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

MAITINE 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 $L^\infty$) and state constraints. The convergence analysis is the main objective.

Let $\Omega$ be a bounded convex open subset of $\mathbb{R}^d$ ($d \geq 1$), with Lipschitz boundary $\Gamma$. 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 $a_{jk} \in C^{1+\beta}(\Omega)$ (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:

$$\frac{\partial y}{\partial t} + Ay + f(y) + \partial \kappa(y - \psi) \ni v \quad \text{in } \Omega \times ]0, T[,$$

$$\frac{\partial y}{\partial n} + by = 0 \quad \text{on } \Gamma \times ]0, T[,$$

$$y(0) = y_o \quad \text{in } \Omega,$$  \hspace{1cm} (1.1)

where $y_o \in H^1(\Omega) \cap C(\overline{\Omega})$, $\psi \in L^q(Q)$ 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 $\kappa$ at $r \in \mathbb{R}$ and $\kappa$ 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:

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

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 Orléans, 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 Orléans, France, and INRIA-Rocquencourt, France. Email - Housnaa.Zidani@inria.fr
We make the following additional assumptions:

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

\[
C_o \leq f'(y) \leq \tilde{C}_o, \quad \forall y \in \mathbb{R}.
\]

(H2) \( q > \frac{d}{2} + 1 \).

(H3) \( y_o \in W^{2, q}_0(\Omega) \cap C(\bar{\Omega}) \cap W^{1, q}_0(\Omega) \).

(H4) \( \psi \in L^q(Q) \) is defined everywhere on \( Q \) and \( y_o(x) \geq \psi(x, 0) \) for any \( x \in \Omega \) (compatibility assumption for the initial condition).

(H5) \( F, L, G : \Omega \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \) 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) \( \ell : \Omega \times \mathbb{R} \rightarrow \mathbb{R} \) 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 \((P)\). 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 \((\tilde{P})\) equivalent to \((P)\), with constraints on both the control variable and state variable as coupled state/control constraints. We cannot avoid a first approximation of \((\tilde{P})\) 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 \( 1 \leq s \leq \infty \), the norms in the spaces \( L^s(\Omega) \), \( L^s(\Gamma) \), \( L^s(Q) \), \( L^s(\Sigma) \) 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 \((\cdot, \cdot)_\Omega\) and \((\cdot, \cdot)_{H^1}\) while \((\cdot, \cdot)_{H^1}^*\) denotes the canonical duality pairing between \(H^1(\Omega)\) and \((H^1(\Omega))^*\).

The Hilbert space \(W(0, T; H^1(\Omega), (H^1(\Omega))^*) = \{ y \in L^2(0, T; H^1(\Omega)) \mid \frac{dy}{dt} \in L^2(0, T; (H^1(\Omega))^*) \}\) endowed with its usual norm, will be denoted by \(W(0, T)\). The space \(W^{2,1,q}(Q)\) is the usual space:

\[
W^{2,1,q}(Q) = \{ y \in L^q(Q) \mid \frac{\partial y}{\partial t}, D_xy, D^2_{x,y} \in L^q(\Omega) \}.
\]

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

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

for any \(\varphi \in K\) where

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

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 \(C(\overline{Q}) \cap W^{2,1,q}(Q)\). 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:

\[
\begin{aligned}
\frac{\partial y}{\partial t} + Ay + f(y) &= v + \theta \quad \text{in } Q, \\
\frac{\partial y}{\partial n_A} + by &= 0 \quad \text{on } \Sigma, \\
y(0) &= y_o \quad \text{in } \Omega, \\
y(x, t) - \psi(x, t) &\geq 0 \quad \forall (x, t) \in Q, \\
\theta(x, t) &\geq 0 \quad \text{a.e. } (x, t) \in Q,
\end{aligned}
\]

(1.2a)

ii) There exists \(C_1 = C_1(T, \Omega, d, q, C_o, m_o, \|y_o\|_{C(\overline{\Omega})}) > 0\) such that, for any \((v, \theta) \in L^q(Q) \times L^q(Q)\),

\[
\|y_{v, \theta}\|_{W^{2,1,q}(Q)} + \|y_{v, \theta}\|_{C(\overline{\Omega})} \leq C_1 (1 + \|v\|_{q, Q} + \|\theta\|_{q, Q})
\]

(1.3)

iii) For every \(\varepsilon, K > 0\), there exists \(\alpha > 0\), and \(C'_4 = C'_4(T, \Omega, d, q, C_o, m_o, \|y_o\|_{C(\overline{\Omega})}, \varepsilon, K, \alpha) > 0\) such that, for any \((v, \theta) \in L^q(Q) \times L^q(Q)\), \(y_{v, \theta}\) belongs to \(C^{\alpha, \frac{q}{q-2}}(\overline{\Omega} \times [\varepsilon, T])\) and obeys

\[
\|y_{v, \theta}\|_{C^{\alpha, \frac{q}{q-2}}(\overline{\Omega} \times [\varepsilon, T])} \leq C'_4
\]

(1.4)

\[
\text{if } \|v\|_{q, Q} + \|\theta\|_{q, Q} \leq K.
\]

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 $C(\overline{Q})$ (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 $(v, \theta)$ 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 $(v_n, \theta_n)$. 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 $Q \times [\varepsilon, T]$. Then, there exists $y \in C(\overline{Q}) \cap W^{2.1,q}(Q)$ such that $(y_n)_n$ converges to $y$ weakly in $W^{2.1,q}(Q)$ and strongly in $C(\overline{Q})$. By direct calculations, we can check that $y$ is the solution of (1.2a) corresponding to $(v, \theta)$.

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:

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

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

Minimize $J(y, v)$ subject to:

$$\frac{\partial y}{\partial t} + Ay + f(y) = v + \theta \text{ in } Q, \quad \frac{\partial y}{\partial n_A} + by = 0 \text{ on } \Sigma, \quad y(., 0) = y_0 \text{ in } \Omega \quad (2.2a)$$

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

$$(v, \theta) \in U_{ad} \times \Theta_{ad} \quad (\text{“Pure” control constraints}) \quad (2.2c)$$

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

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

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

where $\tilde{M}$ is a constant which is allowed to be very large and must be greater than the norm of the control $\theta$ 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 $(\tilde{\mathcal{P}})$ when we replace (2.2c) by:

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

In the sequel we do not care about the existence of an optimal solution to $(\mathcal{P})$ (or $(\tilde{\mathcal{P}})$): 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 $(\tilde{y}, \tilde{v}, \tilde{\theta})$. 


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

\[
\frac{\partial y}{\partial t} + Ay + f(y) = g \quad \text{in } Q, \quad \frac{\partial y}{\partial n_A} + by = 0 \quad \text{in } \Sigma, \quad y(0) = y_0 \quad \text{in } \Omega, \tag{3.1}
\]

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

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

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

\[
\forall y, z \in H^1(\Omega) \times H^1(\Omega) \quad \mathcal{A}(y, z) = \int_{\Omega} \sum_{j,k} a_{jk}(x)D_ky(x)D_jz(x) \, dx.
\]

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 \( \Omega \)).

3.1.1. Grid for \( \Omega \). 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:

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

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 \( \eta, \gamma \) such that:

\[
\frac{h_K}{\rho_K} \leq \eta \quad \text{and} \quad \frac{h}{\rho_K} \leq \gamma \quad \text{for all } K \in \mathcal{F}_h \quad \text{and all } h > 0.
\]

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

\textbf{Remark 3.1.} \textit{Since \( \Omega \) is bounded, the discretization parameter \( h \) is necessarily less than some constant \( h_\Omega \) which only depends on \( \Omega \).}

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 \( \Omega \) correspond with the ones of \( K \), and such that the \( (d+1)-\text{th} \) face is the part of \( \Gamma \) limited by the \( d \) other faces. We denote by \( \tilde{\mathcal{F}}_h \) the family composed by these simplices \( \tilde{K} \) and the simplices contained inside \( \Omega \). Hence we have: \( \overline{\Omega} = \bigcup_{K \in \tilde{\mathcal{F}}_h} K \).

To any such triangulation \( \tilde{\mathcal{F}}_h \) we associate the finite dimensional following space

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

Let \( \{x_j\}_{j=1}^{N_h} \) be the set of all nodes of \( \tilde{\mathcal{F}}_h \) on \( \overline{\Omega}_h \). Let \( \varphi_j(\cdot) \) 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 \( a(\cdot, \cdot) \) defined on \( H^1(\Omega) \) by

\[
a(w, z) := \mathcal{A}(w, z) + C_o \langle w, z \rangle_\Omega + b \langle w, z \rangle_\Gamma \quad \forall w, z \in H^1(\Omega),
\]
where $C_o$ is the constant given in (H1). We can suppose, without loss of generality [15], that
\[
a(z, z) \geq \frac{m_a}{2} \|z\|^2_{H^1(\Omega)} \quad \text{for any } z \in H^1(\Omega).
\] (3.3)

Let $E_h$ be the operator of $H^1(\Omega)$ on $Y_h$ which associates to any $z \in H^1(\Omega)$ the unique element $E_hz$ of $Y_h$ such that for all $\chi \in Y_h$:
\[
\sum_{i,j=1}^{d} \int_{\Omega} a_{ij}(x) D_j(z - E_hz) D_l \chi \, dx + \int_{\Omega} C_o(z - E_hz) \chi \, dx + \int_{\Gamma} b(z - E_hz) \chi \, ds = 0.
\] (3.4)

We observe that this operator satisfies:
\[
a(E_hz, \chi) = a(z, \chi) \quad \text{for any } \chi \in Y_h \text{ and any } z \in H^1(\Omega).
\]

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 $\Omega$ such that for every $h < h_\Omega$, we have
\[
\|E_hz - z\|_{2,\Omega} \leq C_2 \, h \, \|z\|_{H^1(\Omega)} \quad \forall z \in H^1(\Omega).
\]

**3.1.2. Grid for $\Gamma$.** We denote $B_h$ the triangulation of $\Gamma$, inducted by the triangulation $\tilde{F}_h$.

**3.1.3. Partition of $[0, T]$.** Let $N$ be a positive integer. We consider the uniform partition of $[0, T]$ defined by:
\[
t_0 = 0 < t_1 < t_2 < \cdots < t_N = T,
\]
\[
t_i = i\tau \quad (i = 0, \ldots, N) \quad \text{where } \tau := \frac{T}{N}.
\]

For any $i = 1, \ldots, N$, we denote by $\chi_i$ 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:
\[
y^i_h(\cdot) := \sum_{j=1}^{N_e} Y^i_j \varphi_j(\cdot) \in Y_h, \quad j = 1, \ldots, N_e
\] (3.5)

where $Y^i_j = y(x_j, t_i)$ (we remark that $Y^i_j = y(x_j, t_i) \in \mathbb{R}$ is well defined since $y \in C(\mathbb{Q})$.)

Now we derive the discrete analog of (3.2) by means of which we shall define the approximate solution:

**Theorem 3.1.** The system (3.6) admits a unique solution $\{y^i_h\}_{i=1}^N$ 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 \( y_\delta \) for (3.1) using:

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

The function \( y_\delta \) 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 \( g \in L^3(Q) \) satisfying \( \|g\|_{0,Q} \leq M_o \), the solution \( \{y_h^i\}_{i=0,\ldots,N} \) to (3.6) corresponding to \( g \) satisfies:

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

**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 \( \{1,\ldots,N\} \)); we obtain

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

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

which can also be written as follows:

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

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

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

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

\[
K_1 \sum_{i=1}^{i_o} \int_{t_{i-1}}^{t_i} \|g(t)\|_{2,\Omega} \|y_h^i\|_{H^1(\Omega)} \, dt + \sum_{i=1}^{i_o} \tau |f(0)| \|y_h^i\|_{H^1(\Omega)}.
\]
where $K_1$ only depends on $\Omega$ and $d$. With the identity

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

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

$$2m_o \sum_{i=1}^{l_o} \tau \|y_h^i\|^2_{H^1(\Omega)} + \|y_h^o\|^2_{\mathcal{L},\Omega} + \sum_{i=1}^{l_o} \|y_h^i - y_h^{i-1}\|^2_{\mathcal{L},\Omega} \leq$$

$$\|y_h^o\|^2_{\mathcal{L},\Omega} + m_o \sum_{i=1}^{l_o} \tau \|y_h^i\|^2_{H^1(\Omega)} + \sum_{i=1}^{l_o} \|y_h^i - y_h^{i-1}\|^2_{\mathcal{L},\Omega} \leq K_3,$$

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

$$\sum_{i=1}^{l_o} \tau \|y_h^i\|^2_{H^1(\Omega)} + \|y_h^o\|^2_{\mathcal{L},\Omega} + \sum_{i=1}^{l_o} \|y_h^i - y_h^{i-1}\|^2_{\mathcal{L},\Omega} \leq K_3,$$

for any $i_o \in \{1, \ldots, N\}$, 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 $h_\Omega$. This completes the proof. 

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

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

**Proof -** We set $e = y - y_h$, $y^i = y(t_i)$, $y_h^i = y_h(t_i)$, $e^i = y^i - y_h^i$, and $I_i = [t_{i-1}, t_i]$ for $i \geq 1$.

We also set $\partial x^i = \frac{z^i - z_i-1}{\tau}$, $\tau^i := \frac{1}{\tau} \int_{I_i} z(t) dt$ for $i \geq 1$, and $\tau^0 = z_0$ for $i = 0$. In particular we have $\tau^i = \tau^i - y_h^i$. In what follows, $C$ is a constant independent of $h, \tau$ which may depend on $M_o, m_o, C_1, C_2, C_3$.

Setting $\varphi = \tau E_b^\varphi$ in (3.6) yields:

$$\tau \langle \partial y_h^i, E_b^\varphi \rangle_\Omega + \tau \left( A(y_h^i, E_b^\varphi) + \langle f(y_h^i), E_b^\varphi \rangle_\Omega + \langle by_h^i, E_b^\varphi \rangle_\Gamma \right) = \tau \langle \varphi^i, E_b^\varphi \rangle_\Omega.$$  \hspace{1cm} (3.8)

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

$$\frac{d}{dt} \langle y(t), z \rangle_\Omega + A(y(t), z) + \langle f(y(t)), z \rangle_\Omega + \langle by(t), z \rangle_\Gamma = \langle g(t), z \rangle_\Omega.$$  \hspace{1cm} (3.9)

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

$$\tau \langle \partial y^i, \varphi^i \rangle_\Omega + \tau \left( A(\varphi^i, \varphi^i) + \langle \varphi^i, \varphi^i \rangle_\Omega + \langle b \varphi^i, \varphi^i \rangle_\Gamma \right) = \tau \langle \varphi^i, \varphi^i \rangle_\Omega.$$
We subtract (3.8) from (3.9) and sum up from $i = 1$ to $i = i_o$ ($1 \leq i_o \leq N$):

\[
(I) + (II) := \sum_{i=1}^{i_o} \tau (\partial(y^i - y_{h}^{i}), \tau^i)_{\Omega} + \sum_{i=1}^{i_o} \tau (a(y^i, \tau^i) - a(y_{h}^{i}, E_h \tau^i))
\]

\[
= \sum_{i=1}^{i_o} \tau (\partial y_{h}^{i}, (E_h - I) \tau^i)_{\Omega} + \sum_{i=1}^{i_o} \tau (C_o(y_{h}^{i}, \tau^i) - C_o(y_{h}^{i}, E_h \tau^i))_{\Omega}
\]

\[
+ \sum_{i=1}^{i_o} \tau \left( (f(y_{h}^{i}), E_h \tau^i)_{\Omega} - (f^i, \tau^i)_{\Omega} \right) + \sum_{i=1}^{i_o} \tau (\partial(y^i - \bar{y}^i), \tau^i)_{\Omega} =: (III) + (IV) + (V) + (VI).
\]

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

\[
(I) = \sum_{i=1}^{i_o} \tau (\partial(y^i - y_{h}^{i}), \tau^i)_{\Omega}
\]

\[
= \sum_{i=1}^{i_o} \tau (\partial \tau^i, \tau^i)_{\Omega} + \sum_{i=1}^{i_o} \tau (\partial(y^i - \bar{y}^i), \tau^i)_{\Omega} =: (I)_1 + (I)_2.
\]

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

\[
2 \sum_{i=1}^{i_o} \tau (\tau^i - \tau^{i-1}, \tau^i)_{\Omega} = \frac{1}{2} \| \tau^{i} \|^2_{2, \Omega} - \frac{1}{2} \| \tau^{0} \|^2_{2, \Omega} + \frac{1}{2} \sum_{i=1}^{i_o} \| \tau^{i} - \tau^{i-1} \|^2_{2, \Omega}.
\]

we obtain:

\[
(I)_1 \geq C \| \tau^{i} \|^2_{2, \Omega} - C \| \tau^{0} \|^2_{2, \Omega} + C \sum_{i=1}^{i_o} \| \tau^{i} - \tau^{i-1} \|^2_{2, \Omega}.
\]

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

\[
\| y(t) - \bar{y} \|_{2, \Omega} \leq \sqrt{\tau} \| \frac{\partial y}{\partial t} \|_{L^2(\Omega \times I)} \quad \forall i \in \{1,..,N\} \text{ and } t \in T_i.
\]

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

\[
| (I)_2 | = | (y^o - \bar{y}^o, \tau^{i_o})_{\Omega} - \sum_{i=2}^{i_o} (y^{i-1} - \bar{y}^{i-1}, \tau^i - \tau^{i-1})_{\Omega} |
\]

\[
\leq C \sqrt{\tau} \| \frac{\partial y}{\partial t} \|_{L^2(\Omega \times I)} \| \tau^{i_o} \|^2_{2, \Omega} + C \sum_{i=2}^{i_o} \sqrt{\tau} \| \frac{\partial y}{\partial t} \|_{L^2(\Omega \times I)} \| \tau^i - \tau^{i-1} \|^2_{2, \Omega}
\]

\[
\leq \epsilon \| \tau^{i_o} \|^2_{2, \Omega} + \epsilon \sum_{i=1}^{i_o} \| \tau^i - \tau^{i-1} \|^2_{2, \Omega} + \frac{C}{\epsilon} \tau \| \frac{\partial y}{\partial t} \|_{L^2(\Omega \times I)}^2
\]

\[
\leq \epsilon \| \tau^{i_o} \|^2_{2, \Omega} + \epsilon \sum_{i=1}^{i_o} \| \tau^i - \tau^{i-1} \|^2_{2, \Omega} + \frac{C}{\epsilon} \tau^2 C_1^2 (1 + M_o)^2
\]

\[
\leq \epsilon \| \tau^{i_o} \|^2_{2, \Omega} + \epsilon \sum_{i=1}^{i_o} \| \tau^i - \tau^{i-1} \|^2_{2, \Omega} + \frac{C}{\epsilon} \tau
\]
where $\varepsilon > 0$ will be chosen at the end of estimates. From the definition of $E_h$, it follows:

\[
(II) = \sum_{i=1}^{i_0} \tau \left( a(\tilde{\gamma}^i, \tilde{v}^i) - a(y^i_h, E_h \tilde{v}^i) \right) = \sum_{i=1}^{i_0} \tau a(\tilde{v}^i, E_h \tilde{v}^i) = \sum_{i=1}^{i_0} \tau a(\tilde{v}^i, \tilde{v}^i)
\]

\[
\geq C \sum_{i=1}^{i_0} \tau \| \tilde{v}^i \|^2_{H^1(\Omega)}.
\]

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

\[
|III| = \left| \sum_{i=1}^{i_0} \tau \langle \partial \tilde{y}^i_h, (E_h - I) \tilde{v}^i \rangle \right| \leq C_2 \sum_{i=1}^{i_0} \| y^i_h - y^i_h \|^2_{L^2} \| (E_h - I) \tilde{v}^i \|_{L^2} \| \tilde{v}^i \|_{H^1(\Omega)}
\]

\[
\leq \varepsilon \tau \sum_{i=1}^{i_0} \| \tilde{v}^i \|^2_{H^1(\Omega)} + \frac{C_2 C_3}{4\varepsilon} \frac{h^2}{\tau}
\]

\[
\leq \varepsilon \tau \sum_{i=1}^{i_0} \| \tilde{v}^i \|^2_{H^1(\Omega)} + \frac{C}{\varepsilon} \frac{h^2}{\tau}.
\]

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

\[
|IV| \leq |C_0 \sum_{i=1}^{i_0} \tau \langle \tilde{y}^i_h - y^i_h, \tilde{v}^i \rangle \|_{\Omega}| + \sum_{i=1}^{i_0} \tau \langle (I - E_h) \tilde{v}^i \rangle_{\Omega}|
\]

\[
\leq |C_0 \sum_{i=1}^{i_0} \tau \| \tilde{v}^i \|^2_{L^2} + \| C_0 \sum_{i=1}^{i_0} \tau \| y^i_h \|_{L^2} \| (I - E_h) \tilde{v}^i \|_{L^2} |
\]

\[
\leq \sum_{i=1}^{i_0} \tau \| \tilde{v}^i \|^2_{L^2} + \sum_{i=1}^{i_0} \tau \| \tilde{v}^i \|^2_{L^2} + \varepsilon \sum_{i=1}^{i_0} \tau \| \tilde{v}^i \|^2_{L^2} + C \varepsilon h^2 \sum_{i=1}^{i_0} \tau \| y^i_h \|^2_{L^2} |
\]

\[
\leq \sum_{i=1}^{i_0} \tau \| \tilde{v}^i \|^2_{L^2} + \varepsilon \sum_{i=1}^{i_0} \tau \| \tilde{v}^i \|^2_{L^2} + C \varepsilon h^2.
\]

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:

\[
|V| \leq \left| \sum_{i=1}^{i_0} \tau \langle f(y^i_h), (I - E_h) \tilde{v}^i \rangle_{\Omega} \right| + \left| \sum_{i=1}^{i_0} \tau \langle f(y^i_h), (I - E_h) \tilde{v}^i \rangle_{\Omega} \right|
\]

\[
\leq \sum_{i=1}^{i_0} \left| \int_t^{t_{i_0}} \| f(y^i_h) - f(y(t)) \| \| \tilde{v}^i \|_{L^2} \right| + \sum_{i=1}^{i_0} \tau \langle f(y^i_h), (I - E_h) \tilde{v}^i \rangle_{\Omega} |
\]

\[
\leq \tilde{C}_0 \sum_{i=1}^{i_0} \left| \int_t^{t_{i_0}} \| f(y^i_h) - f(y(t)) \| \| \tilde{v}^i \|_{L^2} + C \sum_{i=1}^{i_0} \tau \| y^i_h \|_{L^2} + 1 \| (I - E_h) \tilde{v}^i \|_{L^2} \right|
\]

\[
\leq \frac{\varepsilon}{\tilde{C}} \sum_{i=1}^{i_0} \tau \| \tilde{v}^i \|^2_{L^2} + \frac{C}{\varepsilon} \int_0^{t_{i_0}} \left| \int_t^{t_{i_0}} \| f(y^i_h) - f(y(t)) \| \| \tilde{v}^i \|_{L^2} + C \sum_{i=1}^{i_0} \tau \| y^i_h \|_{L^2} + 1 \| (I - E_h) \tilde{v}^i \|_{L^2} \right|
\]

\[
\leq \frac{\varepsilon}{\tilde{C}} \sum_{i=1}^{i_0} \tau \| \tilde{v}^i \|^2_{L^2} + \frac{C}{\varepsilon} \varepsilon h^2 + \frac{C}{\varepsilon} \sum_{i=1}^{i_0} \tau \| \tilde{v}^i \|^2_{L^2}.
\]
Since \( \|g\|_{q, \Omega} \leq M \), we have:

\[
\|v\|_{2, \Omega} + \sum_{i=1}^{i_0} \|v_i - v_{i-1}\|_{2, \Omega} \leq M \left( \int_{t_{i-1}}^{t_i} \|\dot{y}(t)\|_{2, \Omega}^2 \, dt \right)^{1/2} + \epsilon \sum_{i=1}^{i_0} \tau \|\varepsilon\|_{H^1(\Omega)}^2.
\]

To summarize, we have proved the following estimate:

\[
\|v\|_{2, \Omega} + \sum_{i=1}^{i_0} \|v_i - v_{i-1}\|_{2, \Omega} + \sum_{i=1}^{i_0} \tau \|\varepsilon\|_{H^1(\Omega)}^2 \leq C \left( \tau + \frac{h^2}{\tau} + h^2 \right) + C \int_{t_{i-1}}^{t_i} \|\varepsilon(t)\|_{2, \Omega}^2 \, dt + C \sum_{i=1}^{i_0} \tau \|\varepsilon\|_{2, \Omega}^2.
\]

Let us observe that for \( i \in \{1, ..., N\} \), we have:

\[
\|v(t)\|_{2, \Omega}^2 \leq 2\|y(t) - \mathcal{Y}\|_{2, \Omega}^2 + 2\|\varepsilon - y_h\|_{2, \Omega}^2.
\]

Now we set \( C \varepsilon = \frac{1}{2} \) 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 \( i_0 \in \{1, ..., N\} \), we have:

\[
\|v(t)\|_{2, \Omega}^2 + \sum_{i=1}^{i_0} \tau \|\varepsilon\|_{H^1(\Omega)}^2 \leq C (\tau + \frac{h^2}{\tau} + h^2) + C \int_{t_{i-1}}^{t_i} \|\varepsilon(t)\|_{2, \Omega}^2 \, dt + C \sum_{i=1}^{i_0} \tau \|\varepsilon\|_{2, \Omega}^2.
\]  

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

\[
\|v\|_{2, \Omega} + \sum_{i=1}^{i_0} \tau \|\varepsilon\|_{H^1(\Omega)}^2 \leq C (\tau + \frac{h^2}{\tau} + h^2) + C \sum_{i=1}^{i_0} \tau \|\varepsilon\|_{2, \Omega}^2.
\]

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

\[
\sup_{t \in [0, 1]} \|v(t)\|_{2, \Omega} + \sum_{i=1}^{i_0} \tau \|\varepsilon\|_{H^1(\Omega)}^2 \leq C (\tau + \frac{h^2}{\tau} + h^2).
\]

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

\[
a_n + b_n \leq c_n + \sum_{k=0}^{n-1} \tau a_k \quad \forall n \geq 1,
\]

with \(a_0 + b_0 \leq c_0\). Then, the following inequality holds:

\[
a_n + b_n \leq c_n \exp(\tau n) \quad \forall n.
\]

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


4.1. 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 \(C(\overline{\Omega})\), 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 \(\delta\) is small. Therefore, we are obliged to relax these constraints with respect to some \(\varepsilon\) 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:

\[
(\mathcal{P}^\varepsilon) \quad \inf \left\{ J(y,v,\theta) \mid (y,v,\theta) \in \mathcal{C}(\overline{Q}) \times V_{ad}, \ (y,v,\theta) \text{ satisfies } (2.2a) \text{ and } \int_Q (y - \psi)^- \ dx \ dt \leq \varepsilon, \left( \int_Q (F^+(x,t,y(x,t)))^2 \ dx \ dt \right)^{1/2} \leq \varepsilon \right\}.
\]

Here \(F^+\) denotes the nonnegative part of \(F\) (and \(y^-\) the negative part of \(y\)). Since \(V_{ad}\) is bounded in \(L^2(Q) \times L^2(Q)\) and the state equation is linear with respect to the control variables, the problem \((\mathcal{P})\) is weakly stable in the following sense.

Proposition 4.1. We have \(\lim_{\varepsilon \to 0} (\inf (\mathcal{P}^\varepsilon)) = \inf (\mathcal{P}) = \inf (\mathcal{P})\).

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 \((\mathcal{P})\) is feasible for \((\mathcal{P}^\varepsilon)\)). Let \((y_n, v_n, \theta_n)_{n \geq 1}\) be a minimizing sequence of \((\mathcal{P}^\varepsilon)\). The sequence \((v_n, \theta_n)_{n \geq 1}\) is included in \(V_{ad}\) and is bounded in \(L^q \times L^q(Q)\); therefore there exists a subsequence, still denoted by \((v_n, \theta_n)_{n \geq 1}\), 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 \(C(\overline{Q})\) of \((y_n)_n\) to the solution \(y_\varepsilon\) of (2.2a)
corresponding to \((v_z, \theta_z)\). On the other hand, since \((y_n, v_n, \theta_n)\) is feasible for \((\mathcal{P}^c)\), we have:

\[
\int_Q ((y_n - \psi)^{-})^2 \, dx \, dt \leq \varepsilon, \quad \frac{1}{Q} \left( F^+(x, t, y_n(x, t)) \right) \, dx \, dt \leq \varepsilon, \quad \left( \int_Q (y_n - \psi) \theta_n \, dx \, dt \right)^2 \leq \varepsilon,
\]

for every \(n \geq 1\). By passing to the limit, when \(n \to \infty\), we deduce that

\[
\int_Q ((y_{\varepsilon} - \psi)^{-})^2 \, dx \, dt \leq \varepsilon, \quad \frac{1}{Q} \left( F^+(x, t, y_{\varepsilon}(x, t)) \right) \, dx \, dt \leq \varepsilon, \quad \left( \int_Q (y_{\varepsilon} - \psi) \theta_{\varepsilon} \, dx \, dt \right)^2 \leq \varepsilon,
\]

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

\[
\inf(\mathcal{P}^c) \leq J(y_{\varepsilon}, v_{\varepsilon}) \leq \liminf_{n} J(y_n, v_n) = \inf(\mathcal{P}^c).
\]

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

**Step 2:** Now we study the asymptotic behavior of \((y_{\varepsilon}, v_{\varepsilon}, \theta_{\varepsilon})\). Again, since \(V_{ad}\) is bounded in \(L^3(Q) \times L^2(Q)\), 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 \(L^3(Q)\), and \(\theta_{\varepsilon_j} \to \bar{\theta}\) in \(L^2(Q)\) when \(j \to \infty\). If we denote by \(y_{\varepsilon_j}\) 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 \(\bar{Q}\). By the same arguments as in Step 1, we can check that \((\bar{y}, \bar{v}, \bar{\theta})\) is feasible for \((\bar{\mathcal{P}})\). Using again the convexity of \(J\) with respect to the second variable and the feasibility of \((\bar{y}, \bar{v}, \bar{\theta})\) for \((\bar{\mathcal{P}})\), we get

\[
\inf(\bar{\mathcal{P}}) \leq J(\bar{y}, \bar{v}) \leq \liminf_{j \to \infty} J(y_{\varepsilon_j}, v_{\varepsilon_j}) = \lim_{j \to \infty} \inf(\mathcal{P}^{\varepsilon_j}) \leq \inf(\bar{\mathcal{P}}),
\]

which concludes the proof. \(\blacksquare\)

**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 \(\tau\) (space and time discretization steps), we set \(\delta = (h, \tau)\) and we consider the space \(\mathcal{V}_{\delta}\) defined by:

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

Any function \(v_{\delta}\) of \(\mathcal{V}_{\delta}\) may be written as

\[
v_{\delta}(x, t) = \sum_{i=1}^{N} \sum_{j=1}^{N_x} V^i_j \chi_\delta(t) \chi_{K_j}(x) \tag{4.2}
\]

where \(\chi_{K_j}\) is the characteristic function of \(K_j\) and \(V^i_j \in \mathbb{R}\). Any function \(v\) of \(L^q(Q)\) can be approximated by \(v_{\delta}\) as in \((4.2)\) where

\[
V^i_j = \frac{1}{|Q_{ij}|} \int_{Q_{ij}} v(x, t) \, dx \, dt \quad \text{for any } i = 1, \ldots, N, \quad j = 1, \ldots, N_x,
\]
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 \( v_\delta \in V_\delta \) 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 \( Q_{\delta_k} = Q_{i_j, i_k} \) 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

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

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

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

and we obtain

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

We conclude with the Lebesgue dominated convergence theorem.

The end of the proposition follows from the uniqueness of the weak limit. \( \square \)

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

\[
\begin{cases}
  y_\delta(x,t) = \sum_{i=1}^N \chi_i(t) \, y_i^h(x) \\
  \quad \forall (x,t) \in \Omega \times [0,T],
  \quad y_\delta(\cdot,0) = E_h y_0.
\end{cases}
\]

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

\[
V^i = (V_j^i)_{j=1,...,N_e} \in \mathbb{R}^{N_e}, \quad V = (V^i)_{i=1,...,N} \in \mathbb{R}^{N_{tot}},
\]

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

We define now the well known mass and stiffness matrices:

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

With these notations, we have

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

(4.4)
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

$$
\begin{cases}
\text{For } i = 1, \ldots, N, \text{ and } j = 1, \ldots, N_e \\
\frac{1}{\tau} \sum_{k=1}^{N_e} \left( \int_{\Omega} \varphi_k \varphi_j \, dx \right) (Y^i_k - Y^{i-1}_k) + \sum_{k=1}^{N_e} A(\varphi_k, \varphi_j) Y^i_k \\
+ \int_{\Omega} f(y_i^k) \varphi_j \, dx + \sum_{k=1}^{N_e} \left( \int_{\Gamma} b \varphi_k \varphi_j \, ds \right) Y^i_k = \frac{1}{\tau} \int_{t_{i-1}}^{t_i} \int_{\Omega} g(\cdot, t) \varphi_j \, dx \, dt \\
y_0^i := E_h y_0.
\end{cases}
$$

Let us set $[A] = [A(\varphi_k, \varphi_j)]_{1 \leq k, j \leq N_e}$ and $[B] = [\int_{\Gamma} b \varphi_k \varphi_j \, ds]_{1 \leq k, j \leq N_e}$, and for every $i = 1, \ldots, N$

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

Then the above relation is equivalent to

$$
\begin{cases}
\text{For } i = 1, \ldots, N \\
([M] + \tau[A] + \tau[B]) Y^i + \tau \Phi(Y^i) = [M] Y^{i-1} + H_i(g), \\
y_0^i := E_h y_0.
\end{cases}
$$

(4.5)

Now we may define the discretized problem corresponding to $(P^\varepsilon)$. For any $(v_\delta, \theta_\delta)$ in $V_\delta \times V_\delta$ 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

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

and

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

For any $\varepsilon > 0$, we define

$$
(P^\varepsilon_\delta)
$$

$$
\begin{align*}
\text{minimize} & \quad J(y_\delta, v_\delta, \theta_\delta) \\
\text{with} & \quad (v_\delta, \theta_\delta) \in V_{ad, \delta} \text{ and } y_\delta = y_\delta(v_\delta, \theta_\delta) \\
\int_Q ((y_\delta - \psi_\delta)^2 \, dx \, dt \leq \varepsilon, \\
\int_Q (F^+(x, t, y_\delta))^2 \, dx \, dt \leq \varepsilon,
\end{align*}
$$

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 $(P^\varepsilon_\delta)$ 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 \( D^\delta \) of \( (P^\delta) \) is non-empty.

Let \( (\bar{y}, \bar{v}, \bar{\theta}) \) be a solution of \( (P) \) 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

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

Theorem 3.2 yields

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

and by theorem 1.2 we conclude that \( \|y_\delta - \bar{y}\|_{L^\infty(L^2)} \rightarrow 0 \) as \( \delta = (h, \tau) \rightarrow 0 \) with \( \frac{h}{\tau} = O(1) \).

Since \( \bar{y} \) satisfies the constraints:

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

the function \( y_\delta \) satisfies

\[
\int_Q ((y_\delta - \psi)^-)^2 \, dx \, dt \leq \varepsilon, \quad \int_Q (F^+(x, t, y_\delta))^2 \, dx \, dt \leq \varepsilon, \quad \left( \int_Q \theta_\delta (y_\delta - \psi_\delta) \, dx \, dt \right)^2 \leq \varepsilon,
\]

as soon as \( \delta = (h, \tau) \) is small enough, say \( \delta \leq \delta^* \) where \( \delta^* \) depends on \( \varepsilon \). Therefore, \( D^\delta \) is non-empty for such \( \delta \).

Since \( D^\delta \) 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 \( (P^\delta) \) to the solution of the continuous problem \( (P^\varepsilon) \).

**Theorem 4.1.** We fix \( \varepsilon > 0 \). If \( (v_\delta^\varepsilon, \theta_\delta^\varepsilon)_{\delta \leq \delta^*} \) denotes a solution to \( (P^\delta) \), one can extract a subsequence weakly converging towards \( (v_\varepsilon, \theta_\varepsilon) \) in \( L^3(0, T) \times L^3(Q) \), where \( (v_\varepsilon, \theta_\varepsilon) \) is a solution to \( (P^\varepsilon) \). In addition we get

\[
\lim_{\delta \rightarrow 0} \inf (P^\delta) = \inf (P^\varepsilon).
\]

Proof - The sequence \( (v_\delta^\varepsilon, \theta_\delta^\varepsilon)_{\delta \leq \delta^*} \) belongs to \( V_{ad} \) and is bounded in \( L^q(Q) \times L^q(Q) \) (uniformly with respect to \( \delta \)); therefore, there exists a subsequence (still denoted \( (v_\delta^\varepsilon, \theta_\delta^\varepsilon)_{\delta} \)) and \( (v_\varepsilon, \theta_\varepsilon) \in V_{ad} \), such that \( (v_\delta^\varepsilon, \theta_\delta^\varepsilon) \) weakly converges towards \( (v_\varepsilon, \theta_\varepsilon) \) in \( L^3(Q) \times L^3(Q) \) as \( \delta \rightarrow 0 \).

Let \( y_\delta^\varepsilon \) be the solution to (4.3)-(4.5) associated to \( v_\delta^\varepsilon + \theta_\delta^\varepsilon \), and \( y_\varepsilon \) be the solution of (2.2a) associated to \( (v_\varepsilon, \theta_\varepsilon) \). We know with Theorem 3.2, that

\[
\lim_{\delta \rightarrow 0} \|y_\delta^\varepsilon - y(v_\delta^\varepsilon, \theta_\delta^\varepsilon)\|_{L^2(Q)} = 0,
\]

where \( y(v_\delta^\varepsilon, \theta_\delta^\varepsilon) \) is the solution to the (continuous) state-equation (2.2a) associated to \( (v_\delta^\varepsilon, \theta_\delta^\varepsilon) \). In addition,

\[
\lim_{\delta \rightarrow 0} \|y_\delta^\varepsilon - y_\varepsilon\|_{L^2(Q)} = 0,
\]
using the compactness result of Theorem 1.2.

Hence $y_\delta^\epsilon$ strongly converges towards $y_\epsilon$ in $L^2(Q)$ (and even in $L^\infty(0,T;L^2(\Omega))$). Since $(y_\delta^\epsilon, v_\delta^\epsilon, \theta_\delta^\epsilon)$ satisfies the constraints of the problem ($P^\epsilon_\delta$), by passing to the limit when $\delta$ tends to 0 and taking the Lipschitz continuity of $F$ and the continuity of $\psi$ into account (see (H4)-(H5)), we obtain

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

and

$$\left( \int_Q (y_\epsilon - \psi) \theta_\epsilon dx dt \right)^2 \leq \epsilon.$$

Therefore $(y_\epsilon, v_\epsilon, \theta_\epsilon)$ is feasible for $(P^\epsilon)$. 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 $\epsilon$ 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 $D^\epsilon_\delta$ is non empty (i.e. the discretized feasible domain corresponding to $\epsilon = 0^+$), that is, for example if we precisely assume

\[ (H6) \quad \forall \delta \leq \delta^* \quad \exists (y_\delta^0, v_\delta^0, \theta_\delta^0) \in D^\epsilon_\delta. \]

5. **Conclusion.** Our purpose was to compute the solution to problem $(P)$. 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 $(P^\epsilon_\delta)$. It is still a significant challenge because of the (bilinear) complementarity constraint

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

that can be (formally) written

$$\Theta^T M (Y - \Psi) \leq \epsilon,$$

where $\Theta, \Psi$ 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 complementarity constraint. Therefore, the convergence analysis is not obvious. This work will be developed in a forthcoming paper, where we will present numerical examples.

**REFERENCES**


18


