Asociación Argentina
de Mecánica Computacional

Mecánica Computacional Vol XXIX, págs. 4707-4720 (artículo completo)
Eduardo Dvorkin, Marcela Goldschmit, Mario Storti (Eds.)
Buenos Aires, Argentina, 15-18 Noviembre 2010

NUMERICAL STUDY OF MIXED LEAST-SQUARES FINITE ELEMENT FORMULATIONS FOR TRANSIENT ADVECTION-DIFFUSION EQUATIONS

Regina C. Leal Toledo[a,b] **and Vitoriano Ruas**[b,c]

[a]*Departamento de Ciência da Computação, Universidade Federal Fluminense, rua Passo da Pátria 156, Bloco E, 3^o andar, Niterói, Rio de Janeiro state, CEP 24210-240, Brazil, leal@ic.uff.br*

[b]*Programa d Pós-graudução em Ciência da Computação, Universidade Federal Fluminense, rua Passo da Pátria 156, Bloco E, 3^o andar, Niterói, Rio de Janeiro state, CEP 24210-240, Brazil, vitoriano.ruas@ic.uff.br*

[c]*UPMC Univ. Paris 6, UMR 7190, Institut Jean Rond d'Alembert / CNRS, 4 place Jussieu, couloir 55-65, 4^e étage, 75252 Paris cedex 05, France, vitoriano.ruas@upmc.fr*

Keywords: Advection-diffusion, Crank-Nicholson, Finite elements, Galerkin, Least-squares, Reaction, Transient.

Abstract. A mixed finite element scheme designed for solving the time-dependent advection-diffusion equations expressed in terms of both the primal unknown and its flux, incorporating or not a reaction term, is studied. Once a time discretization of the Crank-Nicholson type is performed, the resulting system of equations allows for a stable approximation of both fields, by means of classical Lagrange continuous piecewise polynomial functions of arbitrary degree, in any space dimension. Convergence results in the mean square sense in space for the primal unknown and its gradient, together with the flux variable and its divergence, and in appropriate senses in time applying to this pair of fields are given. Numerical experiments illustrate the performance of the scheme, while allowing to check the optimality of the convergence results.

1 INTRODUCTION

The numerical solution of advection-diffusion equations is known to be a delicate problem in many respects, even when they are linear. This is particularly the case of simulations at high Péclet numbers, due to the need to capture sharp gradients of the solution close to boundary layers. In the framework of finite element approximations several approaches to handle this problem in a satisfactory manner have been adopted since the seventies, and in this respect we would like to quote the contributions of Heinrich cf. Heinrich et al. (1978) and Baba and Tabata cf. Baba and Tabata (1981). The celebrated procedure introduced in the early eighties by Hughes and Brooks Brooks and Hughes (1982) is still widely in use. It is based on the modification of the standard Galerkin formulation by introducing stabilizing numerical diffusion in the streamline direction. This is generally known as the SUPG technique, which gave rise the several variants since then, such as Galerkin least-squares formulations. Other stable methods are suitable for the explicit solution of transient problems, such as the one long exploited by Kawahara and collaborators see e.g. Kawahara and Hirano (1983) in the eighties and the nineties. These can be viewed as adaptions of the Lax or the Lax-Wendroff schemes for finite difference discretizations in space and time to a finite element environment, to be used more specifically in the framework of a standard piecewise linear Galerkin approximation. Recently the second author and collaborators exploited this idea by modifying the method in such a way that it gives rise to convergent approximations in the maximum norm, even for non uniform meshes cf. Ruas et al. (2009).

Later on several formulations of the Petrov-Galerkin or the least-squares type were sudied for dealing with this mixed problem, mostly in the stationary case see e.g. Pehlivanov et al. (1994). A one field alternative to all these methods was proposed more recently by Carneiro de Araujo and the second author see e.g. Ruas and Carneiro de Araujo (2009). The main feature of their method is a quadratic interpolation of the primal field of the Hermite type incorporating the mean fluxes across element interfaces as degrees of freedom. Numerical experiments with these elements and classical mixed methods of comparable order showed that the former are globally more accurate.

Anyhow except for a few contributions such as Yang (2002), the numerical analysis of mixed methods in the time-dependent case seems to have been overlooked, specially in the case where advection plays a significant role. In this respect the contribution of the first author and collaborators based on a mixed space-time least squares formulation cf. Novo et al. (2006) showed to be very effective from the computational point of view, while allowing for the use of space interpolations of arbitrary order of both fields. In this paper we endeavour to give convergence results that hold for a slightly modified version of this method, in which the time dependence is handled by means of the classical Crank-Nicholson scheme, while keeping essentially the same least-squares formulation, as far as space is concerned. Another interesting point of the present contribution is the fact that the stability and convergence results do not rely on the fact that a reaction term appears in the equations. In fact the case of the advection-diffusion-reaction equations is treated here as a mere by-product of our analysis for advection-diffusion problems.

Differently from most works on the convergence of finite element methods for time-dependent problems, the analysis carried out in this paper endeavours to address in a clear manner the two steps to be demonstrated in order to establish the convergence in the sense of certain norms, of a numerical method for solving differential equations. Indeed similarly to Ruas et al. (2009) and Carneiro de Araujo et al. (2010), in accordance with the celebrated Lax Equivalence Theorem cf. Lax and Richtmyer (1956), we do this by proving separately the method's consistency and stability in the same norms as convergence is supposed to hold.

2 PROBLEM STATEMENT AND NOTATIONS

Let us consider the time-dependent advection-diffusion problem including or not a reaction term, defined in a domain $\Omega \times (0,T)$, where Ω is a bounded subset of $\Re^N$, $N = 1,2$ or 3 with boundary $\partial\Omega$ and T is a finite time, described as follows:
We wish to determine a scalar valued function $u(\mathbf{x},t)$ satisfying:

$$\left\{\begin{array}{l} \partial_t u - \nabla \cdot K\nabla u + \mathbf{w}\cdot\nabla u + \sigma u = f \ \text{ in } \Omega \times (0,T) \\ u = g \ \text{ on } \Gamma_0 \times (0,T) \\ K\nabla u \cdot \vec{\nu} = 0 \ \text{ on } \Gamma_1 \times (0,T) \\ u = u^0 \quad \text{in } \Omega \text{ for t = 0} \end{array}\right. \tag{1}$$

where $\vec{\nu}$ is the unit outer normal vector on $\partial\Omega$, Γ_0 and Γ_1 are two disjoint portions of $\partial\Omega$, $\partial_t\mathcal{G}$ represents the first order time derivative of a scalar or vector valued function $\mathcal{G}$ and ∇ denotes the gradient operator. The measure of Γ_1 may be null but the one of Γ_0 is assumed to be strictly positive.
Throughout this paper the *dot* product, besides being used to denote the standard inner product of $\Re^N$, in a term like $\nabla\cdot\mathbf{a}$ represents the divergence of a vector valued function $\mathbf{a}$, and in a term like $\mathbf{a}\cdot\nabla$ the operator $\displaystyle\sum_{i=1}^{N} a_i \frac{\partial}{\partial x_i}$.
In (1) K is a constant symmetric diffusivity tensor assumed to be positive-definite , and $\mathbf{w}$ is a given advective velocity assumed to be solenoidal and to belong to $[\mathcal{C}^m(\bar{\Omega})]^N$ for a certain $m \in I\!N^*$. We further assume that $\mathbf{w}$ satisfies the condition $\mathbf{w}\cdot\vec{\nu} \geq 0$ on Γ_1 (in this sense Γ_1 is viewed as a part of $\partial\Omega$ containing only portions of either the outlet or slip walls surrounding the region Ω, in which an incompressible fluid flows with velocity $\mathbf{w}$). σ in turn is a non-negative constant coefficient standing for an eventual reactive phenomenon associated with the advection-diffusion process under study.
The data f and g are respectively, a given forcing function belonging to $L^\infty[(0,T);L^2(\Omega)]$ cf. Fujita et al. (2001) and a prescribed value on $\Gamma_0 \times (0,T)$. For the sake of simplicity and without essential losses in our analytical results, we take in this work $g \equiv 0$. We further assume that $u^0 \in H^1(\Omega)$ and $\nabla\cdot K\nabla u^0 \in L^2(\Omega)$. We shall require additional regularity on both u^0 and f to be specified later on.

Let us define two Hilbert spaces for natural norms cf. Ciarlet (1978) and Girault and Raviart (1986) which will play a key role in all the sequel, namely, $V := \{v \ / \ v \in H^1(\Omega), v_{/\Gamma_0} = 0\}$ and $\mathbf{Q} := \{\mathbf{q} \ /\mathbf{q} \in \mathbf{H}(div,\Omega), \ \mathbf{q}_{/\Gamma_1}\cdot\vec{\nu} = 0\}$ cf. Girault and Raviart (1986). We further introduce the flux variable $\mathbf{p} := -K\nabla u$, which belongs to $\mathbf{Q}$ by assumption. Then given a strictly positive constant α we set (1) in the following equivalent mixed variational form of the least-squares type, where (A,B) denotes the standard inner product of two scalar or vector valued functions A and B in $L^2(\Omega)$, and $(A,B)_K$ represents (KA,B), A and B being two vector valued functions, and I denotes the identity operator.

$$\left\{\begin{array}{l} \text{Find } u(\cdot,t) \in V \text{ with } \partial_t u(\cdot,t) \in L^2(\Omega) \text{ and } u = u^0 \text{ in } \Omega \text{ for } t = 0, \\ \text{together with } \mathbf{p} \in \mathbf{Q} \text{ with } \mathbf{p} := -K\nabla u^0 \text{ in } \Omega \text{ for } t = 0, \\ \text{such that } \forall v \in V \text{ and } \forall \mathbf{q} \in \mathbf{Q} \text{ we have for every } t \in (0,T): \\ \\ (\partial_t u + \nabla\cdot\mathbf{p} + [\mathbf{w}\cdot\nabla + \sigma I]u, v + \alpha\{\nabla\cdot\mathbf{q} + [\mathbf{w}\cdot\nabla + \sigma I]v\}) \\ +(\nabla u + K^{-1}\mathbf{p}, \nabla v + K^{-1}\mathbf{q})_K = \ (f, v + \alpha\{\nabla\cdot\mathbf{q} + [\mathbf{w}\cdot\nabla + \sigma I]v\}). \end{array}\right. \tag{2}$$

Notice that $(\cdot,\cdot)_K$ is an inner product on $L^2(\Omega)^N$ and the associated norm denoted by $\| \cdot \|_K$

is equivalent to the standard norm of this space, which incidentally will be denoted henceforth by $\| \cdot \|$ even in the scalar case. More specifically it holds,

$$\lambda \parallel A \parallel^2 \leq \parallel A \parallel_K^2 \leq \mu \parallel A \parallel^2, \ \ \forall A \in L^2(\Omega)^N \tag{3}$$

where λ and μ are respectively the smallest eigenvalue and the largest eigenvalue of K.

Remark 1: If equation (1) is written in dimensionless form λ^{-1} represents the so-called Péclet number. ■

Remark 2: Strictly speaking in formulation (2) the Neumann boundary condition on Γ_1 for u need not to be enforced as the Dirichlet boundary condition $\mathbf{p} \cdot \vec{\nu} = 0$. Indeed this condition is implictly satisfied by the field $K\nabla u$, and hence it is useless to prescribe it to $\mathbf{p}$ too. However we do this here not only because such a condition is perfectly compatible with fields in $\mathbf{H}(div, \Omega)$ cf. Girault and Raviart (1986), but also because we need it for the analyses carried out hereafter. ■

Throughout this paper we denote by $\| \cdot \|_{m,p,\Omega}$ the standard norm of Sobolev space $W^{m,p}(\Omega)$ cf. Adams (1975) with $m \in I\!N$ and $p \in \Re$, $p \geq 1$. $W^{m,2}(\Omega)$ is commonly denoted by $H^m(\Omega)$ for $m \neq 0$. The standard norm of $H^m(\Omega)$ is simply denoted by $\| \cdot \|_{m,\Omega}$ except for $m = 0$, all those notations applying to scalar or vector versions of the corresponding spaces.

3 SPACE-TIME DISCRETIZATION

An adaption of the well-known Crank-Nicholson scheme for the time discretization of parabolic equations in terms of a single field is used to discretize in time equation (1). More specifically, given an integer M, $M > 1$, we define a time step $\Delta t = T/M$. This leads to a partition of $[0,T]$ into M intervals I_n of equal length Δt, namely $I_n := ([n-1]\Delta t, n\Delta t)$, for $n = 1, 2, \ldots, M$. Then setting $f^r = f(r\Delta t)$ for any real number $r \in [0, M]$, starting from u^0 and $\mathbf{p}^0$, for $n = 1, 2, \ldots, M$ we determine an approximation of $(u[n\Delta t]; \mathbf{p}[n\Delta t])$ denoted by $(u^n; \mathbf{p}^n)$, as the solution of:

$$\left\{ \begin{array}{l} \dfrac{u^n - u^{n-1}}{\Delta t} + \nabla \cdot \dfrac{\mathbf{p}^n + \mathbf{p}^{n-1}}{2} + (\mathbf{w} \cdot \nabla + \sigma I) \dfrac{u^n + u^{n-1}}{2} = f^{n-1/2} \ \text{ in } \Omega \\ \dfrac{1}{2}[K\nabla u^n + \mathbf{p}^n + K\nabla u^{n-1} + \mathbf{p}^{n-1}] = \vec{0} \text{ in } \Omega \\ u^n = 0 \text{ on } \Gamma_0 \\ K\nabla u^n \cdot \vec{\nu} = 0 \text{ on } \Gamma_1. \end{array} \right. \tag{4}$$

In this work we will deal with the counterpart of equation (1) discretized in time, namely, a least squares formulation of system (4) analogous to (2) written as follows:

$$\left\{\begin{array}{l}\text{Starting from } u^0 \in V \text{ and } \mathbf{p}^0 = -K\nabla u^0 \in \mathbf{Q}, \text{for } n = 1, 2, \ldots, M, \\ \text{ find } u^n \in V \text{ and } \mathbf{p}^n \in \mathbf{Q} \text{ such that } \forall v \in V \text{ and } \forall \mathbf{q} \in \mathbf{Q} : \\ \\ (u^n + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{p}^n + [\mathbf{w} \cdot \nabla + \sigma I]u^n\}, v + \alpha\{\nabla \cdot \mathbf{q} + [\mathbf{w} \cdot \nabla + \sigma I]v\}) \\ +\Delta t(\nabla u^n + K^{-1}\mathbf{p}^n, \nabla v + K^{-1}\mathbf{q})_K/2 \\ = \ \Delta t(f^{n-1/2}, v + \alpha\{\nabla \cdot \mathbf{q} + [\mathbf{w} \cdot \nabla + \sigma I]v\}) \\ +(u^{n-1} - \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{p}^{n-1} + [\mathbf{w} \cdot \nabla + \sigma I]u^{n-1}\}, v + \alpha\{\nabla \cdot \mathbf{q} + [\mathbf{w} \cdot \nabla + \sigma I]v\}) \\ -\Delta t(\nabla u^{n-1} + K^{-1}\mathbf{p}^{n-1}, \nabla v + K^{-1}\mathbf{q})_K/2 \end{array}\right. \tag{5}$$

A straightforward application of the Lax-Milgram Theorem, in all similar to the one considered in Section 5 for the fully discrete version of the stationary analogue of (5), establishes that this system has a unique solution. By inspection it is also easy to see that the pair $(u^n, \mathbf{p}^n)$ satisfying (4) is this solution.

Remark 3: Both (4) and (5) are equivalent to the following time discretization of (1):

$$\left\{\begin{array}{l} \dfrac{u^n - u^{n-1}}{\Delta t} - (\nabla \cdot K\nabla - \mathbf{w} \cdot \nabla - \sigma I)\dfrac{u^n + u^{n-1}}{2} = f^{n-1/2} \quad \text{in } \Omega \\ u^n = 0 \text{ on } \Gamma_0 \\ K\nabla u^n \cdot \vec{\nu} = 0 \text{ on } \Gamma_1. \blacksquare \end{array}\right.$$

Now for the sake of simplicity we assume that Ω is an interval if $N = 1$, a polygon if $N = 2$ and a polyhedron if $N = 3$.

In so doing we consider an analogue of (5) discretized in space defined as follows:
Let $\mathcal{T}_h$ be a partition of Ω into intervals for $N = 1$, into triangles or convex quadrilaterals for $N = 2$, and into tetrahedra or convex hexahedra with quadrilateral faces for $N = 3$, with maximum edge length equal to h. We assume that $\mathcal{T}_h$ satisfies the usual compatibility conditions for finite element meshes, and that it belongs to a quasi-uniform family of partitions. We also assume that both Γ_0 and Γ_1 are such that they can be completely covered by the union of edges for $N = 2$ or faces for $N = 3$, of elements belonging to $\mathcal{T}_h$. If $\mathcal{T}_h$ consists of N-simplices, for every $E \in \mathcal{T}_h$ $R_k(E)$ denotes the space of polynomials of degree less than or equal to k, and otherwise $R_k(E)$ denotes the space of functions defined as transforms of polynomials defined in a unit square or cube $\hat{E}$ of degree less than or equal to k in each of the N space variables, through the N-linear mapping from $\hat{E}$ onto E.
In so doing for any $k \in I\!\!N^*$ we introduce the following spaces associated with $\mathcal{T}_h$:

$$S_{h,k} := \left\{ v \,|\ \ v \in C^0(\bar{\Omega}) \text{ and } v_{/E} \in R_k(E),\ \forall E \in \mathcal{T}_h \right\},$$

$$\mathbf{Q}_h := \{\mathbf{q} \in \mathbf{Q} \,|\, \forall i \ q_i \in S_{h,l}\} \ \text{ for } l \in I\!\!N^*,$$

$$V_h := S_{h,j} \cap V \text{ for } j \in I\!\!N^*.$$

Then letting u_h^0 and $\mathbf{p}_h^0$ be the standard V_h-interpolate of u^0 and the $\mathbf{Q}_h$- interpolate of $\mathbf{p}^0$, respectively, we set the following problem to approximate (5) (or yet (4)), for every n, $n =$

$1, 2, \ldots, M$:

$$\left\{\begin{array}{l} \text{Starting from } u_h^0 \in V_h \text{ and } \mathbf{p}_h^0 \in \mathbf{Q}_h, \text{for } n = 1, 2, \ldots, M, \\ \text{ find } u_h^n \in V_h \text{ and } \mathbf{p}_h^n \in \mathbf{Q}_h \text{ such that } \forall v \in V_h \text{ and } \forall \mathbf{q} \in \mathbf{Q}_h : \\ \\ (u_h^n + \dfrac{\Delta t}{2}\{\nabla\cdot\mathbf{p}_h^n + [\mathbf{w}\cdot\nabla + \sigma I]u_h^n\}, v + \alpha\{\nabla\cdot\mathbf{q} + [\mathbf{w}\cdot\nabla + \sigma I]v\}) \\ +\Delta t(\nabla u_h^n + K^{-1}\mathbf{p}_h^n, \nabla v + K^{-1}\mathbf{q})_K/2 \\ = \;\Delta t(f^{n-1/2}, v + \alpha\{\nabla\cdot\mathbf{q} + [\mathbf{w}\cdot\nabla + \sigma I]v\}) \\ +(u_h^{n-1} - \dfrac{\Delta t}{2}\{\nabla\cdot\mathbf{p}_h^{n-1} + [\mathbf{w}\cdot\nabla + \sigma I]u_h^{n-1}\}, v + \alpha\{\nabla\cdot\mathbf{q} + [\mathbf{w}\cdot\nabla + \sigma I]v\}) \\ -\Delta t(\nabla u_h^{n-1} + K^{-1}\mathbf{p}_h^{n-1}, \nabla v + K^{-1}\mathbf{q})_K/2 \end{array}\right. \tag{6}$$

As a simple application of the Lax-Milgram Theorem the following result holds:

Proposition 1: Problem (6) has a unique solution for every Δt. ■

In general scheme (6) is not suitable for practical purposes since the exact integration of terms involving $\mathbf{w}$ and f is out of reach. That is why in principle we must resort to numerical integration of such terms, which will be interpreted here as the replacement of $\mathbf{w}$ or f^r with their standard interpolates $\mathbf{w}_h$ and f_h^r in $[S_{h,l+1}]^N$ and $S_{h,j}$ respectively, to be specified later on. In this manner we are led to a new approximate problem instead of (6). Denoting its solution in the same way as the one of problem (6) for simplicity, this problem is stated as follows:

$$\left\{\begin{array}{l} \text{Starting from } u_h^0 \in V_h \text{ and } \mathbf{p}_h^0 \in \mathbf{Q}_h, \text{for } n = 1, 2, \ldots, M, \\ \text{ find } u_h^n \in V_h \text{ and } \mathbf{p}_h^n \in \mathbf{Q}_h \text{ such that } \forall v \in V_h \text{ and } \forall \mathbf{q} \in \mathbf{Q}_h : \\ \\ (u_h^n + \dfrac{\Delta t}{2}\{\nabla\cdot\mathbf{p}_h^n + [\mathbf{w}_h\cdot\nabla + \sigma I]u_h^n\}, v + \alpha\{\nabla\cdot\mathbf{q} + [\mathbf{w}_h\cdot\nabla + \sigma I]v\}) \\ +\Delta t(\nabla u_h^n + K^{-1}\mathbf{p}_h^n, \nabla v + K^{-1}\mathbf{q})_K/2 \\ = \;\Delta t(f_h^{n-1/2}, v + \alpha\{\nabla\cdot\mathbf{q} + [\mathbf{w}_h\cdot\nabla + \sigma I]v\}) \\ +(u_h^{n-1} - \dfrac{\Delta t}{2}\{\nabla\cdot\mathbf{p}_h^{n-1} + [\mathbf{w}_h\cdot\nabla + \sigma I]u_h^{n-1}\}, v + \alpha\{\nabla\cdot\mathbf{q} + [\mathbf{w}_h\cdot\nabla + \sigma I]v\}) \\ -\Delta t(\nabla u_h^{n-1} + K^{-1}\mathbf{p}_h^{n-1}, \nabla v + K^{-1}\mathbf{q})_K/2 \end{array}\right. \tag{7}$$

Problem (7) is well-posed too, owing to the definitions of $\mathbf{w}_h$ and f_h^r.

4 STABILITY

Remark 4: Throughout the remainder of this work the letter C combined or not with other symbols will represent different strictly positive constants independent of Δt and h. ■

In this Section we proceed to the stability analysis of scheme (7). For this purpose it is convenient to assume that we are solving a more general problem, namely:

$$\left\{\begin{array}{l} \text{Starting from } u_h^0 \in V_h \text{ and } \mathbf{p}_h^0 \in \mathbf{Q}_h, \text{for } n=1,2,\ldots,M, \\ \text{ find } u_h^n \in V_h \text{ and } \mathbf{p}_h^n \in \mathbf{Q}_h \text{ such that } \forall v \in V_h \text{ and } \forall \mathbf{q} \in \mathbf{Q}_h: \\ \\ (u_h^n + \dfrac{\Delta t}{2}\{\nabla\cdot\mathbf{p}_h^n + [\mathbf{w}_h\cdot\nabla + \sigma I]u_h^n\}, v + \alpha\{\nabla\cdot\mathbf{q} + [\mathbf{w}_h\cdot\nabla + \sigma I]v\}) \\ +\Delta t(\nabla u_h^n + K^{-1}\mathbf{p}_h^n, \nabla v + K^{-1}\mathbf{q})_K/2 \ = \\ \Delta t\{L_h^{n-1/2}(v + \alpha\{\nabla\cdot\mathbf{q} + \mathbf{w}_h\cdot\nabla + \sigma I]v\}) + \sqrt{\Delta t}G_h^{n-1/2}(\nabla v + K^{-1}\mathbf{q})\} \\ +(u_h^{n-1} - \dfrac{\Delta t}{2}\{\nabla\cdot\mathbf{p}_h^{n-1} + [\mathbf{w}_h\cdot\nabla + \sigma I]u_h^{n-1}\}, v + \alpha\{\nabla\cdot\mathbf{q} + [\mathbf{w}_h\cdot\nabla + \sigma I]v\}) \\ -\Delta t(\nabla u_h^{n-1} + K^{-1}\mathbf{p}_h^{n-1}, \nabla v + K^{-1}\mathbf{q})_K/2 \end{array}\right. \tag{8}$$

We assume that $L_h^{n-1/2}$ and $G_h^{n-1/2}$ are linear functionals satisfying:

$$\left\{\begin{array}{l} L_h^{n-1/2}(d) \leq |L_h^{n-1/2}| \parallel d \parallel, \ \forall d \in \mathcal{D}_h, \\ G_h^{n-1/2}(\mathbf{d}) \leq |G_h^{n-1/2}| \parallel \mathbf{d} \parallel_K, \ \forall \mathbf{d} \in \mathbf{D}_h, \end{array}\right. \tag{9}$$

where $|\cdot|$ denotes the standard functional norm and $\mathcal{D}_h$ is the function space that equals the direct sum of V_h, the space spanned by the first order derivatives of functions in V_h and the space spanned by the first order derivatives of components of vector fields in $\mathbf{Q}_h$. together with its first order derivatives, and $\mathbf{D}_h = \{\mathbf{d}|d_i \in \mathcal{D}_h, i=1,\ldots,N\}$.
Notice that in the problem we are solving we have $L_h^{n-1/2}(v) = (f_h^{n-1/2}, v)\ \forall v \in L^2(\Omega)$ and $G_h^{n-1/2} \equiv 0$.

In our stability analysis the following quantity is needed:

$$D := \parallel \nabla\cdot\mathbf{w}_h \parallel_{0,\infty,\Omega} \tag{10}$$

Notice that $\nabla\cdot\mathbf{w} = 0$. Hence assuming that the regularity $\mathbf{w} \in [H^{l+2}(\Omega)]^N$ holds, if m is the largest integer such that $1 \leq m < l+2-N/2$ we have $\mathbf{w} \in [\mathcal{C}^m(\bar{\Omega})]^N$ cf. Adams (1975). In this way, according to standard approximation results, there exists a constant $C_{\mathbf{w}}$ such that,

$$D \leq \sqrt{N} \parallel \mathbf{w} - \mathbf{w}_h \parallel_{1,\infty,\Omega} \leq C_{\mathbf{w}} h^{m-1} \parallel \mathbf{w} \parallel_{m,\infty,\Omega}. \tag{11}$$

For problem (refStab1) the following stability result holds:

<u>Theorem 1:</u> Taking $\alpha = \Delta t/2$, setting $\gamma := \ \max\{\max[2\sigma, 4] + \dfrac{D}{4}, \dfrac{19W^2}{2\delta\lambda}, 4 + \dfrac{2W^2}{\lambda}, \dfrac{3W^2}{\lambda}\}$ with $\delta = (3-\sqrt{5})/4$, where $W := C_l \sup_\Omega |\mathbf{w}|$, and assuming that $\gamma\Delta t \leq \dfrac{1}{2}$, the following stability result holds for scheme (8):

$$\left\{\begin{array}{c} \forall n \leq M: \qquad\qquad \parallel u_h^n \parallel^2 + \dfrac{\lambda\Delta t}{2} \parallel \nabla u_h^n \parallel^2 \\ +\Delta t \displaystyle\sum_{i=1}^{n} \left[\dfrac{\parallel u_h^i - u_h^{i-1} \parallel^2}{8\Delta t} + \dfrac{\delta}{\mu} \parallel \mathbf{p}_h^i + \mathbf{p}_h^{i-1} \parallel^2 + \dfrac{\Delta t}{24} \parallel \nabla\cdot(\mathbf{p}_h^i + \mathbf{p}_h^{i-1}) \parallel^2\right] \\ \leq (3e^2)^{\gamma T}\left[\parallel u_h^0 \parallel^2 + \dfrac{\mu\Delta t}{2} \parallel \nabla u_h^0 \parallel^2 + \Delta t \displaystyle\sum_{i=1}^{n}\{|L_h^{i-1/2}|^2 + |G_h^{i-1/2}|^2\}\right]. \blacksquare \end{array}\right. \tag{12}$$

As an immediate consequence of Theorem 1 we have

Corollary 1: Provided $\gamma\Delta t \leq 1/2$, stability holds in the following sense for scheme (7):

$$\left\{ \begin{array}{c} \forall n \leq M: \parallel u_h^n \parallel^2 + \dfrac{\lambda \Delta t}{2} \parallel \nabla u_h^n \parallel^2 \\ +\Delta t \displaystyle\sum_{i=1}^{n} \left[\dfrac{\parallel u_h^i - u_h^{i-1} \parallel^2}{8\Delta t} + \dfrac{\delta}{\mu} \parallel \mathbf{p}_h^i + \mathbf{p}_h^{i-1} \parallel^2 + \dfrac{\Delta t}{24} \parallel \nabla \cdot (\mathbf{p}_h^i + \mathbf{p}_h^{i-1}) \parallel^2 \right] \\ \leq (3e^2)^{\gamma T} \left[\parallel u_h^0 \parallel^2 + \dfrac{\mu \Delta t}{2} \parallel \nabla u_h^0 \parallel^2 + \Delta t \displaystyle\sum_{i=1}^{n} ||f_h^{i-1/2}||^2 \right] .\blacksquare \end{array} \right. \tag{13}$$

5 CONSISTENCY

Henceforth we take $\alpha = \Delta t/2$, and consider that Δt is related to h in a fixed manner, namely,

$$\Delta t = C_\Delta h^\tau, \tag{14}$$

where τ is a strictly positive real number. In so doing we introduce three powers of h, namely,

$$\rho_0 = \min\{l + \min[2, \tau], j + \min[1, \tau/2]\},$$
$$\rho_1 = \min\{l + \min[1, \tau/2], j\},$$
$$\rho_2 = \min\{l + \min[1 - \tau/2, 0], j - \tau/2]\}.$$

Moreover from this Section on we assume that Ω is convex.

As an additional preparatory step to address the convergence of our scheme, we establish in this Section that it is consistent in an appropriate sense. For this purpose we define in $[0, T]$ a pair $(\tilde{u}_h(t); \tilde{\mathbf{p}}_h(t)) \in V_h \times \mathbf{Q}_h$ as a sort of projection $(\tilde{u}_h; \tilde{\mathbf{p}}_h)$ of $u(t)$ and $\mathbf{p}(t) = -K\nabla u(t)$. Such a projection is defined by solving a stationary problem in all similar to (7), taking a function f on the right hand side defined upon $u(t)$ for each value of $t \in [0, T]$. Actually assuming that $\forall t \in [0, T]$ $u(t) \in H^k(\Omega)$ with $k = \max\{l + 1, j + 1\}$, since $\mathbf{w}_h$ is uniformly bounded in $[W^{1,\infty}(\Omega)]^N$ we can prove that $\forall t \in [0, T]$:

$$\parallel [\mathbf{p} - \tilde{\mathbf{p}}_h](t) \parallel + \parallel [u - \tilde{u}_h](t) \parallel_{1,\Omega} \leq \hat{C}_1 h^{\rho_1} \parallel u(t) \parallel_{k,\Omega} \tag{15}$$

$$\text{and} \parallel \nabla \cdot [\mathbf{p} - \tilde{\mathbf{p}}_h](t) \parallel \leq \hat{C}_2 h^{\rho_2} \parallel u(t) \parallel_{k,\Omega} . \tag{16}$$

The pair $(\tilde{u}_h(t); \tilde{\mathbf{p}}_h(t))$ is continuously differentiable with respect to t in $[0, T]$, since the datum u is continuously differentiable with respect to time in $[0, T]$. Then clearly enough the pair $(\partial_t \tilde{u}_h(t); \partial_t \tilde{\mathbf{p}}_h(t))$ is well-defined in $V_h \times \mathbf{Q}_h$ for every $t \in [0, T]$. Moreover by well-known arguments cf. Thomee (1997) we can prove that it is precisely the unique solution of our stationary projection problem when the datum $u(t)$ is replaced by $\partial_t u$, since none of the data $\mathbf{w}$, K and σ depend on t. Moreover assuming that both $u(t)$ and $\partial_t u(t)$ belong to $H^k(\Omega)$ $\forall t \in [0, T]$, provided Ω is convex we can prove in a rather standard manner that $\forall t \in [0, T]$:

$$\left\{ \begin{array}{l} \parallel [u - \tilde{u}_h](t) \parallel \leq \hat{C}_0 h^{\rho_0} \parallel u(t) \parallel_{k,\Omega} \\ \parallel [\partial_t u - \partial_t \tilde{u}_h](t) \parallel \leq \hat{C}_0 h^{\rho_0} \parallel \partial_t u(t) \parallel_{k,\Omega} . \end{array} \right. \tag{17}$$

Next we define $(\tilde{u}_h^n;\tilde{\mathbf{p}}_h^n) \in V_h \times \mathbf{Q}_h$ as the pair $(\tilde{u}_h(t);\tilde{\mathbf{p}}_h(t))$ for $t = n\Delta t$ and $n = 0, 1, \ldots, M$.

Then we apply scheme (8) to the pair $(\tilde{u}_h^n;\tilde{\mathbf{p}}_h^n) \in V_h \times \mathbf{Q}_h$, assuming that u_h^{n-1} and $\mathbf{p}_h^{n-1}$ are replaced by $\tilde{u}_h^{n-1} \in V_h$ and $\tilde{\mathbf{p}}_h^{n-1} \in \mathbf{Q}_h$ respectively, for $n = 1, 2, \ldots, M$. In so doing we determine the residuals in (8), when $(u_h^n;\mathbf{p}_h^n)$ is replaced by $(\bar{u}_h^n;\bar{\mathbf{p}}_h^n) := (\tilde{u}_h^n - u_h^n, \tilde{\mathbf{p}}_h^n - \mathbf{p}_h^n)$ and $(u_h^{n-1};\mathbf{p}_h^{n-1})$ is replaced by $(\bar{u}_h^{n-1};\bar{\mathbf{p}}_h^{n-1}) := (\tilde{u}_h^{n-1} - u_h^{n-1};\tilde{\mathbf{p}}_h^{n-1} - \mathbf{p}_h^{n-1})$.

By definition we have for a given $n \geq 1$:

$$\left\{ \begin{array}{l} \forall v \in V_h \text{ and } \forall \mathbf{q} \in \mathbf{Q}_h : \\ \Delta t\{\mathcal{R}_h^{n-1/2}(v + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{q} + [\mathbf{w}_h \cdot \nabla + \sigma I]v\}) + \sqrt{\Delta t}[\mathcal{S}_h^{n-1/2}(\nabla v) + \mathcal{P}_h^{n-1/2}(\mathbf{q})]\} \\ = \ (\bar{u}_h^n - \bar{u}_h^{n-1}, v + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{q} + [\mathbf{w}_h \cdot \nabla + \sigma I]v\}) \\ +\dfrac{\Delta t}{2}\{(\nabla \cdot \bar{\mathbf{p}}_h^n + [\mathbf{w}_h \cdot \nabla + \sigma I]\bar{u}_h^n, v + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{q} + [\mathbf{w}_h \cdot \nabla + \sigma I]v\}) \\ +(\nabla \bar{u}_h^n + K^{-1}\bar{\mathbf{p}}_h^n, \nabla v + K^{-1}\mathbf{q})_K/2 \\ +(\nabla \cdot \bar{\mathbf{p}}_h^{n-1} + [\mathbf{w}_h \cdot \nabla + \sigma I]\bar{u}_h^{n-1}, v + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{q} + [\mathbf{w}_h \cdot \nabla + \sigma I]v\}) \\ +(\nabla \bar{u}_h^{n-1} + K^{-1}\bar{\mathbf{p}}_h^{n-1}, \nabla v + K^{-1}\mathbf{q})_K/2\}; \end{array} \right. \tag{18}$$

where $\mathcal{R}_h^{n-1/2}$ is a functional representing the residual in the first equation of (8) and $\mathcal{S}_h^{n-1/2}$ stands for the residual associated with the second equation of (8).

Setting $\tilde{u}^r := u(r\Delta t)$ and $\tilde{\mathbf{p}}^r := \mathbf{p}(r\Delta t)$ for $r \in [0, M]$, and noticing that $\nabla \tilde{u}^r + K^{-1}\tilde{\mathbf{p}}^r = \vec{0}$ $\forall r \in [0, M]$, from the definition of $(\tilde{u}_h^n;\tilde{\mathbf{p}}_h^n)$, and noticing that $(u_h^n;\mathbf{p}_h^n)$ is the solution of (7), after some manipulations we can rewrite (18) for a given $n \geq 1$ as follows:

$$\left\{ \begin{array}{l} \forall v \in V_h \text{ and } \forall \mathbf{q} \in \mathbf{Q}_h : \\ \Delta t\{\mathcal{R}_h^{n-1/2}(v + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{q} + [\mathbf{w}_h \cdot \nabla + \sigma I]v\}) + \sqrt{\Delta t}\mathcal{S}_h^{n-1/2}(\nabla v + K^{-1}\mathbf{q})\} \\ = \ (\tilde{u}_h^n - \tilde{u}_h^{n-1}, v + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{q} + [\mathbf{w}_h \cdot \nabla + \sigma I]v\}) \\ +\dfrac{\Delta t}{2}\{(\nabla \cdot \tilde{\mathbf{p}}^n + [\mathbf{w}_h \cdot \nabla + \sigma I]\tilde{u}^n, v + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{q} + [\mathbf{w}_h \cdot \nabla + \sigma I]v\}) \\ +(\nabla \cdot \tilde{\mathbf{p}}^{n-1} + [\mathbf{w}_h \cdot \nabla + \sigma I]\tilde{u}^{n-1}, v + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{q} + [\mathbf{w}_h \cdot \nabla + \sigma I]v\})\} \\ -\Delta t(f_h^{n-1/2}, v + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{q} + [\mathbf{w}_h \cdot \nabla + \sigma I]v\}). \end{array} \right. \tag{19}$$

In view of (19) $\mathcal{S}_h^{n-1/2}$ is readily seen to be the null functional. On the other hand we observe that $\mathcal{R}_h^{n-1/2}$ is expressed in terms of a single function denoted by $F_h^{n-1/2}$. In short we have:

$$\left\{ \begin{array}{l} \forall v \in V_h \text{ and } \forall \mathbf{q} \in \mathbf{Q}_h : \\ \mathcal{R}_h^{n-1/2}(v + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{q} + [\mathbf{w}_h \cdot \nabla + \sigma I]v\}) \\ = \ (F_h^{n-1/2}, v + \dfrac{\Delta t}{2}\{\nabla \cdot \mathbf{q} + [\mathbf{w}_h \cdot \nabla + \sigma I]v\}). \end{array} \right. \tag{20}$$

In this way after estimating (in the L^2-norm) the function $F_h^{n-1/2}$, using (17), (15) and (16) we are led to:

Proposition 5: Let $f \in L^\infty[(0,T); H^{j+1}(\Omega)]$, $\mathbf{w} \in \{H^{l+2}(\Omega)\}^N$, $\partial_{ttt} u \in L^\infty[(0,T); H^1(\Omega)]$ and u, $\partial_t u \in L^\infty[(0,T); H^k(\Omega)]$ for $k = \max\{l+1, j+1\}$. Then we have:

$$|\mathcal{R}_h^{n-1/2}| \leq C_R(\mathbf{w})\; g_R(u)(\Delta t^2 + h^{\rho_0}) \tag{21}$$

where $g_R(u) = \max\{ \underset{s\in(0,T)}{\sup}\ \text{ess}\ \| \nabla \partial_{ttt} u(s) \|, \underset{s\in(0,T)}{\sup}\ \text{ess}\ [\| f(s) \|_{j+1,\Omega} + \| \partial_t u(s) \|_{k,\Omega} + \| u(s) \|_{k,\Omega}]\}$.■

6 CONVERGENCE

In this Section we establish the convergence of the method in natural norms outlined in (12). In this aim we first suppose that in problem (8) $L_h^{n-1/2} = \mathcal{R}_h^{n-1/2}$ and $G_h^{n-1/2} = \mathcal{S}_h^{n-1/2}$. Then the following result holds:

Proposition 6: Let $\Delta t = T/M$ with $M > 2\gamma T$ where γ is defined in the statement of Theorem 1. Then we have $\forall n \leq M$:

$$\left\{ \begin{array}{l} \| \bar{u}_h^n \|^2 + \dfrac{\lambda \Delta t}{2} \| \nabla \bar{u}_h^n \|^2 \\ +\Delta t \displaystyle\sum_{i=1}^{n} \left[\dfrac{\| \bar{u}_h^i - \bar{u}_h^{i-1} \|^2}{8\Delta t} + \dfrac{\delta}{\mu} \| \bar{\mathbf{p}}_h^i + \bar{\mathbf{p}}_h^{i-1} \|^2 + \dfrac{\Delta t}{24} \| \nabla \cdot (\bar{\mathbf{p}}_h^i + \bar{\mathbf{p}}_h^{i-1}) \|^2 \right] \\ \leq (3e^2)^{\gamma T} \left\{ \Delta t \displaystyle\sum_{i=1}^{n} |\mathcal{R}_h^{i-1/2}|^2 + \| \tilde{u}_h^0 - \mathcal{I}_{h,j}(u^0) \|^2 + \dfrac{\mu \Delta t}{2} \| \nabla[\tilde{u}_h^0 - \mathcal{I}_{h,j}(u^0)] \|^2 \right\} .■ \end{array} \right. \tag{22}$$

Corollary 2: Under the same assumptions as in Propositions 5 and 6, if $u^0 \in H^k(\Omega)$, Δt is given by (14) and h is small enough for the inequality $\gamma \Delta t \leq 1/2$ to hold, we have $\forall n \leq M$:

$$\left\{ \begin{array}{l} \left\{ \| \bar{u}_h^n \|^2 + \dfrac{\lambda \Delta t}{2} \| \nabla \bar{u}_h^n \|^2 \right. \\ \left. +\Delta t \displaystyle\sum_{i=1}^{n} \left[\dfrac{\| \bar{u}_h^i - \bar{u}_h^{i-1} \|^2}{8\Delta t} + \dfrac{\delta}{\mu} \| \bar{\mathbf{p}}_h^i + \bar{\mathbf{p}}_h^{i-1} \|^2 + \dfrac{\Delta t}{24} \| \nabla \cdot (\bar{\mathbf{p}}_h^i + \bar{\mathbf{p}}_h^{i-1}) \|^2 \right] \right\}^{1/2} \\ \leq \bar{C}(\mathbf{w}) e^{\gamma T}\, h^{\rho_3}\, [g_R(u) + \| u^0 \|_{k,\Omega}]. \end{array} \right. \tag{23}$$

where $g_R(u)$ is defined in the statement of Proposition 5 and ρ_3 is given by:

$$\rho_3 = \min\{2\tau, l + \min[2,\tau], j + \min[1, \tau/2]\}.■ \tag{24}$$

We are now ready to give our convergence results. Recalling that $(\tilde{u}^n; \tilde{\mathbf{p}}^n) := (u[n\Delta t]; \mathbf{p}[n\Delta t])$ first we have:

Theorem 2: Let the assumptions of Corollary 2 hold. If Δt fulfills (14) and h is small enough for the inequality $\gamma \Delta t \leq 1/2$ to hold, $\exists\, \mathcal{C}_1$ and $\mathcal{C}_2$ depending only on τ, l, j, Ω, T, K, $\mathbf{w}$ and σ

such that $\forall n \leq M$ the following estimates apply:

$$\left\{ \begin{array}{l} \| \ \tilde{u}^n - u_h^n \ \| \leq \mathcal{C}_1 \ h^{\eta_0} \ g_R(u); \\ \| \ \nabla(\tilde{u}^n - u_h^n) \ \| \leq \mathcal{C}_2 \ h^{\eta_1} \ [g_R(u) + \| \ u^0 \ \|_{k,\Omega}]; \\ \eta_0 = \rho_3 \text{ and } \eta_1 = \eta_0 - \tau/2. \end{array} \right. \tag{25}$$

As a consequence of Theorem 2, convergence in the sense of $L^2(\Omega)$ of u_h^n to $u(t)$, and of ∇u_h^n to $\nabla u(t)$ as n goes to ∞ and h goes to zero is established, provided $t = n\Delta t$ remains fixed. Next we give another result stating the convergence of $\mathbf{p}_h^n$ to $\mathbf{p}$ and of $\nabla \cdot \mathbf{p}_h^n$ to $\nabla \cdot \mathbf{p}$ in a weaker sense. More precisely we mean the sense of the discrete norm of $L^2[(0,T); L^2(\Omega)]$ denoted by $\| \ \cdot \ \|_M$ given by,

$\| \ \mathcal{G} \ \|_M := [\sum_{n=1}^{M} \Delta t \ \| \ \mathcal{G}^{n-1/2} \ \|^2]^{1/2}$, with $\mathcal{G}^r = \mathcal{G}(r\Delta t)$, for $r \in [0, M]$, $\mathcal{G}$ being a scalar (resp. vector) valued function belonging to $C^0\{[0,T]; L^2(\Omega)\}$ (resp. $\{C^0\{[0,T]; L^2(\Omega)\}\}^N$).

Here it is particularly handy to define two functions $\mathbf{p}_h(\mathbf{x}, t)$ and $\tilde{\mathbf{p}}_h(\mathbf{x}, t)$, both constant in each interval I_n, whose values for every $t \in I_n$ are $(\mathbf{p}_h^n + \mathbf{p}_h^{n-1})/2$, $(\tilde{\mathbf{p}}_h^n + \tilde{\mathbf{p}}_h^{n-1})/2$ respectively, for $n = 1, 2, \ldots, M$.

Theorem 3: Under the same regularity assumptions on the solution of (1) made in Theorem 2 and for $M > 2T\gamma$, Δt being given by (14) $\exists$ $\mathcal{C}_3$ and $\mathcal{C}_4$ depending only on τ, l, j, Ω, T, K, $\mathbf{w}$ and σ such that:

$$\left\{ \begin{array}{l} \| \ \mathbf{p} - \mathbf{p}_h \ \|_M \leq \mathcal{C}_3 h^{\eta_1} \ [g_R(u) + \| \ u^0 \ \|_{k,\Omega}]; \\ \| \ \nabla \cdot (\mathbf{p} - \mathbf{p}_h) \ \|_M \leq \mathcal{C}_4 h^{\eta_2} \ [g_R(u) + \| \ u^0 \ \|_{k,\Omega}] \\ \text{with } \eta_2 = \eta_1 - \tau/2. \end{array} \right. \tag{26}$$

Finally we give an a priori error estimate in the standard norm of $L^2[(0,T); L^2(\Omega)]$ cf. Fujita et al. (2001) denoted here more simply by $\| \ \cdot \ \|_{0,T,\Omega}$, applying to the time-derivative of u approximated by a function u_h obtained with our method. More specifically, $u_h(\mathbf{x}, t)$ is defined to be the function that varies linearly with t in each interval I_n for every n, and whose value for $t = n\Delta t$ is u_h^n, $n = 1, 2, \ldots, M$, $\forall \mathbf{x} \in \Omega$.

Theorem 4: Under the same regularity assumptions on the solution of (1) made in Theorem 2 and for $M > 2T\gamma$, Δt being given by (14) $\exists$ $\mathcal{C}_5$ depending only on τ, l, j, Ω, T, K, $\mathbf{w}$ and σ such that,

$$\| \ \partial_t[u - u_h] \ \|_{0,T,\Omega} \leq \mathcal{C}_5 h^{\eta_3} \ f_R(u). \tag{27}$$

where $f_R(u) = g_R(u) + \| \ u^0 \ \|_{k,\Omega} + \| \ \partial_t u \ \|_{L^2[(0,T);H^k(\Omega)]} + \| \ \partial_{tt} u \ \|_{L^2[(0,T);H^1(\Omega)]}$ and $\eta_3 = \min\{\tau, l + \min[2 - \tau/2, \tau/2], j + \min[1 - \tau/2, 0]\}$. ∎

Remark 5: As far as numerical integration is concerned it is possible to refine our error analysis. This could be achieved by resorting to numerical integration formulae applying to the products $(\mathbf{w} \cdot \nabla d, e)$ and (f, d) for $d, e \in \mathcal{D}_h$ instead of $\mathbf{w}$ and f alone, and to the well-established theory on variational crimes cf. Strang and Fix (1973). However by no means this would change the essence of the convergence results presented in this work.∎

7 NUMERICAL EXPERIMENTS

In order to check the optimality of the error estimates obtained in this work, the authors carried out some experiments with their method, by solving a test problem with known analytical solution with piecewise linear finite element representations of both unknown fields.
More specifically equation (1) is approximated in the domain $\Omega \times (0, T)$, where Ω is the unit square $(0, 1)^2$ and $T = 1$. We present below some relevant results for the case where Γ_0 is the portion of $\partial\Omega$ given by $xy = 0$, taking $K = I$ and $\mathbf{w}(x, y) = (y; x)/2$.
We consider an exact analytical solution given by:

$$\begin{cases} u(x,y,t) &= \qquad (x - x^2/2)(y - y^2/2)e^{-t}; \\ \mathbf{p}(x,y,t) &= ([x-1][y-y^2/2]; [y-1][x-x^2/2])e^{-t}. \end{cases}$$

In so doing the forcing function f is simply set to be equal to $u_t + \nabla \cdot \mathbf{p} - \mathbf{w} \cdot \mathbf{p}$.

We solved this problem with uniform triangular meshes obtained by first subdividing Ω into L^2 equal squares with edge length $h = 1/L$, each one of them being in turn subdivided into two triangles by taking their diagonals parallel to the line $x = y$.
We display in Tables 1 and 2 below absolute approximation errors for increasing values of L, by taking $M = L$, i.e. $\Delta t = h$. In Table 1 we give errors in the norm of $L^2(\Omega)$ of the approximate values of u and ∇u for $t = T$. In Table 2 we supply the errors of $\mathbf{p}$, $\nabla \cdot \mathbf{p}$ and u_t in the norm of $L^2[(0, T); L^2(\Omega)]$ (in both discrete and continuous versions according to the field).

L	u	∇u
8	0.5887054E-03	0.1232884E-01
16	0.1461917E-03	0.5910155E-02
32	0.4061363E-04	0.2898990E-02
64	0.4013044E-04	0.1474399E-02

Table 1: Absolute errors of u and ∇u in the norm of $L^2(\Omega)$ for $t = 1$.

L	$\mathbf{p}$	$\nabla \cdot \mathbf{p}$	u_t
8	0.2562423E-02	0.2061569E-01	0.1511747E-02
16	0.6670678E-03	0.1060378E-01	0.3973692E-03
32	0.1790714E-03	0.5377365E-02	0.1118547E-03
64	0.1207285E-03	0.2714279E-02	0.1243903E-03

Table 2: Absolute errors in the $L^2[(0, 1); L^2(\Omega)]$-norm of $\mathbf{p}$, $\nabla \cdot \mathbf{p}$ and u_t.

As one can infer from both tables, convergence rates observed from the numerical results are in perfect agreement with the theoretical predictions, as far as the gradient of u is concerned. This is also roughly the case of $\mathbf{p}$, since the observed convergence rate, though greater than $3/2$ for lower values of L gets closer to 1 as L increases. On the other hand a convergence rate better than $1/2$ (about 1), can be reported for the divergence of $\mathbf{p}$. Less clear observations apply to both u and u_t, since the numerical convergence rates for small values of L are rather close to 2 and $7/4$ respectively, instead of $3/2$ and 1, but deteriorate significantly as L increases. Notice that this effect cannot be explained by increasing roundoff errors as the value of L becomes larger, since it does not reflect on errors of other fields, such as the L^2 errors of ∇u.

8 CONCLUSIONS

To conclude we would like to comment the results obtained in this work:
Convergence in space in the $H^1(\Omega)$ and $\mathbf{H}(div, \Omega)$ norms was demonstrated for both fields involved in the mixed formulation, approximated by finite elements of arbitrary order in space. Provided $\tau \leq 2$ such results are optimal in the $L^\infty[(0,T); H^1(\Omega)]$-norm for the primal variable u, and in the $L^2[(0,T); L^2(\Omega)]$-norm for the flux variable $\mathbf{p}$. On the other hand only sub-optimal results in the $L^\infty[(0,T); L^2(\Omega)]$-norm hold for the primal variable if $\Delta t = \mathcal{O}(h)$. If $\Delta t = \mathcal{O}(h^2)$ optimality is recovered or maintained in all the error estimates above, but in this case we have to deal with a stringent and not so natural condition on Δt for a scheme of the Crank-Nicholson type. For this reason among others, the choice $\tau > 2$ should be discarded. In global terms our results indicate that the best orders of convergence are attained for $j = l+1$, $l = 3$ and $\Delta t = \mathcal{O}(h^2)$, but this choice is certainly not reasonable at all from the computational point of view. More realistically if one sticks to popular piecewise linear approximations of both fields, that is if $j = l = 1$, and takes $\Delta t = \mathcal{O}(h)$, the following orders of convergence have been demonstrated:

$$\begin{cases} \| \ [u - u_h](n\Delta t) \ \| = \mathcal{O}(h^{3/2}), \ \forall n; \\ \| \ \nabla[u - u_h](n\Delta t) \ \| = \mathcal{O}(h), \ \forall n; \\ \{\displaystyle\int_0^T \| \ [\mathbf{p} - \mathbf{p}_h](s) \ \|^2 \, ds\}^{1/2} = \mathcal{O}(h); \\ \{\displaystyle\int_0^T \| \ \nabla \cdot [\mathbf{p} - \mathbf{p}_h](s) \ \|^2 \, ds\}^{1/2} = \mathcal{O}(h^{1/2}); \\ \{\displaystyle\int_0^T \| \ \partial_t[u - u_h](s) \ \|^2 \, ds\}^{1/2} = \mathcal{O}(h); \end{cases}$$

No significant discrepancies between the above theoretical predictions and the numerical results exhibited in the previous section were found, except perharps for the divergence of $\mathbf{p}$. Notice that a persistent lack of optimality in the L^2 error estimates for the primal variable u is to be found in other authors' works on least-squares finite element methods for time-dependent problems see. e.g. Yang (2002). In this work the authors pointed out in an explicit manner the source of this phenomenon, namely, the sub-optimal estimate (17).

A final word of clarification about the assumptions on the data K, $\mathbf{w}$ and σ is in order: Although the analysis becomes technically much more complicated if they depend on space and time, under reasonable hypotheses on the regularity of these data, the same qualitative convergence results as in the case studied here should hold. Actually the authors are currently exploiting their numerical approach in the framework of time-dependent advection-diffusion problems of practical interest, in which all those data vary in both space and time. Corresponding results will be reported shortly in Leal-Toledo and Ruas (2010).

Acknowledgements: The author V. Ruas gratefully acknowledges the financial support provided by the agency CNPq through grant 307996/2008-5.

REFERENCES

Adams R. *Sobolev Spaces*. Academic Press, New York, 1975.

Baba M. and Tabata M. On a conservative upwind finite element scheme for convective-diffusion equations. *RAIRO Analyse Numérique*, 15, 1981.

Brooks A. and Hughes T. The streamline upwind/petrov-galerkin formulation for convection dominated flows with particular emphasis on the navier-stokes equations. *Computer Methods in Applied Mechanics and Engineering*, 32:199–259, 1982.

Carneiro de Araujo J., Gomes P., and Ruas V. Study of a finite element method for the time-dependent generalized stokes system associated with viscoelastic flow. *Journal of Computational and Applied Mathematics*, 234:2562–2577, 2010.

Ciarlet P. *The Finite Element Method for Ellipic Problems*. North-Holland, Amsterdam, 1978.

Fujita H., Saito N., and Suzuki T. *Operator Theory and Numerical Methods*. North Holland, New York, 2001.

Girault V. and Raviart P. *Finite Element Methods for Navier-Stokes Equations*. Springer-Verlag, Berlin, 1986.

Heinrich J., Huyakorn P., Zienkiewicz O., and Mitchell A. An upwind finite element scheme for two-dimensional convective transport equation. *International Journal for Numerical Methods in Engineering*, 12:187–190, 1978.

Kawahara M. and Hirano H. A finite element method for high reynolds number viscous fluid flow using two step explicit scheme. *International Journal for Numerical Methods in Fluids*, 3:137–163, 1983.

Lax P. and Richtmyer R. Survey of the stability of linear finite difference equations. *Communications in Pure and Applied Mathematics*, 9:267–293, 1956.

Leal-Toledo R. and Ruas V. Numerical solution of transient advection-diffusion equations by a mixed least-squares finite element method. *To appear*, 2010.

Novo C., Leal-Toledo R., Toledo E., and Martins L. Discontinuous mixed space-time least-squares formulation for transient advection-diffusion-reaction equations. *Mecánica Computacional, A.Cardona, N.Nigro, V.Sonzogni and M.Storti eds.*, pages 1113–1125, 2006.

Pehlivanov A., Carey G., and Lazarov R. Least-squares mixed finite elements for second-order elliptic problems. *SIAM Journal of Numerical Analysis*, 31, 1994.

Ruas V., Brasil Jr. A., and Trales P. An explicit method for convection-diffusion equations. *Japan Journal of Industrial and Applied Mathematics*, 26, 2009.

Ruas V. and Carneiro de Araujo J. A quadratic triangle of the hermite type for second order elliptic problem. *Zeitschrift fuer Angewandte Mathematik und Mechanik*, 89, 2009.

Strang G. and Fix G. *An Analysis of the Finite Element Method*. Prentice Hall, Englewood Cliffs, 1973.

Thomee V. *Galerkin Finite Element Methods for Parabolic Problems*. Springer Series in Computational Mathematics 25, Springer, Berlin, 1997.

Yang D. Least-squares finite element methods for nonlinear parabolic problems. *Journal of Computational Mathematics*, 20, 2002.