A Finite Element Formulation for the Stokes Problem Allowing Equal Velocity-Pressure Interpolation

R. Codina

J. Blasco

A Finite Element Formulation for the Stokes Problem Allowing Equal Velocity-Pressure Interpolation

Ramon Codina

Jordi Blasco

Submitted for publication to the journal

Computer Methods in Applied Mechanics and Engineering

A FINITE ELEMENT FORMULATION
FOR THE STOKES PROBLEM ALLOWING
EQUAL VELOCITY-PRESSURE INTERPOLATION

Ramon Codina and Jordi Blasco

Escola Tècnica Superior d'Enginyers de Camins, Canals i Ports
Universitat Politècnica de Catalunya
Gran Capità s/n, Edifici C1, 08034 Barcelona, Spain

CONTENTS

A FINITE ELEMENT FORMULATION FOR THE STOKES PROBLEM ALLOWING EQUAL VELOCITY–PRESSURE INTERPOLATION

Ramon Codina and Jordi Blasco

Escola Tècnica Superior d'Enginyers de Camins, Canals i Ports
Universitat Politècnica de Catalunya
Gran Capità s/n, Edifici C1, 08034 Barcelona, Spain

Summary

In this paper we study a variational formulation of the Stokes problem that accomodates the use of equal velocity-pressure finite element interpolations. The motivation of this method relies on the analysis of a class of fractional-step methods for the Navier-Stokes equations for which it is known that equal interpolations yield good numerical results. The reason for this turns out to be the difference between two discrete Laplacian operators computed in a different manner. The formulation of the Stokes problem considered here aims to reproduce this effect. From the analysis of the finite element approximation of the problem we obtain stability and optimal error estimates using velocity-pressure interpolations satisfying a compatibility condition much weaker than the inf-sup condition of the standard formulation. In particular, this condition is fulfilled by the most common equal order interpolations.

1 Introduction

The choice of the velocity and pressure spaces for the finite element approximation of the Stokes problem is of major importance. The standard Galerkin approach necessitates an interpolation for both fields satisfying the classical inf-sup or Babuška-Brezzi stability condition (see, e.g., Ref. [1]). Elements satisfying it have been blamed to be complicated and expensive in practice, especially in three dimensional problems. This being unavoidable or not, the fact is that several numerical methods have been recently developed with the goals of either using equal interpolations or stabilizing simple elements, such as the Q_1/P_0 pair (multilinear velocity, piecewise constant pressure). Examples of the first group are the methods of Brezzi & Douglas [2], Douglas & Wang [3] and the popular Galerkin/least-squares (GLS) technique of Hughes *et al.* [4, 5]. Fortin & Boivin [6] and Silvester & Kechkar [7] developed stabilization techniques for the Q_1/P_0 element, and similar ideas

can also be found in Refs. [8, 9]. The analysis of a rather general stabilization technique is presented in the paper of Franca & Stenberg [10].

On the other hand, it has been observed in practice that some fractional-step methods for the incompressible Navier-Stokes equations that employ a pressure Poisson equation in the projection step allow to use equal interpolation. This is in general true for methods that segregate the pressure and compute it via a Poisson equation (see for example Refs. [11–13]). In the fractional-step method presented by Zienkiewicz & Codina in Ref. [14], this fact was intuitively explained by the presence of a nonzero matrix multiplying the pressure in the continuity equation. This matrix is the difference between two discrete Laplacian matrices computed in a different way. It turns out that it is positive semi-definite [15], thus explaining in part why equal interpolation is possible.

In this paper we present a new formulation for the Stokes problem whose motivation is to have the same stabilization properties as the fractional-step methods just mentioned, even though the equations to be solved are very different. We introduce a new vector variable which, in the discrete problem, is the projection onto the space of continuous vector functions of the pressure gradient. This results in an important increase in the number of nodal unknowns, making the applicability of the method limited from the computational standpoint. Nevertheless, iterative strategies may be devised to make the method more efficient, although we shall not pursue this in this paper.

To analyze the stability of the finite element approximation, we introduce a technique based on the decomposition of the vector space that contains both velocities and pressure gradients into three orthogonal subspaces. We prove stability for each of the components of the pressure gradient separately. In order to bound one of these components we are led to an inf-sup condition for stability, similar to that obtained for the classical Galerkin approximation but much weaker. In particular, it is satisfied by most of the common equal order interpolations. To prove this fact, we use a macroelement technique similar to that presented by Stenberg in Ref. [16] (see also Ref. [17]). Once stability is established, optimal error estimates are proved under the usual regularity assumptions.

We have organized the paper as follows. The formulation we propose is described in Section 2. After stating the problem and introducing some notation, we describe a type of fractional-step methods that motivate the method whose analysis is undertaken in Section 3. In Section 4 we present some very simple numerical tests and make some remarks concerning the implementation of the method and its relationship with the GLS technique. Finally, we draw some conclusions.

2 The Stokes problem reformulated

2.1 Statement of the problem

Let us first consider the classical Stokes problem for an incompressible fluid. Let Ω be an open, bounded and polyhedral domain of $\mathbb{R}^d$, where $d = 2$ or 3 is the number of space dimensions, and $\Gamma = \partial\Omega$ its boundary. The Stokes problem consists in finding a velocity

$\boldsymbol{u}$ and a pressure p such that

$$-\nu\Delta\boldsymbol{u} + \nabla p = \boldsymbol{f} \qquad\qquad \text{in } \Omega, \qquad\qquad (1)$$
$$\nabla\cdot\boldsymbol{u} = 0 \qquad\qquad \text{in } \Omega, \qquad\qquad (2)$$
$$\boldsymbol{u} = 0 \qquad\qquad \text{on } \Gamma, \qquad\qquad (3)$$

where ν is the kinematic viscosity and $\boldsymbol{f}$ is the force vector. We have considered the homogeneous Dirichlet boundary condition (3) for simplicity.

To write the weak form of problem (1)–(3) we need to introduce some notation. As usual, we denote by $H^m(\omega)$ the Sobolev space of mth order in a set ω, consisting of functions whose distributional derivatives of order up to m belong to $L^2(\omega)$, and by $H_0^1(\omega)$ the subspace of $H^1(\omega)$ of functions with zero trace on Γ. A bold character is used for the vector counterpart of these spaces. The L^2 scalar product is denoted by $(\cdot,\cdot)_\omega$, and the H^m norm by $\|\cdot\|_{m,\omega}$. The subscript m is omitted when $m = 0$ and so is ω when it is Ω.

Let us now consider the spaces

$$V = H_0^1(\Omega), \qquad Q = \left\{ q \in L^2(\Omega) \;\middle|\; \int_\Omega q \; d\Omega = 0 \right\}, \qquad\qquad (4)$$

and the bilinear forms

$$a(\boldsymbol{u},\boldsymbol{v}) = \nu(\nabla\boldsymbol{u}, \nabla\boldsymbol{v}), \qquad b(q,\boldsymbol{v}) = (q, \nabla\cdot\boldsymbol{v}), \qquad\qquad (5)$$

with $\boldsymbol{u}, \boldsymbol{v} \in V$ and $q \in Q$. If $< \cdot, \cdot >$ denotes the duality pairing between V and its topological dual space V' where $\boldsymbol{f}$ is assumed to belong, the weak form of problem (1)–(3) consists in finding $(\boldsymbol{u}, p) \in V \times Q$ such that

$$a(\boldsymbol{u},\boldsymbol{v}) - b(p,\boldsymbol{v}) = <\boldsymbol{f},\boldsymbol{v}> \qquad\qquad \forall \boldsymbol{v} \in V, \qquad\qquad (6)$$
$$b(q,\boldsymbol{u}) = 0 \qquad\qquad \forall q \in Q. \qquad\qquad (7)$$

Existence and uniqueness of solution to this problem follows from the coercivity of a in $V \times V$, which is a consequence of the Poincaré-Friedrics inequality, and from the inf-sup or Babuška-Brezzi condition. These conditions can be written as follows: there exist positive constants K_a and K_b such that

$$a(\boldsymbol{v},\boldsymbol{v}) \geq K_a\|\boldsymbol{v}\|_1 \quad \forall \boldsymbol{v} \in V, \qquad\qquad (8)$$
$$\inf_{q\in Q_1} \sup_{v\in V_1} b(q,\boldsymbol{v}) \geq K_b, \qquad\qquad (9)$$

where Q_1 and V_1 are defined as

$$Q_1 = \{q \in Q \mid \|q\| = 1\}, \qquad V_1 = \{\boldsymbol{v} \in V \mid \|\boldsymbol{v}\|_1 = 1\}. \qquad\qquad (10)$$

Condition (9) holds true for V and Q given by Eq. (4).

If instead of having $\boldsymbol{f} \in V' = \boldsymbol{H}^{-1}(\Omega)$ we require $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ and Γ is sufficiently smooth, it is known that the solution of problem (6) & (7) verifies $\boldsymbol{u} \in V \cap \boldsymbol{H}^2(\Omega)$ and

3

$p \in Q \cap H^1(\Omega)$, that is, the regularity of the solution increases (see, e.g., Ref. [18]). Also, the duality $< f, v >$ in Eq. (6) can be replaced by (f, v). In the case of polygonal Γ that we consider, we need to require explicitly $p \in H^1(\Omega)$. This is the situation that we consider throughout in this paper.

Let $\mathcal{T}_h$ denote a finite element partition of the domain Ω of diameter h. For simplicity, we assume that all the element domains $K \in \mathcal{T}_h$ are the image of a reference element $\hat{K}$ through a polynomial mapping F_K, affine for simplicial elements, bilinear for quadrilaterals and trilinear for hexahedra. On $\hat{K}$ we define the polynomial spaces $\hat{V} = [R_{k_v}(\hat{K})]^d$ and $\hat{Q} = R_{k_q}(\hat{K})$, where, as usual, $R_k = P_k$ for simplicial elements and $R_k = Q_k$ for quadrilaterals and hexahedra. From $\hat{V}$ and $\hat{Q}$ we construct the finite element spaces

$$Q_h = \left\{ q_h \in C^0(\Omega) \mid q_h|_K = \hat{q} \circ F_K^{-1}, \ \hat{q} \in \hat{Q}, \ K \in \mathcal{T}_h \right\}, \tag{11}$$

$$V_h = \left\{ v_h \in [C^0(\Omega)]^n \mid v_h|_K = \hat{v} \circ F_K^{-1}, \ \hat{v} \in \hat{V}, \ K \in \mathcal{T}_h \right\}, \tag{12}$$

$$V_{h,0} = \{ v_h \in V_h \mid v_h|_\Gamma = 0 \}. \tag{13}$$

Notice that both the velocity and pressure finite element spaces V_h and Q_h are referred to the same partition and both are made up with continuous functions. The case $k_v = k_q - 1$ corresponds to Taylor-Hood type elements. In what follows, we put special emphasis on the case of equal interpolation $k_v = k_q$.

The discrete finite element counterpart of problem (6) & (7) can now be written as follows: find $(u_h, p_h) \in V_{h,0} \times Q_h$ such that

$$a(u_h, v_h) - b(p_h, v_h) = (f, v_h) \qquad \forall v_h \in V_{h,0}, \tag{14}$$

$$b(q_h, u_h) = 0 \qquad \forall q_h \in Q_h. \tag{15}$$

2.2 On a class of fractional-step methods

In order to motivate the method to be introduced in the following section, let us first describe the application of the classical fractional-step method of Chorin [19] and Temam [20] to the transient version of problem (1)–(3), that is,

$$\frac{\partial u}{\partial t} - \nu \Delta u + \nabla p = f, \tag{16}$$

$$\nabla \cdot u = 0. \tag{17}$$

These equations must be supplied with initial and boundary conditions, although they are irrelevant for what follows.

Consider a partition of the time interval into time steps of size Δt and denote by a superscript the time step counter. With u^n known, the classical fractional-step method consists in finding an intermediate velocity $u^{n+1/2}$ as the solution of the equation

$$\frac{1}{\Delta t} \left(u^{n+1/2} - u^n \right) - \nu \Delta u^{n+1/2} = f, \tag{18}$$

4

followed by the projection of $u^{n+1/2}$ onto the space of solenoidal vector fields. This leads to solving

$$\frac{1}{\Delta t}\left(u^{n+1} - u^{n+1/2}\right) + \nabla p^{n+1} = 0, \tag{19}$$

$$\nabla \cdot u^{n+1} = 0. \tag{20}$$

A common approach for solving problem (19)–(20) is to take the divergence of Eq. (19) and make use of Eq. (20), to yield a Poisson equation for the pressure, namely,

$$\Delta p^{n+1} = \frac{1}{\Delta t}\nabla \cdot u^{n+1/2}. \tag{21}$$

Once this equation is solved, Eq. (19) can be used to obtain u^{n+1}, thus uncoupling the calculation of the velocity and the pressure, which is one of the reasons for the success of fractional-step methods.

Once a finite element space discretization is performed, the matrix form of Eqs. (19)–(21) will be

$$\frac{1}{\Delta t}M\left(U^{n+1} - U^{n+1/2}\right) + GP^{n+1} = 0, \tag{22}$$

$$-G^{t}U^{n+1} = 0, \tag{23}$$

$$\Delta t L P^{n+1} - G^{t}U^{n+1/2} = 0. \tag{24}$$

In these equations, we use capital letters to denote the vectors of nodal unknowns of the corresponding lower case variables, M is the mass matrix, G is the matrix coming from the gradient term and L is the one coming from the Laplacian.

If the intermediate velocity $U^{n+1/2}$ is eliminated in Eq. (24) using Eq. (22), we obtain

$$-G^{t}U^{n+1} + \Delta t\left(L - G^{t}M^{-1}G\right)P^{n+1} = 0. \tag{25}$$

Therefore, we see that, whereas at the continuous level it is equivalent to use either Eqs. (19) & (20) or Eqs. (19) & (21), at the discrete level there is a difference between using Eqs. (22) & (23) and Eqs. (22) & (24). The latter choice corresponds a modification of the continuity Eq. (23) to Eq. (25). The term $L - G^{t}M^{-1}G$ may be understood as the difference between two discrete Laplacian operators computed in a different manner. This matrix turns out to be positive semi-definite [15], which increases the stability of the numerical method. Thus, the benefit of using Eqs. (22) & (24) is more than just uncoupling the velocity and pressure computations.

The matrices appearing in these expressions should in fact be modified according to the boundary conditions imposed on (18) and (19) & (20), a point that we have deliberately omitted since it does not affect our discussion and boundary conditions are always controversial when using fractional-step methods.

2.3 Modified discrete problem

We are now in a position to present the finite element formulation that we propose. The idea is to recover the stabilization properties of the fractional-step method discussed above.

The term $G^t M^{-1} G P$ can be obtained by taking first the gradient of the pressure, projecting it onto the discrete space of velocities and then taking the divergence of the resulting vector field.

Let $\alpha > 0$ be given. The modification of problem (14) & (15) that we consider is as follows: find $(u_h, \tilde{u}_h, p_h) \in V_{h,0} \times V_h \times Q_h$ such that

$$
\begin{aligned}
a(u_h, v_h) - b(p_h, v_h) &= (f, v_h) & \forall v_h \in V_{h,0}, & \quad (26) \\
\alpha(\nabla p_h, \nabla q_h) - \alpha(\tilde{u}_h, \nabla q_h) + b(q_h, u_h) &= 0 & \forall q_h \in Q_h, & \quad (27) \\
-(\nabla p_h, \tilde{v}_h) + (\tilde{u}_h, \tilde{v}_h) &= 0 & \forall \tilde{v}_h \in V_h. & \quad (28)
\end{aligned}
$$

Observe that the vector $\tilde{u}_h$ is precisely the projection of ∇p_h onto V_h.

If we denote by K the matrix coming from the viscous term (i.e., from a) and introduce a subscript naught to refer to the non prescribed degrees of freedom, the matrix version of this problem is

$$
\begin{aligned}
KU + G_0 P &= F_0, & (29) \\
\alpha L P - \alpha G^t \tilde{U} - G_0^t U &= 0, & (30) \\
-GP + M\tilde{U} &= 0. & (31)
\end{aligned}
$$

Eliminating $\tilde{U}$ from Eq. (31) and inserting it in Eq. (30) it is found that

$$
\alpha \left(L - G^t M^{-1} G \right) P - G_0^t U = 0, \tag{32}
$$

an equation similar to Eq. (25).

Problem (26)–(28) is consistent, in the sense that the solution of the continuous problem satisfies it. If this solution is sufficiently smooth, the original problem (1)–(3) may be replaced by

$$
\begin{aligned}
-\nu \Delta u + \nabla p &= f & \text{in } \Omega, & \quad (33) \\
-\alpha \left(\Delta p - \nabla \cdot \tilde{u} \right) + \nabla \cdot u &= 0 & \text{in } \Omega, & \quad (34) \\
-\nabla p + \tilde{u} &= 0 & \text{in } \Omega, & \quad (35) \\
u &= 0 & \text{on } \Gamma, & \quad (36) \\
\frac{\partial p}{\partial n} - n \cdot \tilde{u} &= 0 & \text{on } \Gamma, & \quad (37)
\end{aligned}
$$

where n is the unit outward normal to Γ. This problem is exactly equivalent to problem (1)–(3). Since problem (26)–(28) can be thought of as the discretization of the weak form of problem (33)–(37), we can expect the correct behavior of the pressure near the boundary. We shall come back to this point in section 4.

3 Numerical analysis

3.1 Preliminaries

In this section we analyze problem (26)–(28). We prove that the solution is stable under a mild condition for the velocity and pressure finite element spaces. After this, we give optimal error estimates for the unknowns.

First we need to introduce some notation. Let us consider the bilinear form on $(V_{h,0} \times Q_h \times V_h)^2$ defined as

$$
\begin{aligned}
\mathcal{B}(u_h, p_h, \tilde{u}_h; v_h, q_h, \tilde{v}_h) = & \, a(u_h, v_h) - b(p_h, v_h) \\
& + \alpha(\nabla p_h, \nabla q_h) - \alpha(\tilde{u}_h, \nabla q_h) + b(q_h, u_h) \\
& - \alpha(\nabla p_h, \tilde{v}_h) + \alpha(\tilde{u}_h, \tilde{v}_h).
\end{aligned}
\tag{38}
$$

Problem (26)–(28) can be written now as: find $(u_h, p_h, \tilde{u}_h) \in V_{h,0} \times Q_h \times V_h$ such that

$$
\mathcal{B}(u_h, p_h, \tilde{u}_h; v_h, q_h, \tilde{v}_h) = (f, v_h) \qquad \forall (v_h, q_h, \tilde{v}_h) \in V_{h,0} \times Q_h \times V_h.
\tag{39}
$$

We assume that the family of finite element partitions $\{\mathcal{T}_h\}_{h>0}$ is quasi-uniform, that is, there exists a constant $\sigma > 0$ such that for all $h > 0$

$$
\min\{\operatorname{diam}(B_K) \mid K \in \mathcal{T}_h\} \geq \sigma \max\{\operatorname{diam}(K) \mid K \in \mathcal{T}_h\},
\tag{40}
$$

where B_K is the largest ball contained in $K \in \mathcal{T}_h$. Condition (40) is needed in order to have the following inverse estimate (see, e.g., Ref. [21]): there exist a constant $C > 0$ such that

$$
\|v_h\|_1 \leq \frac{C}{h}\|v_h\|_0 \qquad \forall v_h \in V_h.
\tag{41}
$$

From now onwards we use C, possibly with subscripts, to denote a positive constant independent of the mesh size, not necessarily the same at different occurrences.

A possible modification of the bilinear form $\mathcal{B}$ defined in Eq. (38) could be to define the parameter α and the terms that it multiplies elementwise, that is, $\alpha(\nabla p_h, \nabla q_h)$ could be replaced by

$$
\sum_{K \in \mathcal{T}_h} \alpha_K (\nabla p_h, \nabla q_h)_K,
\tag{42}
$$

and similarly for the rest of terms affected by α. This modification would allow to replace condition (40) by the weaker condition of nondegeneracy of the family $\{\mathcal{T}_h\}_{h>0}$, since only the elementwise version of the inverse estimate (41) would be needed (see Ref. [21]).

Let ∇Q_h denote the space of vector functions which are gradients of elements of Q_h and consider the vector space

$$
E_h := V_h + \nabla Q_h = E_{h,1} \oplus E_{h,2} \oplus E_{h,3},
\tag{43}
$$

where $E_{h,i}$, $i = 1, 2, 3$, are three mutually L^2 orthogonal subspaces defined as

$$E_{h,1} := V_{h,0}, \tag{44}$$

$$E_{h,2} := V_{h,0}^{\perp} \cap V_h, \tag{45}$$

$$E_{h,3} := V_h^{\perp}. \tag{46}$$

Let us denote by $P_{h,i}$ the orthogonal projection from E_h to $E_{h,i}$, and $P_{h,ij} := P_{h,i} + P_{h,j}$, $i, j = 1, 2, 3$. Also, we denote $E_{h,ij} := E_{h,i} \oplus E_{h,j}$. In order to prove that the pressure gradient in problem (39) is stable, we shall bound independently the three terms in the decomposition

$$\nabla p_h = P_{h,1}(\nabla p_h) + P_{h,2}(\nabla p_h) + P_{h,3}(\nabla p_h). \tag{47}$$

Finally, to obtain error estimates for the solution of problem (39) we shall make use of the approximation properties of the spaces $V_{h,0}$, Q_h and V_h. These can be written as follows. If $v \in H^r(\Omega) \cap V$, $r \geq 1$, and $q \in H^s(\Omega) \cap Q$, $s \geq 1$, there exist $\Pi_{h,1}(v) \in V_{h,0}$, $\Pi_{h,2}(q) \in Q_h$ and $\Pi_{h,3}(\nabla q) \in V_h$ such that

$$\|v - \Pi_{h,1}(v)\|_m \leq C_1 h^{k_1} \|v\|_{k_1}, \qquad k_1 = \min\{r, k_v + 1\} - m, \tag{48}$$

$$\|q - \Pi_{h,2}(q)\|_m \leq C_2 h^{k_2} \|q\|_{k_2}, \qquad k_2 = \min\{s, k_q + 1\} - m, \tag{49}$$

$$\|\nabla q - \Pi_{h,3}(\nabla q)\|_m \leq C_3 h^{k_3} \|\nabla q\|_{k_3}, \qquad k_3 = \min\{s - 1, k_v + 1\} - m. \tag{50}$$

3.2 Stability

We now prove that the solution of problem (39) is stable. For the pressure gradient, the three components appearing in Eq. (47) are bounded separately. The bound for the first one can be obtained independently of α, whereas the third component can be bounded only if $\alpha > 0$. Thus, the stability provided by the method in comparison with the standard problem (14) & (15) is precisely in the control over the term $P_{h,3}(\nabla p_h)$, that is, the component of the pressure gradient orthogonal to the space of continuous vector fields V_h.

The second component in Eq. (47) deserves special attention. It depends on the properties of the finite element spaces, and not on the problem actually solved. For the moment, we assume that there is a positive constant K_2' such that

$$\|\nabla q_h\| \leq K_2' \|P_{h,13}(\nabla q_h)\| \qquad \forall q_h \in Q_h, \tag{51}$$

which means that $\|P_{h,2}(\nabla q_h)\|$ can be bounded in terms of $\|P_{h,13}(\nabla q_h)\|$. In the next subsection we show that this is similar to the inf-sup condition of the standard problem, although much weaker and, in particular, verified when equal interpolation is used.

We also need to make an assumption on the behavior of α in terms of h: there is a constant α_0, independent of h, such that

$$\alpha \geq \alpha_0 h^2. \tag{52}$$

Under all these assumptions we can prove the following:

THEOREM 1. *Suppose that the family of finite element partitions $\{\mathcal{T}_h\}_{h>0}$ is such that the inverse estimate (41) and condition (51) hold, and suppose also that α satisfies (52). Then, there exists a unique solution to problem (39) that verifies the stability estimate*

$$\|\boldsymbol{u}_h\|_1 + h\|\nabla p_h\| \le C\|\boldsymbol{f}\| \tag{53}$$

for a constant C independent of h.

PROOF: Since the problem is finite-dimensional, it is enough to prove that (53) holds. From the definition of the bilinear form $\mathcal{B}$ in Eq. (38) it is easy to see that

$$\mathcal{B}(\boldsymbol{u}_h, p_h, \tilde{\boldsymbol{u}}_h; \boldsymbol{u}_h, p_h, \tilde{\boldsymbol{u}}_h) = a(\boldsymbol{u}_h, \boldsymbol{u}_h) + \alpha\|\nabla p_h - \tilde{\boldsymbol{u}}_h\|^2 = (\boldsymbol{f}, \boldsymbol{u}_h) \le \|\boldsymbol{f}\|\|\boldsymbol{u}_h\|_1 \tag{54}$$

From the coercivity of the bilinear form a (Eq. (8)) it follows that

$$\|\boldsymbol{u}_h\|_1 \le \frac{1}{K_a}\|\boldsymbol{f}\|. \tag{55}$$

On the other hand, Eq. (28) can now be written as $\tilde{\boldsymbol{u}}_h = P_{h,12}(\nabla p_h)$, and therefore from Eq. (54) it follows that

$$\alpha\|P_{h,3}(\nabla p_h)\|^2 = \alpha\|\nabla p_h - \tilde{\boldsymbol{u}}_h\|^2 \le \|\boldsymbol{f}\|\|\boldsymbol{u}_h\|_1$$

and, from estimate (55),

$$\|P_{h,3}(\nabla p_h)\| \le \frac{1}{\sqrt{\alpha K_a}}\|\boldsymbol{f}\|. \tag{56}$$

On the other hand, from Eq. (26) we have that

$$\begin{aligned}
\|P_{h,1}(\nabla p_h)\|^2 &= (\nabla p_h, P_{h,1}(\nabla p_h)) \\
&= (\boldsymbol{f}, P_{h,1}(\nabla p_h)) - a(\boldsymbol{u}_h, P_{h,1}(\nabla p_h)) \\
&\le \|\boldsymbol{f}\|\|P_{h,1}(\nabla p_h)\| + N_a\|\boldsymbol{u}_h\|_1\|P_{h,1}(\nabla p_h)\|_1 \\
&\le \left(\|\boldsymbol{f}\| + C\frac{N_a}{h}\frac{\|\boldsymbol{f}\|}{K_a}\right)\|P_{h,1}(\nabla p_h)\|,
\end{aligned} \tag{57}$$

where we have called N_a the norm of a and we have used the inverse estimate (41).
Estimate (53) follows now from (51), (55)–(57) and the assumption (52) on α. $\qquad\square$

3.3 A weakened inf-sup condition

The previous stability estimate, as well as the error estimate in section 3.4, depend on whether condition (51) holds or not. This condition is equivalent to the existence of a constant $K_2 > 0$ such that

$$\inf_{q_h \in Q_h} \sup_{v_h \in E_{h,13}} \frac{(\nabla q_h, v_h)}{\|\nabla q_h\|\|v_h\|} \ge K_2. \tag{58}$$

The equivalence between conditions (51) and (58) is easy to prove. In particular, it is found that the constant K_2 in Eq. (58) may be taken as $1/K_2'$, where K_2' is the constant in Eq. (51).

Condition (58) is similar to the standard Babuška-Brezzi condition for the discrete problem, that is, the discrete version of condition (9). The only difference is the space where v_h runs: it is $E_{h,13}$, and not only $E_{h,1}$ as it happens with the standard condition. This is possible due to the fact that control over $\|P_{h,3}(\nabla p_h)\|$ is provided by the formulation itself, without having to rely on a compatibility condition on the velocity and pressure finite element spaces. Thus, condition (58) is weaker than the standard one.

This section is devoted to show that condition (58) holds under mild conditions over the finite element interpolation and, in particular, that it is satisfied when using some equal interpolations. To this end, we apply a macroelement technique similar to that of Refs. [16, 17], from which we take part of our notation.

For each h, let $\mathcal{M}_h$ be a collection of macroelements covering Ω, a macroelement M being the union of one or more element domains of $\mathcal{T}_h$. 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 $G_M : M_0 \longrightarrow M$ such that:

(i) $G_M(M_0) = M$,

(ii) If $M_0 = \bigcup_{j=1}^m K_{0,j}$, then $M = \bigcup_{j=1}^m G_M(K_{0,j})$, where $K_{0,j} \in \mathcal{T}_{h_0}, j = 1, ..., m$.

(iii) $G_{M|K_0} = F_K \circ F_{K_0}^{-1}$, where $K = G_M(K_0)$ and F_K and F_{K_0} are the mappings from the reference element $\hat{K}$ to $K \in \mathcal{T}_h$ and to $K_0 \in \mathcal{T}_{h_0}$, respectively, introduced earlier.

Note 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 $\mathcal{E}_1, ..., \mathcal{E}_{n_c}$.

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 $\mathcal{T}_h$ 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 (51) holds in a macroelement, then it also holds in Ω:

LEMMA 1. If there exists a constant $C > 0$ such that

$$\|\nabla q_h\|_M \leq C\|P_{M,13}(\nabla q_h)\|_M \qquad \forall q_h \in Q_h, \tag{59}$$

for all $M \in \mathcal{M}_h$, then condition (51) holds for a constant K_2' independent of h.

PROOF: Let $q_h \in Q_h$ and let $v_{M,i}$ 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

$$v_h = \sum_M v_M = \sum_M \left(v_{M,1} + v_{M,3} \right), \tag{60}$$

Clearly, $v_{M,1} \in E_{M,1} \subset E_{h,1}$ $\forall M$ and thus $\sum_M v_{M,1} \in E_{h,1}$. Let $v_{h,12} \in E_{h,12}$. Since

$v_{h,12}|_M \in E_{M,12} = E_{M,3}^\perp$ (orthogonality in E_M) we have that

$$\int_\Omega v_{h,12} \cdot \left(\sum_M v_{M,3} \right) \, d\Omega = \sum_M \int_M v_{h,12}|_M \cdot v_{M,3} \, d\Omega = 0,$$

that is, $\sum_M v_{M,3} \in E_{h,12}^\perp = E_{h,3}$. Therefore, v_h in Eq. (60) 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 $\|v_h\|$:

$$\|v_h\|^2 = \int_\Omega \left(\sum_M v_M \right)^2$$

$$= \int_\Omega \left[\sum_M (v_M)^2 + 2 \sum_{M \neq M', M \cap M' \neq \emptyset} v_M \cdot v_{M'} \right] d\Omega$$

$$\leq \sum_M \|v_M\|^2 + 2 \sum_{M \neq M', M \cap M' \neq \emptyset} \|v_M\| \|v_{M'}\|$$

$$\leq \sum_M \|v_M\|^2 + \sum_{M \neq M', M \cap M' \neq \emptyset} \left(\|v_M\|^2 + \|v_{M'}\|^2 \right)$$

$$\leq (1 + N_M N_K) \sum_M \|v_M\|^2$$

$$\leq (1 + N_M N_K) \sum_M \|\nabla q_h\|_M^2$$

$$\leq (1 + N_M N_K) N_M \|\nabla q_h\|^2,$$

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

$$\|v_h\| \leq C_0 \|\nabla q_h\|. \tag{61}$$

On the other hand, from (59) it follows that

$$\int_\Omega \nabla q_h \cdot v_h \, d\Omega = \sum_M \int_M \nabla q_h \cdot v_M$$

$$= \sum_M \|P_{M,13}(\nabla q_h)\|^2$$

$$\geq \frac{1}{C'^2} \sum_M \|\nabla q_h\|_M^2$$

$$\geq \frac{1}{C^2} \|\nabla q_h\|^2. \tag{62}$$

But, using inequality (61),

$$\int_\Omega \nabla q_h \cdot v_h \, d\Omega = \int_\Omega P_{h,13}(\nabla q_h) \cdot v_h \, d\Omega \leq C_0 \|P_{h,13}(\nabla q_h)\| \|\nabla q_h\|. \tag{63}$$

The lemma follows combining inequalities (62) and (63) with $K_2' = C_0 C^2$. $\qquad\square$

The next step is to give sufficient conditions for property (59) to hold. First we give a rather technical lemma whose proof is omitted:

LEMMA 2. *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 such that*

$$\lim_{\mu\to 0}\left[\sup_{y_\mu\in Y_\mu}\inf_{y\in Y}\operatorname{dist}(y_\mu,y)\right] = \lim_{\mu\to 0}\left[\sup_{y\in Y}\inf_{y_\mu\in Y_\mu}\operatorname{dist}(y_\mu,y)\right] = 0. \tag{64}$$

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 the second argument. Then

$$\lim_{\mu\to 0}\left[\inf_{x\in X}\sup_{y_\mu\in Y_\mu}|f_\mu(x,y_\mu)|\right] = \inf_{x\in X}\sup_{y\in Y}|f(x,y)|. \tag{65}$$

This result is used now to prove the following:

LEMMA 3. *Let $\mathcal{E}_i$ be one of the equivalence classes introduced above, $i\in\{1,2,...,n_c\}$, and suppose that the following condition holds:*

$$\exists M_0\in\mathcal{E}_i \text{ such that } \quad \forall q\in Q_{M_0} \int_{M_0}\nabla q\cdot v\, dM = 0 \ \ \forall v\in E_{M_0,13} \Rightarrow \nabla q = 0. \tag{66}$$

Then, there exists a constant $C_i>0$ such that, for all $M\in\mathcal{E}_i$,

$$\|\nabla q\|_M \le C_i\|P_{M,13}(\nabla q)\|_M \quad \forall q\in Q_M. \tag{67}$$

PROOF: Let us consider the following function defined on the class $\mathcal{E}_i$:

$$\beta(M) = \inf_{q\in Q_M}\sup_{v\in E_{M,13}}\frac{(\nabla q,v)_M}{\|\nabla q\|_M\|v\|_M}. \tag{68}$$

Inequality (67) is equivalent to saying that $\beta(M)\ge 1/C_i$ for all $M\in\mathcal{E}_i$. This can be proved as the equivalence between (51) and (58).

From assumption (66) it is easy to see that $\beta(M)>0$ for all $M\in\mathcal{E}_i$. 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 $\{\mathcal{T}_h\}_{h>0}$ (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, $M' \in \mathcal{E}_i$. We want to show that $\beta(M') \to \beta(M)$ as $M' \to M$. Let $G : M \to M'$ be the homeomorphism that relates M and M'. We denote its Jacobian matrix (piecewise continuous) by DG. Let also

$$J' := \max_{x' \in M'} |DG^{-1}|(x'), \qquad j' := \min_{x' \in M'} |DG^{-1}|(x'), \qquad (69)$$

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 (69) depend on the macroelement M' and tend to 1 as $M' \to M$, that is, as $G \to I$.

Let us write the function β as

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

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

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

$$\int_{M'} \nabla'q' \cdot v' \, dM' = \int_M \nabla q \cdot DG^{-1} \cdot v |DG| \, dM,$$

$$\int_{M'} v' \cdot v' \, dM' = \int_M v \cdot v |DG| \, dM,$$

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

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

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

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

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

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

which is in general $\neq 0$ since $v_{12} \circ G^{-1} |DG| \notin E_{M',12}$ if $|DG|$ is not continuous. Therefore, if $S_G = G^*S'$ then $S_G \neq S$.

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

$$\beta(M') = \inf_{q \in Q_{M,0}} \sup_{v \in S_G} f_G(\nabla q, v).$$

Now we use lemma 2 to prove the continuity of β. Let

$$Z = \left\{ v \in E_M \mid \frac{1}{2} \leq \|v\|_M \leq 2 \right\}.$$

We have that

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

and thus $\sqrt{j'} \leq \|G^* v\|_M \leq \sqrt{J'}$, with j' and J' defined in Eq. (69). If we take M' sufficiently close to M, $j' > 1/4$ and $J' < 4$, so that $S \subset Z$ and $S_G \subset Z$.

It is now easy to prove that $f(\nabla q, v)$ is uniformly continuous in the second argument in $Q_{M,0} \times Z$ and that $f_G(\nabla q, v)$ converges uniformly to $f(\nabla q, v)$ in $Q_{M,0} \times Z$. To apply lemma 2 it remains to check condition (64) with $Y = S$ and $Y_\mu = S_G$, the parameter μ being now replaced by the function G and $\mu \to 0$ by $G \to I$.

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

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

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

$$\sup_{v_G \in S_G} \inf_{v \in S} \mathrm{dist}(v_G, v) \to 0 \quad \text{as} \quad G \to I. \tag{70}$$

Also, given $\bar{v} = v_1 + v_3 \in S$, with $v_1 \in E_{M,1}$ and $v_3 \in E_{M,3}$, let

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

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

$$\sup_{v \in S} \inf_{v_G \in S_G} \mathrm{dist}(v_G, v) \to 0 \quad \text{as} \quad G \to I. \tag{71}$$

From (70) and (71) it may be concluded that hypothesis (64) holds in the present situation and ultimately that the function β defined in Eq. (68) is continuous, which is what had to be proved. $\qquad \square$

Combining lemmas 1 and 3 we obtain the following result:

THEOREM 2. Suppose that for all the equivalence classes $\mathcal{E}_i$, $i = 1, ..., n_c$ of macroelements of $\{\mathcal{T}_h\}_{h>0}$ condition (66) holds. Then, there exists a constant $K_2 > 0$, independent of h, for which the inf-sup condition (58) is verified.

PROOF: Let $C = \min\{C_1, ..., C_{n_c}\}$, where C_i is the constant for the equivalence class $\mathcal{E}_i$ established by lemma 3. 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 1. The theorem follows from the equivalence between (51) and (58). $\qquad\qquad\square$

From this result we see that condition (66) is the key for proving that the finite element formulation is stable. Again, it is similar to the condition obtained in Ref. [16], the only difference being the space where the function $\boldsymbol{v}$ runs: $\boldsymbol{E}_{M,1}$ in that reference, $\boldsymbol{E}_{M,13}$ in our case.

Next we prove that condition (66) is verified in a simple case using equal interpolation, namely, using complete polynomials of degree $k \geq 1$ for simplicial elements. According to theorem 2, we prove it on arbitrary classes of macroelements. The only restriction on them is specified next. The macroelement technique can also be applied to other cases of interest, such as the use of tensor product polynomials for quadrilaterals and hexahedra.

PROPOSITION 1. Suppose that $k_v = k_q = k$ and that $\hat{K}$ is a simplex. Let $\mathcal{E}$ 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 (66) is satisfied on $\hat{M}$.

PROOF: We prove condition (66) by imposing continuity of ∇q on $\hat{M}$ rather than orthogonality to $E_{\hat{M},3}$, due to the difficulty of characterizing this space. Orthogonality to $E_{\hat{M},1}$ is enforced directly.

Let us consider the case of linear elements ($k = 1$) first. For a given $q \in Q_{\hat{M}}$, ∇q is constant on each element $K \subset \hat{M}$; if we assume ∇q is continuous, it must be constant on $\hat{M}$. Since we have assumed the existence of at least one vertex P interior to $\hat{M}$, orthogonality of ∇q with respect to all velocity fields which take arbitrary values on P and zero at the nodes of $\partial\hat{M}$ implies the vanishing of ∇q.

Let us now turn to the case of higher order elements ($k > 1$). Given $q \in Q_{\hat{M}}$, for each $K \subset \hat{M}$ one has that the components of $\nabla q|_K$ belong to $P_{k-1}(K)$. Thus, if these components are continuous, they can be determined by their values at the nodes of $\hat{M}$ corresponding to a piecewise interpolation with polynomials of degree $k - 1$. Let n_{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 of ∇q with respect to all continuous vector functions that take arbitrary values at the nodes of $\text{Int}(\hat{M})$ are linearly independent restrictions, it is enough to prove that $n_{\text{int}} \geq n_{k-1}$.

Let us 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. Suppose first that all the elements have at most one edge on $\partial\hat{M}$. If

n_{ele} is the number of elements in $\hat{M}$ with one edge on the boundary, n_{ver} the number of edges with only one node on the boundary and n_{edg} the number of edges on $\partial \hat{M}$, we have that

$$n_{\text{int}} - n_{k-1} \geq n_{\text{ele}} \times [(k-1)(k-2)/2 - (k-2)(k-3)/2]$$
$$+ n_{\text{ver}} \times [(k+1) - 1 - k] + n_{\text{edg}} \times [0 - (k-2)]$$
$$\geq n_{\text{ele}} - n_{\text{edg}},$$

which is ≥ 0 due to the assumption on $\hat{M}$. If now we include the 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 we still have that $n_{\text{int}} - n_{k-1} \geq 0$.

Finally, in the three-dimensional case each tetrahedron $K \subset \hat{M}$ has $(k-1)(k-2) \times (k-3)/6$ nodes of $P_k(K)$ on $\text{Int}(K)$, whereas $(k-2)(k-3)(k-4)/6$ of $P_{k-1}(K)$. As before, consider first 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_{ver} the number of edges with one node on $\partial \hat{M}$, n_{bou} the number of faces on $\partial \hat{M}$ and n_{edg} the number of edges on $\partial \hat{M}$, we have that

$$n_{\text{int}} - n_{k-1} \geq n_{\text{ele}} \times [(k-1)(k-2)(k-3)/6 - (k-2)(k-3)(k-4)/6]$$
$$+ n_{\text{fac}} \times [(k-1)(k-2)/2 - (k-2)(k-3)/2] + n_{\text{ver}} \times [(k+1) - 1 - k]$$
$$+ n_{\text{bou}} \times [0 - (k-2)(k-3)/2] + n_{\text{edg}} \times [0 - (k-2)]$$
$$\geq n_{\text{ele}} + n_{\text{fac}} - n_{\text{edg}} - n_{\text{bou}},$$

which is again ≥ 0. If now we consider elements with two faces on $\partial \hat{M}$, for each of these elements 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 $(k-2)(k-3)(k-4)/6 + (k-2)(k-3) + (k-2)$. $\qquad \square$

3.4 Convergence

Once the stability of the finite element approximation of the Stokes problem given by Eqs. (26)–(28) or, equivalently, by Eq. (39) has been established, we turn now our attention to the convergence analysis. We give error estimates in the H^1 norm for both velocity and pressure. In this particular case, this is the 'natural' norm for proving these error estimates, since it is the norm in which we have proven stability. After this one can prove convergence in the L^2 norm by using classical duality arguments as in Ref. [2] for the GLS method.

THEOREM 3. Suppose that the family of finite element partitions $\{\mathfrak{T}_h\}_{h>0}$ and the associated finite element spaces V_h and Q_h are such that the inverse estimate (41) and the stability condition (58) hold. Assume also that the parameter α in Eq. (27) is such that

$$\alpha_0 h^2 \leq \alpha \leq \alpha_1 h^2, \qquad (72)$$

with α_0, α_1 independent of h. Then the solution of problem (26)–(28) satisfies the error estimate

$$\|u - u_h\|_1 + h\|\nabla p - \nabla p_h\| + h\|\nabla p - \tilde{u}_h\| \leq C E(h) \qquad (73)$$

where

$$E(h) := \inf_{v_h \in V_{h,0}} \|u - v_h\|_1 + \frac{1}{h} \inf_{v_h \in V_{h,0}} \|u - v_h\| + \inf_{q_h \in Q_h} \|p - q_h\|$$
$$+ h \inf_{q_h \in Q_h} \|\nabla p - \nabla q_h\| + h \inf_{\tilde{v}_h \in V_h} \|\nabla p - \tilde{v}_h\|$$

and u and p are the solution of the continuous problem (6) & (7).

PROOF: Since problem (39) is consistent we have that

$$\mathcal{B}(u, p, \nabla p; v_h, q_h, \tilde{v}_h) = (f, v_h) \qquad \forall (v_h, q_h, \tilde{v}_h) \in V_{h,0} \times Q_h \times V_h.$$

Substracting this equation from Eq. (39) and taking as test functions $(v_h - u_h, q_h - p_h, \tilde{v}_h - \tilde{u}_h) \in V_{h,0} \times Q_h \times V_h$ we obtain

$$\mathcal{B}(u - u_h, p - p_h, \nabla p - \tilde{u}_h; u - u_h, p - p_h, \nabla p - \tilde{u}_h)$$
$$= \mathcal{B}(u - u_h, p - p_h, \nabla p - \tilde{u}_h; u - v_h, p - q_h, \nabla p - \tilde{v}_h), \tag{74}$$

for all $(v_h, q_h, \tilde{v}_h) \in V_{h,0} \times Q_h \times V_h$. Using the expression of the form $\mathcal{B}$ given in Eq. (38), from Eq. (74) it is found that

$$a(u - u_h, u - u_h) + \alpha(\tilde{u}_h - \nabla p_h, \tilde{u}_h - \nabla p_h) = a(u - u_h, u - v_h)$$
$$+ (\nabla p - \nabla p_h, u - v_h) + b(p - q_h, u - u_h) + \alpha(\tilde{u}_h - \nabla p_h, \tilde{v}_h - \nabla q_h).$$

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

$$\|u - u_h\|_1^2 + \frac{\alpha}{K_a}\|\tilde{u}_h - \nabla p_h\|^2 \leq C \left[\|u - u_h\|_1\|u - v_h\|_1 + \|\nabla p - \nabla p_h\|\|u - v_h\|\right.$$
$$\left. + \|p - q_h\|\|u - u_h\|_1 + \alpha\|\tilde{u}_h - \nabla p_h\|\|\tilde{v}_h - \nabla q_h\|\right]. \tag{75}$$

Let us denote by $E_m(\cdot)$ the error in the H^m norm of either u, p or ∇p and $I_m(u) := \|u - v_h\|_m$, $I_0(p) := \|p - q_h\|$, $I_0(\nabla p) := \|\nabla p - \nabla q_h\|$ and $I_0(\tilde{u}) := \|\nabla p - \tilde{v}_h\|$. Also, let $G := \|\tilde{u}_h - \nabla p_h\|$. We can thus write Eq. (75) as

$$E_1^2(u) + \frac{\alpha}{K_a}G^2 \leq C\left[E_1(u)I_1(u) + E_0(\nabla p)I_0(u) + I_0(p)E_1(u) + \alpha G\|\tilde{v}_h - \nabla q_h\|\right]. \tag{76}$$

Since

$$\|\tilde{v}_h - \nabla q_h\| \leq \|\tilde{v}_h - \nabla p\| + \|\nabla p - \nabla q_h\| = I_0(\tilde{u}) + I_0(\nabla p),$$

from Eq. (76) we obtain

$$E_1^2(u) + \frac{\alpha}{K_a}G^2 \leq C\left[E_1(u) + hE_0(\nabla p) + \alpha^{1/2}G\right]$$
$$\times \max\left\{I_1(u), I_0(p), \frac{1}{h}I_0(u), \alpha^{1/2}I_0(\tilde{u}), \alpha^{1/2}I_0(\nabla p)\right\}. \tag{77}$$

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

$$
\begin{aligned}
E_0(\nabla p) &\leq \|\nabla p - P_{h,12}(\nabla q_h)\| + \|P_{h,12}(\nabla q_h) - \nabla p_h\| \\
&\leq \|\nabla p - P_{h,12}(\nabla q_h)\| + \|P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)\| \\
&\quad + \|P_{h,2}(\nabla q_h) - P_{h,2}(\nabla p_h)\| + \|P_{h,3}(\nabla p_h)\|.
\end{aligned}
\tag{78}
$$

Using now the stability condition (51) we obtain

$$
\begin{aligned}
\|P_{h,2}(\nabla q_h) - P_{h,2}(\nabla p_h)\| &\leq C\|P_{h,13}(\nabla q_h) - P_{h,13}(\nabla p_h)\| \\
&\leq C\|P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)\| \\
&\quad + C\|P_{h,3}(\nabla q_h)\| + C\|P_{h,3}(\nabla p_h)\|.
\end{aligned}
\tag{79}
$$

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 Eq. (79) it is found that

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

Using this inequality in the estimate (78) we get

$$
\begin{aligned}
E_0(\nabla p) &\leq (1+C)\|\nabla p - P_{h,12}(\nabla q_h)\| + (1+C)\|P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)\| \\
&\quad + (1+C)\|P_{h,3}(\nabla p_h)\| + C\|\nabla p - \nabla q_h\|.
\end{aligned}
\tag{80}
$$

Let us bound now the different terms in Eq. (80). If we still denote by $P_{h,12}$ the extension of the projection onto $E_{h,12} = V_h$ from the whole space $L^2(\Omega)$, we have that

$$
\|\nabla p - P_{h,12}(\nabla q_h)\| \leq \|\nabla p - P_{h,12}(\nabla p)\| + \|P_{h,12}(\nabla p) - P_{h,12}(\nabla q_h)\|.
\tag{81}
$$

Since

$$
(\nabla p - P_{h,12}(\nabla p), \tilde{v}_h) = 0 \qquad \forall \tilde{v}_h \in V_h,
$$

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

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

that is,

$$
\|\nabla p - P_{h,12}(\nabla p)\| \leq I_0(\tilde{u}).
\tag{82}
$$

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

$$\|P_{h,12}(\nabla p) - P_{h,12}(\nabla q_h)\| \leq \|P_{h,12}\|\|\nabla p - \nabla q_h\|$$
$$\leq I_0(\nabla p). \tag{83}$$

Using inequalities (82) and (83) in (81) we obtain

$$\|\nabla p - P_{h,12}(\nabla q_h)\| \leq I_0(\tilde{u}) + I_0(\nabla p). \tag{84}$$

The second term in Eq. (80) can be bounded using the first equation of the problem, that is, Eq. (26), and making use of the inverse estimate (41):

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

where N_a is the norm of a. Therefore

$$\|P_{h,1}(\nabla q_h) - P_{h,1}(\nabla p_h)\| \leq C\frac{N_a}{h}E_1(u) + I_0(\nabla p). \tag{85}$$

The third term in Eq. (80) is $(1 + C)G$ and the last one is $CI_0(\nabla p)$. So, using bounds (84) and (85) in (80) we obtain

$$E_0(\nabla p) \leq C\left[I_0(\tilde{u}) + I_0(\nabla p) + \frac{N_a}{h}E_1(u) + G \right], \tag{86}$$

and using this in (77) we get

$$\begin{aligned}
E_1^2(u) + \frac{\alpha}{K_a}G^2 &\leq C\left[E_1(u) + hI_0(\nabla p) + hI_0(\tilde{u}) + (h + \alpha^{1/2})G \right] \\
&\times \max\left\{ I_1(u), I_0(p), \frac{1}{h}I_0(u), \alpha^{1/2}I_0(\tilde{u}), \alpha^{1/2}I_0(\nabla p) \right\}. \tag{87}
\end{aligned}$$

From the behavior assumed for the parameter α, Eq. (87) implies that there exist constants C_1 and C_2 such that

$$E_1(u) \leq C_1 \max\left\{ I_1(u), \frac{1}{h}I_0(u), hI_0(\nabla p), I_0(p), hI_0(\tilde{u}) \right\}, \tag{88}$$

$$G \leq \frac{C_2}{h} \max\left\{ I_1(u), \frac{1}{h}I_0(u), hI_0(\nabla p), I_0(p), hI_0(\tilde{u}) \right\}. \tag{89}$$

Equation (88) is the error estimate for the velocity. Using (88) and (89) in (86) we obtain the error estimate for the pressure:

$$hE_0(\nabla p) \leq C_3 \max \left\{ I_1(u), \frac{1}{h}I_0(u), hI_0(\nabla p), I_0(p), hI_0(\tilde{u}) \right\}. \tag{90}$$

On the other hand,

$$\|\tilde{u}_h - \nabla p\| = \|\nabla p_h - \nabla p - P_{h,3}(\nabla p_h)\| \leq E_0(\nabla p) + G. \tag{91}$$

The theorem follows combining inequalities (88)–(91). $\qquad\square$

Clearly, estimate (73) is optimal. From the approximation properties (48)–(50) it follows that if $u \in H^r(\Omega) \cap V$, $r \geq 1$, and $p \in H^s(\Omega) \cap Q$, $s \geq 1$, then the error function $E(h)$ in Eq. (73) behaves like h^k, with $k = \min\{r-1, s, k_v, k_q+1\}$.

It is also remarkable that we have had to use the fact that $\alpha_0 h^2 \leq \alpha \leq \alpha_1 h^2$, whereas to prove stability in Theorem 1 we only used that $\alpha_0 h^2 \leq \alpha$. Thus, the behavior of α is dictated by the stability and convergence analysis. To make it dimensional, we take it as $\alpha = \alpha_0 h^2/\nu$, where α_0 is a dimensionless parameter.

4 Numerical tests

In this section we present two simple numerical examples of the solution of problem (26)–(28). In the implementation on the computer, we have solved this problem iteratively, first updating u_h from Eq. (26) using a guess for p_h, then updating p_h from Eq. (27) using a guess for $\tilde{u}_h$ and the current u_h and finally updating $\tilde{u}_h$ from Eq. (28) using the p_h just computed. Although the performance of this scheme has not been completely satisfactory, the problems to be solved in this iterative process are very simple, all of them requiring the solution of algebraic systems with symmetric and positive-definite matrices. We have solved them using the conjugate gradient method.

4.1 A test with analyical solution

The purpose of this first test is to check numerically the convergence rates predicted by theorem 3, that is, the convergence of u_h to u in the H^1 norm and the convergence of ∇p_h and $\tilde{u}_h$ to ∇p in the L^2 norm. For that purpose we consider the test problem presented in Ref. [22], in which Ω is the unit square and the force term is selected so that the solution of problem (1)–(3) with $\nu = 1$ is $u = (u_x, u_y)$, with $u_x = x^2(1 - x^2)(2y - 6y^2 + 4y^3)$ and $u_y = (-2x + 6x^2 - 4x^3)y^2(1 - y)^2$, and $p = x - x^2$.

We have solved the problem using the P_1 and the P_2 elements. We have also solved the standard problem (14) & (15) using the mixed P_2/P_1 element (continuous quadratic
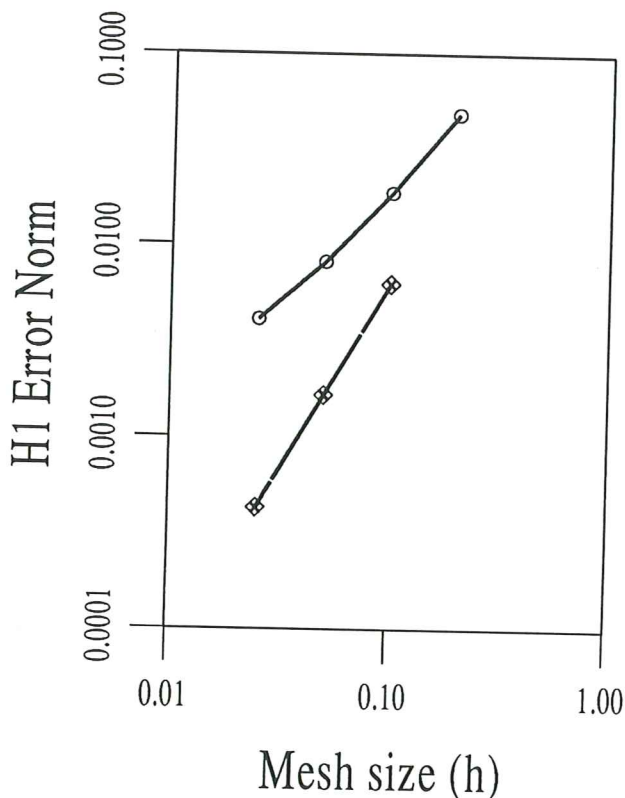
Figure 1. Convergence of $\|\nabla u - \nabla u_h\|$. $\circ$: P_1 element; $\diamond$: P_2 element; $\times$: mixed P_2/P_1 element.

velocities, continuous linear pressures), which satisfies the discrete counterpart of the inf-sup condition (9). All the finite element meshes that we have employed are uniform.

We have plotted in Fig. 1 the convergence of the velocity. As expected, the rate of convergence is 1 for the P_1 element and 2 for the P_2 and the mixed P_2/P_1 elements. In this case, these last two elements give the same error for the three meshes that we have used.

Figure 2 shows the convergence of the pressure gradient. For the P_2 and the mixed P_2/P_1 elements the rate of convergence is 1, being the absolute error of the latter greater than that of the former. What is not predicted by Theorem 3 is the convergence of the pressure gradient for the P_1 element observed in Fig. 2. Notice that for the finest mesh the rate of convergence found for the first three meshes is lost.

Convergence of the projected pressure gradient ($\tilde{u}_h$) is very similar to that of the pressure gradient itself for the P_1 and P_2 elements. It is shown in Fig. 3.

4.2 Behavior of the pressure near the boundary

This second example is intented to discuss a misbehavior of the pressure near the boundary using the GLS formulation, as described in Ref. [23]. In essence, the GLS method consists in replacing $\tilde{u}_h$ in Eq. (27) by $g_h := f + \nu \Delta u_h$ and evaluating the integrals involved in the L^2 inner product element by element (see Ref. [4]). In other words, instead of using
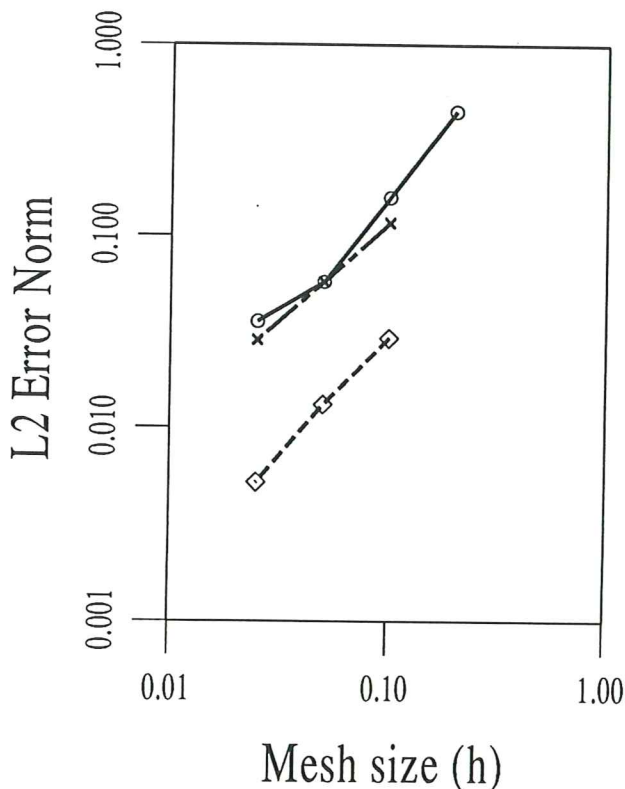
Figure 2. Convergence of $\|\nabla p - \nabla p_h\|$. o: P_1 element; ◇: P_2 element; ×: mixed P_2/P_1 element.

the projection of the pressure gradient onto the space of continuous vector functions, the expression resulting from the differential form of the momentum equation (that is, Eq. (1)) evaluated on each element is employed.

Although the rate of convergence of the method in the H^1 and L^2 norms is optimal [2, 4], the pressure may be poorly approximated near the boundary. Suppose that $f = 0$ and that the flow is induced by a non-homogeneous Dirichlet condition. Then, $g_h = \nu \Delta u_h$ and this approximates $\nu \Delta u$ within the elements only using polynomials of order $k \geq 2$. In the case of linear elements, $g_h = 0$. If we take $\tilde{u}_h = 0$ in problem (33)–(37), it is clear that the pressure verifies (weakly) the condition $\partial p_h / \partial n = 0$ on Γ (see Eq. (37)), which is wrong. Therefore, we may expect an incorrect pressure near the boundary, especially using linear elements. To overcome this, a modification of the GLS method including a boundary term was presented in Ref. [23]. This problem does not appear in the formulation introduced in this paper.

One of the numerical examples of Ref. [23] consists in the solution of a Poiseuille flow in a trapezoidal domain. We have also solved this problem using two meshes of P_1 elements with 13×13 and 25×25 nodes uniformly distributed along the sides. The first mesh is shown in Fig. 4. For this problem, $f = 0$ and a parabolic velocity profile is prescribed at both the inlet and the outlet; on the top and bottom edges $u = 0$ is prescribed. The pressure gradient in this case must be constant.

Pressure contours computed with both the GLS method and the formulation presented in this paper are shown in Fig. 5. In spite of the improvement obtained with the mesh
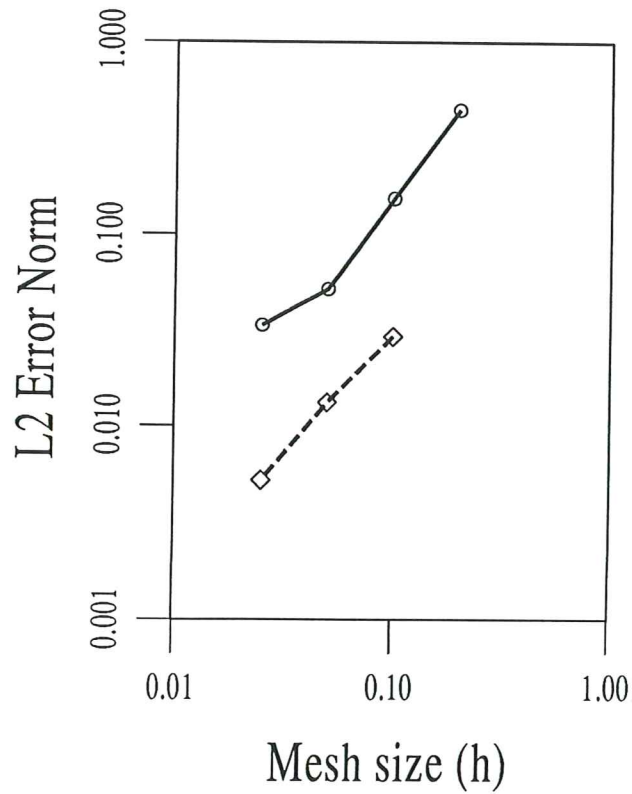
Figure 3. Convergence of $\|\nabla p - \tilde{u}_h\|$. $\circ$: P_1 element; $\diamond$: P_2 element.
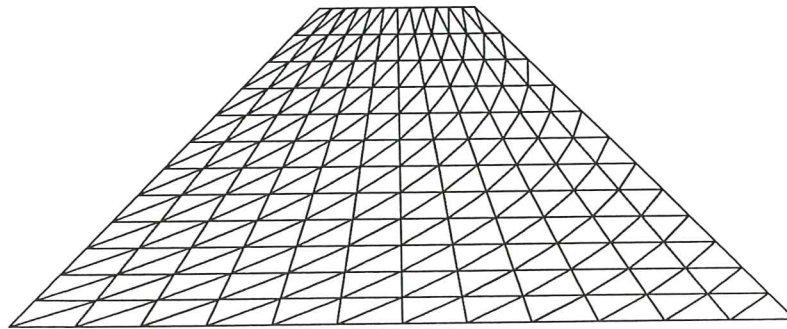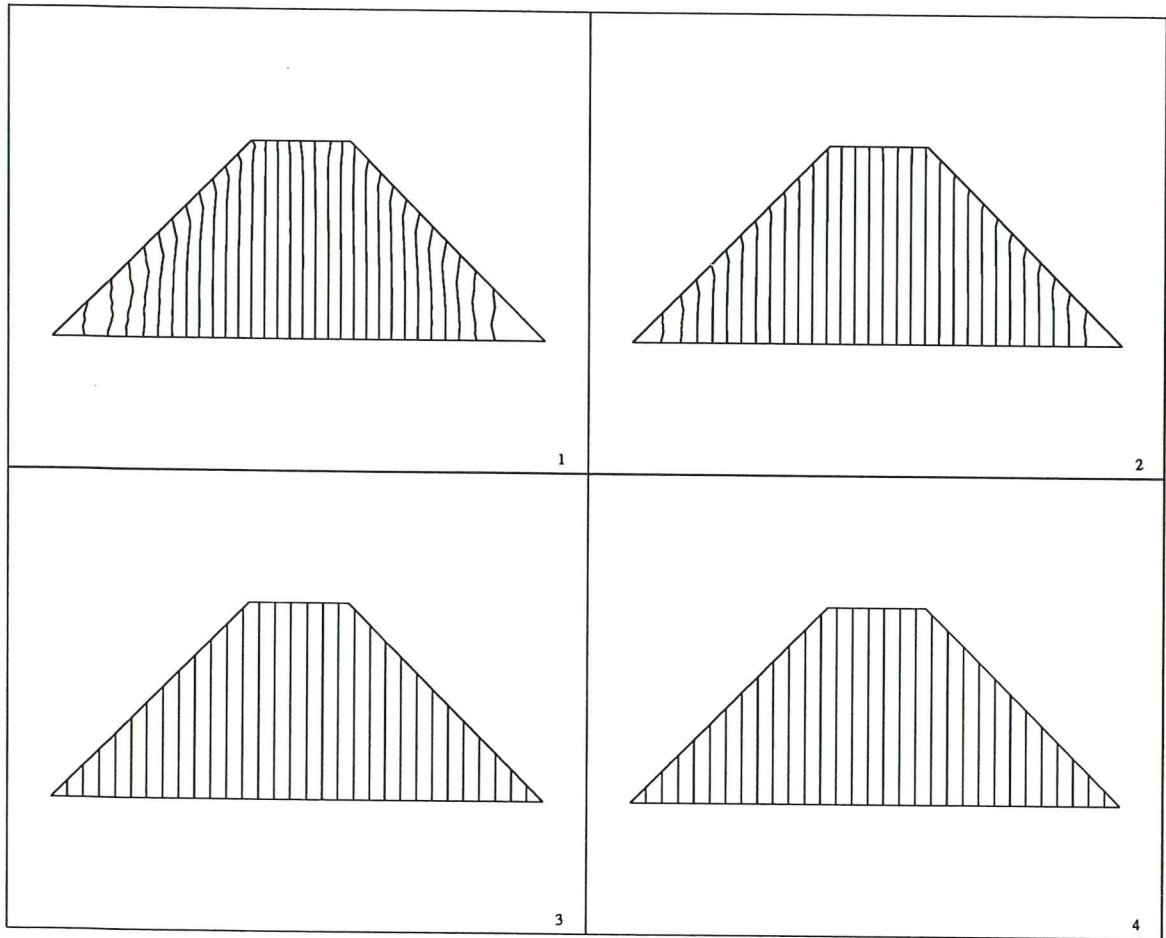

Figure 4. Finite element mesh using 12×12 P_1 elements for the Poiseuille flow example.

of 25×25 nodes with respect to that of 13×13 nodes, pressure contours using the GLS method are wrong near the boundary, whereas results solving problem (26)–(28) are correct on both meshes.

Figure 5. Pressure contours for the Poiseuille flow example. (1): GLS method, coarse mesh; (2): GLS method, fine mesh; (3): Present method, coarse mesh; (4): Present method, fine mesh.

5 Conclusions

We have analyzed in this paper a finite element formulation for the Stokes problem that has a compatibility restriction for the velocity and pressure interpolations weaker than that of the standard approach. We have seen that this restriction can be formulated in terms of a condition that involves only macroelements, that is, assemblies or 'patches' of elements, and that is easy to check. In particular, it is verified using equal interpolation for the velocity and the pressure. From this compatibility condition we have proved stability and obtained optimal error estimates. Also, the formulation depends on an algorithmic parameter α whose dependence on the mesh diameter has been dictated by this convergence analysis.

The development of efficient numerical methods for solving the Stokes problem from the formulation that we have presented remains open. Although the straight solution of problem (29)–(31) is unacceptable from the computational point of view, we believe that the idea of using the projection of the pressure gradient onto the velocity space can be used for the design of practical numerical algorithms, especially in the context of iterative

schemes for the Navier-Stokes equations.

We think that the fact that the formulation presented here allows equal velocity-pressure interpolation makes it interesting by itself. However, in our opinion the real interest of our analysis relies on the fact that it explains why equal interpolation is possible in some commonly used fractional step methods and, in general, in any method that in terms of a primitive u-p approach introduces the difference between the two discrete Laplacians appearing in Eq. (32).

References

[1] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer-Verlag, 1991.

[2] F. Brezzi and J. Douglas, 'Stabilized mixed methods for the Stokes problem', *Numer. Math.*, **53**, 225–235, (1988).

[3] J. Douglas and J. Wang, 'An absolutely stabilized finite element method for the Stokes problem', *Math. Comput.*, **52**, 495–508, (1989).

[4] T.J.R. Hughes, L.P. Franca and M. Balestra, 'A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuška-Brezzi condition: a stable Petrov-Galerkin formulation for the Stokes problem accommodating equal-order interpolations', *Comput. Meths. Appl. Mech. Engrg.*, **59**, 85–99, (1986).

[5] L. Franca, T.J.R. Hughes, A.F.D. Loula and I. Miranda, 'A new family of stable elements for nearly incompressible elasticity based on a mixed Petrov-Galerkin finite element formulation', *Numer. Math.*, **53**, 123–141 (1988).

[6] M. Fortin, and S. Boivin, 'Iterative stabilization of the bilinear velocity-constant pressure element', *Int. J. Numer. Meth. Fluids*, **10**, 125–140 (1990).

[7] P.J. Silvester and N. Kechkar, 'Stabilised bilinear-constant velocity-pressure finite elements for the conjugate gradient solution of the Stokes problem', *Comput. Meths. Appl. Mech. Engrg.*, **79**, 71–86, (1990).

[8] J. Atanga and D. Silvester, 'Iterative methods for stabilized mixed velocity-pressure finite elements', *Int. J. Numer. Meth. Fluids*, **14**, 71–81 (1992).

[9] C. Vincent and R. Boyer, 'A preconditioned conjugate gradient Uzawa-type method for the solution of the Stokes problem by mixed Q1-P0 stabilized finite elements', *Int. J. Numer. Meth. Fluids*, **14**, 289–298 (1992).

[10] L. Franca and R. Stenberg, 'Error analysis of some Galerkin least-squares methods for the elasticity equations', *SIAM J. Numer. Anal.*, **28**, 1680–1697 (1991).

[11] G.E. Schneider, G.D. Raithby and M.M. Yovanovich, 'Finite element analysis of incompressible fluid flow incorporating equal order pressure and velocity interpolation', in *Numerical Methods in Laminar and Turbulent Flow*, C. Taylor, K. Morgan, C.A. Brebbia (eds.), Pentech Press, Plymouth, 1978.

[12] M. Kawahara and K. Ohmiya, 'Finite element analysis of density flow using the velocity correction method', *Int. J. Num. Meth. Fluids*, **5**, 981–993 (1985).

[13] J.G. Rice and R.J. Schnipke, 'An equal-order velocity-pressure formulation that does not exhibit spurious pressure modes', *Comput. Meth. Appl. Mech. Engng.*, **58**, 135–149 (1986).

[14] O.C. Zienkiewicz and R. Codina, 'A general algorithm for compressible and incompressible flow. Part I: the split, characteristic based scheme' *Int. J. Num. Meth. Fluids*, **20**, 869–885 (1995).

[15] R. Codina, M. Vázquez and O.C. Zienkiewicz, 'A fractional step method for compressible flows: boundary conditions and incompressible limit', in the proceedings of the *IX International Conference on Finite Elements in Fluids*, Venezia, Italy, 1995.

[16] R. Stenberg, 'Analysis of mixed finite element methods for the Stokes problem: a unified approach', *Math. Comput.*, **42**, 9–23 (1984).

[17] R. Stenberg, 'A technique for analysing finite element methods for viscous incompressible flow', *Int. J. Numer. Meth. Fluids*, **11**, 935–948 (1990).

[18] R. Temam, *Navier-Stokes equations*, (North-Holland, 1984).

[19] A.J. Chorin, 'Numerical solution of the Navier-Stokes equations', *Math. Comput.*, **22**, 745–762 (1968).

[20] R. Temam, 'Sur l'approximation de la solution des équations de Navier-Stokes par la méthode des pas fractionaires (I)', *Arch. Rat. Mech. Anal.*, **32**, 135–153 (1969).

[21] S.C. Brenner and L.R. Scott, *The mathematical theory of finite element methods*, (Springer-Verlag, 1994).

[22] J.T. Oden and O.P. Jacquotte, 'Stability of some mixed finite element methods for Stokesian flows', *Comput. Meth. Appl. Mech. Engng.*, **43**, 231–247 (1984).

[23] J.J. Droux and T.J.R. Hughes, 'A boundary integral modification of the Galerkin least squares formulation for the Stokes problem', *Comput. Meth. Appl. Mech. Engng.*, **113**, 173–182 (1994).