A FULLY DISCRETE APPROXIMATION FOR CONTROL PROBLEMS GOVERNED BY PARABOLIC VARIATIONAL INEQUALITIES

MAïTINE BERGOUNIOUX* AND HOUSNAA ZIDANI†

Abstract. In this work we consider a numerical approximation of an optimal control problem governed by variational inequalities. We use a total discretization scheme: implicit Euler discretization with respect to the time variable and finite element method for the space variable, and give convergence results.

Key words. Error estimates, Discrete Approximations for Control Problems, State constraints, Unbounded controls.

AMS Classification: 49J20, 49M25

1. Introduction.

1.1. Formulation of the problem. This paper deals with numerical approximation of optimal control problems for semilinear parabolic variational inequalities with controls in L^p (not necessarily in L^∞) and state constraints. The convergence analysis is the main objective.

Let Ω be a bounded convex open subset of $\mathbb{R}^d$ ($d \geq 1$), with Lipschitz boundary Γ. We fix $T > 0$, and put $Q = \Omega\times]0,T[$, $\Sigma = \Gamma\times]0,T[$.

Let A be a second order uniformly elliptic operator:

$$Ay = -\sum_{j,k=1}^{d} D_j(a_{jk}(x)D_k y),$$

where the coefficients $a_{jk} \in \mathcal{C}^{1+\beta}(\overline{\Omega})$ (with $\beta > 0$) satisfy

$$\sum_{j,k} a_{jk}(x)\chi_j\chi_k \geq m_o|\chi|^2 \quad \text{for all } \chi \in \mathbb{R}^d \text{ and all } x \in \Omega$$

with $m_o > 0$. Consider the following parabolic variational inequality:

$$\frac{\partial y}{\partial x} + Ay + f(y) + \partial\kappa(y-\psi) \ni v \;\text{ in } \Omega\times]0,T[, \quad \frac{\partial y}{\partial n_A} + by = 0 \;\text{ on } \Gamma\times]0,T[, \quad y(0) = y_o \;\text{ in } \Omega, \tag{1.1}$$

where $y_o \in H^1(\Omega) \cap \mathcal{C}(\overline{\Omega})$, $\psi \in L^q(Q)$ is defined everywhere on Q, $b \in \mathbb{R}^+$, $f : \mathbb{R} \to \mathbb{R}$, and the control variable v is distributed. $\partial\kappa(y)(x,t) = \partial\kappa(y(x,t))$, $\partial\kappa(r)$ is the subdifferential of the function κ at $r \in \mathbb{R}$ and κ is the indicator function of $\mathbb{R}^+$:

$$\kappa(r) = \begin{cases} 0 & \text{if } r \geq 0 \\ +\infty & \text{else ,} \end{cases}$$

Define the following control constraint set:

$$U_{ad} = \left\{ v \in L^q(\Omega\times]0,T[) \mid \|v\|_{q,\Omega\times]0,T[} \leq M \right\},$$

where M is a fixed positive number. The control problem is defined by

*Département de Mathématiques, Université d'Orléans, Laboratoire MAPMO, UFR Sciences, BP 6759, 45067 Orlans, France, and INRIA-Rocquencourt, France. Email - Maitine.Bergounioux@labomath.univ-orleans.fr

† Département de Mathématiques, Université d'Orléans, Laboratoire MAPMO, UFR Sciences, BP 6759, 45067 Orlans, France, and INRIA-Rocquencourt, France. Email - Housnaa.Zidani@inria.fr

$$(\mathcal{P})\quad \begin{cases} \text{minimize } J(y,v) = \displaystyle\int_0^T\int_\Omega L(x,t,y(x,t)\,dx\,dt + \int_0^T\int_\Omega G(x,t,v(x,t))\,dxdt + \int_\Omega \ell(x,y(x,T))\,dx \\ \\ \text{subject to } \ v \in U_{ad}, \quad (y,v) \text{ satisfies (1.1)}, \quad F(x,t,y(x,t)) \le 0 \ \ \forall (x,t) \in \overline{\Omega}\times[0,T]. \end{cases}$$

We make the following additional assumptions:

(H1) $f : \mathbb{R} \longrightarrow \mathbb{R}$ is Lipschitz on $\mathbb{R}$ and of class $\mathcal{C}^1$. Hereafter, we denote by $C_o, \widetilde{C}_o \in \mathbb{R}$ real numbers such that

$$C_o \le f'(y) \le \widetilde{C}_o, \quad \forall y \in \mathbb{R}.$$

(H2) $q > \dfrac{d}{2} + 1$.
(H3) $y_o \in W_o^{2-\frac{1}{2q},q}(\Omega)(\subset \mathcal{C}(\overline{\Omega}) \cap W_o^{1,q}(\Omega))$.
(H4) $\psi \in L^q(Q)$ is defined everywhere on $\overline{Q}$ and $y_o(x) \ge \psi(x,0)$ for any $x \in \Omega$ (compatibility assumption for the initial condition).
(H5) F, L, $G : \Omega\times\mathbb{R}\times\mathbb{R} \longrightarrow \mathbb{R}$ are continuous. For any $(x,t) \in \Omega\times\mathbb{R}$, the function $F(x,t,\cdot)$ is Lipschitz continuous, and the function $G(x,t,\cdot)$ is convex.
(H6) $\ell : \Omega\times\mathbb{R} \longrightarrow \mathbb{R}$ is continuous. For any $x \in \Omega$, the function $\ell(x,\cdot)$ is Lipschitz.

The main goal of the present paper is to study a numerical approximation for $(\mathcal{P})$. Because of the variational inequality and state constraints this is delicate. First we use an idea based on the formulation of (1.1) as in [1] with a slackness variable and the regularity of its solution. Then we obtain a problem $(\widetilde{\mathcal{P}})$ equivalent to $(\mathcal{P})$, with constraints on both the control variable and state variable as coupled state/control constraints. We cannot avoid a first approximation of $(\widetilde{\mathcal{P}})$ that allows to relax some constraints. Otherwise, we would not be able to prove that the discretized formulation of the problem has at least one solution. Then we study the discretization of the relaxed problem and we shall prove that the discretized solutions are "close enough" the continuous one.

The main difficulty arising in the numerical study of optimal control problems governed by parabolic equations, is related to the nature of the state equation [10, 11, 20]. Indeed, error discretization estimations (either with respect to the time variable or to the space variable) are often established under regularity assumptions of the time-derivative of the solution [9, 14, 19]. In the case of optimal control problems, we cannot suppose that these assumptions are satisfied since it may change the nature of the problem under consideration.

There are many works on parabolic equations discretization process. We just mention here the book of Thomée [19], where error estimates are given for equations with smooth data. We also mention the papers of Choudury [6, 7], where some optimal error estimates are given for linear parabolic equations with nonsmooth data and with Dirichlet boundary conditions.

Here we are interested in the full discretization of a semilinear parabolic equation with nonsmooth data and with Robin-type boundary condition. We derive an error estimate in $L^\infty(L^2)$-norm (see theorem 3.2). For this, we use some techniques similar to the ones already used by Nochetto-Verdi [13].

Notation . For any s such that $1 \le s \le \infty$, the norms in the spaces $L^s(\Omega)$, $L^\tau(\Gamma)$, $L^s(Q)$, $L^\tau(\Sigma)$ are denoted by $\|\cdot\|_{s,\Omega}$, $\|\cdot\|_{s,\Gamma}$, $\|\cdot\|_{s,Q}$, $\|\cdot\|_{s,\Sigma}$. The inner products of $L^2(\Omega)$ and $L^2(\Gamma)$ are denoted respectively by $\langle\ ,\ \rangle_\Omega$ and $\langle\ ,\ \rangle_\Gamma$ while $\langle\ ,\ \rangle_{H^1}$ denotes the canonical duality pairing between $H^1(\Omega)$ and $(H^1(\Omega))'$.

The Hilbert space $W(0,T;H^1(\Omega),(H^1(\Omega))') = \{y \in L^2(0,T;H^1(\Omega)) \mid \dfrac{dy}{dt} \in L^2(0,T;(H^1(\Omega))')\}$ endowed with its usual norm, will be denoted by $W(0,T)$. The space $W^{2,1,q}(Q)$ is the usual space:

$$W^{2,1,q}(Q) = \{y \in L^q(Q) \mid \frac{\partial y}{\partial t},\ D_x y,\ D^2_{xx} y \in L^q(Q)\}.$$

In the sequel, we denote by C_i, for $i \in \mathbb{N} - \{0\}$, generic constants occurring in the estimates given in propositions.

1.2. Preliminaries. The weak solution of (1.1) associated to v is defined as the function $y \in W(0,T)$ satisfying $y(0) = y_o$, and

$$\int_0^T \langle \frac{\partial y}{\partial t}(t), y(t) - \varphi(t)\rangle_{H^1}\, dt + \int_Q \sum_{j,k} a_{jk}(x) D_k y D_j (y-\varphi)\, dx\, dt + \int_Q f(y)(y-\varphi)\, dx\, dt$$
$$+ \int_\Sigma b\, y(y-\varphi)\, ds dt \le \int_Q v\,(y-\varphi)\, dx\, dt,$$

for any $\varphi \in \mathcal{K}$ where

$$\mathcal{K} = \{\ \varphi \in L^2(0,T;H^1(\Omega))\ \mid\ \varphi \ge \psi \text{ a.e. in } Q\ \}\ .$$

Theorem 1.1. *Assume (H1)-(H4). For every $u \in L^q(Q)$, the variational inequality (1.1) admits a unique weak solution y_v in $\mathcal{C}(\overline{Q}) \cap W^{2,1,q}(Q)$. Moreover, the following holds.*

i) For every $v \in L^q(Q)$, there exists $\theta \in L^q(Q)$, such that y_v is also the solution of:

$$\begin{cases} \dfrac{\partial y}{\partial t} + Ay + f(y) &= v + \theta \quad \text{in } Q\ , \\ \dfrac{\partial y}{\partial n_A} + by &= 0 \quad \text{on } \Sigma\ , \qquad \text{and} \\ y(0) &= y_o \quad \text{in } \Omega\ , \end{cases} \tag{1.2a}$$

$$\begin{cases} y(x,t) - \psi(x,t) \ge 0 & \forall (x,t) \in Q, \\ \theta(x,t) \ge 0 & \text{a.e. } (x,t) \in Q, \\ (y(x,t) - \psi(x,t))\ \theta(x,t) = 0 & \text{a.e. } (x,t) \in Q, \end{cases} \tag{1.2b}$$

ii) There exists $C_1 = C_1(T,\Omega,d,q,C_o,m_o,\|y_o\|_{\mathcal{C}(\overline{\Omega})}) > 0$ such that, for any $(v,\theta) \in L^q(Q) \times L^q(Q)$, the weak solution $y_{v\theta}$ of (1.2a) associated to (v,θ) satisfies:

$$\|y_{v\theta}\|_{W^{2,1,q}(Q)} + \|y_{v\theta}\|_{\mathcal{C}(\overline{Q})} \le C_1\,(1 + \|v\|_{q,Q} + \|\theta\|_{q,Q}) \tag{1.3}$$

iii) For every $\varepsilon, K > 0$, there exists $\alpha > 0$, and $C_1' = C_1'(T,\Omega,d,q,C_o,m_o,\|y_o\|_{\mathcal{C}(\overline{\Omega})},\varepsilon,K,\alpha) > 0$ such that, for any $(v,\theta) \in L^q(Q) \times L^q(Q)$, $y_{v\theta}$ belongs to $\mathcal{C}^{\alpha,\frac{\alpha}{2}}(\overline{\Omega} \times [\varepsilon,T])$ and obeys

$$\|y_{v\theta}\|_{\mathcal{C}^{\alpha,\frac{\alpha}{2}}(\overline{\Omega}\times[\varepsilon,T])} \le C_1' \tag{1.4}$$

if $\|v\|_{q,Q} + \|\theta\|_{q,Q} \le K$.

Proof - The proof of the above results is similar to the one given in [1]. ■

In addition we shall use the following compactness result.

Theorem 1.2. *The mapping $(v,\theta) \mapsto y(v,\theta)$ where $y(v,\theta)$ is the solution of (1.2a), is sequentially continuous from $L^q(Q) \times L^q(Q)$, endowed with the weak-$L^q(Q) \times L^q(Q)$ topology, into $\mathcal{C}(\overline{Q})$ (strong topology).*

Proof - Let $\{(v_n,\theta_n)\}_n$ be a convergent sequence for the weak-$L^q(Q) \times L^q(Q)$ topology, and let (v,θ) be the weak limit of $(v_n,\theta_n)_n$. For any $n \in \mathbb{N}$, we denote by y_n the solution $y(v_n,\theta_n)$ of (1.2a) associated to (v_n,θ_n). From Theorem 1.1, we know that the sequence $(y_n)_n$ is bounded in $C(\overline{Q}) \cap W^{2,1,q}(Q)$. Furthermore, for every $\varepsilon > 0$, $(y_n)_n$ is bounded in a Hölder space on $\overline{\Omega} \times [\varepsilon, T]$. Then, there exists $y \in \mathcal{C}(\overline{Q}) \cap W^{2,1,q}(Q)$ such that $(y_n)_n$ converges to y weakly in $W^{2,1,q}(Q)$ and strongly in $\mathcal{C}(\overline{Q})$. By direct calculations, we can check that y is the solution of (1.2a) corresponding to (v,θ). ■

2. An equivalent problem to $(\mathcal{P})$. Using the previous result we may replace the state inequation by a system of equations involving a new control variable which is the Lagrange multiplier associated to the variational inequality. Therefore, the control problem $(\mathcal{P})$ turns to be a "standard" optimal control problem governed by a state equation and involving additional constraints on both the state and the control functions . More precisely, consider a new set of controls:

$$\Theta_{ad} = \{\theta \in L^q(Q) \mid \theta \geq 0 \text{ a.e. in } Q\}\,. \tag{2.1}$$

With Theorem 1.1, we see that problem $(\mathcal{P})$ is equivalent to the following one $(\widetilde{\mathcal{P}})$:

$$\text{Minimize} \quad J(y,v) \quad \text{subject to} \quad :$$

$$\frac{\partial y}{\partial t} + Ay + f(y) = v + \theta \text{ in } Q, \qquad \frac{\partial y}{\partial n_A} + by = 0 \text{ on } \Sigma, \qquad y(.,0) = y_o \text{ in } \Omega \tag{2.2a}$$

$$F(x,t,y(x,t)) \leq 0,\ y(x,t) \geq \psi(x,t) \text{ in } \overline{Q} \qquad \text{("Pure" state constraints)} \tag{2.2b}$$

$$(v,\theta) \in U_{ad} \times \Theta_{ad} \qquad \text{("Pure" control constraints)} \tag{2.2c}$$

$$\int_Q (y(x,t) - \psi(x,t))\ \theta(t,x)\ dx\ dt = 0 \quad \text{(Mixed State/Control integral constraints)}\ . \tag{2.2d}$$

We are going to perform a numerical study of this problem rather than the previous (genuine) one. To discretize the problem $(\widetilde{\mathcal{P}})$ and get some convergence results, we need to bound the new control function (see [2] for example) to use compactness properties. Therefore we set

$$V_{ad} = \left\{(v,\theta) \in U_{ad} \times \Theta_{ad} \mid \|\theta\|_{q,Q} \leq \widetilde{M}\right\}, \tag{2.3}$$

where $\widetilde{M}$ is a constant which is allowed to be very large and must be greater than the norm of the control $\bar{\theta}$ corresponding to an optimal solution of $(\mathcal{P})$ (as in [2]). From a numerical point of view this may be the largest constant allowed by the computer. The problem $(\mathcal{P})$ is still equivalent to $(\widetilde{\mathcal{P}})$ when we replace (2.2c) by:

$$(u,\theta) \in V_{ad}. \tag{2.4}$$

In the sequel we do not care about the existence of an optimal solution to $(\mathcal{P})$ (or $(\widetilde{\mathcal{P}})$): one can refer to [1]. Since we are interested in the numerical approximation of these problems we assume from now that such an optimal solution exists and we call it $(\bar{y},\bar{v},\bar{\theta})$.

3. Discretization of the state equation. We first give some results about the discretization of the following parabolic equation:

$$\frac{\partial y}{\partial t} + Ay + f(y) = g \quad \text{in } Q, \qquad \frac{\partial y}{\partial n_A} + by = 0 \quad \text{in } \Sigma, \qquad y(0) = y_o \quad \text{in } \Omega, \tag{3.1}$$

where g belongs to $L^q(Q)$. The weak solution y of (3.1) belongs to $W^{2,1,q}(Q)$, and satisfies

$$\frac{d}{dt}\langle y(t), \chi\rangle_{H^1} + \mathcal{A}(y(t), \chi) + \int_\Omega f(y(x,t))\chi(x)\,dx + \int_\Gamma by(s,t)\chi(s)\,ds = \int_\Omega g\chi(x)\,dx \tag{3.2}$$

for every $\chi \in H^1(\Omega)$ and a.e. $t \in]0, T[$. The bilinear form $\mathcal{A}$ is defined as follows

$$\forall\, y, z \ \in H^1(\Omega) \times H^1(\Omega) \qquad \mathcal{A}(y, z) = \int_\Omega \sum_{j,k} a_{jk}(x) D_k y(x)\, D_j z(x)\, dx\ .$$

3.1. Discretization and approximating spaces. Now, we make the discretization process precise: we use a finite difference scheme for the time variable (implicit Euler method) and a finite element approximation for the space variable (in Ω).

3.1.1. Grid for Ω. Let $(\mathcal{F}_h)_h$ be a family of triangulations of $\overline{\Omega}$ into closed d-simplices. To any simplex $K \in \mathcal{F}_h$, we associate two parameters:

- h_K:=diam(K),
- ρ_K:=sup $\{$ diam(S) $\mid S$ is a ball contained in $K\}$.

We suppose that $h = \max_{K\in\mathcal{F}_h} h_K$ and that $(\mathcal{F}_h)_h$ is regular in the following sense ([8], p. 132):
i) There exist two positive numbers η, γ such that:

$$\frac{h_K}{\rho_K} \le \eta \quad \text{and} \quad \frac{h}{\rho_K} \le \gamma \qquad \text{for all } K \in \mathcal{F}_h \quad \text{and all } h > 0.$$

ii) We set $\overline{\Omega}_h = \bigcup_{K\in\mathcal{F}_h} K$, Ω_h its interior (in general, $\Omega_h \neq \Omega$) and Γ_h its boundary. We assume that $\overline{\Omega}_h$ is convex.

Remark 3.1. *Since Ω is bounded, the discretization parameter h is necessarily less than some constant h_Ω which only depends on Ω.*
To every simplex $K \in \mathcal{F}_h$ dealing with the boundary, we associate a "curved" simplex $\tilde{K} \in \overline{\Omega}$ such that the d interior faces to Ω correspond with the ones of K, and such that the $(d+1) - th$ face is the part of Γ limited by the d other faces. We denote by $\widetilde{\mathcal{F}}_h$ the family composed by these simplexes $\tilde{K}$ and the simplexes contained inside Ω. Hence we have: $\overline{\Omega} = \bigcup_{K\in\widetilde{\mathcal{F}}_h} K$.
To any such triangulation $\widetilde{\mathcal{F}}_h$ we associate the finite dimensional following space

$$\mathcal{Y}_h = \left\{ z \in \mathcal{C}(\overline{\Omega}) : z|_K \text{ is affine for any } K \in \widetilde{\mathcal{F}}_h \right\}.$$

Let $\{x_j\}_{j=1}^{N_e}$ be the set of all nodes of $\widetilde{\mathcal{F}}_h$ on $\overline{\Omega}_h$. Let $\varphi_j(.)$ be the basis function associated to the node x_j ($\varphi_j \in \mathcal{Y}_h$, $\varphi_j(x_k) = 1$ if $k = j$, $\varphi_j(x_k) = 0$, if $k \neq j$).
In order to analyze the error we perform if we consider the approximation y_h of the system (3.6) instead of the (exact) solution y of (3.1), we consider the bilinear form $a(.,.)$ defined on $H^1(\Omega)$ by

$$a(w, z) := \mathcal{A}(w, z) + C_o\langle w, z\rangle_\Omega + b\,\langle w, z\rangle_\Gamma \qquad \forall w, z \in H^1(\Omega),$$

where C_o is the constant given in (H1). We can suppose, without loss of generality [15], that

$$a(z,z) \geq \frac{m_o}{2}\|z\|^2_{H^1(\Omega)} \qquad \text{for any} \quad z \in H^1(\Omega). \tag{3.3}$$

Let E_h be the operator of $H^1(\Omega)$ on $\mathcal{Y}_h$ which associates to any $z \in H^1(\Omega)$ the unique element $E_h z$ of $\mathcal{Y}_h$ such that for all $\chi \in \mathcal{Y}_h$:

$$\sum_{l,j=1}^{d} \int_\Omega a_{lj}(x) D_j(z - E_h z) D_l \chi \, dx + \int_\Omega C_o(z - E_h z)\chi \, dx + \int_\Gamma b\,(z - E_h z)\chi \, ds = 0. \tag{3.4}$$

We observe that this operator satisfies:

$$a(E_h z, \chi) = a(z,\chi) \quad \text{for any } \chi \in \mathcal{Y}_h \;\text{ and any }\; z \in H^1(\Omega)\ .$$

Therefore it is a projection operator and we have the following classical approximation results [8]:

PROPOSITION 3.1. *There exists $C_2 > 0$ depending on Ω such that for every $h < h_\Omega$, we have*

$$\|E_h z - z\|_{2,\Omega} \leq C_2\, h\, \|z\|_{H^1(\Omega)} \qquad \forall z \in H^1(\Omega).$$

3.1.2. Grid for Γ. We denote $\mathcal{B}_h$ the triangulation of Γ, inducted by the triangulation $\widetilde{\mathcal{F}}_h$.

3.1.3. Partition of $[0,T]$. Let N be a positive integer. We consider the uniform partition of $[0,T]$ defined by:

$$t_o = 0 < t_1 < t_2 < \cdots < t_N = T,$$

$$t_i = i\tau \quad (i = 0,\ldots,N), \quad \text{where} \quad \tau := \frac{T}{N}.$$

For any $i = 1,\ldots,N$, we denote by χ_i the characteristic function of $]\,t_{i-1}, t_i]$.
The exact value $y^i = y(\cdot, t_i)$ of the weak solution y of (3.2) at time t_i $(i = 1,\ldots,N)$ will be approximated by:

$$y_h^i(\cdot) := \sum_{j=1}^{N_e} Y_j^i \varphi_j(.) \ \in \mathcal{Y}_h, \quad j = 1,..,N_e \tag{3.5}$$

where $Y_j^i = y(x_j, t_i)$ (we remark that $Y_j^i = y(x_j,t_i) \in \mathbb{R}$ is well defined since $y \in \mathcal{C}(\overline{Q})$.)
Now we derive the discrete analog of (3.2) by means of which we shall define the approximate solution:

$$\begin{cases} \text{For } i = 1,\ldots,N \text{ and for any } \varphi \in \mathcal{Y}_h \\ \displaystyle\int_\Omega \frac{y_h^i - y_h^{i-1}}{\tau}\varphi\, dx + \mathcal{A}(y_h^i,\varphi) + \int_\Omega f(y_h^i)\,\varphi\, dx + \int_\Gamma b\, y_h^i \varphi\, ds = \frac{1}{\tau}\int_{t_{i-1}}^{t_i}\int_\Omega g(x,t)\,\varphi\, dx\, dt \\ y_h^o := E_h y_o. \end{cases} \tag{3.6}$$

THEOREM 3.1. *The system* (3.6) *admits a unique solution $\{y_h^i\}_{i=1}^N$ of the form (3.5).*
Proof - The proof of this theorem is based on compactness and monotony arguments (see [18], Chapter 5). ■

From now, we set $\delta = (h, \tau)$ (space-step, time-step), and we define the "discretized" solution y_δ for (3.1) using:

$$y_\delta(\cdot, t) = \sum_{i=1}^{N} \chi_i(t)\, y_h^i(\cdot) \quad \text{on }]\,0, T], \quad \text{and } y_\delta(\cdot, 0) = y_h^o \ . \tag{3.7}$$

The function y_δ belongs to $L^\infty(Q)$ and

$$\int_Q y_\delta^2(x,t)\, dx\, dt = \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} \int_\Omega (y_\delta(x,t))^2 dx\, dt = \tau \sum_{i=1}^{N} \int_\Omega y_h^i(x)^2 dx = \tau \sum_{i=1}^{N} \|y_h^i\|_{2,\Omega}^2 \ .$$

Remark 3.2. *In the scheme (3.6), the time discretization is the implicit regressive Euler discretization, which is known to be unconditionally stable [12, p.107].*

3.2. Error estimates for state equation discretization.

Lemma 3.1. *(Stability result)*

For every $M_o > 0$, there exists a positive constant $C_3 = C_3(\Omega, d, \sigma, C_o, m_o, \|y_o\|_{H^1(\Omega)}, M_o, h_\Omega)$ such that for any $h \in]\,0, h_\Omega[$, $\tau \in]0, T[$, and $g \in L^q(Q)$ satisfying $\|g\|_{q,Q} \le M_o$, the solution $\{y_h^i\}_{i=0,\dots,N}$ to (3.6) corresponding to g satisfies:

$$\tau \sum_{i=1}^{N} \|y_h^i\|_{H^1(\Omega)}^2 + \max_{1 \le i \le N} \|y_h^i\|_{2,\Omega}^2 + \sum_{i=1}^{N} \|y_h^i - y_h^{i-1}\|_{2,\Omega}^2 \le C_3$$

Proof - We take $\varphi := \tau\, y_h^i$ in (3.6) and sum from $i = 1$ to $i = i_o$ (i_o being a generic index in $\{1, .., N\}$); we obtain

$$\sum_{i=1}^{i_o} \int_\Omega (y_h^i - y_h^{i-1}) y_h^i\, dx + \sum_{i=1}^{i_o} \tau \mathcal{A}(y_h^i, y_h^i) + \sum_{i=1}^{i_o} \tau \int_\Omega f(y_h^i) y_h^i\, dx + \sum_{i=1}^{i_o} \tau \int_\Gamma b[y_h^i]^2\, ds$$

$$= \sum_{i=1}^{i_o} \int_{t_{i-1}}^{t_i} \int_\Omega g(x,t) y_h^i\, dx\, dt,$$

which can also be written as follows:

$$\sum_{i=1}^{i_o} \int_\Omega (y_h^i - y_h^{i-1}) y_h^i dx + \sum_{i=1}^{i_o} \tau \mathcal{A}(y_h^i, y_h^i) + \sum_{i=1}^{i_o} \tau \int_\Omega (f(y_h^i) - f(0)) y_h^i\, dx$$

$$+ \sum_{i=1}^{i_o} \tau \int_\Gamma b[y_h^i]^2\, ds = \sum_{i=1}^{i_o} \int_{t_{i-1}}^{t_i} \int_\Omega g(x,t) y_h^i\, dx\, dt - \sum_{i=1}^{i_o} \tau \int_\Omega f(0) y_h^i\, dx.$$

Taking (H1) into account, and using (3.3), we deduce from the above equality :

$$\sum_{i=1}^{i_o} \int_\Omega (y_h^i - y_h^{i-1}) y_h^i\, dx + m_o \sum_{i=1}^{i_o} \tau \|y_h^i\|_{H^1(\Omega)}^2 \le$$

$$K_1 \Big(\sum_{i=1}^{i_o} \int_{t_{i-1}}^{t_i} \|g(t)\|_{2,\Omega} \|y_h^i\|_{H^1(\Omega)}\, dt + \sum_{i=1}^{i_o} \tau\, |f(0)|\, \|y_h^i\|_{H^1(\Omega)} \Big),$$

where K_1 only depends on Ω and d. With the identity

$$2\sum_{i=1}^{i_o}(a_i - a_{i-1})a_i = a_{i_o}^2 - a_o^2 + \sum_{i=1}^{i_o}(a_i - a_{i-1})^2,$$

and the inequality $2\,ab \le \varepsilon a^2 + \frac{b^2}{\varepsilon}$, we obtain:

$$2m_o \sum_{i=1}^{i_o} \tau \|y_h^i\|^2_{H^1(\Omega)} + \|y_h^{i_o}\|^2_{2,\Omega} + \sum_{i=1}^{i_o} \|y_h^i - y_h^{i-1}\|^2_{2,\Omega} \le$$

$$\|y_h^o\|^2_{2,\Omega} + m_o \sum_{i=1}^{i_o} \tau \|y_h^i\|^2_{H^1(\Omega)} + \frac{2K_1}{m_o}[K_2\|g\|^2_{q,Q} + |f(0)|^2],$$

where $K_2 = K_2(\Omega, d, \sigma)$ is such that $\|\chi\|^2_{2,\Sigma} \le K_2\|\chi\|^2_{\sigma,\Sigma}$ for all $\chi \in L^\sigma(\Sigma)$. Thanks to proposition 3.1 and the bound $\|g\|_{q,Q} \le M_o$, we finally obtain

$$\sum_{i=1}^{i_o} \tau \|y_h^i\|^2_{H^1(\Omega)} + \|y_h^{i_o}\|^2_{2,\Omega} + \sum_{i=1}^{i_o} \|y_h^i - y_h^{i-1}\|^2_{2,\Omega} \le K_3,$$

for any $i_o \in \{1, .., N\}$, where K_3 only depends on $K_1, K_2, \Omega, M, m_o, C_o, \sigma, f(0), \|y_o\|_{H^1(\Omega)}$, and h_Ω. This completes the proof. ■

THEOREM 3.2. *Let $\delta = (h, \tau)$ be in $]\,0, h_\Omega[\times]0, T[$. Let $g \in L^q(Q)$ be such that $\|g\|_{q,Q} \le M_o$ (for $M_o > 0$), and let y and y_δ be the solutions of (3.1) and (3.6)-(3.7) associated to g. There exists a positive constant $C_4 = C_4(T, \Omega, q, m_o, C_o, \widetilde{C}_o, f(0), \|y_o\|_{H^1}, M_o, h_\Omega, C_1, C_2, C_3)$ independent of h, τ, g such that:*

$$\|y_\delta - y\|^2_{L^\infty(0,T;L^2(\Omega))} + \sum_{i=1}^{N} \frac{1}{\tau} \| \int_{I_i} (y_\delta(t) - y(t))\, dt\, \|^2_{H^1(\Omega)}\, dt \le C_4 \tau (1 + \frac{h^2}{\tau^2}).$$

Proof - We set $e = y - y_\delta$, $y^i = y(t_i)$, $y_h^i = y_\delta(t_i)$, $e^i = y^i - y_h^i$, and $I_i =]t_{i-1}, t_i]$ for $i \ge 1$.
We also set $\partial z^i = \dfrac{z^i - z^{i-1}}{\tau}$, $\overline{z}^i := \dfrac{1}{\tau}\displaystyle\int_{I_i} z(t)\,dt$ for $i \ge 1$, and $\overline{z}^o = z_o$ for $i = 0$. In particular we have $\overline{e}^i = \overline{y}^i - y_h^i$. In what follows, C is a constant independent of h, τ which may depend on M_o, m_o, C_1, C_2, C_3. Setting $\varphi = \tau\, E_h \overline{e}^i$ in (3.6) yields :

$$\tau\langle \partial y_h^i, E_h\overline{e}^i\rangle_\Omega + \tau\left(\mathcal{A}(y_h^i, E_h\overline{e}^i) + \langle f(y_h^i), E_h\overline{e}^i\rangle_\Omega + \langle b y_h^i, E_h\overline{e}^i\rangle_\Gamma\right) = \tau\langle \overline{g}^i, E_h\overline{e}^i\rangle_\Omega. \tag{3.8}$$

On the other hand, for any $z \in H^1(\Omega)$, we have:

$$\frac{d}{dt}\langle y(t), z\rangle_\Omega + \mathcal{A}(y(t), z) + \langle f(y(t)), z\rangle_\Omega + \langle b y(t), z\rangle_\Gamma = \langle g(t), z\rangle_\Omega.$$

As $y \in W^{1,2,q}(Q)$, then $y^i \in H^1(\Omega)$ and we may choose $z = \overline{e}^i$ in the previous equality. Then integrating over I_i gives:

$$\tau\langle \partial y^i, \overline{e}^i\rangle_\Omega + \tau\left(\mathcal{A}(\overline{y}^i, \overline{e}^i) + \langle \overline{f}^i, \overline{e}^i\rangle_\Omega + \langle b\overline{y}^i, \overline{e}^i\rangle_\Gamma\right) = \tau\langle \overline{g}^i, \overline{e}^i\rangle_\Omega. \tag{3.9}$$

We subtract (3.8) from (3.9) and sum up from $i=1$ to $i=i_o$ $(1 \le i_o \le N)$:

$$\begin{aligned}
(\mathrm{I}) + (\mathrm{II}) &:= \sum_{i=1}^{i_o} \tau \langle \partial (y^i - y_h^i), \overline{e}^i \rangle_\Omega + \sum_{i=1}^{i_o} \tau \left(a(\overline{y}^i, \overline{e}^i) - a(y_h^i, E_h \overline{e}^i) \right) \\
&= \sum_{i=1}^{i_o} \tau \langle \partial y_h^i, (E_h - I)\overline{e}^i \rangle_\Omega + \sum_{i=1}^{i_o} \tau \left(C_o \langle \overline{y}^i, \overline{e}^i \rangle_\Omega - C_o \langle y_h^i, E_h \overline{e}^i \rangle_\Omega \right) \\
&\quad + \sum_{i=1}^{i_o} \tau \Big(\langle f(y_h^i), E_h \overline{e}^i \rangle_\Omega - \langle \overline{f}^i, \overline{e}^i \rangle_\Omega \Big) + \sum_{i=1}^{i_o} \tau \langle \overline{g}^i, (I - E_h)\overline{e}^i \rangle_\Omega \\
&=: (\mathrm{III}) + (\mathrm{IV}) + (\mathrm{V}) + (\mathrm{VI}).
\end{aligned}$$

We now estimate each term of the above equality. First, we write (I) in the following form:

$$\begin{aligned}
(\mathrm{I}) &= \sum_{i=1}^{i_o} \tau \langle \partial (y^i - y_h^i), \overline{e}^i \rangle_\Omega \\
&= \sum_{i=1}^{i_o} \tau \langle \partial \overline{e}^i, \overline{e}^i \rangle_\Omega + \sum_{i=1}^{i_o} \tau \langle \partial (y^i - \overline{y}^i), \overline{e}^i \rangle_\Omega =: (\mathrm{I})_1 + (\mathrm{I})_2.
\end{aligned}$$

Using the following identity (for any F bilinear symmetric):

$$2 \sum_{i=1}^{i_o} F(x_i - x_{i-1}, x_i) = F(x_{i_o}, x_{i_o}) - F(x_o, x_o) + \sum_{i=1}^{i_o} F(x_i - x_{i-1}, x_i - x_{i-1}), \tag{3.10}$$

we obtain:

$$(\mathrm{I})_1 = \sum_{i=1}^{i_o} \langle \overline{e}^i - \overline{e}^{i-1}, \overline{e}^i \rangle_\Omega = \frac{1}{2} \|\overline{e}^{i_o}\|_{2,\Omega}^2 - \frac{1}{2} \|\overline{e}^0\|_{2,\Omega}^2 + \frac{1}{2} \sum_{i=1}^{i_o} \|\overline{e}^i - \overline{e}^{i-1}\|_{2,\Omega}^2.$$

From proposition 3.1, we get :

$$(\mathrm{I})_1 \ge C \|\overline{e}^{i_o}\|_{2,\Omega}^2 - C\, h^2 + C \sum_{i=1}^{i_o} \|\overline{e}^i - \overline{e}^{i-1}\|_{2,\Omega}^2.$$

Dealing with the estimate of $(\mathrm{I})_2$, we observe first that the following inequality holds:

$$\|y(t) - \overline{y}^i\|_{2,\Omega} \le \sqrt{\tau} \|\frac{\partial y}{\partial t}\|_{L^2(\Omega \times I_i)} \qquad \forall i \in \{1, .., N\} \;\text{ and }\; t \in \overline{I_i}. \tag{3.11}$$

With $\overline{y}^o = y^o$, (3.11), and (1.3), we have

$$\begin{aligned}
|(\mathrm{I})_2| &= |\langle y^{i_o} - \overline{y}^{i_o}, \overline{e}^{i_o} \rangle_\Omega - \sum_{i=2}^{i_o} \langle y^{i-1} - \overline{y}^{i-1}, \overline{e}^i - \overline{e}^{i-1} \rangle_\Omega | \\
&\le C \sqrt{\tau} \|\frac{\partial y}{\partial t}\|_{L^2(\Omega \times I_{i_o})} \, \|\overline{e}^{i_o}\|_{2,\Omega} + C \sum_{i=2}^{i_o} \sqrt{\tau} \|\frac{\partial y}{\partial t}\|_{L^2(\Omega \times I_{i-1})} \, \|\overline{e}^i - \overline{e}^{i-1}\|_{2,\Omega} \\
&\le \varepsilon \|\overline{e}^{i_o}\|_{2,\Omega}^2 + \varepsilon \sum_{i=1}^{i_o} \|\overline{e}^i - \overline{e}^{i-1}\|_{2,\Omega}^2 + \frac{C}{\varepsilon} \tau \|\frac{\partial y}{\partial t}\|_{L^2(\Omega \times I_{i-1})}^2 \\
&\le \varepsilon \|\overline{e}^{i_o}\|_{2,\Omega}^2 + \varepsilon \sum_{i=1}^{i_o} \|\overline{e}^i - \overline{e}^{i-1}\|_{2,\Omega}^2 + \frac{C}{\varepsilon} \tau \|\frac{\partial y}{\partial t}\|_{L^2(Q)}^2 \\
&\le \varepsilon \|\overline{e}^{i_o}\|_{2,\Omega}^2 + \varepsilon \sum_{i=1}^{i_o} \|\overline{e}^i - \overline{e}^{i-1}\|_{2,\Omega}^2 + \frac{C}{\varepsilon} \tau C_1^2 (1 + M_o)^2 \\
&\le \varepsilon \|\overline{e}^{i_o}\|_{2,\Omega}^2 + \varepsilon \sum_{i=1}^{i_o} \|\overline{e}^i - \overline{e}^{i-1}\|_{2,\Omega}^2 + \frac{C}{\varepsilon} \tau
\end{aligned}$$

where $\varepsilon > 0$ will be chosen at the end of estimates. From the definition of E_h, it follows:

$$\begin{aligned}(\text{II}) &= \sum_{i=1}^{i_o} \tau \left(a(\overline{y}^i, \overline{e}^i) - a(y_h^i, E_h \overline{e}^i) \right) = \sum_{i=1}^{i_o} \tau a(\overline{e}^i, E_h \overline{e}^i) = \sum_{i=1}^{i_o} \tau a(\overline{e}^i, \overline{e}^i) \\ &\geq C \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|^2_{H^1(\Omega)}.\end{aligned}$$

The term (III) is bounded by means of proposition 3.1 and of lemma 3.1 as follows:

$$\begin{aligned}|(\text{III})| &= \left| \sum_{i=1}^{i_o} \tau \langle \partial y_h^i, (E_h - I)\overline{e}^i \rangle_\Omega \right| \leq C_2 \sum_{i=1}^{i_o} \|y_h^i - y_h^{i-1}\|_{2,\Omega} \; h \|\overline{e}^i\|_{H^1(\Omega)} \\ &\leq \varepsilon \; \tau \sum_{i=1}^{i_o} \|\overline{e}^i\|^2_{H^1(\Omega)} + \frac{C_2^2 C_3}{4\varepsilon} \; \frac{h^2}{\tau} \\ &\leq \varepsilon \, \tau \sum_{i=1}^{i_o} \|\overline{e}^i\|^2_{H^1(\Omega)} + \frac{C}{\varepsilon} \; \frac{h^2}{\tau}.\end{aligned}$$

We bound (IV) using the same arguments as for (III):

$$\begin{aligned}|(\text{IV}))| &\leq |C_o \sum_{i=1}^{i_o} \tau \langle \overline{y}^i - y_h^i, \overline{e}^i \rangle_\Omega | + |C_o \sum_{i=1}^{i_o} \tau \langle y_h^i, (I - E_h)\overline{e}^i \rangle_\Omega | \\ &\leq |C_o| \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|^2_{2,\Omega} + |C_o| \sum_{i=1}^{i_o} \tau \|y_h^i\|_{2,\Omega} \; \|(I - E_h)\overline{e}^i\|_{2,\Omega} \\ &\leq |C_o| \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|^2_{2,\Omega} + |C_o| C_2 \sum_{i=1}^{i_o} \tau \|y_h^i\|_{H^1(\Omega)} \; h \|\overline{e}^i\|_{H^1(\Omega)} \\ &\leq C \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|^2_{2,\Omega} + \varepsilon \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|^2_{H^1(\Omega)} + \frac{C}{\varepsilon} h^2 \sum_{i=1}^{i_o} \tau \|y_h^i\|^2_{H^1(\Omega)} \\ &\leq C \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|^2_{2,\Omega} + \varepsilon \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|^2_{H^1(\Omega)} + \frac{C}{\varepsilon} h^2.\end{aligned}$$

Now we bound the term involving non linearities. We rewrite (V) in a suitable manner. In addition, we use the fact that f is a Lipschitz function and the proposition 3.1, which lead to:

$$\begin{aligned}|(\text{V})| &\leq | \sum_{i=1}^{i_o} \tau \langle f(y_h^i) - \overline{f}^i, \overline{e}^i \rangle_\Omega | + | \sum_{i=1}^{i_o} \tau \langle f(y_h^i), (I - E_h)\overline{e}^i \rangle_\Omega | \\ &\leq \sum_{i=1}^{i_o} |\langle \int_{I_i} (f(y_h^i) - f(y(t)))\, dt, \overline{e}^i \rangle_\Omega | + \sum_{i=1}^{i_o} \tau |\langle f(y_h^i), (I - E_h)\overline{e}^i \rangle_\Omega | \\ &\leq \widetilde{C}_o \sum_{i=1}^{i_o} \| \int_{I_i} |y_h^i - y(t)| \, dt \|_{2,\Omega} \; \|\overline{e}^i\|_{2,\Omega} + C \sum_{i=1}^{i_o} \tau \left(\|y_h^i\|_{2,\Omega} + |f(0)| \right) \|(E_h - I)\overline{e}^i\|_{2,\Omega} \\ &\leq \frac{\varepsilon}{2} \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|^2_{2,\Omega} + \frac{C}{\varepsilon} \int_0^{t_{i_o}} \|e(t)\|^2_{2,\Omega} \, dt + C \sum_{i=1}^{i_o} \tau (\|y_h^i\|_{2,\Omega} + 1) \; h \|\overline{e}^i\|_{H^1(\Omega)} \\ &\leq \frac{\varepsilon}{2} \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|^2_{H^1(\Omega)} + \frac{C}{\varepsilon} \int_0^{t_{i_o}} \|e(t)\|^2_{2,\Omega} + \frac{C\,(C_3 + 1)}{\varepsilon} h^2 + \frac{\varepsilon}{2} \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|^2_{H^1(\Omega)} \\ &\leq \frac{C}{\varepsilon} \int_0^{t_{i_o}} \|e(t)\|^2_{2,\Omega} \, dt + \frac{C}{\varepsilon} h^2 + \varepsilon \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|^2_{H^1(\Omega)}.\end{aligned}$$

Since $\|g\|_{q,Q} \le M$, we have:

$$
\begin{aligned}
|(\mathrm{VI})| = \left| \sum_{i=1}^{i_o} \tau \langle \overline{g}^i, (E_h - I)\overline{e}^i \rangle_\Omega \right| &\le C \sum_{i=1}^{i_o} \left(\int_{I_i} \|g(t)\|_{2,\Omega}^2 \, dt \right)^{1/2} h \, \sqrt{\tau} \|\overline{e}^i\|_{H^1(\Omega)} \\
&\le \frac{C}{\varepsilon} h^2 \sum_{i=1}^{i_o} \|g\|_{2,I_i\times\Omega}^2 + \varepsilon \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|_{H^1(\Omega)}^2 \le \frac{C}{\varepsilon} h^2 + \varepsilon \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|_{H^1(\Omega)}^2 .
\end{aligned}
$$

To summarize, we have proved the following estimate:

$$
\begin{aligned}
&\|\overline{e}^{i_o}\|_{2,\Omega}^2 + \sum_{i=1}^{i_o} \|\overline{e}^i - \overline{e}^{i-1}\|_{2,\Omega}^2 + \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|_{H^1(\Omega)}^2 \le \\
&\frac{C}{\varepsilon} (\tau + \frac{h^2}{\tau} + h^2) + \frac{C}{\varepsilon} \int_0^{t_{i_o}} \|e(t)\|_{2,\Omega}^2 \, dt + C \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|_{2,\Omega}^2 \\
&+ C\varepsilon \|\overline{e}^{i_o}\|_{2,\Omega}^2 + C\varepsilon \sum_{i=1}^{i_o} \|\overline{e}^i - \overline{e}^{i-1}\|_{2,\Omega}^2 + 4C\varepsilon \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|_{H^1(\Omega)}^2 . \qquad (3.12)
\end{aligned}
$$

Now we set $C\,\varepsilon = \frac{1}{6}$ so that the last three terms of the right-hand side in inequality (3.12) are controlled by the terms of the left-hand side of (3.12). Thus, for any $i_o \in \{1, ..., N\}$, we have:

$$
\|\overline{e}^{i_o}\|_{2,\Omega}^2 + \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|_{H^1(\Omega)}^2 \le C\,(\tau + \frac{h^2}{\tau} + h^2) + C \int_0^{t_{i_o}} \|e(t)\|_{2,\Omega}^2 \, dt + C \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|_{2,\Omega}^2 . \qquad (3.13)
$$

Let us observe that for $i \in \{1, ..., N\}$ and $t \in]t_{i-1}, t_i]$, we have:

$$
\begin{aligned}
\|e(t)\|_{2,\Omega}^2 &\le 2\|y(t) - \overline{y}^i\|_{2,\Omega}^2 + 2\|\overline{y}^i - y_h^i\|_{2,\Omega}^2 \\
&\le 2\tau \|\frac{\partial y}{\partial t}\|_{L^2(Q)}^2 + 2\|\overline{e}^i\|_{2,\Omega}^2 \\
&\le 2\tau C_1^2 (M_o + 1)^2 + 2\|\overline{e}^i\|_{2,\Omega}^2 \\
&\le 2\, \left(C\tau + \|\overline{e}^i\|_{2,\Omega}^2 \right) . \qquad (3.14)
\end{aligned}
$$

This observation with (3.13) yields:

$$
\|\overline{e}^{i_o}\|_{2,\Omega}^2 + \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|_{H^1(\Omega)}^2 \le C\,(\tau + \frac{h^2}{\tau} + h^2) + C \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|_{2,\Omega}^2 . \qquad (3.15)
$$

Using a discrete Gronwall inequality (see Lemma 3.2), we obtain:

$$
\|\overline{e}^{i_o}\|_{2,\Omega}^2 + \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|_{H^1(\Omega)}^2 \le C\,(\tau + \frac{h^2}{\tau} + h^2) \qquad (3.16)
$$

From (3.14), we have $\sup_{t\in]0,1]} \|\overline{e}(t)\|_{2,\Omega}^2 \le C \left(\max_{1\le i\le N} \|\overline{e}^i\|_{2,\Omega}^2 + \tau \right)$. This estimate together with (3.16) gives:

$$
\sup_{t\in]0,1]} \|e(t)\|_{2,\Omega}^2 + \sum_{i=1}^{i_o} \tau \|\overline{e}^i\|_{H^1(\Omega)}^2 \le C\,(\tau + \frac{h^2}{\tau} + h^2).
$$

This completes the proof. ■

Lemma 3.2. (*Discrete Gronwall inequality*).
Let $(a_n)_n$ and $(b_n)_n$ be two sequences of positive numbers, let $(c_n)_n$ be an increasing sequence of positive numbers, and let $\tau > 0$ such that

$$a_n + b_n \le c_n + \sum_{k=0}^{n-1} \tau a_k \qquad \forall n \ge 1,$$

with $a_o + b_o \le c_o$. Then, the following inequality holds:

$$a_n + b_n \le c_n \exp(\tau n) \quad \forall n.$$

Proof - We can find this version of Gronwall's inequality in [22] p. 43 (see also [18]). ■

4. Approximation of the optimal control problem.

4.1. Relaxation of the constraints. In [4], in order to study the approximation of optimal control problems governed by elliptic equations, the state equation is discretized with a finite element method and the state constraint is replaced by a finite number of constraints at the nodes of the discretization. Notice that error estimates for discretizations of elliptic equations (in dimension 3 or less than 3) can be obtained for the norm of $\mathcal{C}(\overline{\Omega})$, which allows to obtain the convergence of the discretization scheme analyzed in [4] and [5].

In the case of parabolic equations, we are only able to establish error estimates for the $L^\infty(L^2)$-norm, and thus we cannot deal with the state constraints as in [4, 5]. Therefore, we must give an integral form for these pointwise constraints. In addition, we must ensure that the discretized problem has a solution, i.e. the "discretized" feasible domain is non empty. A good candidate to be an element of this discretized feasible domain is of course the discrete approximation of the solution to the continuous problem (which satisfies the continuous constraints). Unfortunately, the constraints are not necessarily satisfied a priori for the discretized solution even if the discretization step δ is small. Therefore, we are obliged to relax these constraints with respect to some ε small enough (see the proof of proposition 4.3 below).
Thus, we consider a (continuous) penalized problem $(\mathcal{P}^\varepsilon)$ defined for $\varepsilon > 0$ as follows:

$$(\mathcal{P}^\varepsilon) \quad \inf \left\{ J(y,v,\theta) \;\middle|\; (y,v,\theta) \in \mathcal{C}(\overline{Q}) \times V_{ad},\ (y,v,\theta) \text{ satisfies (2.2}a\text{) and} \right.$$
$$\left. \int_Q \left((y-\psi)^-\right)^2 dx\,dt \le \varepsilon, \int_Q (F^+(x,t,y(x,t)))^2\,dx\,dt \le \varepsilon, \left(\int_Q (y-\psi)\,\theta\,dx\,dt\right)^2 \le \varepsilon \right\}.$$

Here F^+ denotes the nonnegative part of F (and y^- the negative part of y). Since V_{ad} is bounded in $L^q(Q) \times L^q(Q)$ and the state equation is linear with respect to the control variables, the problem $(\widetilde{\mathcal{P}})$ is weakly stable to the right in the following sense.

Proposition 4.1. *We have* $\lim\limits_{\varepsilon \to 0} (\inf(\mathcal{P}^\varepsilon)) = \inf(\widetilde{\mathcal{P}}) = \inf(\mathcal{P})$.
Proof - Step 1: We fix $\varepsilon > 0$ and we prove that the problem $(\mathcal{P}^\varepsilon)$ has at least one solution $(y_\varepsilon, v_\varepsilon, \theta_\varepsilon)$. It's clear that the feasible domain of $(\mathcal{P}^\varepsilon)$ is non empty (any solution of $(\widetilde{\mathcal{P}})$ is feasible for $(\mathcal{P}^\varepsilon)$). Let $(y_n, v_n, \theta_n)_{n\ge1}$ be a minimizing sequence of $(\mathcal{P}^\varepsilon)$. The sequence $(v_n, \theta_n)_n$ is included in V_{ad} and is bounded in $L^q \times L^q(Q)$; therefore there exists a subsequence, still denoted by $(v_n, \theta_n)_n$, converging to some $(v_\varepsilon, \theta_\varepsilon)$ in $L^q \times L^q(Q)$. Since V_{ad} is convex closed, it's also weakly closed and then $(v_\varepsilon, \theta_\varepsilon)$ belongs to V_{ad}. In addition, theorem 1.2 gives the strong convergence in $\mathcal{C}(\overline{Q})$ of $(y_n)_n$ to the solution y_ε of (2.2a)

corresponding to $(v_\varepsilon, \theta_\varepsilon)$. On the other hand, since (y_n, v_n, θ_n) is feasible for $(\mathcal{P}^\varepsilon)$, we have:

$$\int_Q \left((y_n-\psi)^-\right)^2 \, dx\, dt \le \varepsilon, \ \int_Q (F^+(x,t,y_n(x,t)))^2 \, dx\, dt \le \varepsilon, \ \left(\int_Q (y_n-\psi)\,\theta_n \, dx\, dt\right)^2 \le \varepsilon,$$

for every $n \ge 1$. By passing to the limit, when $n \to \infty$, we deduce that

$$\int_Q \left((y_\varepsilon-\psi)^-\right)^2 \, dx\, dt \le \varepsilon, \ \int_Q (F^+(x,t,y_\varepsilon(x,t)))^2 \, dx\, dt \le \varepsilon, \ \left(\int_Q (y_\varepsilon-\psi)\,\theta_\varepsilon \, dx\, dt\right)^2 \le \varepsilon,$$

and then $(y_\varepsilon, v_\varepsilon, \theta_\varepsilon)$ is feasible for the problem $(\mathcal{P}^\varepsilon)$. By the continuity of J with respect to y and the convexity and continuity assumption with respect to v (see (H5)-(H6)), we conclude:

$$\mathrm{Inf}(\mathcal{P}^\varepsilon) \le J(y_\varepsilon, v_\varepsilon) \le \liminf_n J(y_n, v_n) = \mathrm{Inf}(\mathcal{P}^\varepsilon).$$

Therefore, $(y_\varepsilon, v_\varepsilon, \theta_\varepsilon)$ is a solution of $(\mathcal{P}^\varepsilon)$.

Step 2: Now we study the asymptotic behavior of $(y_\varepsilon, v_\varepsilon, \theta_\varepsilon)_\varepsilon$. Again, since V_{ad} is bounded in $L^q(Q) \times L^q(Q)$, there exists a sequence $\{\varepsilon_j\}_{j=1}^\infty$ and $(\bar v, \bar\theta) \in V_{ad}$ such that $\varepsilon_j \to 0$, $v_{\varepsilon_j} \to \bar v$ weakly in $L^q(Q)$, and $\theta_{\varepsilon_j} \to \bar\theta$ in $L^q(Q)$ when $j \to \infty$. If we denote by y_{ε_j} and $\bar y$ the states associated respectively to $(v_{\varepsilon_j}, \theta_{\varepsilon_j})$ and to $(\bar v, \bar\theta)$, Theorem 1.2 yields that $y_{\varepsilon_j} \to \bar y$ uniformly in $\overline{Q}$. By the same arguments as in Step 1, we can check that $(\bar y, \bar v, \bar\theta)$ is feasible for $(\widetilde{\mathcal{P}})$. Using again the convexity of $\tilde J$ with respect to the second variable and the feasibility of $(\bar y, \bar v, \bar\theta)$ for $(\widetilde{\mathcal{P}})$, we get

$$\inf(\widetilde{\mathcal{P}}) \le J(\bar y, \bar v) \le \liminf_{j\to\infty} J(y_{\varepsilon_j}, v_{\varepsilon_j}) = \lim_{j\to\infty} \inf(\mathcal{P}^{\varepsilon_j}) \le \inf(\widetilde{\mathcal{P}}),$$

which concludes the proof. ■

REMARK 4.1. *The stability considered in the above proposition has been already introduced by many authors (see [4, 5] for example) in order to study the approximation of optimal control problems governed by elliptic equations.*

REMARK 4.2. *We may remark that the relaxation of the bilinear integral constraint is needed anyway if a related Lagrange multiplier is expected [2]. Indeed, this constraint is too stressing, so that usual regularity conditions (see [24] for example) cannot be satisfied. Therefore, it is not possible to ensure the existence of a Lagrange multiplier associated to this constraint.*

4.2. The discretized problem. We recall that for any h and τ (space and time discretization steps), we set $\delta = (h, \tau)$ and we consider the space $\mathcal{V}_\delta$ defined by:

$$\mathcal{V}_\delta = \{v_\delta \in L^q(Q) \mid v_\delta|_{K\times]t_{i-1},t_i]} \text{ is constant for any } K \in \widetilde{\mathcal{F}}_h \text{ and any } i = 1,\dots,N\}. \tag{4.1}$$

Any function v_δ of $\mathcal{V}_\delta$ may be written as

$$v_\delta(x,t) = \sum_{i=1}^{N} \sum_{j=1}^{N_e} V_j^i \, \chi_i(t)\, \chi_{K_j}(x) \tag{4.2}$$

where χ_{K_j} is the characteristic function of K_j and $V_j^i \in \mathbb{R}$. Any function v of $L^q(Q)$ can be approximated by v_δ as in (4.2) where

$$V_j^i = \frac{1}{|Q_{ij}|} \int_{Q_{ij}} v(x,t)\, dx\, dt \qquad \text{for any } i = 1,\dots,N, \ j = 1,\dots,N_e \ ,$$

where $Q_{ij} = K_j \times]t_{i-1}, t_i]$ and $|Q_{ij}|$ is the measure of Q_{ij}.

PROPOSITION 4.2. *Assume $v \in L^q(Q)$ and $v_\delta \in \mathcal{V}_\delta$ is defined as above. Then*

$$\lim_{\delta \to 0} \|v_\delta - v\|_{q,Q} = 0 \ .$$

Proof - Let $(x,t) \in Q$ where $v(x,t)$ makes sense and $Q_{\delta_k} = Q_{i_\delta\, j_\delta}$ a sequence of discretized cells which tends to $\{(x,t)\}$ as $\delta \to 0$. Extending v by 0 outside Q we have $v \in L^1(\mathbb{R}^{d+1})$ and a classical result (see [17] for example) yields that

$$\frac{1}{|Q_{\delta_k}|} \int_{Q_{\delta_k}} v(\xi, s)\ d\xi\, ds \to v(x,t) \ \text{ as } \delta \to 0 \ .$$

Therefore v_{δ_k} converges to v almost everywhere, and $v_{\delta_k}^q$ converges to v^q as well. In addition, with Hölder inequality we get

$$|V_j^i|^q \le \frac{1}{|Q_{ij}|} \int_{Q_{ij}} |v(x,t)|^q \ dx\, dt \ ,$$

and we obtain

$$\|v_\delta\|_{q,Q}^q = \sum_{i=1}^{N} \sum_{j=1}^{N_e} |Q_{ij}|\, |V_j^i|^q \le \sum_{i=1}^{N} \sum_{j=1}^{N_e} \int_{Q_{ij}} |v(x,t)|^q \ dx\, dt \ \le \|v\|_{q,Q}^q \ .$$

We conclude with the Lebesgue dominated convergence theorem.
The end of the proposition follows from the uniqueness of the weak limit. ■

Similarly, we recall that, for $\delta = (h, \tau)$, the discretized solution of the state equation is defined as

$$\left\{ \begin{array}{ll} y_\delta(x,t) = \displaystyle\sum_{i=1}^{N} \chi_i(t)\, y_h^i(x) = \sum_{i=1}^{N} \sum_{j=1}^{N_e} Y_j^i\, \chi_i(t)\, \varphi_j(x) & \text{in } \Omega \times]0,T], \\ y_\delta(\cdot, 0) = E_h y_o. & \end{array} \right. \tag{4.3}$$

We call $\mathcal{Y}_\delta$ the finite dimensional space which basis is $(\chi_i \varphi_j)_{i=1,\ldots,N, j=1,\ldots,N_e}$. This space dimension is $N_{tot} = N * N_e$. In the sequel we shall set

$$V^i = (V_j^i)_{j=1,\ldots,N_e} \in \mathbb{R}^{N_e} \ , \ V = (V^i)_{i=1,\ldots,N} \in \mathbb{R}^{N_{tot}} \ ,$$

$$Y^i = (Y_j^i)_{j=1,\ldots,N_e} \in \mathbb{R}^{N_e} \ , \ Y = (Y^i)_{i=1,\ldots,N} \in \mathbb{R}^{N_{tot}} \ .$$

We define now the well known mass and stiffness matrices:

$$[M] = \left[\int_\Omega \varphi_i(x)\, \varphi_j(x)\, dx \right]_{1 \le i,j \le N_e} \quad \text{and} \quad [R] = \left[\int_\Omega D\varphi_i(x)\, D\varphi_j(x)\, dx \right]_{1 \le i,j \le N_e} . \tag{4.4}$$

With these notations, we have

$$\|y_h^i\|_{H^1(\Omega)}^2 = (Y^i)^\top ([M] + [R])\, Y^i \ ,$$

where $Z^\top$ denotes the transposed vector of Z. We recall that $[M]$ and $[R]$ are symmetric, definite positive. Let us detail the discretized equation (3.6). It is equivalent to

$$\left\{\begin{array}{l} \text{For } i=1,\dots,N \text{ and } j=1,\dots,N_e \\ \displaystyle \frac{1}{\tau}\sum_{k=1}^{N_e}\left(\int_\Omega \varphi_k\varphi_j\,dx\right)(Y_k^i - Y_k^{i-1}) + \sum_{k=1}^{N_e}\mathcal{A}(\varphi_k,\varphi_j)Y_k^i \\ \displaystyle +\int_\Omega f(y_h^i)\,\varphi_j\,dx + \sum_{k=1}^{N_e}\left(\int_\Gamma b\,\varphi_k\,\varphi_j\,ds\right)Y_k^i = \frac{1}{\tau}\int_{t_{i-1}}^{t_i}\int_\Omega g(\cdot,t)\,\varphi_j\,dx\,dt \\ y_h^o := E_h y_o. \end{array}\right.$$

Let us set

$$[A] = [\mathcal{A}(\varphi_k,\varphi_j)]_{1\le k,j\le N_e} \quad \text{and} \quad [B] = [\int_\Gamma b\varphi_k\varphi_j\,ds]_{1\le k,j\le N_e}\ ,$$

and for every $i=1,\dots,N$

$$\Phi(Y^i) = \left(\int_\Omega f(y_h^i(x))\,\varphi_j(x)\,dx\right)_{j=1,\dots,N_e}, \quad H_i(g) = \left(\int_{t_{i-1}}^{t_i}\int_\Omega g(x,t)\,\varphi_j(x)\,dx\,dt\right)_{j=1,\dots,N_e}.$$

Then the above relation is equivalent to

$$\left\{\begin{array}{l} \text{For } i=1,\dots,N \\ ([M]+\tau[A]+\tau[B])\,Y^i + \tau\,\Phi(Y^i) = [M]Y^{i-1} + H_i(g)\ , \\ y_h^o := E_h y_o. \end{array}\right. \tag{4.5}$$

Now we may define the discretized problem corresponding to $(\mathcal{P}^\varepsilon)$. For any (v_δ,θ_δ) in $\mathcal{V}_\delta\times\mathcal{V}_\delta$ we denote by $y_\delta(v_\delta,\theta_\delta)$ the solution of (4.3)-(4.5) associated to $g=v_\delta+\theta_\delta$. We set

$$V_{ad,\delta} = V_{ad}\cap(\mathcal{V}_\delta\times\mathcal{V}_\delta)\ ,$$

and

$$\psi_\delta(x,t) = \sum_{i=1}^{N}\sum_{j=1}^{N_e}\Psi_j^i\,\chi_i(t)\,\varphi_j(x) \quad \text{where } \Psi_j^i = \psi(x_j,t_i)\ .$$

For any $\varepsilon>0$, we define

$$(\mathcal{P}_\delta^\varepsilon) \qquad \left\{\begin{array}{l} \text{minimize} \quad J(y_\delta,v_\delta,\theta_\delta) \\ \text{with } (v_\delta,\theta_\delta)\in V_{ad,\delta} \text{ and } y_\delta = y_\delta(v_\delta,\theta_\delta) \\ \displaystyle \int_Q ((y_\delta-\psi_\delta)^-)^2\,dx\,dt \le \varepsilon, \int_Q (F^+(x,t,y_\delta))^2\,dx\,dt\le\varepsilon, \\ \displaystyle \left(\int_Q \theta_\delta\,(y_\delta-\psi_\delta)\,dx\,dt\right)^2 \le \varepsilon. \end{array}\right.$$

We may now ascertain that the discretized problem has a solution :

Proposition 4.3. *Given $\varepsilon>0$, there exists $\delta^*=(h^*,\tau^*)$ such that for all $\delta=(h,\tau)\in]\,0,h^*]\times]\,0,\tau^*]$ the problem $(\mathcal{P}_\delta^\varepsilon)$ has a solution.*

Proof - Let us fix $\varepsilon > 0$ and let us prove that there exists $\delta^* = (h^*, \tau^*)$ such that for all $\delta \in]\,0, h^*] \times]\,0, \tau^*]$, $\frac{h}{\tau} = O(1)$, the feasible domain $\mathcal{D}^\varepsilon_\delta$ of $(\mathcal{P}^\varepsilon_\delta)$ is non empty.
Let $(\bar{y}, \bar{v}, \bar{\theta})$ be a solution of $(\widetilde{\mathcal{P}})$ and let $(v_\delta, \theta_\delta)$ be the discrete approximation of $(\bar{v}, \bar{\theta})$ which is known to belong to $V_{ad,\delta}$ and converge to $(\bar{v}, \bar{\theta})$ in $L^q(Q) \times L^q(Q)$ (proposition 4.2). Let $y_\delta = y_\delta(v_\delta, \theta_\delta)$ be the solution of (4.3)-(4.5) associated to $g = v_\delta + \theta_\delta$, and $y(v_\delta, \theta_\delta)$ the solution of (2.2a) associated to $(v_\delta, \theta_\delta)$. Since we have

$$\|y_\delta - \bar{y}\|^2_{L^\infty(0,T;L^2(\Omega))} \le \|y_\delta - y(v_\delta, \theta_\delta)\|^2_{L^\infty(0,T;L^2(\Omega))} + \|y(v_\delta, \theta_\delta) - \bar{y}\|^2_{L^\infty(0,T;L^2(\Omega))},$$

Theorem 3.2 yields

$$\|y_\delta - \bar{y}\|^2_{L^\infty(0,T;L^2(\Omega))} \le C_4 \tau (1 + \frac{h^2}{\tau^2}) + \|y(v_\delta, \theta_\delta) - \bar{y}\|^2_{L^\infty(0,T;L^2(\Omega))}$$

and by theorem 1.2 we conclude that $\|y_\delta - \bar{y}\|_{L^\infty(L^2)} \to 0$ as $\delta = (h, \tau) \to 0$ with $\dfrac{h}{\tau} = O(1)$.
Since $\bar{y}$ satisfies the constraints:

$$\int_Q ((\bar{y} - \psi)^-)^2 \, dx\, dt = 0, \quad \int_Q (F^+(x, t, \bar{y}))^2 \, dx\, dt = 0, \left(\int_Q \bar{\theta}\, (\bar{y} - \psi) \, dx\, dt \right)^2 = 0,$$

the function y_δ satisfies

$$\int_Q ((y_\delta - \psi_\delta)^-)^2 \, dx\, dt \le \varepsilon, \int_Q (F^+(x, t, y_\delta))^2 \, dx\, dt \le \varepsilon, \left(\int_Q \theta_\delta\, (y_\delta - \psi_\delta) \, dx\, dt \right)^2 \le \varepsilon,$$

as soon as $\delta = (h, \tau)$ is small enough, say $\delta \le \delta^*$ where δ^* depends on ε. Therefore, $\mathcal{D}^\varepsilon_\delta$ is non empty for such δ.
Since $\mathcal{D}^\varepsilon_\delta$ is closed and nonempty and thanks to (H4)-(H6), the end of the proof is classical. ■

We end this section with a convergence result of the solution to the discretized problem $(\mathcal{P}^\varepsilon_\delta)$ to the solution of the continuous problem $(\mathcal{P}^\varepsilon)$.

THEOREM 4.1. *We fix $\varepsilon > 0$. If $(v^\varepsilon_\delta,\ \theta^\varepsilon_\delta)_{\delta \le \delta^*}$ denotes a solution to $(\mathcal{P}^\varepsilon_\delta)$, one can extract a subsequence weakly converging towards $(v_\varepsilon, \theta_\varepsilon)$ in $L^q(Q) \times L^q(Q)$, where $(v_\varepsilon, \theta_\varepsilon)$ is a solution to $(\mathcal{P}^\varepsilon)$. In addition we get*

$$\lim_{\delta \to 0}\ \inf(\mathcal{P}^\varepsilon_\delta) = \inf\ (\mathcal{P}^\varepsilon)\ .$$

Proof - The sequence $(v^\varepsilon_\delta, \theta^\varepsilon_\delta)_\delta$ belongs to V_{ad} and is bounded in $L^q(Q) \times L^q(Q)$ (uniformly with respect to δ); therefore, there exists a subsequence (still denoted $(v^\varepsilon_\delta, \theta^\varepsilon_\delta)_\delta$) and $(v_\varepsilon, \theta_\varepsilon) \in V_{ad}$, such that $(v^\varepsilon_\delta, \theta^\varepsilon_\delta)_\delta$ weakly converges towards $(v_\varepsilon, \theta_\varepsilon)$ in $L^q(Q) \times L^q(Q)$ as $\delta \to 0$.
Let y^ε_δ be the solution to (4.3)-(4.5) associated to $v^\varepsilon_\delta + \theta^\varepsilon_\delta$, and y_ε be the solution of (2.2a) associated to $(v_\varepsilon, \theta_\varepsilon)$. We know with Theorem 3.2, that

$$\lim_{\delta \to 0} \|y^\varepsilon_\delta - y(v^\varepsilon_\delta, \theta^\varepsilon_\delta)\|_{2,Q} = 0$$

where $y(v^\varepsilon_\delta, \theta^\varepsilon_\delta)$ is the solution to the (continuous) state-equation (2.2a) associated to $(v^\varepsilon_\delta, \theta^\varepsilon_\delta)$. In addition,

$$\lim_{\delta \to 0} \|y(v^\varepsilon_\delta, \theta^\varepsilon_\delta) - y_\varepsilon\|_{2,Q} = 0\ ,$$

using the compactness result of Theorem 1.2 .
Hence y_δ^ε strongly converges towards y_ε in $L^2(Q)$ (and even in $L^\infty(0,T;L^2(\Omega))$). Since $(y_\delta^\varepsilon, v_\delta^\varepsilon, \theta_\delta^\varepsilon)$ satisfies the constraints of the problem $(\mathcal{P}_\delta^\varepsilon)$, by passing to the limit when δ tends to 0 and taking the Lipschitz continuity of F and the continuity of ψ into account (see (H4)-(H5)), we obtain

$$\int_Q (F^+(x,t,y_\varepsilon))^2\,dx\,dt \le \varepsilon\ , \quad \int_Q ((y_\varepsilon-\psi)^-)^2\,dx\,dt \le \varepsilon,$$

and

$$\left(\int_Q (y_\varepsilon-\psi)\,\theta_\varepsilon\,dx\,dt\right)^2 \le \varepsilon\ .$$

Therefore $(y_\varepsilon, v_\varepsilon, \theta_\varepsilon)$ is feasible for $(\mathcal{P}^\varepsilon)$. The end of the proof is a consequence of the semi-continuity on $L^q(Q)$-weak of the functional J to minimize. ■

REMARK 4.3. *The relaxation of the constraints via ε is imposed by the necessity to ensure the non vacuity of the discretized feasible domain. Of course, this is useless if we are able to ascertain that $\mathcal{D}_\delta^o$ is non empty (i.e. the discretized feasible domain corresponding to "$\varepsilon = 0$"), that is, for example if we precisely assume*

$$(H6) \qquad \forall \delta \le \delta^* \qquad \exists (y_\delta^o, v_\delta^o, \theta_\delta^o) \in \mathcal{D}_\delta^o\ .$$

5. Conclusion. Our purpose was to compute the solution to problem $(\mathcal{P})$. The first step was to discretize the problem and to give appropriate error estimates : this is done in the present paper. Next step is to formulate a numerical algorithm that allows to solve $(\mathcal{P}_\delta^\varepsilon)$. It is still a significant challenge because of the (bilinear) complementarity constraint

$$\left(\int_Q \theta_\delta\,(y_\delta-\psi_\delta)\,dx\,dt\right)^2 \le \varepsilon,$$

that can be (formally) written

$$\Theta^t\,M(Y-\Psi) \le \varepsilon\ ,$$

where Θ, Ψ and Y are the component vectors for discretized functions and M is a real matrix.

Standard SQP methods cannot be used directly and have to be adapted together with a convex approximation of this complentarity constraint. Therefore, the convergence analysis is not obvious. This work will be developed in a forthcoming paper, where we will present numerical examples.

REFERENCES

[1] **M. Bergounioux, H.Zidani**, *Pontryagin Maximum Principle for Optimal control of Variational Inequalities,* SIAM Journal on Control and Optimization, Vol. 37, n^o 4, pp. 1273-1290, 1999
[2] **M. Bergounioux**, *Optimal Control of an Obstacle Problem,* Applied Mathematics and Optimization, 36, pp 147-172 (1997).
[3] **J. Burger, M. Pogu**, *Functional and Numerical Solution of a Control Problem Originating from Heat Transfer*, J. Optim. Theo. Appl., Vol. 68 (1991), p. 49-73.
[4] **E. Casas**, *Analisis de la Convergencia en la Approximacion Numerica de Problemas de Control con Restricciones Sobre el Estado*, Actas del XII Cedya II Congresso de Matematica Aplicada Oviedo, 1991, p. 301-306.
[5] **E. Casas**, *The Relaxation Theory Applied to Optimal Control Problems of Semilinear Elliptic Equations*, in System Modelling and Optimization, J. Dolezal and J. Fidler Eds., Chapman and Hall, 1996, p. 187-194.

[6] **G. Choudury**, *Fully Discrete Galerkin Approximations of Parabolic Boundary-Value Problems with Nonsmooth Boundary Data*, Numer. Math., Vol 57(1990), p. 179-203.

[7] **G. Choudury**, *Fully Discrete Galerkin Approximations of Parabolic Boundary-Value Problems with Nonsmooth Boundary Data*, Appl. Math. Optim., Vol. 31 (1995), p. 41-55.

[8] **P. G. Ciarlet**, The Finite Element Method for Elliptic Problems, North Holland, 1987.

[9] **K. Eriksson, C. Johnson**, *Adaptive Finite Element Methods for Parabolic Problems IV: A Nonlinear Model Problem*, SIAM J. Numer. Anal. Vol. 32 (1995), p. 1750-1763.

[10] **I. Lasiecka,** *Ritz-Galerkin Approximation of the Time Optimal Boundary Control Problem for Parabolic Systems with Dirichlet Boundary Conditions,* SIAM J. Cont. Optim., Vol. 22 (1984), p. 477-500.

[11] **U. Mackenroth**, *Numerical Solution of Some Parabolic Boundary Control Problems by Finite Elements,* Control Problems for Systems Described by Partial Differential Equations and Applications, I. Lasiecka and R. Triggiani Eds., Lecture Notes in Control and Information Sciences, Vol. 97 (1987), p. 325-335.

[12] **P. Neittaanmäki, D. Tiba**, Optimal Control of Nonlinear Parabolic Systems, Marcel Dekker Eds., 1994.

[13] **R. H. Nochetto, C. Verdi**, *Approximation of Degenerate Parabolic Problems Using Numerical Integration*, SIAM J. Num. Anal., Vol. 25 (1988), p. 784-814.

[14] **A. K. Pani**, *An H^1-Galerkin Mixed Finite Element Method for Parabolic Partial Differential Equations*, SIAM J. Num. Anal., Vol. 35 (1998), p. 712-727.

[15] **J.-P. Raymond, H. Zidani**, *Hamiltonian Pontryagin's Principles for Control Problems Governed by Semilinear Parabolic Equations,* Appl. Math. and Optim., Vol. 39 (1999), no. 2, p. 143–177.

[16] **T. Roubicek**, *A Stable Approximation of a Constrained Optimal Control for Continuous Casting,* Num. Funct. Anal. and Optim., Vol. 13 (1992), p. 487-494.

[17] **W. Rudin**, *Analyse réelle et complexe*, Masson, Paris (1978).

[18] **M. Sibony**, Analyse Numérique III, Itérations et Approximations, Hermann, Paris 1988.

[19] **V. Thomée** Galerkin finite element methods for parabolic problems, Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1997.

[20] **F. Tröltzsch,** *Semidiscrete Ritz-Galerkin Approximation of Nonlinear Parabolic Boundary Control Problems - Strong Convergence of Optimal Controls*, Appl. Math.Optim., Vol. 29 (1994), p. 309-329.

[21] **C. Verdi,** *Optimal Error Estimates for an Approximation of Degenerate Parabolic Problems*, Numer. Funct. Anal. and Optim., Vol. 9 (1987), p. 657-670.

[22] **A. Zenisek**, Nonlinear Elliptic and Evolution Problems and their Finite Element Approximations, Academic Press, 1990.

[23] **H. Zidani**, *Optimal Control Problem for Semilinear Parabolic Equation: Optimality Conditions and Numerical Approximations*, Thèse, Toulouse, 1996.

[24] **J. Zowe and S. Kurcyusz**, *Regularity and stability for the mathematical programming problem in Banach spaces*, Applied Mathematics and Optimization, 5, pp.49–62, 1979.