TOTAL ERROR ESTIMATES OF MIXED FINITE ELEMENT METHODS FOR NONLINEAR REACTION-DIFFUSION EQUATIONS

M. Isabel Asensio[†] and Luis Ferragut [*]
Department of Applied Mathematics. University of Salamanca
Plaza de la Merced s/n, 37008 Salamanca, Spain.
[†]e-mail: mas@gugu.usal.es
[*]e-mail: ferragut@gugu.usal.es

Abstract
We present some total error estimates of mixed finite element methods for parabolic problems with nonlinear diffusive and reactive terms. For the spatial discretization we use a mesh of rectangular elements with sides parallel to the axes, and standard Raviart-Thomas finite elements. Three temporal discrete schemes are studied, a completely implicit Euler scheme, a semi-implicit Euler scheme, and a completely explicit Euler scheme. As an application, we present some results of the explicit scheme with lowest order Raviart-Thomas finite elements for a model of natural combustion.

Keywords - Error estimates, mixed finite element methods, reaction-diffusion equations.

1. INTRODUCTION

Nonlinear reaction-diffusion equations occur in numerous applications (Fife, 1979), especially hydrologic and bio-geochemical phenomena (Huyakorn & Pinder, 1983; Murray, 1993;), for example in biological systems (Ruuth, 1995), or hydro-geological systems (Yifeng et al., 1994, Yutiang et al., 1995). We can find examples in very different applications. The authors have developed a mathematical model for natural combustion (Montenegro et al. 1997, Asensio & Ferragut, 2000) whose analysis has given rise to the results presented in this paper.

The equations have the form

$$\frac{\partial u}{\partial t} - \vec{\nabla} \cdot \left(a(u) \vec{\nabla} u \right) = f(u) \quad in\, \Omega \times (0,T) \tag{1}$$

where Ω is an open bounded domain in $\Re^2$ with piecewise C^1 boundary Γ, for example, polygonal. We assume homogeneous Newmann boundary conditions, that is,

$$a(u)\vec{\nabla} u \cdot \vec{n}\big|_{\Gamma} = 0 \tag{2}$$

We also assume that the nonlinear reaction function is continuous, smooth and satisfies a globally Lipschitz condition: there exists a positive constant $L_f>0$ such that

$$|f(x) - f(y)| \leq L_f |x - y| \quad for\, all\, x, y \in \Re \tag{3}$$

We assume as well that the nonlinear diffusive function is continuous, smooth, bounded, positive, and verifies a global Lipschitz condition, that is, there exist three positive constants α, μ, $L_a>0$, such that

$$|a(x) - a(y)| \leq L_a |x - y| \quad for\, all\, x, y \in \Re \tag{4}$$

$$0 < \alpha \leq a(x) \leq \mu < \infty \quad for\, all\, x \in \Re \tag{5}$$

Thus $A(x)=(a(x))^{-1}$ is well defined, bounded and verifies a Lipschitz condition with constant $L_A = L_a / \alpha^2$.

In the natural combustion model mentioned before, the reaction function is an exponential function of the form $e^{u/1+\varepsilon u}$, where ε is the nondimensional inverse of the activation energy, and the diffusive term represents the radiation and conduction phenomena.

Introducing the flux $\vec{p} = a(u)\vec{\nabla} u$ as a new variable, the problem to solve is

$$\frac{\partial u}{\partial t} - \vec{\nabla} \cdot \vec{p} = f(u) \quad in\, \Omega \times (0,T) \tag{6}$$

$$A(u)\vec{p} = \vec{\nabla} u \quad in\, \Omega \times (0,T) \tag{7}$$

$$\vec{p} \cdot \vec{n}\big|_{\Gamma} = 0 \quad on\, \Gamma \times (0,T) \tag{8}$$

with initial conditions $u(0) = u_I$, and the corresponding initial condition for $\vec{p}$ obtained from the equation (7).

In some applications it is important to approximate the flux $\vec{p}$ and u with comparable accuracy. Moreover, we need to accommodate the nonlinear reaction term in order to uncouple this nonlinearity. These reasons have moved us towards a mixed finite-element scheme to discretize the system (6-8) in space.

For the temporal discretization we propose and analyse three different schemes: a completely implicit Euler scheme, a semi-implicit Euler scheme, and a completely explicit Euler scheme.

2. DESCRIPTION OF THE METHOD

Assuming the solution of (6), (7) and (8) is sufficiently regular, the mixed variational problem to solve is to find $u(t) \in L^2(\Omega)$ and $\vec{p}(t) \in H_0(div;\Omega)$ such that for all $t \in (0,T)$

$$\left(\frac{\partial u(t)}{\partial t}, v\right) - \left(\vec{\nabla} \cdot \vec{p}(t), v\right) = \left(f(u(t)), v\right) \quad \forall v \in L^2(\Omega) \tag{9}$$

$$\left(A(u(t))\,\vec{p}(t), \vec{q}\right) + \left(u(t), \vec{\nabla} \cdot \vec{q}\right) = 0 \quad \forall \vec{q} \in H_0(div;\Omega) \tag{10}$$

with the corresponding initial conditions.

The space of functions for scalar functions is $L^2(\Omega)$, so we relax the conditions over the solution *u*, but we increase the conditions over the flux $\vec{p}$ by imposing $\vec{p}(t) \in H_0(div;\Omega)$, where

$$H_0(div;\Omega) = \left\{\vec{q} \in H(div;\Omega) : \vec{q} \cdot \vec{n} = 0 \; en \; \Gamma\right\}, \tag{11}$$

with $\vec{n}$ normal unit exterior vector to the boundary Γ, and

$$H(div;\Omega) = \left\{\vec{q} \in \left(L^2(\Omega)\right)^d : \vec{\nabla} \cdot \vec{q} \in L^2(\Omega)\right\}. \tag{12}$$

Thus, the boundary condition (8) is implicit in $\vec{p}(t) \in H_0(div;\Omega)$.

For simplicity, we have used the scalar product notation to represent the corresponding integrals,

$$(u,v) = \int_\Omega uv\,d\vec{x} \quad \forall u, v \in L^2(\Omega). \tag{13}$$

$$(\vec{p}, \vec{q}) = \int_\Omega \vec{p} \cdot \vec{q}\,d\vec{x} \quad \forall \vec{p}, \vec{q} \in H_0(div;\Omega). \tag{14}$$

On a rectangular domain $\Omega=(0,L_x)\times(0,L_y)$, let $\Delta_x=\{0=x_0<x_1<...<x_N=L_x\}$ be a grid on the *x*-axis, and let $\Delta_y=\{0=y_0<y_1<...<y_N=L_y\}$ be a grid on the *y*-axis. Thus $\Delta_x\times\Delta_y$ is a grid on Ω of rectangles having sides parallel to the axes. If the mesh size of this rectangular grid is *h*, we denote it by Δ_h, and each rectangle of the grid by τ. Let $\{\Delta_h : h\in H\}$ be a regular family of this kind of grids on Ω.

We use standard Raviart-Thomas (Raviart & Thomas, 1977) approximation subspaces for s≥1, $L_h^{(s-1)}$ and $E_{0,h}^{(s)}$, for *u* and $\vec{p}$ respectively. The space $L_h^{(s-1)}$ consists of scalar functions that are piecewise polynomials of degree *s*-1 on Δ_x and on Δ_y,

$$L_h^{(s-1)} = \left\{v \in H^1(\Omega) : \forall \tau \in \Delta_h, v|_\tau \in P_{s-1,s-1}(\tau)\right\} \tag{15}$$

The space $E_{0,h}^{(s)}$ consists of vector-valued functions whose first components are piecewise polynomials of degree *s* on Δ_x, and piecewise polynomials of degree *s*-1 on Δ_y, and whose second components are piecewise polynomials of degree *s*-1 on Δ_x, and piecewise polynomials of degree *s* on Δ_y,

$$E_{0,h}^{(s)} = \left\{\vec{q} \in H_0(div;\Omega) : \forall \tau \in \Delta_h, \vec{q}|_\tau \in P_{s,s-1}(\tau) \times P_{s-1,s}(\tau)\right\} \tag{16}$$

We denote by $\text{☞}_h^{(s)}$ the corresponding interpolation operator.

Given these spaces, we seek an approximate $u_h(t) \in L_h^{(s-1)}$ and an approximate $\vec{p}_h(t) \in E_{0,h}^{(s)}$ such that for all $t\in(0,T)$

$$\left(\frac{\partial u_h(t)}{\partial t}, v_h\right) - \left(\vec{\nabla} \cdot \vec{p}_h(t), v_h\right) = \left(f(u_h(t)), v_h\right) \quad \forall v_h \in L_h^{(s-1)} \tag{17}$$

$$\left(A(u_h(t))\,\vec{p}_h(t), \vec{q}_h\right) + \left(u_h(t), \vec{\nabla} \cdot \vec{q}_h\right) = 0 \quad \forall \vec{q}_h \in E_{0,h}^{(s)} \tag{18}$$

with initial conditions $u_h(0)=u_{I,h}$, where $u_{I,h}$ is an approximation of $u(0)$ on $L_h^{(s-1)}$. Note that $u_h(0)$ determines $\vec{p}_h(0)$ by equation (16).

Using an implicit Euler scheme, the problem to solve, with the standard notation for progressive and regressive differences, is to calculate u_h^n and $\vec{p}_h^n$ for $n\geq 1$ verifying

$$\left(\bar{\partial}_t u_h^n, v_h\right)-\left(\vec{\nabla}\cdot\vec{p}_h^n, v_h\right)=\left(f\left(u_h^n\right), v_h\right) \quad \forall v_h\in L_h^{(s-1)} \tag{19}$$

$$\left(A\left(u_h^n\right)\vec{p}_h^n, \vec{q}_h\right)+\left(u_h^n, \vec{\nabla}\cdot\vec{q}_h\right)=0 \quad \forall \vec{q}_h\in E_{0,h}^{(s)} \tag{20}$$

Note that equations (19) and (20) are coupled. To solve this system of equations we use an algorithm of type Uzawa: let $u_h^{n,0}=u_h^{n-1}$, given $u_h^{n,\gamma}$, we calculate $u_h^{n,\gamma+1}$ for $\gamma=0,\ldots,N_\gamma$ - 1, verifying

$$\left(A\left(u_h^{n,\gamma}\right)\vec{p}_h^{n,\gamma}, \vec{q}_h\right)+\left(u_h^{n,\gamma}, \vec{\nabla}\cdot\vec{q}_h\right)=0 \quad \forall \vec{q}_h\in E_{0,h}^{(s)} \tag{21}$$

$$\left(\frac{u_h^{n,\gamma+1}-u_h^{n-1}}{k}, v_h\right)-\left(\vec{\nabla}\cdot\vec{p}_h^{n,\gamma}, v_h\right)=\left(f\left(u_h^{n,\gamma+1}\right), v_h\right) \quad \forall v_h\in L_h^{(s-1)} \tag{22}$$

and then $u_h^n=u_h^{n,N_\gamma}$. By choosing N_γ=1, we are actually using a semi-implicit scheme: to calculate u_h^n and $\vec{p}_h^n$ for $n\geq 1$ verifying

$$\left(\bar{\partial}_t u_h^n, v_h\right)-\left(\vec{\nabla}\cdot\vec{p}_h^{n-1}, v_h\right)=\left(f\left(u_h^n\right), v_h\right) \quad \forall v_h\in L_h^{(s-1)} \tag{23}$$

$$\left(A\left(u_h^n\right)\vec{p}_h^n, \vec{q}_h\right)+\left(u_h^n, \vec{\nabla}\cdot\vec{q}_h\right)=0 \quad \forall \vec{q}_h\in E_{0,h}^{(s)} \tag{24}$$

Applying a complete explicit Euler scheme, the problem to solve is to calculate u_h^n and $\vec{p}_h^n$ for $n\geq 1$, verifying

$$\left(\bar{\partial}_t u_h^n, v_h\right)-\left(\vec{\nabla}\cdot\vec{p}_h^{n-1}, v_h\right)=\left(f\left(u_h^{n-1}\right), v_h\right) \quad \forall v_h\in L_h^{(s-1)} \tag{25}$$

$$\left(A\left(u_h^n\right)\vec{p}_h^n, \vec{q}_h\right)+\left(u_h^n, \vec{\nabla}\cdot\vec{q}_h\right)=0 \quad \forall \vec{q}_h\in E_{0,h}^{(s)} \tag{26}$$

3. SEMIDISCRETE ERROR ESTIMATES

We break down the semidiscrete error into the form (Thomée, 1984)

$$u_h(t)-u(t)=\left(u_h(t)-\tilde{u}_h(t)\right)+\left(\tilde{u}_h(t)-u(t)\right)=\vartheta(t)+\rho(t) \tag{27}$$

$$\vec{p}_h(t)-\vec{p}(t)=\left(\vec{p}_h(t)-\tilde{\vec{p}}_h(t)\right)+\left(\tilde{\vec{p}}_h(t)-\vec{p}(t)\right)=\vec{z}(t)+\vec{r}(t) \tag{28}$$

$\left(\tilde{u}_h(t), \tilde{\vec{p}}_h(t)\right)$ being the elliptic projection on $L_h^{(s-1)}\times E_{0,h}^{(s)}$ of the exact solution $\left(u(t), \vec{p}(t)\right)$ defined by

$$\left(\vec{\nabla}\cdot\left(\vec{p}(t)-\tilde{\vec{p}}_h(t)\right), v_h\right)=0 \quad \forall v_h\in L_h^{(s-1)} \tag{29}$$

$$\left(A(u)\left(\vec{p}(t)-\tilde{\vec{p}}_h(t)\right), \vec{q}_h\right)+\left(u(t)-\tilde{u}_h(t), \vec{\nabla}\cdot\vec{q}_h\right)=0 \quad \forall \vec{q}_h\in E_{0,h}^{(s)} \tag{30}$$

***Lemma* 1.** If $(u,\vec{p})$ are sufficiently regular, then $\forall t\in(0,T)$ we have

$$\begin{aligned}&\|\rho(t)\|_{0,\Omega}\leq C(u(t))h^s,\\ &\|\vec{r}(t)\|_{H(div;\Omega)}\leq C(u(t))h^s.\end{aligned} \tag{31}$$

❑

Proof.

As $\left(u(t), \vec{p}(t)\right)$ is solution of the continuous problem given by (9) and (10), we have

$$\left(A(u(t))\vec{p}(t), \vec{q}\right)+\left(u(t), \vec{\nabla}\cdot\vec{q}\right)=0 \quad \forall \vec{q}\in H_0(div;\Omega) \tag{32}$$

In particular, this is also true $\forall \vec{q}_h \in E_{0,h}^{(s)} \subset H_0(div;\Omega)$, so $\left(\widetilde{u}_h(t), \widetilde{\vec{p}}_h(t)\right)$ verifies

$$\left(\vec{\nabla} \cdot \widetilde{\vec{p}}_h(t), v_h\right) = \left(\vec{\nabla} \cdot \vec{p}(t), v_h\right) \quad \forall v_h \in L_h^{(s-1)} \tag{33}$$

$$\left(A(u(t))\widetilde{\vec{p}}_h(t), \vec{q}_h\right) + \left(\widetilde{u}_h(t), \vec{\nabla} \cdot \vec{q}_h\right) = 0 \quad \forall \vec{q}_h \in E_{0,h}^{(s)} \tag{34}$$

Then, $\left(\widetilde{u}_h(t), \widetilde{\vec{p}}_h(t)\right)$ is a solution of the following discrete elliptic problem for $f = \vec{\nabla} \cdot \vec{p}(t)$ and $\kappa = A(u(t))$, where by the properties of $A(\cdot\,)$, $1/\mu \le \kappa = A(u(t)) \le 1/\alpha$,

$$\left(\vec{\nabla} \cdot \widetilde{\vec{p}}_h(t), v_h\right) = (f, v_h) \quad \forall v_h \in L_h^{(s-1)} \tag{35}$$

$$\left(\kappa\, \widetilde{\vec{p}}_h(t), \vec{q}_h\right) + \left(\widetilde{u}_h, \vec{\nabla} \cdot \vec{q}_h\right) = 0 \quad \forall \vec{q}_h \in E_{0,h}^{(s)} \tag{36}$$

for which the corresponding continuous elliptic problem is

$$\left(\vec{\nabla} \cdot \widetilde{\vec{p}}(t), v\right) = (f, v) \quad \forall v \in L^2(\Omega)/\Re \tag{37}$$

$$\left(\kappa \widetilde{\vec{p}}(t), \vec{q}\right) + \left(\widetilde{u}(t), \vec{\nabla} \cdot \vec{q}\right) = 0 \quad \forall \vec{q} \in H_0(div;\Omega) \tag{38}$$

Obviously $(u(t), \vec{p}(t))$ is a solution of this problem. Let us see that this problem verifies the continuous "inf-sup" Babuska-Brezzi condition to prove the existence and uniqueness of its solution. As $\vec{p}(t) \in H_0(div;\Omega)$, by the divergence theorem, it follows that $f \in L^2(\Omega)/\Re$. We define the following space of functions

$$V = \left\{\vec{q} \in H_0(div;\Omega) : \int_\Omega v \vec{\nabla} \cdot \vec{q} = 0 \;\forall v \in L^2(\Omega)/\Re\right\} \tag{39}$$

The bilinear forms that define the continuous problem given by (37) and (38) are

$$a(\vec{p}, \vec{q}) = \int_\Omega \kappa\, \vec{p} \cdot \vec{q} \quad \forall \vec{p}, \vec{q} \in H_0(div;\Omega) \tag{40}$$

$$b(\vec{q}, v) = \int_\Omega v\, \vec{\nabla} \cdot \vec{q} \quad \forall \vec{q} \in H_0(div;\Omega), v \in L^2(\Omega)/\Re \tag{41}$$

which are trivially continuous, and $a(\cdot,\cdot)$ is V-elliptic.

We use the following auxiliary problem to prove the continuous "inf-sup" condition via a standard technique (Roberts & Thomas, 1991): for each $v \in L^2(\Omega)/\Re$, let φ_v be the unique element of $H^1(\Omega)/\Re$ verifying

$$\Delta \varphi_v = v \quad in\, \Omega \tag{42}$$

$$\vec{\nabla} \varphi_v \cdot \vec{n}\big|_\Gamma = 0 \quad on\, \Gamma \tag{43}$$

For a certain constant C, we have $\left\|\vec{\nabla}\varphi_v\right\|_{0,\Omega} \le C\|v\|_{0,\Omega}$, and the "inf-sup" condition follows for $\beta = 1/\left(1 + C^2\right)^{1/2}$

$$\sup_{\vec{q} \in X} \frac{b(\vec{q}, v)}{\|\vec{q}\|_{H(div;\Omega)}} \ge \frac{\|v\|_{0,\Omega}^2}{\left(\left\|\vec{\nabla}\varphi_v\right\|_{0,\Omega}^2 + \|\varphi_v\|_{0,\Omega}^2\right)^{1/2}} \ge \frac{1}{\left(1 + C^2\right)^{1/2}} \|v\|_{0,\Omega} \tag{44}$$

The space of functions $V_h = \left\{\psi_h \in E_{0,h}^{(s)} : b(\psi_h, \mu_h) = 0 \,\forall \mu_h \in L_h^{(s-1)}\right\}$ verifies $V_h \subset V$, so the bilinear form $a(\cdot,\cdot)$ is also V_h-elliptic. Moreover, the interpolation operator $\Pi_h^{(s)}$ satisfies

$$\forall \vec{q} \in H_0(div;\Omega), \forall v_h \in L_h^{(s-1)}, \int_\Omega \vec{\nabla}\left(\vec{q} - \Pi_h^{(s)}\vec{q}\right) v_h = 0 \tag{45}$$

$$\forall \vec{q} \in H_0(div;\Omega), \left\|\Pi_h^{(s)}\vec{q}\right\|_{H(div;\Omega)} \le C\|\vec{q}\|_{H(div;\Omega)} \tag{46}$$

hence, the discrete "inf-sup" Babuska-Brezzi condition follows and this assures the existence and uniqueness of the solution of the discrete elliptic problem given by (35) and (36), and also the error bound (Roberts & Thomas, 1991)

$$\left\|\vec{p}(t)-\tilde{\vec{p}}_h(t)\right\|_{H(div;\Omega)}+\left\|u(t)-\tilde{u}_h(t)\right\|_{0,\Omega}\le C\left\{\inf_{\vec{q}_h\in E_{0,h}^{(s)}}\left\|\vec{p}-\vec{q}_h\right\|_{H(div;\Omega)}+\inf_{v_h\in L_h^{(s-1)}}\left\|u-v_h\right\|_{0,\Omega}\right\} \quad (47)$$

On the other side, for $\vec{p}(t)$ sufficiently regular, we have

$$\left\|\vec{p}(t)-\mathcal{E}_h^{(s)}\vec{p}(t)\right\|_{0,\Omega}\le Ch^s\left|\vec{p}(t)\right|_{s,\Omega}$$

$$\left\|\vec{\nabla}\cdot\left(\vec{p}(t)-\mathcal{E}_h^{(s)}\vec{p}(t)\right)\right\|_{0,\Omega}\le Ch^s\left|\vec{\nabla}\cdot\vec{p}(t)\right|_{s,\Omega} \quad (48)$$

which gives a bound in the norm of $H(div;\Omega)$,

$$\left\|\vec{p}(t)-\mathcal{E}_h^{(s)}\vec{p}(t)\right\|_{H(div;\Omega)}\le Ch^s\left(\left|\vec{p}(t)\right|_{s,\Omega}+\left|\vec{\nabla}\cdot\vec{p}(t)\right|_{s,\Omega}\right) \quad (49)$$

In the same way, for $u(t)$ sufficiently regular, we have

$$\left\|u(t)-\mathcal{L}_h^{(s-1)}u(t)\right\|_{0,\Omega}\le Ch^s\left|u(t)\right|_{s,\Omega} \quad (50)$$

Then, we are dare by taking $\vec{q}_h=\mathcal{E}_h^{(s)}\vec{p}(t)$ and $v_h=\mathcal{L}_h^{(s-1)}u(t)$.

■

Lemma 2. If $(u,\vec{p})$ are sufficiently regular, then $\forall t\in(0,T)$ we have

$$\left\|\rho'(t)\right\|_{0,\Omega}\le C(u(t))h^s,$$

$$\left\|\vec{r}'(t)\right\|_{H(div;\Omega)}\le C(u(t))h^s. \quad (51)$$

❑

Proof.

The reasoning in this case is not so simple as in the previous lemma. Let us see the corresponding elliptic problem in this case: differentiating (10) with respect to the time variable, we have $\forall t\in(0,T)$,

$$\left(A(u(t))\,\vec{p}'(t),\vec{q}\right)+\left(u'(t),\vec{\nabla}\cdot\vec{q}\right)=-\left(\left(A(u(t))\right)'\vec{p}(t),\vec{q}\right)\quad\forall\vec{q}\in H_0(div;\Omega) \quad (52)$$

Differentiating (35) and (36) in the same way,

$$\left(\vec{\nabla}\cdot\tilde{\vec{p}}_h'(t),v_h\right)=\left(\vec{\nabla}\cdot\vec{p}'(t),v_h\right)\quad\forall v_h\in L_h^{(s-1)} \quad (53)$$

$$\left(A(u(t))\tilde{\vec{p}}_h'(t),\vec{q}_h\right)+\left(\tilde{u}_h'(t),\vec{\nabla}\cdot\vec{q}_h\right)=-\left(\left(A(u(t))\right)'\tilde{\vec{p}}_h(t),\vec{q}_h\right)\quad\forall\vec{q}_h\in E_h^{(s)} \quad (54)$$

Thus in this case the spaces of functions and the bilinear form $b(\cdot,\cdot)$ are the same, but now $f=\vec{\nabla}\cdot\vec{p}'(t)$ and we cannot be certain that it belongs to $L^2(\Omega)/\Re$.

We are going to use the following general result (Brezzi and Fortin, 1991): let X and M be two Hilbert spaces over $\Re$, $\|\cdot\|_X$ and $\|\cdot\|_M$ their respective norms, X' and M' the corresponding dual spaces, and $\langle\cdot,\cdot\rangle_X$ and $\langle\cdot,\cdot\rangle_M$ the duality products. Let $X_h\subset X$ and $M_h\subset M$ be certain finite-dimensional subspaces. Let $b(\cdot,\cdot):X\times M - \Re$ be a continuous bilinear form. $\mathcal{B}\in\mathcal{L}(X,M')$ denotes the linear operator associated to this bilinear form, and $\mathcal{B}_h\in\mathcal{L}(X_h,M_h')$ the corresponding discrete operator. Given $\varphi\in X$, we define the afinne variety

$$V_h(\mathcal{B}\varphi)=\left\{\Psi_h\in X_h:\langle\mathcal{B}_h\Psi_h,\mu_h\rangle_M=\langle\mathcal{B}\varphi,\mu_h\rangle_M\quad\forall\mu_h\in M_h\right\} \quad (55)$$

If it verifies the discrete "inf-sup" Babuska-Brezzi condition, that is, if there exists a constant $\beta^*>0$ such that $\forall\mu_h\in M_h$

$$\sup_{\psi_h \in X_h} \frac{b(\psi_h, \mu_h)}{\|\psi_h\|_X} \geq \beta^* \|\mu_h\|_M \tag{56}$$

then

$$\inf_{z_h \in V_h(\varphi)} \|z_h - \varphi\|_X \leq \left(\frac{\|b\|}{\beta^*} + 1 \right) \inf_{z_h \in X_h} \|z_h - \varphi\|_X \tag{57}$$

In our particular case, we define the following afinne variety,

$$V_h(\vec{\nabla} \cdot \vec{p}'(t)) = \{ \vec{z}_h \in E_h^{(s)} : (\vec{\nabla} \cdot \vec{z}_h, v_h) = (\vec{\nabla} \cdot \vec{p}'(t), v_h) \quad \forall v_h \in L_h^{(s-1)} \} \tag{58}$$

For each $\vec{z}_h \in V_h(\vec{\nabla} \cdot \vec{p}'(t))$, let $\vec{w}_h = \tilde{\vec{p}}_h'(t) - \vec{z}_h$. It is clear that $\vec{w}_h \in V_h$,

$$(\vec{\nabla} \cdot \vec{w}_h, v_h) = (\vec{\nabla} \cdot \tilde{\vec{p}}_h'(t), v_h) - (\vec{\nabla} \cdot \vec{z}_h, v_h) = (\vec{\nabla} \cdot \vec{p}'(t), v_h) - (\vec{\nabla} \cdot \vec{p}'(t), v_h) = 0 \tag{59}$$

Then, taking norms and using (52) and (54), it follows that

$$\begin{aligned} &\mu \frac{1}{\gamma} \|\vec{w}_h\|_{0,\Omega}^2 = \frac{1}{\mu} (\vec{w}_h, \vec{w}_h) \leq (A(u(t))\vec{w}_h, \vec{w}_h) = (A(u(t))\tilde{\vec{p}}_h'(t), \vec{w}_h) - (A(u(t))\vec{z}_h, \vec{w}_h) = \\ &= -\left((A(u(t)))' \vec{r}(t), \vec{w}_h \right) + (A(u(t))(\vec{p}'(t) - \vec{z}_h), \vec{w}_h) + (u'(t) - v_h, \vec{\nabla} \cdot \vec{w}_h) \leq \\ &\leq \left\| (A(u(t)))' \vec{r}(t) \right\|_{0,\Omega} \|\vec{w}_h\|_{0,\Omega} + \|A(u(t))(\vec{p}'(t) - \vec{z}_h)\|_{0,\Omega} \|\vec{w}_h\|_{0,\Omega} + (u'(t) - v_h, \vec{\nabla} \cdot \vec{w}_h) \end{aligned} \tag{60}$$

Now, realigning, dividing by $\|\vec{w}_h\|_{0,\Omega}$, by the properties of $A(u(t))$, we have

$$\begin{aligned} &\|\vec{w}_h\|_{0,\Omega} \leq \mu \left\{ \left\| (A(u(t)))' \vec{r}(t) \right\|_{0,\Omega} + \|A(u)(\vec{p}'(t) - \vec{z}_h)\|_{0,\Omega} + \frac{(u'(t) - v_h, \vec{\nabla} \cdot \vec{w}_h)}{\|\vec{w}_h\|_{0,\Omega}} \right\} \leq \\ &\frac{\mu L_a}{\alpha^2} \|\vec{r}(t)\|_{0,\Omega} + \frac{\mu}{\alpha} \|\vec{p}'(t) - \vec{z}_h\|_{0,\Omega} + \mu \sup_{\vec{s}_h \in V_h} \frac{(u'(t) - v_h, \vec{\nabla} \cdot \vec{s}_h)}{\|\vec{s}_h\|_{0,\Omega}} \end{aligned} \tag{61}$$

Since $\vec{w}_h \in V_h$, we have $\|\vec{w}_h\|_{0,\Omega} = \|\vec{w}_h\|_{H(div;\Omega)}$, so $\forall \vec{z}_h \in V_h(\vec{\nabla} \cdot \vec{p}'(t))$ and $\forall v_h \in M_h$ it follows that

$$\begin{aligned} &\|\vec{r}'(t)\|_{H(div;\Omega)} = \|\tilde{\vec{p}}_h'(t) - \vec{p}'(t)\|_{H(div;\Omega)} \leq \|\tilde{\vec{p}}_h'(t) - \vec{z}_h\|_{H(div;\Omega)} + \|\vec{z}_h - \vec{p}'(t)\|_{H(div;\Omega)} = \\ &= \|\vec{w}_h\|_{H(div;\Omega)} + \|\vec{z}_h - \vec{p}'(t)\|_{H(div;\Omega)} \leq \\ &\leq \frac{\mu L_a}{\alpha^2} \|\vec{r}(t)\|_{H(div;\Omega)} + \left(1 + \frac{\mu}{\alpha}\right) \|\vec{p}'(t) - \vec{z}_h\|_{H(div;\Omega)} + \mu \sup_{\vec{s}_h \in V_h} \frac{(u'(t) - v_h, \vec{\nabla} \cdot \vec{s}_h)}{\|\vec{s}_h\|_{H(div;\Omega)}} \end{aligned} \tag{62}$$

Taking the minimum over the corresponding spaces, we have

$$\begin{aligned} \|\vec{r}'(t)\|_{H(div;\Omega)} \leq \frac{\mu L_a}{\alpha^2} \|\vec{r}(t)\|_{H(div;\Omega)} &+ \left(1 + \frac{\mu}{\alpha}\right) \inf_{\vec{z}_h \in V_h(\vec{\nabla} \cdot \vec{p}')} \|\vec{p}'(t) - \vec{z}_h\|_{H(div;\Omega)} + \\ &+ \mu \inf_{v_h \in M_h} \sup_{\vec{s}_h \in V_h} \frac{(u'(t) - v_h, \vec{\nabla} \cdot \vec{s}_h)}{\|\vec{s}_h\|_{H(div;\Omega)}} \end{aligned} \tag{63}$$

Now, by the general result described before for this afinne variety $V_h(\vec{\nabla} \cdot \vec{p}'(t))$,

$$\begin{aligned} \|\vec{r}'(t)\|_{H(div;\Omega)} \leq \frac{\mu L_a}{\alpha^2} \|\vec{r}(t)\|_{H(div;\Omega)} &+ \left(1 + \frac{\mu}{\alpha}\right)\left(1 + \frac{1}{\beta^*}\right) \inf_{\vec{z}_h \in X_h} \|\vec{p}'(t) - \vec{z}_h\|_{H(div;\Omega)} + \\ &+ \mu \inf_{v_h \in M_h} \|u'(t) - v_h\|_{0,\Omega} \end{aligned} \tag{64}$$

Then, if $u'(t)$ and $\vec{p}'(t)$ are sufficiently smooth, using the bound obtained for

$\|\vec{r}(t)\|_{H(div;\Omega)}$ in the previous lemma, taking $\vec{z}_h = \Pi^{(s)}_{0,h}\vec{p}'(t)$ and $v_h = P_h^{(s-1)}u'(t)$, we conclude

$$\begin{aligned}\|\vec{r}'(t)\|_{H(div;\Omega)} &\le \frac{\mu L_a}{\alpha^2}\|\vec{r}(t)\|_{H(div;\Omega)} + \left(1+\frac{\mu}{\alpha}\right)\left(1+\frac{1}{\beta^*}\right)\left\|\vec{p}'(t) - \Pi^{(s)}_{0,h}\vec{p}'(t)\right\|_{H(div;\Omega)} \\ &+ \mu\left\|u'(t) - P_h^{(s-1)}u'(t)\right\| \le C(u(t))h^s\end{aligned} \tag{65}$$

In order to bound $\|\rho'(t)\|_{0,\Omega}$, from equations (52) and (54), it follows that

$$\begin{aligned}&\left(\widetilde{u}_h'(t) - v_h, \vec{\nabla}\cdot\vec{q}_h\right) = -\left((A(u(t)))'\,\widetilde{\vec{p}}_h(t), \vec{q}_h\right) - \left(A(u)\widetilde{\vec{p}}'(t)_h, \vec{q}_h\right) - \left(v_h, \vec{\nabla}\cdot\vec{q}_h\right) = \\ &= \left((A(u(t)))'\left(\vec{p}(t) - \widetilde{\vec{p}}_h(t)\right), \vec{q}_h\right) + \left(A(u(t))\left(\vec{p}'(t) - \widetilde{\vec{p}}_h'(t)\right), \vec{q}_h\right) + \left(u'(t) - v_h, \vec{\nabla}\cdot\vec{q}_h\right)\end{aligned} \tag{66}$$

As it verifies the discrete "inf-sup" condition, we can write

$$\begin{aligned}&\|\widetilde{u}_h'(t) - v_h\|_{0,\Omega} \le \frac{1}{\beta^*}\sup_{\vec{q}_h\in X_h}\frac{\left(\widetilde{u}_h'(t) - v_h, \vec{\nabla}\cdot\vec{q}_h\right)}{\|\vec{q}_h\|_{H(div;\Omega)}} = \\ &= \frac{1}{\beta^*}\sup_{\vec{q}_h\in X_h}\frac{1}{\|\vec{q}_h\|_{H(div;\Omega)}}\left\{\left((A(u(t)))'\left(\vec{p}(t) - \widetilde{\vec{p}}_h(t)\right), \vec{q}_h\right) + \left(A(u(t))\left(\vec{p}'(t) - \widetilde{\vec{p}}_h'(t)\right), \vec{q}_h\right) + \left(u'(t) - v_h, \vec{\nabla}\cdot\vec{q}_h\right)\right\} \le \\ &\le \frac{1}{\beta^*}\left\{\left\|(A(u(t)))'\left(\vec{p}(t) - \widetilde{\vec{p}}_h(t)\right)\right\|_{0,\Omega} + \left\|A(u(t))\left(\vec{p}'(t) - \widetilde{\vec{p}}'(t)_h\right)\right\|_{0,\Omega} + \sup_{\vec{q}_h\in X_h}\frac{\|u'(t) - v_h\|_{0,\Omega}\|\vec{\nabla}\cdot\vec{q}_h\|_{0,\Omega}}{\|\vec{q}_h\|_{H(div;\Omega)}}\right\} \le \\ &\le \frac{1}{\beta^*}\left\{\frac{L_a}{\alpha^2}\|\vec{r}(t)\|_{H(div;\Omega)} + \frac{1}{\alpha}\|\vec{r}'(t)\|_{H(div;\Omega)} + \|u'(t) - v_h\|_{0,\Omega}\right\}\end{aligned} \tag{67}$$

So, $\forall v_h \in M_h$ we have

$$\begin{aligned}\|\rho'(t)\|_{0,\Omega} &= \|\widetilde{u}_h'(t) - u'(t)\|_{0,\Omega} \le \|\widetilde{u}_h'(t) - v_h\|_{0,\Omega} + \|v_h - u'(t)\|_{0,\Omega} \le \\ &\le \frac{L_a}{\beta^*\alpha^2}\|\vec{r}(t)\|_{H(div;\Omega)} + \frac{1}{\beta^*\alpha}\|\vec{r}'(t)\|_{H(div;\Omega)} + \left(1+\frac{1}{\beta^*}\right)\|u'(t) - v_h\|_{0,\Omega}\end{aligned} \tag{68}$$

For $u(t)$, $\vec{p}(t)$, $u'(t)$ and $\vec{p}'(t)$ sufficiently smooth, taking $v_h = P_h^{(s-1)}u'(t)$, by the previous bounds, we conclude that $\forall t \in (0,T)$

$$\|\rho'(t)\|_{0,\Omega} \le \frac{L_a}{\beta^*\alpha^2}\|\vec{r}(t)\|_{H(div;\Omega)} + \frac{1}{\beta^*\alpha}\|\vec{r}(t)\|_{H(div;\Omega)} + \left(1+\frac{1}{\beta^*}\right)\left\|u'(t) - P_h^{(s-1)}u'(t)\right\|_{0,\Omega} \le C(u(t))h^s \tag{69}$$

■

4. TOTAL ERROR ESTIMATES

In order to estimate the total error for any of the three completely discrete schemes described before, we break down the total error into the form

$$u_h^n - u(nk) = \left(u_h^n - \widetilde{u}_h^n\right) + \left(\widetilde{u}_h^n - u(nk)\right) = \vartheta^n + \rho^n \tag{70}$$

$$\vec{p}_h^n - \vec{p}(nk) = \left(\vec{p}_h^n - \widetilde{\vec{p}}_h^n\right) + \left(\widetilde{\vec{p}}_h^n - \vec{p}(nk)\right) = \vec{z}^n + \vec{r}^n \tag{71}$$

with $\widetilde{u}_h^n = \widetilde{u}_h(nk)$ and $\widetilde{\vec{p}}_h^n = \widetilde{\vec{p}}_h(nk)$, where $\left(\widetilde{u}_h(t), \widetilde{\vec{p}}_h(t)\right)$ is the elliptic projection of the exact solution at time *t* given by (29) and (30). Then, for each time step $t_n = nk$, $\left(\widetilde{u}_h^n, \widetilde{\vec{p}}_h^n\right)$ satisfies

$$\left(\vec{\nabla}\cdot\left(\vec{p}(nk)-\tilde{\vec{p}}_h^n\right),v_h\right)=0 \quad \forall v_h\in L_h^{(s-1)} \tag{72}$$
$$\left(A(u(nk))\left(\vec{p}(nk)-\tilde{\vec{p}}_h^n\right),\vec{q}_h\right)+\left(u(nk)-\tilde{u}_h^n,\vec{\nabla}\cdot\vec{q}_h\right)=0 \quad \forall\vec{q}_h\in E_{0,h}^{(s)} \tag{73}$$

Since $(u(t),\vec{p}(t))$ is the solution of the continuous problem given by (9) and (10), for each time step $t_n=nk$, we have

$$\left(A(u(nk))\vec{p}(nk),\vec{q}\right)+\left(u(nk),\vec{\nabla}\cdot\vec{q}\right)=0 \quad \forall\vec{q}\in H_0(div;\Omega) \tag{74}$$

This is also true $\forall\vec{q}_h\in E_{0,h}^{(s)}\subset H_0(div;\Omega)$, so $\left(\tilde{u}_h^n,\tilde{\vec{p}}_h^n\right)$ verifies

$$\left(\vec{\nabla}\cdot\tilde{\vec{p}}_h^n,v_h\right)=\left(\vec{\nabla}\cdot\vec{p},v_h\right) \quad \forall v_h\in L_h^{(0)} \tag{75}$$
$$\left(A(u(nk))\tilde{\vec{p}}_h^n,\vec{q}_h\right)+\left(\tilde{u}_h^n,\vec{\nabla}\cdot\vec{q}_h\right)=0 \quad \forall\vec{q}_h\in E_{0,h}^{(1)} \tag{76}$$

Then, by lemma 1, it follows that

$$\left\|\rho^n\right\|_{0,\Omega}\le C(u(nk))h^s \tag{77}$$
$$\left\|\vec{r}^{\,n}\right\|_{H(div;\Omega)}\le C(u(nk))h^s \tag{78}$$

So it just remains to bound ϑ^n and $\vec{z}^{\,n}$ for each of the three schemes.

Theorem 1. Under the appropriate regularity assumptions, the implicit Euler scheme, for small enough *k*, verifies

$$\left\|u_h^n-u(nk)\right\|_{0,\Omega}\le C(u)\left(\left\|u_{I,h}-u_I\right\|_{0,\Omega}+h^s+k\right) \tag{79}$$
$$\left(\sum_{j=1}^{n}k\left\|\vec{p}_h^{\,j}-\vec{p}(nj)\right\|_{0,\Omega}^2\right)^{1/2}\le C(u)\left(\left\|u_{I,h}-u_I\right\|_{0,\Omega}+h^s+k\right) \tag{80}$$

❑

Proof.

The definition of ϑ^n and $\vec{z}^{\,n}$ gives

$$\left(\overline{\partial}_t u_h^n,v_h\right)=\left(\overline{\partial}_t\vartheta^n,v_h\right)+\left(\overline{\partial}_t\tilde{u}_h^n,v_h\right) \tag{81}$$
$$\left(\vec{\nabla}\cdot\vec{p}_h^n,v_h\right)=\left(\vec{\nabla}\cdot\vec{z}^{\,n},v_h\right)+\left(\vec{\nabla}\cdot\tilde{\vec{p}}_h^n,v_h\right) \tag{82}$$

Since $\left(u_h^n,\vec{p}_h^{\,n}\right)$ is the solution of the discrete parabolic problem, and $(u(nk),\vec{p}(nk))$ is the solution of the continuous parabolic problem at time t_n, from the equations (9) and (19) it follows that

$$\left(\overline{\partial}_t u_h^n,v_h\right)=\left(\vec{\nabla}\cdot\vec{p}_h^n,v_h\right)+\left(f\left(u_h^n\right),v_h\right) \tag{83}$$
$$\left(u_t(nk),v_h\right)=\left(\vec{\nabla}\cdot\vec{p}(nk),v_h\right)+\left(f(u(nk)),v_h\right) \tag{84}$$

realigning and taking into account the equations (33) and (34),

$$\left(\overline{\partial}_t\vartheta^n,v_h\right)=\left(\vec{\nabla}\cdot\vec{z}^{\,n},v_h\right)+\left(f\left(u_h^n\right)-f(u(nk)),v_h\right)+\left(u'(nk)-\overline{\partial}_t\tilde{u}_h^n,v_h\right) \tag{85}$$

On the other side, by equations (10) and (20),

$$\left(A\left(u_h^n\right)\vec{z}^{\,n},\vec{q}_h\right)+\left(\vartheta^n,\vec{\nabla}\cdot\vec{q}_h\right)=\left(\left(A(u(nk))-A\left(u_h^n\right)\right)\tilde{\vec{p}}_h^n,\vec{q}_h\right) \tag{86}$$

Taking $v_h=\vartheta^n$ in (85) and $\vec{q}_h=\vec{z}^{\,n}$ in (86),

$$\left(\overline{\partial}_t\vartheta^n,\vartheta^n\right)=\left(\vec{\nabla}\cdot\vec{z}^{\,n},\vartheta^n\right)+\left(f\left(u_h^n\right)-f(u(nk)),\vartheta^n\right)+\left(u'(nk)-\overline{\partial}_t\tilde{u}_h^n,\vartheta^n\right) \tag{87}$$
$$\left(A\left(u_h^n\right)\vec{z}^{\,n},\vec{z}^{\,n}\right)+\left(\vartheta^n,\vec{\nabla}\cdot\vec{z}^{\,n}\right)=\left(\left(A(u(nk))-A\left(u_h^n\right)\right)\tilde{\vec{p}}_h^n,\vec{z}^{\,n}\right) \tag{88}$$

and realigning,

$$\begin{aligned}&\left(\overline{\partial}_t\vartheta^n,\vartheta^n\right)+\left(A\left(u_h^n\right)\vec{z}^{\,n},\vec{z}^{\,n}\right)=\left(f\left(u_h^n\right)-f(u(nk)),\vartheta^n\right)+\\&+\left(u'(nk)-\overline{\partial}_t\tilde{u}_h^n,\vartheta^n\right)+\left(\left(A(u(nk))-A\left(u_h^n\right)\right)\tilde{\vec{p}}_h^n,\vec{z}^{\,n}\right)\end{aligned} \tag{89}$$

By the properties of $A(u)$ and $f(u)$, this becomes

$$\left(A\left(u_h^n\right)\vec{z}^{\,n},\vec{z}^{\,n}\right)\geq\frac{1}{\mu}\left\|\vec{z}^{\,n}\right\|_{0,\Omega}^2 \tag{90}$$

$$\begin{aligned}\left(f\left(u_h^n\right)-f(u(nk)),\vartheta^n\right)&\leq\frac{1}{2}L_f^2\left\|u_h^n-u(nk)\right\|_{0,\Omega}^2+\frac{1}{2}\left\|\vartheta^n\right\|_{0,\Omega}^2\leq\\&\leq\frac{1}{2}L_f^2\left(\left\|\vartheta^n\right\|_{0,\Omega}^2+\left\|\rho^n\right\|_{0,\Omega}^2\right)+\frac{1}{2}\left\|\vartheta^n\right\|_{0,\Omega}^2\end{aligned} \tag{91}$$

$$\left(u'(nk)-\overline{\partial}_t\widetilde{u}_h^n,\vartheta^n\right)\leq\frac{1}{2}\left\|u'(nk)-\overline{\partial}_t\widetilde{u}_h^n\right\|_{0,\Omega}^2+\frac{1}{2}\left\|\vartheta^n\right\|_{0,\Omega}^2 \tag{92}$$

$$\begin{aligned}\left(\left(A(u(nk))-A\left(u_h^n\right)\right)\widetilde{\vec{p}}_h^{\,n},\vec{z}^{\,n}\right)&\leq\frac{\mu}{4}\left\|\left(A(u(nk))-A\left(u_h^n\right)\right)\widetilde{\vec{p}}_h^{\,n}\right\|_{0,\Omega}^2+\frac{1}{\mu}\left\|\vec{z}^{\,n}\right\|_{0,\Omega}^2\leq\\&\leq\frac{\mu L_A^2}{4}\left(\left\|\rho^n\right\|_{0,\Omega}^2+\left\|\vartheta^n\right\|_{0,\Omega}^2\right)\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{L_\infty}^2+\frac{1}{\mu}\left\|\vec{z}^{\,n}\right\|_{0,\Omega}^2\end{aligned} \tag{93}$$

Realigning and taking into account that $\left(\overline{\partial}_t\vartheta^n,\vartheta^n\right)=\frac{1}{2}\overline{\partial}_t\left\|\vartheta^n\right\|_{0,\Omega}^2+\frac{k}{2}\left\|\overline{\partial}_t\vartheta^n\right\|_{0,\Omega}^2$, it follows

$$\begin{aligned}\overline{\partial}_t\left\|\vartheta^n\right\|_{0,\Omega}^2+k\left\|\overline{\partial}_t\vartheta^n\right\|_{0,\Omega}^2+\frac{2}{\mu}\left\|\vec{z}^{\,n}\right\|_{0,\Omega}^2&\leq\left(L_f^2+\frac{\mu L_A^2}{2}\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{L_\infty}^2\right)\left\|\rho^n\right\|_{0,\Omega}^2+\\&+\left\|u'(nk)-\overline{\partial}_t\widetilde{u}_h^n\right\|_{0,\Omega}^2+\left(L_f^2+\frac{\mu L_A^2}{2}\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{L_\infty}^2+2\right)\left\|\vartheta^n\right\|_{0,\Omega}^2+\frac{2}{\mu}\left\|\vec{z}^{\,n}\right\|_{0,\Omega}^2\end{aligned} \tag{94}$$

Simplifying, if $k\left\|\overline{\partial}_t\vartheta^n\right\|_{0,\Omega}^2\geq 0$, we can write

$$\begin{aligned}\overline{\partial}_t\left\|\vartheta^n\right\|_{0,\Omega}^2&\leq\left(L_f^2+\frac{\mu L_A^2}{2}\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{L_\infty}^2\right)\left\|\rho^n\right\|_{0,\Omega}^2+\\&+\left\|u'(nk)-\overline{\partial}_t\widetilde{u}_h^n\right\|_{0,\Omega}^2+\left(L_f^2+\frac{\mu L_A^2}{2}\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{L_\infty}^2+2\right)\left\|\vartheta^n\right\|_{0,\Omega}^2\end{aligned} \tag{95}$$

By the definition of the regresive difference, and denoting

$$C=L_f^2+\frac{\mu L_A^2}{2}\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{L_\infty}^2+2 \tag{96}$$

$$R_n=\left(L_f^2+\frac{\mu L_A^2}{2}\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{L_\infty}^2\right)\left\|\rho^n\right\|_{0,\Omega}^2+\left\|u'(nk)-\overline{\partial}_t\widetilde{u}_h^n\right\|_{0,\Omega}^2 \tag{97}$$

it follows that

$$\left\|\vartheta^n\right\|_{0,\Omega}^2\leq\frac{1}{1-kC}\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2+\frac{k}{1-kC}R_n \tag{98}$$

For $k\leq\frac{1}{2C}$, we have $\frac{1}{1-kC}\leq 1+2kC\leq e^{2kC}$ and $\frac{k}{1-kC}\leq 2k$, so we can write,

$$\left\|\vartheta^n\right\|_{0,\Omega}^2\leq e^{2kC}\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2+2kR_n \tag{99}$$

and renaming $C=e^{2knC}$ this becomes

$$\left\|\vartheta^n\right\|_{0,\Omega}^2\leq C\left\|\vartheta^0\right\|_{0,\Omega}^2+2kC\sum_{j=1}^{n}R_j \tag{100}$$

It remains to study the different terms of R_j:

i) $\quad \left\| u'(jk) - \overline{\partial}_t \tilde{u}_h^j \right\|_{0,\Omega} \le C(u)\left(h^s + k\right)$.

Effectively, we can expand the first term in the following way

$$u'(jk) - \overline{\partial}_t \tilde{u}_h^j = u'(jk) - \overline{\partial}_t u(jk) + \overline{\partial}_t u(jk) - \overline{\partial}_t \tilde{u}_h^j = \underbrace{u'(jk) - \overline{\partial}_t u(jk)}_{(1)} + \underbrace{\overline{\partial}_t \rho^j}_{(2)} \qquad (101)$$

where,

$$\begin{aligned} &(1) \quad u'(jk) - \overline{\partial}_t u(jk) = \frac{1}{k}\left[ku'(jk) - \left(u(jk) - u((j-1)k)\right)\right] = \\ &= \frac{1}{k}\int_{(j-1)k}^{jk} (s-(j-1)k)u''(s)ds \end{aligned} \qquad (102)$$

Then,

$$\left\| u'(jk) - \overline{\partial}_t u(jk) \right\|_{0,\Omega} = \left\| \frac{1}{k}\int_{(j-1)k}^{jk} (s-(j-1)k)u''(s)ds \right\|_{0,\Omega} \le C(u)k \qquad (103)$$

On the other side,

$$(2) \quad \overline{\partial}_t \rho^j = \frac{1}{k}\int_{(j-1)k}^{jk} \rho'(s)ds \qquad (104)$$

and by lemma 2, we have

$$\left\| \overline{\partial}_t \rho^j \right\|_{0,\Omega} = \left\| \frac{1}{k}\int_{(j-1)k}^{jk} \rho'(s)ds \right\|_{0,\Omega} \le C(u)h^s \qquad (105)$$

ii) $\quad \left\| \rho^j \right\|_{0,\Omega} \le C(u)h^s$ by lemma 2.

iii) $\quad \left\| \tilde{\vec{p}} j \right\|_{0,\infty,\Omega} \le C(u)$

Let $\boldsymbol{\pi}_h$ be the following interpolation operator over $H(div;\Omega)$: as $\vec{p} = a(u)\vec{\nabla} u$, we define $g(u) = \int_0^u a(\xi)d\xi$, so $g(u)$ is differentiable almost everywhere in the domain, $g'(u) = a(u)$ and $\vec{p} = \vec{\nabla} g(u)$; we now define $\boldsymbol{\pi}_h$ such that $\boldsymbol{\pi}_h \vec{p}(jk) = \vec{\nabla} I_h g(u(jk))$ where I_h is a linear interpolation operator. This operator verifies

$$\left\| \vec{p}(jk) - \boldsymbol{\pi}_h \vec{p}(jk) \right\|_{0,\Omega} = \left\| \vec{\nabla} g(u(jk)) - \vec{\nabla} I_h g(u(jk)) \right\|_{0,\Omega} \le C(u(jk))h \qquad (106)$$

$$\left\| \boldsymbol{\pi}_h \vec{p}(jk) \right\|_{0,\infty,\Omega} \le C\left\| \vec{p}(jk) \right\|_{0,\infty,\Omega} \qquad (107)$$

And we can reason as follows using lemma 1,

$$\begin{aligned} &\left\| \tilde{\vec{p}}_h^j \right\|_{L\infty} \le \left\| \tilde{\vec{p}}_h^j - \boldsymbol{\pi}_h \vec{p}(jk) \right\|_{0,\infty,\Omega} + \left\| \boldsymbol{\pi}_h \vec{p}(jk) \right\|_{0,\infty,\Omega} \le \\ &\le Ch^{-1}\left\| \tilde{\vec{p}}_h^j - \boldsymbol{\pi}_h \vec{p}(jk) \right\|_{0,\Omega} + \left\| \boldsymbol{\pi}_h \vec{p}(jk) \right\|_{0,\infty,\Omega} \le \\ &\le Ch^{-1}\left(\left\| \tilde{\vec{p}}_h^j - \vec{p}(jk) \right\|_{0,\Omega} + \left\| \vec{p}(jk) - \boldsymbol{\pi}_h \vec{p}(jk) \right\|_{0,\Omega} \right) + C\left\| \vec{p}(jk) \right\|_{0,\infty,\Omega} \le \\ &\le Ch^{-1}\left(\left\| \vec{r}(jk) \right\|_{0,\Omega} + C(u)h \right) + C\left\| \vec{p}(jk) \right\|_{0,\infty,\Omega} \le C(u) \end{aligned} \qquad (108)$$

All this together shows that $R_j \le C(u)\left(h^{2s} + k^2\right)$, and then

$$2kC\sum_{j=1}^{n} R_j \le 2kCnC(u)\left(h^2 + k^2\right) \le C(u)\left(h^{2s} + k^2\right) \tag{109}$$

Then, by lemma 1,

$$\begin{aligned} \left\|\vartheta^0\right\|_{0,\Omega}^2 &\le \left\|u_h^0 - u(0)\right\|_{0,\Omega}^2 + \left\|u(0) - \widetilde{u}_h^0\right\|_{0,\Omega}^2 = \\ &= \left\|u_{I,h} - u_I\right\|_{0,\Omega}^2 + \left\|\rho^0\right\|_{0,\Omega}^2 \le \left\|u_{I,h} - u_I\right\|_{0,\Omega}^2 + C(u)h^{2s} \end{aligned} \tag{110}$$

So this becomes

$$\left\|\vartheta^n\right\|_{0,\Omega} \le C(u)\left(\left\|u_{I,h} - u_I\right\|_{0,\Omega} + h^s + k\right) \tag{111}$$

Finally, by lemma 1 and the previous inequality (111), it follows that

$$\left\|u_h^n - u(nk)\right\|_{0,\Omega} \le \left\|\vartheta^n\right\|_{0,\Omega} + \left\|\rho^n\right\|_{0,\Omega} \le C(u)\left(\left\|u_{I,h} - u_I\right\|_{0,\Omega} + h^s + k\right) \tag{112}$$

On the other side, modifying slightly inequality (93), we can write

$$\left(\left(A(u(nk)) - A(u_h^n)\right)\widetilde{\vec{p}}_h^{\,n}, \vec{z}^{\,n}\right) \le \frac{\mu L_A^2}{2}\left(\left\|\rho^n\right\|_{0,\Omega}^2 + \left\|\vartheta^n\right\|_{0,\Omega}^2\right)\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{0,\infty,\Omega}^2 + \frac{1}{2\mu}\left\|\vec{z}^{\,n}\right\|_{0,\Omega}^2 \tag{113}$$

so,

$$\begin{aligned} \overline{\partial}_t\left\|\vartheta^n\right\|_{0,\Omega}^2 + \frac{1}{2\mu}\left\|\vec{z}^{\,n}\right\|_{0,\Omega}^2 &\le \left(L_f^2 + \mu L_A^2\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{L_\infty}^2\right)\left\|\rho^n\right\|_{0,\Omega}^2 + \\ &+ \left\|u'(nk) - \overline{\partial}_t\widetilde{u}_h^n\right\|_{0,\Omega}^2 + \left(L_f^2 + \mu L_A^2\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{L_\infty}^2 + 2\right)\left\|\vartheta^n\right\|_{0,\Omega}^2 \end{aligned} \tag{114}$$

Denoting now,

$$\begin{aligned} C &= L_f^2 + \mu L_A^2\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{0,\infty,\Omega}^2 + 2 \\ R_n &= \left(L_f^2 + \mu L_A^2\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{0,\infty,\Omega}^2\right)\left\|\rho^n\right\|_{0,\Omega}^2 + \left\|u'(nk) - \overline{\partial}_t\widetilde{u}_h^n\right\|_{0,\Omega}^2 \end{aligned} \tag{115}$$

and reasoning as before, for $k \,:\, 1/2C$, and renaming C if necessary, we have

$$\left\|\vartheta^n\right\|_{0,\Omega}^2 + \frac{2}{\mu}\sum_{j=1}^{n} k\left\|\vec{z}^{\,j}\right\|_{0,\Omega}^2 \le C\left\|\vartheta^0\right\|_{0,\Omega}^2 + 2kC\sum_{j=1}^{n} R_j \tag{116}$$

so,

$$\sum_{j=1}^{n} k\left\|\vec{z}^{\,j}\right\|_{0,\Omega}^2 \le C(u)\left(\left\|u_{I,h} - u_I\right\|_{0,\Omega}^2 + h^{2s} + k^2\right) \tag{117}$$

Finally, we conclude by lemma 1,

$$\left(\sum_{j=1}^{n} k\left\|\vec{p}_h^{\,j} - \vec{p}(nj)\right\|_{0,\Omega}^2\right)^{1/2} \le C(u)\left(\left\|u_{I,h} - u_I\right\|_{0,\Omega} + h^s + k\right) \tag{118}$$

■

Theorem 2. Under the appropriate regularity assumptions, the semi-implicit Euler scheme and the explicit Euler scheme verify for $k \,:\, Ch^2$

$$\left\|u_h^n - u(nk)\right\|_{0,\Omega} \le C(u)\left(\left\|u_{I,h} - u_I\right\|_{0,\Omega} + h^s + k\right) \tag{119}$$

$$\left(\sum_{j=1}^{n} k\left\|\vec{p}_h^{\,j} - \vec{p}(nj)\right\|_{0,\Omega}^2\right)^{1/2} \le C(u)\left(\left\|u_{I,h} - u_I\right\|_{0,\Omega} + h^s + k\right) \tag{120}$$

❑

Proof.

Reasoning as in the previous proof, we have for the explicit scheme,

$$\left(\overline{\partial}_t \vartheta^n, v_h\right) = \left(\vec{\nabla} \cdot \vec{z}^{\,n-1}, v_h\right) + \left(f\left(u_h^{n-1}\right) - f\left(u((n-1)k)\right), v_h\right) + \\ + \left(u'((n-1)k) - \overline{\partial}_t \widetilde{u}_h^n, v_h\right) \tag{121}$$

$$\left(A\left(u_h^{n-1}\right)\vec{z}^{\,n-1}, \vec{q}_h\right) + \left(\vartheta^{n-1}, \vec{\nabla} \cdot \vec{q}_h\right) = \left(\left(A(u((n-1)k)) - A\left(u_h^{n-1}\right)\right)\widetilde{\vec{p}}_h^{\,n-1}, \vec{q}_h\right) \tag{122}$$

Taking, $\vec{q}_h = \vec{z}^{\,n-1}$ and $v_h = \vartheta^{n-1}$,

$$\left(\overline{\partial}_t \vartheta^n, \vartheta^{n-1}\right) + \left(A\left(u_h^{n-1}\right)\vec{z}^{\,n-1}, \vec{z}^{\,n-1}\right) = \left(f\left(u_h^{n-1}\right) - f\left(u((n-1)k)\right), \vartheta^n\right) + \\ + \left(u'((n-1)k) - \overline{\partial}_t \widetilde{u}_h^n, \vartheta^n\right) + \left(\left(A(u((n-1)k)) - A\left(u_h^{n-1}\right)\right)\widetilde{\vec{p}}_h^{\,n-1}, \vec{z}^{\,n-1}\right) \tag{123}$$

As in the implicit scheme, inequalities (90) (91) (92) and (113) are also true.

On the other side, in this case it follows that

$$\left(\overline{\partial}_t \vartheta^n, \vartheta^{n-1}\right) = \frac{1}{2}\overline{\partial}_t \left\|\vartheta^n\right\|_{0,\Omega}^2 - \frac{k}{2}\left\|\overline{\partial}_t \vartheta^n\right\|_{0,\Omega}^2 \tag{124}$$

Then,

$$\overline{\partial}_t \left\|\vartheta^n\right\|_{0,\Omega}^2 + \frac{1}{\mu}\left\|\vec{z}^{\,n-1}\right\|_{0,\Omega}^2 \le k\left\|\overline{\partial}_t \vartheta^n\right\|_{0,\Omega}^2 + 2\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + L_f^2\left(\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + \left\|\rho^{n-1}\right\|_{0,\Omega}^2\right) + \\ + \left\|u'((n-1)k) - \overline{\partial}_t \widetilde{u}_h^n\right\|_{0,\Omega}^2 + \mu L_A^2\left(\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + \left\|\rho^{n-1}\right\|_{0,\Omega}^2\right)\left\|\widetilde{\vec{p}}_h^{\,n-1}\right\|_{L\infty}^2 \tag{125}$$

Let us write (121) with $v_h = \overline{\partial}_t \vartheta^n$,

$$\left(\overline{\partial}_t \vartheta^n, \overline{\partial}_t \vartheta^n\right) = \left(\vec{\nabla} \cdot \vec{z}^{\,n-1}, \overline{\partial}_t \vartheta^n\right) + \left(f\left(u_h^{n-1}\right) - f\left(u((n-1)k)\right), \overline{\partial}_t \vartheta^n\right) + \\ + \left(u'((n-1)k) - \overline{\partial}_t \widetilde{u}_h^n, \overline{\partial}_t \vartheta^n\right) \tag{126}$$

Taking norms and reasoning as before we obtain

$$\left\|\overline{\partial}_t \vartheta^n\right\|_{0,\Omega}^2 \le 4\left\|\vec{\nabla} \cdot \vec{z}^{\,n-1}\right\|_{0,\Omega}^2 + 4L_f^2\left(\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + \left\|\rho^{n-1}\right\|_{0,\Omega}^2\right) + 4\left\|u'((n-1)k) - \overline{\partial}_t \widetilde{u}_h^n\right\|_{0,\Omega}^2 \tag{127}$$

so we conclude

$$\overline{\partial}_t \left\|\vartheta^n\right\|_{0,\Omega}^2 + \frac{1}{\mu}\left\|\vec{z}^{\,n-1}\right\|_{0,\Omega}^2 \le 4k\left\|\vec{\nabla} \cdot \vec{z}^{\,n-1}\right\|_{0,\Omega}^2 + 2\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + (1+4k)L_f^2\left(\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + \left\|\rho^{n-1}\right\|_{0,\Omega}^2\right) + \\ + (1+4k)\left\|u'((n-1)k) - \overline{\partial}_t \widetilde{u}_h^n\right\|_{0,\Omega}^2 + \mu L_A^2\left(\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + \left\|\rho^{n-1}\right\|_{0,\Omega}^2\right)\left\|\widetilde{\vec{p}}_h^{\,n-1}\right\|_{L\infty}^2 \tag{128}$$

On the other side, $\left\|\nabla \cdot \vec{z}^{\,n-1}\right\|_{0,\Omega}^2 \le Ch^{-2}\left\|\vec{z}^{\,n-1}\right\|_{0,\Omega}^2$, which allows us to balance the terms in $\vec{z}^{\,n-1}$ in the previous equation choosing $k \le h^2/4\mu C$. Then, for this time step, realigning in (128),

$$\left\|\vartheta^n\right\|_{0,\Omega}^2 \le \left(1 + \left(2 + L_f^2 + \mu L_A^2\left\|\widetilde{\vec{p}}_h^{\,n-1}\right\|_{L_\infty}^2\right)k + 4L_f^2 k^2\right)\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + \\ + k\left(\left(L_f^2 + \mu L_A^2\left\|\widetilde{\vec{p}}_h^{\,n-1}\right\|_{L_\infty}^2\right)\left\|\rho^{n-1}\right\|_{0,\Omega}^2 + \left\|u'((n-1)k) - \overline{\partial}_t \widetilde{u}_h^n\right\|_{0,\Omega}^2\right) + \\ + k^2\left(4L_f^2\left\|\rho^{n-1}\right\|_{0,\Omega}^2 + 4\left\|u'((n-1)k) - \overline{\partial}_t \widetilde{u}_h^n\right\|_{0,\Omega}^2\right) \tag{129}$$

Taking k small enough, we can assume $k^2 \le k$. We have also proved that $\left\|\widetilde{\vec{p}}_h^{\,n}\right\|_{L_\infty} \le C(u)$. Then, denoting now,

$$C = 2 + 5L_f^2 + \mu L_A^2\left\|\widetilde{\vec{p}}_h^{\,n-1}\right\|_{L_\infty}^2 \tag{130}$$

$$R_n = \left(5L_f^2 + \mu L_A^2 \left\|\tilde{\vec{p}}_h^{\,n-1}\right\|_{L_\infty}^2\right)\left\|\rho^{n-1}\right\|_{0,\Omega}^2 + 5\left\|u'((n-1)k) - \overline{\partial}_t \tilde{u}_h^n\right\|_{0,\Omega}^2 \tag{131}$$

we can write

$$\left\|\vartheta^n\right\|_{0,\Omega}^2 \le e^{kC}\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + kR_n \tag{132}$$

Reasoning as in the implicit case and renaming C if necessary, it follows that

$$\left\|\vartheta^n\right\|_{0,\Omega}^2 \le C\left\|\vartheta^0\right\|_{0,\Omega}^2 + kC\sum_{j=1}^{n} R_j \tag{133}$$

The only term of R_j different from the implicit case is $\left\|u'((j-1)k) - \overline{\partial}_t \tilde{u}_h^j\right\|_{0,\Omega}^2$. We can expand this term in a similar way

$$\begin{aligned} u'((j-1)k) - \overline{\partial}_t \tilde{u}_h^j &= u'((j-1)k) - \overline{\partial}_t u(jk) + \overline{\partial}_t u(jk) - \overline{\partial}_t \tilde{u}_h^j = \\ &= u'((j-1)k) - \partial_t u((j-1)k) + \overline{\partial}_t \rho^n \end{aligned} \tag{134}$$

By the definition of regressive difference, one has

$$\begin{aligned} u'((j-1)k) - \partial_t u((j-1)k) &= \frac{1}{k}\left[ku'((j-1)k) - (u(jk) - u((j-1)k))\right] = \\ &= \frac{1}{k}\int_{(j-1)k}^{jk} (s - jk)u''(s)ds \end{aligned} \tag{135}$$

Then,

$$\left\|u'((j-1)k) - \partial_t u((j-1)k)\right\|_{0,\Omega} = \left\|\frac{1}{k}\int_{(j-1)k}^{jk} (s - jk)u''(s)ds\right\|_{0,\Omega} \le C(u)k \tag{136}$$

so again $R_j \le C(u)\left(h^{2s} + k^2\right)$, and we conclude for the explicit scheme

$$\left\|u_h^n - u(nk)\right\|_{0,\Omega} \le C(u)\left(\left\|u_{I,h} - u_I\right\|_{0,\Omega} + h^s + k\right) \tag{137}$$

Taking $k \le h^2/8\mu C$ instead of $k \le h^2/4\mu C$, we can balance only a part of $\left\|\vec{z}^{\,n-1}\right\|_{0,\Omega}^2$, so it becomes

$$\begin{aligned} \left\|\vartheta^n\right\|_{0,\Omega}^2 + \frac{1}{2\mu}\left\|\vec{z}^{\,n-1}\right\|_{0,\Omega}^2 &\le \left(1 + \left(2 + L_f^2 + \mu L_A^2\left\|\tilde{\vec{p}}_h^{\,n-1}\right\|_{L_\infty}^2\right)k + 4L_f^2 k^2\right)\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + \\ &+ k\left(\left(L_f^2 + \mu L_A^2\left\|\tilde{\vec{p}}_h^{\,n-1}\right\|_{L_\infty}^2\right)\left\|\rho^{n-1}\right\|_{0,\Omega}^2 + \left\|u'((n-1)k) - \overline{\partial}_t \tilde{u}_h^n\right\|_{0,\Omega}^2\right) + \\ &+ k^2\left(4L_f^2\left\|\rho^{n-1}\right\|_{0,\Omega}^2 + 4\left\|u'((n-1)k) - \overline{\partial}_t \tilde{u}_h^n\right\|_{0,\Omega}^2\right) \end{aligned} \tag{138}$$

Reasoning as before, we obtain

$$\left\|\vartheta^n\right\|_{0,\Omega}^2 + \frac{1}{2\mu}\sum_{j=1}^{n-1} k\left\|\vec{z}^{\,j}\right\|_{0,\Omega}^2 \le C\left\|\vartheta^0\right\|_{0,\Omega}^2 + kC\sum_{j=1}^{n} R_j \tag{139}$$

so,

$$\sum_{j=1}^{n} k\left\|\vec{z}^{\,j}\right\|_{0,\Omega}^2 \le C(u)\left(\left\|u_{I,h} - u_I\right\|_{0,\Omega}^2 + h^{2s} + k^2\right) \tag{140}$$

and with the boundary of $\left\|\vec{r}^{\,n}\right\|_{0,\Omega}^2$ given by lemma 1, we conclude for the explicit scheme

$$\left(\sum_{j=1}^{n} k\left\|\vec{p}_h^{\,j} - \vec{p}(nj)\right\|_{0,\Omega}^2\right)^{1/2} \le C(u)\left(\left\|u_{I,h} - u_I\right\|_{0,\Omega} + h^s + k\right) \tag{141}$$

For the semi-implicit scheme, the equation (122) remains true , and the equation in ϑ^n is obtained as in the previous cases,

$$\begin{aligned}\left(\overline{\partial}_t\vartheta^n, v_h\right) = \left(\vec{\nabla}\cdot\vec{z}^{\,n-1}, v_h\right) + \left(f\left(u_h^n\right) - f\left(u((n-1)k)\right), v_h\right) + \\ + \left(u'((n-1)k) - \overline{\partial}_t\tilde{u}_h^n, v_h\right)\end{aligned} \tag{142}$$

Taking $\vec{q}_h = \vec{z}^{\,n-1}$ in (122) and $v_h = \vartheta^{n-1}$ in (142), it follows that

$$\begin{aligned}\left(\overline{\partial}_t\vartheta^n, \vartheta^{n-1}\right) + \left(A\left(u_h^{n-1}\right)\vec{z}^{\,n-1}, \vec{z}^{\,n-1}\right) = \left(f\left(u_h^n\right) - f\left(u((n-1)k)\right), \vartheta^n\right) + \\ + \left(u'((n-1)k) - \overline{\partial}_t\tilde{u}_h^n, \vartheta^n\right) + \left(\left(A\left(u((n-1)k)\right) - A\left(u_h^{n-1}\right)\right)\tilde{\vec{p}}_h^{\,n-1}, \vec{z}^{\,n-1}\right)\end{aligned} \tag{143}$$

Reasoning as in the explicit case,

$$\begin{aligned}\overline{\partial}_t\left\|\vartheta^n\right\|_{0,\Omega}^2 + \frac{1}{\mu}\left\|\vec{z}^{\,n-1}\right\|_{0,\Omega}^2 \le k\left\|\overline{\partial}_t\vartheta^n\right\|_{0,\Omega}^2 + 2\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + L_f^2\left\|u_h^n - u((n-1)k)\right\|_{0,\Omega}^2 + \\ + \left\|u'((n-1)k) - \overline{\partial}_t\tilde{u}_h^n\right\|_{0,\Omega}^2 + \mu L_A^2\left(\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + \left\|\rho^{n-1}\right\|_{0,\Omega}^2\right)\left\|\tilde{\vec{p}}_h^{\,n-1}\right\|_{L\infty}^2\end{aligned} \tag{144}$$

Writing the equation (143) for $v_h = \overline{\partial}_t\vartheta^n$, similarly to the explicit case, we obtain the following bound for $\left\|\overline{\partial}_t\vartheta^n\right\|_{0,\Omega}$,

$$\left\|\overline{\partial}_t\vartheta^n\right\|_{0,\Omega}^2 \le 4\left\|\vec{\nabla}\cdot\vec{z}^{\,n-1}\right\|_{0,\Omega}^2 + 4L_f^2\left\|u_h^n - u((n-1)k)\right\|_{0,\Omega}^2 + 4\left\|u'((n-1)k) - \overline{\partial}_t\tilde{u}_h^n\right\|_{0,\Omega}^2 \tag{145}$$

which substituted in (144), gives as before, for $k \le \frac{h^2}{4\gamma C}$, the following bound,

$$\begin{aligned}\overline{\partial}_t\left\|\vartheta^n\right\|_{0,\Omega}^2 \le 2\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + (1+4k)L_f^2\left\|u_h^n - u((n-1)k)\right\|_{0,\Omega}^2 + \\ + (1+4k)\left\|u'((n-1)k) - \overline{\partial}_t\tilde{u}_h^n\right\|_{0,\Omega}^2 + \mu L_A^2\left(\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + \left\|\rho^{n-1}\right\|_{0,\Omega}^2\right)\left\|\tilde{\vec{p}}_h^{\,n-1}\right\|_{L\infty}^2\end{aligned} \tag{146}$$

Expanding $\left\|u_h^n - u((n-1)k)\right\|_{0,\Omega} \le \left\|\vartheta^n\right\|_{0,\Omega} + \left\|\rho^n\right\|_{0,\Omega} + k\left\|\overline{\partial}_t u(nk)\right\|_{0,\Omega}$, realigning and reasoning as before we obtain

$$(1-kC)\left\|\vartheta^n\right\|_{0,\Omega}^2 \le (1+kC)\left\|\vartheta^{n-1}\right\|_{0,\Omega}^2 + kR_n \tag{147}$$

where,

$$C = max\left\{5L_f^2, 2 + \mu L_A^2\left\|\tilde{\vec{p}}_h^{\,n-1}\right\|_{L_\infty}^2\right\} \tag{148}$$

$$R_n = \mu L_A^2\left\|\tilde{\vec{p}}_h^{\,n-1}\right\|_{L_\infty}^2\left\|\rho^{n-1}\right\|_{0,\Omega}^2 + 5L_f^2\left\|\rho^n\right\|_{0,\Omega}^2 + 5\left\|u'((n-1)k) - \overline{\partial}_t\tilde{u}_h^n\right\|_{0,\Omega}^2 + 5L_f^2\left\|\overline{\partial}_t u(nk)\right\|_{0,\Omega}^2 \tag{149}$$

Again, for $k \le \frac{1}{2C}$, and renaming C if necessary, we have

$$\left\|\vartheta^n\right\|_{0,\Omega}^2 \le C\left\|\vartheta^0\right\|_{0,\Omega}^2 + 2kC\sum_{j=1}^{n} R_j \tag{150}$$

Now, the only different term in R_j is $\left\|\overline{\partial}_t u(nk)\right\|_{0,\Omega}^2 \le C(u(nk))$, so here also,

$$\left\|u_h^n - u(nk)\right\|_{0,\Omega} \le C(u(nk))\left(\left\|u_{I,h} - u_I\right\|_{0,\Omega} + h + k\right) \tag{151}$$

For the bound on $\left\|\vec{z}^{\,n-1}\right\|_{0,\Omega}^2$, we reason as in the explicit case and obtain

$$\sum_{j=1}^{n} k\left\|\vec{z}^{\,j}\right\|_{0,\Omega}^2 \le C(u)\left(\left\|u_{I,h} - u_I\right\|_{0,\Omega}^2 + h^{2s} + k^2\right) \tag{152}$$

and with the bound of $\left\|\vec{r}^{\,n}\right\|_{0,\Omega}^{2}$ given by lemma 1, we again obtain

$$\left(\sum_{j=1}^{n} k\left\|\vec{p}_h^{\,j}-\vec{p}(nj)\right\|_{0,\Omega}^{2}\right)^{1/2} \leq C(u)\left(\left\|u_{I,h}-u_I\right\|_{0,\Omega}+h^s+k\right) \tag{153}$$

■

5. APPLICATIONS

We now illustrate the schemes described in this paper with an example of the natural combustion model developed by the authors. This simplified model is based on conservation laws (Cox, 1995; Williams, 1985; Zeldovich et al., 1985) and nondimensionalized in a rational manner in order to elucidate the significant parameters (Bebernes and Eberly, 1989). Let $\Omega=(0,L_x)\times(0,L_y)$ be the rectangular nondimensional domain, $u(x,y,t)$ represents the nondimensional temperature and $\beta(x,y,t)$ the nondimensional mass fraction of solid fuel,

$$\frac{\partial u}{\partial \tau}-\vec{\nabla}\cdot\left(K(u)\vec{\nabla}_{\xi\eta}u\right)=f(u,\beta) \tag{154}$$

$$\frac{d\beta}{dt}=g(u,\beta) \tag{155}$$

where,

$$K(u)=\kappa(1+\varepsilon u)^3+1$$

$$f(u,\beta)=s(u)^{+}\beta\, e^{u/1+\varepsilon u}-\alpha u \tag{156}$$

$$g(u,\beta)=-\frac{\varepsilon}{q}s(u)^{+}\beta\, e^{u/1+\varepsilon u}$$

with ε the nondimensional inverse of the activation energy, q the nondimensional reation heat, κ the nondimensional inverse of the conductivity coefficient, α the nondimensional natural convection coefficient and

$$s(u)^{+}=\left(sig(u-u_{ref})\right)^{+}=\begin{cases}1 & if\ u\geq u_{ref}\\ 0 & otherwise\end{cases} \tag{157}$$

which represents the phase change from the endothermic to the exothermic phase at nondimensional temperature u_{ref} (Asensio and Ferragut, 2000).

The initial conditions are $u(x,y,0)=u_I(x,y)$ and $\beta(x,t,0)=\beta_I(x,y)$, and the boundary condition is $K(u)\vec{\nabla}u\cdot\vec{n}\big|_{\Gamma}=0$.

For this application we have chosen the finite subspaces described before for s=1, that is, the lowest order Raviart-Thomas finite element, $L_h^{(0)}$ and $E_{0,h}^{(1)}$. Thus the discrete solution is not continuous on the edges of the rectangles of the mesh, but the approximation of the heat flux is continuous on the inter-elements boundaries.

The base functions chosen for the space $L_h^{(0)}$ are

$$\varphi_k(\vec{x})=\begin{cases}1 & if\ \vec{x}\in\tau_k\\ 0 & if\ other\end{cases} \tag{158}$$

where k=0,...,L-1, with L=dim $L_h^{(0)}$ equal to the number of rectangles of the mesh.

The base functions chosen for the space $E_{0,h}^{(1)}$ are $\vec{\phi}_n(\vec{x})$ for n=0,...,E_0, verifying

$$\int_{l_m} \vec{\phi}_n(\vec{x}) \cdot \vec{\gamma}_m ds = \delta_{n,m} \quad n,m = 0,\ldots,E_0 \tag{159}$$

with l_n the n-th edge of the mesh, $\vec{\gamma}_m$ the unit normal vector to the m-th edge with the prearranged orientation shown in figure 1, and E_0=dim $E_{0,h}^{(1)}$ equal to the number of the interior edges of the mesh.

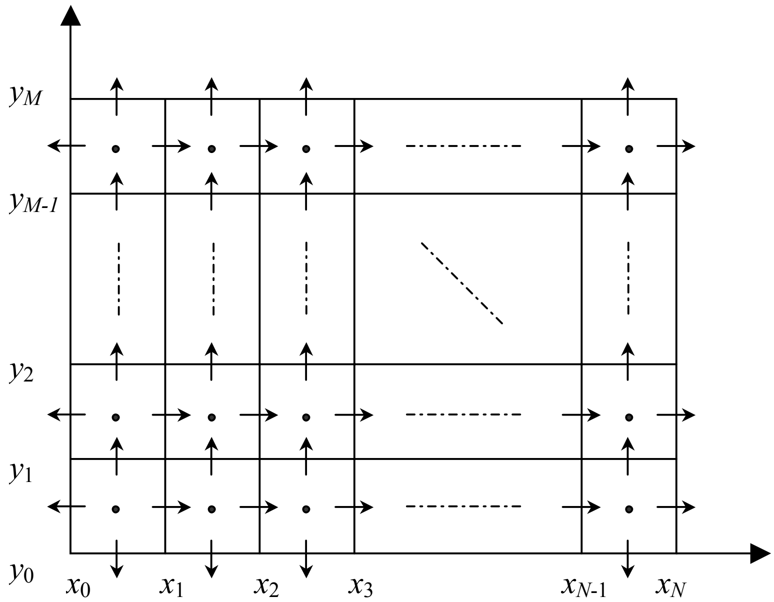

Figure 1. *Mesh and normal vectors.*

The equations of the corresponding semidiscrete variational mixed problem, using the base functions described before, are

$$\sum_{i=0}^{L-1} \frac{\partial u_i(t)}{\partial t}(\varphi_i, \varphi_j) - \sum_{l=0}^{E_0-1} p_l(t)\left(\vec{\nabla}\cdot\vec{\phi}_l, \varphi_j\right) = \left(f(u_h, \beta_h), \varphi_j\right) \quad j = 0,\ldots,L \tag{160}$$

$$\sum_{l=0}^{E_0-1} p_l(t)\left(\left(K(u_h)\right)^{-1}\vec{\phi}_l, \vec{\phi}_m\right) + \sum_{i=0}^{L-1} u_i(t)\left(\varphi_i, \vec{\nabla}\cdot\vec{\phi}_m\right) = 0 \quad m = 0,\ldots,E_0 - 1 \tag{161}$$

$$\sum_{i=0}^{L-1} \frac{d\beta_i(t)}{dt}(\varphi_i, \varphi_j) = \left(g(u_h, \beta_h), \varphi_j\right) \quad j = 0,\ldots,L-1 \tag{162}$$

where $u_i(t)$, $\beta_i(t)$ and $p_i(t)$ are the coordinates of the approximate solution $u_h(t)$, $\beta_h(t)$ and $\vec{p}_h(t)$ on the base functions described. Initial conditions $u_{I,i}$ and $\beta_{I,i}$ for i=0,..., L-1, are the coordinates of $u_h(0) = u_{I,h}$ and $\beta_h(0) = \beta_{I,h}$ on the base functions chosen, where $u_{I,h}$ and $\beta_{I,h}$ are an approximation of u_I and β_I on $L_h^{(0)}$. Notice that $u_h(0)$ determines $\vec{p}_h(0)$ by equation (161).

We define by $\mathsf{U}(t) = (u_0(t),\ldots,u_{L-1}(t))^t$, $\beta(t) = (\beta_0(t),\ldots,\beta_{L-1}(t))^t$ and $\mathsf{P}(t) = (p_0(t),\ldots,p_{E_0-1}(t))^t$ the vector coordinates,. For $i,j = 0,\ldots,L-1$ y

$l, m = 0, \ldots, E_0 - 1$, the system matrices are

$$
\begin{aligned}
&\mathsf{I}_{i,j} = (\varphi_i, \varphi_j) = \int_\Omega \varphi_i(\vec{x})\varphi_j(\vec{x})\,d\vec{x} \\
&\mathsf{B}_{l,j} = (\vec{\nabla}\cdot\vec{\phi}_l, \varphi_j) = \int_\Omega \vec{\nabla}\cdot\vec{\phi}_l(\vec{x})\varphi_j(\vec{x})\,d\vec{x} \\
&\mathsf{A}(\mathsf{U}(t))_{l,m} = \left((K(u_h))^{-1}\vec{\phi}_l, \vec{\phi}_m\right) = \int_\Omega (K(u_h))^{-1}\vec{\phi}_l(\vec{x})\vec{\phi}_m(\vec{x})\,d\vec{x}
\end{aligned}
\tag{163}
$$

and the vectors,

$$
\begin{aligned}
&\mathsf{F}(\mathsf{U}(t), \beta(t))_j = (f(u_h, \beta_h), \varphi_j) = \int_\Omega f(u_h, \beta_h)\varphi_j(\vec{x})\,d\vec{x} \\
&\mathsf{G}(\mathsf{U}(t), \beta(t))_j = (g(u_h, \beta_h), \varphi_j) = \int_\Omega g(u_h, \beta_h)\varphi_j(\vec{x})\,d\vec{x}
\end{aligned}
\tag{164}
$$

With this notation, the problem to solve is

$$
\begin{aligned}
&\mathsf{I}\cdot\frac{\partial}{\partial t}\mathsf{U}(t) - \mathsf{B}\cdot\mathsf{P}(t) = \mathsf{F}(\mathsf{U}(t), \beta(t)) \\
&\mathsf{A}(\mathsf{U}(t))\cdot\mathsf{P}(t) + \mathsf{B}^t\cdot\mathsf{U}(t) = 0 \\
&\mathsf{I}\cdot\frac{d}{dt}\beta(t) = \mathsf{G}(\mathsf{U}(t), \beta(t))
\end{aligned}
\tag{165}
$$

Before the application of time discretization, we analyse some of the terms of the equations, taking into account the space discretization chosen.

$$
(\varphi_i, \varphi_j) = \int_\Omega \varphi_i(\vec{x})\varphi_j(\vec{x})\,d\vec{x} = \sum_{k=0}^{L-1}\int_{\tau_k}\varphi_i(\vec{x})\varphi_j(\vec{x})\,d\vec{x} = area(\tau_i)\delta_{i,j} \tag{166}
$$

Since u_h and β_h are constants over each rectangle of the mesh, we have, for both reactive terms, f and g,

$$
\begin{aligned}
(f(u_h, \beta_h), \varphi_j) &= \int_\Omega f(u_h, \beta_h)\varphi_j(\vec{x})\,d\vec{x} = \sum_{k=0}^{L-1}\int_\Omega f(u_h, \beta_h)\varphi_j(\vec{x})\,d\vec{x} = \\
&= \sum_{k=0}^{L-1}\int_{\tau_j} f(u_h, \beta_h)\,d\vec{x} = area(\tau_j) f(u_j, \beta_j)
\end{aligned}
\tag{167}
$$

The equations for the semi-implicit scheme are

$$
\begin{aligned}
&\sum_{i=0}^{L-1}\overline{\partial}_t u_i^n(\varphi_i, \varphi_j) - \sum_{l=0}^{E_0-1} p_l^{n-1}(\vec{\nabla}\cdot\vec{\phi}_l, \varphi_j) = (f(u_h^n, \beta_h^n), \varphi_j) && j = 0, \ldots, L-1 \\
&\sum_{l=0}^{E_0-1} p_l^n\left((K(u_h^n))^{-1}\vec{\phi}_l, \vec{\phi}_m\right) + \sum_{i=0}^{L-1} u_i^n(\varphi_i, \vec{\nabla}\cdot\vec{\phi}_m) = 0 && m = 0, \ldots, E_0-1 \\
&\sum_{i=0}^{L-1}\overline{\partial}_t \beta_i^n(\varphi_i, \varphi_j) = (g(u_h^n, \beta_h^n), \varphi_j) && j = 0, \ldots, L-1
\end{aligned}
\tag{168}
$$

Thus we have to solve a linear system with matrix coefficients depending on time via the temperature, but the nonlinearity of the reactive term remains uncoupled, so we have to solve for each time step and over each rectangle of the mesh the corresponding nonlinear equation for the temperature. Notice that the equation of the mass fraction of solid fuel is linear in this variable. The mixed finite element method allows the uncoupling of the nonlinearity of the reactive term and adapting the algorithm for parallel computers.

If we recover the two ordinary differential equations of nondimensional temperature and mass fraction of solid fuel over each rectangle τ_k of the mesh and each time interval $[t_{n-1}, t_n]$,

$$\begin{aligned}&\frac{du_k(t)}{dt}-\frac{1}{area(\tau)}(\mathrm{BP}^n)_k+\alpha u_k(t)=\beta_k(t)e(u_k(t))\\&\frac{d\beta_k(t)}{dt}=-\frac{\varepsilon}{q}\beta_k(t)e(u_k(t))\end{aligned}\tag{169}$$

given $u_k(t_{n-1})=u_k^{n-1}$ and $\beta_k(t_{n-1})=\beta_k^{n-1}$, we want to calculate $u_k(t_n)$ and $\beta_k(t_n)$ verifying (169). To solve this system of ordinary differential equations we can use an adaptive method, that is, beginning with an initial time step $\Delta t = t_n - t_{n-1}$ and reducing it only in the necessary elements.

The equations to solve for the explicit scheme are

$$\begin{aligned}&\sum_{i=0}^{L-1}\overline{\partial}_t\, u_i^n(\varphi_i,\varphi_j)-\sum_{l=0}^{E_0-1}p_l^{n-1}(\vec{\nabla}\cdot\vec{\phi}_l,\varphi_j)=\left(f(u_h^{n-1},\beta_h^{n-1}),\varphi_j\right) && j=0,\ldots,L-1\\&\sum_{l=0}^{E_0-1}p_l^n\left((K(u_h^n))^{-1}\vec{\phi}_l,\vec{\phi}_m\right)+\sum_{i=0}^{L-1}u_i^n(\varphi_i,\vec{\nabla}\cdot\vec{\phi}_m)=0 && m=0,\ldots,E_0-1\\&\sum_{i=0}^{L-1}\overline{\partial}_t\,\beta_i^n(\varphi_i,\varphi_j)=\left(g(u_h^n,\beta_h^n),\varphi_j\right) && j=0,\ldots,L-1\end{aligned}\tag{170}$$

In order to improve the stability of the scheme, we can maintain implicit all the linear terms of *f* and *g*. For each time step $n\geq 1$, this scheme involves:

1º) to solve the following linear system of dimension E_0,

$$\mathrm{A}(\mathrm{U}^{n-1})\mathrm{P}^{n-1}=-\mathrm{B}^t\mathrm{U}^{n-1}\tag{171}$$

2º) to calculate for each $j=0,\ldots,L-1$,

$$u_j^n=\frac{1}{1+\Delta t\alpha}\left(u_j^{n-1}+\frac{\Delta t}{area(\tau)}(\mathrm{BP}^{n-1})_j+\Delta t\,\beta_j^{n-1}e(u_j^{n-1})\right)\tag{172}$$

3º) to calculate for each $j=0,\ldots,L-1$,

$$\beta_j^n=\frac{\beta_j^{n-1}}{1+\Delta t\frac{\varepsilon}{q}e(u_j^n)}\tag{173}$$

The results of figure 2 correspond to a square domain of side 100, and a regular mesh of 1600 elements, so the size of the linear system given by (171) is 3120. The time step in this example is 10^{-4}. The value of the parameters justified by their physical meaning are $\varepsilon=0.03$, $q=2.5$, $\alpha=5$, $\kappa=2.5$, $u_{ref}=2$. The initial nondimensional mass fraction of solid fuel has a random distribution, black colour means no fuel and clear green means maximum of fuel. We assume one initial focus of high temperature at the centre of the domain. The fire front spread has the expected form, depending on the density of the fuel. The nonlinear diffusive term in the temperature equation represents the heat transfer by radiation and allows the fire front to pass through areas with no fuel.

6. CONCLUSIONS

Different numerical schemes for solving nonlinear reaction-diffusion problems have been presented. A mixed finite element technique allows to uncouple the nonlinearity of the reactive term and this makes the adapting of the algorithms for parallel computers possible. Total error estimates for three different temporal schemes and Raviart-Thomas finite elements have been obtained.

As an application we have presented some results from the mathematical model of natural combustion developed by the authors.

REFERENCES

1. Fife, P.C. (1979): *Mathematical aspects of reacting and diffusion systems*. Lecture Notes in Biomathematics, 28. Springer-Verlag, New York.
2. Huyakorn, P.S. and Pinder G.F. (1983): *Computational methods in subsurface flow*. Academic Press, New York.
3. Ruuth, S.J. (1995):"Implicit-explicit methods for reaction-diffusion problems in patern formation". *Journal of mathematical Biology*, vol. 34, pp. 148-176.
4. Yifeng Wang, Nahon, D. and Merino, E. (1994): "Dynamic model of the genesis of calcatres replacing silicate rocks in semi-arid regions". *Geochimica et Cosmochimica Acta*, Vol. 58, No. 23, pp. 5131-5145.
5. Yutiang Wang, Yifeng Wang and Merino, E. (1995): Dynamic weathering model: Constrains required by coupled dissolution and pseudomorphic replacement. *Geochimica et Cosmochimica Acta*, Vol. 59, No. 8, pp. 1559-1570.
6. Murray, J. (1993):*Mathematical biology*, 2nd edition. Springer-Verlag, New York.
7. Montenegro, R., Plaza, A., Ferragut, L. and Asensio, M.I. (1997): "Application of a nonlinear evolution model to fire propagation". *Nonlinear Analysis, Theory, methods & Applications*, vol. 30, No. 5, pp. 2873-2882.
8. Asensio, M.I. and Ferragut, L. (2000): "On a wildland fire model with radiation". *International Journal for Numerical Methods in Engineering*, (in review).
9. Raviart, P.A. and Thomas, J.M. (1977): A mixed finite element method for second order elliptic problems. Mathematical Aspects of Finite Element Methods, I. Galliani and E. Magenes, eds. Lecture Notes in Math. 606, Springer-Verlag, berlin, pp. 292-315.
10. Roberts, J.E. and Thomas, J.-M., (1991): *Mixed and Hybrid Methods*. Handbook of Numerical Analysis, Vol. II, Finite Element Methods (Part 1). Ed. P.G. Ciarlet and J.L. Lions. Elsevier Science Publishers B.V. (North- Holland)
11. Thomée, V. (1984): *Galerkin Finite Element Methods for Parabolic Problems*. Lecture Notes in Methematics, 1054. Springer-Verlag.
12. Brezzi, F. and Fortin, M. (1991): *Mixed and Hybrid Finite Element Methods*. Springer Series in Computational Mathematics, 15. Springer-Verlag.
13. Cox, G. (1995): *Combustion fundamentals of fire*. Academic Press.
14. Williams, F.A.(1985):*Combustion Theory*.2nd ed.Benjamin Cummings,Menlo Park.
15. Zeldovich, Y.B., Barenblatt, G.I., Librovich, V.B. and Makhviladze, G.M. (1985): *The Mathematical Theory of Combustion and Explosions*. Consultants Boureau.
16. Bebernes, J. and Eberly, D. (1989): *Mathematical Problems from Combustion Theory*. Applied Mathematical Science, vol. 83. Springer-Verlag.

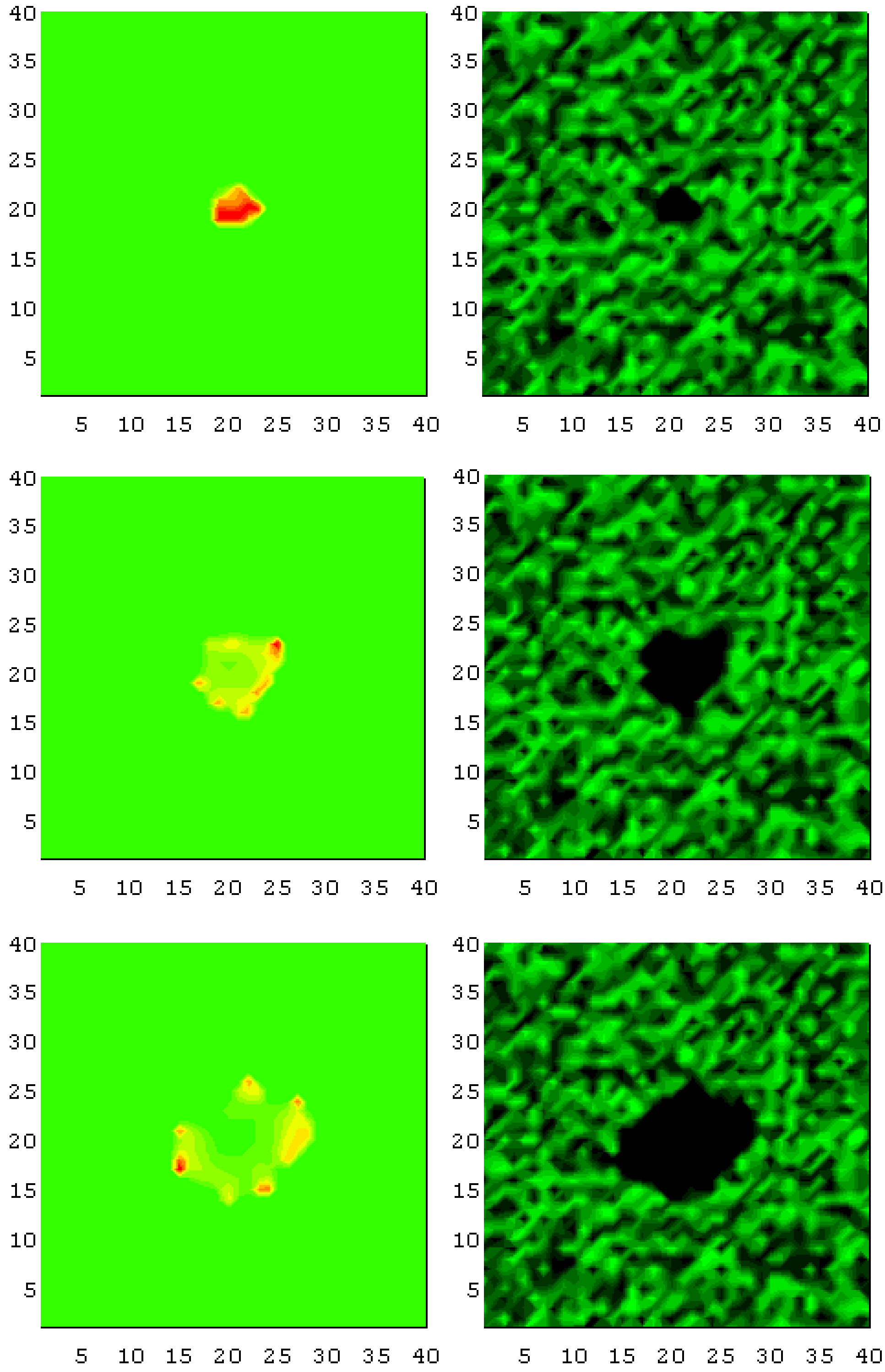

Figure 2. Nondimensional temperature and nondimensional mass fraction of solid fuel at t=0.1, 0.3 and 0.5.