A priori error estimates of a finite element method for an isothermal phase-field model related to the solidification process of a binary alloy

D. KESSLER[1] AND J.-F. SCHEID[2]

[1] *Département de Mathématiques, Ecole Polytechnique Fédérale, 1015 Lausanne, Switzerland,*
daniel.kessler@epfl.ch

[2] *Université Henri Poincaré Nancy 1, Institut Elie Cartan, B.P. 239, 54506 Vandoeuvre-les-Nancy Cédex, France,*
jean-francois.scheid@iecn.u-nancy.fr

Abstract

We introduce a piecewise linear finite element scheme with semi-implicit time discretization for an evolutionary phase-field system modelling the isothermal solidification process of a binary alloy. This system can be written in a vectorial form as a non-linear parabolic system. The convergence of the scheme with error estimate is then proved by introducing a generalized vectorial elliptic projector.

1. Introduction

The phase-field model we consider describes the isothermal solidification process of a binary alloy. It involves the relative concentration c of one component with respect to the mixture and an order parameter ϕ called the phase-field, which accounts for the solidification state of the alloy by taking values between 0 (in a pure solid phase) and 1 (in a pure liquid phase). The model we study is very similar to the Warren and Boettinger model, see Warren and Boettinger (1995). It has been already introduced in former publications (see e.g. Kessler (2001); Kessler, Krüger, Rappaz and Scheid (2000); Kessler, Krüger and Scheid (1998); Rappaz and Scheid (2000)) and has successfully been used to simulate dendritic growth, see Krüger, Picasso and Scheid (2001). A recapitulation of the modelling is out of the scope of this paper, and we will immediatly introduce the mathematical problem. Let Ω be an open subset of $\mathbb{R}^2$ with smooth boundary $\partial\Omega$ and a unit normal vector ν. For $T > 0$, the time evolution of $\phi = \phi(x, t)$ and

$c = c(x,t)$ for $x \in \Omega$ and $t \in [0,T]$, is governed by the following equations :

$$(P)\left\{\begin{array}{ll} \dfrac{\partial \phi}{\partial t} = M\Delta\phi + F_1(\phi) + cF_2(\phi) & \text{in } \Omega\times(0,T), \\ \dfrac{\partial c}{\partial t} = \operatorname{div}\Big(D_1(\phi)\nabla c + D_2(c,\phi)\nabla\phi\Big) & \text{in } \Omega\times(0,T), \\ \dfrac{\partial \phi}{\partial \nu} = \dfrac{\partial c}{\partial \nu} = 0 & \text{on } \partial\Omega\times(0,T), \\ \phi(0)=\phi_0, \quad c(0)=c_0 & \text{in } \Omega, \end{array}\right.$$

where M is a positive constant. The nonlinear functions F_i, D_i, $i = 1,2$ in Problem (P) satisfy the following assumptions :

$(\mathcal{H})$
- $\cdot$ F_1, F_2 and D_1, D_2 are Lipschitz and bounded functions.
- $\cdot$ D_1 is a positive function bounded below by a positive constant D_s.

The aim of this paper is to analyse a numerical scheme for (P) based on a finite element space discretization and a semi-implicit time scheme. For technical reasons, we restrict ourselves to the bidimensional space case. We obtain optimal error estimates for the scheme we consider and we prove that the scheme is unconditionaly convergent. Error estimates for finite element methods have been performed in the past for thermal pure element phase-field models (see Chen and Hoffmann (1994)), but the nonlinearities in such models are different from those of the solutal model we consider in this paper. We also mention Barrett and Blowey (1998) where an isothermal phase separation model is described by a coupled system of Cahn-Hilliard equations.

In order to study the convergence of the numerical scheme we first rewrite Problem (P) in a convenient vectorial form. Problem (P) can be read as a uniformly parabolic system for the auxiliary vectorial variable $\vec{u} = (\phi, \alpha c)$ where the positive parameter α has to be chosen small enough. The idea to get optimal error estimates is mainly based on the introduction of a generalized vectorial projector related to the vectorial form of Problem (P).

In section 2, we introduce the vectorial form of (P) and we specify the mathematical framework. The numerical scheme we study is stated in section 3. The main result of this paper is established in section 4 where an optimal error estimate is derived. To this end, we first define in subsection 4.1 a generalized vectorial elliptic projector for which some error bounds are obtained. This projector will be useful in the next subsection 4.2 to prove the convergence result. Finally, we present in section 5 some numerical results that confirm the theoritical prediction.

2. Mathematical problem in vectorial form

We transform Problem (P) to a vectorial form by defining the vectorial variable $\vec{u} = (\phi, \alpha c)^T$ where α is an arbitrary positive parameter that will be fixed later.

Then Problem (P) reads as a vectorial problem of the form : Find $\vec{u}(x,t) \in \mathbb{R}^2$ such that

$$(P_V) \begin{cases} \dfrac{\partial \vec{u}}{\partial t} = \operatorname{div}\left(D(\vec{u})\nabla\vec{u}\right) + \vec{F}(\vec{u}) & \text{in} \quad \Omega \times (0,T), \qquad (1) \\ \dfrac{\partial \vec{u}}{\partial \nu} = 0 & \text{on} \quad \partial\Omega \times (0,T), \qquad (2) \\ \vec{u}(0) = \vec{u}_0 & \text{in} \quad \Omega, \qquad (3) \end{cases}$$

where the 2×2 triangular matrix D is given by :

$$D(\vec{u}) = \begin{pmatrix} M & 0 \\ \alpha D_2(c,\phi) & D_1(\phi) \end{pmatrix} \quad \text{and} \quad \vec{F}(\vec{u}) = \begin{pmatrix} F_1(\phi) + cF_2(\phi) \\ 0 \end{pmatrix}$$

and where $\operatorname{div}\left(D(\vec{u})\nabla\vec{u}\right) := \displaystyle\sum_{i=1,2} \frac{\partial}{\partial x_i}\left(D(\vec{u})\,\frac{\partial}{\partial x_i}\,\vec{u}\right)$.

Assumption $(\mathcal{H})$ related to Problem (P) leads to the following assumptions for the vectorial problem (P_V) :

(A1) $\vec{F}$ is a 2-vector of Lipschitz bounded components. We call $\mathcal{L}_{\vec{F}}$ the maximum of the components' Lipschitz constants.

(A2) D is a 2x2 lower triangular matrix whose coefficients are given by $d_{11} = M > 0$, $d_{12} = 0$, $d_{21} = \alpha D_2(\vec{u})$ and $d_{22} = D_1(\vec{u}) \geqslant D_s > 0$. The functions $D_1(\vec{u})$ and $D_2(\vec{u})$ are Lipschitz bounded functions. We call D_M the maximum of the components' absolute bounds and $\mathcal{L}_D$ the maximum of the components' Lipschitz constants.

Since $M > 0$ and $D_1(\vec{u}) \geqslant D_s > 0$ for all $\vec{u} \in \mathbb{R}^2$, we can choose the parameter α small enough for $D(\vec{u})$ to be definite positive uniformly with $\vec{u}$. Indeed, if we choose $\alpha < 2(MD_s)^{1/2}/\left\|D_2\right\|_\infty$ where $\left\|D_2\right\|_\infty = \sup_{\vec{u}\in\mathbb{R}^2}|D_2(\vec{u})|$, then it can be shown that $\vec{v}^T D(\vec{u})\vec{v} \geqslant \min(M, D_s)\vec{v}^T\vec{v}$ for all $\vec{u}$, $\vec{v} \in \mathbb{R}^2$. So in addition to (A1) and (A2), it is plainly justified to make an extra assumption on the positiveness of the matrix.

(A3) The matrix $D(\vec{u})$ is definite positive uniformly with $\vec{u}$ i.e. there exists a constant $D_m > 0$ independent of $\vec{u}$ such that $\vec{v}^T D(\vec{u})\vec{v} \geqslant D_m \vec{v}^T\vec{v}$ for all $\vec{u}$, $\vec{v} \in \mathbb{R}^2$.

Remark : As we pointed out previously, assumption (A3) is fulfilled if α is chosen small enough.

Now, we recall some basic properties about vectorial calculus that will be useful later on. First, throughout this article we shall denote by (:) the double scalar product in $\mathbb{R}^2 \otimes \mathbb{R}^2$, such that

$$A\nabla\vec{v} : \nabla\vec{w} = \sum_{i,j=1,2}\ \sum_{\mu,\nu=1,2} A_{\mu\nu}\,\frac{\partial v_\mu}{\partial x_i}\,\frac{\partial w_\nu}{\partial x_j}, \qquad \vec{v}, \vec{w} \in \mathbb{R}^2. \qquad (4)$$

It is then clear that for all matrix $A \in \mathcal{M}_{2\times2}(\mathbb{R})$ with bounded components, we have

$$\int_\Omega A\nabla\vec{v} : \nabla\vec{w} = \int_\Omega A^T\nabla\vec{w} : \nabla\vec{v}, \qquad \text{for all } \vec{v}, \vec{w} \in H^1(\Omega, \mathbb{R}^2). \tag{5}$$

Furthermore, for any matrix $A \in \mathcal{M}_{2\times2}(\mathbb{R})$ with bounded and Lipschitz components, the following Green's formula

$$\int_\Omega A\nabla\vec{v} : \nabla\vec{w} = -\int_\Omega \operatorname{div}(A\nabla\vec{v}).\vec{w} + \int_{\partial\Omega} A\frac{\partial\vec{v}}{\partial\nu}.\vec{w} \tag{6}$$

holds for all $\vec{v} \in H^2(\Omega, \mathbb{R}^2)$ and $\vec{w} \in H^1(\Omega, \mathbb{R}^2)$.

From the analytical point of view, the well-posedness of the original problem (P) has been investigated in Rappaz and Scheid (2000), under assumption $(\mathcal{H})$. These results applied to the vectorial form lead in particular to an existence and uniqueness result for (P_V). Under assumptions (A1)-(A3) and if the initial data $\vec{u}_0 = (u_{0,1}, u_{0,2})^T \in H^2(\Omega) \times H^1(\Omega)$ satisfy $\dfrac{\partial u_{0,1}}{\partial\nu} = 0$ on $\partial\Omega$, then for any $T > 0$, there exists a unique solution $\vec{u} = (u_1, u_2)^T$ of Problem (P_V) such that $u_1 \in C^0([0,T]; H^2(\Omega)) \cap H^1(\Omega \times (0,T))$ and $u_2 \in C^0([0,T]; H^1(\Omega)) \cap H^1(0,T; L^2(\Omega))$.

Finally, let us indicate that the solution $\vec{u}$ satisfies the following variational formulation :

$$(P_\mathcal{V}) \left\{ \begin{array}{ll} \displaystyle\int_\Omega \frac{\partial\vec{u}}{\partial t}.\vec{v} + \int_\Omega D(\vec{u})\nabla\vec{u} : \nabla\vec{v} = \int_\Omega \vec{F}(\vec{u}).\vec{v}, \qquad \forall\vec{v} \in H^1(\Omega, \mathbb{R}^2), \ \text{a.e. } t \in (0,T) & (7) \\ \vec{u}(0) = \vec{u}_0. & (8) \end{array} \right.$$

This variational formulation will be useful for the expression of the numerical scheme in the next section and consequently for the error analysis.

3. Numerical scheme

From now, we shall assume that the domain Ω is a convex polygonal subset of $\mathbb{R}^2$. We approximate Problem (P_V) by a $\mathbb{P}_1$-finite element in space, semi-implicit in time discretization. To begin with, let us introduce some notations. We denote by $\mathcal{T}_h$ a regular triangulation of the domain Ω (see Ciarlet (1978)), where h is the diameter of the biggest triangle in $\mathcal{T}_h$ and we define the space $V_h = \{v_h \in C^0(\overline{\Omega});\quad v_h|_K \in \mathbb{P}_1(K), \forall K \in \mathcal{T}_h\}$ and $V_h^2 = V_h \times V_h$. For a given integer $N \geqslant 1$, we denote by $\tau = T/N$ the time step and by $t^n = n\tau$, the current time for $n = 0, ..., N$. We consider the approximation $\vec{u}_h^n$ of the exact solution $\vec{u}(t_n)$. For the rest of the article, we choose an initial data $\vec{u}_0$ belonging to $H^2(\Omega; \mathbb{R}^2)$ so that it is a continuous function. We denote by r_h the Lagrange interpolation operator on V_h and note that $r_h\vec{u}_0$ is well defined. Based on the variational formulation $(P_{\mathcal{V}})$, we now introduce an approximate problem for $\vec{u}_h^n$:

$$(P_{h,\tau})\begin{cases} \text{For } n=1,\dots,N, \text{ find } \vec{u}_h^n \in V_h^2 \text{ such that for all } \vec{v}_h \in V_h^2, \\ \displaystyle\int_\Omega \frac{\vec{u}_h^n - \vec{u}_h^{n-1}}{\tau}.\vec{v}_h + \int_\Omega D(\vec{w}_h^{\theta n})\nabla\vec{u}_h^n : \nabla\vec{v}_h = \int_\Omega \vec{F}(\vec{u}_h^{n-1}).\vec{v}_h, & (9) \\ \vec{u}_h^0 = r_h\vec{u}_0, & (10) \end{cases}$$

where $\theta \in [0,1]$ and the vector $\vec{w}_h^{\theta n}$ is defined from $\vec{u}_h^n = (u_{1h}^n, u_{2h}^n)^T$ and $\vec{u}_h^{n-1}$ by

$$\vec{w}_h^{\theta n} = \begin{pmatrix} (1-\theta)u_{1h}^{n-1} + \theta u_{1h}^n \\ u_{2h}^{n-1} \end{pmatrix}. \tag{11}$$

It is easy to see that for all $\theta \in [0,1]$, the discrete problem $(P_{h,\tau})$ has a unique solution. This is due to the fact that the matrix $D(\vec{u})$ is lower triangular and that the component $(D(\vec{u}))_{11} = M$ is a constant. So, at first, from $\vec{u}_h^{n-1} = (u_{1h}^{n-1}, u_{2h}^{n-1})^T$ we determine u_{1h}^n by solving the equation (9) with $\vec{v}_h = (v_h, 0)^T$; then since the second component of $\vec{w}_h^{\theta n}$ does not depend on $\vec{u}_h^n$ at all, we determine u_{2h}^n by solving the equation (9) with $\vec{v}_h = (0, v_h)^T$. Also, for any $\theta \in [0,1]$ we do not have to solve nonlinear algebraic equations at each time step, while still granting inconditional convergence, as we will see in the next section. Lastly, note that in the approximate problem $(P_{h,\tau})$ we consider, no numerical integration is taken into account.

4. Convergence result

The following theorem states the main result of this paper, concerning the convergence of the solution $\vec{u}_h^n$ of the discrete problem $(P_{h,\tau})$ to the exact solution $\vec{u}$ of the continuous problem (P_V). We need some extra assumptions on the triangulation of Ω. We assume that

(A4) The triangulation $\mathcal{T}_h$ verifies an inverse assumption i.e. there exists a constant β such that $\forall K \in \mathcal{T}_h$, $\dfrac{h}{h_K} \leqslant \beta$, where h_K stands for the diameter of the triangle K.

THEOREM 4.1: *Let assumptions* (A1),(A2) *and* (A3) *be fulfilled. Suppose that the triangulation $\mathcal{T}_h$ satisfies the inverse assumption* (A4). *If the solution $\vec{u}$ of Problem (P_V) belongs to $H^1\big(0,T;H^2(\Omega,\mathbb{R}^2)\cap W^{1,\infty}(\Omega,\mathbb{R}^2)\big)$, then there exist two positive constants C and τ^* independent of h and τ such that for $0<\tau\leqslant\tau*$,*

$$\max_{0\leqslant n\leqslant N} \|\vec{u}(t^n) - \vec{u}_h^n\|_{L^2(\Omega,\mathbb{R}^2)} \leqslant C(h^2+\tau). \tag{12}$$

The proof of Theorem 4.1 will be given on subsection 4.2. It is based on the introduction of a generalized vectorial elliptic projector, which is defined and studied on the next subsection 4.1.

4.1. A generalized vectorial elliptic projector

We will introduce a vectorial elliptic projector which is a generalization of the scalar elliptic projector used for instance by Thomée (1991). Through this section, we deal with a 2×2 matrix which is not assumed to be triangular nor symmetric. In particular results of section 4.1.2 about the properties of the vectorial projector is valid for a general 2×2 matrix.

Definition: Let $D(x)$ be a 2×2 matrix of bounded functions, positive definite uniformly with $x \in \Omega$. We define the generalized vectorial elliptic projector (GVP)

$$\begin{aligned} \pi_h \quad : \quad H^1(\Omega, \mathbb{R}^2) &\longrightarrow V_h^2 \\ \vec{u} &\longmapsto \pi_h \vec{u} \end{aligned}$$

by the relation $\displaystyle\int_\Omega D\nabla(\vec{u} - \pi_h\vec{u}) : \nabla\vec{v}_h + \int_\Omega (\vec{u} - \pi_h\vec{u}).\vec{v}_h = 0, \quad \forall \vec{v}_h \in V_h.$ (13)

The Lax-Milgram lemma ensures that π_h is well-defined. Note that the second term in the left-hand side of equation (13) is necessary to account for Neumann boundary conditions for $\vec{u}$ in our problem.

4.1.1. Time-dependent GVP

We now consider a time dependent matrix that will depend on both space $x \in \Omega$ and time $t \in [0, T]$, and we define a time dependent generalized vectorial projector. We will assume that :

(H1) $D \in C^0\left([0, T]; L^\infty\left(\Omega, \mathcal{M}_{2\times 2}(\mathbb{R})\right)\right)$.

(H2) D is uniformly positive definite i.e. there exists a constant β independent of x and t such that $\vec{\xi}^T D(x, t)\vec{\xi} \geqslant \beta\, \vec{\xi}^T\vec{\xi}$ for all $\vec{\xi} \in \mathbb{R}^2$ and $x \in \Omega$, $t \in [0, T]$.

We introduce a time dependent bilinear form in $H^1(\Omega, \mathbb{R}^2)$ defined for all $t \in [0, T]$ by

$$\vec{u}, \vec{v} \in H^1(\Omega, \mathbb{R}^2) \longmapsto a_t(\vec{u}, \vec{v}) = \int_\Omega D(t)\nabla\vec{u} : \nabla\vec{v} + \int_\Omega \vec{u}.\vec{v}, \tag{14}$$

Under assumptions (H1)-(H2), it is straightforward that the bilinear form $a_t(.,.)$ is coercive and continuous on $H^1(\Omega, \mathbb{R}^2)$ uniformly with t, *i.e.* one can exhibit coercivity and continuity constants which are independent of t. Lax-Milgram's lemma then allows us to generalize Definition 4.1 as the next definition.

Definition: Under assumptions (H1)-(H2), we define the time-dependent generalized vectorial elliptic projector (GVP)

$$\begin{aligned} \pi_h \quad : \quad C^0\left([0, T]; H^1(\Omega, \mathbb{R}^2)\right) &\longrightarrow L^\infty\left(0, T; V_h^2\right) \\ \vec{u} &\longmapsto \pi_h \vec{u} \end{aligned}$$

by the relation $\quad a_t\left(\vec{u}(t) - \pi_h\vec{u}(t), \vec{v}_h\right) = 0, \quad \forall \vec{v}_h \in V_h^2, \quad \forall t \in [0, T].$ (15)

4.1.2. Properties of the time-dependent GVP

We will now give some properties for the time-dependent GVP. We derive error bounds that will be key ingredients for the a priori estimates on section 4.2 for the proof of Theorem 4.1.

PROPOSITION 4.1: *Under assumptions* (H1)-(H2), *if in addition* $D \in L^\infty(0,T;W^{1,\infty}(\Omega,\mathcal{M}_{2\times 2}(\mathbb{R})))$ *and* $\vec{u} \in C^0([0,T];H^2(\Omega,\mathbb{R}^2))$ *then* $\pi_h\vec{u} \in C^0([0,T];V_h^2)$ *and there exists a positive constant* C *independent of* h, *such that*

$$\|\vec{u}-\pi_h\vec{u}\|_{L^\infty(0,T;L^2(\Omega,\mathbb{R}^2))} + h\,\|\vec{u}-\pi_h\vec{u}\|_{L^\infty(0,T;H^1(\Omega,\mathbb{R}^2))} \leqslant Ch^2. \tag{16}$$

PROPOSITION 4.2: *Under assumptions* (H1)-(H2), *if in addition* $D \in L^\infty(0,T;W^{1,\infty}(\Omega,\mathcal{M}_{2\times 2}(\mathbb{R}))) \cap H^1(0,T;L^\infty(\Omega,\mathcal{M}_{2\times 2}(\mathbb{R})))$ *and* $\vec{u} \in H^1(0,T;H^2(\Omega,\mathbb{R}^2))$ *then* $\pi_h\vec{u} \in H^1(0,T;V_h^2)$ *and there exists a positive constant* C *independent of* h, *such that*

$$\left\|\frac{\partial}{\partial t}(\vec{u}-\pi_h\vec{u})\right\|_{L^2(0,T;L^2(\Omega,\mathbb{R}^2))} + h\left\|\frac{\partial}{\partial t}(\vec{u}-\pi_h\vec{u})\right\|_{L^2(0,T;H^1(\Omega,\mathbb{R}^2))} \leqslant Ch^2. \tag{17}$$

PROPOSITION 4.3: *Let assumptions* (H1)-(H2) *be fulfilled. Suppose that the triangulation* $\mathcal{T}_h$ *satisfies the inverse assumption* (A4). *If* $\vec{u} \in H^1(0,T;H^2(\Omega,\mathbb{R}^2)) \cap L^\infty(0,T;W^{1,\infty}(\Omega,\mathbb{R}^2))$ *then there exists a positive constant* C *independent of* h, *such that*

$$\|\nabla\pi_h\vec{u}\|_{L^\infty(0,T;L^\infty(\Omega,\mathbb{R}^2))} \leqslant C. \tag{18}$$

Remark : Propositions 4.1 and 4.2 are still valid in space dimension 3. However, Proposition 4.3 is not available in space dimension greater than 2. Indeed, the constant C in (18) depends on $h^{1-d/2}$ where d is the space dimension.

Now we deal with the proofs of the three propositions.

4.1.3. Proof of the properties of the time-dependent GVP

We will need a lemma for prooving the properties of the GVP. This result extends a regularity result from Ladyzhenskaya and Ural'tseva (1968) from scalar elliptic problems to elliptic systems.

LEMMA 4.1: *Let* $A \in W^{1,\infty}(\Omega,\mathcal{M}_{2\times 2}(\mathbb{R}))$ *be a uniformly positive definite matrix and let* $\vec{f} \in L^2(\Omega,\mathbb{R}^2)$. *Then the solution* $\vec{w} \in H^1(\Omega,\mathbb{R}^2)$ *to the equation*

$$\int_\Omega A\nabla\vec{w} : \nabla\vec{v} + \int_\Omega \vec{w}.\vec{v} = \int_\Omega \vec{f}.\vec{v}, \quad \forall \vec{v} \in H^1(\Omega,\mathbb{R}^2), \tag{19}$$

is actually in $H^2(\Omega,\mathbb{R}^2)$ *and satisfies* $\dfrac{\partial\vec{w}}{\partial\nu} = 0$ *a.e. on* $\partial\Omega$. *Furthermore, there exists a constant* C *independent of* $\vec{f}$ *such that*

$$\|\vec{w}\|_{H^2(\Omega,\mathbb{R}^2)} \leqslant C\left\|\vec{f}\right\|_{L^2(\Omega,\mathbb{R}^2)}. \tag{20}$$

Proof of Lemma 4.1

Let $\vec{f} = (f_1, f_2) \in L^2(\Omega) \times L^2(\Omega)$. According to Lax-Milgram's lemma, there exists a unique solution $\vec{w} = (w_1, w_2) \in H^1(\Omega) \times H^1(\Omega)$ to the equation (19). We note $A = (a_{ij})_{1 \leqslant i,j \leqslant 2}$ with $a_{ij} \in W^{1,\infty}(\Omega)$. Since A is uniformly positive definite, there are three positive constants β_i, $i = 1, 2, 3$, such that $0 < \beta_1 \leqslant a_{11} \leqslant \beta_2$ and $a_{11}a_{22} - a_{12}a_{21} \geqslant \beta_3 > 0$. Under the lemma's assumptions, Ladyzhenskaya and Ural'tseva's result (see Ladyzhenskaya and Ural'tseva (1968), p.188) tells us that there exists a unique $\tilde{w}_2 \in H^2(\Omega)$ such that

$$\begin{aligned} -\operatorname{div}\left(\left(a_{22} - \frac{a_{12}}{a_{11}} a_{21}\right) \nabla \tilde{w}_2\right) + \tilde{w}_2 \\ = f_2 - \frac{a_{12}}{a_{11}}(f_1 - w_1) + \nabla\left(\frac{a_{12}}{a_{11}}\right) . (a_{11} \nabla w_1 + a_{21} \nabla w_2) . \end{aligned} \tag{21}$$

For the same reasons, once $\tilde{w}_2 \in H^2(\Omega)$ is given, there exists a unique $\tilde{w}_1 \in H^2(\Omega)$ such that

$$-\operatorname{div}(a_{11} \nabla \tilde{w}_1) + \tilde{w}_1 = f_1 + \operatorname{div}(a_{21} \nabla \tilde{w}_2) . \tag{22}$$

Let $v \in H^1(\Omega)$. If we substract a weak form of (21) with v as a test function, from equation (19) with $\vec{v} = (a_{12}/a_{11}\, v, -v)$, we find that

$$\int_\Omega \left(a_{22} - \frac{a_{12}}{a_{11}} a_{21}\right) \nabla(w_2 - \tilde{w}_2) . \nabla v + (w_2 - \tilde{w}_2) v = 0, \quad \forall v \in H^1(\Omega). \tag{23}$$

If we now substract a weak form of (22) with v as a test function, from equation (19) with $\vec{v} = (v, 0)$, we find that

$$\int_\Omega a_{11} \nabla(w_1 - \tilde{w}_1) . \nabla v + a_{21} \nabla(w_2 - \tilde{w}_2) . \nabla v + (w_1 - \tilde{w}_1) v = 0, \quad \forall v \in H^1(\Omega). \tag{24}$$

By first chosing $v = w_2 - \tilde{w}_2$ in (23) and then $v = w_1 - \tilde{w}_1$ in (24), we conclude that $w_1 \equiv \tilde{w}_1$ and $w_2 \equiv \tilde{w}_2$. Therefore $\vec{w} = (w_1, w_2) \in H^2(\Omega, \mathbb{R}^2)$ and $\vec{w}$ satisfies

$$-\operatorname{div}(A \nabla \vec{w}) + \vec{w} = \vec{f} \qquad \text{a.e. in } \Omega, \tag{25}$$

$$\frac{\partial \vec{w}}{\partial \nu} = 0 \qquad \text{a.e. on } \partial\Omega. \tag{26}$$

Using the assumptions on the matrix A, it follows that there exists a constant $\hat{C} > 0$ such that

$$\| \Delta \vec{w} \|_{L^2(\Omega, \mathbb{R}^2)} \leqslant \hat{C} \left(\| \vec{w} \|_{H^1(\Omega, \mathbb{R}^2)} + \left\| \vec{f} \right\|_{L^2(\Omega, \mathbb{R}^2)} \right) . \tag{27}$$

On the other hand, we obtain from (19) with $\vec{v} = \vec{w}$ that there exists a constant $\bar{C} > 0$ such that

$$\| \vec{w} \|_{H^1(\Omega, \mathbb{R}^2)} \leqslant \bar{C} \left\| \vec{f} \right\|_{L^2(\Omega, \mathbb{R}^2)} \tag{28}$$

Combining (27), (28) and a well-known result from elliptic theory (see Lions and Magenes (1968), chap. 2, Th. 5.1), we find the estimate (20) stated in the lemma. ■

Remark : Note that this proove could be generalized to an elliptic system with more than two unknowns.

Now, we are able to proove the three propositions.

Proof of Proposition 4.1

• We note $r_h : C^0(\Omega, \mathbb{R}^2) \to V_h^2$ the $\mathbb{P}_1$-Lagrange interpolation operator on V_h. It is well known (see e.g. Ciarlet (1978)) that the interpolation error on H^1 norm can be estimated by

$$\|\vec{w} - r_h\vec{w}\,\|_{H^1(\Omega,\mathbb{R}^2)} \leqslant C\,h\,|\vec{w}|_{H^2(\Omega,\mathbb{R}^2)}, \quad \forall \vec{w} \in H^2(\Omega, \mathbb{R}^2), \tag{29}$$

where $|.|_{H^2(\Omega,\mathbb{R}^2)}$ denotes the H^2 semi-norm and C is a positive constant independent of $\vec{w}$ and h.

With the previously introduced notations, we can write that for all $t \in [0, T]$, using first the coercivity of a_t, then Definition 4.2 and finally the continuity of a_t,

$$\begin{aligned}
\beta\,\|\vec{u}(t) - \pi_h\vec{u}(t)\,\|_{H^1}^2 &\leqslant a_t\left(\vec{u}(t) - \pi_h\vec{u}(t), \vec{u}(t) - \pi_h\vec{u}(t)\right) && (30)\\
&\leqslant a_t\left(\vec{u}(t) - \pi_h\vec{u}(t), \vec{u}(t) - r_h\vec{u}(t)\right) && (31)\\
&\leqslant \eta\,\|\vec{u}(t) - \pi_h\vec{u}(t)\,\|_{H^1}\,\|\vec{u}(t) - r_h\vec{u}(t)\,\|_{H^1}, && (32)
\end{aligned}$$

where β and η are positive constants. Using the interpolation error estimate (29), we find that

$$\|\vec{u}(t) - \pi_h\vec{u}(t)\,\|_{L^\infty(0,T;H^1(\Omega,\mathbb{R}^2))} \leqslant C_1 h, \tag{33}$$

where C_1 depends on $\|\vec{u}\,\|_{L^\infty(0,T;H^2(\Omega,\mathbb{R}^2))}$ and is independent of h.

• For the L^2-error estimate, we use Aubin-Nitsche's technique, by introducing the dual problem to the definition of $\pi_h\vec{u}(t)$. We define, for a fixed $t \in [0, T]$, the auxiliary function $\vec{w} \in H^1(\Omega, \mathbb{R}^2)$ as the solution to the adjoint equation :

$$a_t(\vec{v}, \vec{w}) = \int_\Omega \left(\vec{u}(t) - \pi_h\vec{u}(t)\right).\vec{v}, \quad \text{for all } \vec{v} \in H^1(\Omega, \mathbb{R}^2). \tag{34}$$

Once again, Lax-Milgram's lemma ensures that $\vec{w}$ is well-defined. Using assumption (H2), the regularity of D and $\vec{u}$, and Lemma 4.1 with $A = D^T$, we obtain that $\vec{w} \in H^2(\Omega, \mathbb{R}^2)$ and that there exists a constant C_2 independent of $\vec{u}$ and h, such that

$$|\vec{w}|_{H^2(\Omega,\mathbb{R}^2)} \leqslant C_2 \, \| \vec{u} - \pi_h \vec{u} \, \|_{L^2(\Omega,\mathbb{R}^2)} \,. \tag{35}$$

From equation (34), Definition 4.2 and the continuity of a_t, we find that

$$\begin{aligned}
\beta \, \| \vec{u}(t) - \pi_h \vec{u}(t) \, \|_{L^2}^2 &= a_t \left(\vec{u}(t) - \pi_h \vec{u}(t), \vec{w} \right) && (36)\\
&= a_t \left(\vec{u}(t) - \pi_h \vec{u}(t), \vec{w} - r_h \vec{w} \right) && (37)\\
&\leqslant \eta \, \| \vec{u}(t) - \pi_h \vec{u}(t) \, \|_{H^1} \, \| \vec{w} - r_h \vec{w} \, \|_{H^1} && (38)
\end{aligned}$$

Using result (33), interpolation estimate (29) and the dual H^2-bound (35), we find that there exists a positive constant C_3 such that

$$\| \vec{u}(t) - \pi_h \vec{u}(t) \, \|_{L^2(\Omega,\mathbb{R}^2)} \leqslant C_3 h^2, \tag{39}$$

and since this last inequality is valid for any fixed $t \in [0, T]$, we obtain

$$\| \vec{u} - \pi_h \vec{u} \, \|_{L^\infty(0,T;L^2(\Omega,\mathbb{R}^2))} \leqslant C_3 h^2, \tag{40}$$

where C_3 depends on $\| \vec{u} \, \|_{L^\infty(0,T;H^2(\Omega,\mathbb{R}^2))}$ and is independent of h.

• Finally, using the coercivity of $a_t(.,.)$, Definition 4.2 for π_h, Hölder's inequality and the above estimates (33), it can be easily proved under the proposition's assumptions that

$$\lim_{s \to t} \| \pi_h \vec{u}(t) - \pi_h \vec{u}(s) \, \|_{H^1(\Omega,\mathbb{R}^2)} = 0, \quad \forall t \in [0, T], \tag{41}$$

i.e. that $\pi_h \vec{u} \in C^0 \left([0, T]; V_h^2 \right)$. ■

Proof of Proposition 4.2

• We take advantage of the fact that V_h is a finite dimensional space. Let's call $\pi_i(t)$ for $i = 1, \ldots, 2n_h$, the coordinates of $\pi_h \vec{u}(t)$ in a basis of V_h^2 defined by a set of linearly independent elements $\{\vec{\psi}_1, \ldots, \vec{\psi}_{2n_h}\}$, where n_h is the dimension of V_h, i.e.

$$\pi_h \vec{u}(t) = \sum_{i=1}^{2n_h} \pi_i(t) \vec{\psi}_i \tag{42}$$

Definition 4.2 can then be translated as :

$$A(t) \vec{\pi}(t) = \vec{b}(t), \qquad \forall t \in [0, T], \tag{43}$$

where
$$\begin{aligned}
\vec{\pi}(t) &= (\pi_j(t))_{1\leqslant j\leqslant 2n_h}\,,\\
\vec{b}(t) &= \left(b_k(t) = \int_\Omega D(t)\nabla\vec{u}(t) : \nabla\vec{\psi}_k + \int_\Omega \vec{u}(t).\vec{\psi}_k\right)_{1\leqslant k\leqslant 2n_h}, \quad \text{and}\\
\vec{A}(t) &= \left(a_{kj}(t) = \int_\Omega D(t)\nabla\vec{\psi}_j : \nabla\vec{\psi}_k + \int_\Omega \vec{\psi}_j.\vec{\psi}_k\right)_{1\leqslant k,j\leqslant 2n_h}.
\end{aligned}$$

By the existence and unicity of $\vec{\pi}(t)$ (Lax-Milgram's lemma), we know that $A(t)$ is invertible for all $t\in[0,T]$ and we get

$$\vec{\pi}(t) = A^{-1}(t)\vec{b}(t), \qquad \forall t\in[0,T]. \tag{44}$$

Since $H^1(0,T)\subset C^0([0,T])$, and $A(t)$ is invertible for all $t\in[0,T]$, it is clear that $\vec{b}\in H^1(0,T;\mathbb{R}^{2n_h})$ and $A^{-1}\in H^1(0,T;\mathcal{M}_{2n_h\times 2n_h}(\mathbb{R}))$, and therefore $\vec{\pi}\in H^1(0,T;\mathbb{R}^{2n_h})$. Thus we have that $\pi_h\vec{u}\in H^1(0,T;V_h^2)$.

• We now derivate equation (15) with respect to t and we obtain

$$\begin{aligned}
&\int_\Omega D'(t)\nabla\left(\vec{u}(t)-\pi_h\vec{u}(t)\right) : \nabla\vec{v}_h + \int_\Omega D(t)\nabla\frac{\partial}{\partial t}\left(\vec{u}(t)-\pi_h\vec{u}(t)\right) : \nabla\vec{v}_h\\
&+\int_\Omega \frac{\partial}{\partial t}\left(\vec{u}(t)-\pi_h\vec{u}(t)\right).\vec{v}_h \;=\; 0, \qquad \text{for all } \vec{v}_h\in V_h^2, \quad \text{a.e. in } (0,T),
\end{aligned} \tag{45}$$

where $D'(t)$ stands for the matrix of the time derivatives of the components of $D(t)$.

With similar steps as (29)-(33), it is then easy to prove that

$$\left\|\frac{\partial}{\partial t}\left(\vec{u}-\pi_h\vec{u}\right)\right\|_{L^2(0,T;H^1(\Omega,\mathbb{R}^2))} \leqslant C_4 h, \tag{46}$$

where C_4 depends on $\|\vec{u}\|_{L^2(0,T;H^2(\Omega,\mathbb{R}^2))}$, $\left\|\frac{\partial\vec{u}}{\partial t}\right\|_{L^2(0,T;H^2(\Omega,\mathbb{R}^2))}$ and $\|D'\|_{L^2(0,T;L^\infty(\Omega,\mathbb{R}^2))}$ but is independent of h.

• In order to get L^2-error estimate, again we use Aubin-Nitsche's technique. This time we define $\vec{w}(t)\in H^1(\Omega,\mathbb{R}^2)$ as the solution to the adjoint equation :

$$a_t(\vec{v},\vec{w}(t)) = \int_\Omega \frac{\partial}{\partial t}\left(\vec{u}(t)-\pi_h\vec{u}(t)\right).\vec{v}, \quad \forall\vec{v}\in H^1(\Omega,\mathbb{R}^2), \quad \text{a.e. } t\in(0,T). \tag{47}$$

Thus $\vec{w}(t)$ is well-defined a.e. in $(0,T)$ and applying Lemma 4.1, we find that $\vec{w}(t)\in H^2(\Omega,\mathbb{R}^2)$ and $\dfrac{\partial\vec{w}(t)}{\partial\nu}=0$ a.e. on $\partial\Omega$, for a.e. $t\in(0,T)$ and that there exists a constant $C_5>0$ such that,

$$\|\vec{w}(t)\|_{H^2(\Omega,\mathbb{R}^2)} \leqslant C_5\left\|\frac{\partial}{\partial t}\left(\vec{u}(t)-\pi_h\vec{u}(t)\right)\right\|_{L^2(\Omega,\mathbb{R}^2)}, \quad \text{a.e. } t\in(0,T). \tag{48}$$

Using (47) and (45), we find that, a.e. in $(0, T)$,

$$\begin{aligned} \left\| \frac{\partial}{\partial t} \left(\vec{u}(t) - \pi_h \vec{u}(t)\right) \right\|^2_{L^2(\Omega,\mathbb{R}^2)} = \; & a_t \left(\frac{\partial}{\partial t} \left(\vec{u}(t) - \pi_h \vec{u}(t)\right), \vec{w}(t) - r_h \vec{w}(t) \right) \\ & + \int_\Omega D'(t) \nabla(\vec{u}(t) - \pi_h \vec{u}(t)) : \nabla(\vec{w}(t) - r_h \vec{w}(t)) \\ & - \int_\Omega D'(t) \nabla(\vec{u}(t) - \pi_h \vec{u}(t)) : \nabla \vec{w}(t). \end{aligned} \tag{49}$$

Applying Green's formula (6) with property (5) to the last term of the right-hand side, and using the continuity of a_t, we find that, a.e. in $(0, T)$,

$$\begin{aligned} \left\| \frac{\partial}{\partial t} \left(\vec{u}(t) - \pi_h \vec{u}(t)\right) \right\|^2_{L^2(\Omega,\mathbb{R}^2)} \leqslant \; & \| \vec{w}(t) - r_h \vec{w}(t) \|_{H^1(\Omega,\mathbb{R}^2)} \left(\eta \left\| \frac{\partial}{\partial t} \left(\vec{u}(t) - \pi_h \vec{u}(t)\right) \right\|_{H^1(\Omega,\mathbb{R}^2)} \right. \\ & \left. + \| D'(t) \|_{L^\infty(\Omega,\mathcal{M}_{2\times 2})} \| \vec{u}(t) - \pi_h \vec{u}(t) \|_{H^1(\Omega,\mathbb{R}^2)} \right) \\ & + \int_\Omega \left(\nabla D'(t)^T \nabla \vec{w} + D'(t)^T \Delta \vec{w} \right) . \left(\vec{u}(t) - \pi_h \vec{u}(t)\right) . \end{aligned} \tag{50}$$

Integrating on time and using Cauchy-Schwartz's inequality, we find that

$$\begin{aligned} & \left\| \frac{\partial}{\partial t} \left(\vec{u} - \pi_h \vec{u}\right) \right\|^2_{L^2(0,T;L^2(\Omega,\mathbb{R}^2))} \\ & \quad \leqslant \| \vec{w} - r_h \vec{w} \|_{L^2(0,T;H^1(\Omega,\mathbb{R}^2))} \left(\eta \left\| \frac{\partial}{\partial t} \left(\vec{u} - \pi_h \vec{u}\right) \right\|_{L^2(0,T;H^1(\Omega,\mathbb{R}^2))} \right. \\ & \quad \left. + \; \| D' \|_{L^2(0,T;L^\infty(\Omega,\mathcal{M}_{2\times 2}))} \| \vec{u} - \pi_h \vec{u} \|_{L^\infty(0,T;H^1(\Omega,\mathbb{R}^2))} \right) \\ & \quad + 2 \| D' \|_{L^2(0,T;W^{1,\infty}(\Omega,\mathcal{M}_{2\times 2}))} \| \vec{w} \|_{L^2(0,T;H^2(\Omega,\mathbb{R}^2))} \| \vec{u} - \pi_h \vec{u} \|_{L^\infty(0,T;L^2(\Omega,\mathbb{R}^2))} . \end{aligned} \tag{51}$$

Using then (46) and Proposition 4.1, we obtain that there exists a positive constant C_6 which depends on $\| D \|_{H^1(0,T;W^{1,\infty}(\Omega,\mathcal{M}_{2\times 2}))}$ but independent of h such that

$$\left\| \frac{\partial}{\partial t} \left(\vec{u} - \pi_h \vec{u}\right) \right\|^2_{L^2(0,T;L^2(\Omega,\mathbb{R}^2))} \leqslant C_6 \left(h \| \vec{w} - r_h \vec{w} \|_{L^2(0,T;H^1(\Omega,\mathbb{R}^2))} + h^2 \| \vec{w} \|_{L^2(0,T;H^2(\Omega,\mathbb{R}^2))} \right) . \tag{52}$$

From interpolation estimate (29) and the H^2-norm estimate (48) together with (52), we conclude that there exists a positive constant C_7 independent of h such that

$$\left\| \frac{\partial}{\partial t} \left(\vec{u} - \pi_h \vec{u}\right) \right\|_{L^2(0,T;L^2(\Omega,\mathbb{R}^2))} \leqslant C_7 \, h^2. \tag{53}$$

■

Proof of Proposition 4.3:

Using assumption `(A4)`, we can write the following inverse inequality in V_h (see Ciarlet (1978), p.140) : there exists a positive constant C independent of h such that

$$\|\nabla v_h\|_{L^\infty(\Omega,\mathbb{R})} \leqslant Ch^{-1}\|\nabla v_h\|_{L^2(\Omega,\mathbb{R})}, \quad \forall v_h \in V_h. \tag{54}$$

Therefore, since $\pi_h\vec{u}(t) - r_h\vec{u}(t) \in V_h^2$, we have for a.e $t \in (0,T)$:

$$\|\nabla(\pi_h\vec{u}(t) - r_h\vec{u}(t))\|_{L^\infty(\Omega,\mathbb{R}^2)} \leqslant Ch^{-1}\|\nabla(\pi_h\vec{u}(t)-r_h\vec{u}(t))\|_{L^2(\Omega,\mathbb{R}^2)} \tag{55}$$

$$\leqslant Ch^{-1}\big(\|\nabla(\pi_h\vec{u}(t) - \vec{u}(t))\|_{L^2(\Omega,\mathbb{R}^2)} + \|\nabla(\vec{u}(t) - r_h\vec{u}(t))\|_{L^2(\Omega,\mathbb{R}^2)}\big). \tag{56}$$

Then using interpolation estimate (29) and (33), we infer that for a.e. $t \in (0,T)$

$$\|\nabla(\pi_h\vec{u}(t) - r_h\vec{u}(t))\|_{L^\infty(\Omega,\mathbb{R}^2)} \leqslant C_8, \tag{57}$$

where C_8 is independent of h, and depends on $\|\vec{u}\|_{L^\infty(0,T;H^2(\Omega,\mathbb{R}^2))}$.

On the other hand, we can estimate a $W^{1,\infty}$-interpolation error for $\vec{u}$. For the Lagrange interpolation operator, we have that (see Ciarlet (1978), p.121) there exists a constant C independent of h and $\vec{u}$ such that, for a.e. $t \in (0,T)$,

$$\|\vec{u}(t) - r_h\vec{u}(t)\|_{W^{1,\infty}(\Omega,\mathbb{R}^2)} \leqslant C\|\vec{u}(t)\|_{W^{1,\infty}(\Omega,\mathbb{R}^2)}, \tag{58}$$

and therefore

$$\begin{aligned}\|\nabla r_h\vec{u}(t)\|_{L^\infty(\Omega,\mathbb{R}^2)} &\leqslant \|\nabla(r_h\vec{u}(t) - \vec{u}(t))\|_{L^\infty(\Omega,\mathbb{R}^2)} + \|\nabla\vec{u}(t)\|_{L^\infty(\Omega,\mathbb{R}^2)} \\ &\leqslant (1+C)\|\vec{u}(t)\|_{W^{1,\infty}(\Omega,\mathbb{R}^2)}.\end{aligned} \tag{59}$$

Finally, using (57) and (59), we find that there exists a constant C_9 independent of h such that for a.e. $t \in (0,T)$

$$\begin{aligned}\|\nabla\pi_h\vec{u}(t)\|_{L^\infty(\Omega,\mathbb{R}^2)} &\leqslant \|\nabla(\pi_h\vec{u}(t) - r_h\vec{u}(t))\|_{L^\infty(\Omega,\mathbb{R}^2)} + \|\nabla r_h\vec{u}(t)\|_{L^\infty(\Omega,\mathbb{R}^2)} \\ &\leqslant C_9.\end{aligned} \tag{60}$$

Proposition 4.3 is then proved. ■

Remark : For the more general case of space dimension $d \leqslant 3$, estimate (54) goes actually as $h^{-d/2}$, and then rather than inequality (60), we get an estimate depending on $h^{1-d/2}$. So for $d = 3$, the constant is not bounded with h.

4.2. Proof of the convergence result

• First of all, let us remark that the GVP assumptions (H1),(H2) as well as the regularity assumptions in Propositions 4.1, 4.2 and 4.3, are implied by the assumptions (A1),(A2),(A3) and the regularity assumption on the exact solution $\vec{u}$ of (P_V) in Theorem 4.1 with $D = D(\vec{u})$. So we can define $\pi_h\vec{u} \in H^1(0,T;V_h^2)$ and use the three GVP properties for $\pi_h\vec{u}$ given by Propositions 4.1, 4.2 and 4.3. Also, from now on, $\|.\|_0$ will denote the norm of $L^2(\Omega,\mathbb{R}^2)$, Q_n the space-time domain $(t^{n-1},t^n)\times\Omega$ and $\overline{g}^n = \frac{1}{\tau}\int_{t^{n-1}}^{t^n} g(t)dt$ the average of an integrable function g on $[t^{n-1},t^n]$. Finally, let us define for $n = 0,\dots N$,

$$\delta\vec{u}_h^n = \pi_h\vec{u}(t^n) - \vec{u}_h^n. \tag{61}$$

From the numerical scheme (9), for all $\vec{v}_h \in V_h^2$ and for $n = 1,\dots,N$, we have :

$$\begin{aligned}\int_\Omega \left(\delta\vec{u}_h^n - \delta\vec{u}_h^{n-1}\right).\vec{v}_h + \tau\int_\Omega D(\vec{w}_h^{\theta n})\nabla\delta\vec{u}_h^n : \nabla\vec{v}_h &= \int_\Omega \left(\pi_h\vec{u}(t^n) - \pi_h\vec{u}(t^{n-1})\right).\vec{v}_h \\ &+ \tau\int_\Omega D(\vec{w}_h^{\theta n})\nabla\pi_h\vec{u}(t^n) : \nabla\vec{v}_h \\ &- \tau\int_\Omega \vec{F}(\vec{u}_h^{n-1}).\vec{v}_h\ .\end{aligned} \tag{62}$$

Furthermore, since both $\vec{u}$ and $\pi_h\vec{u}$ are in $H^1(0,T;L^2(\Omega,\mathbb{R}^2))$, we have

$$\begin{aligned}\int_\Omega \left(\pi_h\vec{u}(t^n) - \pi_h\vec{u}(t^{n-1})\right).\vec{v}_h &= \int_\Omega \left(\pi_h\vec{u}(t^n) - \vec{u}(t^n)\right).\vec{v}_h - \int_\Omega \left(\pi_h\vec{u}(t^{n-1}) - \vec{u}(t^{n-1})\right).\vec{v}_h \\ &+ \int_\Omega \left(\vec{u}(t^n) - \vec{u}(t^{n-1})\right).\vec{v}_h \\ &= \tau\int_\Omega \overline{\frac{\partial}{\partial t}\left(\pi_h\vec{u} - \vec{u}\right)}^n .\vec{v}_h + \tau\int_\Omega \overline{\frac{\partial\vec{u}}{\partial t}}^n .\vec{v}_h\ .\end{aligned} \tag{63}$$

Now, using equation (7) of the exact problem $(P_{\mathcal{V}})$, we deduce that

$$\int_\Omega \overline{\frac{\partial\vec{u}}{\partial t}}^n .\vec{v}_h = \int_\Omega \overline{\vec{F}(\vec{u})}^n .\vec{v}_h - \int_\Omega \overline{D(\vec{u})\nabla\vec{u}}^n : \nabla\vec{v}_h\ . \tag{64}$$

Then from equations (63) and (64) together with equation (62), we obtain

$$\begin{aligned}&\int_\Omega \left(\delta\vec{u}_h^n - \delta\vec{u}_h^{n-1}\right).\vec{v}_h + \tau\int_\Omega D(\vec{w}_h^{\theta n})\nabla\delta\vec{u}_h^n : \nabla\vec{v}_h \\ &\quad= \tau\int_\Omega \overline{\frac{\partial}{\partial t}\left(\pi_h\vec{u} - \vec{u}\right)}^n .\vec{v}_h - \tau\int_\Omega \overline{D(\vec{u})\nabla\vec{u}}^n : \nabla\vec{v}_h \\ &\quad+\tau\int_\Omega \left(\overline{\vec{F}(\vec{u})}^n - \vec{F}(\vec{u}_h^{n-1})\right).\vec{v}_h + \tau\int_\Omega D(\vec{w}_h^{\theta n})\nabla\pi_h\vec{u}(t^n) : \nabla\vec{v}_h\ .\end{aligned} \tag{65}$$

Moreover, by the definition (15) of the GVP, we get for all $\vec{v}_h \in V_h^2$:

$$\int_\Omega \overline{D(\vec{u})\nabla\vec{u}}^n : \nabla\vec{v}_h = \int_\Omega \overline{D(\vec{u})\nabla\pi_h\vec{u}}^n : \nabla\vec{v}_h + \int_\Omega \overline{(\pi_h\vec{u}(t) - \vec{u}(t))}^n .\vec{v}_h . \quad (66)$$

Then using (66) in equation (65), we obtain that :

$$\begin{aligned}
&\int_\Omega \left(\delta\vec{u}_h^n - \delta\vec{u}_h^{n-1}\right).\vec{v}_h + \tau\int_\Omega D(\vec{w}_h^{\theta n})\nabla\delta\vec{u}_h^n : \nabla\vec{v}_h \\
&\quad = \tau\int_\Omega \overline{\frac{\partial}{\partial t}(\pi_h\vec{u} - \vec{u})}^n .\vec{v}_h + \tau\int_\Omega \overline{(\vec{u} - \pi_h\vec{u})}^n .\vec{v}_h \\
&\quad +\tau\int_\Omega \left(D(\vec{w}_h^{\theta n})\nabla\pi_h\vec{u}(t^n) - \overline{D(\vec{u})\nabla\pi_h\vec{u}}^n\right) : \nabla\vec{v}_h \\
&\quad +\tau\int_\Omega \left(\overline{\vec{F}(\vec{u})}^n - \vec{F}(\vec{u}_h^{n-1})\right).\vec{v}_h, \qquad \text{for all } \vec{v}_h \in V_h^2.
\end{aligned} \quad (67)$$

We may now choose $\vec{v}_h = \delta\vec{u}_h^n$ in equation (67). Using assumption `(A3)` and applying Cauchy-Schwartz and Young's inequalities five times to equation (67), we get the following inequality, valid for all $\varepsilon_0, \ldots, \varepsilon_4 > 0$ and for $n = 1, \ldots, N$:

$$\begin{aligned}
&(1 - (\varepsilon_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3)) \left\|\delta\vec{u}_h^n\right\|_0^2 - \frac{1}{4\varepsilon_0}\left\|\delta\vec{u}_h^{n-1}\right\|_0^2 + \tau(D_m - \varepsilon_4)\left\|\nabla\delta\vec{u}_h^n\right\|_0^2 \\
&\quad \leqslant \frac{\tau^2}{4\varepsilon_1}\left\|\overline{\frac{\partial}{\partial t}(\pi_h\vec{u} - \vec{u})}^n\right\|_0^2 + \frac{\tau^2}{4\varepsilon_2}\left\|\overline{\vec{u} - \pi_h\vec{u}}^n\right\|_0^2 + \frac{\tau^2}{4\varepsilon_3}\left\|\overline{\vec{F}(\vec{u})}^n - \vec{F}(\vec{u}_h^{n-1})\right\|_0^2 \\
&\quad +\frac{\tau}{4\varepsilon_4}\left\|D(\vec{w}_h^{\theta n})\nabla\pi_h\vec{u}(t^n) - \overline{D(\vec{u})\nabla\pi_h\vec{u}}^n\right\|_0^2 .
\end{aligned} \quad (68)$$

• We must now estimate the four terms of the right-hand side of inequality (68) before using the discrete Gronwall's lemma to obtain the final estimate of $\|\delta\vec{u}_h^n\|_0^2$.

i) The first term in the right-hand side of (68) can be immediatly estimated using Cauchy-Schwartz's inequality. We get :

$$\begin{aligned}
\left\|\overline{\frac{\partial}{\partial t}(\pi_h\vec{u} - \vec{u})}^n\right\|_0^2 &= \int_\Omega \left|\frac{1}{\tau}\int_{t^{n-1}}^{t^n}\frac{\partial}{\partial t}(\pi_h\vec{u} - \vec{u})dt\right|^2 dx \\
&\leqslant \int_\Omega \frac{1}{\tau}\int_{t^{n-1}}^{t^n}\left|\frac{\partial}{\partial t}(\pi_h\vec{u} - \vec{u})\right|^2 dt\, dx \\
&\leqslant \frac{1}{\tau}\left\|\frac{\partial}{\partial t}(\vec{u} - \pi_h\vec{u})\right\|_{L^2(Q_n)}^2 ,
\end{aligned} \quad (69)$$

where $|.|$ stands for the vectorial norm.

ii) In a similar way, the second term can be estimated as

$$\left\| \overline{\vec{u} - \pi_h \vec{u}}^{\,n} \right\|_0^2 \leqslant \frac{1}{\tau} \left\| \vec{u} - \pi_h \vec{u} \right\|_{L^2(Q_n)}^2 . \tag{70}$$

iii) The third term of (68) can be read as

$$\left\| \overline{\vec{F}(\vec{u})}^{\,n} - \vec{F}(\vec{u}_h^{n-1}) \right\|_0^2 = \int_\Omega \left| \frac{1}{\tau} \int_{t^{n-1}}^{t^n} \left(\vec{F}(\vec{u}(t)) - \vec{F}(\vec{u}_h^{n-1}) \right) dt \right|^2 dx. \tag{71}$$

Then we use Cauchy-Schwartz's inequality and the Lipschitz assumption (A1) on $\vec{F}$ in order to get

$$\begin{aligned} \left\| \overline{\vec{F}(\vec{u})}^{\,n} - \vec{F}(\vec{u}_h^{n-1}) \right\|_0^2 &\leqslant \int_\Omega \frac{1}{\tau} \int_{t^{n-1}}^{t^n} \left| \vec{F}(\vec{u}(t)) - \vec{F}(\vec{u}_h^{n-1}) \right|^2 dt dx \\ &\leqslant \frac{\mathcal{L}_{\vec{F}}^2}{\tau} \int_\Omega \int_{t^{n-1}}^{t^n} \left| \vec{u}(t) - \vec{u}_h^{n-1} \right|^2 dt dx \\ &\leqslant \frac{2\mathcal{L}_{\vec{F}}^2}{\tau} \left(\int_\Omega \int_{t^{n-1}}^{t^n} \left| \vec{u}(t) - \vec{u}(t^{n-1}) \right|^2 + \int_\Omega \int_{t^{n-1}}^{t^n} \left| \vec{u}(t^{n-1}) - \vec{u}_h^{n-1} \right|^2 \right). \end{aligned} \tag{72}$$

Now, since we have the following relation $\vec{u}(t) - \vec{u}(t^{n-1}) = \displaystyle\int_{t^{n-1}}^{t^n} \frac{\partial \vec{u}}{\partial t}(s)\, ds$, for all $t \in [t^{n-1}, t^n]$, it is easy to see that

$$\int_\Omega \int_{t^{n-1}}^{t^n} \left| \vec{u}(t) - \vec{u}(t^{n-1}) \right|^2 dt \leqslant \tau^2 \left\| \frac{\partial \vec{u}}{\partial t} \right\|_{L^2(Q_n)}^2 . \tag{73}$$

Then we deduce from (72) and (73) that

$$\left\| \overline{\vec{F}(\vec{u})}^{\,n} - \vec{F}(\vec{u}_h^{n-1}) \right\|_0^2 \leqslant 2\tau \mathcal{L}_{\vec{F}}^2 \left\| \frac{\partial \vec{u}}{\partial t} \right\|_{L^2(Q_n)}^2 + 2\mathcal{L}_{\vec{F}}^2 \left\| \vec{u}(t^{n-1}) - \vec{u}_h^{n-1} \right\|_0^2 . \tag{74}$$

Finally, introducing the projector $\pi_h \vec{u}(t^{n-1})$ in the above estimate, we obtain

$$\left\| \overline{\vec{F}(\vec{u})}^{\,n} - \vec{F}(\vec{u}_h^{n-1}) \right\|_0^2 \leqslant 2\tau \mathcal{L}_{\vec{F}}^2 \left\| \frac{\partial \vec{u}}{\partial t} \right\|_{L^2(Q_n)}^2 + 4\mathcal{L}_{\vec{F}}^2 \left\| \vec{u}(t^{n-1}) - \pi_h \vec{u}(t^{n-1}) \right\|_0^2 + 4\mathcal{L}_{\vec{F}}^2 \left\| \delta \vec{u}_h^{n-1} \right\|_0^2 . \tag{75}$$

iv) The final term of (68) requires a little bit more work to estimate. Using Cauchy-Schwartz and Young's inequalities, we can separate it in two terms as

$$\begin{aligned} \left\| D(\vec{w}_h^{\theta n}) \nabla \pi_h \vec{u}(t^n) - \overline{D(\vec{u}) \nabla \pi_h \vec{u}}^{\,n} \right\|_0^2 &\leqslant 2 \left\| \left(D(\vec{w}_h^{\theta n}) - \overline{D(\vec{u})}^{\,n} \right) \nabla \pi_h \vec{u}(t^n) \right\|_0^2 \\ &\quad + 2 \left\| \overline{D(\vec{u})}^{\,n} \nabla \pi_h \vec{u}(t^n) - \overline{D(\vec{u}) \nabla \pi_h \vec{u}}^{\,n} \right\|_0^2 . \end{aligned} \tag{76}$$

- We will start by estimating the first right-hand term of (76) :

$$\left\| \left(D(\vec{w}_h^{\theta n}) - \overline{D(\vec{u})}^n \right) \nabla \pi_h \vec{u}(t^n) \right\|_0^2 \leqslant 2 \left\| \nabla \pi_h \vec{u} \right\|^2_{L^\infty(0,T;L^\infty(\Omega,\mathbb{R}^2))} \sum_{i,j=1}^{2} \left\| D_{ij}(\vec{w}_h^{\theta n}) - \overline{D_{ij}(\vec{u})}^n \right\|_0^2, \tag{77}$$

where D_{ij} stands for the components of matrix D.

We now introduce an auxiliary function

$$\vec{w}^\theta(t) = \begin{pmatrix} (1-\theta)u_1(t-\tau) + \theta u_1(t) \\ u_2(t-\tau) \end{pmatrix}, \quad \text{for } t \geqslant \tau. \tag{78}$$

For all combinations of $i, j = 1, 2$ and for $n \geqslant 1$, we have

$$\left\| D_{ij}(\vec{w}_h^{\theta n}) - \overline{D_{ij}(\vec{u})}^n \right\|_0^2 \leqslant 2 \left(\left\| D_{ij}(\vec{w}_h^{\theta n}) - D_{ij}(\vec{w}^\theta(t^n)) \right\|_0^2 + \left\| D_{ij}(\vec{w}^\theta(t^n)) - \overline{D_{ij}(\vec{u})}^n \right\|_0^2 \right). \tag{79}$$

Let us estimate the first term in the right-hand side of (79). By the use of Lipschitz assumption **(A2)** on the matrix D, we have

$$\left\| D_{ij}(\vec{w}_h^{\theta n}) - D_{ij}(\vec{w}^\theta(t^n)) \right\|_0^2 \leqslant \mathcal{L}_D^2 \left\| \vec{w}_h^{\theta n} - \vec{w}^\theta(t^n) \right\|_0^2. \tag{80}$$

Moreover, Definition (78) for $\vec{w}^\theta$ leads to

$$\begin{aligned} \left| \vec{w}_h^{\theta n} - \vec{w}^\theta(t^n) \right|^2 &= \left((1-\theta)(u_{1h}^{n-1} - u_1(t^{n-1}) + \theta(u_{1h}^n - u_1(t^n)) \right)^2 + \left(u_{2h}^{n-1} - u_2(t^{n-1}) \right)^2 \\ &\leqslant 2 \left(\left| \vec{u}_h^{n-1} - \vec{u}(t^{n-1}) \right|^2 + \theta^2 \left| \vec{u}_h^n - \vec{u}(t^n) \right|^2 \right), \end{aligned} \tag{81}$$

from what we deduce that

$$\begin{aligned} \left\| \vec{w}_h^{\theta n} - \vec{w}^\theta(t^n) \right\|_0^2 \leqslant\ & 4 \left(\left\| \delta \vec{u}_h^{n-1} \right\|_0^2 + \left\| \vec{u}(t^{n-1}) - \pi_h \vec{u}(t^{n-1}) \right\|_0^2 \right. \\ & \left. + \theta^2 \left\| \delta \vec{u}_h^n \right\|_0^2 + \theta^2 \left\| \vec{u}(t^n) - \pi_h \vec{u}(t^n) \right\|_0^2 \right). \end{aligned} \tag{82}$$

Thus from (80) and (82), and since $0 \leqslant \theta \leqslant 1$ we obtain

$$\begin{aligned} \left\| D_{ij}(\vec{w}_h^{\theta n}) - D_{ij}(\vec{w}^\theta(t^n)) \right\|_0^2 \leqslant\ & 4\mathcal{L}_D^2 \left(\left\| \delta \vec{u}_h^{n-1} \right\|_0^2 + \theta^2 \left\| \delta \vec{u}_h^n \right\|_0^2 + \left\| \vec{u}(t^{n-1}) - \pi_h \vec{u}(t^{n-1}) \right\|_0^2 \right. \\ & \left. + \left\| \vec{u}(t^n) - \pi_h \vec{u}(t^n) \right\|_0^2 \right). \end{aligned} \tag{83}$$

Now, we estimate the second term of the right-hand side of (79). First we have

$$\left\| D_{ij}(\vec{w}^\theta(t^n)) - \overline{D_{ij}(\vec{u})}^n \right\|_0^2 = \int_\Omega \left(\frac{1}{\tau} \int_{t^{n-1}}^{t^n} \left(D_{ij}(\vec{w}^\theta(t^n)) - D_{ij}(\vec{u}(t)) \right) dt \right)^2 dx. \tag{84}$$

Then using Cauchy-Schwartz's inequality and the Lipschitz assumption (A2) on the matrix D, we get

$$\left\| D_{ij}(\vec{w}^{\theta}(t^n)) - \overline{D_{ij}(\vec{u})}^n \right\|_0^2 \leqslant \frac{\mathcal{L}_D^2}{\tau} \left\| \vec{w}^{\theta}(t^n) - \vec{u} \right\|_{L^2(Q_n)}^2 . \tag{85}$$

Furthermore we have for all $t \in [t^{n-1}, t^n]$

$$\left| \vec{w}^{\theta}(t^n) - \vec{u}(t) \right|^2 = \big((1-\theta) u_1(t^{n-1}) + \theta u_1(t^n) - u_1(t) \big)^2 + \big(u_2(t^{n-1}) - u_2(t) \big)^2 . \tag{86}$$

Thus remarking that $(1-\theta)u_1(t^{n-1})+\theta u_1(t^n)-u_1(t) = (1-\theta)\int_t^{t^{n-1}} \frac{\partial u_1}{\partial t}(s)ds + \theta \int_t^{t^n} \frac{\partial u_1}{\partial t}(s)ds$ and $u_2(t^{n-1}) - u_2(t) = \int_t^{t^{n-1}} \frac{\partial u_2}{\partial t}(s)ds$, we deduce by Cauchy-Schwartz's inequality that, since $0 \leqslant \theta \leqslant 1$,

$$\left| \vec{w}^{\theta}(t^n) - \vec{u}(t) \right|^2 \leqslant \tau \int_{t^{n-1}}^{t^n} \left| \frac{\partial \vec{u}}{\partial t} \right|^2 dt, \quad \text{for all } t \in [t^{n-1}, t^n]. \tag{87}$$

Then from (85) and (87), we obtain

$$\left\| D_{ij}(\vec{w}^{\theta}(t^n)) - \overline{D_{ij}(\vec{u})}^n \right\|_0^2 \leqslant \tau \mathcal{L}_D^2 \left\| \frac{\partial \vec{u}}{\partial t} \right\|_{L^2(Q_n)}^2 . \tag{88}$$

- The second right-hand term of (76) is estimated as follows. We have

$$\left\| \overline{D(\vec{u})}^n \nabla \pi_h \vec{u}(t^n) - \overline{D(\vec{u}) \nabla \pi_h \vec{u}}^n \right\|_0^2 = \int_{\Omega} \left| \frac{1}{\tau} \int_{t^{n-1}}^{t^n} D(\vec{u}(t)) \left(\nabla \pi_h \vec{u}(t^n) - \nabla \pi_h \vec{u}(t) \right) dt \right|^2 dx. \tag{89}$$

By Cauchy-Schwartz's inequality and the boundeness assumption of D in (A2), we obtain that

$$\left\| \overline{D(\vec{u})}^n \nabla \pi_h \vec{u}(t^n) - \overline{D(\vec{u}) \nabla \pi_h \vec{u}}^n \right\|_0^2 \leqslant \frac{D_M^2}{\tau} \left\| \nabla \pi_h \vec{u}(t^n) - \nabla \pi_h \vec{u} \right\|_{L^2(Q_n)}^2 . \tag{90}$$

Since for all $t \in [t^{n-1}, t^n]$ we have $\nabla \pi_h \vec{u}(t^n) - \nabla \pi_h \vec{u}(t) = \int_{t^{n-1}}^{t^n} \nabla \frac{\partial \pi_h \vec{u}}{\partial t}(s)ds$, we deduce, using Cauchy-Schwartz's inequality and estimate (90), that

$$\left\| \overline{D(\vec{u})}^n \nabla \pi_h \vec{u}(t^n) - \overline{D(\vec{u}) \nabla \pi_h \vec{u}}^n \right\|_0^2 \leqslant \tau D_M^2 \left\| \nabla \frac{\partial \pi_h \vec{u}}{\partial t} \right\|_{L^2(Q_n)}^2 \tag{91}$$

- Grouping steps (76)-(91), we find that

$$\begin{aligned}
&\left\| D(\vec{w}_h^{\theta n})\nabla\pi_h\vec{u}(t^n) - \overline{D(\vec{u})\nabla\pi_h\vec{u}}^n \right\|_0^2 \\
&\qquad \leqslant 128\mathcal{L}_D^2 \left\| \nabla\pi_h\vec{u} \right\|^2_{L^\infty(0,T;L^\infty(\Omega,\mathbb{R}^2))} \left(\left\| \vec{u}(t^{n-1}) - \pi_h\vec{u}(t^{n-1}) \right\|_0^2 \right. \\
&\qquad\qquad \left. + \left\| \vec{u}(t^n) - \pi_h\vec{u}(t^n) \right\|_0^2 + \left\| \delta\vec{u}_h^{n-1} \right\|_0^2 + \theta^2 \left\| \delta\vec{u}_h^n \right\|_0^2 \right) \\
&\qquad\quad + 32\tau\mathcal{L}_D^2 \left\| \nabla\pi_h\vec{u} \right\|^2_{L^\infty(0,T;L^\infty(\Omega,\mathbb{R}^2))} \left\| \frac{\partial\vec{u}}{\partial t} \right\|^2_{L^2(Q_n)} \\
&\qquad\quad + 2\tau D_M^2 \left\| \nabla\frac{\partial\pi_h\vec{u}}{\partial t} \right\|^2_{L^2(Q_n)} .
\end{aligned} \tag{92}$$

Let us note

$$K_1 = 128\mathcal{L}_D^2 \left\| \nabla\pi_h\vec{u} \right\|^2_{L^\infty(0,T;L^\infty(\Omega,\mathbb{R}^2))} \quad \text{and} \quad K_2 = \max\left(\frac{K_1}{4}, 2D_M^2 \right). \tag{93}$$

Then we have

$$\begin{aligned}
&\left\| D(\vec{w}_h^{\theta n})\nabla\pi_h\vec{u}(t^n) - \overline{D(\vec{u})\nabla\pi_h\vec{u}}^n \right\|_0^2 \\
&\qquad \leqslant K_1 \left(2\left\| \vec{u} - \pi_h\vec{u} \right\|^2_{L^\infty(0,T;L^2(\Omega,\mathbb{R}^2))} + \left\| \delta\vec{u}_h^{n-1} \right\|_0^2 + \theta^2 \left\| \delta\vec{u}_h^n \right\|_0^2 \right) \\
&\qquad\quad + K_2\tau \left(\left\| \frac{\partial\vec{u}}{\partial t} \right\|^2_{L^2(Q_n)} + \left\| \nabla\frac{\partial\pi_h\vec{u}}{\partial t} \right\|^2_{L^2(Q_n)} \right) .
\end{aligned} \tag{94}$$

• We can now go back to inequality (68). We choose $\varepsilon_0 = \dfrac{1}{2}$, $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = \dfrac{\tau}{3}$ and $\varepsilon_4 = D_m$. Then from estimates (69), (70), (75) together with inequality (68), we get for $n = 1, ..., N$

$$\begin{aligned}
&(\frac{1}{2} - \tau)\left\| \delta\vec{u}_h^n \right\|_0^2 - \frac{1}{2}\left\| \delta\vec{u}_h^{n-1} \right\|_0^2 \\
&\qquad \leqslant \frac{3}{4}\left\| \vec{u} - \pi_h\vec{u} \right\|^2_{H^1(t^{n-1},t^n;L^2(\Omega,\mathbb{R}^2))} + \left(\frac{3}{2}\mathcal{L}_{\vec{F}}^2 + \frac{K_2}{4D_m} \right)\tau^2 \left\| \frac{\partial\vec{u}}{\partial t} \right\|^2_{L^2(Q_n)} \\
&\qquad\quad + \frac{K_2}{4D_m}\tau^2 \left\| \nabla\frac{\partial\pi_h\vec{u}}{\partial t} \right\|^2_{L^2(Q_n)} + \left(3\mathcal{L}_{\vec{F}}^2 + \frac{K_1}{2D_m} \right)\tau \left\| \vec{u} - \pi_h\vec{u} \right\|^2_{L^\infty(0,T;L^2(\Omega,\mathbb{R}^2))} \\
&\qquad\quad + \left(3\mathcal{L}_{\vec{F}}^2 + \frac{K_1}{4D_m} \right)\tau \left\| \delta\vec{u}_h^{n-1} \right\|_0^2 + \theta^2\frac{K_1}{4D_m}\tau \left\| \delta\vec{u}_h^n \right\|_0^2 .
\end{aligned} \tag{95}$$

Thus we have

$$\begin{aligned}(\frac{1}{2}-\mu_1\tau)\left\|\delta\vec{u}_h^n\right\|_0^2-(\frac{1}{2}+\mu_2\tau)\left\|\delta\vec{u}_h^{n-1}\right\|_0^2 \\ \leqslant \frac{3}{4}\left\|\vec{u}-\pi_h\vec{u}\right\|^2_{H^1(t^{n-1},t^n;L^2(\Omega,\mathbb{R}^2))}+K_3\tau^2\left(\left\|\frac{\partial\vec{u}}{\partial t}\right\|^2_{L^2(Q_n)}+\left\|\nabla\frac{\partial\pi_h\vec{u}}{\partial t}\right\|^2_{L^2(Q_n)}\right) \\ +K_4\tau\left\|\vec{u}-\pi_h\vec{u}\right\|^2_{L^\infty(0,T;L^2(\Omega,\mathbb{R}^2))},\end{aligned} \tag{96}$$

where we have put

$$\mu_1=1+\frac{\theta^2K_1}{4D_m},\qquad \mu_2=3\mathcal{L}_{\vec{F}}^2+\frac{K_1}{4D_m} \tag{97}$$

and

$$K_3=\frac{3}{2}\mathcal{L}_{\vec{F}}^2+\frac{K_2}{4D_m},\qquad K_4=3\mathcal{L}_{\vec{F}}^2+\frac{K_1}{2D_m}. \tag{98}$$

Now, let us define

$$\tau^*=\frac{1}{4\mu_1}>0. \tag{99}$$

Remark that τ^* depends on $\left\|\nabla\pi_h\vec{u}\right\|_{L^\infty(0,T;L^\infty(\Omega,\mathbb{R}^2))}$ but thanks to Proposition 4.3 the constant τ^* is independent of h. In that way, for all $0<\tau\leqslant\tau^*$, we have

$$\frac{1}{2}-\mu_1\tau\geqslant\frac{1}{4}. \tag{100}$$

In addition, it is straightforward to prove that for all $0<\tau\leqslant\tau^*$, we have

$$(\frac{1}{2}+\mu_2\tau)\leqslant(\frac{1}{2}-\mu_1\tau)(1+\mu\tau) \tag{101}$$

where

$$\mu=4(\mu_1+\mu_2). \tag{102}$$

Remark also that μ does not depend on h and τ.

Then using (100) and (101), we deduce from (96) that for all $n=1,\dots,N$ and $0<\tau\leqslant\tau^*$,

$$\left\|\delta\vec{u}_h^n\right\|_0^2-(1+\mu\tau)\left\|\delta\vec{u}_h^{n-1}\right\|_0^2\leqslant\lambda_n \tag{103}$$

where

$$\begin{aligned}\lambda_n &= 3\left\|\vec{u}-\pi_h\vec{u}\right\|^2_{H^1(t^{n-1},t^n;L^2(\Omega,\mathbb{R}^2))}+4K_3\tau^2\left(\left\|\frac{\partial\vec{u}}{\partial t}\right\|^2_{L^2(Q_n)}+\left\|\nabla\frac{\partial\pi_h\vec{u}}{\partial t}\right\|^2_{L^2(Q_n)}\right) \\ &\quad +4K_4\tau\left\|\vec{u}-\pi_h\vec{u}\right\|^2_{L^\infty(0,T;L^2(\Omega,\mathbb{R}^2))}.\end{aligned} \tag{104}$$

Now, we sum inequality (103) over n, in order to get

$$\|\delta\vec{u}_h^n\|_0^2 \leqslant \left\|\delta\vec{u}_h^{n-1}\right\|_0^2 + \sum_{k=1}^{n}\lambda_k + \mu\tau\sum_{k=1}^{n}\left\|\delta\vec{u}_h^{k-1}\right\|_0^2, \tag{105}$$

for all $1 \leqslant n \leqslant N$ and $0 < \tau \leqslant \tau^*$. We can then use the discrete Gronwall's lemma (see for instance Quarteroni and Valli (1991), §1.4) on inequality (105) and find that, for $n = 1, \ldots, N$,

$$\|\delta\vec{u}_h^n\|_0^2 \leqslant \left(\|\delta\vec{u}_h^0\|_0^2 + \sum_{k=1}^{n}\lambda_k\right)\exp(\mu T). \tag{106}$$

Furthermore, using the definition (104) of λ_k, we have that for all $1 \leqslant n \leqslant N$

$$\begin{aligned}\sum_{k=1}^{n}\lambda_k \leqslant\ & 3\,\|\vec{u}-\pi_h\vec{u}\,\|_{H^1(0,T;L^2(\Omega,\mathbb{R}^2))}^2 \\ & + 4K_3\tau^2\left(\left\|\frac{\partial\vec{u}}{\partial t}\right\|_{L^2(0,T;L^2(\Omega,\mathbb{R}^2))}^2 + \left\|\nabla\frac{\partial\vec{u}}{\partial t}\right\|_{L^2(0,T;L^2(\Omega,\mathbb{R}^2))}^2\right. \\ & \left. + \left\|\nabla\frac{\partial\pi_h\vec{u}}{\partial t} - \nabla\frac{\partial\vec{u}}{\partial t}\right\|_{L^2(0,T;L^2(\Omega,\mathbb{R}^2))}^2\right) \\ & + 4K_4T\,\|\vec{u}-\pi_h\vec{u}\,\|_{L^\infty(0,T;L^2(\Omega,\mathbb{R}^2))}^2 .\end{aligned} \tag{107}$$

Then it is plain, using Propositions 4.1 and 4.2 for the properties of the time dependent GVP, that there exists a positive constant C_1 independent of h and τ, such that for all $1 \leqslant n \leqslant N$

$$\sum_{k=1}^{n}\lambda_k \leqslant C_1(h^4+\tau^2). \tag{108}$$

Finally, using (10), (16) and (29), we find that there exists a constant C_2 independent of h and τ such that

$$\left\|\delta\vec{u}_h^0\right\|_0 \leqslant C_2h^2. \tag{109}$$

Therefore, using inequalities (108) and (109) together in inequality (106), we find that there exists a constant C_3 independent of h and τ such that, for any $0 < \tau \leqslant \tau^*$,

$$\|\delta\vec{u}_h^n\|_0 \leqslant C_3(h^2+\tau), \quad \text{for } n = 1, \ldots, N. \tag{110}$$

• We complete the proof of the convergence result by writing

$$\|\vec{u}(t^n) - \vec{u}_h^n\|_0 \leqslant \|\vec{u}(t^n) - \pi_h \vec{u}(t^n)\|_0 + \|\delta \vec{u}_h^n\|_0 \tag{111}$$

for all $1 \leqslant n \leqslant N$. Then we use Proposition 4.1 and estimate (110) to conclude that there exists a constant C independent of h and τ such that

$$\|\vec{u}(t^n) - \vec{u}_h^n\|_0 \leqslant C(h^2 + \tau), \tag{112}$$

for all $0 < \tau \leqslant \tau^*$ and $n = 1, \dots, N$. The theorem is then proved. ■

5. Numerical tests

Numerical tests have been performed on an adimensional problem using data from Kessler, Krüger and Scheid (1998) and Warren and Boettinger (1995). We refer to them for a complete physical description. We define the nonlinear terms F_1, F_2 and D_1, D_2 in Problem (P), for ϕ and c in the interval $[0,1]$. Outside this interval, all the terms are truncated to constant values. As we shall see, this gives Lipschitz and bounded nonlinear terms and so we are in the framework of the previous analysis. We choose (for ϕ and c in $[0,1]$) :

$$F_1(\phi) = \alpha_1 g'(\phi) + \beta_1 p'(\phi) \quad \text{and} \quad F_2(\phi) = \alpha_2 g'(\phi) + \beta_2 p'(\phi),$$

where α_i, β_i are model parameters linked to physical characteristics of the binary alloy we will consider and g be the polynomial double-well type function defined by $g(\phi) = \phi^2(1-\phi^2)$ and the related polynom $p = \int_0^\phi g(s)ds / \int_0^1 g(s)ds = \phi^3(6\phi^2 - 15\phi^4 + 10)$. On the other hand, we choose

$$D_1(\phi) = D + p(\phi)(1-D) \quad \text{and} \quad D_2(c,\phi) = \gamma c(1-c) D_1(\phi) F_2(\phi),$$

where D stands for the ratio of the diffusive coefficients in the pure solid and liquid phases. The parameter γ is linked to properties of the materials.

Notice that with those definitions, the terms F_i vanish for $\phi = 0$ and $\phi = 1$. Moreover these terms are taken to be zero when ϕ is outside of interval $[0,1]$. In that way, we obtain Lipschitz and bounded functions. The same process is applied for D_1 which is then always positive, and for D_2. In fact, a maximum principle holds (see Rappaz and Scheid (2000)) which guarantees that if initial data belong to $[0,1]$ then the same holds for the solution at any time and then the troncature procedure is justified.

The physical example we consider is a Ni-Cu alloy. The numerical values of physical parameters are given in Kessler, Krüger and Scheid (1998) and Warren and Boettinger (1995) and we report them in Table 1, for a problem adimensionalized in space relatively to a domain characteristic length $l = 2 \cdot 10^{-4} cm$ and in time relatively to the liquid diffusion characteristic time l/D_l, where $D_l = 10^{-5} cm^2/s$ is the physical diffusivity coefficient in the liquid phase.

These parameters have been derived from physical characteristics of the Ni-Cu alloy considering that the interface thickness is of order $\delta = 10^{-5} cm$, which is higher than what would be physically expected, but allows for reasonable calculation meshes and time steps (the α_i parameters go as $1/\delta^2$ and would become extremely high with a smaller value for δ).

Table 1: Values of the physical parameters

M	α_1	α_2	β_1	β_2	γ	D
63.5	$-3.23 \cdot 10^{11}$	$1.20 \cdot 10^{6}$	$-1.64 \cdot 10^{9}$	$3.17 \cdot 10^{9}$	$-8.85 \cdot 10^{-11}$	10^{-5}

We choose $\theta = 1$ in the scheme (9) of $(P_{h,\tau})$. The adimensional problem is defined in the unit square, and we fix a final adimensional time $t_f = 10^{-5}$ (higher times can create stability problems due to the stiffness of the source terms, i.e. the high values of their Lipshitz constants). We then construct an exact and explicit solution. We add right hand sides to equations in Problem (P) so that given functions $\phi_e(x, y, t)$ and $c_e(x, y, t)$ (defined for (x, y) in $[0, 1] \times [0, 1]$ and $t \in [0, t_f]$) are then solutions.

For a first test we choose the infinitely derivable functions

$$\phi_e(x, y, t) = c_e(x, y, t) = \frac{1}{2}\left(1 + \sin\left(2\pi\left(x + \frac{t}{t_f}\right)\right)\sin\left(2\pi\left(y + \frac{t}{t_f}\right)\right)\right)$$

We choose to relate the time step τ to the mesh size h of a regular mesh by the relationship $\tau = 40h^2$. Let us denote by $e_h = \max_{1 \leqslant n \leqslant N} \|\vec{u}(t^n) - \vec{u}_h^n\|_{L^2(\Omega,\mathbb{R}^2)}$ the error between the exact solution $\vec{u} = (\phi_e, c_e)$ and the computed solution $\vec{u}_h$.

We are interested in the local slope of the error with respect to h in logarithmic scale, which we define by

$$s_j = \frac{\ln(e_{h_j}) - \ln(e_{h_{j-1}})}{\ln(h_j) - \ln(h_{j-1})},$$

where h_{j-1} and h_j are choices of mesh sizes for two consecutive calculations, and e_{h_j} and $e_{h_{j-1}}$ the corresponding computed errors.

The results of these tests are given in table 2. Note that the slopes s_j take values very close to 2. This simple test therefore confirms our theoretical result of convergence order $h^2 + \tau$, with a very regular test function.

Nevertheless, on physical simulations, the solutions are not as regular as the product of two sines, and their main feature is that their values change very fast on regions of length scale δ (let us remind that δ was one of the model parameters, discussed earlier in this section). For this reason, we now present a second numerical test, with test functions reproducing the features of the

Table 2: Errors and convergence order for very regular test functions

j	h_j	e_{h_j}	s_j
1	$5.000E-2$	$3.700E-1$	
2	$2.500E-2$	$9.332E-2$	1.988
3	$1.667E-2$	$4.158E-2$	1.994
4	$1.250E-2$	$2.343E-2$	1.994
5	$1.000E-2$	$1.500E-2$	1.996
6	$8.333E-3$	$1.043E-2$	1.997
7	$7.143E-3$	$7.662E-3$	1.998
8	$6.250E-3$	$5.868E-3$	1.998
9	$5.556E-3$	$4.637E-3$	1.998

physical solutions, yet regular enough to be in the scope of our convergence theorem.

We define $\rho(t) = 0.15 + 0.5\dfrac{t}{t_f}$ and we choose

$$\phi_e(\vec{x}, t) = \begin{cases} 0 & \text{if } r(\vec{x}) < \rho(t), \\ 0.5\left(1 - \cos\left(\dfrac{r(\vec{x}) - \rho(t)}{2\delta}\pi\right)\right) & \text{if } \rho(t) < r(\vec{x}) < \rho(t) + 2\delta, \\ 1 & \text{if } \rho(t) + 2\delta < r(\vec{x}), \end{cases}$$

and

$$c_e(\vec{x}, t) = \begin{cases} 0 & \text{if } r(\vec{x}) < \rho(t), \\ 0.3 + 0.2\left(1 - \cos\left(\dfrac{r(\vec{x}) - \rho(t)}{\delta}\pi\right)\right) & \text{if } \rho(t) < r(\vec{x}) < \rho(t) + \delta, \\ 0.7 - 0.1\left(1 - \cos\left(\dfrac{r(\vec{x}) - \rho(t) - \delta}{\delta}\pi\right)\right) & \text{if } \rho(t) < r(\vec{x}) + \delta < \rho(t) + 2\delta, \\ 1 & \text{if } \rho(t) + 2\delta < r(\vec{x}), \end{cases}$$

where $r(\vec{x})$ is the distance between $\vec{x}$ and the center of Ω. The isovalues of the solution are expanding concentric circles with a boundary layer of width 2δ.

We follow the same procedure as for the previous tests. The results are given in table 3. Again this test confirms the theoretical result of convergence order $h^2 + \tau$.

6. Conclusion

In this paper we have obtained error estimates of a finite element method applied to a coupled system of non-linear evolution equations. These equations are related to a phase-field model for the solidification of a binary alloy. The main idea is to reduce the equations to a non-linear parabolic system and then introduce a generalized vectorial projector based on the elliptic part of the system operator. We derive projection errors in several norms. Error estimates of the piecewise linear finite element method we used are obtained by comparing the

Table 3: Errors and convergence order for test functions similar to physical solutions

j	h_j	e_{h_j}	s_j
1	$5.000E-2$	$2.561E-1$	
2	$2.500E-2$	$5.071E-2$	2.337
3	$1.667E-2$	$2.282E-2$	1.969
4	$1.250E-2$	$1.295E-2$	1.969
5	$1.000E-2$	$8.355E-3$	1.965
6	$8.333E-3$	$5.810E-3$	1.992
7	$7.143E-3$	$4.281E-3$	1.981
8	$6.250E-3$	$3.284E-3$	1.987
9	$5.556E-3$	$2.596E-3$	1.995

approximate solution with the generalized projection of the exact solution at every time step in the L^2 norm. It is shown that error estimates are of order 2 in the space mesh size h and of order 1 in the time step τ. In addition, there is no condition connecting h and τ. Numerical tests supply results which are in good agreement with the theoretical predictions.

Phase-field models are characterized by a small thickness region where an order parameter goes from 0 to 1 (the transition layer liquid/solid). Unfortunately, in the error estimates the constants depend on the inverse of this thickness so that constants become large when the thickness decreases. So in practice we need to have a small mesh size and time step in the transition region. Let us mention that a posteriori error estimates have also been performed on this model (see Krüger, Picasso and Scheid (2001)) and provided a criteria for the refinement of the mesh, thus allowing numerical calculations to be precise enough in the transition layer without refining the mesh in the whole domain, which would require too large calculation times.

Finally, let us remark that our analysis should be applicable to more general non-linear parabolic systems in $\mathbb{R}^n$ provided that the $n \times n$-matrix of the system is triangular and can be reduced (as we did it with our 2×2 matrix) to a definite positive matrix.

Acknowledgements

We thank Prof. J. Rappaz for initiating and managing the phase field research group at the Dept. of Mathematics of EPFL. We acknowledge the Swiss National Science Foundation for finantially supporting this project. We also thank Prof. M. Rappaz and the Dept. of Materials Science, EPFL, for an ongoing collaboration with our research work.

References

Barrett J.W., Blowey J.F., (1998). Finite element approximation of a model for phase separation of a multi-component alloy with a concentration-dependent mobility matrix, *IMA J. Numer. Anal.*, **18**, 287–328.

Chen Z., Hoffmann K.-H., (1994). An error estimate for a finite-element scheme for a phase field model, *IMA J. Numer. Anal.*, **14**, 243–255.

Ciarlet P.G., (1978). *The finite element method for elliptic problems*, North-Holland,

Kessler D., (2001). Sharp interface limits of a thermodynamically consistent solutal phase-field model, *to appear in J. Crystal Growth*,

Kessler D., Krüger O., Rappaz J., Scheid J.-F., (2000). A phase-field model for the isothermal solidification process of a binary alloy, *Computer Assisted Mechanics and Engineering Sciences*, **7**, 279–288.

Kessler D., Krüger O., Scheid J.-F., (1998). Construction d'un modèle de champ de phase à température homogène pour la solidification d'un alliage binaire, *DMA-EPFL internal report*,

Krüger O., Picasso M., Scheid J.-F., A posteriori error estimates and adaptive finite elements for a nonlinear parabolic problem related to solidification, *to appear in Applied Numerical Mathematics*,

Ladyzhenskaya O.A., Ural'tseva N.N., (1968). *Linear and quasilinear elliptic equations*, Academic Press,

Lions J.-L., Magenes E., (1968). *Problèmes aux limites non homogènes et applications,* Vol. 1, Dunod, Paris,

Quarteroni A., Valli A., (1991). *Numerical approximation of partial differential equations*, Springer-Verlag,

Rappaz J., Scheid J.-F., (2000). Existence of solutions to a phase-field model for the solidification process of a binary alloy, *Math. Methods Appl. Sci.*, **23**, 491–513.

Thomée V., (1991). *Galerkin Finite Element Methods for Parabolic Problems*, Springer, 2$^{\text{nd}}$ edition,

Warren J.A., Boettinger W.J., (1995). Prediction of dendritic growth and microsegregation patterns in a binary alloy using the phase-field model, *Acta metall. mater.*, **43**, 689–703.