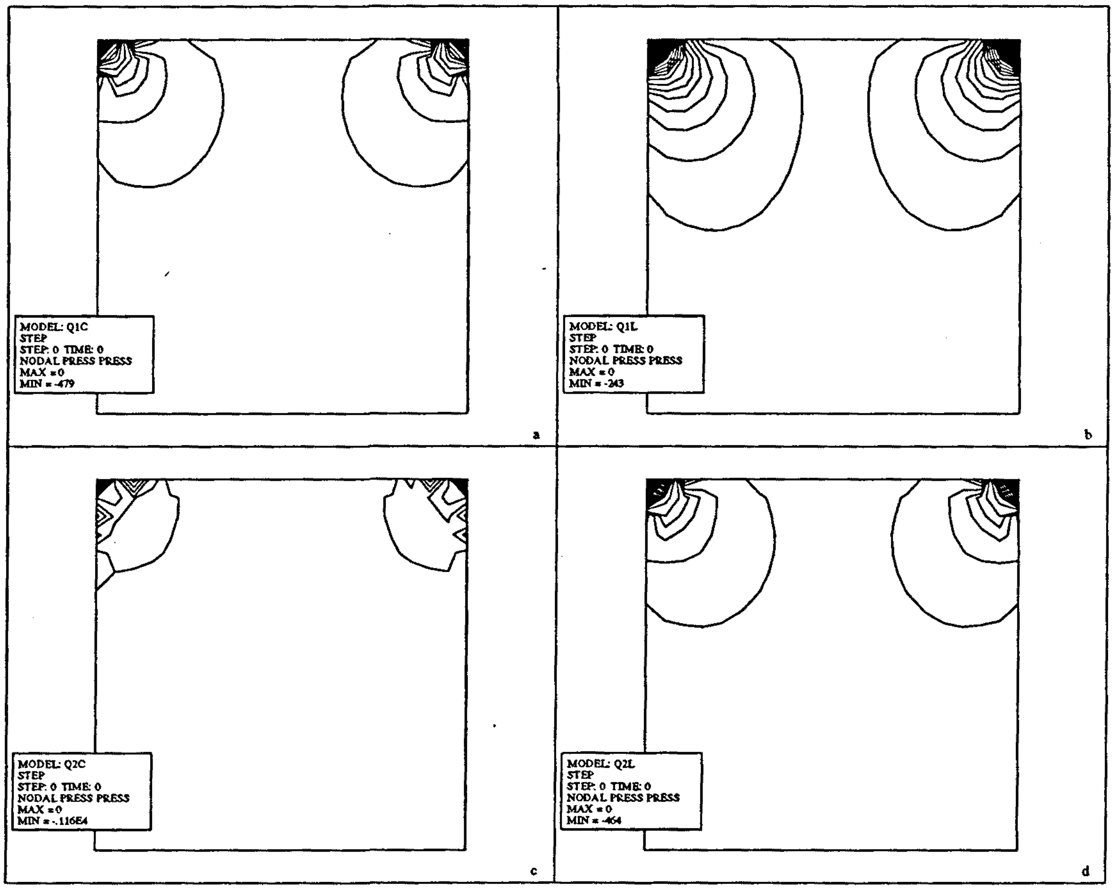


Figure 2.6: Cavity flow, uniform 21×21 mesh, quadrilateral element, pressure contours: a) $Q_1 - C$; b) $Q_1 - L$; c) $Q_2 - C$; d) $Q_2 - L$.

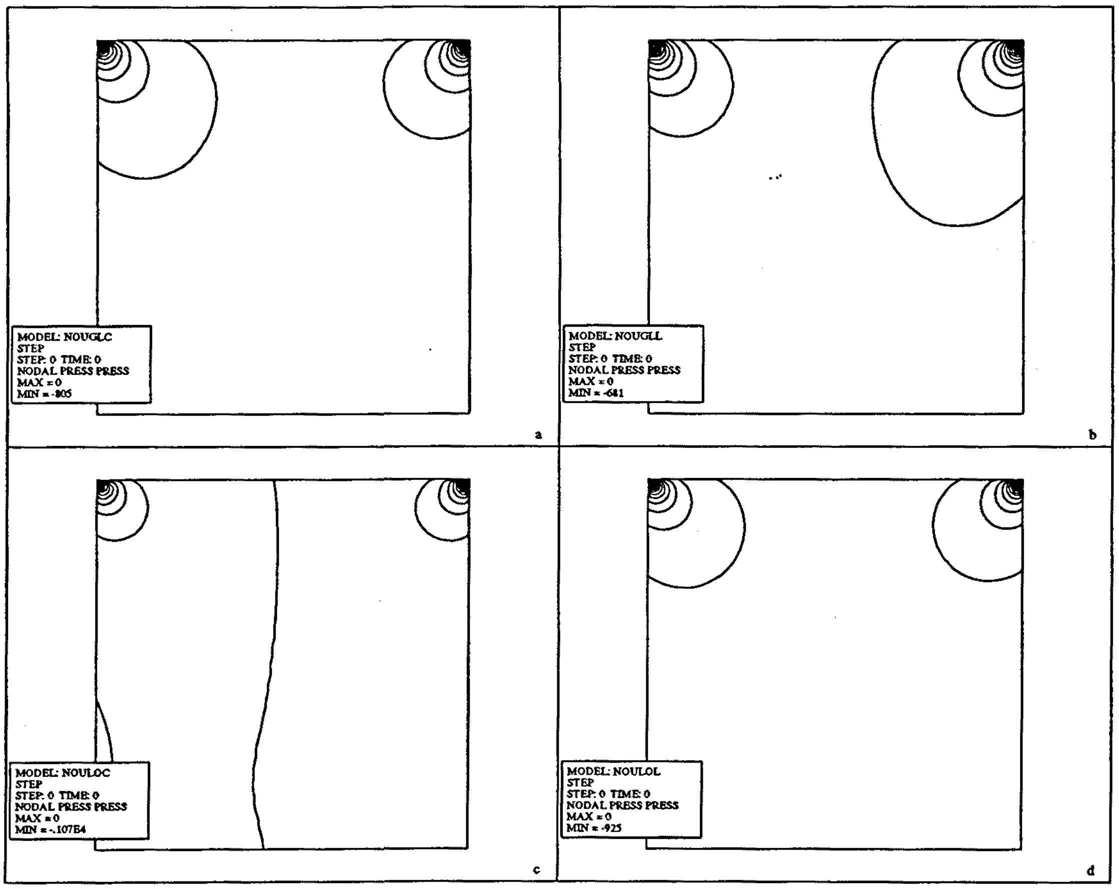


Figure 2.7: Cavity flow, nonuniform 39 × 39 mesh, pressure contours: a) global α, consistent mass matrix; b) global α, lumped mass matrix; c) local α, consistent mass matrix; d) local α, lumped mass matrix.

$$\begin{aligned} u_x &= x^2(1-x^2)(2y-6y^2+4y^3) \\ u_y &= (2x-6x^2+4x^3)y^2(1-y^2) \end{aligned} \tag{2.109}$$

for $0 \leq x, y \leq 1$, and the pressure solution is then:

$$p = x - x^2 \tag{2.110}$$

(so that p vanishes at the top right corner). With this test case, we first performed a study of the influence of the parameter α_0 on the exact errors of velocity, pressure and pressure gradient solutions in the L^2 and H_0^1 norms on a uniform mesh; we will see, in particular, that there are minimum values of these errors at critical values of the parameter, which we select. Uniform meshes are used here so that a study of the order of error with respect to a characteristic mesh size h can be provided, using the optimal values of the parameter α_0 just obtained. The expected orders of accuracy for all the variables, norms and elements have been found.

The first results we present were obtained with a uniform 21 × 21 mesh, using the coupled block Gauss–Seidel method with a consistent mass matrix and a tolerance of $\epsilon_{\text{cou}} = 10^{-5}$. In Figure 2.8 we show the variation of the exact error of the velocity in $\mathbf{L}^2(\Omega)$ and $\mathbf{H}_0^1(\Omega)$, the pressure in $L^2(\Omega)$ and the pressure gradient in $\mathbf{L}^2(\Omega)$, both with respect to ∇p_h and $\mathbf{w}_h$, as a function of the coefficient α_0, for the elements P_1, Q_1, P_2 and Q_2. It can be seen that minimum values for the pressure error are attained for values of α_0 in a range close to the optimal values of the GLS method for each element type: 1/3 for linear and bilinear elements and 1/9 for quadratic and biquadratic ones; these are the values that we have used up to now, and that we adopt in what follows. Nevertheless, minimum errors for the pressure gradient are achieved at a larger value of α_0 than the critical one. The velocity errors are less sensitive to variations of α_0, but tend to be minimized close to the optimal values of the GLS method.

The variation of α_0 does not only affect the precision of the method but also the convergence rates of the iterative scheme. It is seen in Figure 2.9 that the number of iterations required for convergence grows drastically with an increase of α_0 beyond the critical values, while small values of α_0 yield rapidly convergent schemes, at the expense of a loss of precision.

Having found optimal values of the parameter α_0, we then checked numerically the theoretical orders of accuracy of the velocity, pressure and pressure gradient solutions as a function of the mesh size h, as given by Theorems 2.1 and 2.2.; we summarize these orders of error in Table 2.5 for reference. To this end, we solved the reformulated Stokes problem 2.21–2.22–2.23 for this test case on three uniform meshes with 11 × 11, 21 × 21 and 41 × 41 nodes, respectively, and with the four elements considered up to now. The value of α was always computed as $\alpha = \alpha_0 \dfrac{h^2}{4\nu}$, and α_0 was taken as the

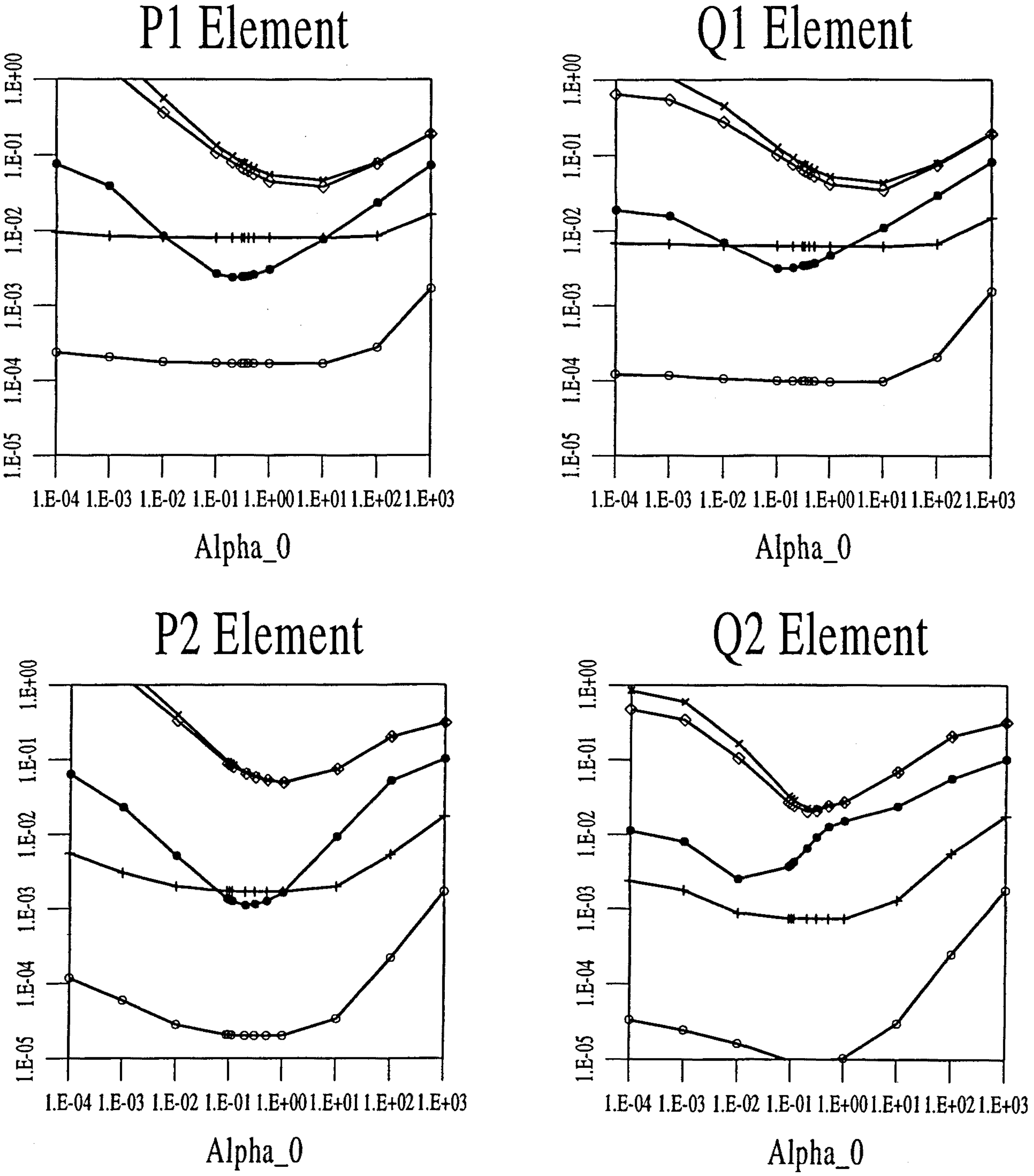


Figure 2.8: Error variation with α_0: $\circ = |\mathbf{u} - \mathbf{u}_h|$; $+ = ||\mathbf{u} - \mathbf{u}_h||$; $\bullet = |p - p_h|$; $\times = |\nabla p - \nabla p_h|$; $\diamond = |\nabla p - \mathbf{w}_h|$.

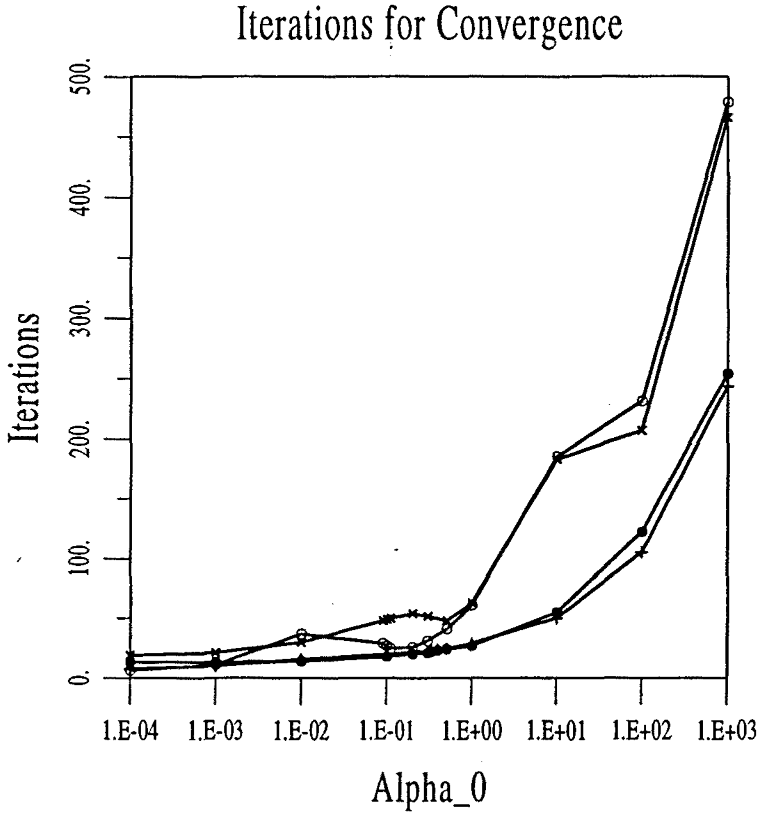


Figure 2.9: Iteration for convergence with α_0: + P_1 Element; • Q_1 Element; o P_2 Element; × Q_2 Element.

optimal values just established. The results were obtained with the coupled block-Gauss-Seidel scheme 2.103-2.104-2.105 with a consistent mass matrix and a tolerance of $\epsilon_{cou} = 10^{-3}$; we also tried higher and lower values of the tolerance: in the first case, larger errors were found, whereas in the second the precision did not improve (but rather degraded due to round-off errors). We present these results in Figures 2.10 to 2.14, where we have included the errors for a P_2P_1 mixed interpolation of the Stokes problem for comparison. The linear regression coefficients computed for these lines are given in Table 2.6.

It can be observed that the optimal orders of accuracy predicted in Table 2.5, and even higher orders for the pressure solution, are achieved with our

Element	$\|\mathbf{u}-\mathbf{u}_h\|$	$\|p-p_h\|$	$\|\|\mathbf{u}-\mathbf{u}_h\|\|$	$\|\nabla(p-p_h)\|$	$\|\nabla p-\mathbf{w}_h\|$
P_1, Q_1	2	1	1	0	0
P_2, Q_2	3	2	2	1	1

Table 2.5: Theoretical orders of error in the mesh size h.

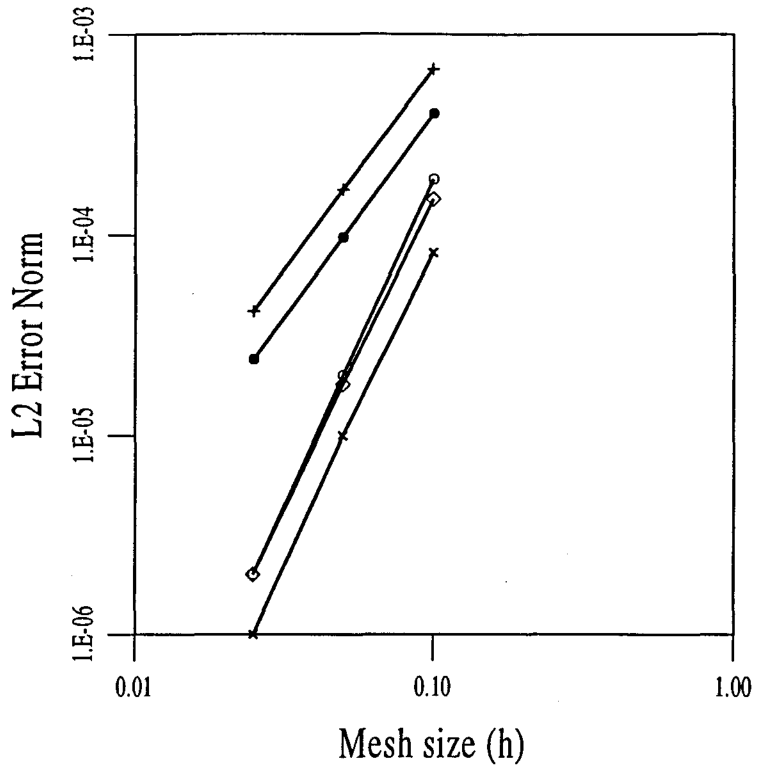


Figure 2.10: Velocity error in L^2: $+$ P_1 Element; $\bullet$ Q_1 Element; $\circ$ P_2 Element; $\times$ Q_2 Element; $\diamond$ Mixed P_2P_1 Element.

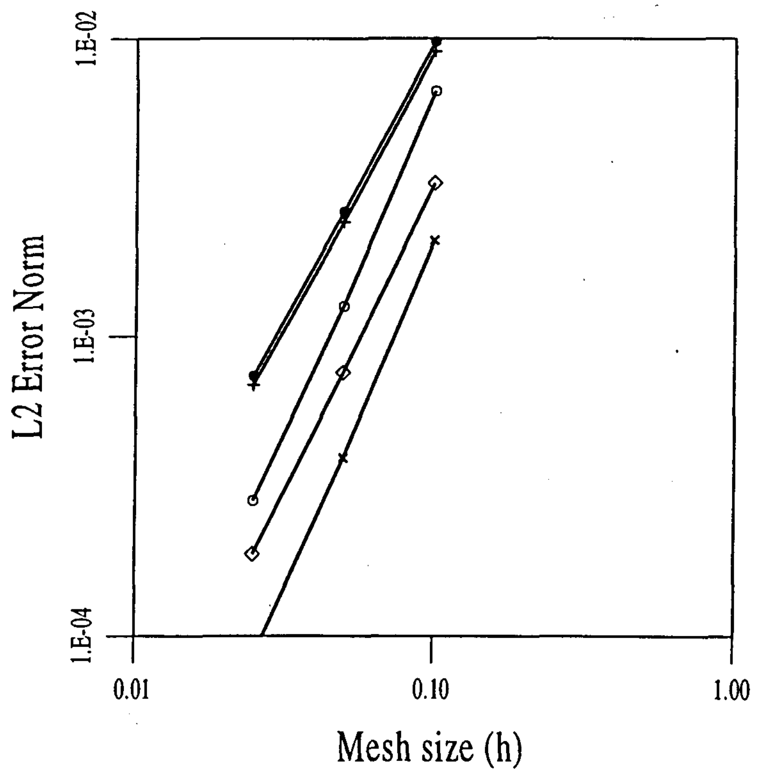


Figure 2.11: Pressure error in L^2: $+$ P_1 Element; $\bullet$ Q_1 Element; $\circ$ P_2 Element; $\times$ Q_2 Element; $\diamond$ Mixed P_2P_1 Element.

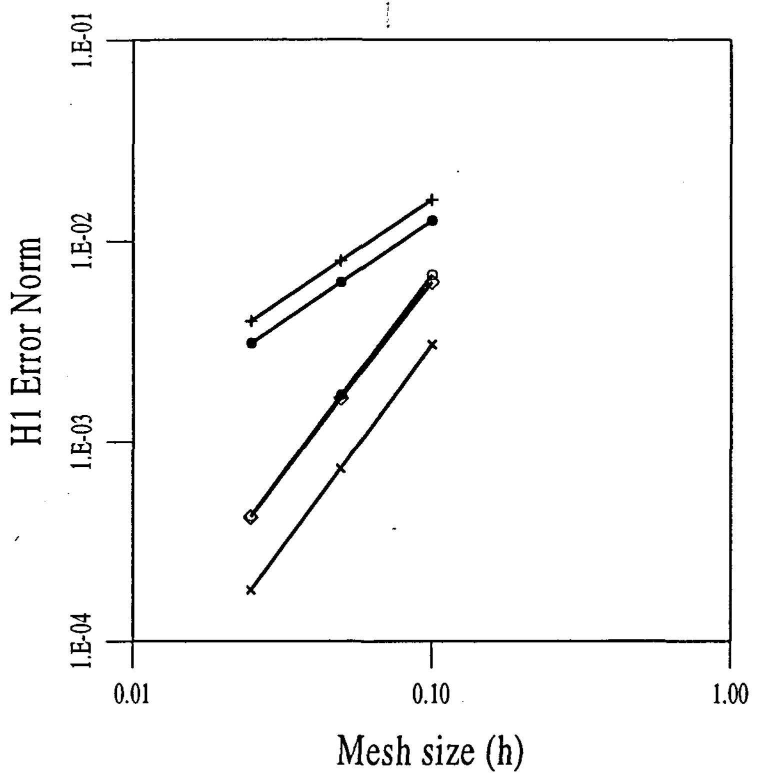


Figure 2.12: Velocity error in H^1: + P_1 Element; • Q_1 Element; ∘ P_2 Element; × Q_2 Element; ◇ Mixed P_2P_1 Element.

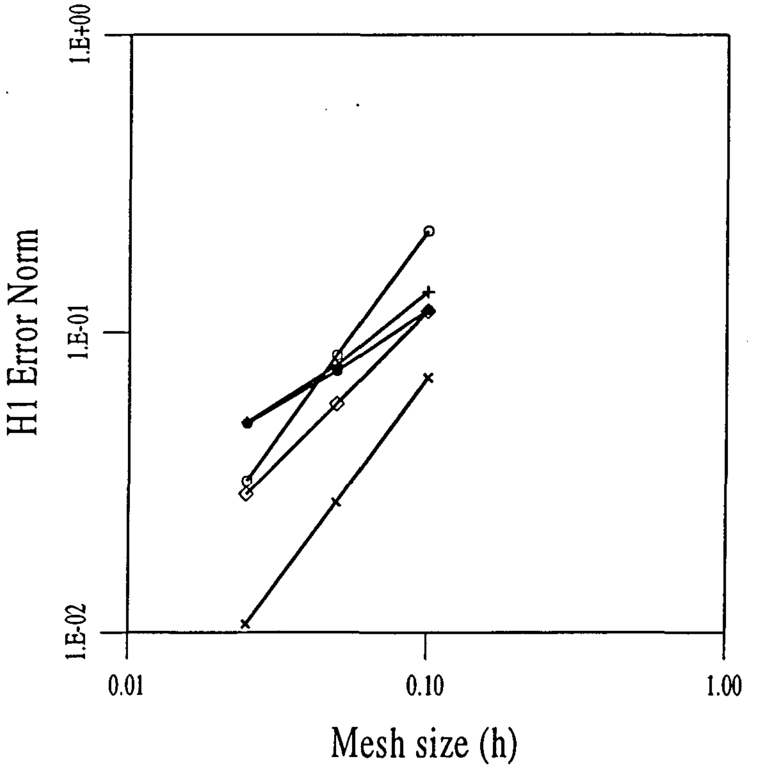


Figure 2.13: Pressure error in H^1: + P_1 Element; • Q_1 Element; ∘ P_2 Element; × Q_2 Element; ◇ Mixed P_2P_1 Element.

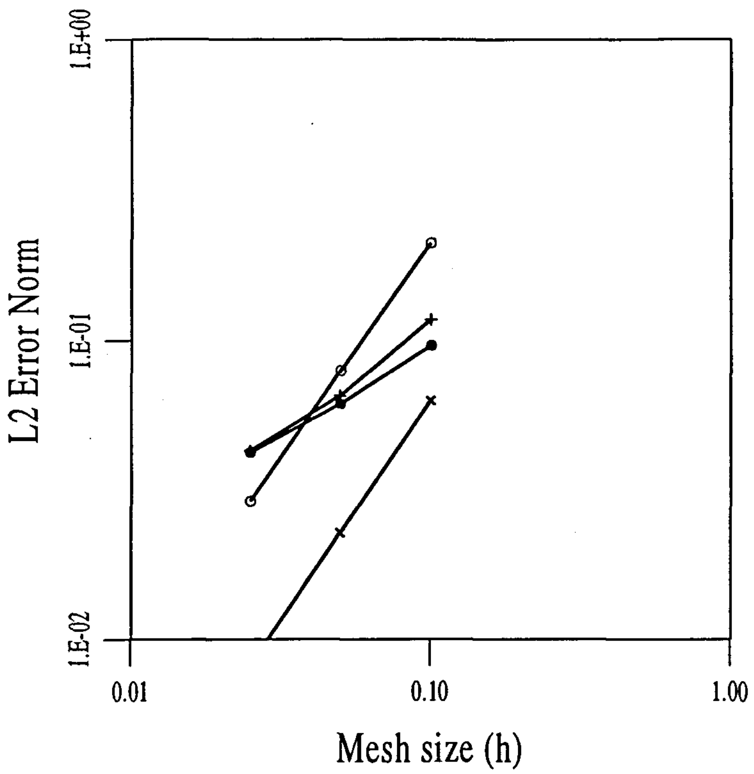


Figure 2.14: Pressure gradient error in L^2: $+$ P_1 Element; • Q_1 Element; o P_2 Element; $\times$ Q_2 Element.

Element	$\lvert\mathbf{u}-\mathbf{u}_h\rvert$	$\lvert p-p_h\rvert$	$\lVert\mathbf{u}-\mathbf{u}_h\rVert$	$\lvert\nabla p-\nabla p_h\rvert$	$\lvert\nabla p-\mathbf{w}_h\rvert$
P_1	2.0	1.9	1.0	0.7	0.7
Q_1	2.0	1.9	1.0	0.6	0.6
P_2	3.3	2.3	2.0	1.4	1.4
Q_2	3.2	2.3	2.0	1.4	1.5

Table 2.6: Oden's flow: linear regression coefficients for different errors.

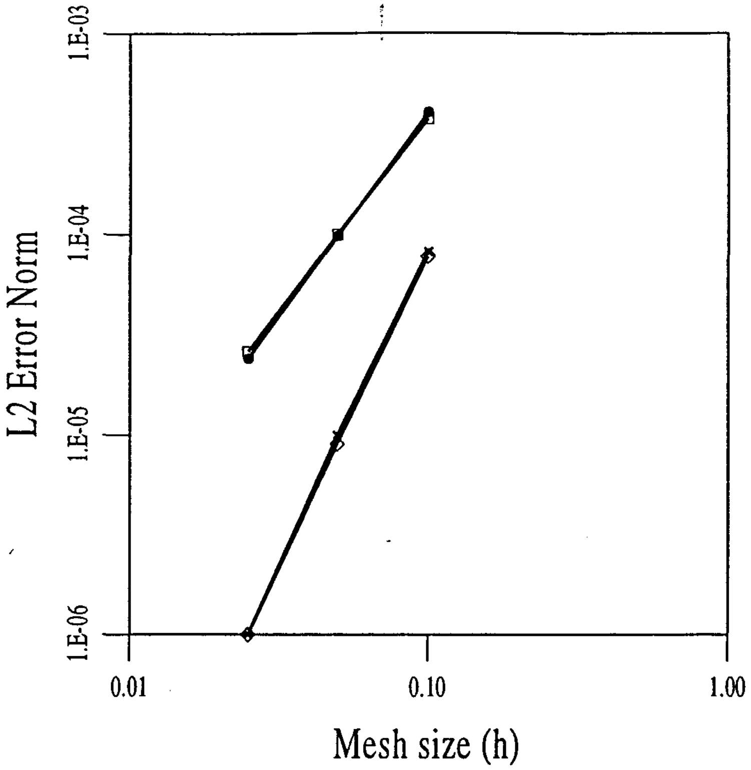


Figure 2.15: Velocity error in L^2: • Q_1 Element; × Q_2 Element; □ = GLS method, Q_1 Element; ◇ GLS method, Q_2 Element.

method for all variables and norms. Moreover, both the discrete pressure gradient and the gradient of the discrete pressure seem to converge for the P_1 and Q_1 elements to ∇p, a fact which is not predicted by the theory.

We then compared the accuracy results obtained with our method to those of the GLS formulation. We show in Figures 2.15 to 2.18 the errors computed for the solutions on quadrilateral elements, both the Q_1 and the Q_2, for the GLS method and ours. The velocity solution is the same for the two methods with both elements; the pressure solution, however, is slightly more accurate for the GLS method than ours when using bilinear elements, at least for the present value of the parameter α_0; for biquadratic elements, however, our method seems to be assymptotically more accurate.

2.6.3 Behaviour of the pressure near the boundary

This last example is intended to discuss a misbehaviour of the pressure near the boundary which appears when using the GLS method, as described by J.J. Droux and T.J.R. Hughes in [31]. Although this method is optimal both in H^1 and L^2 norms, the pressure may be poorly approximated near the boundary for linear elements. This is so because in this case the term $\nu\Delta\mathbf{u}_h$ vanishes identically on element interiors, so that a wrong boundary condition $\mathbf{n}\cdot(\nabla p_h - \mathbf{f}) = 0$ is being imposed weakly (see the modification of the GLS method in [31] to overcome this difficulty).

In our reformulated method the boundary condition $\mathbf{n}\cdot\nabla p \;-\; \mathbf{n}\cdot\mathbf{w} = 0$

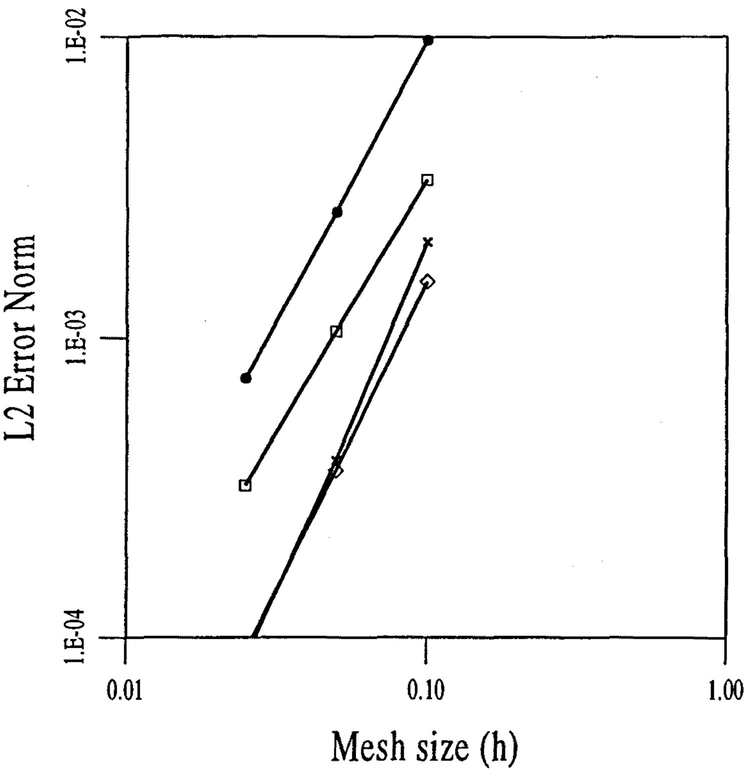


Figure 2.16: Pressure error in L^2: • Q_1 Element; × Q_2 Element; □ = GLS method, Q_1 Element; ◇ GLS method, Q_2 Element.

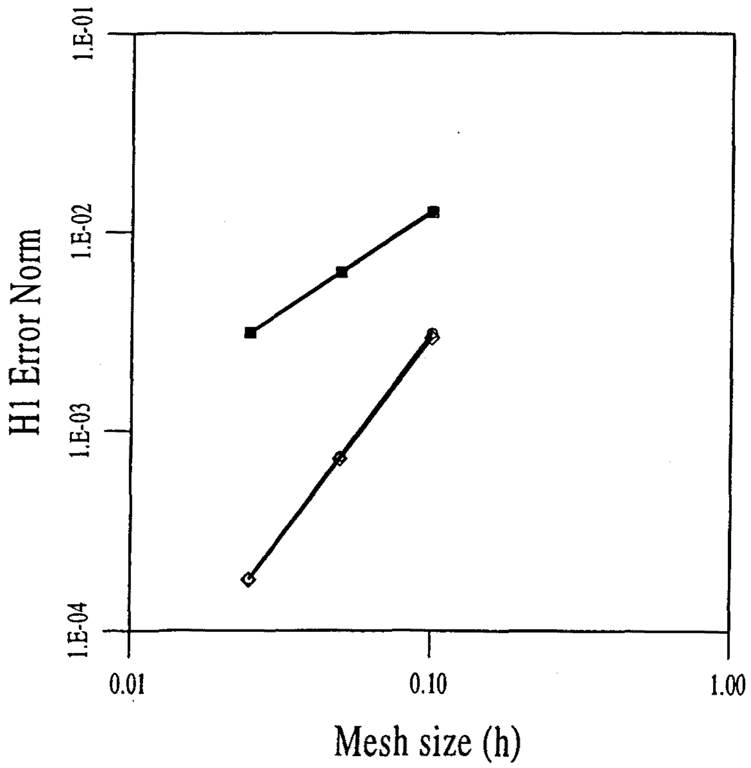


Figure 2.17: Velocity error in H^1: • Q_1 Element; × Q_2 Element; □ = GLS method, Q_1 Element; ◇ GLS method, Q_2 Element.

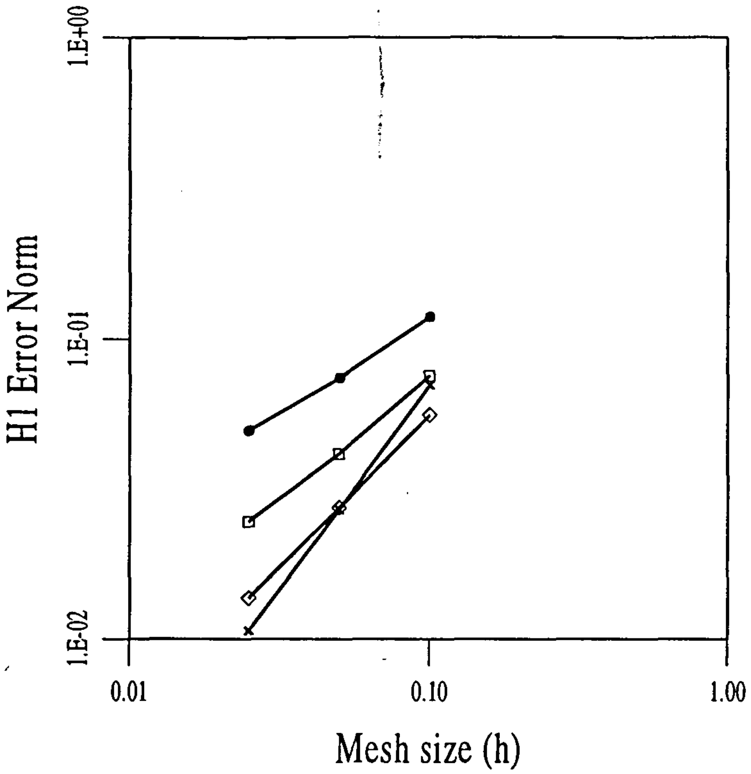


Figure 2.18: Pressure error in H^1: • Q_1 Element; × Q_2 Element; □ = GLS method, Q_1 Element; ◇ GLS method, Q_2 Element.

enforced weakly is consistent with the original Stokes problem, so that correct behaviour of the pressure near the boundary was expected. The same test case as in [31] was considered, consisting of fully developed Poiseuille flow on a 2-dimensional trapezoidal domain. We solved this problem on two meshes of P_1 elements, with 13 × 13 and 25 × 25 nodes uniformly distributed along the sides. The first mesh is shown in Figure 2.19. For this problem, there is no external force, a parabolic velocity profile is prescribed both at the inlet and outlet and a solid wall condition is imposed on the top and bottom edges. The pressure gradient in this case is constant and horizontal.

The pressure contours obtained for the GLS method and the reformulated method on both meshes are shown in Figure 2.20. Although they improve with mesh refinement, the pressure results for the GLS method are not correct near the boundary, whereas the results for our method are exact on both meshes.

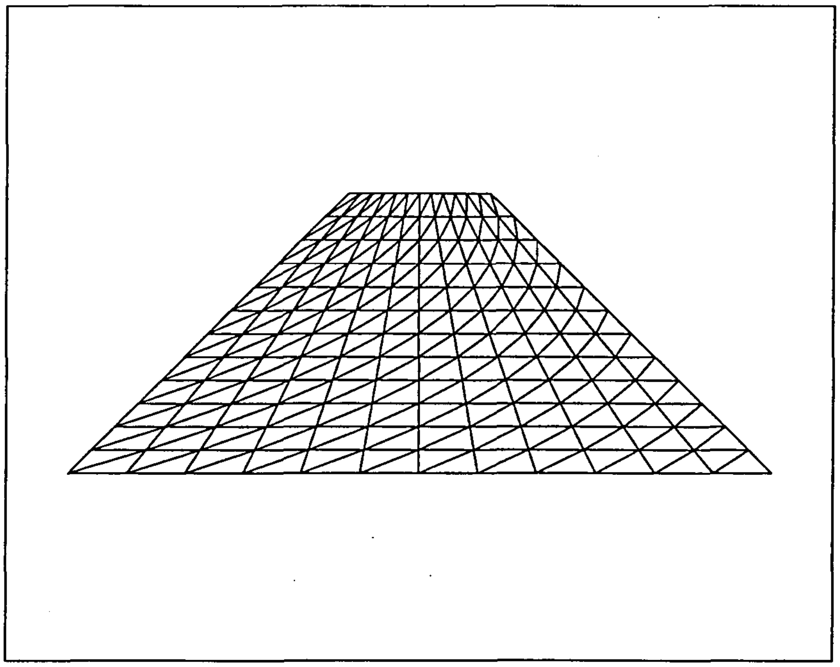


Figure 2.19: Trapezoidal domain, coarse mesh.

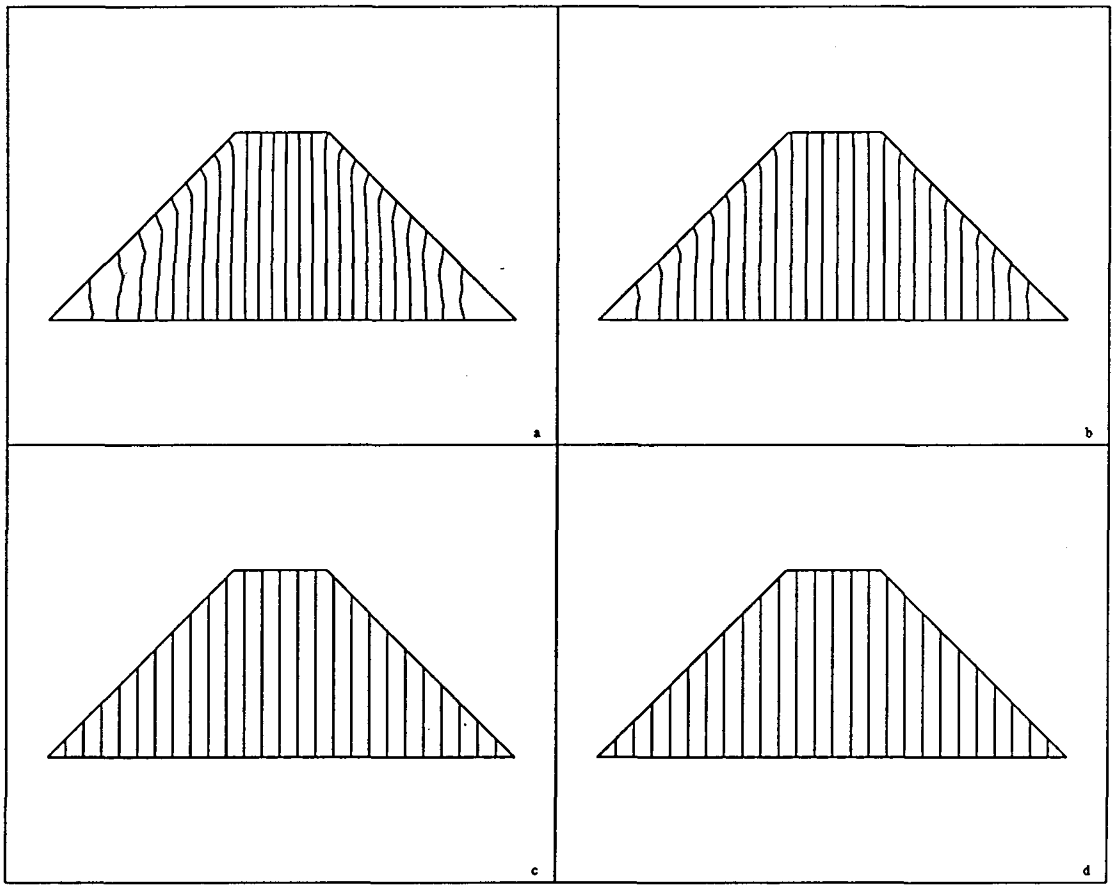


Figure 2.20: Trapezoidal domain, pressure contours: a) GLS method, coarse mesh; b) GLS method, fine mesh; c) Present method, coarse mesh; d) Present method, fine mesh.

Chapter 3

Reformulated Navier–Stokes equations

The object of this Chapter is to extend the reformulated method studied in Chapter 2 to the steady, incompressible Navier–Stokes equations 1.12. The difference of this equations with the Stokes problem 1.13 is the appearance of a nonlinear term embracing convective effects, which was neglected in 1.13.

The incompressible Navier–Stokes equations model a large number of flow situations, and are used in many practical applications. Moreover, they are the 'next stage' towards the study of the unsteady Navier–Stokes equations: they are still affected by the incompressibility condition and they introduce the difficulties relative to the nonlinearity, but not the time evolution yet.

The development and study of numerical methods to approximate the solution of these equations has received much attention in the last decades. Besides incompressibility, they have to deal with the treatment of the nonlinearity of the problem and the advective–diffusive character of the equations, which is specilly hard for high Reynolds number flows. Thus, nonlinear solvers and, in some cases, techniques to stabilize the convection, are required to approximate these equations, as well as adequate treatment of incompressibility.

We review a few basic facts about the steady, incompressible Navier–Stokes equations in Section 3.1, concerning the existence, uniqueness and approximation of solutions. In Section 3.2, we present the extention of the reformulated method to this problem, while in 3.3 we prove stability and optimal convergence of the method to the solution of the equations, assuming uniqueness of such a solution and the stability condition 2.28 of the linear, reformulated Stokes problem. We then consider several possibilities for the iterative solution of the resulting nonlinear system of discrete equations, which we present in Section 3.4. Finally, we show some numerical results obtained with this method on three test cases, including a numerical convergence study which confirms the optimal error estimates proved theoretically.

3.1 The steady, incompressible Navier Stokes equations

We recall here the steady, incompressible Navier–Stokes equations for reference, with homogeneous Dirichlet boundary conditions:

$$\begin{aligned}(\mathbf{u}\cdot\nabla)\mathbf{u} - \nu\Delta\mathbf{u} + \nabla p &= \mathbf{f} \quad \text{in } \Omega \\ \nabla\cdot\mathbf{u} &= 0 \quad \text{in } \Omega \\ \mathbf{u} &= 0 \quad \text{on } \Gamma\end{aligned} \tag{3.1}$$

Complete studies of this equation system can be found, among others, in [71], [105] and [43]. We mainly follow these last two references here. With the definitions of the operators b and c given in Section 1.2, and given $\mathbf{f} \in \mathbf{H}^{-1}(\Omega)$, the weak form of these equations consists of finding $\mathbf{u} \in \mathbf{H}_0^1(\Omega)$ and $p \in L_0^2(\Omega)$ such that:

$$\begin{aligned}c(\mathbf{u},\mathbf{u},\mathbf{v}) + \nu(\nabla\mathbf{u},\nabla\mathbf{v}) + b(\mathbf{v},p) &= <\mathbf{f},\mathbf{v}>, \quad \forall\mathbf{v} \in \mathbf{H}_0^1(\Omega) \\ b(\mathbf{u},q) &= 0, \qquad \forall q \in L_0^2(\Omega)\end{aligned} \tag{3.2}$$

Since we are assuming Ω bounded and Lipschitz continuous, problem 3.2 has at least one solution (see [43]), which satisfies 3.1 in distribution sense. Uniqueness does not hold in general, but it holds for *sufficiently small data* or *sufficiently large viscosity.* The precise form of these statements may be written in different ways; following [105], we take it as the following condition, where from now onwards we assume that $\mathbf{f} \in \mathbf{L}^2(\Omega)$:

$$\Upsilon \doteq \frac{C_\Omega C_{111}|\mathbf{f}|}{\nu^2} < 1 \tag{3.3}$$

The constant C_Ω was introduced in 1.14 and C_{111} is the constant appearing in the standard continuity condition of the trilinear form c.

Under the assumption 3.3, the solution $(\mathbf{u},p)$ of 3.1 is unique. If the homogeneous boundary condition is replaced by a nonhomogeneous condition:

$$\mathbf{u} = \tilde{\mathbf{u}} \quad \text{on } \Gamma$$

where $\tilde{\mathbf{u}}$ satisfies the null flux condition 1.9, similar existence and uniqueness results can be obtained, the latter under a condition similar to 3.3.

Standard Galerkin finite element approximation of the Navier–Stokes problem 3.2 is subject to the same compatibility restrictions as the Stokes problem: the *inf–sup* condition 1.27 should hold for standard optimal convergence results. This is the case for the Q_2P_1 element; for the popular Q_1P_0 element, however, macroelement techniques may be used again.

We introduce here the matrix form of a discretization of the Navier–Stokes equations 3.2 by the Galerkin finite element method. In the notation used up to now, the discrete version of 3.2 can be written as:

$$\begin{array}{rcl} A(U)U \ + \ KU \ + \ G_0P & = & F \\ G_0^tU & = & 0 \end{array} \qquad (3.4)$$

where $A(U)$ is the convective matrix with a given (nodal) velocity field U. Some iterative method should be used to find a solution of the nonlinear problem 3.4; standard schemes for this problem are Picard's iteration and Newton–Raphson's method, which are first and second order schemes respectively, apart from methods of gradient type (see [43]). One drawback of Newton–Raphson's method is that the initial approximation used in it should belong to the attraction basin of the solution for the scheme to converge. These two schemes take the following form in this context:

- Picard's method:

$$\begin{array}{rcl} A(U^{i-1})U^i \ + \ KU^i \ + \ G_0P^i & = & F \\ G_0^tU^i & = & 0 \end{array}$$

- Newton–Raphson's method:

$$\begin{array}{rcl} A(U^{i-1})U^i \ + \ A(U^i)U^{i-1} + \ KU^i \ + \ G_0P^i & = & F \ + \ A(U^{i-1})U^{i-1} \\ G_0^tU^i & = & 0 \end{array}$$

We will use these two approximations in Section 3.4.

3.2 Development of the method

3.2.1 The continuous problem

We present our extension of the reformulated method for the Stokes problem 2.14 to the Navier–Stokes problem 3.1 with homogeneous boundary conditions. Our analysis is valid for any trilinear form $\tilde{c}$ defined on $(\mathbf{H}_0^1(\Omega))^3$ which is skew–symmetric in its last two arguments and continuous; it is also restricted to a class of Navier-Stokes problems with some additional regularity conditions, which we define next in a similar way to the Stokes case.

Definition 3.1: *the steady, incompressible Navier–Stokes equation* 3.1 *is called regular if its solutions satisfy* $\mathbf{u} \in \mathbf{H}^2(\Omega)$ *and* $p \in H^1(\Omega)$ *whenever* $\mathbf{f} \in \mathbf{L}^2(\Omega)$*, and there exists a constant* $C_r > 0$ *such that:*

$$||\mathbf{u}||_2 \ + \ ||p||_1 \ \leq \ C_r \, |\mathbf{f}|$$

As in the linear case, this is too restrictive for our purposes. We only need the pressure gradient to be in $L^2(\Omega)$; we call this case again a p–regular Navier–Stokes problem:

Definition 3.2: *the steady, incompressible Navier–Stokes equation* 3.1 *is called p-regular if its solutions satisfy $p \in H^1(\Omega)$ whenever $\mathbf{f} \in \mathbf{L}^2(\Omega)$*

The reformulation of problem 3.1 that we propose is the following, where we adopt the skew–symmetric form of the convective operator:

$$\begin{array}{rcll} \frac{1}{2}(\nabla\cdot\mathbf{u})\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} - \nu\Delta\mathbf{u} + \nabla p & = & \mathbf{f} & \text{in } \Omega \\ \nabla p - \mathbf{w} & = & 0 & \text{in } \Omega \\ \nabla\cdot\mathbf{u} + \alpha(-\Delta p + \nabla\mathbf{w}) & = & 0 & \text{in } \Omega \\ \mathbf{u} & = & 0 & \text{on } \Gamma \\ \mathbf{n}\cdot\nabla p - \mathbf{n}\cdot\mathbf{w} & = & 0 & \text{on } \Gamma \end{array} \tag{3.5}$$

where, again, $\alpha > 0$. Calling again $V_0 = \mathbf{H}_0^1(\Omega)$, $Q = H^1(\Omega)/\mathrm{I\!R}$ and $V = \mathbf{L}^2(\Omega)$, and assuming $\mathbf{f} \in \mathbf{L}^2(\Omega)$, the weak form of this problem consists of finding $\mathbf{u} \in V_0$, $p \in Q$ and $\mathbf{w} \in V$ such that:

$$\begin{array}{rcl} \tilde{c}(\mathbf{u},\mathbf{u},\mathbf{v}) + \nu(\nabla\mathbf{u},\nabla\mathbf{v}) + (\nabla p,\mathbf{v}) & = & (\mathbf{f},\mathbf{v}),\ \forall\mathbf{v}\in V_0 \\ (\nabla\cdot\mathbf{u},q) + \alpha(\nabla p,\nabla q) - \alpha(\mathbf{w},\nabla q) & = & 0,\quad \forall q\in Q \\ (\nabla p,\mathbf{y}) - (\mathbf{w},\mathbf{y}) & = & 0,\quad \forall\mathbf{y}\in V \end{array} \tag{3.6}$$

We prove that in case the Navier–Stokes problem is p–regular and under the uniqueness condition $\Upsilon < 1$, problem 3.6 has a unique solution, which is the solution of the original problem:

Proposition 3.1: *assume that the Navier–Stokes problem is p–regular and that condition* 3.3 *holds; then, there exists a unique solution $(\mathbf{u},p,\mathbf{w}) \in V_0 \times Q \times V$ of* 3.6*, where $(\mathbf{u},p)$ is the unique solution of* 3.1 *and $\mathbf{w} = \nabla p$ in* $\mathbf{L}^2(\Omega)$.

PROOF: existence is a consequence of the properties of the solution $(\mathbf{u},p)$ assumed. To prove uniqueness, we define a form $\tilde{D}$ on $(V_0 \times Q \times V)^2$ as:

$$\tilde{D}(\mathbf{u},p,\mathbf{w};\mathbf{v},q,\mathbf{y}) = \nu(\nabla\mathbf{u},\nabla\mathbf{v}) + (\nabla p,\mathbf{v}) + \tilde{c}(\mathbf{u},\mathbf{u},\mathbf{v}) + (\nabla\cdot\mathbf{u},q)$$

$$+\alpha(\nabla p,\nabla q) - \alpha(\mathbf{w},\nabla q) - \alpha(\nabla p,\mathbf{y}) + \alpha(\mathbf{w},\mathbf{y})$$

which is quadratic in its first argument and linear in its second. Problem 3.6 can then be written as:

$$\tilde{D}(\mathbf{u},p,\mathbf{w};\mathbf{v},q,\mathbf{y}) = (\mathbf{f},\mathbf{v}), \qquad \forall(\mathbf{v},q,\mathbf{y}) \in (V_0 \times Q \times V)$$

The coercivity of the linear problem is preserved due to the skew-symmetry of the operator $\tilde{c}$:

$$\tilde{D}(\mathbf{u},p,\mathbf{w};\mathbf{u},p,\mathbf{w}) = \nu||\mathbf{u}||^2 \ + \ \alpha|\nabla p - \mathbf{w}|^2, \quad \forall(\mathbf{u},p,\mathbf{w}) \in (V_0 \times Q \times V)$$

so that the stability estimate:

$$||\mathbf{u}|| \leq \frac{C_\Omega|\mathbf{f}|}{\nu} \tag{3.7}$$

holds for any solution. Let now $(\mathbf{u}_*,p_*,\mathbf{w}_*)$ and $(\mathbf{u}_{**},p_{**},\mathbf{w}_{**})$ be two such solutions. We call $(\bar{\mathbf{u}},\bar{p},\bar{\mathbf{w}})$ their difference, so that for all $(\mathbf{v},q,\mathbf{y}) \in (V_0 \times Q \times V)$:

$$\tilde{D}(\bar{\mathbf{u}},\bar{p},\bar{\mathbf{w}};\mathbf{v},q,\mathbf{y}) = -\tilde{c}(\mathbf{u}_{**},\bar{\mathbf{u}},\mathbf{v}) \ - \ \tilde{c}(\bar{\mathbf{u}},\mathbf{u}_*,\mathbf{v}) \ + \ \tilde{c}(\bar{\mathbf{u}},\bar{\mathbf{u}},\mathbf{v}) \tag{3.8}$$

Thus:

$$\nu||\bar{\mathbf{u}}||^2 \ + \ \alpha|\nabla\bar{p} - \bar{\mathbf{w}}|^2 \ = \ \tilde{D}(\bar{\mathbf{u}},\bar{p},\bar{\mathbf{w}};\bar{\mathbf{u}},\bar{p},\bar{\mathbf{w}}) \ = \ -\tilde{c}(\bar{\mathbf{u}},\mathbf{u}_*,\bar{\mathbf{u}})$$

Therefore:

$$\nu||\bar{\mathbf{u}}||^2 \ \leq \ C_{111}||\bar{\mathbf{u}}||^2||\mathbf{u}_*|| \ \leq \ C_{111}\frac{C_\Omega|\mathbf{f}|}{\nu}||\bar{\mathbf{u}}||^2,$$

so that condition 3.3 implies $\bar{\mathbf{u}} = 0$, that is, $\mathbf{u}_* = \mathbf{u}_{**}$. The continuous LBB condition ensures that there exists $\mathbf{v} \in V_0$ such that:

$$\beta|\bar{p}|^2 \ \leq \ \frac{(\nabla\bar{p},\mathbf{v})}{||\mathbf{v}||} \ = \ 0$$

according to 3.8 with $\bar{\mathbf{u}} = 0$; this implies $p_* = p_{**}$ in Q. Finally, $\bar{\mathbf{w}} = \nabla\bar{p} = 0$ yields $\mathbf{w}_* = \mathbf{w}_{**}$, so that the solution is indeed unique. □

3.2.2 The discrete problem

Let now $V_{h,0} \subset V_0$, $Q_h \subset Q$ and $V_h \subset V$ be finite dimensional subspaces associated to a discretization of Ω into finite elements, indexed by $h > 0$. The discrete version of 3.6 reads:

$$\begin{aligned} \tilde{c}(\mathbf{u}_h,\mathbf{u}_h,\mathbf{v}_h) \ + \ \nu(\nabla\mathbf{u}_h,\nabla\mathbf{v}_h) \ + \ (\nabla p_h,\mathbf{v}_h) \ &= \ (\mathbf{f},\mathbf{v}_h), \forall\mathbf{v}_h \in V_{h,0} \\ (\nabla\cdot\mathbf{u}_h,q_h) \ + \ \alpha(\nabla p_h,\nabla q_h) \ - \ \alpha(\mathbf{w}_h,\nabla q_h) \ &= \ 0, \ \forall q_h \in Q_h \\ (\nabla p_h,\mathbf{y}_h) \ - \ (\mathbf{w}_h,\mathbf{y}_h) \ &= \ 0, \ \forall\mathbf{y}_h \in V_h \end{aligned} \tag{3.9}$$

We prove existence of a solution of 3.9 as the limit of an iterative Picard's method. Given $\mathbf{u}_h^0 \in V_{h,0}$ arbitrary, we generate a sequence $(\mathbf{u}_h^i, p_h^i, \mathbf{w}_h^i) \in V_{h,0} \times Q_h \times V_h$ such that:

$$\begin{array}{rcl} \tilde{c}(\mathbf{u}_h^{i-1}, \mathbf{u}_h^i, \mathbf{v}_h) + \nu(\nabla \mathbf{u}_h^i, \nabla \mathbf{v}_h) + (\nabla p_h^i, \mathbf{v}_h) & = & (\mathbf{f}, \mathbf{v}_h), \forall \mathbf{v}_h \in V_{h,0} \\ (\nabla \cdot \mathbf{u}_h^i, q_h) + \alpha(\nabla p_h^i, \nabla q_h) - \alpha(\mathbf{w}_h^i, \nabla q_h) & = & 0, \ \forall q_h \in Q_h \\ (\nabla p_h^i, \mathbf{y}_h) - (\mathbf{w}_h^i, \mathbf{y}_h) & = & 0, \ \forall \mathbf{y}_h \in V_h \end{array} \quad (3.10)$$

This linear problem has a unique solution if the interpolation satisfies condition 2.28:

Proposition 3.2: *assume that $V_{h,0}$, Q_h and V_h satisfy 2.28. Then, given $\mathbf{u}_h^{i-1} \in V_{h,0}$, 3.10 has a unique solution $(\mathbf{u}_h^i, p_h^i, \mathbf{w}_h^i) \in V_{h,0} \times Q_h \times V_h$.*

PROOF: problem 3.10 can be seen as a reformulated Stokes problem; the bilinear form associated to it is defined on $(V_{h,0} \times Q_h \times V_h)^2$ by:

$$\tilde{D}_i(\mathbf{u}_h, p_h, \mathbf{w}_h; \mathbf{v}_h, q_h, \mathbf{y}_h) = \nu(\nabla \mathbf{u}_h, \nabla \mathbf{v}_h) + (\nabla p_h, \mathbf{v}_h) + \tilde{c}(\mathbf{u}_h^{i-1}, \mathbf{u}_h, \mathbf{v}_h)$$

$$+(\nabla \cdot \mathbf{u}_h, q_h) + \alpha(\nabla p_h, \nabla q_h) - \alpha(\mathbf{w}_h, \nabla q_h) - \alpha(\nabla p_h, \mathbf{y}_h) + \alpha(\mathbf{w}_h, \mathbf{y}_h)$$

This form satisfies the same coercivity condition as the linear problem, since, for all $(\mathbf{u}_h, p_h, \mathbf{w}_h) \in (V_{h,0} \times Q_h \times V_h)$:

$$\tilde{D}_i(\mathbf{u}_h, p_h, \mathbf{w}_h; \mathbf{u}_h, p_h, \mathbf{w}_h) = \nu||\mathbf{u}_h||^2 + \alpha|\nabla p_h - \mathbf{w}_h|^2,$$

Problem 3.10 then reads:

$$\tilde{D}_i(\mathbf{u}_h, p_h, \mathbf{w}_h; \mathbf{v}_h, q_h, \mathbf{y}_h) = (\mathbf{f}, \mathbf{v}_h), \quad \forall(\mathbf{v}_h, q_h, \mathbf{y}_h) \in (V_{h,0} \times Q_h \times V_h)$$

so that any solution satisfies the stability estimate:

$$||\mathbf{u}_h|| \leq \frac{C_\Omega |\mathbf{f}|}{\nu} \quad (3.11)$$

Moreover:

$$\alpha|P_{h,3}(\nabla p_h^i)|^2 = \alpha|\nabla p_h^i - \mathbf{w}_h|^2 \leq (\mathbf{f}, \mathbf{u}_h^i) \leq \frac{C_\Omega^2 |\mathbf{f}|^2}{\nu}$$

so that:

$$|P_{h,3}(\nabla p_h^i)| \le \frac{C_\Omega |\mathbf{f}|}{(\alpha\nu)^{1/2}} \tag{3.12}$$

Since problem 3.10 is linear and finite dimensional, it is sufficient to show that the homogeneous problem has a unique solution. Setting $\mathbf{f} = 0$ in 3.11 and 3.12, we get $\mathbf{u}_h^i = 0$ and $P_{h,3}(\nabla p_h^i) = 0$. Taking $\mathbf{v}_h = P_{h,1}(\nabla p_h^i)$ in 3.10, given that $|P_{h,1}(\nabla p_h^i)|^2 = (\nabla p_h^i, P_{h,1}(\nabla p_h^i))$, we obtain $P_{h,1}(\nabla p_h^i) = 0$, and, by 2.28, $\nabla p_h^i = 0$. Finally $\mathbf{w}_h^i = P_{h,12}(\nabla p_h^i) = 0$. □

We next prove that the sequence of iterates converges to a solution of 3.9. In particular, this establishes the existence of such a solution. For this purpose, we require condition 2.28 and the inverse inequality 1.23 to hold; we will also assume the uniqueness condition $\Upsilon < 1$, which will be shown later to be a sufficient condition for uniqueness here too.

Proposition 3.3: *assume that the discretization Θ_h of Ω is uniformly regular, so that the inverse inequality* 1.23 *holds, that the interpolation satisfies condition* 2.28 *and that* 3.3 *also holds. Then, for arbitrary $\mathbf{u}_h^0 \in V_{h,0}$, the sequence of Picard iterates $(\mathbf{u}_h^i, p_h^i, \mathbf{w}_h^i)$ converges to a solution $(\mathbf{u}_h, p_h, \mathbf{w}_h)$ of* 3.9.

PROOF: substracting 3.10 for i and $i-1$, we get:

$$\begin{aligned}
\nu(\nabla(\mathbf{u}_h^i - \mathbf{u}_h^{i-1}), \nabla \mathbf{v}_h) &+ (\nabla(p_h^i - p_h^{i-1}), \mathbf{v}_h) \qquad (3.13)\\
+\, \tilde{c}(\mathbf{u}_h^{i-1} - \mathbf{u}_h^{i-2}, \mathbf{u}_h^{i-1}, \mathbf{v}_h) &+ \tilde{c}(\mathbf{u}_h^{i-1}, \mathbf{u}_h^i - \mathbf{u}_h^{i-1}, \mathbf{v}_h) = 0, \ \forall\, \mathbf{v}_h \in V_{h,0}\\
(\nabla\cdot(\mathbf{u}_h^i - \mathbf{u}_h^{i-1}), q_h) &+ \alpha(\nabla(p_h^i - p_h^{i-1}), \nabla q_h) \qquad (3.14)\\
&- \alpha(\mathbf{w}_h^i - \mathbf{w}_h^{i-1}, \nabla q_h) = 0, \ \forall q_h \in Q_h\\
(\nabla(p_h^i - p_h^{i-1}), \mathbf{y}_h) &- (\mathbf{w}_h^i - \mathbf{w}_h^{i-1}, \mathbf{y}_h) = 0, \ \forall \mathbf{y}_h \in V_h \quad (3.15)
\end{aligned}$$

Taking $\mathbf{v}_h = \mathbf{u}_h^i - \mathbf{u}_h^{i-1}$ in 3.13, $q_h = p_h^i - p_h^{i-1}$ in 3.14, $\mathbf{y}_h = -\alpha(\mathbf{w}_h^i - \mathbf{w}_h^{i-1})$ in 3.15 and adding them up, we find:

$$\begin{aligned}
\nu||\mathbf{u}_h^i - \mathbf{u}_h^{i-1}||^2 &+ \alpha\,|\nabla(p_h^i - p_h^{i-1}) - (\mathbf{w}_h^i - \mathbf{w}_h^{i-1})|^2 \qquad (3.16)\\
&+ \tilde{c}(\mathbf{u}_h^{i-1} - \mathbf{u}_h^{i-2}, \mathbf{u}_h^{i-1}, \mathbf{u}_h^i - \mathbf{u}_h^{i-1}) = 0
\end{aligned}$$

Thus:

$$\begin{aligned}
\nu||\mathbf{u}_h^i - \mathbf{u}_h^{i-1}||^2 &\le C_{111}\, ||\mathbf{u}_h^{i-1} - \mathbf{u}_h^{i-2}||\, ||\mathbf{u}_h^{i-1}||\, ||\mathbf{u}_h^i - \mathbf{u}_h^{i-1}||\\
&\le \frac{C_\Omega\, C_{111}\, |\mathbf{f}|}{\nu}\, ||\mathbf{u}_h^{i-1} - \mathbf{u}_h^{i-2}||\, ||\mathbf{u}_h^i - \mathbf{u}_h^{i-1}||
\end{aligned}$$

or equivalently:

$$||\mathbf{u}_h^i - \mathbf{u}_h^{i-1}|| \leq \Upsilon ||\mathbf{u}_h^{i-1} - \mathbf{u}_h^{i-2}||$$

and, by induction:

$$||\mathbf{u}_h^i - \mathbf{u}_h^{i-1}|| \leq ||\mathbf{u}_h^1 - \mathbf{u}_h^0|| \, \Upsilon^{i-1} \leq C\,\Upsilon^i$$

The hypothesis $\Upsilon < 1$ ensures the convergence of $\mathbf{u}_h^i$ in the finite dimensional space $V_{h,0}$. If we now take $\mathbf{v}_h = P_{h,1}(\nabla(p_h^i - p_h^{i-1}))$ in 3.13, we get:

$$\begin{array}{rcl}
|P_{h,1}(\nabla(p_h^i - p_h^{i-1}))|^2 & = & (\nabla(p_h^i - p_h^{i-1}), P_{h,1}(\nabla(p_h^i - p_h^{i-1}))) \\
 & = & -\,\nu(\nabla(\mathbf{u}_h^i - \mathbf{u}_h^{i-1}), P_{h,1}(\nabla(p_h^i - p_h^{i-1}))) \\
 & - & \tilde{c}(\mathbf{u}_h^{i-1} - \mathbf{u}_h^{i-2}, \mathbf{u}_h^{i-1}, P_{h,1}(\nabla(p_h^i - p_h^{i-1})) \\
 & - & \tilde{c}(\mathbf{u}_h^{i-1}, \mathbf{u}_h^i - \mathbf{u}_h^{i-1}, P_{h,1}(\nabla(p_h^i - p_h^{i-1})) \\
 & \leq & \Big[\nu\,||\mathbf{u}_h^i - \mathbf{u}_h^{i-1}|| + C_{111}\,||\mathbf{u}_h^{i-1} - \mathbf{u}_h^{i-2}||\,||\mathbf{u}_h^{i-1}|| \\
 & + & C_{111}\,||\mathbf{u}_h^{i-1}||\,||\mathbf{u}_h^i - \mathbf{u}_h^{i-1}||\Big]\,||P_{h,1}(\nabla(p_h^i - p_h^{i-1}))|| \\
 & \leq & \dfrac{C}{h}\,(\Upsilon^i + \Upsilon^{i-1})\,|P_{h,1}(\nabla(p_h^i - p_h^{i-1}))|
\end{array}$$

so that:

$$|P_{h,1}(\nabla(p_h^i - p_h^{i-1}))| \leq \frac{C}{h}\Upsilon^{i-1} \tag{3.17}$$

From 3.16 we also get:

$$\begin{array}{rcl}
|P_{h,3}(\nabla(p_h^i - p_h^{i-1}))|^2 & = & \alpha\,|\nabla(p_h^i - p_h^{i-1}) - (\mathbf{w}_h^i - \mathbf{w}_h^{i-1})|^2 \\
 & \leq & C_{111}\,||\mathbf{u}_h^{i-1}||\,||\mathbf{u}_h^{i-1} - \mathbf{u}_h^{i-2}||\,||\mathbf{u}_h^i - \mathbf{u}_h^{i-1}|| \\
 & \leq & C\,\Upsilon^{2i-1}
\end{array}$$

which implies:

$$|P_{h,3}(\nabla(p_h^i - p_h^{i-1}))| \leq C\Upsilon^{i-1} \tag{3.18}$$

From 3.17, 3.18, 2.28 and 3.3, convergence of p_h^i to some p_h in Q_h is established. Finally, $\mathbf{w}_h^i = P_{h,12}(\nabla p_h^i)$ also converges to some $\mathbf{w}_h \in V_h$. Taking the limit of 3.10 when i tends to infinity, we prove that $(\mathbf{u}_h, p_h, \mathbf{w}_h)$ is a solution of 3.9. □

The proof of Proposition 3.3 shows in particular that Picard's method is a first order scheme for this problem.

We have proved, in particular, that under the inverse inequality 1.23, the compatibility condition on the interpolation 2.28 also required for the linear problem and the uniqueness condition 3.3 of the standard continuous problem, a solution of the discrete reformulated problem 3.9 exists, which can be obtained as the limit of the Picard iteration 3.10 starting from an arbitrary $\mathbf{u}_h^0$. To end this Section, we establish the uniqueness of such a solution under the same conditions:

Proposition 3.4: *assume that* 1.23, 2.28 *and* 3.3 *hold. Then, the solution of* 3.9 *is unique.*

PROOF: the proof is essentially the same as in Proposition 3.1. Any solution satisfies:

$$\tilde{D}(\mathbf{u}_h, p_h, \mathbf{w}_h; \mathbf{v}_h, q_h, \mathbf{y}_h) = (\mathbf{f}, \mathbf{v}_h), \quad \forall(\mathbf{v}_h, q_h, \mathbf{y}_h) \in (V_0 \times Q \times V)$$

and therefore:

$$||\mathbf{u}_h|| \le \frac{C_\Omega |\mathbf{f}|}{\nu}$$

If $(\mathbf{u}_{h,*}, p_{h,*}, \mathbf{w}_{h,*})$ and $(\mathbf{u}_{h,**}, p_{h,**}, \mathbf{w}_{h,**})$ are two solutions, and we call again $(\bar{\mathbf{u}}_h, \bar{p}_h, \bar{\mathbf{w}}_h)$ their difference, we find again that for all $(\mathbf{v}_h, q_h, \mathbf{y}_h) \in (V_{h,0} \times Q_h \times V_h)$:

$$\tilde{D}(\bar{\mathbf{u}}_h, \bar{p}_h, \bar{\mathbf{w}}_h; \mathbf{v}_h, q_h, \mathbf{y}_h) = -\tilde{c}(\mathbf{u}_{h,**}, \bar{\mathbf{u}}_h, \mathbf{v}_h) - \tilde{c}(\bar{\mathbf{u}}_h, \mathbf{u}_{h,*}, \mathbf{v}_h) + \tilde{c}(\bar{\mathbf{u}}_h, \bar{\mathbf{u}}_h, \mathbf{v}_h) \tag{3.19}$$

This implies:

$$\nu||\bar{\mathbf{u}}_h||^2 + \alpha|\nabla\bar{p}_h - \bar{\mathbf{w}}_h|^2 = \tilde{D}(\bar{\mathbf{u}}_h, \bar{p}_h, \bar{\mathbf{w}}_h; \bar{\mathbf{u}}_h, \bar{p}_h, \bar{\mathbf{w}}_h) = -\tilde{c}(\bar{\mathbf{u}}_h, \mathbf{u}_{h,*}, \bar{\mathbf{u}}_h)$$

and:

$$\nu||\bar{\mathbf{u}}_h||^2 \le \frac{C_{111} C_\Omega |\mathbf{f}|}{\nu} ||\bar{\mathbf{u}}_h||^2,$$

so that 3.3 implies $\bar{\mathbf{u}}_h = 0$. Taking now $\mathbf{v}_h = P_{h,1}(\nabla\bar{p}_h)$, $q_h = 0$ and $\mathbf{y}_h = 0$ in 3.19, we find $P_{h,1}(\nabla\bar{p}_h) = 0$, whereas $\mathbf{v}_h = 0$, $q_h = p_h$ and $\mathbf{y}_h = \bar{\mathbf{w}}_h$ yields $P_{h,3}(\nabla\bar{p}_h) = \nabla\bar{p}_h - \bar{\mathbf{w}}_h = 0$. Condition 2.28 then ensures that $\nabla\bar{p}_h = 0$, and finally $\bar{\mathbf{w}}_h = P_{h,12}(\nabla\bar{p}_h) = 0$ □

In summary, we have proved existence and uniqueness of a discrete solution of the reformulated Navier–Stokes problem 3.9, assuming the weak compatibility condition 2.28, the 'classical' uniqueness condition 3.3 and some regularity of the mesh: the discrete LBB condition is not required at all.

3.3 Stability and convergence of the method

We extend now the stability and convergence analysis performed in Section 2.3 for the reformulated Stokes problem to the Navier–Stokes case. The conditions needed for this purpose are, again, the stability condition 2.28, the inverse inequality 1.23, some regularity of the solution of the original problem and a certain behaviour if the coefficient α in terms of the mesh size h; for this nonlinear problem, however, we will also require the uniqueness condition 3.3.

3.3.1 Stability

Let us first establish a stability estimate, which is essentially underlying the existence results already proved:

Proposition 3.5: *assume that* 1.23, 2.28 *and* 3.3 *hold; assume also that* α *satisfies* 2.29. *Then, the solution* $(\mathbf{u}_h, p_h, \mathbf{w}_h)$ *of* 3.9 *satisfies the stability estimate:*

$$|||(\mathbf{u}_h, p_h, \mathbf{w}_h)||| \leq C\,|\mathbf{f}|, \tag{3.20}$$

for some constant C *independent of* h*, where* $|||\,.\,|||$ *is the mesh dependent norm defined in* 2.31.

PROOF: we proof that the Picard iterates $(\mathbf{u}_h^i, p_h^i, \mathbf{w}_h^i)$ satisfy 3.20, so that this will also hold for $(\mathbf{u}_h, p_h, \mathbf{w}_h)$ by passing to the limit. According to 3.11, we have that $||\mathbf{u}_h^i|| \leq \dfrac{C_\Omega|\mathbf{f}|}{\nu}$, and by 3.12 and the assumption on α, $|P_{h,3}(\nabla p_h^i)| \leq \dfrac{C}{h}\,|\mathbf{f}|$. Moreover:

$$\begin{aligned}
|P_{h,1}(\nabla p_h^i)|^2 &= (\nabla p_h^i, P_{h,1}(\nabla p_h^i)) \qquad (3.21)\\
&= (\mathbf{f}, P_{h,1}(\nabla p_h^i)) \;-\; \nu(\nabla \mathbf{u}_h^i, \nabla P_{h,1}(\nabla p_h^i))\\
&- \tilde{c}(\mathbf{u}_h^{i-1}, \mathbf{u}_h^i, P_{h,1}(\nabla p_h^i))\\
&\leq |\mathbf{f}|\,|P_{h,1}(\nabla p_h^i)| \;+\; \nu\,||\mathbf{u}_h^i||\,||P_{h,1}(\nabla p_h^i)||\\
&+ C_{111}\,||\mathbf{u}_h^{i-1}||\,||\mathbf{u}_h^i||\,||P_{h,1}(\nabla p_h^i)||\\
&\leq \left[\,|\mathbf{f}| + \nu\frac{C_\Omega|\mathbf{f}|}{\nu}\,\frac{C}{h} \;+\; C_{111}\frac{C_\Omega^2|\mathbf{f}|^2}{\nu^2}\frac{C}{h}\right]|P_{h,1}(\nabla p_h^i)|\\
&\leq C\,\frac{|\mathbf{f}|}{h}\,|P_{h,1}(\nabla p_h^i)|
\end{aligned}$$

according to 3.10, 1.23, the continuity of $\tilde{c}$ and the Schwarz inequality. This yields $|P_{h,1}(\nabla p_h^i)| \leq \dfrac{C}{h}|\mathbf{f}|$. Condition 2.28 then ensures that $|\nabla p_h^i| \leq \dfrac{C}{h}|\mathbf{f}|$ and since $|\mathbf{w}_h^i| = |P_{h,12}(\nabla p_h^i)| \leq |\nabla p_h^i|$, we prove 3.20 for $(\mathbf{u}_h^i, p_h^i, \mathbf{w}_h^i)$, and, by taking the limit when i tends to infinity, for $(\mathbf{u}_h, p_h, \mathbf{w}_h)$. □

3.3.2 Convergence in natural norms

To prove optimal convergence of the discrete solution of 3.9 to the continuous solution in natural norms, we repeat the analysis of the linear problem, only modified to account for the effects of the convective term. We will therefore focus only on these effects here.

Theorem 3.1: *assume that the Navier–Stokes problem* 3.1 *is p–regular, and that conditions* 1.23, 2.28 *and* 3.3 *hold; assume also that α satisfies* 2.36. *Then, the solution $(\mathbf{u}_h, p_h, \mathbf{w}_h)$ of* 3.9 *satisfies:*

$$|||(\mathbf{u}-\mathbf{u}_h, p-p_h, \nabla p-\mathbf{w}_h)||| \le C\,E(h) \tag{3.22}$$

for some sonstant $C > 0$ independent of h, where $(\mathbf{u}, p)$ is the solution of 3.1 *and the interpolation error function $E(h)$ was defined in* 2.38.

PROOF: for any $(\mathbf{v}_h, q_h, \mathbf{y}_h) \in V_{h,0} \times Q \times V$, we have:

$$\begin{array}{rl}
\multicolumn{2}{l}{\tilde{D}(\mathbf{u}-\mathbf{u}_h, p-p_h, \nabla p-\mathbf{w}_h; \mathbf{v}_h, q_h, \mathbf{y}_h)} \\
= & \nu(\nabla \mathbf{u}, \nabla \mathbf{v}_h) + (\nabla p, \mathbf{v}_h) + \tilde{c}(\mathbf{u}, \mathbf{u}, \mathbf{v}_h) \qquad (3.23) \\
- & \nu(\nabla \mathbf{u}_h, \nabla \mathbf{v}_h) - (\nabla p_h, \mathbf{v}_h) + \tilde{c}(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}_h) \\
- & \tilde{c}(\mathbf{u}, \mathbf{u}_h, \mathbf{v}_h) \ - \ \tilde{c}(\mathbf{u}_h, \mathbf{u}, \mathbf{v}_h) \\
= & 2\tilde{c}(\mathbf{u}_h, \mathbf{u}_h, \mathbf{v}_h) - \tilde{c}(\mathbf{u}, \mathbf{u}_h, \mathbf{v}_h) \\
- & \tilde{c}(\mathbf{u}_h, \mathbf{u}, \mathbf{v}_h) \\
= & \tilde{c}(\mathbf{u}_h - \mathbf{u}, \mathbf{u}_h, \mathbf{v}_h) + \tilde{c}(\mathbf{u}_h, \mathbf{u}_h - \mathbf{u}, \mathbf{v}_h)
\end{array}$$

This implies, by the linearity of $\tilde{D}$ in its second argument, that:

$$\begin{array}{rl}
\multicolumn{2}{l}{\tilde{D}(\mathbf{u}-\mathbf{u}_h, p-p_h, \nabla p-\mathbf{w}_h; \mathbf{u}-\mathbf{u}_h, p-p_h, \nabla p-\mathbf{w}_h)} \\
= & \tilde{D}(\mathbf{u}-\mathbf{u}_h, p-p_h, \nabla p-\mathbf{w}_h; \mathbf{u}-\mathbf{v}_h, p-q_h, \nabla p-\mathbf{y}_h) \quad (3.24) \\
+ & \tilde{c}(\mathbf{u}_h-\mathbf{u}, \mathbf{u}_h, \mathbf{v}_h-\mathbf{u}_h) \ + \ \tilde{c}(\mathbf{u}_h, \mathbf{u}_h-\mathbf{u}, \mathbf{v}_h-\mathbf{u}_h)
\end{array}$$

for any $(\mathbf{v}_h, q_h, \mathbf{y}_h) \in (V_{h,0} \times Q_h \times V_h)$. Coercivity of $\tilde{D}$ implies that:

$$\text{LHS of 3.24} = \nu||\mathbf{u}-\mathbf{u}_h||^2 + \alpha|\nabla p_h - \mathbf{w}_h|^2$$

whereas:

$$\begin{array}{rl}
\text{RHS of 3.24} = & \nu(\nabla(\mathbf{u}-\mathbf{u}_h), \nabla(\mathbf{u}-\mathbf{v}_h)) + (\nabla(p-p_h), \mathbf{u}-\mathbf{v}_h) \ (3.25) \\
+ & (\nabla(q_h - p), \mathbf{u}-\mathbf{u}_h) + \alpha(\mathbf{w}_h - \nabla p_h, \mathbf{y}_h - \nabla q_h) \\
+ & \tilde{c}(\mathbf{u}-\mathbf{u}_h, \mathbf{u}-\mathbf{u}_h, \mathbf{u}-\mathbf{v}_h) \ + \ \tilde{c}(\mathbf{u}_h-\mathbf{u}, \mathbf{u}_h, \mathbf{v}_h-\mathbf{u}_h) \\
+ & \tilde{c}(\mathbf{u}_h, \mathbf{u}_h-\mathbf{u}, \mathbf{v}_h-\mathbf{u}_h)
\end{array}$$

For the linear part, the same argument as in Section 2.3 is valid here, so that it will not be repeated. As for the quadratic terms, we have:

$$\begin{aligned}
QT &= \tilde{c}(\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{v}_h) \qquad (3.26)\\
&+ \tilde{c}(\mathbf{u}_h-\mathbf{u},\mathbf{u}_h,\mathbf{v}_h-\mathbf{u}_h) + \tilde{c}(\mathbf{u}_h,\mathbf{u}_h-\mathbf{u},\mathbf{v}_h-\mathbf{u}_h)\\
&= \tilde{c}(\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{v}_h) + \tilde{c}(\mathbf{u}-\mathbf{u}_h,\mathbf{u}_h,\mathbf{u}-\mathbf{v}_h)\\
&+ \tilde{c}(\mathbf{u}-\mathbf{u}_h,\mathbf{u}_h,\mathbf{u}_h-\mathbf{u}) + \tilde{c}(\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{v}_h)\\
&+ \tilde{c}(\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h,\mathbf{u}_h-\mathbf{u})\\
&= \tilde{c}(\mathbf{u}-\mathbf{u}_h,\mathbf{u},\mathbf{u}-\mathbf{v}_h) + \tilde{c}(\mathbf{u}-\mathbf{u}_h,\mathbf{u}_h,\mathbf{u}_h-\mathbf{u})\\
&+ \tilde{c}(\mathbf{u}_h,\mathbf{u}-\mathbf{u}_h,\mathbf{u}-\mathbf{v}_h)\\
&\leq \frac{2C_{111}C_\Omega|\mathbf{f}|}{\nu}||\mathbf{u}-\mathbf{u}_h||\,||\mathbf{u}-\mathbf{v}_h||\\
&+ \frac{C_{111}C_\Omega|\mathbf{f}|}{\nu}||\mathbf{u}-\mathbf{u}_h||^2
\end{aligned}$$

according to the stability estimates derived above for the continuous and discrete solutions. The first term of 3.26 is the product of an error of the method and an interpolation error, and it can be included in the convergence analysis of the linear problem. The second term can be passed to the left–hand–side, yielding $(\nu - \frac{C_{111}C_\Omega|\mathbf{f}|}{\nu})\,||\mathbf{u}-\mathbf{u}_h||^2$. Condition 3.3 ensures that this coefficient is positive, and the analysis of the linear problem can then be repeated. □

3.3.3 Convergence in L^2-norm

We finally prove that the error estimates derived for the discrete solution of the reformulated Navier–Stokes problem can be improved by an order in h in the norm of $L^2(\Omega)$, in a similar way to the Stokes problem. For the velocity error estimates, we will need an auxiliary problem which we study first. In a similar way to [21], we call $\mathbf{y} \in \mathbf{H}_0^1(\Omega)$ and $\chi \in L_0^2(\Omega)$ the solution of the following linear problem:

$$\begin{aligned}
\nu((\mathbf{y},\mathbf{v})) - \tilde{c}(\mathbf{u}_h,\mathbf{y},\mathbf{v}) &+ \tilde{c}(\mathbf{v},\mathbf{u},\mathbf{y})\\
- (\chi,\nabla\cdot\mathbf{v}) &= (\mathbf{u}-\mathbf{u}_h,\mathbf{v}), \quad \forall\mathbf{v}\in\mathbf{H}_0^1(\Omega) \qquad (3.27)\\
(\nabla\cdot\mathbf{y},q) &= 0, \quad \forall q\in L_0^2(\Omega)
\end{aligned}$$

Existence and uniqueness of a solution to this problem is guaranteed by the stability estimate of the continuous solution $\mathbf{u}$, the uniqueness condition 3.3 and the continuous LBB condition:

Lemma 3.1: *assume that* 3.3 *holds; then, problem* 3.27 *has a unique solution* $(\mathbf{y},\chi)$.

PROOF: taking $\mathbf{v} \in Y$, the first equation of 3.3 can be written as:

$$a^{aux}(\mathbf{y},\mathbf{v}) \doteq \nu((\mathbf{y},\mathbf{v})) - \tilde{c}(\mathbf{u}_h,\mathbf{y},\mathbf{v}) + \tilde{c}(\mathbf{v},\mathbf{u},\mathbf{y}) = (\mathbf{u}-\mathbf{u}_h,\mathbf{v}) \quad (3.28)$$

The bilinear form a^{aux} is continuous in $Y \times Y$, due to the continuity of the trilinear form $\tilde{c}$ on $(\mathbf{H}^1_0(\Omega))^3$ and the stability properties 3.7 and 3.11 of the continuous and discrete solutions $\mathbf{u}$ and $\mathbf{u}_h$, respectively. Skew–symmetry of $\tilde{c}$ implies that, for all $\mathbf{v} \in Y$:

$$\begin{aligned} a^{aux}(\mathbf{v},\mathbf{v}) &= \nu\,||\mathbf{v}||^2 - \tilde{c}(\mathbf{v},\mathbf{v},\mathbf{u}) \geq \nu\,||\mathbf{v}||^2 - C_{111}\,||\mathbf{v}||^2\,||\mathbf{u}|| \\ &\geq (\nu - \frac{C_{111}C_\Omega|\mathbf{f}|}{\nu})\,||\mathbf{v}||^2 \end{aligned}$$

and the uniqueness condition 3.3 ensures that this coefficient is possitive, so that a^{aux} is coercive on Y. The Lax–Milgram theorem establishes existence and uniqueness of a solution of 3.28 in Y, and the continuous LBB condition 1.25 that of χ in $L^2_0(\Omega)$. □

This result is now used to obtain improved error estimates for the velocity and pressure in the space $L^2(\Omega)$, assuming more regularity on the domain:

Theorem 3.2: *assume that the domain Ω is such that the Stokes problem* 2.13 *is regular, that conditions* 1.23, 2.28 *and* 3.3 *hold, and that α satisfies* 2.36. *Then, the solution $(\mathbf{u}_h, p_h, \mathbf{w}_h)$ of* 3.9 *satisfies:*

$$|\mathbf{u}-\mathbf{u}_h| + h\,|p-p_h|_{L^2_0(\Omega)} \leq C\,h\,E(h) \quad (3.29)$$

where $(\mathbf{u}, p)$ is the solution of 3.1.

PROOF: the proof is essentially the same as in that of Theorem 2.2 in Section 2.3, so that only the modifications related to the convective term will be specified. The regularity now assumed on Ω implies that the estimates 2.64 hold for the solution $(\mathbf{y},\chi)$ of the auxiliary problem 3.27. Let now $\mathbf{y}_h$ and χ_h be optimal order approximations satisfying 2.65 and T_i (i=1,2,3) the three terms into which $|\mathbf{u}-\mathbf{u}_h|^2$ can be split (see the proof of Theorem 2.2); we only need to account for T_2, which is now:

$$\begin{aligned} T_2 &= \nu(\nabla \mathbf{y}_h, \nabla(\mathbf{u}-\mathbf{u}_h)) - \tilde{c}(\mathbf{u}_h,\mathbf{y},\mathbf{u}-\mathbf{u}_h) + \tilde{c}(\mathbf{u}-\mathbf{u}_h,\mathbf{u},\mathbf{y}) \\ &= -(\nabla(p-p_h),\mathbf{y}_h) + \tilde{c}(\mathbf{u}_h,\mathbf{u}_h,\mathbf{y}_h) - \tilde{c}(\mathbf{u},\mathbf{u},\mathbf{y}_h) \\ &- \tilde{c}(\mathbf{u}_h,\mathbf{y},\mathbf{u}-\mathbf{u}_h) + \tilde{c}(\mathbf{u}-\mathbf{u}_h,\mathbf{u},\mathbf{y}) \\ &= (\nabla(p-p_h),\mathbf{y}-\mathbf{y}_h) + \tilde{c}(\mathbf{u}_h,\mathbf{u}_h-\mathbf{u},\mathbf{y}_h) - \tilde{c}(\mathbf{u}_h,\mathbf{u}_h-\mathbf{u},\mathbf{y}) \\ &+ \tilde{c}(\mathbf{u}_h-\mathbf{u},\mathbf{u},\mathbf{y}_h) - \tilde{c}(\mathbf{u}_h-\mathbf{u},\mathbf{u},\mathbf{y}) \end{aligned}$$

$$\begin{aligned}
&= (\nabla(p-p_h),\mathbf{y}-\mathbf{y}_h) + \tilde{c}(\mathbf{u}_h,\mathbf{u}_h-\mathbf{u},\mathbf{y}_h-\mathbf{y})\\
&+ \tilde{c}(\mathbf{u}_h-\mathbf{u},\mathbf{u},\mathbf{y}_h-\mathbf{y})\\
&\le |\nabla(p-p_h)|\,|\mathbf{y}-\mathbf{y}_h| + C_{111}\,||\mathbf{u}_h-\mathbf{u}||\,(||\mathbf{u}_h||+||\mathbf{u}||)\,||\mathbf{y}-\mathbf{y}_h||\\
&\le C\,h^2\,||\mathbf{y}||_2\,|\nabla(p-p_h)| + C\,h\,||\mathbf{u}_h-\mathbf{u}||\,||\mathbf{y}||_2\\
&\le C\,h\,(h\,|\nabla(p-p_h)| + ||(\mathbf{u}-\mathbf{u}_h)||)\,|(\mathbf{u}-\mathbf{u}_h)|
\end{aligned}$$

and the error estimate for the velocity is established. As for the pressure, we call again $\mathbf{z}$ and ξ the solution of the Stokes problem 2.66 and $\mathbf{z}_h$ an approximation of $\mathbf{z}$ satisfying 2.68; we now have:

$$\begin{aligned}
|p-p_h|^2 &= (p-p_h,p-p_h) = (\nabla\cdot\mathbf{z},p-p_h)\\
&= (\nabla\cdot(\mathbf{z}-\mathbf{z}_h),p-p_h) - (\mathbf{z}_h,\nabla(p-p_h))\\
&= -(\mathbf{z}-\mathbf{z}_h,\nabla(p-p_h)) + \nu(\nabla(\mathbf{u}-\mathbf{u}_h),\nabla\mathbf{z}_h)\\
&+ \tilde{c}(\mathbf{u},\mathbf{u},\mathbf{z}_h) - \tilde{c}(\mathbf{u}_h,\mathbf{u}_h,\mathbf{z}_h)\\
&= -(\mathbf{z}-\mathbf{z}_h,\nabla(p-p_h)) + \nu(\nabla(\mathbf{u}-\mathbf{u}_h),\nabla(\mathbf{z}_h-\mathbf{z}))\\
&+ (\nabla(\mathbf{u}-\mathbf{u}_h),\nabla\mathbf{z}) + \tilde{c}(\mathbf{u},\mathbf{u}-\mathbf{u}_h,\mathbf{z}_h)\\
&+ \tilde{c}(\mathbf{u}-\mathbf{u}_h,\mathbf{u}_h,\mathbf{z}_h)\\
&= -(\mathbf{z}-\mathbf{z}_h,\nabla(p-p_h)) + \nu(\nabla(\mathbf{u}-\mathbf{u}_h),\nabla(\mathbf{z}_h-\mathbf{z}))\\
&+ (\nabla(\mathbf{u}-\mathbf{u}_h),\nabla\mathbf{z}) + \tilde{c}(\mathbf{u},\mathbf{u}-\mathbf{u}_h,\mathbf{z}_h-\mathbf{z})\\
&+ \tilde{c}(\mathbf{u},\mathbf{u}-\mathbf{u}_h,\mathbf{z}) + \tilde{c}(\mathbf{u}-\mathbf{u}_h,\mathbf{u}_h,\mathbf{z}_h-\mathbf{z})\\
&+ \tilde{c}(\mathbf{u}-\mathbf{u}_h,\mathbf{u}_h,\mathbf{z})\\
&\le |\mathbf{z}-\mathbf{z}_h|\,|\nabla(p-p_h)| + \nu\,||\mathbf{u}-\mathbf{u}_h||\,(||\mathbf{z}-\mathbf{z}_h|| + ||\mathbf{z}||)\\
&+ C_{111}\,\frac{C_\Omega|\mathbf{f}|}{\nu}\,|\mathbf{u}-\mathbf{u}_h|\,(2\,||\mathbf{z}-\mathbf{z}_h|| + 2\,||\mathbf{z}||)\\
&\le \left[C\,h\,|\nabla(p-p_h)| + C\,||\mathbf{u}-\mathbf{u}_h||\right]\,||\mathbf{z}||\\
&\le C\,(\,h\,|\nabla(p-p_h)| + ||\mathbf{u}-\mathbf{u}_h||\,)\,|p-p_h|
\end{aligned}$$

and the estimate for the pressure is established. □

3.4 Computational aspects

The discrete nonlinear equation system 3.9 can be solved numerically in different ways. We studied and compared several possibilities for its solution, all of which take the form of iterative methods.

With the notation introduced up to now, the equation system 3.9 relative to a finite element discretization of the domain Ω can be written as:

$$\begin{aligned}
A(U)U + KU + G_0P &= F\\
-G_0^tU + \alpha LP - \alpha G^tW &= 0\\
GP - MW &= 0
\end{aligned} \tag{3.30}$$

This equation system is a nonlinear problem for the variables (U, P, W); standard nonlinear solvers may be applied to it, but, once more, it seemed appropriate to take advantage of the particular structure of the system, so as to reduce its size and storage requirements. We decided to develop iterative solvers based on the *coupled block–Gauss–Seidel* method for the linear problem introduced in 2.5.2 (the *uncoupled block–Gauss–Seidel* scheme of 2.5.1 was impractical even for the linear case). We will first present the different alternatives considered, which employ either a Picard or a Newton–Raphson method for the nonlinearity, and then show their performance on a standard test problem. We concentrate here again on the performance of each of the different schemes, putting off to the next Section the analysis of the results actually achieved, which, anyway, were almost the same for all the methods.

All the schemes are presented in the form of an iterative process for the variables U, P and W. The same initializations as in the linear case are assumed.

In the implementation on the computer, we have adopted the standard formulation of the convective term $(\mathbf{u} \cdot \nabla)\mathbf{u}$, as in equation 3.1, although the analysis of the reformulated method for the Navier–Stokes problem was based on a skew-symmetric formulation. Nevertheless, the discrete velocity field in this method is 'nearly' incompressible, since, according to 3.9, we have that $\nabla \cdot \mathbf{u}_h = O(\alpha|\nabla p_h - \mathbf{w}_h|)$. This quantity will be small, given that $\alpha = O(h^2)$ and that both ∇p_h and $\mathbf{w}_h$ approximate ∇p to optimal order in h. The difference between the two formulations of the convective term in this method is negligible.

3.4.1 Nonlinear iterative solvers

We considered several possibilities for the numerical solution of the nonlinear system of equations 3.30, which we explain in what follows.

- Coupled formulation–nonlinearity scheme.

  We considered the possibility of coupling the iterations required to solve the linear problem with the coupled block–Gauss–Seidel scheme, which we say are due to the *formulation*, with those of a nonlinear solver, either by an explicit, Picard's or Newton–Raphson's method.

  The first and simplest alternative consists of evaluating the nonlinear term $A(U)U$ at the previous iteration values, in an explicit way. Thus, a straightforward extension of the coupled block–Gauss–Seidel scheme to the nonlinear case reads:

$$\begin{array}{rcrcl} KU^i & + & G_0P^i & = & F - A(U^{i-1})U^{i-1} \\ -G_0^tU^i & + & \alpha LP^i & = & \alpha G^tW^{i-1} \\ & & MW^i & = & GP^i \end{array}$$

The obvious advantage of this scheme is that the system matrix of the linear problems to be solved at each iteration is the same for all the iterations; it can thus be computed and factorized only once at the beginning of the calculations, if direct methods are to be used for these linear problems. The computational cost of this method is, in principle, similar to that of the linear problem, but for the evaluation of the convective residue $A(U^{i-1})U^{i-1}$ at each iteration and if a similar number of iterations were required for convergence.

The explicit approximation of the convective term is a zero–th order method. Its main disadvantage, however, is that it is highly.unstable, even for relatively low Reynolds number flows; in our computations, we could not go beyond a value of $Re = 10$ with this scheme. This makes this alternative not usable in practical situations.

If a Picard's approximation is used for the nonlinear term, the scheme reads:

$$\begin{array}{rcl} A(U^{i-1})U^i \; + \; KU^i \; + \; G_0P^i & = & F \\ -G_0^tU^i \; + \; \alpha LP^i & = & \alpha G^tW^{i-1} \\ MW^i & = & GP^i \end{array}$$

whereas for Newton–Raphson, it is:

$$\begin{array}{rcl} A(U^{i-1})U^i + A(U^i)U^{i-1} + KU^i + G_0P^i & = & F + A(U^{i-1})U^{i-1} \\ -G_0^tU^i \; + \; \alpha LP^i & = & \alpha G^tW^{i-1} \\ MW^i & = & GP^i \end{array}$$

We expected this coupled schemes to combine the stability properties of the nonlinear solvers with the convergence properties of the method for the linear problem, maybe at the expense of a few extra iterations with respect to the linear problem. However, the matrix for the linear systems for velocity and pressure to be solved at each iteration is not constant any more, and needs being computed and factorized at each iteration. The mass matrix for the pressure gradient equations may, again, be considered consistent or lumped. The convergence criterion for this method is the same as for the linear problem, that is, 2.102.

This method produced acceptable convergence results, but still not comparable to standard nonlinear solvers (see 3.4.2). Other sources of trouble were that the method was unstable for large values of the Reynolds number and that succesive over (or under) relaxation proved inadequate in this case.

- Nested formulation–nonlinearity scheme.

  We considered a second possibility consisting of a pair of nested loops, an *outer* iterative process for the formulation, similar to the coupled block–Gauss–Seidel scheme of the linear problem but with a fully implicit approximation of the nonlinear term, and an *inner* iteration loop to solve this nonlinearity at each of the outer iterations. That is, at the i–th iteration level of the outer scheme, and with the following initializations for the nonlinear solver: $U^{i,0} = U^{i-1}$, $P^{i,0} = P^{i-1}$ and $W^{i,0} = W^{i-1}$, the nested formulation–nonlinearity scheme with a Picard's method for the nonlinearity is:

$$\begin{array}{rcl} A(U^{i,j-1})U^{i,j} + KU^{i,j} + G_0P^{i,j} & = & F \\ -G_0^tU^{i,j} + \alpha LP^{i,j} & = & \alpha G^tW^{i-1} \end{array}$$

  This scheme is iterated in j until the convergence criterion:

$$\max(\frac{|U^{i,j} - U^{i,j-1}|_2}{|U^{i,j}|_2}, \frac{|P^{i,j} - P^{i,j-1}|_2}{|P^{i,j}|_2}) < \epsilon_{\text{nl}} \qquad (3.31)$$

  holds. In that case, we set U^i and P^i equal to the last iteration values, and update W^{i-1} as $MW^i = GP^i$. For Newton–Raphson's case, this scheme becomes:

$$\begin{array}{rcl} A(U^{i,j-1})U^{i,j} + A(U^{i,j})U^{i,j-1} + KU^{i,j} + G_0P^{i,j} & = & F \\ & + & A(U^{i,j-1})U^{i,j-1} \\ -G_0^tU^{i,j} + \alpha LP^{i,j} & = & \alpha G^tW^{i-1} \end{array}$$

  In either case, the outer iteration loop stops when condition 2.102 is satisfied.

  These methods have the disadvantage that at each of the outer iterations, a number of linear systems has to be solved with different system matrices, which have to be computed and factorized every time. Nevertheless, we hoped that as the outer iteration scheme proceeds, the number of inner iterations needed to solve the nonlinearity would decrease, since the initial values for the nonlinear solver progressively approach the solution. This fact could, in principle, make these methods competitive with the previous ones, but that was not the case.

- Nested nonlinearity–formulation scheme.

  Finally, we developed a method in which the inner and outer iterations loops of the previous schemes are performed in reversed order. That is, an outer iterative method is considered to solve the nonlinearity of the full problem, within which an inner iteration scheme of the coupled block–Gauss–Seidel type is used for the formulation of the method. In Picard's case, setting $U^{0,j} = U^{j-1}$, $P^{0,j} = P^{j-1}$ and $W^{0,j} = W^{j-1}$ at the j–th outer iteration level, this means:

$$\begin{array}{rcl} A(U^{j-1})U^{i,j} + KU^{i,j} + G_0P^{i,j} & = & F \\ -G_0^tU^{i,j} + \alpha LP^{i,j} & = & \alpha G^tW^{i-1,j} \\ MW^{i,j} & = & GP^{i,j} \end{array}$$

This scheme is iterated in i until:

$$\mathrm{Err}(U^{i,j}, P^{i,j}, W^{i,j}; U^{i-1,j}, P^{i-1,j}, W^{i-1,j}) < \epsilon_{\mathrm{cou}},$$

where the error function Err was defined in 2.102; the final values of $U^{i,j}$, $P^{i,j}$ and $W^{i,j}$ are then taken as U^j, P^j and U^j, respectively. The outer iteration (in j) proceeds until:

$$\mathrm{Err}(U^j, P^j, W^j; U^{j-1}, P^{j-1}, W^{j-1}) < \epsilon_{\mathrm{nl}}$$

Convergence of this Picard's iteration in j was proved theoretically in Section 3.2.

In Newton–Raphson's case, this becomes:

$$\begin{array}{rcl} A(U^{j-1})U^{i,j} + A(U^{i,j})U^{j-1} + KU^{i,j} + G_0P^{i,j} & = & F + A(U^{j-1})U^{j-1} \\ -G_0^tU^{i,j} + \alpha LP^{i,j} & = & \alpha G^tW^{i-1,j} \\ MW^{i,j} & = & GP^{i,j} \end{array}$$

These methods proved to be superior to any of the schemes previously considered. At each of the outer iterations, a single system matrix needs being computed, which is the same for all the inner iterations (since it does not depend on $(i-1)$). Moreover, the number of inner iterations required to solve the linear formulation decreases as the outer iteration scheme advances, becoming very small in the last stages, as the initial approximation approaches the solution.

In the next Subsection we present a study of the computational performance of all these different iterative schemes on a test problem. In particular, we compare the convergence rates of these methods for the solution of the reformulated Navier–Stokes equations 3.30 among themselves and with the GLS method, which is again one of the most widely used methods for the Navier–Stokes equations existing nowadays; besides, this scheme, like ours, is also formulated in terms of primitive variables and allows the use of equal order interpolations.

The extension of the GLS formulation of the Stokes problem to advective–diffusive equations first and then to a linearized form of the incompressible Navier–Stokes equations was studied by L.P. Franca *et al.* in [35] and [36], and convergence analysis for these methods were also given in [37]. For the full nonlinear, incompressible Navier–Stokes equations, the GLS method can be written as:

$$\begin{aligned}
((\mathbf{u}_h \cdot \nabla)\mathbf{u}_h, \mathbf{v}_h) \;&+\; \nu(\nabla \mathbf{u}_h, \nabla \mathbf{v}_h) \;+\; (\nabla p_h, \mathbf{v}_h) \\
+ \sum_{K \in \theta_h} \alpha_K \;&\Big((\mathbf{u}_h \cdot \nabla)\mathbf{u}_h - \nu \Delta \mathbf{u}_h + \nabla p_h - \mathbf{f}, (\mathbf{u}_h \cdot \nabla)\mathbf{v}_h - \nu \Delta \mathbf{v}_h \Big)_K \\
&= \; (\mathbf{f}, \mathbf{v}_h), \qquad\qquad \forall \mathbf{v}_h \in V_{h,0} \\
(\nabla \cdot \mathbf{u}_h, q_h) \;&+\; \sum_{K \in \theta_h} \alpha_K \Big((\mathbf{u}_h \cdot \nabla)\mathbf{u}_h - \nu \Delta \mathbf{u}_h + \nabla p_h - \mathbf{f}, \nabla q_h \Big) \\
&= \; 0, \qquad\qquad \forall q_h \in Q_h
\end{aligned}$$

(once again, we have omitted term $\tau_2(\nabla \cdot \mathbf{u}_h, \nabla \cdot \mathbf{v}_h))$ present in the original formulation of the method, see [36]). The local parameters α_K are taken here again as $\alpha_K = \alpha_0 \frac{h_K^2}{4\nu}$, where $\alpha_0 = 1/3$ for the P_1 and Q_1 elements.

Notice that this scheme introduces additional nonlinearities in the problem, such as a cubic one or the dependence of the coefficients α_K on $\mathbf{u}_h$. Quadratic convergence rates of Newton–Raphson's scheme are not preserved for this problem, so that a Picard approximation is used for all convective terms in these equations. The proof of convergence of this method to the continuous solution is still an open problem.

3.4.2 Performance of the iterative schemes

As a test problem to evaluate the convergence rates of the different schemes just presented, we considered again the cavity flow problem, this time for the Navier–Stokes equations 3.1. This flow problem was solved at two different Reynolds numbers, Re= 400 and Re= 1000.

We used the nonuniform 39 × 39 mesh of Figure 2.4, with a $P1$ element interpolation. We employed local values of the parameter α. A consistent mass matrix was taken for the pressure gradient in all cases (mass lumping did not affect the convergence of the nonlinearity). These data were selected because they displayed the best results in the linear case, so as to concentrate here on the performance of the nonlinear solvers.

We present the convergence results for the different schemes in Table 3.1 for Re= 400 and in Table 3.2 for Re= 1000 (the explicit scheme diverged in both cases). The methods are coded according to the following notation:

- CFNP: coupled formulation–nonlinearity Picard method

- CFNNR: coupled formulation–nonlinearity Newton–Raphson method
- NFNP: nested formulation–nonlinearity Picard method
- NFNNR: nested formulation–nonlinearity Newton–Raphson method
- NNFP: nested nonlinearity–formulation Picard method
- NNFNR: nested nonlinearity–formulation Newton–Raphson method

We include the performance rates for the GLS method for comparison. We show the number of outer and, when applicable, inner iterations needed for convergence in each case for the following values of the different tolerances:

- CFN: $\epsilon_{\text{cou}} = 10^{-3}$.
- NFN: $\epsilon_{\text{cou}} = 10^{-3}$, $\epsilon_{\text{nl}} = 10^{-4}$.
- NNF: $\epsilon_{\text{cou}} = 10^{-3}$, $\epsilon_{\text{nl}} = 10^{-3}$, for Re= 400.
- NNF: $\epsilon_{\text{cou}} = 10^{-4}$, $\epsilon_{\text{nl}} = 10^{-3}$, for Re= 1000
- GLS: $\epsilon_{\text{nl}} = 10^{-3}$.

which were chosen so that the precision of the solution of the inner iteration loop did not influence the convergence of the outer loop, which was always of 0.1%, and so as to converge fastest in each case (in this sense, in the NNFP case for Re=1000 a relaxation value of 1.3 was used). We also show the total computing time, as a percentage of that of the NNFP scheme. The cases that are not indicated in this Tables diverged.

It can be observed that both in CFN and NFN based schemes the convergence is dominated by the formulation, and rather slow; NFN schemes are specially costly because of the need to form and factorize the system matrix at each of the inner iterations. On the contrary, in NNF schemes the overall convergence is dominated by the nonlinearity, and the convergence of the inner iteration loops becomes faster as the outer loop proceeds; in particular, Newton–Raphson's scheme requires a few iterations less than Picard's, but it is globally more expensive due to the evaluation of some extra terms, and becomes unstable at Re=1000 (we tried starting it after as many as 5 Picard's iterations and it still diverged). Finally, our NNFP scheme proved to be a little more costly than the GLS method, given that it needed one more iteration in both cases and that it has to deal with the inner iteration loop for the coupling with the pressure gradient variable.

Scheme	Outer Iterations	Inner Iterations	Total CPU time
CFNP	17	-	200
CFNNR	17	-	285
NFNP	17	14,14,13,10,10,10 9,6,6,6,6,6 5,5,2,2,2	1461
NNFP	8	16,11,7,5,4,2,2,1	100
NNFNR	6	16,11,5,10,10,1	108
GLS	7	-	85

Table 3.1: Convergence of the nonlinear solvers, Re=400.

Scheme	Outer Iterations	Inner Iterations	Total CPU time
CFNP	67	-	563
NNFP	11	9,15,17,17,13,12 10,8,7,6,5	100
GLS	10	-	82

Table 3.2: Convergence of the nonlinear solvers, Re=1000.

3.4.3 Summary of computational aspects

We have developed several methods for the solution of the nonlinear discrete problem 3.30, all of which take the form of iterative schemes with either a Picard or a Newton–Raphson approximation of the nonlinearity. After several tests on a benchmark problem for two different values of the Reynolds number, it turns out that the most efficient scheme is a system of two nested loops, the outer one being a Picard iteration for the nonlinearity and the inner one a variant of the coupled block–Gauss–Seidel scheme of the linear problem with an additional advective term. This method remained stable and convergent even for moderately convective problems, when Newton–Raphson based schemes diverged due to poor initializations.

A comparison of the performance of this scheme with the GLS method for the same test case indicates that the latter needs about 20% less computational time than the former, in average. However, the reformulated method retains the quadratic convergence of Newton–Raphson's iteration, which, although unstable for high Reynolds numbers, can be very efficient in transient problems. Moreover, we will show in the examples of the next Section that our scheme produces more accurate results than the GLS method on the same mesh. Besides. in one of the numerical examples of the next Section we present a more detailed comparison of the numerical performance of our Nested Nonlinearity–Formulation Picard method with the GLS method.

3.5 Numerical results

We present here some of the results obtained with the reformulated method for the Navier–Stokes equations on three test problems: the cavity flow problem considered in the convergence analysis of the previous Section, a problem with an analytical solution and a problem of Poiseuille flow through a junction of pipes. The second case was intended again to achieving the optimal orders of accuracy in the mesh size proved theoretically for the different variables, norms and element types, while in the third one we give a detailed comparison of the performance of our method with the GLS method.

3.5.1 Cavity flow problem

The results obtained in the convergence analysis of the previous Section for the lid–driven cavity flow problem were identical for the different solution methods. We present the results of the nested nonlinearity–formulation scheme with a Picard approximation of the nonlinearity, for Re=400 (in Figures 3.1 and 3.2) and Re=1000 (in Figures 3.3 and 3.4), in the form of streamlines and pressure contours. Secondary bottom left and right subvortices, commonly found for these values of the Reynolds number, can be observed, whereas no top left vortex is present. These results compare well with benchmark solutions for this problem, such as those of [42] and [88],

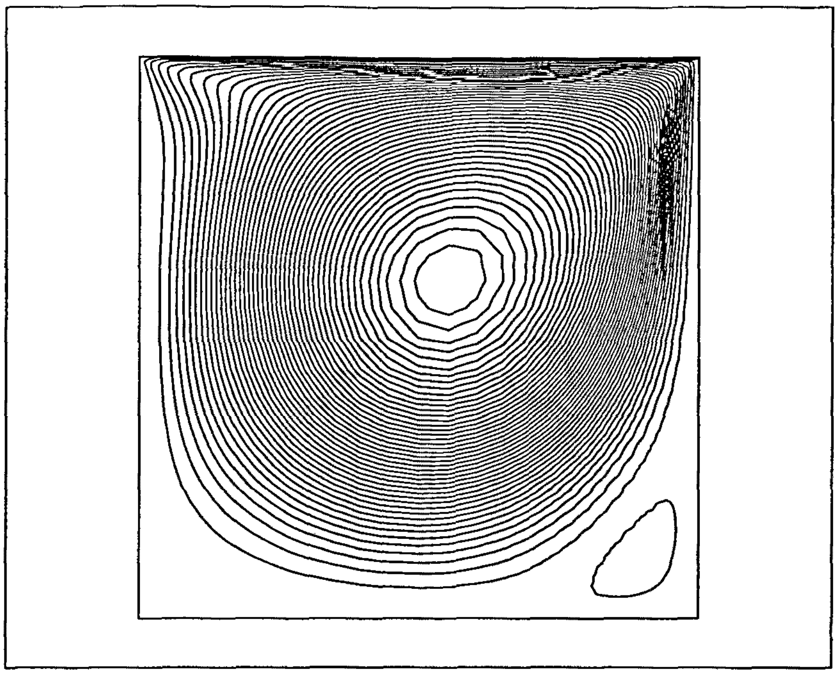


Figure 3.1: Cavity flow, Re=400, streamlines.

and other published solutions such as those of [6], [96] or [99], for Re=400, and [30],[65], [96] or [99], for Re=1000. The pressure results also compare well with those present in these references.

3.5.2 Kovasznay flow

In order to check numerically the optimal orders of accuracy proved theoretically for the different variables and norms in Section 3.3, we considered a problem introduced by Kovasznay (see [69]), modelling laminar flow behind a two dimensional grid, in which an analytical solution of the steady incompressible Navier–Stokes equations with no forcing term is available. The velocity solution $\mathbf{u} = (u, v)$ is given by:

$$\begin{aligned} u(x,y) &= 1 - e^{\lambda x}\cos(2\pi y) \\ v(x,y) &= \frac{\lambda}{2\pi}e^{\lambda x}\sin(2\pi y) \end{aligned} \tag{3.32}$$

for $(x,y) \in \mathbb{R}^2$, whereas the pressure is:

$$p(x,y) = p_0 - \frac{1}{2}e^{\lambda x} \tag{3.33}$$

where p_0 is an arbitrary constant and the parameter λ is given in terms of the Reynolds Re number by:

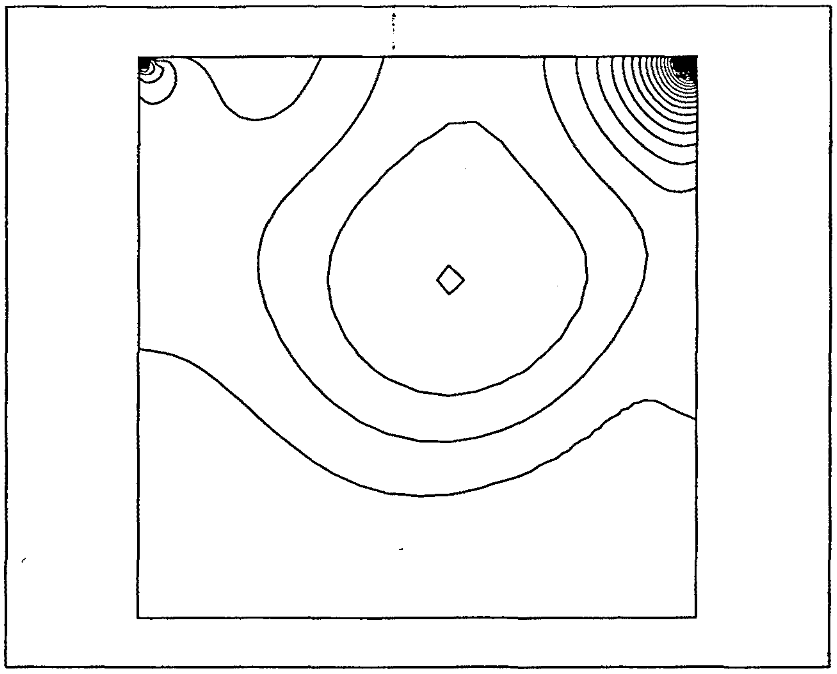


Figure 3.2: Cavity flow, Re=400, pressure contours.

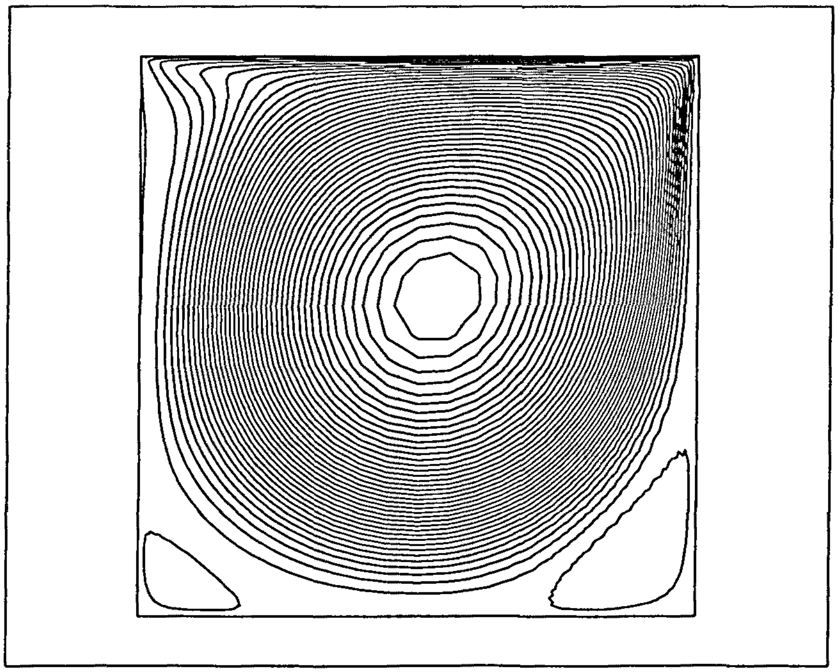


Figure 3.3: Cavity flow, Re=1000, streamlines.

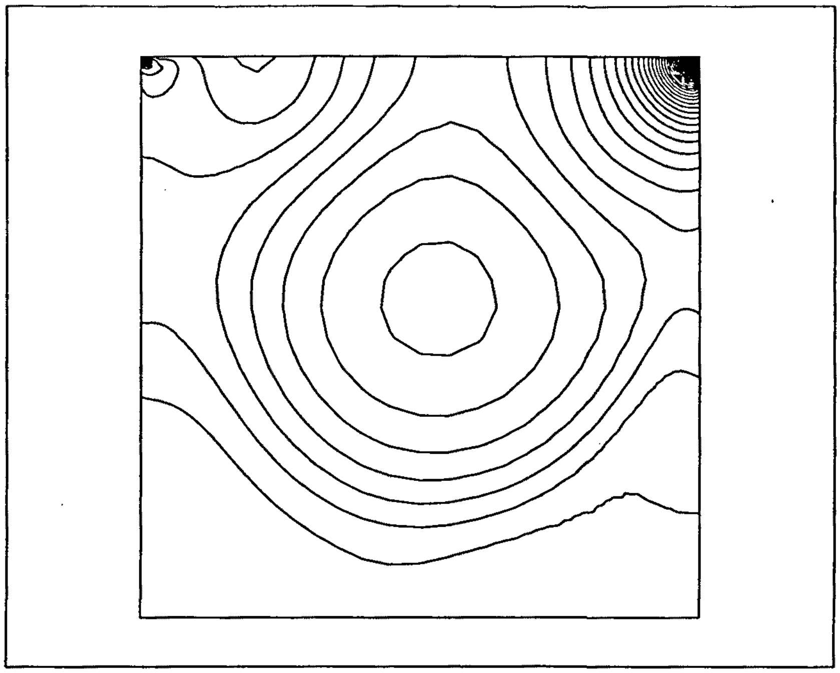


Figure 3.4: Cavity flow, Re=1000, pressure contours.

$$\lambda = \frac{\text{Re}}{2} - (\frac{\text{Re}^2}{4} + 4\pi^2)^{1/2} < 0$$

This flow problem was solved numerically in [93] and [63] for a value of Re= 40. We solved it in the domain $\Omega = [-\frac{1}{2}, 1] \times [-\frac{1}{2}, \frac{1}{2}]$ for that value of the Reynolds number (that is, for $\nu = 0.025$), with the four elements considered and on four different uniform meshes, made up with 19 × 13, 31 × 21, 43 × 29 and 61 × 41 nodes, respectively. In all cases, the solution was obtained by the nested nonlinearity-formulation scheme with a Newton–Raphson approximation of the convective term and a consistent mass matrix for the pressure gradient system, starting from the fluid at rest, but for the prescribed boundary conditions (which were given by the value of the analytical solution at the boundary). The tolerance for convergence in the iteration for nonlinearity was $\epsilon_{\text{nl}} = 10^{-4}$, and the same value was taken for the tolerance of the inner formulation iteration. It took 5 iterations of Newton–Raphson's method in all cases to find the solution. The only exception was the P_2 element case with the finer 61×41 mesh: Newton–Raphson's method diverged in that case; we used Picard's iteration instead, which required 9 iterations to find the solution for the same values of the tolerances. The number of inner iterations decreased with the outer iteration scheme in all cases, from about a hundred in the first iteration, in the worst cases, to one in the fifth iteration in all cases. No relaxation was used for this problem.

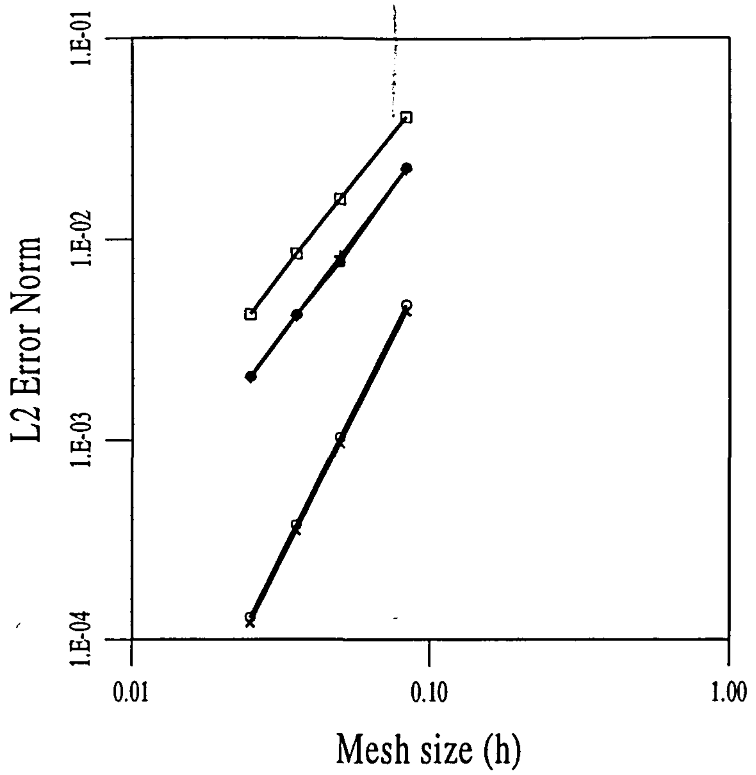


Figure 3.5: Kovasznay flow, velocity error in L^2: + P_1 Element; • Q_1 Element; o P_2 Element; × Q_2 Element; □ = GLS method, Q_1Q_1 element.

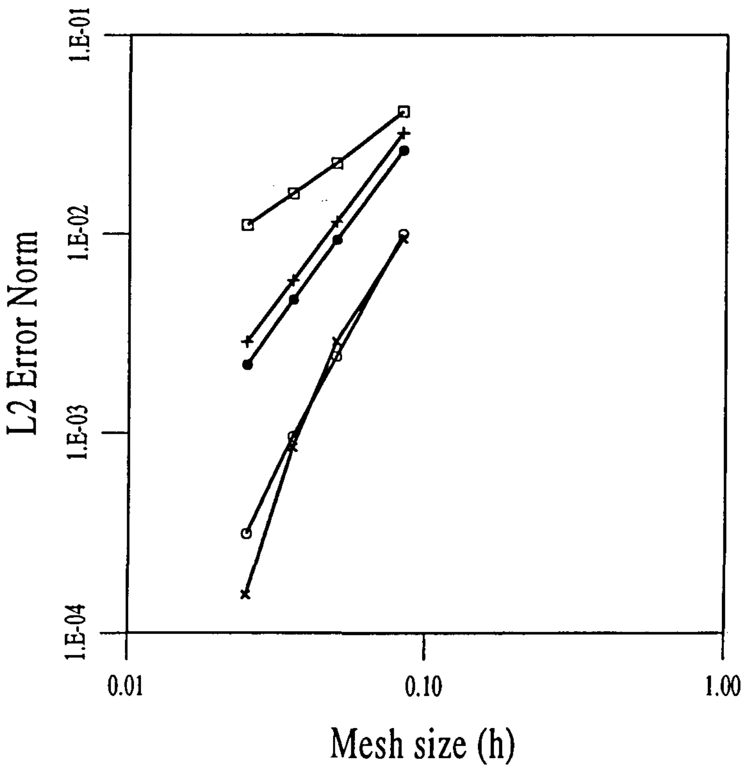


Figure 3.6: Kovasznay flow, pressure error in L^2: + P_1 Element; • Q_1 Element; o P_2 Element; × Q_2 Element; □ = GLS method, Q_1Q_1 element.

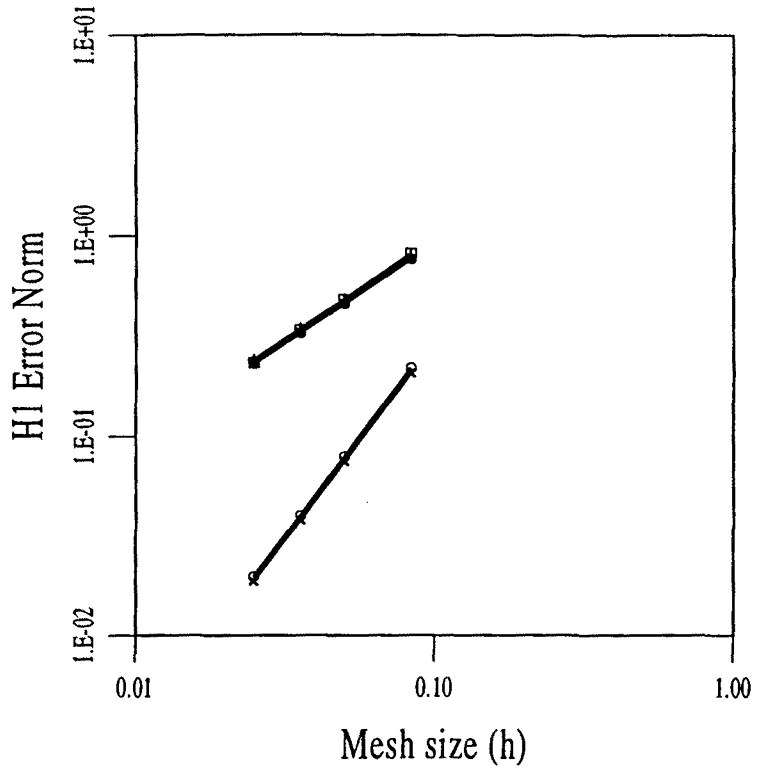


Figure 3.7: Kovasznay flow, velocity error in H^1: + P_1 Element; • Q_1 Element; o P_2 Element; × Q_2 Element; □ = GLS method, Q_1Q_1 element.

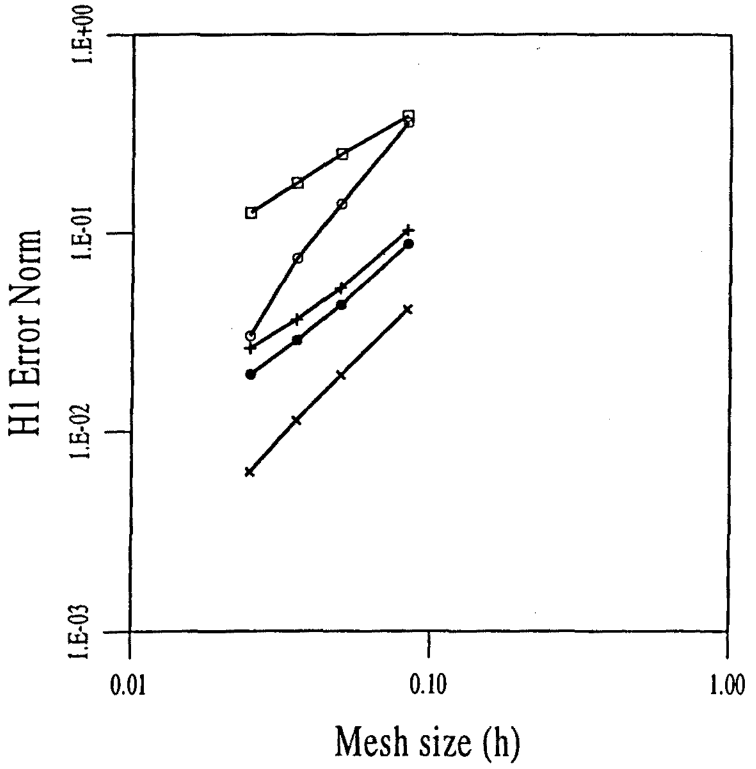


Figure 3.8: Kovasznay flow, pressure error in H^1: + P_1 Element; • Q_1 Element; o P_2 Element; × Q_2 Element; □ = GLS method, Q_1Q_1 element.

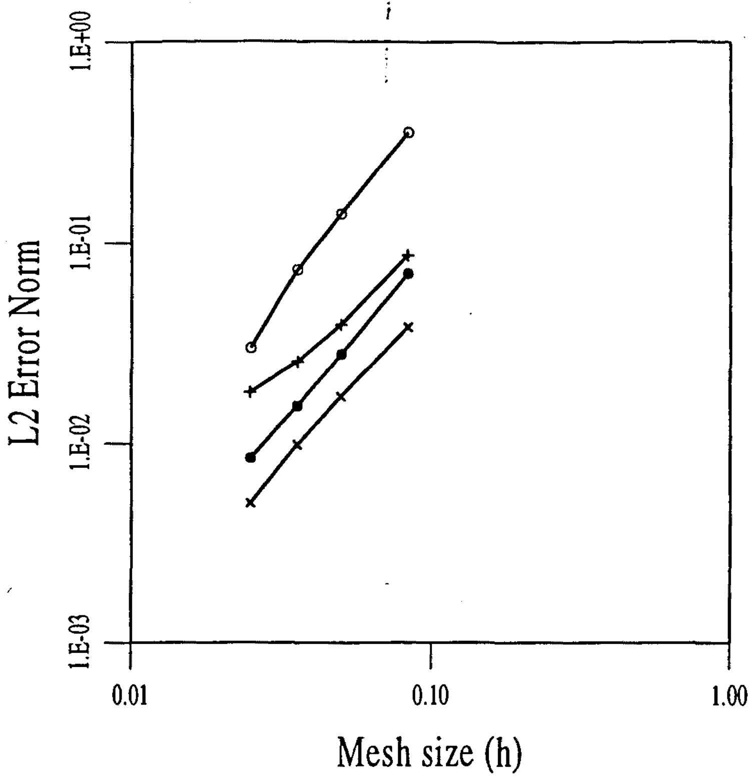


Figure 3.9: Kovasznay flow, pressure gradient error in L^2: + P_1 Element; • Q_1 Element; o P_2 Element; × Q_2 Element.

We then computed the exact errors $|\mathbf{u}-\mathbf{u}_h|$, $||\mathbf{u}-\mathbf{u}_h||$, $|p-p_h|$, $|\nabla p-\nabla p_h|$ and $|\nabla p - \mathbf{w}_h|$ for each mesh and element. Since this is a confined flow problem, we fixed the value of the pressure at the last node to zero, which always corresponded to the top right corner (1.,0.5). We then had to take $p_0 = \frac{1}{2}e^{\lambda}$ in 3.33 so that the analytical solution also satisfied this linear restriction.

The results obtained can be seen in Figures 3.5 to 3.9 as a function of the mesh size, where we have included the errors obtained for the GLS method with a Q_1Q_1 element interpolation for comparison. The linear regression coefficients computed for these lines are shown in Table 3.3. As can be observed, the theoretical orders of accuracy are found in all cases (they are again those of Table 2.5); those of the velocity solution are specially sharp, whereas for the pressure there seems to be a gain of one order of accuracy both in L^2 and H^1. In particular, the convergence of the pressure gradient was not ensured by the theory for linear and bilinear elements, but nevertheless they display at least first order convergence for this variable. We conjecture that this improvements in the accuracy of the pressure solution are a consequence of some superconvergence phenomenon, which is probably due to the extreme regularity of the meshes used (they are all made up of uniform square elements).

Moreover, it is clearly seen that the GLS method, although optimal in

Element	$\lvert \mathbf{u}-\mathbf{u}_h \rvert$	$\lvert p-p_h \rvert$	$\lVert \mathbf{u}-\mathbf{u}_h \rVert$	$\lvert \nabla p-\nabla p_h \rvert$	$\lvert \nabla p-\mathbf{w}_h \rvert$
P_1	2.0	2.0	1.0	1.1	1.3
Q_1	2.0	2.1	1.0	1.2	1.8
P_2	3.0	2.9	2.0	2.0	2.0
Q_2	3.0	3.4	2.0	1.6	1.7

Table 3.3: Kovasznay flow: linear regression coefficients for different errors.

all cases, produces less accurate results than our method, specially for the pressure. Once again, the Q_2 element provides the most accurate results.

We present the numerical solution obtained with the P_1 element on the 31 × 21 mesh in Figures 3.10 and 3.11 for the velocity and the pressure, respectively. All the solutions we computed were almost indistinguishable from one another from a graphic point of view.

3.5.3 Poiseuille flow through a junction of pipes

We finally considered a test problem introduced by J.G. Heywood *et al.* in [56], which consists of a fully developed channel flow in a pipe which bifurcates into two. This problem was considered in [56] to study the effect of the truncation of an unbounded domain and the introduction of artificial boundaries; they were specially concerned about the effect of 'natural' boundary conditions in outflow boundaries, associated to different formulations of the Navier–Stokes equations. Beside this issue of artificial boundary conditions, here we use this problem as a numerical check of the performance of our method with respect to the GLS formulation.

The geometry and mesh used for this problem can be seen in Figure 3.12; we used bilinear quadrilateral elements. The mesh consists of 2076 nodes and 1950 elements. A Poiseuille inflow was prescribed upstream, the no–slip condition was enforced on the channel walls and natural conditions were applied weakly at the two outlets. We also set $\mathbf{f} = \mathbf{0}$ for this problem.

We iterated our NNFP scheme to convergence at a tolerance of $\epsilon_{\text{nl}} = 10^{-3}$, starting from the fluid at rest, but for the inflow boundary condition, with

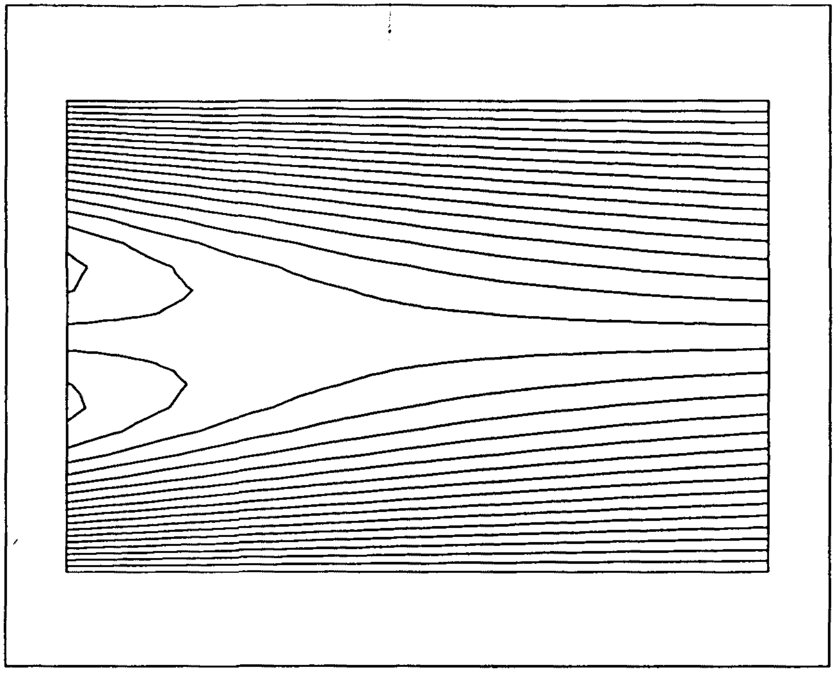


Figure 3.10: Kovasznay flow, streamlines.

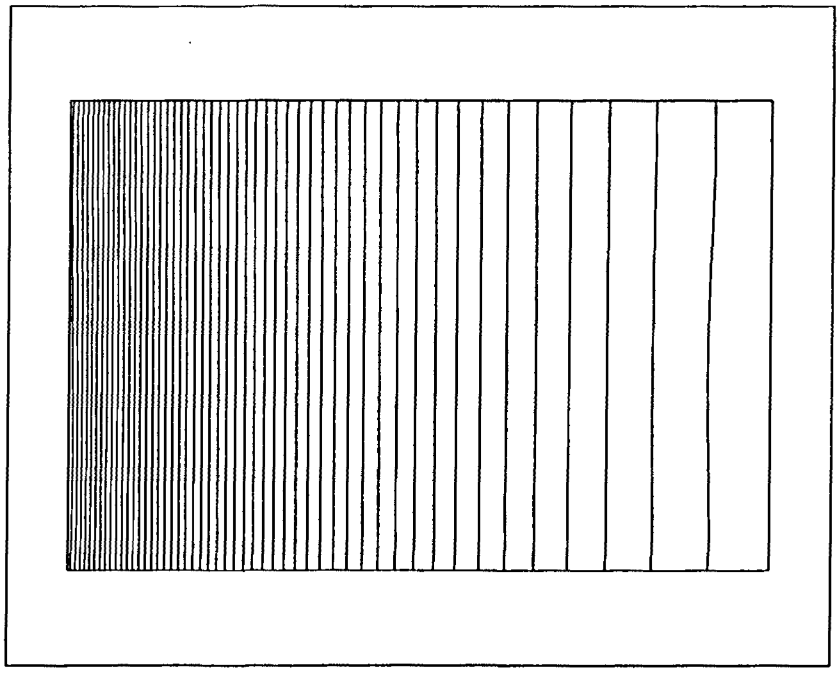


Figure 3.11: Kovasznay flow, pressure contours.

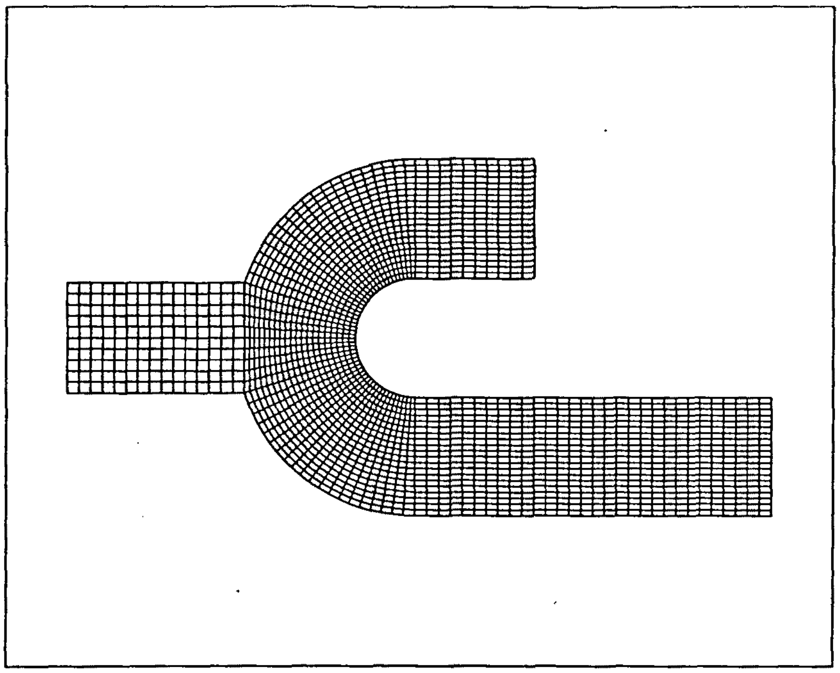


Figure 3.12: Flow through a junction, mesh.

local parameters $\alpha_K = \alpha_0 \frac{h_K^2}{4\nu}$ with $\alpha_0 = 1/3$, and for a Reynolds number of 50 (as in [56]). It took 8 iterations of our scheme to reach convergence, each of which needed $15, 10, 8, 6, 4, 2, 2$ and 1 iterations, respectively, for the inner loop to converge at a tolerance of $\epsilon_{\text{cou}} = 10^{-3}$. We also considered the GLS method with a Picard iteration and the same initial values. It took 7 iterations of this scheme to reach a converged state.

In Table 3.4 we present the results of a study of computing times (in seconds) for different phases of the two methods. The first row displays the average CPU time spent per iteration in the two cases, which is then split into the following four rows. The first one of these shows the average cost of the computation and assembly of the system matrix for the velocity–pressure equation; we have included here the times required for the common terms of the two methods, that is, the Laplacian for velocities and pressure, the gradient and divergence matrices and the advective term $A(U^{i-1})$. The next row shows the time needed for the additional terms with respect to the previous ones: in the GLS method, these correspond to the extra terms in the formulation, whereas in ours they are due to the need to evaluate the pressure residue $\alpha G W^{i-1,j}$ at each of the inner iterations (of which there are 6 in each of the outer ones in average). The average time for the matrix factorization in each iteration and the system solution (once per iteration in the GLS method and 6 times, in average, in ours) is shown in the fourth row; it can be deduced from it that it takes about 0.1 seconds, in average, to perform a

CPU times	GLS method	Present method
Average per iteration	18.2	18.4
Matrix computation	14.7	15.2
Additional terms	1.7	0.6
Matrix fact. and solution	1.7	2.2
Pres. grad. solution	-	0.3
Total time	84	100

Table 3.4: Flow in a junction of pipes: comparison of performance of the two methods.

forward and a backward substitution in this case (the unknowns are reordered at the beginning of the program following a certain renumbering strategy which minimizes the storage requirements of the system). Finally, we show the average time (in the outer iterations) required in our method for the pressure gradient residue formation and solution for all the inner iterations (this gives about 0.04 seconds per forward and backward substitution for this variable in this case). The total computing time, as a percentage of that of our method, is given in the last row.

It can be concluded from these computations that the evaluation of the extra terms in the GLS formulation makes it more costly to form the system matrix for that case in each iteration; but the need to perform the inner iteration loops for the coupling with the pressure gradient makes the average cost per iteration comparable for the two methods. Once again, our scheme needs one more iteration for the overall convergence, and this makes it about 19% more costly than the GLS method, in this example.

We show the results obtained for this problem in Figure 3.13, for both the GLS method and ours. The pressure contours and streamlines are quite satisfactory in both cases. To check the effect of the shorter outflow region introduced in the computational domain, we computed the flux through the

Method	Upper outflow	Lower outflow	Total outflow
GLS	0.36779	0.24369	0.61148
Present	0.36804	0.24344	0.61148

Table 3.5: Flow in a junction of pipes: flux through outflow regions.

upper and lower outflow sections for the two methods. Knowing that the inflow flux was 0.61148, ideally one would like to find half of that flux flowing through each outlet region, that is, 0.30574. The actual results obtained, which can be seen in Table 3.5, are very similar for the two methods; in both cases there is a greater flow through the upper section, but the total flux was conserved very accurately.

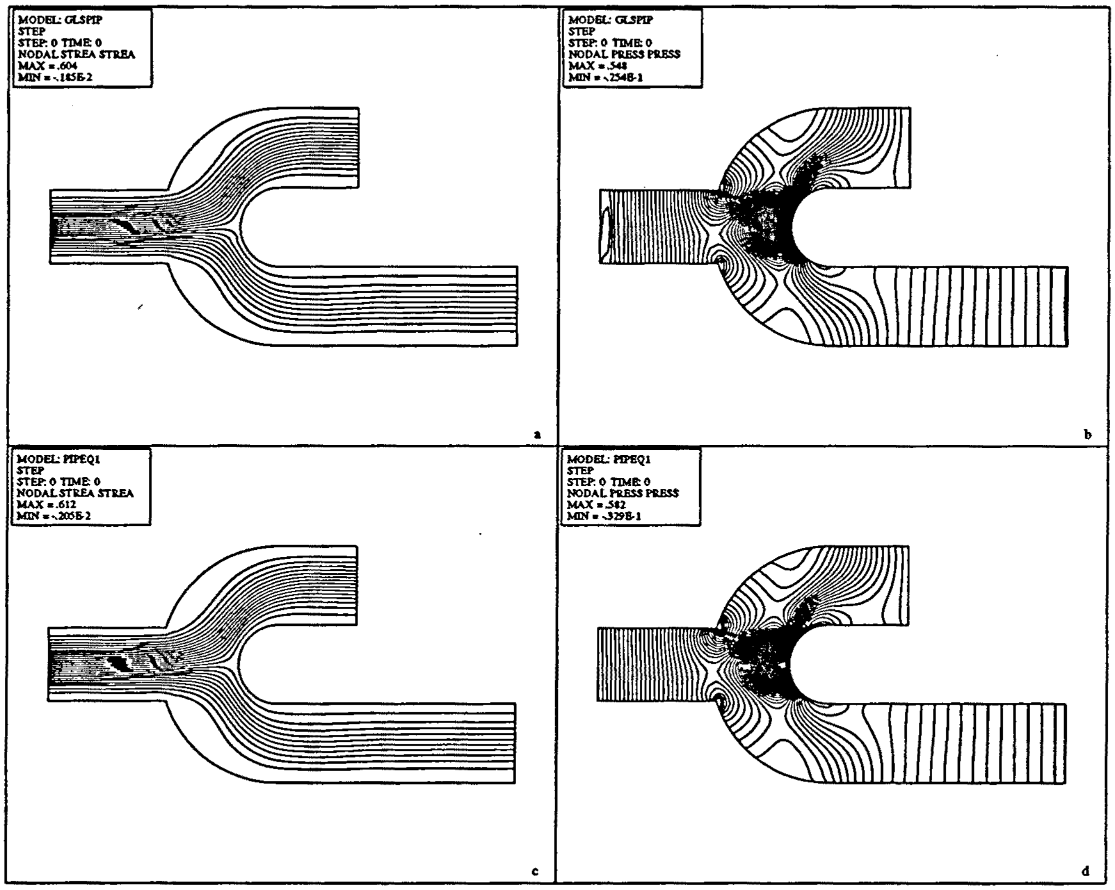


Figure 3.13: Flow through a junction: a) GLS method, streamlines; b) GLS method, pressure contours; c) Present method, streamlines; d) Present method, pressure contours.

Chapter 4

Viscosity splitting fractional step method

In this Chapter we develop and study a fractional–step method for the solution of the unsteady, incompressible Navier–Stokes equations. These equations constitute the full nonlinear, time evolution problem of incompressible viscous flow motion, whose numerical solution is of an undoubtable practical importance. Numerical methods for this problem have to deal with the discretization of both space and time.

The fractional–step method that we consider is mainly intended to overcome the difficulties encountered in projection methods regarding the imposition of boundary conditions. In our method, each time step is decomposed into two substeps, and in each of these the velocity boundary conditions of the continuous problem are enforced. Moreover, incompressibility is split from the nonlinearity, which are the two main difficulties met when solving the Navier–Stokes equations.

Furthermore, this method was introduced during the study of a known predictor–multicorrector algorithm applied to the solution of the unsteady, incompressible Navier–Stokes equations, which was this way shown to belong to the category of fractional–step methods. The study of this algorithm is the object of Chapter 5.

We review some known facts about the unsteady, incompressible Navier–Stokes equations in Section 4.1. In 4.2 we introduce the fractional step method that we consider, and prove the convergence of this method to the continuous solution. Moreover, under some stronger regularity assumptions on the continuous solution, we prove some error estimates for the velocity solution in the case of homogeneous Dirichlet boundary conditions; these estimates show that in this method both the *intermediate* and the *end–of-step* velocities are weakly order 1 accurate in the time step in the space $\mathbf{L}^2(\Omega)$, and weakly order 1/2 in $\mathbf{H}_0^1(\Omega)$; this last fact is possible due to the satisfaction of the correct boundary conditions at the two steps of the scheme. The pressure solution is also shown to be at least order 1/2 accurate. In Section 4.3 we consider a similar fractional step method, this time with pressure correction; similar error estimates are proved for this method for the velocity and pressure solutions. We then make some further remarks on fractional step methods, concerning the dependence of the steady state solution reached with these methods with respect to the time step. Furthermore, in 4.4 we present the fully discrete version of our viscosity splitting, pressure correction method with two different finite element space interpolations, and an efficient implementation of this method, while in 4.5 we show some numerical results obtained with it.

4.1 The unsteady, incompressible Navier Stokes equations

We recall here the standard formulation of the unsteady, incompressible Navier–Stokes equations in primitive velocity–pressure variables, assuming homogeneous Dirichlet boundary conditions for simplicity of exposition:

$$\begin{aligned}\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p - \nu\Delta\mathbf{u} &= \mathbf{f} \text{ in } \Omega\times(0,T)\\ \nabla\cdot\mathbf{u} &= 0 \text{ in } \Omega\times(0,T)\\ \mathbf{u} &= 0 \text{ on } \Gamma\times(0,T)\\ \mathbf{u}(\mathbf{x},0) &= \mathbf{u}_0(\mathbf{x}) \text{ in } \Omega\end{aligned}$$

Complete studies of this equation system can be found in [105], which we mainly follow here, and [71].

With the notation introduced in Chapter 1, the weak form of this problem consists of finding two functions $\mathbf{u} \in L^2(0,T;\mathbf{H}_0^1(\Omega))$ and $p \in L^2(0,T;L^2(\Omega))$ such that, given $\mathbf{f} \in L^2(0,T;\mathbf{H}^{-1}(\Omega))$ and $\mathbf{u}_0 \in H$:

$$\begin{aligned}\frac{d}{dt}(\mathbf{u}(t),\mathbf{v}) + a(\mathbf{u}(t),\mathbf{v}) + c(\mathbf{u}(t),\mathbf{u}(t),\mathbf{v}) + b(\mathbf{v},p(t)) &= (\mathbf{f},\mathbf{v}),\ \forall \mathbf{v}\in\mathbf{H}_0^1(\Omega)\\ b(\mathbf{u}(t),q) &= 0, \qquad \forall q \in L_0^2(\Omega)\\ \mathbf{u}(0) &= \mathbf{u}_0 \qquad\qquad (4.1)\end{aligned}$$

If the dimension of space is $d \leq 4$ and the domain Ω is bounded and Lipschitz continuous, problem 4.1 has at least one solution $(\mathbf{u},p)$, which satisfies $\mathbf{u} \in L^\infty(0,T;H)$ (see [105]). Uniqueness holds in the 2–dimensional case; in fact, if $d = 2$ the solution $(\mathbf{u},p)$ is unique, $\mathbf{u}$ is a.e. equal to a continuous function from $[0,T]$ into H and $\lim_{t\to 0^+}\mathbf{u}(t) = \mathbf{u}_0$ in H. We assume that this continuity result also holds in the three dimensional case.

Uniqueness and more regularity of the solution can be proved by assuming more regularity on the data $\mathbf{f}$ and $\mathbf{u}_0$ and the domain Ω. In fact, according to Heywood and Rannacher (see [54]), if $d \leq 3$ and Ω is such that the Stokes problem is regular, and if one assumes:

A1) $\mathbf{u}_0 \in \mathbf{H}^2(\Omega) \cap Y$ and $\mathbf{f}, \mathbf{f}_t \in L^\infty(0,T;\mathbf{L}^2(\Omega))$.

A2) if $d=3$, $\sup_{t\in[0,T]} \left(||\mathbf{u}(t)||\right) \leq M_1$.

(the subindex t is employed hereafter for $\frac{\partial}{\partial t}$) then the solution $\mathbf{u}$ of 4.1 is unique and satisfies $\mathbf{u} \in C^0(0,T;Y)$, $||\mathbf{u}(t) - \mathbf{u}_0||_2 \longrightarrow 0$ as $t \longrightarrow 0$ and:

R1) $\sup_{t\in[0,T]} \left\{||\mathbf{u}(t)||_2 \,+\, |\mathbf{u}_t(t)| \,+\, |\nabla p(t)|\right\} \,\leq\, C$

R2) $\int_0^T ||\mathbf{u}_t(t)||^2 \, dt \,\leq\, C$

R3) $\int_0^T t|\mathbf{u}_{tt}(t)|^2 \, dt \,\leq\, C$

Condition **A2** is automatically satisfied in the 2-dimensional case. Under the assumptions **A1–A2**, it is also shown in [90] that, according to the modifications introduced in [92]:

R4) $\int_0^T ||\mathbf{u}_{tt}(t)||^2_{Y'} \, dt \,\leq\, C$

These regularity results will be used in the following Sections. As is common practice in this context, we will use repeatedly in our proofs a discrete version of the Gronwall inequality. For the sake of completeness we recall the result here, but we refer to Heywood and Rannacher ([55]) for a proof. The version of this inequality that we shall use is the following:

Lemma 4.1: *let a_i, b_i, c_i, γ_i ($i \in \mathbb{N}$), k and B be positive real numbers such that, for $n \geq 0$:*

$$a_{n+1} \,+\, k\sum_{i=0}^{n+1} b_i \,\leq\, k\sum_{i=0}^{n+1} \gamma_i a_i \,+\, k\sum_{i=0}^{n+1} c_i \,+\, B \tag{4.2}$$

Suppose that $k\gamma_i < 1$ for all i, and set $\sigma_i = (1-k\gamma_i)^{-1}$. Then:

$$a_{n+1} \,+\, k\sum_{i=0}^{n+1} b_i \,\leq\, \exp(k\sum_{i=0}^{n} \sigma_i\gamma_i)\,(k\sum_{i=0}^{n+1} c_i \,+\, B) \tag{4.3}$$

Moreover, as it is deduced from a Remark in page 370 of [55], when the first sum on the right–hand–side of 4.2 extends only to n, then 4.3 holds for all k with $\sigma_i = 1$. In both cases, when all the coefficients γ_i are bounded from above, $k = \delta t$ and $n \leq [T/k]$, the exponential term in the right–hand–side of 4.3 can be bounded by a constant C independent of k. This is the result that we will actually use.

In some of our proofs we will also make use of the operator A^{-1}, defined as the inverse of the Stokes operator $A \doteq -P_H\Delta$. The latter is defined for $\mathbf{u} \in D(A) \doteq Y \cap \mathbf{H}^2(\Omega)$, and is an unbounded, positive, self–adjoint closed operator onto H. Given $\mathbf{u} \in H$, by definition of A, $\mathbf{v} = A^{-1}\mathbf{u}$ is the solution of the following Stokes problem:

$$\begin{aligned} -\Delta \mathbf{v} + \nabla r &= \mathbf{u} \\ \nabla \cdot \mathbf{v} &= 0 \\ \mathbf{v}_{|\Gamma} &= 0 \end{aligned} \tag{4.4}$$

When Ω is such that this problem is regular, there exists a constant $C_1 > 0$ such that:

$$||A^{-1}\mathbf{u}||_s \leq C_1||\mathbf{u}||_{s-2} \quad \text{for } s = 1,2 \tag{4.5}$$

Some inequalities were given by J. Shen in [90] for $(A^{-1}\mathbf{u}, \mathbf{u})$, with $\mathbf{u} \in H$, in terms of $||\mathbf{u}||_{-1}$, and used there to deduce error estimates for the standard projection method. Namely, he had:

$$C_2||\mathbf{u}||^2_{-1} \leq (A^{-1}\mathbf{u}, \mathbf{u}) \leq C_1^2||\mathbf{u}||^2_{-1} \tag{4.6}$$

where C_1 is the constant appearing in 4.5. But, as pointed out by J.L. Guermond in [50] and corrected in [92], the first inequality in not correct and has to be modified to:

$$C_2||\mathbf{u}||^2_{Y'} \leq (A^{-1}\mathbf{u}, \mathbf{u}) \tag{4.7}$$

With this modification, it is claimed in [92] that the results obtained in [90] (and [91]) still hold if the norm $||\mathbf{u}||_{-1}$ is replaced by $||\mathbf{u}||_{Y'}$ throughout the proofs. This is not quite true, since he is still using the inequality 4.5 for $s = 1$. We now show that in 4.5 with $s = 1$ the norm $||\mathbf{u}||_{-1}$ can be replaced by $||\mathbf{u}||_{Y'}$, which is what we actually use in our proofs. Thus, given $\mathbf{u} \in H$ let us call $\mathbf{v} = A^{-1}\mathbf{u}$; we have:

$$\begin{aligned} ||A^{-1}\mathbf{u}||^2 &= ((A^{-1}\mathbf{u}, A^{-1}\mathbf{u})) = ((\mathbf{v}, \mathbf{v})) \\ &= (\mathbf{u}, \mathbf{v}) + (r, \nabla \cdot \mathbf{v}) = (\mathbf{u}, \mathbf{v}) = <\mathbf{u}, \mathbf{v}> \\ &\leq ||\mathbf{u}||_{Y'} ||\mathbf{v}|| = ||\mathbf{u}||_{Y'} ||A^{-1}\mathbf{u}|| \end{aligned}$$

Thus, we have proved that:

$$||A^{-1}\mathbf{u}|| \leq ||\mathbf{u}||_{Y'}$$

We will use this result in what follows.

4.2 Viscosity splitting method

4.2.1 Development of the method

We present here a fractional-step method for the approximation in time of 4.1, in a semidiscrete form. The main purpose of introducing this scheme is to be able to enforce the boundary conditions of the original problem in the two substeps of the method, thus overcoming the difficulties of standard projection methods in this sense, which were explained in Section 1.5. This is achieved by splitting the viscous term into the two substeps. Some of the results presented here can be found in [11]. The method is presented depending on a free parameter $\theta > 0$, which we subsequently fix to 1.

First step: The first step of the method consists of finding, given $\mathbf{u}^n \in Y$, an intermediate velocity $\mathbf{u}^{n+1/2}$ such that:

$$\begin{aligned} \frac{\mathbf{u}^{n+1/2} - \mathbf{u}^n}{\delta t} - \theta\nu\Delta\mathbf{u}^{n+1/2} - (1-\theta)\nu\Delta\mathbf{u}^n + (\mathbf{u}^n \cdot \nabla)\mathbf{u}^{n+1/2} &= \mathbf{f}(t_{n+1}) \\ \mathbf{u}^{n+1/2}|_\Gamma &= 0 \quad (4.8) \end{aligned}$$

where $0 < \theta \leq 1$. The approximation of the convective term may take other forms; the semi–implicit approximation adopted here is taken from [90]. The weak form of 4.8 can be written as:

$$a_\theta^n(\mathbf{u}^{n+1/2}, \mathbf{v}) = l_1(\mathbf{v}), \qquad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega) \qquad (4.9)$$

where the bilinear form a_θ^n is defined by:

$$a_\theta^n(\mathbf{u}, \mathbf{v}) \doteq (\mathbf{u}, \mathbf{v}) + \theta\,\delta t\,\nu((\mathbf{u}, \mathbf{v})) + \delta t((\mathbf{u}^n \cdot \nabla)\mathbf{u}, \mathbf{v}),$$

and the linear form l_1 includes the known terms of 4.8, namely $l_1(\mathbf{v}) \doteq (\mathbf{u}^n, \mathbf{v}) - (1-\theta)\,\delta t\,\nu((\mathbf{u}^n, \mathbf{v})) + \delta t\,(\mathbf{f}(t_{n+1}), \mathbf{v})$. One has that a_θ^n is continuous and coercive with respect to $||\mathbf{u}||$ in $\mathbf{H}_0^1(\Omega)$, due to the skew–symmetric character of the approximation of the convective term (which is in turn a consequence of the solenoidal character of $\mathbf{u}^n$ and the vanishing of $\mathbf{u}^n$ at the boundary) and the presence of the Laplacian term. The form l_1 is continuous on $\mathbf{H}_0^1(\Omega)$ because of the Schwarz and Poincaré inequalities, so that existence and uniqueness of $\mathbf{u}^{n+1/2}$ is established by the Lax–Milgram theorem.

The fully implicit case $\theta = 1$ can be found in the original projection method of R.Temam (see [100]) and in the method of J.Shen ([90]), among others. The Crank–Nicholson case $\theta = 1/2$ is of main importance, since it provides a second order approximation of the viscous term. It is present in higher order methods such as [8], [45], [62] and [65]. The basic difference among these methods is the treatment of the nonlinearity, which is normally second order in time and explicit. The explicit case $\theta = 0$ has also been considered before (see [30] or [73], for instance); we exclude it from the present study because in that case the bilinear form a_0^n is not coercive on $\mathbf{H}_0^1(\Omega)$.

Second step: For the second step of the method, we avoid using the standard projection idea; instead, we include a diffusion term together with incompressibility, which allows the imposition of the full boundary conditions for the velocity. That is, given $\mathbf{u}^{n+1/2}$ from equation 4.9, we look for an *end-of-step* velocity $\mathbf{u}^{n+1}$ and an *end-of-step* pressure p^{n+1} such that:

$$\begin{aligned} \frac{\mathbf{u}^{n+1}-\mathbf{u}^{n+1/2}}{\delta t} - \theta\nu\Delta(\mathbf{u}^{n+1}-\mathbf{u}^{n+1/2}) + \nabla p^{n+1} &= 0 \qquad (4.10)\\ \nabla\cdot\mathbf{u}^{n+1} &= 0 \qquad (4.11)\\ \mathbf{u}^{n+1}|_\Gamma &= 0 \qquad (4.12)\end{aligned}$$

Similar ideas to this scheme can be found in some of the viscosity splitting methods of subsection 1.5.3. The weak form of 4.10–4.11–4.12 consists of finding $\mathbf{u}^{n+1} \in \mathbf{H}_0^1(\Omega)$ and $s^{n+1} = \delta t\, p^{n+1} \in L_0^2(\Omega)$ such that:

$$\begin{aligned} a_\theta(\mathbf{u}^{n+1},\mathbf{v}) + b(\mathbf{v},s^{n+1}) &= l_2(v), \quad \forall \mathbf{v}\in\mathbf{H}_0^1(\Omega)\\ b(\mathbf{u}^{n+1},q) &= 0, \quad \forall q \in L_0^2(\Omega) \qquad (4.13)\end{aligned}$$

where now:

$$a_\theta(\mathbf{u},\mathbf{v}) \doteq (\mathbf{u},\mathbf{v}) + \theta\,\delta t\,\nu((\mathbf{u},\mathbf{v})) \qquad (4.14)$$

is a bilinear, symmetric, continuous form on $\mathbf{H}_0^1(\Omega)$, which is also coercive with respect to $||\mathbf{u}||$, and $l_2(\mathbf{v}) = a_\theta(\mathbf{u}^{n+1/2},\mathbf{v})$ is a known linear continuous map. Problem 4.13 is a mixed problem, in which a_θ is coercive and b satisfies the continuous LBB condition 1.25, so that existence and uniqueness of a solution $(\mathbf{u}^{n+1}, s^{n+1})$ is guaranteed.

REMARK 4.1: By adding 4.8 and 4.10 one gets:

$$\frac{\mathbf{u}^{n+1}-\mathbf{u}^{n}}{\delta t} - \theta\nu\Delta\mathbf{u}^{n+1} - (1-\theta)\nu\Delta\mathbf{u}^n + (\mathbf{u}^n\cdot\nabla)\mathbf{u}^{n+1/2} + \nabla p^{n+1} = \mathbf{f}(t_{n+1}) \qquad (4.15)$$

where the implicit treatment of the viscous term in $\mathbf{u}^n$ and $\mathbf{u}^{n+1}$, and not in the intermediate velocity $\mathbf{u}^{n+1/2}$, can be observed. Moreover, it is clear from 4.15 that, at least for the linear problem, p^{n+1} keeps its meaning as an end-of-step pressure (this is not the case for some fractional step projection methods). The advantage of using a split scheme like 4.8–4.10 rather than a single $(\mathbf{u},p)$ step is the decoupling of the convective effects from incompressibility, which allows the use of suitable approximations for each term.

REMARK 4.2: As in standard projection methods, a Poisson equation can be derived for the pressure to solve 4.10–4.11–4.12. In fact, taking the divergence of 4.10 and using 4.11 yields:

$$\delta t\,\Delta p^{n+1} = (I - \theta\,\delta t\,\nu\Delta)\nabla\cdot\mathbf{u}^{n+1/2} \in H^{-1}(\Omega) \qquad (4.16)$$

sufficient smoothness of the functions involved been assumed. But in order that 4.16 and 4.10 imply 4.11, the incompressibility condition $\nabla\cdot\mathbf{u}^{n+1} = 0$

must also be enforced on the boundary (see [66]), as in the original method of A.J. Chorin (see [22]). Besides, boundary conditions for p^{n+1} cannot be directly derived, and p^{n+1} is subject to integral conditions (see [82]). Therefore, the original grad–div formulation 4.10–4.11 is adopted, which has the advantage of allowing discontinuous pressure approximations and requires of no boundary conditions at all for this variable. One drawback of solving 4.10–4.11 is the need for the spatial approximation chosen to satisfy the discrete LBB condition, a problem that will be encountered in the fully discrete version of the method, and that the velocity and pressure unknowns have to be dealt with at the same time.

REMARK 4.3: This method can also be understood as a projection method in a different sense than the classical one. In the space $\mathbf{H}_0^1(\Omega)$, let us define the norm $[[\mathbf{u}]]$ induced by the scalar product a_θ, which is an equivalent norm to $||\mathbf{u}||$. Recalling the decomposition $\mathbf{H}_0^1(\Omega) = Y \oplus Y^\perp$ of Section 1.2 and the characterization of $Y^\perp$, and calling P_Y the orthogonal projection from $\mathbf{H}_0^1(\Omega)$ onto Y in the norm $[[\mathbf{u}]]$, one has that for any $\mathbf{v} \in \mathbf{H}_0^1(\Omega)$ there exist $\mathbf{u} \in Y$ and $s \in L_0^2(\Omega)$ such that $\mathbf{u} = P_Y(\mathbf{v})$ and $s = (I - P_Y)(\mathbf{v})$, or equivalently, $\mathbf{v} = \mathbf{u} + (-\Delta)^{-1}(\nabla s)$. That is to say, one has that:

$$\begin{array}{rcll} a_\theta(\mathbf{v},\mathbf{w}) & = & a_\theta(\mathbf{u},\mathbf{w}) - (\nabla\cdot\mathbf{w}, s^{n+1}), & \forall \mathbf{w} \in \mathbf{H}_0^1(\Omega) \\ (\nabla\cdot\mathbf{u}^{n+1}, q) & = & 0, & \forall q \in L_0^2(\Omega) \end{array}$$

Equation 4.13 amounts to saying that $\mathbf{u}^{n+1} = P_Y(\mathbf{u}^{n+1/2})$, so that $\mathbf{u}^{n+1}$ is the projection of $\mathbf{u}^{n+1/2}$ onto Y with respect to the norm $[[\mathbf{u}]]$.

For the case $\theta = 1$, we first proved convergence of the intermediate and end–of–step velocities to a continuous solution in the spaces $\mathbf{H}_0^1(\Omega)$ and $\mathbf{L}^2(\Omega)$, in the appropiate sense, in a similar way to the proof of the convergence of the classical projection method given by R. Temam in [100]. The convergence of $\mathbf{u}^{n+1}$ in $\mathbf{H}_0^1(\Omega)$ could not be obtained for the standard projection method, since in that case $\mathbf{u}^{n+1} \notin \mathbf{H}_0^1(\Omega)$ (it does not satisfy the proper boundary condition).

We then obtained weakly order 1 error estimates in the time step for $\mathbf{u}^{n+1/2}$ and $\mathbf{u}^{n+1}$ in $\mathbf{L}^2(\Omega)$ and weakly order 1/2 in $\mathbf{H}_0^1(\Omega)$, following the ideas of J. Shen in [90] and under some more regularity conditions on the continuous solution; the estimates for $\mathbf{u}^{n+1}$ can be improved to strongly order 1 in $\mathbf{L}^2(\Omega)$ and weakly order 1 in $\mathbf{H}_0^1(\Omega)$, under some rather restrictive assumptions; we give the proof of this improvement in an Appendix. We also obtained order 1/2 error estimates for the pressure.

4.2.2 Convergence of the method

We include here our first proof of convergence of the viscosity splitting, fractional step method just considered, which is based on the proof of convergence of the original projection method given by R. Temam in [101] and included in [105]. The proof for our scheme can also be found in [11].

Let us assume that $\mathbf{f} \in L^2(0,T;\mathbf{L}^2(\Omega))$, and consider the weak form of the unsteady, incompressible Navier–Stokes equations 4.1. Its solutions are characterized by satisfying (see [105]) $\mathbf{u} \in L^2(0,T;Y)$ and:

$$\frac{d}{dt}(\mathbf{u}(t),\mathbf{v}) + ((\mathbf{u}(t)\cdot\nabla)\mathbf{u}(t),\mathbf{v}) + \nu((\mathbf{u}(t),\mathbf{v})) = (\mathbf{f}(t),\mathbf{v}), \quad \forall \mathbf{v} \in Y \tag{4.17}$$

We consider the fractional step method 4.8 and 4.10–4.11–4.12 with $\theta = 1$, but with an approximation of the force term $\bar{\mathbf{f}}^n$ which is the time average of $\mathbf{f}$ in $[t_n, t_{n+1}]$ (as is taken in [105]). For this method, and calling $k = \delta t$ to follow the standard notation in this context, we have:

Lemma 4.2: *for all $N = 0,\ldots,[T/k]-1$, the following a priori estimate holds:*

$$\begin{aligned} |\mathbf{u}^{N+1}|^2 &+ \sum_{n=0}^{N}(|\mathbf{u}^{n+1}-\mathbf{u}^{n+1/2}|^2 + |\mathbf{u}^{n+1/2}-\mathbf{u}^n|^2) \\ &+ k\nu\sum_{n=0}^{N}||\mathbf{u}^{n+1}||^2 + k\nu\sum_{n=0}^{N}||\mathbf{u}^{n+1}-\mathbf{u}^{n+1/2}||^2 \le C_1 \end{aligned} \tag{4.18}$$

where $C_1 = |\mathbf{u}^0|^2 + \frac{C_\Omega^2}{\nu}\int_0^T |\mathbf{f}(s)|^2\,ds$ and C_Ω was introduced in 1.14.

PROOF: the proof is similar to that of Lemma 7.1.2 in [105]. Taking the product of 4.8 with $2\,k\,\mathbf{u}^{n+1/2}$ and using the identity $(a-b,2a) = |a|^2 - |b|^2 + |a-b|^2$, we get:

$$\begin{aligned} |\mathbf{u}^{n+1/2}|^2 &- |\mathbf{u}^n|^2 + |\mathbf{u}^{n+1/2}-\mathbf{u}^n|^2 + 2\,\nu\,k||\mathbf{u}^{n+1/2}||^2 \\ &= 2\,k\,(\bar{\mathbf{f}}^n,\mathbf{u}^{n+1/2}) \le 2\,k\,|\bar{\mathbf{f}}^n|\,|\mathbf{u}^{n+1/2}| \\ &\le 2\,k\,|\bar{\mathbf{f}}^n|\,C_\Omega\,||\mathbf{u}^{n+1/2}|| \le \nu\,k\,||\mathbf{u}^{n+1/2}||^2 + k\frac{C_\Omega^2}{\nu}|\bar{\mathbf{f}}^n|^2 \end{aligned}$$

so that:

$$\begin{aligned} |\mathbf{u}^{n+1/2}|^2 &- |\mathbf{u}^n|^2 + |\mathbf{u}^{n+1/2}-\mathbf{u}^n|^2 + \nu\,k||\mathbf{u}^{n+1/2}||^2 \\ &\le k\frac{C_\Omega^2}{\nu}|\bar{\mathbf{f}}^n|^2 \end{aligned} \tag{4.19}$$

Taking now the product of 4.10 with $2k\mathbf{u}^{n+1}$, we get:

$$|\mathbf{u}^{n+1}|^2 - |\mathbf{u}^{n+1/2}|^2 + |\mathbf{u}^{n+1} - \mathbf{u}^{n+1/2}|^2 + k\nu\left(||\mathbf{u}^{n+1}||^2 - ||\mathbf{u}^{n+1/2}||^2 + ||\mathbf{u}^{n+1} - \mathbf{u}^{n+1/2}||^2\right) = 0 \tag{4.20}$$

Adding up 4.19 and 4.20 for $n = 0, \dots, N$, we obtain 4.18 using the fact that (see [105]):

$$k\sum_{n=0}^{N} |\bar{\mathbf{f}}^n|^2 \le \int_0^T |\mathbf{f}(s)|^2\, ds$$

□

Notice the last term appearing in the left–hand side of 4.18, which is not present in [105].

Lemma 4.3: *for every $m, N = 0, \dots, [T/k] - 1$:*

1) $|\mathbf{u}^{m+i/2}|^2 \le C_1,\ i = 1, 2$
2) $k||\mathbf{u}^{m+1/2}||^2 \le C_1/\nu$
3) $\sum_{n=0}^{N} |\mathbf{u}^{n+1} - \mathbf{u}^{n+1/2}|^2 \le C_1$
4) $\sum_{n=0}^{N} |\mathbf{u}^{n+1/2} - \mathbf{u}^{n}|^2 \le C_1$
5) $k\sum_{n=0}^{N} ||\mathbf{u}^{n+1} - \mathbf{u}^{n+1/2}||^2 \le C_1/\nu$
6) $k\sum_{n=0}^{N} ||\mathbf{u}^{n+1}||^2 \le C_1/\nu$

PROOF: the proof is, again, similar to that of [105]. Parts 3) through 6) follow from 4.18. Part 1) with $i = 1$ and part 2) follows from the addition of 4.19 for $n = 0, \dots, m$ and 4.20 for $n = 0, \dots, m - 1$. Finally, part 1) with $i = 2$ is obtained by adding up 4.19 and 4.20 for $n = 0, \dots, m$. □

Notice that the bound 5) was not obtained in [105]. We now define some approximating functions $\mathbf{u}_k^i$ and $\mathbf{u}_k$ in a similar way to [105], which were mentioned in Section 1.5. We introduce a new function $\mathbf{u}_k^3$ which we need for the treatment of the convective term:

$\mathbf{u}_k^1 : [0, T] \to \mathbf{L}^2(\Omega)\ /\ \mathbf{u}_k^1(t) = \mathbf{u}^{n+1/2},\ \ nk \le t < (n+1)k$

$\mathbf{u}_k^2 : [0, T] \to \mathbf{L}^2(\Omega)\ /\ \mathbf{u}_k^2(t) = \mathbf{u}^{n+1},\ \ nk \le t < (n+1)k$

$\mathbf{u}_k^3$: $[0,T] \to \mathbf{L}^2(\Omega)$ / $\mathbf{u}_k^3(t) = \mathbf{u}^n$, $nk \le t < (n+1)k$

$\mathbf{u}_k$: $[0,T] \to \mathbf{L}^2(\Omega)$ / $\mathbf{u}_k$ is continuous, linear on t on each interval $[nk,(n+1)k]$ and $\mathbf{u}_k(t_n) = \mathbf{u}^n$, for $n = 0,\dots,[T/k]$.

These approximating functions $\mathbf{u}_k^i$ and $\mathbf{u}_k$ satisfy, for decreasing k:

Lemma 4.4: *as k tends to zero,*

1) $\mathbf{u}_k^i$ and $\mathbf{u}_k$ are bounded in $\mathbf{L}^\infty(0,T;\mathbf{L}^2(\Omega))$, $i = 1,2,3$
2) $\mathbf{u}_k^i$ and $\mathbf{u}_k$ are bounded in $\mathbf{L}^2(0,T;\mathbf{H}_0^1(\Omega))$, $i = 1,2,3$
3) $(\mathbf{u}_k^2 - \mathbf{u}_k^1)$ and $(\mathbf{u}_k^2 - \mathbf{u}_k^3)$ are bounded in $\mathbf{L}^2(0,T;\mathbf{H}_0^1(\Omega))$

PROOF: these results are a consequence of Lemma 4.3 and the definitions of the functions. □

The main novelty with respect to [105] is now the boundedness of $\mathbf{u}_k^2$ and $\mathbf{u}_k$ in $L^2(0,T;\mathbf{H}_0^1(\Omega))$, together with that of the differences $(\mathbf{u}_k^2 - \mathbf{u}_k^1)$ and $(\mathbf{u}_k^2 - \mathbf{u}_k^3)$. Moreover:

Lemma 4.5:

1) $||\mathbf{u}_k^2 - \mathbf{u}_k^1||_{L^2(0,T;\mathbf{L}^2(\Omega))} \le \sqrt{kC_1}$
2) $||\mathbf{u}_k^2 - \mathbf{u}_k^3||_{L^2(0,T;\mathbf{L}^2(\Omega))} \le \sqrt{4kC_1}$
3) $||\mathbf{u}_k - \mathbf{u}_k^2||_{L^2(0,T;\mathbf{L}^2(\Omega))} \le \sqrt{\dfrac{4kC_1}{3}}$

PROOF: part 1) follows from Lemma 4.3, part 3); 2) results from Lemma 4.3, parts 3) and 4) and the triangle inequality. Finally, 3) is a consequence of the definition of $\mathbf{u}_k$ and Lemma 4.3, parts 3) and 4). □

Following [105], let us now define $\mathbf{f}_k \in L^2(0,T;\mathbf{L}^2(\Omega))$ as $\mathbf{f}_k(t) = \bar{\mathbf{f}}^n$ for $t_n \le t < t_{n+1}$ and $n = 0,\dots,[T/k]-1$. Then:

Lemma 4.6:

$$\begin{aligned}\frac{d}{dt}(\mathbf{u}_k(t),\mathbf{v}) &= -\nu((\mathbf{u}_k^2(t),\mathbf{v})) - c(\mathbf{u}_k^3(t),\mathbf{u}_k^1(t),\mathbf{v}) + (\mathbf{f}_k(t),\mathbf{v}) \\ &\doteq \; < g_k(t),\mathbf{v} >, \qquad \forall \mathbf{v} \in Y, \;\; \forall t \in (0,T) \qquad (4.21)\end{aligned}$$

with g_k bounded in $L^2(0,T;Y')$. In particular, $\mathbf{u}_k$ is a.e. equal to a continuous function from $[0,T]$ into Y.

PROOF: the weak form of 4.8 and 4.10–4.11–4.12 can also be written as:

$$\begin{aligned}(\frac{\mathbf{u}^{n+1/2}-\mathbf{u}^n}{k},\mathbf{v}) + \nu((\mathbf{u}^{n+1/2},\mathbf{v})) &+ c(\mathbf{u}^n,\mathbf{u}^{n+1/2},\mathbf{v}) \qquad (4.22)\\ &= (\bar{\mathbf{f}}^n,\mathbf{v}), \qquad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega)\end{aligned}$$

and:

$$\begin{aligned}(\frac{\mathbf{u}^{n+1}-\mathbf{u}^{n+1/2}}{k},\mathbf{v}) + \nu((\mathbf{u}^{n+1}-\mathbf{u}^{n+1/2},\mathbf{v})) &+ b(\mathbf{v},p^{n+1}) \qquad (4.23)\\ &= 0, \quad \forall \mathbf{v} \in \mathbf{H}_0^1(\Omega)\\ b(\mathbf{u}^{n+1},q) &= 0, \quad \forall q \in L_0^2(\Omega)\end{aligned}$$

respectively. By adding 4.22 and 4.23 for $\mathbf{v} \in Y$, one gets:

$$\begin{aligned}(\frac{\mathbf{u}^{n+1}-\mathbf{u}^n}{k},\mathbf{v}) + \nu((\mathbf{u}^{n+1},\mathbf{v}) &+ c(\mathbf{u}^n,\mathbf{u}^{n+1/2},\mathbf{v})\\ &= <\bar{\mathbf{f}}^n,v> \qquad \forall \mathbf{v} \in Y\end{aligned}$$

so that 4.21 follows from the above definitions. Besides:

$$||g_k(t)||_{Y'} \leq \nu||\mathbf{u}_k^2(t)|| + C_{111}||\mathbf{u}_k^3(t)||\,||\mathbf{u}_k^1(t)|| + |\mathbf{f}_k(t)|$$

where $C_{111} > 0$ is a constant related to the continuity of the trilinear form c (see Section 1.2); the remaining statements are a consequence of Lemma 4.4 and Lemma III.1.1 of [105]. □

The proof of a convergence theorem is now ready:

Theorem 4.1: *let $\mathbf{f} \in L^2(0,T;H)$ and $\mathbf{u}^0 \in Y$. Then, there exists a subsequence k' of k and a solution $\mathbf{u}$ of the Navier–Stokes equations 4.17 such that:*

1) *$\mathbf{u}_{k'}^i$ and $\mathbf{u}_{k'}$ converge to $\mathbf{u}$ in $L^2(0,T;\mathbf{L}^2(\Omega))$ strongly, $i = 1,2,3$.*
2) *$\mathbf{u}_{k'}^i$ and $\mathbf{u}_{k'}$ converge to $\mathbf{u}$ in $L^\infty(0,T;\mathbf{L}^2(\Omega))$ weak-star, $i = 1,2,3$.*
3) *$\mathbf{u}_{k'}^i$ and $\mathbf{u}_{k'}$ converge to $\mathbf{u}$ in $L^2(0,T;\mathbf{H}_0^1(\Omega))$ weakly, $i = 1,2,3$.*

For any other subsequence k'' such that these convergence results hold, $\mathbf{u}$ must be a solution of 4.17.

PROOF: since $\mathbf{u}_k^i$ $(i = 1,2,3)$ and $\mathbf{u}_k$ are bounded in $L^\infty(0,T;\mathbf{L}^2(\Omega))$, there exists a subsequence k' (which can be taken the same for all 4 sequences) and $\mathbf{u}^i$ $(i = 1,2,3)$, $\mathbf{u}^* \in L^\infty(0,T;\mathbf{L}^2(\Omega))$ such that:

$$\begin{aligned} \mathbf{u}^i_{k'} &\longrightarrow \mathbf{u}^i \text{ in } L^\infty(0,T;\mathbf{L}^2(\Omega)) \quad \text{weak} - \text{star} \ (i=1,2,3) \\ \mathbf{u}_{k'} &\longrightarrow \mathbf{u}^* \text{ in } L^\infty(0,T;\mathbf{L}^2(\Omega)) \quad \text{weak} - \text{star} \end{aligned}$$

Since $\mathbf{u}^i_{k'}$ $(i=1,2,3)$ and $\mathbf{u}_{k'}$ are bounded in $L^2(0,T;\mathbf{H}^1_0(\Omega))$, there exists a subsequence of k' (which is also denoted by k') such that:

$$\begin{aligned} \mathbf{u}^i_{k'} &\longrightarrow \mathbf{u}^i \text{ in } L^2(0,T;\mathbf{H}^1_0(\Omega) \quad \text{weakly} \ (i=1,2,3) \\ \mathbf{u}_{k'} &\longrightarrow \mathbf{u}^* \text{ in } L^2(0,T;\mathbf{H}^1_0(\Omega) \quad \text{weakly} \end{aligned}$$

This convergence also holds in $L^2(0,T;\mathbf{L}^2(\Omega))$. Since, by Lemma 4.5:

$$(\mathbf{u}^2_{k'} - \mathbf{u}^3_{k'}),\ (\mathbf{u}^2_{k'} - \mathbf{u}^1_{k'}),\ (\mathbf{u}^2_{k'} - \mathbf{u}_{k'}) \longrightarrow 0 \text{ in } L^2(0,T;\mathbf{L}^2(\Omega)) \quad \text{strongly,}$$

it must be $\mathbf{u}^1 = \mathbf{u}^2 = \mathbf{u}^3 = \mathbf{u}^*$ in $L^\infty(0,T;H) \cap L^2(0,T;Y)$.

Since $\mathbf{u}^2_{k'} \in L^\infty(0,T;H) \cap L^2(0,T;Y)$, one has that $\mathbf{u}^*(t) \in Y$ a.e. in $(0,T)$, and $\mathbf{u}^* \in L^\infty(0,T;H) \cap L^2(0,T;Y)$.

The proof of strong convergence in $L^2(0,T;\mathbf{L}^2(\Omega))$ is the same as in [105], and is therefore omitted. It only remains to show that $\mathbf{u}^*$ is a solution of 4.17. The same argument as in [105] is used, so that the convergence results already proved imply, by taking 4.21 to the limit when k' tends to 0, that:

$$\frac{d}{dt}(\mathbf{u}^*,\mathbf{v}) + \nu((\mathbf{u}^*,\mathbf{v})) + c(\mathbf{u}^*,\mathbf{u}^*,\mathbf{v}) = (\mathbf{f},\mathbf{v}) \quad \forall v \in Y$$

in distribution sense in $(0,T)$, i.e., $\mathbf{u}^*$ satisfies 4.17. This, in turn, implies (see [105]) that $\frac{d\mathbf{u}^*}{dt} \in L^1(0,T;Y')$, $\mathbf{u}^*(0) = \mathbf{u}^0$ weakly in Y and $\mathbf{u}^*$ is a.e. equal to a continuous function from $(0,T)$ into Y. These results ensure that $\mathbf{u}^*$ is a weak solution of 4.17, and the theorem is thus proved. □

In the two dimensional case, one has:

Corollary 4.1: *let $d=2$. Then, the convergence given in Theorem* 4.1 *is of the sequence as a whole.*

PROOF: this result is a consequence of the uniqueness of the solution $\mathbf{u}$ in the two–dimensional case. □

In summary, both the intermediate $\mathbf{u}^{n+1/2}$ and the end–of–step velocities $\mathbf{u}^{n+1}$ have been shown to converge to $\mathbf{u}(t_{n+1})$ in $\mathbf{H}^1_0(\Omega)$, through the functions $\mathbf{u}^1_k$ and $\mathbf{u}^2_k$ respectively. This is an improvement with respect to [105], where $\mathbf{u}^{n+1}$ only converges in $\mathbf{L}^2(\Omega)$.

4.2.3 Error estimates

We now present an error analysis of our viscosity splitting, fractional step method with parameter $\theta = 1$, which follows mainly the ideas of [90] with the modifications introduced in [92].

Let us define the velocity error functions for this method as:

$$\begin{aligned} \mathbf{e}^{n+1} &= \mathbf{u}(t_{n+1}) - \mathbf{u}^{n+1} \\ \mathbf{e}^{n+1/2} &= \mathbf{u}(t_{n+1}) - \mathbf{u}^{n+1/2} \end{aligned}$$

We give an estimate for $\mathbf{e}^{n+1}$ and $\mathbf{e}^{n+1/2}$; in particular, we show that both $\mathbf{u}^{n+1}$ and $\mathbf{u}^{n+1/2}$ are strongly order 1/2 approximations to $\mathbf{u}(t_{n+1})$ in $\mathbf{L}^2(\Omega)$ and weakly order 1/2 in $\mathbf{H}_0^1(\Omega)$.

Lemma 4.7: *if* **A1** *and* **A2** *hold, and if the Stokes problem is regular, then for* $N = 0, \ldots, [T/k] - 1$*:*

$$\begin{aligned} |\mathbf{e}^{N+1}|^2 &+ |\mathbf{e}^{N+1/2}|^2 + \sum_{n=0}^{N} \Big\{ |\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}|^2 + |\mathbf{e}^{n+1/2} - \mathbf{e}^{n}|^2 \Big\} \qquad (4.24) \\ &+ k\nu \sum_{n=0}^{N} \Big\{ ||\mathbf{e}^{n+1}||^2 + ||\mathbf{e}^{n+1/2}||^2 + ||\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}||^2 \Big\} \leq C\,k \end{aligned}$$

PROOF: the first part of the proof is similar to that of [90]. We call $\mathbf{R}^n$ the truncation error defined by:

$$\begin{aligned} \frac{1}{k}(\mathbf{u}(t_{n+1}) - \mathbf{u}(t_n)) &- \nu\Delta(\mathbf{u}(t_{n+1})) + (\mathbf{u}(t_{n+1}) \cdot \nabla)\mathbf{u}(t_{n+1}) + \nabla p(t_{n+1}) \\ &= \mathbf{f}(t_{n+1}) + \mathbf{R}^n \qquad (4.25) \end{aligned}$$

so that:

$$\mathbf{R}^n = \frac{1}{k}\int_{t_n}^{t_{n+1}} (t - t_n)\, \mathbf{u}_{tt}(t)\, dt$$

Subtracting 4.8 from 4.25, we get:

$$\begin{aligned} \frac{1}{k}(\mathbf{e}^{n+1/2} - \mathbf{e}^n) &- \nu\Delta(\mathbf{e}^{n+1/2}) = (\mathbf{u}^n \cdot \nabla)\mathbf{u}^{n+1/2} - (\mathbf{u}(t_{n+1}) \cdot \nabla)\mathbf{u}(t_{n+1}) \\ &+ \mathbf{R}^n - \nabla p(t_{n+1}) \qquad (4.26) \end{aligned}$$

We split the nonlinear terms on the right hand side of 4.26 into three terms as in [90]:

$$\begin{aligned}(\mathbf{u}^n \cdot \nabla)\mathbf{u}^{n+1/2} \;&-\; (\mathbf{u}(t_{n+1}) \cdot \nabla)\mathbf{u}(t_{n+1}) \qquad (4.27)\\ &= \; -(\mathbf{e}^n \cdot \nabla)\mathbf{u}^{n+1/2} \;+\; \Big((\mathbf{u}(t_n) - \mathbf{u}(t_{n+1})) \cdot \nabla\Big)\mathbf{u}^{n+1/2}\\ &- \; (\mathbf{u}(t_{n+1}) \cdot \nabla)\mathbf{e}^{n+1/2}\end{aligned}$$

and then take the inner product of 4.26 with $2k\mathbf{e}^{n+1/2}$ to obtain:

$$\begin{aligned}|\mathbf{e}^{n+1/2}|^2 \;&-\; |\mathbf{e}^n|^2 \;+\; 2\,k\,\nu\,||\mathbf{e}^{n+1/2}||^2 \;+\; |\mathbf{e}^{n+1/2} - \mathbf{e}^n|^2 \qquad (4.28)\\ &= \; 2\,k\; < \mathbf{R}^n, \mathbf{e}^{n+1/2} > \; - \; 2\,k\,(\nabla p(t_{n+1}), \mathbf{e}^{n+1/2})\\ &- \; 2\,k\,c(\mathbf{e}^n, \mathbf{u}^{n+1/2}, \mathbf{e}^{n+1/2}) \;+\; 2\,k\,c(\mathbf{u}(t_n) - \mathbf{u}(t_{n+1}), \mathbf{u}^{n+1/2}, \mathbf{e}^{n+1/2})\\ &- \; 2\,k\,c(\mathbf{u}(t_{n+1}), \mathbf{e}^{n+1/2}, \mathbf{e}^{n+1/2})\end{aligned}$$

We bound each term in the RHS of 4.28 independently:

- Taylor residual term:

$$\begin{aligned}2\,k\; < \mathbf{R}^n, \mathbf{e}^{n+1/2} > \;&\leq\; 2\,k\,||\mathbf{R}^n||_{-1}\,||\mathbf{e}^{n+1/2}||\\ &\leq\; \frac{k\nu}{3}||\mathbf{e}^{n+1/2}||^2 \;+\; C\,k\,||\mathbf{R}^n||_{-1}^2\\ &=\; \frac{k\nu}{3}||\mathbf{e}^{n+1/2}||^2 \;+\; \frac{C}{k}||\int_{t_n}^{t_{n+1}}(t - t_n)\mathbf{u}_{tt}\,dt||_{-1}^2\\ &\leq\; \frac{k\nu}{3}||\mathbf{e}^{n+1/2}||^2\\ &+\; \frac{C}{k}\int_{t_n}^{t_{n+1}}(t - t_n)||\mathbf{u}_{tt}||_{-1}^2\,dt\;\int_{t_n}^{t_{n+1}}(t - t_n)\,dt\\ &\leq\; \frac{k\nu}{3}||\mathbf{e}^{n+1/2}||^2 \;+\; C\,k\int_{t_n}^{t_{n+1}} t\,||\mathbf{u}_{tt}||_{-1}^2\,dt\end{aligned}$$

- Pressure gradient term:

$$\begin{aligned}-2\,k\,(\nabla p(t_{n+1}), \mathbf{e}^{n+1/2}) \;&=\; -2\,k\,(\nabla p(t_{n+1}), \mathbf{e}^{n+1/2} - \mathbf{e}^n)\\ &\leq\; 2\,k\,|\nabla p(t_{n+1})|\;|\mathbf{e}^{n+1/2} - \mathbf{e}^n|\\ &\leq\; \frac{1}{2}|\mathbf{e}^{n+1/2} - \mathbf{e}^n|^2 \;+\; 2\,k^2\,|\nabla p(t_{n+1})|^2\end{aligned}$$

since $\nabla \cdot \mathbf{e}^n = 0$.

- Nonlinear terms:

$$\begin{aligned}
T_1 &= -2\,k\,c(\mathrm{e}^n, \mathbf{u}^{n+1/2}, \mathrm{e}^{n+1/2}) \;=\; 2\,k\,c(\mathrm{e}^n, \mathrm{e}^{n+1/2}, \mathbf{u}^{n+1/2}) \\
&= 2\,k\,c(\mathrm{e}^n, \mathrm{e}^{n+1/2}, \mathbf{u}(t_{n+1})) \\
&\le C\,k\,|\mathrm{e}^n|\;||\mathrm{e}^{n+1/2}||\;||\mathbf{u}(t_{n+1})||_2 \;\le\; C\,k\,|\mathrm{e}^n|\;||\mathrm{e}^{n+1/2}|| \\
&\le \frac{k\nu}{3}||\mathrm{e}^{n+1/2}||^2 \;+\; C\,k\,|\mathrm{e}^n|^2
\end{aligned}$$

$$\begin{aligned}
T_2 &= 2\,k\,c(\mathbf{u}(t_n) - \mathbf{u}(t_{n+1}), \mathbf{u}^{n+1/2}, \mathrm{e}^{n+1/2}) \\
&= -2\,k\,c(\mathbf{u}(t_n) - \mathbf{u}(t_{n+1}), \mathrm{e}^{n+1/2}, \mathbf{u}^{n+1/2}) \\
&= -2\,k\,c(\mathbf{u}(t_n) - \mathbf{u}(t_{n+1}), \mathrm{e}^{n+1/2}, \mathbf{u}(t_{n+1})) \\
&\le C\,k\,|\mathbf{u}(t_n) - \mathbf{u}(t_{n+1})|\;||\mathrm{e}^{n+1/2}||\;||\mathbf{u}(t_{n+1})||_2 \\
&\le C\,k\,|\mathbf{u}(t_n) - \mathbf{u}(t_{n+1})|\;||\mathrm{e}^{n+1/2}|| \\
&\le \frac{k\nu}{3}||\mathrm{e}^{n+1/2}||^2 \;+\; C\,k\,|\int_{t_n}^{t_{n+1}} \mathbf{u}_t\,dt|^2 \\
&\le \frac{k\nu}{3}||\mathrm{e}^{n+1/2}||^2 \;+\; C\,k^2 \int_{t_n}^{t_{n+1}} |\mathbf{u}_t|^2\,dt
\end{aligned}$$

$$T_3 \;=\; -2\,k\,c(\mathbf{u}(t_{n+1}), \mathrm{e}^{n+1/2}, \mathrm{e}^{n+1/2}) \;=\; 0$$

where we have used **R1** and the boundedness and skew–symmetry properties of the trilinear form c.

From all these inequalities we deduce:

$$\begin{aligned}
|\mathrm{e}^{n+1/2}|^2 \;&-\; |\mathrm{e}^n|^2 \;+\; k\,\nu\,||\mathrm{e}^{n+1/2}||^2 \;+\; \frac{1}{2}|\mathrm{e}^{n+1/2} - \mathrm{e}^n|^2 \\
&\le\; C\,k \int_{t_n}^{t_{n+1}} t\,||\mathbf{u}_{tt}||_{-1}^2\,dt \;+\; C\,k^2 \int_{t_n}^{t_{n+1}} |\mathbf{u}_t|^2\,dt \qquad (4.29) \\
&+\; 2\,k^2\,|\nabla p(t_{n+1})|^2 \;+\; C\,k\,|\mathrm{e}^n|^2
\end{aligned}$$

The proof is now different from that of [90]. From 4.10 we have:

$$\frac{\mathrm{e}^{n+1} - \mathrm{e}^{n+1/2}}{k} \;-\; \nu\Delta(\mathrm{e}^{n+1} - \mathrm{e}^{n+1/2}) \;-\; \nabla p^{n+1} \;=\; 0 \qquad (4.30)$$

Taking the inner product of 4.30 with $2k\mathrm{e}^{n+1}$, given that $\nabla \cdot \mathrm{e}^{n+1} = 0$ and that $\mathrm{e}_{|\Gamma}^{n+1} = 0$, we get:

$$\begin{aligned}
|\mathrm{e}^{n+1}|^2 \;-\; |\mathrm{e}^{n+1/2}|^2 \;&+\; |\mathrm{e}^{n+1} - \mathrm{e}^{n+1/2}|^2 \qquad (4.31) \\
&+\; k\,\nu\left(||\mathrm{e}^{n+1}||^2 - ||\mathrm{e}^{n+1/2}||^2 \;+\; ||\mathrm{e}^{n+1} - \mathrm{e}^{n+1/2}||^2\right) = 0
\end{aligned}$$

Adding up 4.29 and 4.31 for $n = 0, \dots, N$, we find:

$$\begin{aligned} |\mathbf{e}^{N+1}|^2 &+ \sum_{n=0}^{N}\Big\{|\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}|^2 + \frac{1}{2}|\mathbf{e}^{n+1/2} - \mathbf{e}^{n}|^2\Big\} \\ &+ k\,\nu \sum_{n=0}^{N}\Big\{||\mathbf{e}^{n+1}||^2 + ||\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}||^2\Big\} \\ &\leq C\,k\Big(\int_0^T t\,||\mathbf{u}_{tt}||_{-1}^2\,dt + k\int_0^T |\mathbf{u}_t|^2\,dt + \sup_{t\in[0,T]}|\nabla p(t)|^2\Big) \\ &+ C\,k\sum_{n=0}^{N}|\mathbf{e}^n|^2 \end{aligned}$$

Applying the discrete Gronwall lemma to the last inequality and using the regularity properties of the solution $(\mathbf{u}, p)$, we obtain:

$$\begin{aligned} |\mathbf{e}^{N+1}|^2 &+ \sum_{n=0}^{N}\Big\{|\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}|^2 + |\mathbf{e}^{n+1/2} - \mathbf{e}^{n}|^2\Big\} \\ &+ k\,\nu \sum_{n=0}^{N}\Big\{||\mathbf{e}^{n+1}||^2 + ||\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}||^2\Big\} \\ &\leq C\,k \end{aligned} \tag{4.32}$$

We still have to prove the bounds for $\mathbf{u}^{n+1/2}$. From 4.31 and the triangle inequality, we get:

$$\begin{aligned} |\mathbf{e}^{N+1/2}|^2 &+ k\,\nu\sum_{n=0}^{N}||\mathbf{e}^{n+1/2}||^2 \\ &\leq |\mathbf{e}^{N+1}|^2 + k\,\nu\,||\mathbf{e}^{N+1}||^2 + |\mathbf{e}^{N+1} - \mathbf{e}^{N+1/2}|^2 \\ &+ k\,\nu\,||\mathbf{e}^{N+1} - \mathbf{e}^{N+1/2}||^2 + 2\,k\,\nu\sum_{n=0}^{N-1}\Big\{||\mathbf{e}^{n+1}||^2 + ||\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}||^2\Big\} \\ &\leq |\mathbf{e}^{N+1}|^2 + 2\,k\,\nu\sum_{n=0}^{N}\Big\{||\mathbf{e}^{n+1}||^2 + ||\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}||^2\Big\} \\ &+ \sum_{n=0}^{N}|\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}|^2 \\ &\leq C\,k \end{aligned}$$

according to 4.32, so that 4.24 follows. □

REMARK 4.4: Lemma 4.7 shows, in particular, that the method provides uniformly stable velocities in $\mathbf{H}_0^1(\Omega)$, that is to say, that there exists a constant $C > 0$ independent of the time step k such that for all $n = 0, \dots, [T/k] - 1$:

$$
\begin{aligned}
||\mathbf{u}^{n+1}|| &\leq C && (4.33)\\
||\mathbf{u}^{n+1/2}|| &\leq C && (4.34)
\end{aligned}
$$

We will use this bounds later on.

We are now in a position to obtain an improved error estimate for the velocity. We will show that $\mathbf{u}^{n+1/2}$ and $\mathbf{u}^{n+1}$ are actually weakly order 1 approximations of the solution in $\mathbf{L}^2(\Omega)$.

Theorem 4.2: *if* **A1** *and* **A2** *hold, and if the Stokes problem is regular, then for* $N = 0, \ldots, [T/k] - 1$ *and small enough* k*:*

$$
k\,\nu \sum_{n=0}^{N} \left(|\mathbf{e}^{n+1}|^2 \; + \; |\mathbf{e}^{n+1/2}|^2 \right) \;\leq\; C\,k^2 \qquad (4.35)
$$

PROOF: let us call $q^{n+1} = p(t_{n+1}) - p^{n+1}$. From 4.15 (with $\theta = 1$) and 4.25, it turns out that:

$$
\begin{aligned}
\frac{1}{k}(\mathbf{e}^{n+1} - \mathbf{e}^{n}) \;&-\; \nu\Delta(\mathbf{e}^{n+1}) \;+\; \nabla q^{n+1} && (4.36)\\
&= \; (\mathbf{u}^n \cdot \nabla)\mathbf{u}^{n+1/2} \;-\; (\mathbf{u}(t_{n+1}) \cdot \nabla)\mathbf{u}(t_{n+1}) \;+\; \mathbf{R}^n
\end{aligned}
$$

We could take the inner product of 4.36 with $2k\mathbf{e}^{n+1}$, which is in Y (and in particular satisfies the proper boundary condition); but then we would need some extra regularity of $\mathbf{e}^{n+1}$, which we cannot prove (see the Appendix). Instead, we take the inner product of 4.36 with $2kA^{-1}\mathbf{e}^{n+1}$, as in [90], and use the self–adjointness of A^{-1} to get:

$$
\begin{aligned}
(\mathbf{e}^{n+1}, A^{-1}\mathbf{e}^{n+1}) \;&-\; (\mathbf{e}^{n}, A^{-1}\mathbf{e}^{n}) \;+\; (\mathbf{e}^{n+1} - \mathbf{e}^{n}, A^{-1}(\mathbf{e}^{n+1} - \mathbf{e}^{n}))\\
&-\; 2\,k\,\nu\,(\Delta\mathbf{e}^{n+1}, A^{-1}\mathbf{e}^{n+1})\\
&=\; 2\,k\,c(\mathbf{u}^{n}, \mathbf{u}^{n+1/2}, A^{-1}\mathbf{e}^{n+1}) \;-\; 2\,k\,c(\mathbf{u}(t_{n+1}), \mathbf{u}(t_{n+1}), A^{-1}\mathbf{e}^{n+1})\\
&+\; 2\,k\; < \mathbf{R}^{n}, A^{-1}\mathbf{e}^{n+1} > \qquad (4.37)
\end{aligned}
$$

The treatment of the term $-2\,k\,\nu\,(\Delta\mathbf{e}^{n+1}, A^{-1}\mathbf{e}^{n+1})$ is simpler in our case than in the standard projection method. In fact, if we take $\mathbf{u} = \mathbf{e}^{n+1}$ in 4.4, we have:

$$
\begin{aligned}
-2\,k\,\nu\,(\Delta\mathbf{e}^{n+1}, A^{-1}\mathbf{e}^{n+1}) \;&=\; 2\,k\,\nu\,(\mathbf{e}^{n+1}, -\Delta(A^{-1}\mathbf{e}^{n+1}))\\
&=\; 2\,k\,\nu\,(\mathbf{e}^{n+1}, \mathbf{e}^{n+1} - \nabla r) \;=\; 2\,k\,\nu\,(\mathbf{e}^{n+1}, \mathbf{e}^{n+1})\\
&=\; 2\,k\,\nu\,|\mathbf{e}^{n+1}|^2
\end{aligned}
$$

since $\nabla \cdot \mathbf{e}^{n+1} = 0$. The right–hand–side terms are bounded as follows. For the Taylor residual term we have:

$$
\begin{aligned}
2\,k < \mathbf{R}^n, A^{-1}\mathbf{e}^{n+1} > \;&\leq\; 2\,k\,||\mathbf{R}^n||_{Y'}\;||A^{-1}\mathbf{e}^{n+1}|| \\
&\leq\; 2\,k\,||\mathbf{e}^{n+1}||_{Y'}\;||\mathbf{R}^n||_{Y'} \\
&\leq\; k\,||\mathbf{e}^{n+1}||^2_{Y'} \;+\; k\,||\mathbf{R}^n||^2_{Y'} \\
&=\; k\,||\mathbf{e}^{n+1}||^2_{Y'} \;+\; k^{-1}\,||\int_{t_n}^{t_{n+1}}(t-t_n)\,\mathbf{u}_{tt}\,dt||^2_{Y'} \\
&\leq\; k\,||\mathbf{e}^{n+1}||^2_{Y'} \;+\; k^{-1}\int_{t_n}^{t_{n+1}}(t-t_n)^2\,dt\int_{t_n}^{t_{n+1}}||\mathbf{u}_{tt}||^2_{Y'}\,dt \\
&\leq\; k\,||\mathbf{e}^{n+1}||^2_{Y'} \;+\; C\,k^2\int_{t_n}^{t_{n+1}}||\mathbf{u}_{tt}||^2_{Y'}\,dt
\end{aligned}
$$

For the nonlinear terms, we use the splitting 4.27 to express them as:

$$
\begin{aligned}
&2k\;\Big(c(\mathbf{u}^n,\mathbf{u}^{n+1/2},A^{-1}\mathbf{e}^{n+1}) \;-\; c(\mathbf{u}(t_{n+1}),\mathbf{u}(t_{n+1}),A^{-1}\mathbf{e}^{n+1})\Big) \\
&=\; 2\,k\,\Big(-c(\mathbf{u}(t_{n+1}),\mathbf{e}^{n+1/2},A^{-1}\mathbf{e}^{n+1}) \;+\; c(\mathbf{u}(t_n)-\mathbf{u}(t_{n+1}),\mathbf{u}^{n+1/2},A^{-1}\mathbf{e}^{n+1}) \\
&-\;\; c(\mathbf{e}^n,\mathbf{u}^{n+1/2},A^{-1}\mathbf{e}^{n+1})\Big)
\end{aligned}
$$

which we call I, II and III, respectively. Then:

$$
\begin{aligned}
\mathrm{I} \;&=\; -2\,k\,c(\mathbf{u}(t_{n+1}),\mathbf{e}^{n+1/2},A^{-1}\mathbf{e}^{n+1}) \\
&=\; 2\,k\,c(\mathbf{u}(t_{n+1}),A^{-1}\mathbf{e}^{n+1},\mathbf{e}^{n+1/2}) \\
&\leq\; C\,k\,||\mathbf{u}(t_{n+1})||_2\;||A^{-1}\mathbf{e}^{n+1}||\;|\mathbf{e}^{n+1/2}| \\
&\leq\; C\,k\,||\mathbf{e}^{n+1}||_{Y'}\;|\mathbf{e}^{n+1/2}| \\
&\leq\; C\,k\,||\mathbf{e}^{n+1}||^2_{Y'} \;+\; \frac{k\nu}{4}|\mathbf{e}^{n+1/2}|^2 \\
&=\; C\,k\,||\mathbf{e}^{n+1}||^2_{Y'} \;+\; \frac{k\nu}{4}\Big\{|\mathbf{e}^{n+1}|^2 \;+\; |\mathbf{e}^{n+1}-\mathbf{e}^{n+1/2}|^2 \\
&+\; k\nu||\mathbf{e}^{n+1}||^2 \;+\; k\nu||\mathbf{e}^{n+1}-\mathbf{e}^{n+1/2}||^2 \;-\; k\nu||\mathbf{e}^{n+1/2}||^2\Big\}
\end{aligned}
$$

where we have used 4.31.

$$
\begin{aligned}
\mathrm{II} \;&=\; 2\,k\,c(\mathbf{u}(t_n)-\mathbf{u}(t_{n+1}),\mathbf{u}^{n+1/2},A^{-1}\mathbf{e}^{n+1}) \\
&\leq\; C\,k\,|\mathbf{u}(t_n)-\mathbf{u}(t_{n+1})|\;||\mathbf{u}^{n+1/2}||\;||A^{-1}\mathbf{e}^{n+1}||_2 \\
&\leq\; C\,k\,|\int_{t_n}^{t_{n+1}}\mathbf{u}_t\,dt|\;|\mathbf{e}^{n+1}| \\
&\leq\; C\,k\,|\int_{t_n}^{t_{n+1}}\mathbf{u}_t\,dt|^2 \;+\; \frac{k\nu}{4}|\mathbf{e}^{n+1}|^2 \\
&\leq\; C\,k^2\int_{t_n}^{t_{n+1}}|\mathbf{u}_t|^2\,dt \;+\; \frac{k\nu}{4}|\mathbf{e}^{n+1}|^2
\end{aligned}
$$

where we have used the bound 4.34. Finally:

$$\begin{aligned}
\mathrm{III} &= -2\,k\,c(\mathbf{e}^n, \mathbf{u}^{n+1/2}, A^{-1}\mathbf{e}^{n+1}) \\
&= 2\,k\,c(\mathbf{e}^n, A^{-1}\mathbf{e}^{n+1}, \mathbf{u}^{n+1/2}) \\
&= 2\,k\,c(\mathbf{e}^n, A^{-1}\mathbf{e}^{n+1}, \mathbf{u}(t_{n+1})) \\
&- 2\,k\,c(\mathbf{e}^n, A^{-1}\mathbf{e}^{n+1}, \mathbf{e}^{n+1/2}) \\
&= \mathrm{III}_a + \mathrm{III}_b
\end{aligned}$$

so that:

$$\begin{aligned}
\mathrm{III}_a &= 2\,k\,c(\mathbf{e}^n, A^{-1}\mathbf{e}^{n+1}, \mathbf{u}(t_{n+1})) \\
&\leq C\,k\,|\mathbf{e}^n|\;||\mathbf{u}(t_{n+1})||_2\;||A^{-1}\mathbf{e}^{n+1}|| \\
&\leq C\,k\,|\mathbf{e}^n|\;||\mathbf{e}^{n+1}||_{Y'} \\
&\leq C\,k\left(|\mathbf{e}^{n+1}| + |\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}| + |\mathbf{e}^{n+1/2} - \mathbf{e}^n|\right)\;||\mathbf{e}^{n+1}||_{Y'} \\
&\leq \frac{k\nu}{4}|\mathbf{e}^{n+1}|^2 + C\,k\left(|\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}|^2 + |\mathbf{e}^{n+1/2} - \mathbf{e}^n|^2\right) \\
&+ C\,k\,||\mathbf{e}^{n+1}||^2_{Y'}
\end{aligned}$$

and:

$$\begin{aligned}
\mathrm{III}_b &= -2\,k\,c(\mathbf{e}^n, A^{-1}\mathbf{e}^{n+1}, \mathbf{e}^{n+1/2}) \\
&\leq C\,k\,|\mathbf{e}^n|\;||A^{-1}\mathbf{e}^{n+1}||_2\;||\mathbf{e}^{n+1/2}|| \\
&\leq C\,k\,|\mathbf{e}^n|\;|\mathbf{e}^{n+1}|\;||\mathbf{e}^{n+1/2}|| \\
&\leq C\,k^{3/2}\,|\mathbf{e}^{n+1}|\;||\mathbf{e}^{n+1/2}|| \\
&\leq \frac{k\nu}{4}|\mathbf{e}^{n+1}|^2 + C\,k^2\,||\mathbf{e}^{n+1/2}||^2
\end{aligned}$$

since, according to Lemma 4.7, $|\mathbf{e}^n| \leq Ck^{1/2}$. All these inequalities yield:

$$\begin{aligned}
(\mathbf{e}^{n+1}, A^{-1}\mathbf{e}^{n+1}) &- (\mathbf{e}^n, A^{-1}\mathbf{e}^n) + (\mathbf{e}^{n+1} - \mathbf{e}^n, A^{-1}(\mathbf{e}^{n+1} - \mathbf{e}^n)) \\
&+ k\,\nu\,|\mathbf{e}^{n+1}|^2 \\
&\leq C\,k^2 \int_{t_n}^{t_{n+1}} ||\mathbf{u}_{tt}||^2_{Y'}\,dt + C\,k^2 \int_{t_n}^{t_{n+1}} |\mathbf{u}_t|^2\,dt \\
&+ C\,k\,||\mathbf{e}^{n+1}||^2_{Y'} + C\,k^2\,||\mathbf{e}^{n+1}||^2 \\
&+ C\,k\,|\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}|^2 + C\,k\,|\mathbf{e}^{n+1/2} - \mathbf{e}^n|^2 \\
&+ C\,k^2\,||\mathbf{e}^{n+1} - \mathbf{e}^{n+1/2}||^2 + C\,k^2\,||\mathbf{e}^{n+1/2}||^2 \qquad (4.38)
\end{aligned}$$

Adding up 4.38 for $n = 0, \ldots, N$, we get:

$$\begin{aligned}
(\mathbf{e}^{N+1}, A^{-1}\mathbf{e}^{N+1}) \;&+\; \sum_{n=0}^{N}(\mathbf{e}^{n+1}-\mathbf{e}^{n}, A^{-1}(\mathbf{e}^{n+1}-\mathbf{e}^{n})) \\
&+\; k\,\nu \sum_{n=0}^{N} |\mathbf{e}^{n+1}|^2 \\
&\leq\; C\,k^2 \int_0^T ||\mathbf{u}_{tt}||_{Y'}^2\,dt \;+\; C\,k^2 \int_0^T |\mathbf{u}_t|^2\,dt \\
&+\; C\,k \sum_{n=0}^{N} ||\mathbf{e}^{n+1}||_{Y'}^2 \;+\; C\,k^2 \sum_{n=0}^{N} ||\mathbf{e}^{n+1}||^2 \\
&+\; C\,k \sum_{n=0}^{N} |\mathbf{e}^{n+1}-\mathbf{e}^{n+1/2}|^2 \;+\; C\,k \sum_{n=0}^{N} |\mathbf{e}^{n+1/2}-\mathbf{e}^{n}|^2 \\
&+\; C\,k^2 \sum_{n=0}^{N} ||\mathbf{e}^{n+1}-\mathbf{e}^{n+1/2}||^2 \;+\; C\,k^2 \sum_{n=0}^{N} ||\mathbf{e}^{n+1/2}||^2
\end{aligned}$$

Using now 4.7, the regularity properties **R1** and **R4** of the continuous solution and the estimates of Lemma 4.7, we get:

$$\begin{aligned}
||\mathbf{e}^{N+1}||_{Y'}^2 \;&+\; \sum_{n=0}^{N} ||\mathbf{e}^{n+1}-\mathbf{e}^{n}||_{Y'}^2 \;+\; k\,\nu \sum_{n=0}^{N} |\mathbf{e}^{n+1}|^2 \\
&\leq\; C\,k^2 \;+\; C\,k \sum_{n=0}^{N} ||\mathbf{e}^{n+1}||_{Y'}^2
\end{aligned}$$

For sufficiently small k, we can apply the discrete Gronwall lemma to the last inequality, and we get:

$$\begin{aligned}
||\mathbf{e}^{N+1}||_{Y'}^2 \;&+\; \sum_{n=0}^{N} ||\mathbf{e}^{n+1}-\mathbf{e}^{n}||_{Y'}^2 \;+\; k\,\nu \sum_{n=0}^{N} |\mathbf{e}^{n+1}|^2 \\
&\leq\; C\,k^2
\end{aligned}$$

and the estimate for $\mathbf{u}^{n+1}$ is proved. For $\mathbf{u}^{n+1/2}$, we have:

$$\begin{aligned}
k\,\nu \sum_{n=0}^{N} |\mathbf{e}^{n+1/2}|^2 \;&\leq\; 2\,k\,\nu \sum_{n=0}^{N}\Big(|\mathbf{e}^{n+1}|^2 + |\mathbf{e}^{n+1}-\mathbf{e}^{n+1/2}|^2\Big) \\
&\leq\; C\,k^2
\end{aligned}$$

due to Lemma 4.7 and the estimate for $\mathbf{u}^{n+1}$, so that 4.35 is proved. □

REMARK 4.5: Shen's proof of 4.35 for the standard projection method in [90] is not quite correct. Apart from the corrections pointed out in [50], he uses the equivalence between $||\mathbf{u}||_{-1}$ and $(A^{-1}\mathbf{u},\mathbf{u})^{1/2}$ as norms on H in an improper way. Apparently, he bounds:

$$||\mathbf{e}^{n+1}||^2_{-1} \ - \ ||\mathbf{e}^n||^2_{-1} \ + \ ||\mathbf{e}^{n+1} - \mathbf{e}^n||^2_{-1}$$

by:

$$C\,(\mathbf{e}^{n+1} - \mathbf{e}^n, 2A^{-1}\mathbf{e}^{n+1}),$$

which cannot be deduced from 4.6.

REMARK 4.6: since we are assuming that the domain Ω is smooth enough for the Stokes problem to be regular, we can assure that our semidiscrete velocities $\mathbf{u}^{n+1/2}$ and $\mathbf{u}^{n+1}$, which are solutions of elliptic problems on Ω, actually belong to $\mathbf{H}^2(\Omega)$. We can improve our error estimates for $\mathbf{u}^{n+1}$ to strongly first order in $\mathbf{L}^2(\Omega)$ and weakly first order in $\mathbf{H}^1_0(\Omega)$ by assuming that $\mathbf{u}^{n+1/2}$ is uniformly bounded in $\mathbf{H}^2(\Omega)$, that is, bounded by a constant independent of k. But we cannot prove this assumption, so we keep our weak order 1 error estimates (which are the same as those obtained in [90] for the standard projection method) and give this improvement in an Appendix.

REMARK 4.7: in Theorem 4.2 we have proved that, in particular:

$$\sum_{n=0}^{N} ||\mathbf{e}^{n+1} - \mathbf{e}^n||^2_{Y'} \ \leq \ C\,k^2$$

But to get some pressure error estimates we would need this inequality in terms of the norm $||.||_{-1}$. This is the reason why the proof presented in [90] of weakly order 1/2 pressure error estimates for the classical projection method is not correct, as explained in [92]; in this last reference this proof is modified for the linear Stokes problem, dropping the nonlinear terms. We could also do so very easily here, but prefer to put off this question to the Appendix, when, using the improved error estimates proved there, we can deal with the full nonlinear problem and still obtain weakly order 1/2 error estimates for the pressure in $L^2_0(\Omega)$.

4.3 A pressure–correction method

4.3.1 Development of the method

We now modify the viscosity splitting method of the previous Section to account for pressure correction. This modified scheme will let us study the predictor–multicorrector algorithm in Chapter 5. Another advantage of using a pressure correction method will be explained in Subsection 4.3.3.

As was seen in some of the methods presented in Section 1.5, such as those of [90], [62] or [81], the basic idea of pressure correction consists of including a pressure gradient term in the first step of the method evaluated at the previous time step, and regarding the Lagrange multiplier of the second step as a pressure increment, rather than an end–of–step pressure in itself. The scheme is started by an arbitrary pressure $\tilde{p}^0$, which we assume belongs to $H^1(\Omega)$; then, the scheme reads:

First step: Given $\tilde{\mathbf{u}}^n \in Y$ and $\tilde{p}^n \in L^2_0(\Omega)$, we seek $\tilde{\mathbf{u}}^{n+1/2}$ such that:

$$\begin{aligned} \frac{\tilde{\mathbf{u}}^{n+1/2}-\tilde{\mathbf{u}}^n}{\delta t} - \theta\nu\Delta\tilde{\mathbf{u}}^{n+1/2} - (1-\theta)\nu\Delta\tilde{\mathbf{u}}^n &+ (\tilde{\mathbf{u}}^n\cdot\nabla)\tilde{\mathbf{u}}^{n+1/2} + \nabla\tilde{p}^n \\ &= \mathbf{f}(t_{n+1}) \qquad (4.39) \\ \tilde{\mathbf{u}}^{n+1/2}|_\Gamma &= 0 \end{aligned}$$

The weak form of this problem consists of finding $\tilde{\mathbf{u}}^{n+1/2} \in \mathbf{H}^1_0(\Omega)$ such that:

$$a^n_\theta(\tilde{\mathbf{u}}^{n+1/2},\mathbf{v}) = \tilde{l}^n_1(\mathbf{v}), \qquad \forall \mathbf{v}\in\mathbf{H}^1_0(\Omega) \qquad (4.40)$$

where now $\tilde{l}^n_1(\mathbf{v}) = l_1(\mathbf{v}) - b(\mathbf{v},\tilde{p}^n)$. Since this linear form is also continuous in $\mathbf{H}^1_0(\Omega)$, once more we have existence and uniqueness of a solution $\tilde{\mathbf{u}}^{n+1/2}$ due to the Lax–Milgram theorem.

Second step: Given now $\tilde{\mathbf{u}}^{n+1/2} \in \mathbf{H}^1_0(\Omega)$, we look for $\tilde{\mathbf{u}}^{n+1}$ and $\tilde{p}^{n+1}$ such that:

$$\begin{aligned} \frac{\tilde{\mathbf{u}}^{n+1}-\tilde{\mathbf{u}}^{n+1/2}}{\delta t} - \theta\nu\Delta(\tilde{\mathbf{u}}^{n+1}-\tilde{\mathbf{u}}^{n+1/2}) &+ \phi\nabla(\tilde{p}^{n+1}-\tilde{p}^n) = 0 \quad (4.41) \\ \nabla\cdot\tilde{\mathbf{u}}^{n+1} &= 0 \qquad (4.42) \\ \tilde{\mathbf{u}}^{n+1}|_\Gamma &= 0 \qquad (4.43) \end{aligned}$$

where $\phi > 0$ is an arbitrary parameter. The weak form of this problem consists of finding $\tilde{\mathbf{u}}^{n+1}$ and $\tilde{s}^{n+1} = \delta t\,\phi\,(\tilde{p}^{n+1}-\tilde{p}^n)$ such that:

$$\begin{aligned} a_\theta(\tilde{\mathbf{u}}^{n+1},\mathbf{v}) + b(\mathbf{v},\tilde{s}^{n+1}) &= l_2(v), \quad \forall\mathbf{v}\in\mathbf{H}^1_0(\Omega) \\ b(\tilde{\mathbf{u}}^{n+1},q) &= 0, \quad \forall q\in L^2_0(\Omega) \qquad (4.44) \end{aligned}$$

which is again a mixed problem.

By adding 4.39 and 4.41 we find:

$$\begin{aligned} \frac{\tilde{\mathbf{u}}^{n+1}-\tilde{\mathbf{u}}^n}{\delta t} - \theta\nu\Delta\tilde{\mathbf{u}}^{n+1} - (1-\theta)\nu\Delta\tilde{\mathbf{u}}^n &+ (\tilde{\mathbf{u}}^n\cdot\nabla)\tilde{\mathbf{u}}^{n+1/2} + \phi\nabla\tilde{p}^{n+1} \\ + (1-\phi)\tilde{p}^n &= \mathbf{f}(t_{n+1}) \qquad (4.45) \end{aligned}$$

where the implicit, but not necessarily fully implicit, character of the approximation of the pressure gradient term can be observed. This allows to choose $\phi = \theta$ to keep the same order of approximation in all the terms of the equation, which is specially relevant if one takes $\theta = 1/2$ to get a second order method (we will see how to obtain second order accuracy in a different way in the next Chapter).

A pressure Poisson equation similar to 4.16 can also be developed in this case, this time for the pressure increment $(\tilde{p}^{n+1} - \tilde{p}^n)$; but again it seems impractical to use it.

The second step of the scheme can still be written as a generalized projection: $\tilde{\mathbf{u}}^{n+1} = P_Y(\tilde{\mathbf{u}}^{n+1/2})$; the Lagrange multiplier associated with it is $\delta t\, \phi(\tilde{p}^{n+1} - \tilde{p}^n)$ this time, rather than $\delta t\, \tilde{p}^{n+1}$ alone.

For this modified scheme we can prove the same first order error estimates for the end–of–step and intermediate velocities as for the original one. Under the assumption of uniformly bounded velocities in $\mathbf{H}^2(\Omega)$, we can improve the estimates for the end–of–step velocities (see the Appendix).

4.3.2 Error estimates

We present an error analysis of our viscosity splitting, pressure correction fractional step method with parameters $\theta = 1$ and arbitrary ϕ, which is similar to that of the previous Section. However, we will need some extra regularity assumption on the semidiscrete pressure solution, which will be stated in what follows.

We define the velocity error functions for this method as:

$$\begin{aligned} \tilde{\mathbf{e}}^{n+1} &= \mathbf{u}(t_{n+1}) - \tilde{\mathbf{u}}^{n+1} \\ \tilde{\mathbf{e}}^{n+1/2} &= \mathbf{u}(t_{n+1}) - \tilde{\mathbf{u}}^{n+1/2} \end{aligned}$$

We give a first estimate for $\tilde{\mathbf{e}}^{n+1}$ and $\tilde{\mathbf{e}}^{n+1/2}$ which shows that both $\tilde{\mathbf{u}}^{n+1}$ and $\tilde{\mathbf{u}}^{n+1/2}$ are strongly order $1/2$ approximations to $\mathbf{u}(t_{n+1})$ in $\mathbf{L}^2(\Omega)$ and weakly order $1/2$ in $\mathbf{H}^1_0(\Omega)$, in a similar way to Lemma 4.7. Since the domain is regular, the solution of the mixed problem 4.41 actually satisfies $\tilde{\mathbf{u}}^{n+1} \in \mathbf{H}^2(\Omega)$ and $(\tilde{p}^{n+1} - \tilde{p}^n) \in H^1(\Omega)$; since we are assuming that $\tilde{p}^0 \in H^1(\Omega)$, this implies by induction that $\tilde{p}^{n+1} \in H^1(\Omega)$. We will assume that the norm of ∇p^{n+1} is bounded uniformly in $\mathbf{L}^2(\Omega)$, that is, that there exists a constant $C > 0$ independent of k such that:

$$|\nabla \tilde{p}^{n+1}| \leq C, \qquad \forall n \geq 0 \tag{4.46}$$

We then have:

<u>Lemma 4.8:</u> *assume that* **A1** *and* **A2** *hold, that the Stokes problem is regular and that* 4.46 *also holds; then for* $N = 0, \ldots, [T/k] - 1$*:*

$$
\begin{aligned}
|\tilde{\mathrm{e}}^{N+1}|^2 \; &+ \; |\tilde{\mathrm{e}}^{N+1/2}|^2 + \sum_{n=0}^{N}\Big\{|\tilde{\mathrm{e}}^{n+1} - \tilde{\mathrm{e}}^{n+1/2}|^2 + |\tilde{\mathrm{e}}^{n+1/2} - \tilde{\mathrm{e}}^{n}|^2\Big\} \qquad (4.47)\\
&+ \; k\,\nu \sum_{n=0}^{N}\Big\{||\tilde{\mathrm{e}}^{n+1}||^2 + ||\tilde{\mathrm{e}}^{n+1/2}||^2 + ||\tilde{\mathrm{e}}^{n+1} - \tilde{\mathrm{e}}^{n+1/2}||^2\Big\} \le C\,k
\end{aligned}
$$

PROOF: the proof is similar to that of Lemma 4.7. Subtracting 4.39 from 4.25, we get:

$$
\begin{aligned}
\frac{1}{k}(\tilde{\mathrm{e}}^{n+1/2} - \tilde{\mathrm{e}}^{n}) \; - \; \nu\Delta(\tilde{\mathrm{e}}^{n+1/2}) \; &= \; (\tilde{\mathbf{u}}^{n}\cdot\nabla)\tilde{\mathbf{u}}^{n+1/2} \; - \; (\mathbf{u}(t_{n+1})\cdot\nabla)\mathbf{u}(t_{n+1})\\
&+ \; \mathbf{R}^{n} \; + \; \nabla(\tilde{p}^{n} - p(t_{n+1})) \qquad (4.48)
\end{aligned}
$$

Taking the inner product of 4.48 with $2k\tilde{\mathrm{e}}^{n+1/2}$ and using the splitting 4.27 of the nonlinear terms, we obtain:

$$
\begin{aligned}
|\tilde{\mathrm{e}}^{n+1/2}|^2 \; &- \; |\tilde{\mathrm{e}}^{n}|^2 + 2\,k\,\nu\,||\tilde{\mathrm{e}}^{n+1/2}||^2 + |\tilde{\mathrm{e}}^{n+1/2} - \tilde{\mathrm{e}}^{n}|^2 \qquad (4.49)\\
&= \; 2\,k\, < \mathbf{R}^{n}, \tilde{\mathrm{e}}^{n+1/2} > \; + 2\,k\,(\nabla(\tilde{p}^{n} - p(t_{n+1})), \tilde{\mathrm{e}}^{n+1/2})\\
&- \; 2\,k\,c(\tilde{\mathrm{e}}^{n}, \tilde{\mathbf{u}}^{n+1/2}, \tilde{\mathrm{e}}^{n+1/2}) \; + \; 2\,k\,c(\mathbf{u}(t_n) - \mathbf{u}(t_{n+1}), \tilde{\mathbf{u}}^{n+1/2}, \tilde{\mathrm{e}}^{n+1/2})\\
&- \; 2\,k\,c(\mathbf{u}(t_{n+1}), \tilde{\mathrm{e}}^{n+1/2}, \tilde{\mathrm{e}}^{n+1/2})
\end{aligned}
$$

We bound each term in the RHS of 4.49 as in Lemma 4.7:

$$
\begin{aligned}
2\,k\, < \mathbf{R}^{n}, \tilde{\mathrm{e}}^{n+1/2} > \; &\le \; \frac{k\nu}{3}||\tilde{\mathrm{e}}^{n+1/2}||^2 \; + \; C\,k \int_{t_n}^{t_{n+1}} t\,||\mathbf{u}_{tt}||_{-1}^2\,dt\\
-2\,k\,c(\tilde{\mathrm{e}}^{n}, \tilde{\mathbf{u}}^{n+1/2}, \tilde{\mathrm{e}}^{n+1/2}) \; &\le \; \frac{k\nu}{3}||\tilde{\mathrm{e}}^{n+1/2}||^2 \; + \; C\,k\,|\tilde{\mathrm{e}}^{n}|^2\\
2\,k\,c(\mathbf{u}(t_n) - \mathbf{u}(t_{n+1}), \tilde{\mathbf{u}}^{n+1/2}, \tilde{\mathrm{e}}^{n+1/2}) \; &\le \; \frac{k\nu}{3}||\tilde{\mathrm{e}}^{n+1/2}||^2 \; + \; C\,k^2 \int_{t_n}^{t_{n+1}} |\mathbf{u}_t|^2\,dt\\
-2\,k\,c(\mathbf{u}(t_{n+1}), \tilde{\mathrm{e}}^{n+1/2}, \tilde{\mathrm{e}}^{n+1/2}) \; &= \; 0
\end{aligned}
$$

As for the pressure gradient term, we use 4.46 to get:

$$
\begin{aligned}
2\,k\,(\nabla(\tilde{p}^{n} - p(t_{n+1})), \tilde{\mathrm{e}}^{n+1/2}) \; &= \; 2\,k\,(\nabla(\tilde{p}^{n} - p(t_{n+1})), \tilde{\mathrm{e}}^{n+1/2} - \tilde{\mathrm{e}}^{n})\\
&\le \; 2\,k\,|\nabla(\tilde{p}^{n} - p(t_{n+1}))|\;|\tilde{\mathrm{e}}^{n+1/2} - \tilde{\mathrm{e}}^{n}|\\
&\le \; \frac{1}{2}|\tilde{\mathrm{e}}^{n+1/2} - \tilde{\mathrm{e}}^{n}|^2 \; + \; C\,k^2\left(|\nabla\tilde{p}^{n}|^2 + |\nabla p(t_{n+1})|^2\right)\\
&\le \; \frac{1}{2}|\tilde{\mathrm{e}}^{n+1/2} - \tilde{\mathrm{e}}^{n}|^2 \; + \; C\,k^2
\end{aligned}
$$

since $\nabla\cdot\tilde{\mathrm{e}}^{n} = 0$.

From all these inequalities we deduce:

$$\begin{aligned}|\tilde{\mathbf{e}}^{n+1/2}|^2 \;&-\; |\tilde{\mathbf{e}}^n|^2 \;+\; k\,\nu\,||\tilde{\mathbf{e}}^{n+1/2}||^2 \;+\; \frac{1}{2}|\tilde{\mathbf{e}}^{n+1/2}-\tilde{\mathbf{e}}^n|^2\\ &\le\; C\,k\int_{t_n}^{t_{n+1}} t\,||\mathbf{u}_{tt}||_{-1}^2\,dt \;+\; C\,k^2\int_{t_n}^{t_{n+1}}|\mathbf{u}_t|^2\,dt \qquad (4.50)\\ &+\; C\,k^2 \;+\; C\,k\,|\tilde{\mathbf{e}}^n|^2\end{aligned}$$

From 4.41 we now have:

$$\frac{\tilde{\mathbf{e}}^{n+1}-\tilde{\mathbf{e}}^{n+1/2}}{k} \;-\; \nu\Delta(\tilde{\mathbf{e}}^{n+1}-\tilde{\mathbf{e}}^{n+1/2}) \;-\; \phi\nabla(\tilde{p}^{n+1}-\tilde{p}^n) \;=\; 0 \qquad (4.51)$$

Taking the inner product of 4.51 with $2k\tilde{\mathbf{e}}^{n+1}$, given that $\nabla\cdot\tilde{\mathbf{e}}^{n+1}=0$ and that $\tilde{\mathbf{e}}^{n+1}_{|\Gamma}=0$, we get an equality similar to 4.31, namely:

$$\begin{aligned}|\tilde{\mathbf{e}}^{n+1}|^2 \;-\; |\tilde{\mathbf{e}}^{n+1/2}|^2 \;&+\; |\tilde{\mathbf{e}}^{n+1}-\tilde{\mathbf{e}}^{n+1/2}|^2 \qquad (4.52)\\ &+\; k\,\nu\left(||\tilde{\mathbf{e}}^{n+1}||^2-||\tilde{\mathbf{e}}^{n+1/2}||^2 \;+\; ||\tilde{\mathbf{e}}^{n+1}-\tilde{\mathbf{e}}^{n+1/2}||^2\right)=0\end{aligned}$$

Adding up 4.50 and 4.52 for $n=0,\ldots,N$, we find:

$$\begin{aligned}|\tilde{\mathbf{e}}^{N+1}|^2 \;&+\; \sum_{n=0}^{N}\Big\{|\tilde{\mathbf{e}}^{n+1}-\tilde{\mathbf{e}}^{n+1/2}|^2 \;+\; \frac{1}{2}|\tilde{\mathbf{e}}^{n+1/2}-\tilde{\mathbf{e}}^n|^2\Big\}\\ &+\; k\,\nu\sum_{n=0}^{N}\Big\{||\tilde{\mathbf{e}}^{n+1}||^2 \;+\; ||\tilde{\mathbf{e}}^{n+1}-\tilde{\mathbf{e}}^{n+1/2}||^2\Big\}\\ &\le\; C\,k\left(\int_0^T t\,||\mathbf{u}_{tt}||_{-1}^2\,dt \;+\; k\int_0^T|\mathbf{u}_t|^2\,dt \;+\; k\right)\\ &+\; C\,k\sum_{n=0}^{N}|\tilde{\mathbf{e}}^n|^2\end{aligned}$$

Applying the discrete Gronwall lemma to the last inequality and using the regularity properties of the solution $\mathbf{u}$, we obtain:

$$\begin{aligned}|\tilde{\mathbf{e}}^{N+1}|^2 \;&+\; \sum_{n=0}^{N}\Big\{|\tilde{\mathbf{e}}^{n+1}-\tilde{\mathbf{e}}^{n+1/2}|^2 \;+\; |\tilde{\mathbf{e}}^{n+1/2}-\tilde{\mathbf{e}}^n|^2\Big\}\\ &+\; k\,\nu\sum_{n=0}^{N}\Big\{||\tilde{\mathbf{e}}^{n+1}||^2 \;+\; ||\tilde{\mathbf{e}}^{n+1}-\tilde{\mathbf{e}}^{n+1/2}||^2\Big\} \qquad (4.53)\\ &\le\; C\,k\end{aligned}$$

We still have to prove the bounds for $\tilde{\mathbf{u}}^{n+1/2}$. Once again, from 4.52 and the triangle inequality, we get:

$$\begin{aligned}
|\tilde{\mathbf{e}}^{N+1/2}|^2 \;&+\; k\,\nu \sum_{n=0}^{N} ||\tilde{\mathbf{e}}^{n+1/2}||^2 \\
&\leq\; |\tilde{\mathbf{e}}^{N+1}|^2 \;+\; k\,\nu\, ||\tilde{\mathbf{e}}^{N+1}||^2 \;+\; |\tilde{\mathbf{e}}^{N+1} - \tilde{\mathbf{e}}^{N+1/2}|^2 \\
&+\; k\,\nu\, ||\tilde{\mathbf{e}}^{N+1} - \tilde{\mathbf{e}}^{N+1/2}||^2 \;+\; 2\,k\,\nu \sum_{n=0}^{N-1} \Big\{ ||\tilde{\mathbf{e}}^{n+1}||^2 + ||\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^{n+1/2}||^2 \Big\} \\
&\leq\; |\tilde{\mathbf{e}}^{N+1}|^2 \;+\; 2\,k\,\nu \sum_{n=0}^{N} \Big\{ ||\tilde{\mathbf{e}}^{n+1}||^2 + ||\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^{n+1/2}||^2 \Big\} \\
&+\; \sum_{n=0}^{N} |\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^{n+1/2}|^2 \\
&\leq\; C\,k
\end{aligned}$$

according to 4.53, so that 4.47 follows. □

Once again we have, in particular, uniformly stable velocities in $\mathbf{H}_0^1(\Omega)$, which will be used later on. We now improve these estimates to weakly first order, in a similar way to Theorem 4.2.

Theorem 4.3: *if* **A1** *and* **A2** *hold, if the Stokes problem is regular and if* 4.46 *also holds, then for* $N = 0, \ldots, [T/k] - 1$ *and small enough* k*:*

$$k\,\nu \sum_{n=0}^{N} \Big(|\tilde{\mathbf{e}}^{n+1}|^2 \;+\; |\tilde{\mathbf{e}}^{n+1/2}|^2 \Big) \;\leq\; C\,k^2 \tag{4.54}$$

PROOF: the proof is similar to that of Theorem 4.2. Let us call $\tilde{q}^{n+1} = p(t_{n+1}) - \phi\tilde{p}^{n+1} - (1-\phi)\tilde{p}^n$. From 4.45 (with $\theta = 1$) and 4.25, we get:

$$\begin{aligned}
\frac{1}{k}(\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^n) \;&-\; \nu\Delta(\tilde{\mathbf{e}}^{n+1}) \;+\; \nabla\tilde{q}^{n+1} \\
&=\; (\mathbf{u}^n \cdot \nabla)\mathbf{u}^{n+1/2} \;-\; (\mathbf{u}(t_{n+1}) \cdot \nabla)\mathbf{u}(t_{n+1}) \;+\; \mathbf{R}^n
\end{aligned} \tag{4.55}$$

Once again, we could take the inner product of 4.55 with $2k\tilde{\mathbf{e}}^{n+1}$, which is in Y (and satisfies the proper boundary condition); but then we would need some extra regularity of $\mathbf{e}^{n+1}$, which we cannot prove (see the Appendix). Instead, we take the inner product of 4.55 with $2kA^{-1}\tilde{\mathbf{e}}^{n+1}$, to get:

$$\begin{aligned}
(\tilde{\mathbf{e}}^{n+1}, A^{-1}\tilde{\mathbf{e}}^{n+1}) \;&-\; (\tilde{\mathbf{e}}^n, A^{-1}\tilde{\mathbf{e}}^n) \;+\; (\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^n, A^{-1}(\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^n)) \\
&-\; 2\,k\,\nu\,(\Delta\tilde{\mathbf{e}}^{n+1}, A^{-1}\tilde{\mathbf{e}}^{n+1}) \\
&=\; 2\,k\,c(\tilde{\mathbf{u}}^n, \tilde{\mathbf{u}}^{n+1/2}, A^{-1}\tilde{\mathbf{e}}^{n+1}) \;-\; 2\,k\,c(\mathbf{u}(t_{n+1}), \mathbf{u}(t_{n+1}), A^{-1}\tilde{\mathbf{e}}^{n+1}) \\
&+\; 2\,k \,<\mathbf{R}^n, A^{-1}\tilde{\mathbf{e}}^{n+1}>
\end{aligned} \tag{4.56}$$

The same treatment as in Theorem 4.2 is given to all the terms in 4.56, yielding:

$$\begin{aligned}
-2\,k\,\nu\,(\Delta\tilde{\mathbf{e}}^{n+1}, A^{-1}\tilde{\mathbf{e}}^{n+1}) &= 2\,k\,\nu\,|\tilde{\mathbf{e}}^{n+1}|^2 \\
2\,k\, <\mathbf{R}^n, A^{-1}\tilde{\mathbf{e}}^{n+1}> &\leq k\,||\tilde{\mathbf{e}}^{n+1}||^2_{Y'} + C\,k^2 \int_{t_n}^{t_{n+1}} ||\mathbf{u}_{tt}||^2_{Y'}\,dt \\
-2\,k\,c(\mathbf{u}(t_{n+1}), \tilde{\mathbf{e}}^{n+1/2}, A^{-1}\tilde{\mathbf{e}}^{n+1}) &= C\,k\,||\tilde{\mathbf{e}}^{n+1}||^2_{Y'} + \frac{k\nu}{4}\Big\{|\tilde{\mathbf{e}}^{n+1}|^2 \\
&+ |\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^{n+1/2}|^2 + k\nu||\tilde{\mathbf{e}}^{n+1}||^2 \\
&+ k\nu||\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^{n+1/2}||^2 - k\nu||\tilde{\mathbf{e}}^{n+1/2}||^2\Big\} \\
2\,k\,c(\mathbf{u}(t_n) - \mathbf{u}(t_{n+1}), \tilde{\mathbf{u}}^{n+1/2}, A^{-1}\tilde{\mathbf{e}}^{n+1}) &\leq C\,k^2 \int_{t_n}^{t_{n+1}} |\mathbf{u}_t|^2\,dt + \frac{k\nu}{4}|\tilde{\mathbf{e}}^{n+1}|^2 \\
-2\,k\,c(\tilde{\mathbf{e}}^n, \tilde{\mathbf{u}}^{n+1/2}, A^{-1}\tilde{\mathbf{e}}^{n+1}) &\leq \frac{k\nu}{4}|\tilde{\mathbf{e}}^{n+1}|^2 \\
&+ C\,k\left(|\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^{n+1/2}|^2 + |\tilde{\mathbf{e}}^{n+1/2} - \tilde{\mathbf{e}}^n|^2\right) \\
&+ C\,k\,||\tilde{\mathbf{e}}^{n+1}||^2_{Y'} \\
&+ \frac{k\nu}{4}|\tilde{\mathbf{e}}^{n+1}|^2 + C\,k^2\,||\tilde{\mathbf{e}}^{n+1/2}||^2
\end{aligned}$$

These inequalities yield:

$$\begin{aligned}
(\tilde{\mathbf{e}}^{n+1}, A^{-1}\tilde{\mathbf{e}}^{n+1}) &- (\tilde{\mathbf{e}}^n, A^{-1}\tilde{\mathbf{e}}^n) + (\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^n, A^{-1}(\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^n)) \\
&+ k\,\nu\,|\tilde{\mathbf{e}}^{n+1}|^2 \\
&\leq C\,k^2 \int_{t_n}^{t_{n+1}} ||\mathbf{u}_{tt}||^2_{Y'}\,dt + C\,k^2 \int_{t_n}^{t_{n+1}} |\mathbf{u}_t|^2\,dt \\
&+ C\,k\,||\tilde{\mathbf{e}}^{n+1}||^2_{Y'} + C\,k^2\,||\tilde{\mathbf{e}}^{n+1}||^2 \\
&+ C\,k\,|\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^{n+1/2}|^2 + C\,k\,|\tilde{\mathbf{e}}^{n+1/2} - \tilde{\mathbf{e}}^n|^2 \\
&+ C\,k^2\,||\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^{n+1/2}||^2 + C\,k^2\,||\tilde{\mathbf{e}}^{n+1/2}||^2 \qquad (4.57)
\end{aligned}$$

Adding up 4.57 for $n = 0, \ldots, N$, we get:

$$\begin{aligned}
(\tilde{\mathbf{e}}^{N+1}, A^{-1}\tilde{\mathbf{e}}^{N+1}) &+ \sum_{n=0}^{N}(\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^n, A^{-1}(\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^n)) \\
&+ k\,\nu \sum_{n=0}^{N} |\tilde{\mathbf{e}}^{n+1}|^2 \\
&\leq C\,k^2 \int_0^T ||\mathbf{u}_{tt}||^2_{Y'}\,dt + C\,k^2 \int_0^T |\mathbf{u}_t|^2\,dt \\
&+ C\,k \sum_{n=0}^{N} ||\tilde{\mathbf{e}}^{n+1}||^2_{Y'} + C\,k^2 \sum_{n=0}^{N} ||\tilde{\mathbf{e}}^{n+1}||^2 \\
&+ C\,k \sum_{n=0}^{N} |\tilde{\mathbf{e}}^{n+1} - \tilde{\mathbf{e}}^{n+1/2}|^2 + C\,k \sum_{n=0}^{N} |\tilde{\mathbf{e}}^{n+1/2} - \tilde{\mathbf{e}}^n|^2
\end{aligned}$$

$$+ \; C\,k^2 \sum_{n=0}^{N} ||\tilde{\mathrm{e}}^{n+1} - \tilde{\mathrm{e}}^{n+1/2}||^2 \; + \; C\,k^2 \sum_{n=0}^{N} ||\tilde{\mathrm{e}}^{n+1/2}||^2$$

Using again 4.7, the regularity properties of the continuous solution and the estimates of Lemma 4.8, we get:

$$\begin{aligned} ||\tilde{\mathrm{e}}^{N+1}||^2_{Y'} \; &+ \; \sum_{n=0}^{N} ||\tilde{\mathrm{e}}^{n+1} - \tilde{\mathrm{e}}^{n}||^2_{Y'} \; + \; k\,\nu \sum_{n=0}^{N} |\tilde{\mathrm{e}}^{n+1}|^2 \\ &\leq \; C\,k^2 \; + \; C\,k \sum_{n=0}^{N} ||\tilde{\mathrm{e}}^{n+1}||^2_{Y'} \end{aligned}$$

For sufficiently small k, we can apply the discrete Gronwall lemma to the last inequality, and we get:

$$\begin{aligned} ||\tilde{\mathrm{e}}^{N+1}||^2_{Y'} \; &+ \; \sum_{n=0}^{N} ||\tilde{\mathrm{e}}^{n+1} - \tilde{\mathrm{e}}^{n}||^2_{Y'} \; + \; k\,\nu \sum_{n=0}^{N} |\tilde{\mathrm{e}}^{n+1}|^2 \\ &\leq \; C\,k^2 \end{aligned}$$

and the estimate for $\tilde{\mathrm{u}}^{n+1}$ is proved. For $\tilde{\mathrm{u}}^{n+1/2}$, we have, once again:

$$\begin{aligned} k\,\nu \sum_{n=0}^{N} |\tilde{\mathrm{e}}^{n+1/2}|^2 \; &\leq \; 2\,k\,\nu \sum_{n=0}^{N} \Big(|\tilde{\mathrm{e}}^{n+1}|^2 + |\tilde{\mathrm{e}}^{n+1} - \tilde{\mathrm{e}}^{n+1/2}|^2 \Big) \\ &\leq \; C\,k^2 \end{aligned}$$

due to Lemma 4.8 and the estimate for $\tilde{\mathrm{u}}^{n+1}$, so that 4.54 is proved. □

REMARK 4.8: once again, since we are assuming that the domain Ω is smooth, we can assure that our semidiscrete velocities $\tilde{\mathrm{u}}^{n+1/2}$ and $\tilde{\mathrm{u}}^{n+1}$, actually belong to $\mathrm{H}^2(\Omega)$. We can also improve our error estimates for $\tilde{\mathrm{u}}^{n+1}$ to strongly first order in $\mathrm{L}^2(\Omega)$ and weakly first order in $\mathrm{H}^1_0(\Omega)$, by assuming that $\tilde{\mathrm{u}}^{n+1/2}$ is uniformly bounded in $\mathrm{H}^2(\Omega)$; we give this improvement in the Appendix.

REMARK 4.9: these error estimates are valid for any value of the parameter ϕ, but are restricted by condition 4.46. However, this condition is less restrictive than the uniform bound for the intermediate velocity in $\mathrm{H}^2(\Omega)$.

REMARK 4.10: we also proved some error estimates for the pressure, but again these depend on the improved estimates for the velocity; we present them in the Appendix.

4.3.3 Dependence of the steady state on the time step

In this subsection we address the issue of whether a steady solution obtained by a fractional-step transient algorithm, when neither the forcing term nor the boundary conditions depend on time, depends on the time step size used to find that solution. We will show that, in general, when pressure correction is used the steady solution is independent of the time step, while if it is not used the final solution may depend on the time step if implicit approximations of the viscous and/or the convective term are employed. This is so because with pressure correction methods the intermediate and end-of-step velocities turn out to be the same at steady state, while in the other methods they do not. We will justify this idea on two methods: the classical projection method and our viscositty splitting method. We drop the boundary conditions for simplicity.

Classical projection method: let us recall here Shen's version of the projection method:

$$\begin{aligned} \frac{\mathbf{u}^{n+1/2} - \mathbf{u}^n}{\delta t} - \nu\Delta\mathbf{u}^{n+1/2} + (\mathbf{u}^n\cdot\nabla)\mathbf{u}^{n+1/2} &= \mathbf{f} \qquad (4.58)\\ \frac{\mathbf{u}^{n+1} - \mathbf{u}^{n+1/2}}{\delta t} + \nabla p^{n+1} &= 0 \qquad (4.59)\end{aligned}$$

A steady state is reached when $\mathbf{u}^{n+1} = \mathbf{u}^n$, $\mathbf{u}^{n+1/2} = \mathbf{u}^{n-1/2}$ and $p^{n+1} = p^n$, which we call $\mathbf{u}$, $\mathbf{u}^{1/2}$ and p, respectively. Adding up 4.58 and 4.59 at steady state yields:

$$-\nu\Delta\mathbf{u}^{1/2} + (\mathbf{u}\cdot\nabla)\mathbf{u}^{1/2} + \nabla p = \mathbf{f}$$

If we now isolate $\mathbf{u}^{1/2}$ from 4.59 and substitute it (formally) into this equation, we get:

$$-\nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p + \delta t(-\nu\Delta\nabla p + (\mathbf{u}\cdot\nabla)\nabla p) = \mathbf{f} \qquad (4.60)$$

It is thus apparent that $(\mathbf{u}, p)$ is not a solution of the steady Navier-Stokes equations 1.12, but of the modified equation 4.60, which depends on δt. The solution $(\mathbf{u}, p)$ will therefore also depend on the time step.

Classical projection method with pressure-correction: let us now consider the pressure-correction projection scheme studied in [90]:

$$\begin{aligned} \frac{\mathbf{u}^{n+1/2} - \mathbf{u}^n}{\delta t} - \nu\Delta\mathbf{u}^{n+1/2} + (\mathbf{u}^n\cdot\nabla)\mathbf{u}^{n+1/2} + \nabla p^n &= \mathbf{f} \quad (4.61)\\ \frac{\mathbf{u}^{n+1} - \mathbf{u}^{n+1/2}}{k} + \phi\nabla(p^{n+1} - p^n) &= 0 \quad (4.62)\end{aligned}$$

At steady state, we first obtain $\mathbf{u}^{n+1/2} = \mathbf{u}$ from 4.62 and then, from 4.61:

$$-\nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p = \mathbf{f}$$

so that $(\mathbf{u}, p)$ is actually a solution of 1.12, and thus independent of δt.

Viscosity splitting method: for the method presented in Section 4.2 with $\theta = 1$:

$$\begin{aligned} \frac{\mathbf{u}^{n+1/2}-\mathbf{u}^n}{\delta t} - \nu\Delta\mathbf{u}^{n+1/2} + (\mathbf{u}^n\cdot\nabla)\mathbf{u}^{n+1/2} &= \mathbf{f} && (4.63)\\ \frac{\mathbf{u}^{n+1}-\mathbf{u}^{n+1/2}}{\delta t} - \nu\Delta(\mathbf{u}^{n+1}-\mathbf{u}^{n+1/2}) + \nabla p^{n+1} &= 0 && (4.64)\end{aligned}$$

We add 4.63 and 4.64 at steady state, and get:

$$-\nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u}^{1/2} + \nabla p = \mathbf{f}$$

By isolating (formally) $\mathbf{u}^{1/2}$ from 4.64 and retaining only first order terms in the time step, we obtain:

$$\begin{aligned} \mathbf{u}^{1/2} &= \mathbf{u} + \delta t\,(I - \delta t\,\nu\,\Delta)^{-1}\nabla p\\ &= \mathbf{u} + O(\delta t)\,\nabla p\end{aligned}$$

so that at steady state:

$$-\nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p + O(\delta t)\Big((\mathbf{u}\cdot\nabla)\nabla p\Big) = \mathbf{f}$$

The steady solution does not satisfy 1.12 but this modified equation, which depends on δt; it is therefore dependent on the time step.

Viscosity splitting method with pressure–correction: finally, we consider the method of this Section, with $\theta = 1$ and arbitrary ϕ:

$$\begin{aligned} \frac{\mathbf{u}^{n+1/2}-\mathbf{u}^n}{\delta t} - \nu\Delta\mathbf{u}^{n+1/2} + (\mathbf{u}^n\cdot\nabla)\mathbf{u}^{n+1/2} + \nabla p^n &= \mathbf{f}\ (4.65)\\ \frac{\mathbf{u}^{n+1}-\mathbf{u}^{n+1/2}}{\delta t} - \nu\Delta(\mathbf{u}^{n+1}-\mathbf{u}^{n+1/2}) + \phi\nabla(p^{n+1}-p^n) &= 0\ (4.66)\end{aligned}$$

At steady state, we find again that $\mathbf{u}^{1/2} = \mathbf{u}$ from 4.66, and then from 4.65:

$$-\nu\Delta\mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} + \nabla p = \mathbf{f}$$

The steady solution satisfies 1.12 and is independent of the time step.

4.4 Computational aspects

We next consider the implementation of our viscosity splitting, pressure correction method with a finite element interpolation of the space variables.

4.4.1 Finite element discretization

We consider space discretizations of our viscosity splitting, pressure correction method 4.39–4.41 with parameters $\theta = 1$ and $\phi = 1$.

Given two finite dimensional spaces $V_h \subset \mathbf{H}_0^1(\Omega)$ and $Q_h \subset L_0^2(\Omega)$, the discrete equivalent to the weak problems 4.40 and 4.44 consists of finding $\mathbf{u}_h^{n+1/2} \in V_h$ such that, given $\mathbf{u}_h^n \in V_h$ and $p_h^n \in Q_h$:

$$\begin{aligned}\frac{1}{\delta t}(\mathbf{u}_h^{n+1/2} - \mathbf{u}_h^n, \mathbf{v}_h) &+ \nu((\mathbf{u}_h^{n+1/2}, \mathbf{v}_h)) + c(\mathbf{u}_h^n, \mathbf{u}_h^{n+1/2}, \mathbf{v}_h) \\ &+ b(\mathbf{v}_h, p_h^n) = (\mathbf{f}(t_{n+1}), \mathbf{v}_h), \quad \forall \mathbf{v}_h \in V_h \quad (4.67)\end{aligned}$$

and $\mathbf{u}_h^{n+1} \in V_h$ and $p_h^{n+1} \in Q_h$ such that:

$$\begin{aligned}\frac{1}{\delta t}(\mathbf{u}_h^{n+1} - \mathbf{u}_h^{n+1/2}, \mathbf{v}_h) + \nu((\mathbf{u}_h^{n+1} - \mathbf{u}_h^{n+1/2}, \mathbf{v}_h)) &+ b(p_h^{n+1} - p_h^n, \mathbf{v}_h) \\ &= 0, \quad \forall \mathbf{v}_h \in V_h \quad (4.68) \\ b(\mathbf{u}_h^{n+1}, q_h) &= 0, \quad \forall q_h \in Q_h\end{aligned}$$

respectively. We are mainly interested in the case when V_h and Q_h are defined through a discretization of Ω into finite elements. In particular, we consider two kinds of quadrilateral elements (in the terminology of two dimensions): the bilinear velocity, constant pressure element (Q_1P_0), which does not satisfy the discrete LBB condition and may develop spurious pressure modes, and the biquadratic velocity, linear pressure element (Q_2P_1), which is div–stable.

4.4.2 Numerical scheme

The matrix form of equations 4.67 and 4.68 is, in the notation used up to now, the following:

$$\begin{aligned}M\frac{U^{n+1/2} - U^n}{\delta t} + KU^{n+1/2} + A(U^n)U^{n+1/2} &+ G_0P^n = F^{n+1} \quad (4.69) \\ M\frac{U^{n+1} - U^{n+1/2}}{\delta t} + K(U^{n+1} - U^{n+1/2}) &+ G_0(P^{n+1} - P^n) \\ &= 0 \quad (4.70) \\ G_0^tU^{n+1} &= 0 \quad (4.71)\end{aligned}$$

The numerical solution of these equations presents some problems. On the one hand, the system matrix for the intermediate velocity equation 4.69

has to be computed and factorized once every time step, due to the implicit approximation of the convective term; moreover, that matrix is not symmetric, since the convective term is skew-symmetric. On the other hand, the coupled system 4.70–4.71 has the structure of a mixed problem, with a zero diagonal term in the incompressibility equation 4.71. Equation 4.70 can be rewritten as:

$$B(U^{n+1} - U^{n+1/2}) \ + \ \delta t G_0(P^{n+1} - P^n) \ = \ 0 \tag{4.72}$$

where $B = M + \delta t K$ was defined in Section 2.1. One can then isolate U^{n+1} from 4.72 and substitute it into 4.71, thus segregating the computation of the pressure from that of the velocity; this yields:

$$(G_0^t B^{-1} G_0)(P^{n+1} - P^n) \ = \ \frac{1}{\delta t} G_0^t U^{n+1/2} \tag{4.73}$$

But the computation of the system matrix for this pressure equation requires of the inversion of a full matrix B, which is prohibitive in most cases. We present an alternative way to solve 4.69–4.70–4.71, which bypasses the problem of inverting the matrix B.

We propose an iterative solution of the discrete equations 4.69–4.70–4.71; in it, each iteration consists of the solution of two diagonal systems and another system with a symmetric, positive (semi)definite matrix, which is simple to compute. This matrix need only be computed and factorized once at the beginning of the calculation; the computational cost of each iteration is, then, only due to the formation of three residual vectors, the solution of two diagonal systems and a backward and forward substitution, if direct methods of solution are used for the pressure system. If few iterations of the proposed scheme are needed, it will be more efficient than solving the original equations 4.69–4.70–4.71, which require of the inversion of the full matrix B once and the computation and factorization of a non–symmetric matrix for the intermediate velocity system, the formation of three right–hand–side vectors and two backward and forward subtitutions every time step. Most of the techniques employed here are adopted from similar ideas within the context of the predictor–multicorrector algorithm to be studied in the next Chapter.

Given the n–th step values U^n and P^n of velocities and pressures, respectively, the iterative procedure starts with the initializations $U_0^{n+1/2} = U^n$, $U_0^{n+1} = U^n$ and $P_0^{n+1} = P^n$ for the values at time t_{n+1}. Then, if $U_i^{n+1/2}$ and U_i^{n+1} are the i–th iteration approximations to $U^{n+1/2}$ and U^{n+1}, respectively, we consider the scheme:

$$M\frac{U_{i+1}^{n+1/2} - U^n}{\delta t} \ + \ KU_i^{n+1/2} \ + \ A(U^n)U_i^{n+1/2} \ + \ G_0 P^n \ = \ F^{n+1} \tag{4.74}$$

$$M\frac{U_{i+1}^{n+1} - U_{i+1}^{n+1/2}}{\delta t} \ + \ K(U_i^{n+1} - U_i^{n+1/2}) \ + \ G_0(P_{i+1}^{n+1} - P^n)$$

$$\begin{aligned} &= 0 && (4.75) \\ G_0^t U_{i+1}^{n+1} &= 0 && (4.76) \end{aligned}$$

At convergence, that is, when $U_{i+1}^{n+1/2} = U_i^{n+1/2}$, $U_{i+1}^{n+1} = U_i^{n+1}$ and $P_{i+1}^{n+1} = P_i^{n+1}$, these values satisfy 4.69–4.70–4.71. The actual stopping criterion that we use is:

$$\max \left(\frac{|U_{i+1}^{n+1} - U_i^{n+1}|_2}{|U_{i+1}^{n+1}|_2}, \frac{|U_{i+1}^{n+1/2} - U_i^{n+1/2}|_2}{|U_{i+1}^{n+1/2}|_2}, \frac{|P_{i+1}^{n+1} - P_i^{n+1}|_2}{|P_{i+1}^{n+1}|_2} \right) \leq \epsilon_{\text{fs}}$$

where $|X|_2$ is again the Euclidean norm of a vector X.

We can also isolate U_{i+1}^{n+1} from 4.75 and substitute it into 4.76, so as to segregate the computation of the pressure from that of the velocity. By doing this, we obtain:

$$(G_0^t M^{-1} G_0)(P_{i+1}^{n+1} - P^n) = \frac{1}{\delta t} G_0^t \Big(U_{i+1}^{n+1/2} - \delta t M^{-1} K (U_i^{n+1} - U_i^{n+1/2}) \Big) \qquad (4.77)$$

To make the scheme computationally efficient, we consider the approximation of the matrix M by its lumped diagonal M^L in all its appearances, which is common practice in similar contexts (see [46], for instance). The computation of the system matrix for 4.77 then becomes feasible, since the inversion it involves is then trivial.

The actual implementation of the scheme, however, is somewhat different. It is given in terms of nodal accelerations $\mathcal{A}$ and time derivatives of elemental pressures, $\dot{P}$. Calling $\mathcal{A}_{i+1}^{n+1/2} = \frac{U_{i+1}^{n+1/2} - U^n}{\delta t}$, $\mathcal{A}_{i+1}^{n+1} = \frac{U_{i+1}^{n+1} - U_{i+1}^{n+1/2}}{\delta t}$ and $\dot{P}_{i+1}^{n+1} = \frac{P_{i+1}^{n+1} - P^n}{\delta t}$, equations 4.74, 4.77 and 4.75 can be written, with the approximation of M by M^L, as:

$$M^L \mathcal{A}_{i+1}^{n+1/2} = R_1$$

with:

$$R_1 = F^{n+1} - K U_i^{n+1/2} - A(U^n) U_i^{n+1/2} - G_0 P^n$$

$$(\delta t)^2 (G_0^t (M^L)^{-1} G_0) \dot{P}_{i+1}^{n+1} = R_p$$

with:

$$R_p = G_0^t \Big(U_{i+1}^{n+1/2} - \delta t^2 (M^L)^{-1} K \mathcal{A}_i^{n+1} \Big)$$

and:

$$U_{i+1}^{n+1/2} = U^n + \delta t \, \mathcal{A}_{i+1}^{n+1}$$

Finally:

$$M^L \mathcal{A}_{i+1}^{n+1} = R_2$$

with:

$$R_2 = -\delta t \left(G_0 \dot{P}_{i+1}^{n+1} + K \mathcal{A}_i^{n+1} \right)$$

The end-of-step values are then corrected as:

$$\begin{aligned} U_{i+1}^{n+1} &= U^n + \delta t\, \mathcal{A}_{i+1}^{n+1/2} + \delta t\, \mathcal{A}_{i+1}^{n+1} \\ &= U_{i+1}^{n+1/2} + \delta t\, \mathcal{A}_{i+1}^{n+1} \\ P_{i+1}^{n+1} &= P^n + \delta t\, \dot{P}_{i+1}^{n+1} \end{aligned}$$

In the next Section we present some numerical results obtained with this scheme.

In the implementation of this method we have adopted the rate–of–deformation tensor formulation of the viscous term $\epsilon(\mathbf{u})$, as that of equation 1.6. This formulation does not assume incompressibility, which is in general not satisfied by the discrete velocity field, and, in outflow boundaries, the natural boundary condition associated with it has the physical meaning of a no stress condition. For the convective term we have employed the standard formulation $(\mathbf{u} \cdot \nabla)\mathbf{u}$.

4.5 Numerical results

We present the results obtained with our viscosity splitting, pressure correction method with parameters $\theta = \phi = 1$ on three test problems. The first one is a test case introduced by van Kan (see [62]), intended to study numerically the order of approximation of the scheme in the time step; the second one is the classical problem of steady flow over a backward facing step, and the third one is the problem of flow around a cylinder.

4.5.1 Numerical accuracy study

As a numerical check for the accuracy properties of the method, we considered a test case introduced by van Kan (see [62]). It consists of the Navier–Stokes flow on a unit square cavity in which an inflow velocity profile is prescribed at the top wall defined by $\mathbf{u}((x,1),t) = (0, -\sin(\pi(x^3 - 3x^2 + 3x))e^{(1-1/t)})$ for $0 \le x \le 1$ and $t > 0$, the bottom and left walls are solid walls and natural boundary conditions are enforced on the right, outlet wall. As in [62], a Reynolds number of 10 was selected, and the fluid was at rest at the start. A uniform mesh consisting of 6 × 6 elements was used for the Q_1P_0 case; in order to compare the results from both interpolations, the same mesh points were used to define a 3 × 3 mesh for the Q_2P_1 element.

δt	$\kappa_1(\delta t)$	$\kappa_2(\delta t)$	$\kappa_p(\delta t)$
1/60	2.1	3.5	2.3
1/64	2.0	3.4	2.4
1/80	2.0	3.0	2.2
1/85	2.0	2.9	2.1

Table 4.1: Van Kan's flow, Q_1P_0 element.

δt	$\kappa_1(\delta t)$	$\kappa_2(\delta t)$	$\kappa_p(\delta t)$
1/75	2.2	3.5	2.5
1/80	2.1	3.3	2.5
1/85	2.1	3.1	2.4

Table 4.2: Van Kan's flow, Q_2P_1 element.

Let's denote by $\kappa_i(\delta t)$ the quotient:

$$\kappa_i(\delta t) = \frac{|U_i(\delta t) - U_i(\frac{1}{2}\delta t)|_2}{|U_i(\frac{1}{2}\delta t) - U_i(\frac{1}{4}\delta t)|_2},$$

where U_i $(i = 1, 2)$ contains the i-th component of the nodal velocities obtained at $t = 1$ with the indicated time–step. Euclidean norms are used for these vectors. Similarly, $\kappa_p(\delta t)$ denotes the same quotient for the elemental pressure (and eventually, pressure spatial derivative) values.

We show in Tables 4.1 and 4.2 the most accurate results obtained with our viscosity splitting, pressure correction method with parameters $\theta = \phi = 1$ for the two different space interpolations. We fixed the value of the tolerance to 10^{-4}; convergence of the iterative scheme was reached in 7 iterations in average for the largest time steps to 4 for the smallest. It can be observed that the scheme is, at least assymptotically, first order accurate in the time step both in velocities and in pressures.

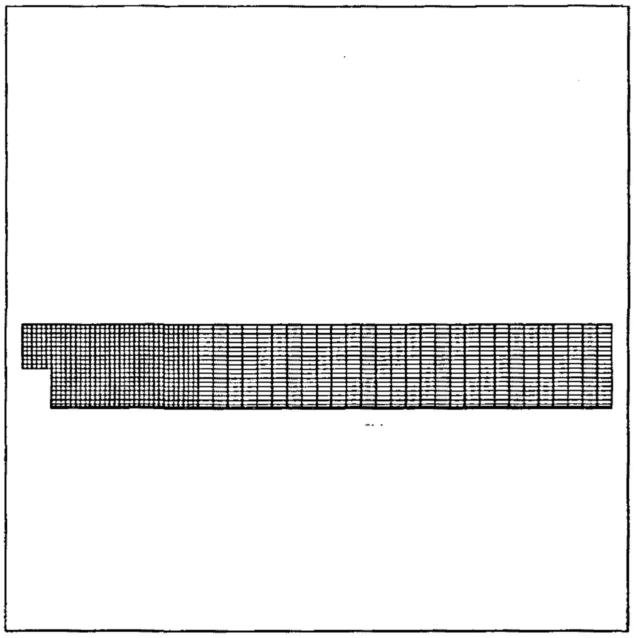


Figure 4.1: Backward facing step, mesh.

4.5.2 Backward facing step

We then studied the well-known problem of the flow over a backward facing step. This problem was extensively studied by B.F. Armaly *et al.* in [2], both experimentally and numerically, and other numerical results have been given by many authors (see [29], [36] or [65], for instance). Here we considered a geometry similar to that of [2], that is, an inflow channel of length 2 and height 1, an expansion ratio of 1 : 1.90 and a total channel length of 40. A Poiseuille parabolic profile was prescribed at the inflow, with a maximum velocity of 1; the top and bottom sides are solid walls, and natural boundary conditions are enforced at the outlet. The mesh used for this problem, which is finer near the step, can be seen in Figure 4.1, where the y-axes has been magnified three times; it consists of 1305 mesh points, which were used to define both the Q_1P_0 and the Q_2P_1 elements. There are 1220 and 305 of such elements, respectively.

We solved this problem for three different values of the Reynolds number: 40, 200 and 400. This was defined upon the average inflow velocity (which is 4/3 for our data), and the inflow channel height. It was obtained experimentally in [2] that in this range of Reynolds numbers the flow is virtually two-dimensional, so that planar numerical models become meaningful. With a time step size of $\delta t = 0.01$, we iterated the scheme 4.74-4.75-4.76 to convergence in each time step with a tolerance of $\epsilon_{\mathrm{fs}} = 10^{-3}$; this was obtained in a very few iterations: 3 or 4 in the first steps, decreasing to 1 in the last steps. A steady state was considered when the accelerations were in the order of 10^{-5}.

We show the results obtained for the different Reynolds numbers in Figures 4.2, for the Q_1P_0 element, and 4.3, for the Q_2P_1, in the form of steady streamlines, where the x-axes has been limited to the range $[0, 10]$. It can be clearly observed in these Figures how the reatachment length of the main vortex increases with increasing Reynolds numbers, a characteristic of the flow which is well known for this problem, since we are working within the laminar Reynolds number range (see [2]). Moreover, the appearance of a secondary separation bubble on the no-step wall at $Re = 400$ can also be observed, which is in good agreement with the experimental results of [2].

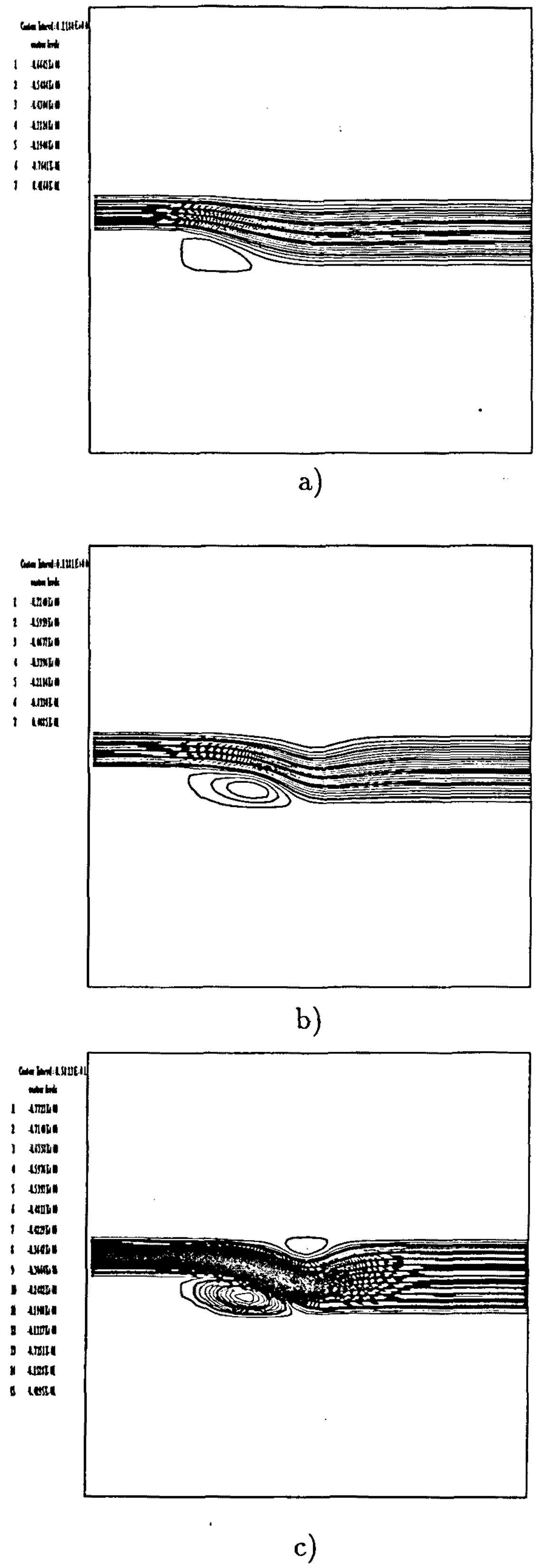
4.5.3 Flow past a circular cylinder

We finally considered the challenging problem of the flow past a circular cylinder, which has attracted the attention of several authors (see [7], [34], [20], [93], [96], [99] or [108], for instance), This has become a compulsory bechmark test for transient algorithms for Navier–Stokes equations.

It is well known that for low values of the Reynolds number, the solution is steady and symmetric about a line parallel to the free-stream flow through a cylinder diameter; a pair of symmetrical eddies develops downstream of the cylinder. But beyond a critical value of Re (which is larger than 40), the steady solution becomes unstable and a periodic solution develops, so that vortex shedding sets in: vortices begin to generate periodically and alternately from each side of the cylinder, and are 'transported' by the flow away from it. This scenario is known in the literature as a von Karman vortex street.

We considered a cylinder of unit diameter and took a computational domain consisting of the rectangle $[0, 21] \times [0, 9]$, the center of the cylinder being situated at the point $(4.5, 4.5)$. These data, however, may not be sufficient to prevent any effect of the introduction of artificial boundaries on the computed solution, as was recently studied in [7], who discussed the influence of the location of the lateral boundaries on the computed flow field; we will see how this may affect our computations. A unit free-stream horizontal velocity was prescribed on the left boundary, whereas natural conditions are enforced on all the others. The mesh used in this case can be seen in Figure 4.4, which consists of 3000 nodes and 2880 of the Q_1P_0 elements.

We first solved the problem for a Reynolds number of 40, which is based upon the free-stream velocity and the cylinder diameter, starting from the fluid at rest but for the prescribed boundary condition. We iterated the scheme 4.74–4.75–4.76 to convergence in each time step with a tolerance of $\epsilon_{fs} = 10^{-2}$, which took an average of 2 iterations. After 1000 steps of size $\delta t = 0.005$, the steady, symmetric solution had been reached, with accelerations in the order of 10^{-4}. It can be seen in Figures 4.5, 4.6 and 4.7, where we show, respectively, the streamlines, the stationary streamlines (that is, the

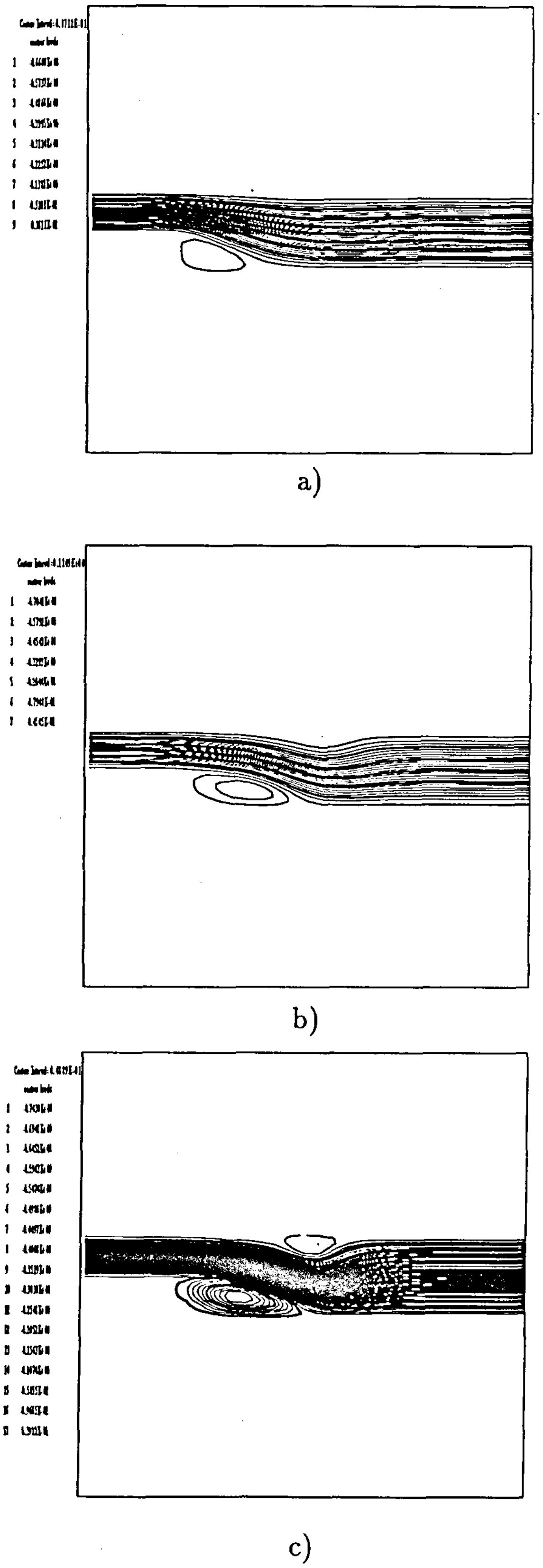


a)

b)

c)

Figure 4.2: Backward facing step, Q_1P_0 element, streamlines: a) $Re = 60$; b) $Re = 200$; c) $Re = 400$.



a)

b)

c)

Figure 4.3: Backward facing step, Q_2P_1 element, streamlines: a) $Re = 60$; b) $Re = 200$; c) $Re = 400$.

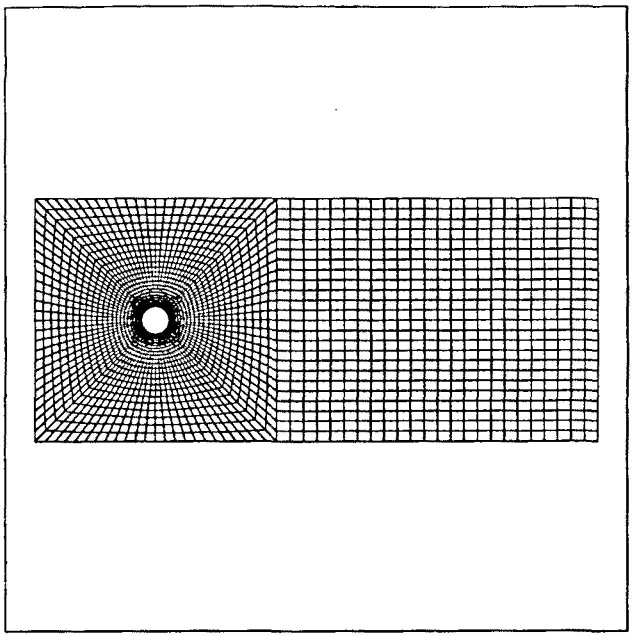


Figure 4.4: Flow past a cylinder, mesh.

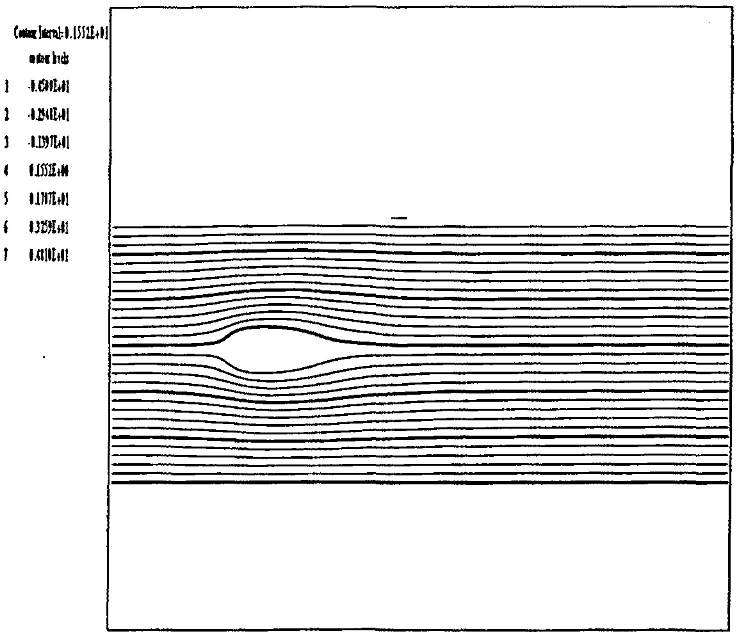


Figure 4.5: Flow past a cylinder, $Re = 40$, streamlines.

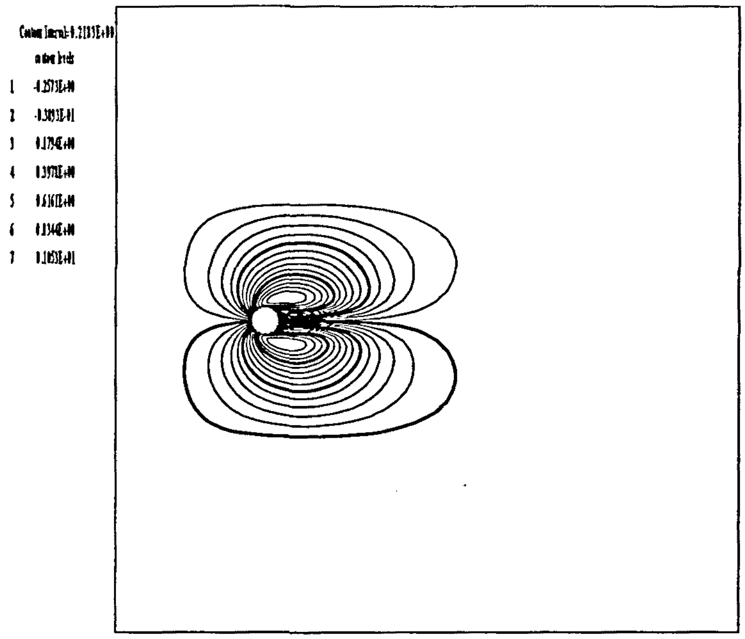


Figure 4.6: Flow past a cylinder, *Re* = 40, stationary streamlines.

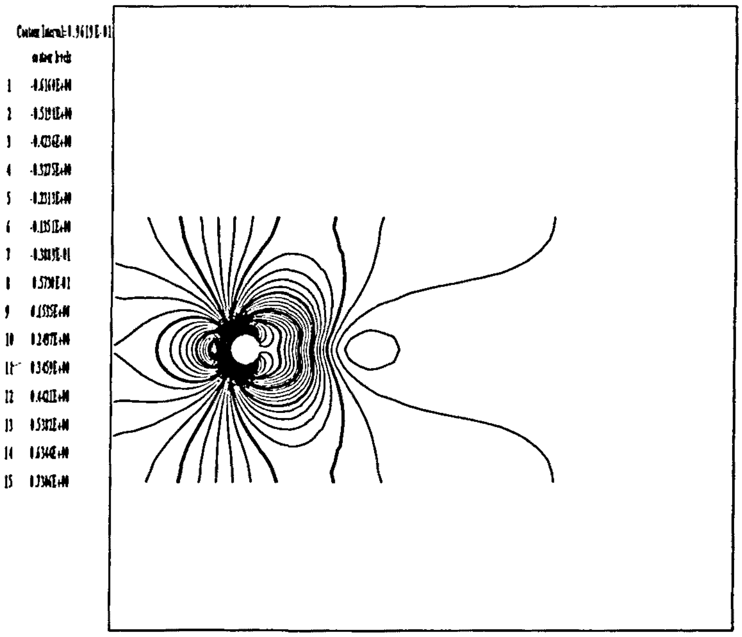


Figure 4.7: Flow past a cylinder, *Re* = 40, nodal pressure contours.

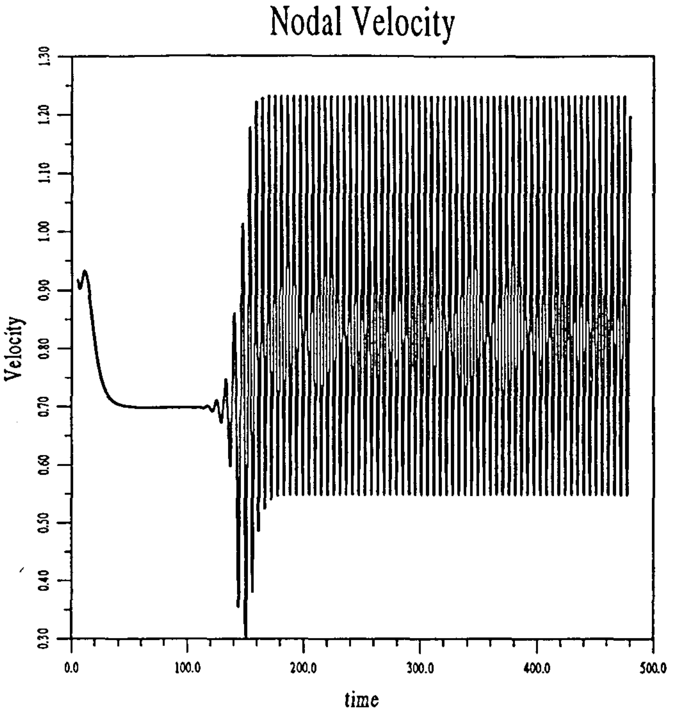


Figure 4.8: Flow past a cylinder, $Re = 100$, nodal velocity history.

streamlines obtained assuming that it is the cylinder that moves with a constant velocity of $(-1,0)$) and the nodal pressure contours obtained from the elemental pressures after a least–squares interpolation process. Symmetry is very accurately achieved.

We then raised the value of the Reynolds number to 100, which is the one commonly used for this problem. We started the computation from the steady solution obtained for $Re = 40$, and performed 19000 steps of size $\delta t = 0.025$; in each of them, 1 or 2 iterations were enough to reach convergence at the same value of the tolerance as before. We found that the solution started oscillating freely at a time near $t = 110$; the final periodicity of the solution was reached by $t = 170$. In Figure 4.8 we show the history of the horizontal velocity at a node situated at the point $(9.0, 5.25)$, that is, downstream of the cylinder and slightly higher. The qualitative change in the solution regime can be clearly observed. In this case, no artificial trick was needed to start up the periodic solution.

The streamlines obtained at the end of the computation ($t = 475$) are shown in Figure 4.9. In Figure 4.10 we plot the stationary streamlines; the wakes behind the cylinder can be clearly seen there. Finally, we show the pressure contours in Figure 4.11. All these results compare very well with other published solutions (see [34], [20], [96] or [108]).

Some flow features are generally used to compare quantitatively the solutions obtained for this problem. Thus, the Strouhal number or adimensional frequency of the solution is one of the most studied quantities; it is defined

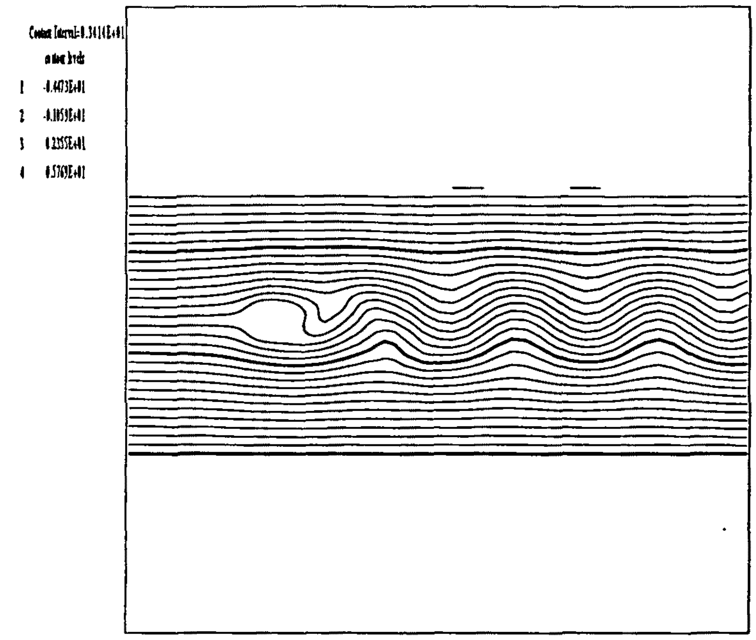


Figure 4.9: Flow past a cylinder, Re = 100, streamlines.

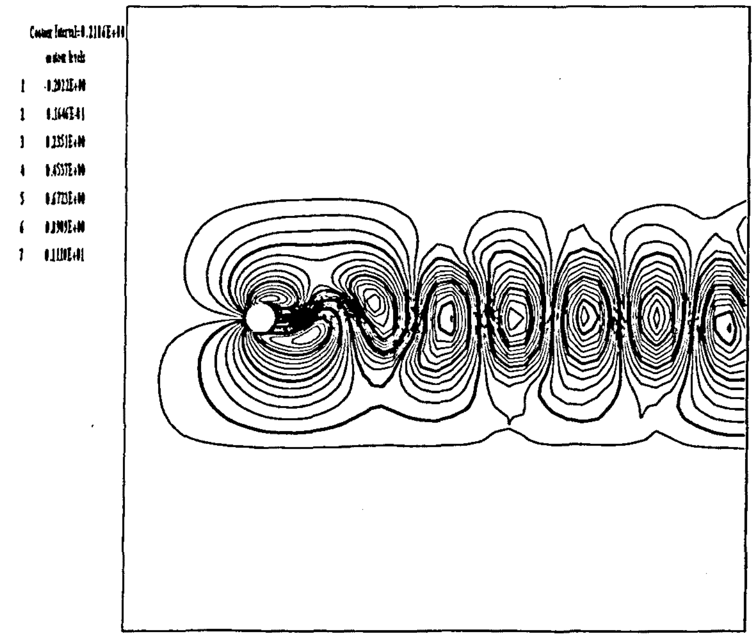


Figure 4.10: Flow past a cylinder, Re = 100, stationary streamlines.

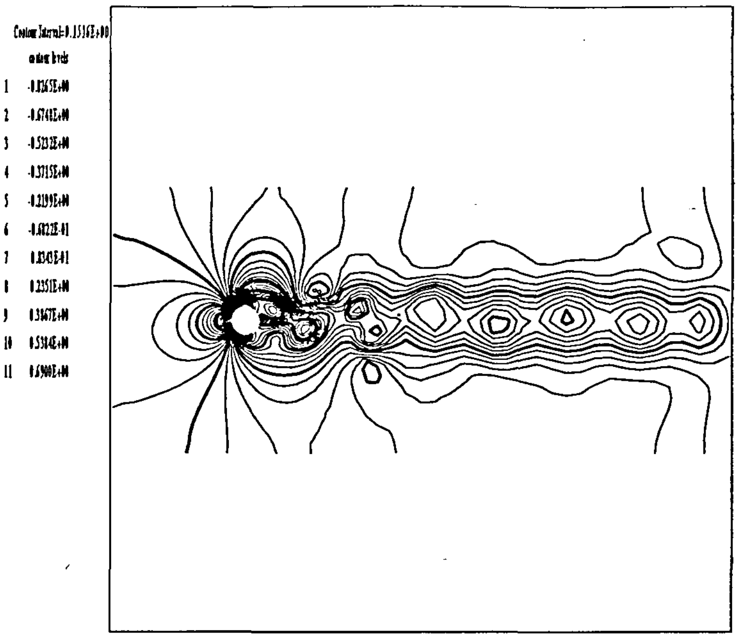


Figure 4.11: Flow past a cylinder, $Re = 100$, nodal pressure contours.

as $St = \dfrac{D}{u_0\tau}$, where D is the cylinder diameter, u_0 is the free–stream velocity (in our case, both equal to 1) and τ is the shedding period of the solution. We performed a Fourier analysis of the nodal velocity signal within the time range [175, 475] (that is, for most of the developed periodic solution) in order to find the dominant frequency of our solution. In Figure 4.12 we show the Fourier spectrum obtained, from which we found a Strouhal number of $St = 0.18667$ (smaller peaks can also be seen at twice and three times that frequency), or equivalently, a period of 5.3571. This period is somewhat smaller than the one generally admitted for this value of the Reynolds number, which is 6, that is, a Strouhal number of $St = 0.16667$ (see [20]). We attibute this discrepancy to the fact that we are using a standard Galerkin finite element interpolation, which is less dissipative than stabilized formulations of the SUPG or GLS type usually employed for this problem. However, discrepances in the value of the Strouhal number depending on the formulation employed were also found by other authors (see [96] or [108]). Moreover, the location of the lateral boundaries in our computational domain may not be far enough from the cylinder to avoid any influence on the solution of the artificial boundary conditions introduced by the formulation; in fact, it was obtained in [7] that at least 12 cylinder diameters on each side of the cylinder are needed to avoid that influence; otherwise, larger Strouhal numbers were obtained. This may be another cause of increase of our computed Strouhal number.

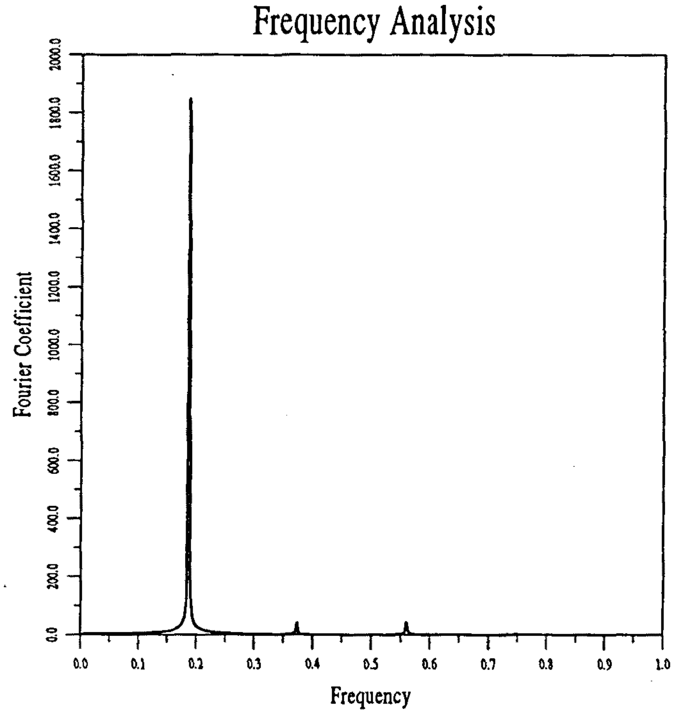


Figure 4.12: Flow past a cylinder, $Re = 100$, Fourier spectrum of the nodal velocity solution.