Pontryagin's Principle for State-Constrained Control Problems Governed by Parabolic Equations with Unbounded Controls.

J. P. Raymond, H. Zidani

Université Paul Sabatier, UMR CNRS MIP, 31062 Toulouse cedex France

Abstract

This paper deals with optimal control problems governed by semilinear parabolic equations with pointwise state constraints and unbounded controls. Under some strong stability assumption, we obtain necessary optimality conditions in the form of a Pontryagin's minimum principle in qualified form. A Pontryagin's principle in nonqualified form is also proved without any stability condition.

Keywords. Optimal control, nonlinear boundary controls, semilinear parabolic equations, state constraints, Pontryagin's minimum principle, unbounded controls.

1 Introduction.

This article concerns control problems for the following parabolic system:

$$\frac{\partial y}{\partial t} + Ay + f(x,t,y) = 0 \quad \text{in } Q, \quad \frac{\partial y}{\partial n_A} + g(s,t,y,v) = 0 \quad \text{on } \Sigma, \quad y(0) = w \text{ in } \Omega, \tag{1}$$

where $\Omega \subset \mathbb{R}^N$, $Q = \Omega\times]0,T[$, $\Sigma = \Gamma\times]0,T[$, $T > 0$, v is a boundary control, w is a control of the initial condition, A is a second order elliptic operator. The control variables, v, w, and the state variable y must satisfy constraints of the form

$$v \in V_{ad} \subset L^\sigma(\Sigma), \qquad w \in W_{ad} \subset C(\overline{\Omega}),$$

$$\phi(y) \in \mathcal{C} \tag{2}$$

where ϕ is a continuous mapping from $C(\overline{Q})$ into $C(\overline{D})$, $\mathcal{C} \subset C(\overline{D})$ is a closed convex subset with nonempty interior, and $\overline{D}$ is a nonempty compact subset of $\overline{Q}$. The control problem is

$$(P) \qquad \inf\{J(y,v,w) \mid y \in W(0,T) \cap C(\overline{Q}), (v,w) \in V_{ad} \times W_{ad}, (y,v,w) \text{ satisfies } (1),(2)\},$$

where the cost functional is defined by

$$J(y,v,w) = \int_Q F(x,t,y(x,t))\,dx\,dt + \int_\Sigma G(s,t,y(s,t),v(s,t))\,ds\,dt + \int_\Omega L(x,y(x,T),w(x))dx. \tag{3}$$

We are mainly interested in optimality conditions for such problems, in the form of Pontryagin's principles. The existence of optimal solutions for (P) is a priori supposed.

In recent years there has been growing interest in optimality conditions for state-constrained control problems governed by partial differential equations (or variational inequalities). This is reflected by an important number of papers on this subject. For convex control problems we refer to [1], [4], [5], [7], [31]. In the case of nonconvex control problems the method of Lagrange multipliers provides optimality conditions, both for bounded and for unbounded controls [3], [11], [37], [32], [33]. When no qualification condition is assumed, optimality conditions are obtained in nonqualified form (optimality conditions of Fritz John type). To get optimality conditions in qualified form with a Lagrange multiplier theorem, a qualification condition such as the Zowe-Kurcyusz regularity condition is needed. (In many problems, this regularity condition corresponds to a Slater type qualification condition, see [37], [32].)

Another method proceeds by penalizing the state constraints and then characterizing optimal solutions of the original problem as ε-solutions of the penalized problems. The characterization of ε-solutions is carried out thanks to the Ekeland's variational principle. By this method optimality conditions are obtained in the form of Pontryagin principles, which are in general more precise than optimality conditions deduced from Lagrange multiplier theorems. Moreover assumptions on the data of the problems (differentiability assumptions, convexity requirement,...) are less restrictive than those necessary for Lagrange multiplier theorems. To the best of our knowledge, except in [23], this method has so far been used only for problems with bounded controls (bounded in time [28], [29], [30] or in space and time [12], [21], [22], [24]).
There is a fundamental reason for this limitation. When we apply Ekeland's principle to obtain Pontryagin's principle, we need a complete metric space, let us say (V_{ad}, d_E) (the space of controls V_{ad}, endowed with the so-called Ekeland's metrics d_E, in order to recover a Pontryagin principle) and a penalized functional $F_\varepsilon(v) = J_\varepsilon(y_v, v)$ (v is the control variable, y_v is the solution of the state equation corresponding to v), which must be lower semicontinuous on (V_{ad}, d_E). To prove this lower semicontinuity property we need some assumptions on J_ε and we have to prove that the mapping $\mathcal{T} : v \longmapsto y_v$ is continuous from (V_{ad}, d_E) into a Banach space Z (which depends on the problem considered). The continuity of $\mathcal{T}$ depends on regularity results for the state equation. In the problems studied in the articles mentioned above, V_{ad} is a subset of some Lebesgue space L^σ with $1 \leq \sigma \leq \infty$, and (thanks to regularity results for partial differential equations) it can be proved that the mapping $\mathcal{T}$ is continuous from L^σ into Z for every $\sigma > \bar{\sigma}$, where $\bar{\sigma}$ depends on the state equation.
If V_{ad} is bounded in L^∞ and if a sequence of controls converges for the Ekeland metric, it can easily be proved that this sequence still converges for the topology of L^σ for any $\sigma > \bar{\sigma}$. Therefore in this case, the mapping $\mathcal{T}$ is continuous and Ekeland's principle can be applied.
If V_{ad} is not bounded in L^∞, convergence in the Ekeland metric does not imply convergence in the Lebesgue space norm, moreover (V_{ad}, d_E) is not necessarily complete. This is the reason why, up to now, in the presence of pointwise state constraints, Pontryagin's principles have only been proved for bounded controls (at least for nonconvex problems, indeed for convex problems the optimality conditions deduced from Lagrange multiplier theorems correspond to Pontryagin principles).

Let us stress that the growth conditions on the integrands and the nonlinear terms in the state equations, postulated in ([30], Chapter 4, Hypothesis 2, p. 130) correspond to bounded controls. The same remark is valid for [12].

In [23], Fattorini and Sritharan prove a Pontryagin principle in nonqualified form for control problems of Navier-Stokes equations in which the controls are not necessarily bounded. Their idea is to work with bounded perturbations (see [23], p. 227). Here we consider a control set of the form

$$V_{ad} = \{v \in L^{\sigma}(\Sigma) \mid v(s,t) \in K_V(s,t) \text{ a.e. in } \Sigma\}$$

where K_V is a measurable multimapping with nonempty and closed values in $\mathcal{P}(\mathbb{R})$ (see Section 2). We do not think that the method developed in [23] can be applied to such a control set. Moreover the method developed in [23] deals with Pontryagin principles in nonqualified form and requires some convexity condition on the cost functional (see Hypothesis 2.9 in [23]).

The purpose of this paper is to extend the method based on Ekeland's principle to problems with unbounded controls. In order to explain the main ideas of this extension let us recall the starting point of the method described above. If $\bar{v}$ is an ε^2-solution of the problem

$$(P_\varepsilon) \qquad inf\{F_\varepsilon(v) \mid v \in V_{ad}\},$$

where F_ε and (V_{ad}, d_E) satisfy the assumptions of the Ekeland's principle, then there exists another ε^2-solution v_ε such that

$$d_E(v_\varepsilon, \bar{v}) \leq \varepsilon \quad \text{and} \quad F_\varepsilon(v_\varepsilon) - F_\varepsilon(v) \leq \varepsilon d(v, v_\varepsilon) \qquad \text{for every } v \in V_{ad}.$$

In order to exploit this optimality condition, v is replaced by some perturbation of v_ε. The methods developed in [8], [12], [20], [22], [24], [29], [30] differ both in their choices of F_ε and in their choices of the perturbations.
Pontryagin principles in qualified form are only obtained in [8], [12], by choosing for F_ε a regularization of an exact penalized functional. A Pontryagin principle is then obtained under a strong stability condition. Pontryagin principles in nonqualified form are obtained in [8] under a weak stability condition by a method of spike perturbations. With an other choice for the penalized functional and other kinds of perturbations, Pontryagin principles in nonqualified form are obtained in [22], [24], [29], [30]. In [22] Fattorini and Murphy use a method of multispike perturbations. The type of perturbations used in [29], [38], [24], [12], [13], [30] can be viewed as a generalization of multispike perturbations, which we call diffuse perturbations. In contrast to spike or multispike perturbations, which are precisely localized, a diffuse perturbation is not localized around some points but it is implicitly defined by some relations (see [12], [29], [38], [40], [24]). The existence of diffuse perturbations satisfying relations a priori defined is proved in [27], [28], [24] and in [12] in a constructive manner. To our knowledge this kind of perturbations has been introduced for the first time by Yao [38] and Li [26]. Connections with Lyapunov's convexity theorem or with Uhl's theorem are clarified in [40], p. 1315

and [28]. We here prove that all the relations needed to define a diffuse perturbation can be obtained as a consequence of the Lyapunov convexity theorem.

Preliminary results related the topic were announced in [34]. The metric space used in [34] is different from the one defined in Section 3.2. This is the reason why a convexity condition (assumption A7) is needed in [34] to ensure some semicontinuity property. Thus, the methods of the present paper improve upon those of [34].

The paper is organized as follows. In the next section we formulate the control problem governed by a semilinear parabolic equation and we state the main results: the weak and strong Pontryagin's principles. In Section 3 we give some regularity results for solutions of the state and adjoint equations. In Section 4, we derive some technical results used in Section 5 to prove the main result stated in Section 2.

2 Assumptions and Main results.

Throughout the sequel Ω denotes a bounded open subset of $\mathbb{R}^N$ ($N \geq 2$) of class $C^{2,\beta}$ for some $0 < \beta \leq 1$ (that is, the boundary Γ of Ω is an $(N-1)$-dimensional manifold of class $C^{2,\beta}$ such that Ω lies locally on one side of Γ. A function is of class $C^{2,\beta}$ if it is of class C^2 and if its second order derivatives are Hölder continuous of exponent β). We denote by q, σ positive numbers satisfying

$$q > N/2+1, \quad \sigma > N+1 \quad \text{and} \quad q\sigma + q > qN + 2\sigma.$$

The differential operator A in equation (1) is defined by

$$Ay(x) = -\sum_{i,j=1}^{N} D_i(a_{ij}(x)D_j y(x)),$$

with coefficients a_{ij} belonging to $C^{1,\beta}(\overline{\Omega})$ and satisfying the conditions

$$a_{ij}(x) = a_{ji}(x) \qquad \text{for every } i,j \in \{1,...,N\}, \qquad m_0|\xi|^2 \leq \sum_{i,j=1}^{N} a_{ij}(x)\xi_j\xi_i \tag{4}$$

for all $x \in \overline{\Omega}$ and all $\xi \in \mathbb{R}^N$, with $0 < m_0$, (D_i denotes the partial derivative with respect to x_i). In equation (1), $\frac{\partial y}{\partial n_A}$ is the conormal derivative of y with respect to A, that is

$$\frac{\partial y}{\partial n_A}(s,t) = \sum_{i,j} a_{ij}(s)D_j y(s,t)n_i(s),$$

where $n = (n_1, ..., n_N)$ is the unit normal to Γ outward Ω.

For every $1 \leq \tau \leq \infty$, the norms in the spaces $L^\tau(\Omega)$, $L^\tau(\Gamma)$, $L^\tau(Q)$, $L^\tau(\Sigma)$ will be denoted by $\|\cdot\|_{\tau,\Omega}$, $\|\cdot\|_{\tau,\Gamma}$, $\|\cdot\|_{\tau,Q}$, $\|\cdot\|_{\tau,\Sigma}$. 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 $W(0,T)$. We also set $\overline{\Omega}_0 = \overline{\Omega} \times \{0\}$ and $\overline{\Omega}_T = \overline{\Omega} \times \{T\}$.

2.1 Assumptions.

(A1) - For every $y \in \mathbb{R}$, $f(\cdot, y)$ is measurable on Q. For almost every $(x,t) \in Q$, $f(x,t,\cdot)$ is of class C^1 on $\mathbb{R}$. The following estimates hold:

$$|f(x,t,0)| \leq M_1(x,t), \qquad C_0 \leq f'_y(x,t,y) \leq M_1(x,t)\eta(|y|),$$

where M_1 belongs to $L^q(Q)$, η is a nondecreasing function from $\mathbb{R}^+$ to $\mathbb{R}^+$ and $C_0 \in \mathbb{R}$. (We have denoted by f'_y the partial derivative of f with respect to y, and in the sequel we adopt the same kind of notation for other functions.)

(A2) - For every $(y,v) \in \mathbb{R}^2$, $g(\cdot, y, v)$ is measurable on Σ. For almost every $(s,t) \in \Sigma$ and every $v \in \mathbb{R}$, $g(s,t,\cdot,v)$ is of class C^1 on $\mathbb{R}$. For almost every $(s,t) \in \Sigma$, $g(s,t,\cdot)$ and $g'_y(s,t,\cdot)$ are continuous on $\mathbb{R} \times \mathbb{R}$. The following estimates hold:

$$|g(s,t,0,v)| \leq M_2(s,t) + m_1|v|, \qquad C_0 \leq g'_y(s,t,y,v) \leq (M_2(s,t) + m_1|v|)\eta(|y|),$$

where M_2 belongs to $L^\sigma(\Sigma)$, $m_1 > 0$, C_0 and η are as in $(A1)$.

(A3) - For every $(y,w) \in \mathbb{R}^2$, $L(\cdot, y, w)$ is measurable on Ω. For almost every $x \in \Omega$, $L(x,\cdot)$ is of class C^1 on $\mathbb{R} \times \mathbb{R}$. The following estimate holds:

$$|L(x,y,w)| + |L'_w(x,y,w)| + |L'_y(x,y,w)| \leq M_3(x)\eta(|w|)\eta(|y|),$$

where $M_3 \in L^1(\Omega)$, and η is as in (A1).

(A4) - For every $y \in \mathbb{R}$, $F(\cdot, y)$ is measurable on Q. For almost every $(x,t) \in Q$, $F(x,t,\cdot)$ is of class C^1 on $\mathbb{R}$. The following estimate holds:

$$|F(x,t,y)| + |F'_y(x,t,y)| \leq M_4(x,t)\eta(|y|),$$

where $M_4 \in L^1(Q)$, and η is as in $(A1)$.

(A5) - For every $(y,v) \in \mathbb{R}^2$, $G(\cdot, y, v)$ is measurable on Σ. For almost every $(s,t) \in \Sigma$ and every $v \in \mathbb{R}$, $G(s,t,\cdot,v)$ is of class C^1 on $\mathbb{R}$. For almost every $(s,t) \in \Sigma$, $G(s,t,\cdot)$ and $G'_y(s,t,\cdot)$ are continuous on $\mathbb{R} \times \mathbb{R}$. The following estimate holds:

$$|G(s,t,y,v)| + |G'_y(s,t,y,v)| \leq (M_5(s,t) + m_1|v|^\sigma)\eta(|y|),$$

where $M_5 \in L^1(\Sigma)$, m_1 and η are as in $(A2)$.

(A6) - The set of constraints on v is defined by

$$V_{ad} = \{v \in L^\sigma(\Sigma) \mid v(s,t) \in K_V(s,t) \text{ for a.e. } (s,t) \in \Sigma\},$$

where K_V is a measurable multimapping with nonempty and closed values in $\mathcal{P}(\mathbb{R})$ (that is, the set of all subsets of $\mathbb{R}$). The constraint on the initial condition is $w \in W_{ad}$, where W_{ad} is a closed convex subset of $C(\overline{\Omega})$.

(A7) - In the state constraint (2), ϕ is a mapping of class C^1 from $C(\overline{Q})$ into $C(\overline{D})$, $\overline{D}$ is a nonempty compact subset of $\overline{Q}$, $\mathcal{C} \subset C(\overline{D})$ is a closed convex subset with nonempty interior in $C(\overline{D})$.

In Section 3.1 we recall an existence and uniqueness result in $W(0,T) \cap C(\overline{Q})$ for equation (1), already proved in [35]. Therefore, the state constraint (2) makes sense because the weak solution of (1) is continuous on $\overline{Q}$. Let us give some examples of state constraints described by (2).

Example 2.1. If we choose $\overline{D} = \overline{\Omega} \times \{T\}$, we have a problem with a terminal state constraint. We may consider $\phi(y) = y|_{\overline{D}}$ ($y|_{\overline{D}}$ is the restriction of y to $\overline{D}$) and $\mathcal{C} = \{z \in C(\overline{D}) \mid ||z - y_T||_{C(\overline{D})} \leq \epsilon\}$ where $\epsilon > 0$ and $y_T \in C(\overline{D})$ are given.

Example 2.2. We consider $\phi(y) = \psi(\cdot, y(\cdot))|_{\overline{D}}$, where $\psi \in C(\overline{Q} \times \mathbb{R})$ is such that ψ'_y (the partial derivative of ψ with respect to y) belongs to $C(\overline{Q} \times \mathbb{R})$, $\mathcal{C} = \{z \in C(\overline{D}) \mid z \leq 0\}$ and $\overline{D}$ is any nonempty compact subset of $\overline{Q}$.

2.2 Strong stability assumption.

For $\gamma \geq 0$, we define the subset $\mathcal{C}_\gamma$ of $C(\overline{D})$ by

$$\mathcal{C}_\gamma = \{\varphi \in C(\overline{D}) \mid \inf_{z \in \mathcal{C}} \|\varphi - z\|_{C(\overline{D})} \leq \gamma\},$$

and we consider the perturbed state constraint

$$\phi(y) \in \mathcal{C}_\gamma. \tag{5}$$

We denote by (P_γ) the following problem :

$$(P_\gamma) \qquad \inf\{J(y,v,w) \mid y \in W(0,T) \cap C(\overline{Q}), (v,w) \in V_{ad} \times W_{ad},\ (y,v,w) \text{ satisfies } (1),(5)\}.$$

Observe that (P) is identical to (P_0). Following [8], [12], [13], we say that (P_γ) is strongly stable on the right if there exist $\tilde{\varepsilon} > 0$ and $\tilde{r} > 0$ such that, for every $\gamma' \in [\gamma, \gamma + \tilde{\varepsilon}]$, we have

$$\inf(P_\gamma) - \inf(P_{\gamma'}) \leq \tilde{r}(\gamma - \gamma').$$

With the additional assumption (A8), a Pontryagin principle for (P) may be obtained in qualified form. Some remarks on (A8) are made after Theorem 2.1.

(A8) (P) is strongly stable on the right.

2.3 Statement of the main result.

We define the boundary Hamiltonian function by:

$$H_\Sigma(s,t,y,v,p,\nu) = \nu G(s,t,y,v) - pg(s,t,y,v)$$

for every $(s,t,y,v,p,\nu) \in \Gamma \times [0,T] \times \mathbb{R}^4$. The main result of this paper is the Pontryagin principle for (P), stated in the following theorem.

Theorem 2.1 *If $(A1)-(A7)$ are fulfilled and if $(\bar y, \bar v, \bar w)$ is a solution of (P), then there exists $\bar p \in L^1(0,T;W^{1,1}(\Omega))$, and there exist $\bar\nu \in \mathbb{R}$, $\bar\mu \in \mathcal{M}(\overline{D})$ (the space of Radon measures on $\overline{D}$) and a measurable subset $\widetilde{\Sigma} \subset \Sigma$ such that*

$$(\bar\nu, \bar\mu) \neq 0, \quad \bar\nu \geq 0, \quad \langle \bar\mu, z - \phi(\bar y)\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})} \leq 0 \quad \forall z \in \mathcal{C}, \tag{6}$$

$$\begin{aligned}
&-\frac{\partial \bar p}{\partial t} + A\bar p + f'_y(x,t,\bar y)\bar p = \bar\nu F'_y(x,t,\bar y) + [\phi'(\bar y)^*\bar\mu]|_Q && \text{in } Q,\\
&\frac{\partial \bar p}{\partial n_A} + g'_y(s,t,\bar y,\bar v)\bar p = \bar\nu G'_y(s,t,\bar y,\bar v) + [\phi'(\bar y)^*\bar\mu]|_\Sigma && \text{on } \Sigma,\\
&\bar p(T) = \bar\nu L'_y(x,\bar y(T),\bar w) + [\phi'(\bar y)^*\bar\mu]|_{\overline{\Omega}_T} && \text{in } \Omega,
\end{aligned} \tag{7}$$

$$\mathcal{L}^N(\widetilde{\Sigma}) = \mathcal{L}^N(\Sigma),$$

$$H_\Sigma(s,t,\bar y(s,t),\bar v(s,t),\bar p(s,t),\bar\nu) = \min_{v\in K_V(s,t)} H_\Sigma(s,t,\bar y(s,t),v,\bar p(s,t),\bar\nu) \qquad \text{for all } (s,t)\in\widetilde{\Sigma}, \tag{8}$$

$$\int_\Omega \bar\nu L'_w(x,\bar y(T),\bar w)(\bar w - w)\,dx + \langle \bar p(0) + [\phi'(\bar y)^*\bar\mu]|_{\overline{\Omega}_0}, \bar w - w\rangle_{\mathcal{M}(\overline{\Omega})\times C(\overline{\Omega})} \leq 0 \quad \text{for all } w \in W_{ad}, \tag{9}$$

where $[\phi'(\bar y)^*\bar\mu]|_Q$ is the restriction of $[\phi'(\bar y)^*\bar\mu]$ to Q, $[\phi'(\bar y)^*\bar\mu]|_\Sigma$ is the restriction of $[\phi'(\bar y)^*\bar\mu]$ to Σ, $[\phi'(\bar y)^*\bar\mu]|_{\overline{\Omega}_T}$ is the restriction of $[\phi'(\bar y)^*\bar\mu]$ to $\overline{\Omega}_T$ and $[\phi'(\bar y)^*\bar\mu]|_{\overline{\Omega}_0}$ is the restriction of $[\phi'(\bar y)^*\bar\mu]$ to $\overline{\Omega}_0$, ($[\phi'(\bar y)^*\bar\mu]$ is the Radon measure on $\overline{Q}$ defined by $z \longmapsto \langle \bar\mu, \phi'(\bar y)z\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})}$ for $z \in C(\overline{Q})$, $\mathcal{L}^N$ denotes the N-dimensional Lebesgue measure).
Moreover, if $(A8)$ is satisfied, we can take $\bar\nu = 1$ in (7), (8), (9).

The meaning of weak solutions for (7), regularity results for $\bar p$, and the definition of $\bar p(0)$ are given in Section 3.

The notion of stability is closely related to the notion of calmness introduced by F. H. Clarke [15]. In the above setting this notion is due to J. V. Burke [9]. It has first been used in control problems by J. F. Bonnans and J. F. Bonnans and E. Casas. We do not know sufficient conditions ensuring that a state-constrained control problem is strongly stable on the right. However, even if (P) is not strongly stable on the right, (P_γ) will be strongly stable for all $\gamma > 0$, except on a subset of $\mathbb{R}^+$ of zero Lebesgue measure [8], [12]. In some situations Pontryagin's principles in qualified form may be derived from a non qualified form. Consider the example described below.

Example 2.3. We suppose that Ω is a connected. The state equation is

$$\frac{\partial y}{\partial t} - \Delta y = 0 \text{ in } Q, \quad \frac{\partial y}{\partial n} + y^4 = v \text{ on } \Sigma, \quad y(0) = y_0 \text{ on } \Omega, \tag{10}$$

where y_0 is a given function in $C(\overline{\Omega})$. We set $V_{ad} = \{v \in L^\sigma(\Sigma) \mid v(s,t) \geq 0 \text{ a.e. on } \Sigma\}$. The state constraints are defined by

$$0 \leq y(x,t) \leq \gamma_d \quad \text{on } \overline{Q},$$

for some given $\gamma_d > 0$. (We suppose that $0 \leq y_0(x) < \gamma_d$ on $\overline{\Omega}$.) Since for $v \in V_{ad}$, the solution y of (10) is nonnegative, then we can restrict the state constraints to $y(x,t) \leq \gamma_d$. We denote by J a cost functional defined as in (3), with $w \equiv y_0$, for which the assumptions (A3) to (A5) are satisfied. We suppose that the control problem

$$(P_{ex}) \qquad \inf\{J(y,v) \mid y \in W(0,T)\cap C(\overline{Q}),\ v \in V_{ad},\ (y,v) \text{ satisfies (10)},\ y(x,t) \le \gamma_d \text{ on } \overline{Q}\},$$

admits a solution $(\bar y, \bar v)$. The adjoint equation (7) for (P_{ex}) corresponding to $(\bar y, \bar v)$ and to $\bar\nu = 0$ is

$$-\frac{\partial \bar p}{\partial t} - \Delta \bar p = \bar\mu_Q \text{ in } Q, \quad \frac{\partial \bar p}{\partial n} + 4\bar y^3 \bar p = \bar\mu_\Sigma \text{ on } \Sigma, \quad \bar p(T) = \bar\mu_{\overline{\Omega}_T} \text{ on } \overline{\Omega},$$

where the measure $\bar\mu = \bar\mu_Q + \bar\mu_\Sigma + \bar\mu_{\overline{\Omega}_T}$ obeys

$$\bar\mu \ge 0, \quad \bar\mu \ne 0, \quad \langle \bar\mu, \bar y - \gamma_d\rangle_{\mathcal{M}(\overline{Q})\times C(\overline{Q})} = 0. \tag{11}$$

(Observe that (11) is nothing else than (6) for (P_{ex}) and $\bar\nu = 0$.) The Pontryagin principle in nonqualified form is expressed:

$$\bar p(s,t)(v - \bar v(s,t)) \ge 0 \quad \text{for all } v > 0 \quad \text{and for almost all } (s,t) \in \Sigma. \tag{12}$$

We set $\bar t = \inf\{t \in [0,T] \mid \bar\mu(\overline{\Omega}\times]t,T]) \ne 0\}$. Following [32] (Theorems 4.2, 4.3 and Remark 4.7), we can define $\bar p_{\overline{\Omega}}(t^+)$ as a function in $L^1(\Omega)$ which satisfies the Green formula

$$\int_{\Omega\times]t,T[} \bar p(\frac{\partial y}{\partial t} - \Delta y)\,dx\,dt + \int_{\Gamma\times]t,T[} \bar p(\frac{\partial y}{\partial n} + 4\bar y^3 y)\,ds\,dt$$

$$= \langle \bar\mu_Q, y\rangle_{\mathcal{M}_b(\Omega\times]t,T[)\times C_b(\Omega\times]t,T[)} + \langle \bar\mu_\Sigma, y\rangle_{\mathcal{M}_b(\Omega\times]t,T[)\times C_b(\Omega\times]t,T[)}$$

$$+\langle p(T), y(T)\rangle_{\mathcal{M}(\overline{\Omega})\times C(\overline{\Omega})} - \langle p_{\overline{\Omega}}(t^+), y(t)\rangle_{\mathcal{M}(\overline{\Omega})\times C(\overline{\Omega})},$$

for all $y \in C^{2,1}(\overline{Q})$ ($C^{2,1}$ means C^2 with respect to x and C^1 with respect to t). With this definition for $\bar p_{\overline{\Omega}}(t^+)$, the function $\bar p$ is, on $\overline{\Omega}\times]0,t[$, the unique solution of

$$-\frac{\partial \bar p}{\partial t} - \Delta \bar p = \bar\mu_{\Omega\times]0,t[} \text{ in } \Omega\times]0,t[, \quad \frac{\partial \bar p}{\partial n} + 4\bar y^3 \bar p = \bar\mu_{\Gamma\times]0,t[} \text{ on } \Gamma\times]0,t[, \quad \bar p(t) = \bar p_{\overline{\Omega}}(t^+) + \bar\mu_{\overline{\Omega}\times\{t\}} \text{ on } \overline{\Omega}.$$

We can easily prove that if $\bar p_{\overline{\Omega}}(t^+) = 0$, then $\mu_{\overline{\Omega}\times]t,T]} = 0$. From the definition of $\bar t$ we see that, for every $\varepsilon > 0$, $\bar p_{\overline{\Omega}}((\bar t - \varepsilon)^+) \ge 0$ and $\bar p_{\overline{\Omega}}((\bar t - \varepsilon)^+) \ne 0$. Let $\hat p$ be the solution of

$$-\frac{\partial p}{\partial t} - \Delta p = 0 \text{ in } \Omega\times]0,\bar t - \varepsilon[, \quad \frac{\partial p}{\partial n} + kp = 0 \text{ on } \Gamma\times]0,\bar t - \varepsilon[, \quad \bar p(t) = \bar p_{\overline{\Omega}}((\bar t - \varepsilon)^+) \text{ on } \overline{\Omega},$$

where $k = \max\{4\bar y^3(x,t) \mid (x,t) \in \overline{Q}\}$. By a comparison principle, we can verify that $\bar p \ge \hat p$ on $\Gamma\times]0,\bar t - \varepsilon[$. Moreover $\hat p \in C^2(\overline{\Gamma} \times [0,\bar t - \varepsilon[)$, $\hat p(\bar t - 2\varepsilon)$ is nonnegative and non identically zero. ¿From the maximum principle for classical solutions of parabolic equations, we deduce that $\hat p > 0$ on $\overline{\Omega} \times [0,\bar t - 2\varepsilon]$. Thus $\bar p(s,t) > 0$ almost everywhere on $\Gamma\times]0,\bar t[$. With (12), this implies $\bar v \equiv 0$ on $\Gamma\times]0,\bar t[$. Thus $0 \le y(x,t) \le \max_{\overline{\Omega}} y_0 < \gamma_d$ on $\overline{\Omega} \times [0,\bar t]$. From (11), it follows $\bar\mu_{\overline{\Omega}\times[0,\bar t]} = 0$. Thus $\bar\mu \equiv 0$ and we get a contradiction. In this simple example we see that the Pontryagin principle in qualified form follows from the Pontryagin principle in nonqualified form.

3 State equation and Adjoint equation.

3.1 Existence, uniqueness and regularity of the state variable.

Theorem 3.1 *Under assumptions (A1) and (A2), if $v \in L^\sigma(\Sigma)$ and $w \in C(\overline{\Omega})$, then the equation (1) admits a unique weak solution y_{vw} in $W(0,T) \cap C(\overline{Q})$. This solution satisfies*

$$\|y_{vw}\|_{\infty,Q} \leq C_1(\|v\|_{\sigma,\Sigma} + \|w\|_{\infty,\Omega} + 1),$$

where $C_1 = C_1(T, \Omega, N, q, \sigma, C_0)$. Moreover, the mapping $(v, w) \longmapsto y_{vw}$ is continuous from $L^\sigma(\Sigma) \times C(\overline{\Omega})$ into $C(\overline{Q})$.

Proof. The existence of a unique weak solution y_{vw} in $W(0,T) \cap C(\overline{Q})$ for equation (1), is proved in ([35], Theorem 3.1). The last part of Theorem can be proved as in ([35], Proposition 4.3). □

Corollary 3.1 *For every $k > 0$ and for every $\varepsilon > 0$, there exist $C_2 = C_2(T, \Omega, N, q, \sigma, C_0, k)$, $C_3 = C_3(T, \Omega, N, q, \sigma, C_0, k, \varepsilon)$ and $\alpha > 0$ such that, for every $(v, w) \in V_{ad} \times W_{ad}$ satisfying $\|v\|_{L^\sigma(\Sigma)} + \|w\|_{\infty,\Omega} \leq k$, the weak solution y_{vw} of (1) corresponding to (v, w) is Hölder continuous on $[\varepsilon, T] \times \overline{\Omega}$ and obeys*

$$\|y_{vw}\|_{C(\overline{Q})} \leq C_2, \quad \|y_{vw}\|_{C^{\alpha, \frac{\alpha}{2}}(\overline{\Omega} \times [\varepsilon, T])} \leq C_3.$$

Moreover, if w is Hölder continuous on $\overline{\Omega}$, then y_{vw} is Hölder continuous on $\overline{Q}$.

Proof. Since y_{vw} belongs to $C(\overline{Q})$, thanks to (A1)-(A2), we see that y_{vw} is also the unique weak solution of

$$\frac{\partial y}{\partial t} + Ay = \tilde{f} \quad \text{in } Q, \quad \frac{\partial y}{\partial n_A} = \tilde{g} \quad \text{on } \Sigma, \quad y(0) = w \quad \text{in } \Omega,$$

where

$$\tilde{f}(\cdot) = -f(\cdot, y_{vw}(\cdot)) \in L^q(Q), \quad \tilde{g}(\cdot) = -g(\cdot, y_{vw}(\cdot), v(\cdot)) \in L^\sigma(\Sigma).$$

If we denote by γ_0 the trace operator from $L^{\sigma'}(0, T; W^{1,\nu}(\Omega))$ into $L^{\sigma'}(0, T; W^{1-\frac{1}{\nu},\nu}(\Gamma))$ (with $\nu = N/(\sigma N - N + 1)$) and by i the continuous embedding from $L^{\sigma'}(0, T; W^{1-\frac{1}{\nu},\nu}(\Gamma))$ into $L^{\sigma'}(0, T; L^{\sigma'}(\Gamma))$, we can write:

$$\langle \tilde{g}, (i \circ \gamma_0) z \rangle_{L^\sigma(\Sigma) \times L^{\sigma'}(\Sigma)} = \langle (\gamma_0^t \circ i^t) \tilde{g}, z \rangle_{L^\sigma(0,T;(W^{1,\nu}(\Omega))') \times L^{\sigma'}(0,T;W^{1,\nu}(\Omega))}$$

for every $z \in L^{\sigma'}(0, T; W^{1,\nu}(\Omega))$. We can identify $(\gamma_0^t \circ i^t)\tilde{g}$ with $(f_0, f_1, ..., f_N) \in (L^\sigma(0, T; L^{\nu'}(\Omega)))^{N+1}$ ($\nu' = \frac{N\sigma}{N-1}$) in the following manner :

$$\langle (\gamma_0^t \circ i^t) \tilde{g}, z \rangle_{L^\sigma(0,T;(W^{1,\nu}(\Omega))') \times L^{\sigma'}(0,T;W^{1,\nu}(\Omega))} = \int_Q (f_0 z + \Sigma_i f_i D_i z) \, dx dt$$

for every $z \in L^{\sigma'}(0, T; W^{1,\nu}(\Omega))$. Therefore, y_{vw} is the weak solution (in the sense of [16]) of the initial boundary value problem

$$\frac{\partial y}{\partial t} - \operatorname{div}(a(x, t, \nabla y)) = \tilde{f} + f_0 \quad \text{in } Q, \qquad a(x, t, \nabla y).n = 0 \quad \text{on } \Sigma, \qquad y(0) = w \quad \text{in } \Omega, \tag{13}$$

where

$$a(x,t,\nabla y) = (\sum_{j=1}^{N} a_{ij} D_j y - f_i)_{i=1,..,N}.$$

Now we can easily verify that assumptions of ([16], Theorem 1.3, Chapter 3) are satisfied by the system (13) and Hölder continuity results of Corollary follow from this theorem. $\square$

3.2 Metric space of controls.

To apply the Ekeland variational principle, we have to define a metric space of controls in order that the mapping $(v,w) \longmapsto y_{vw}$ be continuous from this metric space to $C(\overline{Q})$. Thanks to Theorem 3.1, this continuity condition will be realized if convergence in the metric space of controls implies convergence in $L^\sigma(\Sigma) \times C(\overline{\Omega})$.

In the case when boundary controls are bounded, convergence in (V_{ad}, d_E) (where d_E is the so-called Ekeland distance) implies convergence in $L^\sigma(\Sigma)$. This condition is no longer true for unbounded controls (see [23], p. 227). To overcome this difficulty, we define a new metric space in the following way.

Let $\tilde{v}$ be in V_{ad} (in Section 5, $\tilde{v}$ will be an optimal boundary control that we want to characterize). For $0 < k < \infty$, we define the set:

$$V_{ad}(\tilde{v},k) = \{v \in V_{ad} \mid |v(s,t) - \tilde{v}(s,t)| \leq k \text{ for a.e. } (s,t) \in \Sigma\}.$$

We endow the set $V_{ad}(\tilde{v},k) \times W_{ad}$ with the following metric:

$$d((v_1,w_1),(v_2,w_2)) = \mathcal{L}^N(\{(s,t) \mid v_1(s,t) \neq v_2(s,t)\}) + \|w_1 - w_2\|_{\infty,\Omega}.$$

Remark 3.1. Thanks to [18], we know that the mapping

$$d_E : (v_1,v_2) \longmapsto \mathcal{L}^N(\{(s,t) \mid v_1(s,t) \neq v_2(s,t)\})$$

is a distance on $V_{ad}(\tilde{v},k)$. Moreover, if $(v_n)_n \subset V_{ad}(\tilde{v},k)$, if $v \in V_{ad}(\tilde{v},k)$ and if $\lim_n d_E(v_n,v) = 0$, then $(v_n)_n$ converges to v in $L^\sigma(\Sigma)$. This is no longer true for any sequence $(v_n)_n$ included in V_{ad}.

Lemma 3.1 $(V_{ad}(\tilde{v},k) \times W_{ad}, d)$ *is a complete metric space, and the mapping which associates* $(y_{vw}, J(y_{vw},v,w))$ *with* (v,w) *is continuous from* $(V_{ad}(\tilde{v},k) \times W_{ad}, d)$ *into* $C(\overline{Q}) \times \mathbb{R}$.

Proof. i) To prove that $(V_{ad}(\tilde{v},k) \times W_{ad}, d)$ is a complete metric space, it remains to prove that $(V_{ad}(\tilde{v},k), d_E)$ is complete. Let $(v_n)_n$ be a Cauchy sequence in $(V_{ad}(\tilde{v},k), d_E)$. Following [18], we can prove that $(v_n)_n$ converges for d_E to some measurable function v such that $v(s,t) \in K_V(s,t)$ and $|v(s,t) - \tilde{v}(s,t)| \leq k$ for almost all $(s,t) \in \Sigma$. Therefore $v \in L^\sigma(\Sigma)$ and $v \in V_{ad}(\tilde{v},k)$.

ii) Now, we consider $(v_n,w_n)_{n\geq1} \subset V_{ad}(\tilde{v},k) \times W_{ad}$ and $(v,w) \in V_{ad}(\tilde{v},k) \times W_{ad}$ such that $(v_n,w_n)_n$ converges to (v,w) for the metric d. We denote by y and y_n $(n \geq 1)$ the solution of (1) corresponding respectively to (v,w) and to (v_n,w_n). To prove the continuity result, it remains to prove that the sequence $(y_n, J(y_n,v_n,w_n))_n$ converges to $(y, J(y,v,w))$ in $C(\overline{Q})\times\mathbb{R}$. For this, we remark that $(w_n)_n$ converges to w in $C(\overline{\Omega})$ and $(v_n)_n$ converges to v in $L^\sigma(\Sigma)$. We complete the proof thanks to the continuity assumptions on F, G, L and thanks to the continuity results stated in Theorem 3.1. $\square$

3.3 Adjoint equation.

Let (a,b) be in $L^q(Q)\times L^\sigma(\Sigma)$ with $a\ge C_0$ and $b\ge C_0$. We consider the following terminal boundary value problem :

$$-\frac{\partial p}{\partial t}+Ap+ap=\mu_Q \quad \text{in } Q, \qquad \frac{\partial p}{\partial n_A}+bp=\mu_\Sigma \quad \text{on } \Sigma, \qquad p(T)=\mu_{\overline{\Omega}_T} \quad \text{on } \overline{\Omega}, \tag{14}$$

where $\mu=\mu_Q+\mu_\Sigma+\mu_{\overline{\Omega}_T}$ is a bounded Radon measure on $\overline{Q}\setminus\overline{\Omega}_0$, μ_Q is the restriction of μ to Q, μ_Σ is the restriction of μ to Σ and $\mu_{\overline{\Omega}_T}$ is the restriction of μ to $\overline{\Omega}_T$.

Definition 3.1 *We shall say that* p *is a weak solution of (14) in* $L^1(0,T;W^{1,1}(\Omega))$ *if and only if the two following conditions are verified:*
(i) $ap\in L^1(Q)$ *and* $bp\in L^1(\Sigma)$,
(ii) For every $\varphi\in C^1(\overline{Q})$ *verifying* $\varphi(x,0)=0$ *on* $\overline{\Omega}$, *we have*

$$\int_Q\left\{p\frac{\partial\varphi}{\partial t}+\Sigma_{i,j}a_{ij}D_j\varphi D_ip+a\varphi p\right\}dxdt+\int_\Sigma b\varphi p\,dsdt=\langle\varphi,\mu\rangle_{C_b(\overline{Q}\setminus\overline{\Omega}_0)\times\mathcal{M}_b(\overline{Q}\setminus\overline{\Omega}_0)}.$$

($C_b(\overline{Q}\setminus\overline{\Omega}_0)$ *denotes the space of bounded continuous functions on* $\overline{Q}\setminus\overline{\Omega}_0$, *while* $\mathcal{M}_b(\overline{Q}\setminus\overline{\Omega}_0)$ *denotes the space of bounded Radon measures on* $\overline{Q}\setminus\overline{\Omega}_0$, *that is, the topological dual of* $C_0(\overline{Q}\setminus\overline{\Omega}_0)$.*)*

In all the sequel, we shall say that a pair $(\delta,d)\in\mathbb{R}^2$ verifies the condition $(C_{q\sigma})$ if and only if

$$(C_{q\sigma})\quad\begin{cases}\frac{N\sigma}{\sigma-2}<d\le\frac{N\sigma}{N-1} \ \text{ and } \ \frac{2d}{d-N}<\delta\le\sigma & \text{if } \sigma\le q,\\[2mm] \frac{Nq}{q-2}<d\le\frac{N\sigma}{N-1} \ \text{ and } \ \frac{2d}{d-N}<\delta\le q & \text{if } N\le q<\sigma,\\[2mm] \frac{Nq}{q-2}<d\le\inf(\frac{N\sigma}{N-1},\frac{Nq}{N-q}) \ \text{ and } \ \frac{2d}{d-N}<\delta\le q & \text{if } q<N.\end{cases}$$

Since $q>\frac{N}{2}+1$, $\sigma>N+1$ and $q\sigma+q>qN+2\sigma$, we remark that the set of pairs (δ,d) satisfying $(C_{q\sigma})$ is nonempty. These conditions appear in a natural manner when we study the equation (14). Indeed, if (δ,d) satisfies $(C_{q\sigma})$, if (a,b) belongs to $L^q(Q)\times L^\sigma(\Sigma)$ and if p belongs to $L^{\delta'}(0,T;W^{1,d'}(\Omega))$, then $ap\in L^1(Q)$ and $bp\in L^1(\Sigma)$. We now recall an existence theorem for parabolic equations with measures as data stated in [32].

Theorem 3.2 *Let* (a,b) *be in* $L^q(Q)\times L^\sigma(\Sigma)$ *satisfying* $a\ge C_0$, $b\ge C_0$ *and let* μ *be in* $\mathcal{M}_b(\overline{Q}\setminus\overline{\Omega}_0)$. *Equation (14) admits a unique weak solution* $p\in L^1(0,T;W^{1,1}(\Omega))$. *For every* (δ,d) *satisfying* $(C_{q\sigma})$, $p\in L^{\delta'}(0,T;W^{1,d'}(\Omega))$ *and we have:*

$$\|p\|_{L^{\delta'}(0,T;W^{1,d'}(\Omega))}\le C_4\|\mu\|_{\mathcal{M}_b(\overline{Q}\setminus\overline{\Omega}_0)},$$

where $C_4 = C_4(T, \Omega, N, \delta, d, C_0)$ is independent of a and b.
Moreover, there exists a Radon measure on $\overline{\Omega}$, denoted by $p(0)$ such that :

$$\int_Q p\left\{\frac{\partial y}{\partial t} + Ay + ay\right\} dxdt + \int_\Sigma p\left\{\frac{\partial y}{\partial n_A} + by\right\} dsdt$$
$$= \langle y, \mu\rangle_{C_b(\overline{Q}\setminus\overline{\Omega}_0)\times\mathcal{M}_b(\overline{Q}\setminus\overline{\Omega}_0)} - \langle y(0), p(0)\rangle_{C(\overline{\Omega})\times\mathcal{M}(\overline{\Omega})}$$

for every $y \in Y = \left\{y \in W(0,T)\cap C(\overline{Q}) \mid \frac{\partial y}{\partial t} + Ay \in L^q(Q),\ \frac{\partial y}{\partial n_A} \in L^\sigma(\Sigma)\right\}$.

Remark 3.2. If $p \in L^{\delta'}(0,T; W^{1,d'}(\Omega))$ (where (δ, d) satisfies $(C_{q\sigma})$), and if

$$\mathrm{div}_{xt}\left((\Sigma_j a_{ij} D_j p)_{1\le i\le N}, p\right) = \frac{\partial p}{\partial t} - Ap \quad \text{belongs to } \mathcal{M}_b(Q),$$

then we can define the normal trace of the vector field $((\sum_j a_{ij}D_jp)_{1\le i\le N}, p)$, in the space $W^{\frac{-1}{m},m}(\partial Q)$ (for some $1 < m < \frac{N+1}{N}$). If we denote by $\gamma_n((\sum_j a_{ij}D_jp)_{1\le i\le N}, p)$ this normal trace, we can prove (see Theorem 4.2 in [32]) that this normal trace belongs to $\mathcal{M}(\partial Q)$ and the restriction of $\gamma_n((\sum_j a_{ij}D_jp)_{1\le i\le N}, p)$ to $\overline{\Omega}_T$ is equal to $\mu_{\overline{\Omega}_T}$, the restriction of $\gamma_n((\sum_j a_{ij}D_jp)_{1\le i\le N}, p)$ to Σ is equal to $\mu_\Sigma - bp$ and if $p(0)$ is the measure on $\overline{\Omega}$ which satisfies the Green formula of Theorem 3.2, then $-p(0)$ is the restriction of $\gamma_n((\sum_j a_{ij}D_jp)_{1\le i\le N}, p)$ to $\overline{\Omega}_0$. In fact it can be proved that $p(0)$ belongs to $L^1(\Omega)$ (see Theorem 4.3 in [32]).

Remark 3.3. Let us explain the origin of the condition $(C_{q\sigma})$. In [32], the existence of a weak solution in $L^1(0,T;W^{1,1}(\Omega))$ for equation (14) is proved by duality arguments and an approximation process. The condition $\delta > 2d/(d-N)$ appears to get C^0-regularity results for some adjoint equation associated with (14) (see ([32], Theorem 4.1)). The condition $\delta > 2d/(d-N)$, together with the conditions $p \in L^{\delta'}(0,T, W^{1,d'}(\Omega))$, $a \in L^q(Q)$, $b \in L^\sigma(\Sigma)$, $ap \in L^1(Q)$, $bp \in L^1(\Sigma)$ are equivalent to the condition: "(δ, d) satisfies $(C_{q\sigma})$". In the case when $a \in L^\infty(Q)$ and $b \in L^\infty(\Sigma)$, the condition $(C_{q\sigma})$ can be replaced by the only condition $\delta > 2d/(d-N)$.

4 Existence of diffuse perturbations.

In Section 5, we consider control problems in which the state constraints are penalized. The penalization is chosen in such a way that the solution of (P) that we want to characterize will be an ε-solution of the penalized problem. In order to exploit optimality conditions deduced from the Ekeland's variational principle, we need to construct admissible perturbations of approximate optimal solutions. For this we use a kind of perturbations that we call "diffuse perturbations" and which goes back to Yao [38] and Li [26]. A diffuse perturbation of a control $\bar{v} \in V_{ad}$ is a function v_ρ defined by

$$v_\rho(s,t) = \begin{cases} \bar{v}(s,t) & \text{on } \Sigma\setminus E_\rho,\\ v(s,t) & \text{on } E_\rho,\end{cases}$$

where $v \in V_{ad}$ and E_ρ is some measurable subset of Σ. It is clear that $v_\rho \in V_{ad}$. Contrary to spike perturbations or multispike perturbations, where E_ρ is precisely defined, here E_ρ must satisfy some relations such as (23), (24) and (25). As explained in Lemma 4.1, the existence of E_ρ follows from the Lyapounov convexity Theorem.

To get optimality conditions we need some differential calculus rules, for this type of perturbation, stated in the following theorem.

Theorem 4.1 *Let ρ be positive constant such that $0 < \rho < 1$. For every $v_1, v_2 \in V_{ad}$ and for every $w_1, w_2 \in C(\overline{\Omega})$, there exists a measurable subset $E_\rho \subset \Sigma$ such that:*

$$\mathcal{L}^N(E_\rho) = \rho \mathcal{L}^N(\Sigma), \tag{15}$$

$$\int_{E_\rho} (G(s,t,y_1,v_2) - G(s,t,y_1,v_1))\, dsdt = \rho \int_\Sigma (G(s,t,y_1,v_2) - G(s,t,y_1,v_1))\, dsdt, \tag{16}$$

$$y_\rho = y_1 + \rho z + r_\rho, \quad \textit{with } \lim_{\rho \to 0} \frac{1}{\rho} \|r_\rho\|_{C(\overline{Q})} = 0, \tag{17}$$

$$J(y_\rho, v_\rho, w_\rho) = J(y_1, v_1, w_1) + \rho \Delta J + o(\rho), \tag{18}$$

where v_ρ, w_ρ are the controls defined by:

$$v_\rho(s,t) = \begin{cases} v_1(s,t) & \textit{on } \Sigma \setminus E_\rho \\ v_2(s,t) & \textit{on } E_\rho \end{cases}, \tag{19}$$

$$w_\rho = w_1 + \rho w_2, \tag{20}$$

y_ρ, y_1 are the solutions of (1) corresponding respectively to (v_ρ, w_ρ) and to (v_1, w_1), z is the weak solution of

$$\begin{aligned} &\frac{\partial z}{\partial t} + Az + f'_y(x,t,y_1)z = 0 && \textit{in } Q, \\ &\frac{\partial z}{\partial n_A} + g'_y(s,t,y_1,v_1)z = g(s,t,y_1,v_1) - g(s,t,y_1,v_2) && \textit{on } \Sigma, \\ &z(0) = w_2 && \textit{in } \Omega, \end{aligned} \tag{21}$$

and

$$\Delta J = J'_y(y_1, v_1, w_1)z + J(y_1, v_2, w_1) - J(y_1, v_1, w_1) + \int_\Omega L'_w(x, y_1(T), w_1) w_2\, dx. \tag{22}$$

The proof relies on the following Lemma.

Lemma 4.1 *Let v_1, v_2 be in V_{ad} and let y be in $C(\overline{Q})$. For every $\rho \in]0,1[$, there exists a sequence of measurable subsets $(E^n_\rho)_n$ in Σ such that:*

$$\mathcal{L}^N(E^n_\rho) = \rho \mathcal{L}^N(\Sigma), \tag{23}$$

$$\int_{E^n_\rho} (G(s,t,y,v_1) - G(s,t,y,v_2))\, dsdt = \rho \int_\Sigma (G(s,t,y,v_1) - G(s,t,y,v_2))\, dsdt, \tag{24}$$

$$\frac{1}{\rho} \chi_{E^n_\rho} \rightharpoonup 1 \textit{ weakly star in } L^\infty(\Sigma), \textit{ when } n \to \infty, \tag{25}$$

where $\chi_{E^n_\rho}$ is the characteristic of E^n_ρ.

Remark 4.1. A statement similar to (15), (17), (18) is given in [28], [29], [24] and in [12] (conditions (23), (25) are also stated in [12]). In [28] the proof relies on an extension of Uhl's theorem. The proofs in [24] and in [12] are constructive. Since the existence of E_ρ^n, satisfying together the conditions (23), (24), (25), is not proved neither in [12] nor in [28]. We here give a short proof of Lemma 4.1 based on the Lyapunov convexity theorem.

Proof of Lemma 4.1. We consider a family $(\varphi_n)_n$ dense in $L^1(\Sigma)$. For $n \geq 0$, we set

$$f^n = (1, G(\cdot, y, v_1) - G(\cdot, y, v_2), \varphi_0, \varphi_1, ..., \varphi_n) \in (L^1(\Sigma))^{n+3}.$$

Thanks to Lyapunov's convexity theorem, for every $n \geq 0$ and every $\rho \in]0,1[$, there exists a measurable subset $E_\rho^n \subset \Sigma$ satisfying

$$\int_{E_\rho^n} f^n \, dsdt = \rho \int_\Sigma f^n \, dsdt.$$

Thus, for every $n \geq 0$, E_ρ^n satisfies (23), (24) and

$$\int_{E_\rho^n} \varphi_m \, dsdt = \rho \int_\Sigma \varphi_m \, dsdt, \tag{26}$$

for every $m \in \{0, \cdots, n\}$. Now, for any fixed φ in $L^1(\Sigma)$, we have

$$|\int_\Sigma (\frac{1}{\rho}\chi_{E_\rho^n} - 1)\varphi \, dsdt| \leq |\int_\Sigma (\frac{1}{\rho}\chi_{E_\rho^n} - 1)(\varphi - \varphi_m) \, dsdt| + |\int_\Sigma (\frac{1}{\rho}\chi_{E_\rho^n} - 1)\varphi_m \, dsdt|$$

$$\leq (\frac{1}{\rho} + 1)\|\varphi - \varphi_m\|_{1,\Sigma} + |\int_\Sigma (\frac{1}{\rho}\chi_{E_\rho^n} - 1)\varphi_m \, dsdt|.$$

Since $(\varphi_m)_m$ is dense in $L^1(\Sigma)$, for $\epsilon > 0$ there exists $m_0 > 0$ such that $\|\varphi - \varphi_{m_0}\|_{1,\Sigma} \leq \frac{\epsilon}{\frac{1}{\rho}+1}$. Thanks to (26), for every $n \geq m$, we have $\int_\Sigma (\frac{1}{\rho}\chi_{E_\rho^n} - 1)\varphi_m \, dsdt = 0$. Thus, it follows that

$$\lim_n |\int_\Sigma (\frac{1}{\rho}\chi_{E_\rho^n} - 1)\varphi \, dsdt| = 0,$$

and the proof is complete. □

Proof of Theorem 4.1. The existence of E_ρ satisfying (15), (16) is an easy consequence of Lemma 4.1. The only delicate point is the proof of (17). This kind of result is already given in ([12], Theorem 5.2) and in ([24], Theorem 3.3). Since we deal with unbounded controls and nonmonotone operators, our assumptions are different from those in [12] and in [24]. However, the proof of (17) can be adapted from the ones given in [12] and in [24].

Let ρ be in $]0,1[$ and let $(E_\rho^n)_n$ be the sequence of measurable subsets defined in Lemma 4.1. We set

$$v_\rho^n(s,t) = \begin{cases} v_1(s,t) & \text{on } \Sigma \setminus E_\rho^n \\ v_2(s,t) & \text{on } E_\rho^n \end{cases}, \qquad w_\rho = w_1 + \rho w_2.$$

Let y^n_ρ be the solution of (1) corresponding to (v^n_ρ, w_ρ) and let z be the weak solution of (21). It is clear that $\xi^n_\rho = (y^n_\rho - y_1)/\rho - z$ is the weak solution in $C(\overline{Q}) \cap W(0,T)$ of

$$\frac{\partial \xi}{\partial t} + A\xi + a^n_\rho \xi = f^n_\rho \quad \text{in } Q, \quad \frac{\partial \xi}{\partial n_A} + b^n_\rho \xi = g^n_\rho + h^n_\rho \quad \text{on } \Sigma, \quad \xi(0) = 0 \quad \text{in } \Omega,$$

where

$$\begin{aligned}
a^n_\rho(x,t) &= \int_0^1 f'_y(x,t,(y_1 + \theta(y^n_\rho - y_1))(x,t))\, d\theta, \\
b^n_\rho(s,t) &= \int_0^1 g'_y(s,t,(y_1 + \theta(y^n_\rho - y_1))(s,t), v^n_\rho(s,t))\, d\theta, \\
f^n_\rho &= (f'_y(x,t,y_1) - a^n_\rho)z, \\
g^n_\rho &= (g'_y(s,t,y_1,v_1) - b^n_\rho)z, \\
h^n_\rho &= (1 - \frac{1}{\rho}\chi_{E^n_\rho})(g(s,t,y_1,v_2) - g(s,t,y_1,v_1)),
\end{aligned}$$

and $\chi_{E^n_\rho}$ is the characteristic function of E^n_ρ. We denote by $\xi^{n,1}_\rho$ the solution in $C(\overline{Q}) \cap W(0,T)$ of

$$\frac{\partial \xi}{\partial t} + A\xi + a^n_\rho \xi = f^n_\rho \quad \text{in } Q, \quad \frac{\partial \xi}{\partial n_A} + b^n_\rho \xi = g^n_\rho \quad \text{on } \Sigma, \quad \xi(\cdot,0) = 0 \quad \text{in } \Omega,$$

by $\xi^{n,2}_\rho$ the solution in $C(\overline{Q}) \cap W(0,T)$ of

$$\frac{\partial \xi}{\partial t} + A\xi + a^n_\rho \xi = 0 \quad \text{in } Q, \quad \frac{\partial \xi}{\partial n_A} + b^n_\rho \xi = h^n_\rho \quad \text{on } \Sigma, \quad \xi(\cdot,0) = 0 \quad \text{in } \Omega,$$

and by ζ^n_ρ the solution in $C(\overline{Q}) \cap W(0,T)$ of

$$\frac{\partial \zeta}{\partial t} + A\zeta + a\zeta = 0 \quad \text{in } Q, \quad \frac{\partial \zeta}{\partial n_A} + b\zeta = h^n_\rho \quad \text{on } \Sigma, \quad \zeta(\cdot,0) = 0 \quad \text{in } \Omega,$$

where $a(x,t) = f'_y(x,t,y_1(x,t))$, $b(s,t) = g'_y(s,t,y_1(s,t),v_1(s,t))$. We also have

$$\begin{aligned}
&\frac{\partial(\xi^{n,2}_\rho - \zeta^n_\rho)}{\partial t} + A(\xi^{n,2}_\rho - \zeta^n_\rho) + a^n_\rho(\xi^{n,2}_\rho - \zeta^n_\rho) = (a - a^n_\rho)\zeta^n_\rho && \text{in } Q, \\
&\frac{\partial(\xi^{n,2}_\rho - \zeta^n_\rho)}{\partial n_A} + b^n_\rho(\xi^{n,2}_\rho - \zeta^n_\rho) = (b - b^n_\rho)\zeta^n_\rho && \text{on } \Sigma, \\
&(\xi^{n,2}_\rho - \zeta^n_\rho)(\cdot,0) = 0 && \text{in } \Omega.
\end{aligned}$$

Thanks to ([35], Proposition 3.3), there exists $C = C(T,\Omega,N,q,\sigma,C_0) > 0$ (independent of n and ρ) such that

$$\|\xi^{n,2}_\rho - \zeta^n_\rho\|_{C(\overline{Q})} \le C(\|a - a^n_\rho\|_{q,Q} + \|b - b^n_\rho\|_{\sigma,\Sigma})\|\zeta^n_\rho\|_{C(\overline{Q})}, \tag{27}$$

$$\|\xi^{n,1}_\rho\|_{C(\overline{Q})} \le C(\|f^n_\rho\|_{q,Q} + \|g^n_\rho\|_{\sigma,\Sigma}). \tag{28}$$

The operator $\mathcal{T}$ which associates ζ, the solution in $C(\overline{Q}) \cap W(0,T)$ of

$$\frac{\partial \zeta}{\partial t} + A\zeta + a\zeta = \varphi \quad \text{in } Q, \quad \frac{\partial \zeta}{\partial n_A} + b\zeta = \psi \quad \text{on } \Sigma, \quad \zeta(0) = 0 \quad \text{in } \Omega,$$

with (φ, ψ) is continuous from $L^q(Q) \times L^\sigma(\Sigma)$ into $C^{\alpha,\frac{\alpha}{2}}(\overline{Q})$ for some $0 < \alpha < 1$ (as for Corollary 3.1, this continuity result can be deduced from Theorem 1.3, Chapter 3 in [16]). Since the embedding from $C^{\alpha,\frac{\alpha}{2}}(\overline{Q})$ into $C(\overline{Q})$ is compact, $\mathcal{T}$ may also be considered as a compact operator from $L^q(Q) \times L^\sigma(\Sigma)$ into $C(\overline{Q})$. Because of (25), for every $0 < \rho < 1$ the sequence $(h^n_\rho)_n$ converges to zero for the weak topology of $L^\sigma(\Sigma)$. Therefore, since $\mathcal{T}$ is compact from $L^q(Q) \times L^\sigma(\Sigma)$ into $C(\overline{Q})$, the sequence $(\zeta^n_\rho)_n$ converges to zero in $C(\overline{Q})$. There then exists an integer depending on ρ, denoted by $n(\rho)$, such that

$$\|\zeta^{n(\rho)}_\rho\|_{C(\overline{Q})} \leq \rho. \tag{29}$$

Notice that $(v^{n(\rho)}_\rho)_\rho$ converges to v_1 in $L^\sigma(\Sigma)$ and $(w_\rho)_\rho$ converges to w in $C(\overline{\Omega})$ as ρ tends to zero. Theorem 3.1 yields that $(y^{n(\rho)}_\rho)_\rho$ uniformly converges to y_1 on $\overline{Q}$ as ρ tends to zero. Therefore, thanks to (A1) and (A2), $f^{n(\rho)}_\rho$ and $(a - a^{n(\rho)}_\rho)$ both converge to zero in $L^q(Q)$ when ρ tends to zero and $g^{n(\rho)}_\rho$, $(b - b^{n(\rho)}_\rho)$ both converge to zero in $L^\sigma(\Sigma)$ when ρ tends to zero. Thus, thanks to (27)-(29), we get

$$\lim_{\rho\to 0} \|\xi^{n(\rho)}_\rho\|_{C(\overline{Q})} \leq \lim_{\rho\to 0} \|\xi^{n(\rho),1}_\rho\|_{C(\overline{Q})} + \lim_{\rho\to 0} \|\xi^{n(\rho),2}_\rho - \zeta^{n(\rho)}_\rho\|_{C(\overline{Q})} + \lim_{\rho\to 0} \|\zeta^{n(\rho)}_\rho\|_{C(\overline{Q})} = 0.$$

Now we set $E_\rho = E^{n(\rho)}_\rho$, $v_\rho = v^{n(\rho)}_\rho$ and $\frac{1}{\rho} r_\rho = \xi^{n(\rho)}_\rho$. Conditions (15) to (17) are clearly satisfied; moreover, taking into account (16), (17) and the definition of (v_ρ, w_ρ), we easily verify (18).

□

5 Proof of Pontryagin's principle.

5.1 Penalized problem.

We first give the proof of optimality conditions in qualified form (the case $\bar{\nu} = 1$ in Theorem 2.1). The proof of the nonqualified form can be obtained with slight modifications that we give in Section 5.3. For notational simplicity, throughout what follows we set

$$H_\Sigma(s, t, y, v, p, 1) = H_\Sigma(s, t, y, v, p)$$

for every $(s, t, y, v, p) \in \Gamma \times [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$.

Following [28], [29], since $C(\overline{D})$ is separable, there exists a norm $|\cdot|_{C(\overline{D})}$, which is equivalent to the norm $\|.\|_{C(\overline{D})}$ such that $(C(\overline{D}), |\cdot|_{C(\overline{D})})$ is strictly convex and $\mathcal{M}(\overline{D})$, endowed with the dual norm of $|\cdot|_{C(\overline{D})}$ (denoted by $|\cdot|_{\mathcal{M}(\overline{D})}$), is also strictly convex (see [17], Corollary 2 p. 148, or Corollary 2 p. 167). We define the distance function to $\mathcal{C}$ (for the new norm $|\cdot|_{C(\overline{D})}$) by

$$d_{\mathcal{C}}(\varphi) = \inf_{z\in\mathcal{C}} |\varphi - z|_{C(\overline{D})}.$$

Since $\mathcal{C}$ is convex, then $d_{\mathcal{C}}$ is convex and Lipschitz of rank 1, and we have

$$\limsup_{\substack{\rho\searrow 0,\\ \varphi'\to\varphi}} \frac{d_{\mathcal{C}}(\varphi' + \rho z) - d_{\mathcal{C}}(\varphi')}{\rho} = \max\{\langle \xi, z\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})} \mid \xi \in \partial d_{\mathcal{C}}(\varphi)\} \tag{30}$$

for every $\varphi, z \in C(\overline{D})$, where $\partial d_{\mathcal{C}}$ is the subdifferential in the sense of convex analysis (see [15]). Therefore, for a given $\varphi \in C(\overline{D})$ we have

$$\langle \xi, z-\varphi\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})} + d_{\mathcal{C}}(\varphi) \le d_{\mathcal{C}}(z) \quad \text{for every } \xi \in \partial d_{\mathcal{C}}(\varphi) \text{ and every } z \in C(\overline{D}), \tag{31}$$

$$|\xi|_{\mathcal{M}(\overline{D})} \le 1 \quad \text{for every } \xi \in \partial d_{\mathcal{C}}(\varphi).$$

Moreover it is proved in ([29], Lemma 3.4) that, since $\mathcal{C}$ is a closed convex subset of $C(\overline{D})$, for every $\varphi \notin \mathcal{C}$, and every $\xi \in \partial d_{\mathcal{C}}(\varphi)$, $|\xi|_{\mathcal{M}(\overline{D})} = 1$. Since $\partial d_{\mathcal{C}}(\varphi)$ is convex in $\mathcal{M}(\overline{D})$ and $(\mathcal{M}(\overline{D}), |\cdot|_{\mathcal{M}(\overline{D})})$ is strictly convex, then if $\varphi \notin \mathcal{C}$, $\partial d_{\mathcal{C}}(\varphi)$ is a singleton and $d_{\mathcal{C}}$ is Gâteaux-differentiable at φ.

Let $(\bar y, \bar v, \bar w)$ be a solution of problem (P). Thanks to (A8), we prove in the proposition below that $(\bar y, \bar v, \bar w)$ is also a local solution of some related penalized problems.

Proposition 5.1 *For every $k > 0$, there exists $\lambda = \lambda(k)$ such that $(\bar y, \bar v, \bar w)$ is a solution of the following problem :*

$$(P^{r,k}) \qquad \inf\{J_r(y,v,w) \mid y \in W(0,T)\cap C(\overline{Q}),\ (v,w) \in (V_{ad}(\bar v,k)\times W_{ad}) \cap \mathcal{B}^d_{\lambda(k)}(\bar v,\bar w)$$
$$\textit{and } (y,v,w) \textit{ satisfies } (1)\}$$

where

$$J_r(y,v,w) = J(y,v,w) + r d_{\mathcal{C}}(\phi(y)),$$

$\mathcal{B}^d_{\lambda(k)}(\bar v,\bar w) \subset V_{ad}\times W_{ad}$ is the closed ball centered on $(\bar v,\bar w)$ and with radius $\lambda(k)$ (for the distance d), r only depends on the constant $\tilde r$ given in (A8) and on $\overline{D}$.

Proof. From (A8), there exist $\tilde\varepsilon > 0$ and $\tilde r > 0$ such that

$$\inf(P) = \inf\{J(y_{vw},v,w) + \tilde r\gamma \mid (v,w)\in V_{ad}\times W_{ad},\ \phi(y_{vw}) \in \mathcal{C}_\gamma,\quad \gamma\in[0,\tilde\varepsilon]\}.$$

Now by writing

$$\inf(P) = \inf\{\inf\{J(y_{vw},v,w) + \tilde r\gamma \mid\ \phi(y_{vw})\in\mathcal{C}_\gamma,\quad \gamma\in[0,\tilde\varepsilon]\} \mid (v,w)\in V_{ad}\times W_{ad}\},$$

we get

$$\inf(P) = \inf\{J(y_{vw},v,w) + \tilde r \inf_{z\in\mathcal{C}} \|\phi(y_{vw}) - z\|_{C(\overline{D})} \mid (v,w)\in V_{ad}\times W_{ad},\ \phi(y_{vw})\in\mathcal{C}_{\tilde\varepsilon}\}.$$

Since the norms $|\cdot|_{C(\overline{D})}$ and $\|.\|_{C(\overline{D})}$ are equivalent, there exist $r \ge \tilde r$ and $0 < \varepsilon \le \tilde\varepsilon$ such that

$$\inf\{J(y_{vw},v,w) + \tilde r \inf_{z\in\mathcal{C}} \|\phi(y_{vw}) - z\|_{C(\overline{D})} \mid (v,w)\in V_{ad}\times W_{ad},\ \phi(y_{vw})\in\mathcal{C}_{\tilde\varepsilon}\} \le$$
$$\inf\{J(y_{vw},v,w) + r d_{\mathcal{C}}(\phi(y_{vw})) \mid (v,w)\in V_{ad}\times W_{ad},\ d_{\mathcal{C}}(\phi(y_{vw}))\le\varepsilon\}.$$

Moreover, taking (A6) and Lemma 3.1 into account, there exists $\lambda(k) > 0$ such that

$$d_{\mathcal{C}}(\phi(y_{vw})) \le \varepsilon \quad \text{for every } (v,w) \in (V_{ad}(\bar v,k)\times W_{ad}) \cap \mathcal{B}^d_{\lambda(k)}(\bar v,\bar w).$$

Thus

$$\inf(P) \leq \inf\{J(y_{vw}, v, w) + r d_{\mathcal{C}}(\phi(y_{vw})) \mid (v, w) \in (V_{ad}(\bar{v}, k) \times W_{ad}) \cap \mathcal{B}^d_{\lambda(k)}(\bar{v}, \bar{w})\}$$
$$= \inf(P^{r,k}) \leq J(\bar{y}, \bar{v}, \bar{w}) = \inf(P). \qquad \square$$

Now we set

$$J_{r,n}(y, v, w) = J(y, v, w) + r[(d_{\mathcal{C}}(\phi(y)))^2 + n^{-2}]^{\frac{1}{2}},$$

and we denote by $(P_n^{r,k})$ the problem

$$\inf\{J_{r,n}(y, v, w) \mid y \in W(0, T) \cap C(\overline{Q}),\ (v, w) \in (V_{ad}(\bar{v}, k) \times W_{ad}) \cap \mathcal{B}^d_{\lambda(k)}(\bar{v}, \bar{w})$$
$$\text{and } (y, v, w) \text{ satisfies (1)}\}.$$

Proposition 5.2 *For every $k > 0$,*

$$\inf(P^{r,k}) = \lim_{n\to\infty} \inf(P_n^{r,k}) \quad \textit{and} \quad \lim_{n\to\infty} J_{r,n}(y, v, w) = J_r(y, v, w)$$

for every $y \in C(\overline{Q})$ and every $(v, w) \in L^\sigma(\Sigma) \times C(\overline{\Omega})$. Moreover $(\bar{y}, \bar{v}, \bar{w})$ is a ε^2-solution of $(P_n^{r,k})$ with $\varepsilon^2 = rn^{-1}$.

Proof. The first part of the proof is immediate if we observe that

$$J_r(y, v, w) \leq J_{r,n}(y, v, w) \leq J_r(y, v, w) + rn^{-1}$$

for every $(y, v, w) \in C(\overline{Q}) \times L^\sigma(\Sigma) \times C(\overline{\Omega})$. Moreover, since $(\bar{y}, \bar{v}, \bar{w})$ is solution of $(P^{r,k})$, with the previous inequalities, we get

$$\begin{aligned} J_{r,n}(\bar{y}, \bar{v}, \bar{w}) &\leq J_r(\bar{y}, \bar{v}, \bar{w}) + rn^{-1} \\ &\leq J_r(y, v, w) + rn^{-1} \\ &\leq J_{r,n}(y, v, w) + rn^{-1}, \end{aligned}$$

for every $y \in C(\overline{Q})$ and every $(v, w) \in (V_{ad}(\bar{v}, k) \times W_{ad}) \cap \mathcal{B}^d_{\lambda(k)}(\bar{v}, \bar{w})$ such that (y, v, w) obeys (1).

$\square$

5.2 Proof of Theorem 2.1 (Pontryagin principle in qualified form).

Let k be a positive constant. Thanks to Proposition 5.2, for every $n \geq 1$, $(\bar{y}, \bar{v}, \bar{w})$ is an ε_n^2-solution of $(P_n^{r,k})$, with $\varepsilon_n^2 = rn^{-1}$. For every $k > 0$, we choose $n(k)$ such that

$$\varepsilon_{n(k)} = (\frac{r}{n(k)})^{\frac{1}{2}} \leq \min(\frac{1}{k^{2\sigma}}, \frac{\lambda(k)}{2}).$$

The metric space $((V_{ad}(\bar{v},k) \times W_{ad}) \cap \mathcal{B}^d_{\lambda(k)}(\bar{v},\bar{w}), d)$ is complete and the functional $(v,w) \longmapsto J_{r,n(k)}(y_{vw},v,w)$ is continuous on this metric space. Thanks to the Ekeland's variational principle, for every $k \geq 1$, there exists $(v_k,w_k) \in (V_{ad}(\bar{v},k) \times W_{ad}) \cap \mathcal{B}^d_{\lambda(k)}(\bar{v},\bar{w})$ such that

$$d((v_k,w_k),(\bar{v},\bar{w})) \leq \varepsilon_{n(k)}, \tag{32}$$

$$J_{r,n(k)}(y_k,v_k,w_k) \leq J_{r,n(k)}(y_{vw},v,w) + \varepsilon_{n(k)} d((v_k,w_k),(v,w)) \tag{33}$$

for every $(v,w) \in (V_{ad}(\bar{v},k) \times W_{ad}) \cap \mathcal{B}^d_{\lambda(k)}(\bar{v},\bar{w})$ (y_k and y_{vw} being the states corresponding respectively to (v_k,w_k) and to (v,w)). The proof is split into five parts.

Step 1. Approximate optimality conditions for the boundary control v_k satisfying (32), (33).
For fixed v_0 in V_{ad}, we denote by v_{0k} ($k > 0$) the function in $V_{ad}(\bar{v},k)$ defined by

$$v_{0k}(s,t) = \begin{cases} v_0(s,t) & \text{if } |v_0(s,t) - \bar{v}(s,t)| \leq k, \\ \bar{v}(s,t) & \text{if not.} \end{cases} \tag{34}$$

Applying Theorem 4.1, we deduce the existence of measurable sets E^k_ρ, such that $\mathcal{L}^N(E^k_\rho) = \rho\mathcal{L}^N(\Sigma)$,

$$y^k_{1,\rho} = y_k + \rho z^1_k + r^k_\rho, \quad \lim_{\rho\to 0} \frac{1}{\rho}\|r^k_\rho\|_{C(\overline{Q})} = 0, \tag{35}$$

$$J(y^k_{1,\rho}, v^k_{1,\rho}, w^k_{1,\rho}) = J(y_k,v_k,w_k) + \rho\Delta J^1_k + o(\rho), \tag{36}$$

where $v^k_{1,\rho}$ and $w^k_{1,\rho}$ are defined by

$$v^k_{1,\rho}(s,t) = \begin{cases} v_k(s,t) & \text{on } \Sigma \setminus E^k_\rho, \\ v_{0k}(s,t) & \text{on } E^k_\rho, \end{cases} \qquad w^k_{1,\rho} = w_k, \tag{37}$$

$y^k_{1,\rho}$ is the state corresponding to $(v^k_{1,\rho}, w^k_{1,\rho})$, z^1_k is the weak solution of

$$\begin{aligned} &\frac{\partial z^1_k}{\partial t} + Az^1_k + f'_y(x,t,y_k)z^1_k = 0 && \text{in } Q, \\ &\frac{\partial z^1_k}{\partial n_A} + g'_y(s,t,y_k,v_k)z^1_k = g(s,t,y_k,v_k) - g(s,t,y_k,v_{0k}) && \text{on } \Sigma, \\ &z^1_k(\cdot,0) = 0 && \text{in } \Omega, \end{aligned}$$

and

$$\begin{aligned} \Delta J^1_k = &\int_Q F'_y(x,t,y_k(x,t))z^1_k(x,t)\,dxdt + \int_\Sigma G'_y(s,t,y_k(s,t),v_k(s,t))z^1_k(s,t)\,dsdt + \\ &+ \int_\Sigma [G(s,t,y_k(s,t),v_{0k}(s,t)) - G(s,t,y_k(s,t),v_k(s,t))]\,dsdt \\ &+ \int_\Omega L'_y(x,y_k(x,T),w_k(x))z^1_k(x)\,dx. \end{aligned}$$

On the other hand, we have

$$d((v^k_{1,\rho},w^k_{1,\rho}),(\bar{v},\bar{w})) \leq d((v^k_{1,\rho},w^k_{1,\rho}),(v_k,w_k)) + d((v_k,w_k),(\bar{v},\bar{w}))$$

$$\leq \mathcal{L}^N(E^k_\rho) + \varepsilon_{n(k)} \leq \rho\mathcal{L}^N(\Sigma) + \varepsilon_{n(k)}.$$

There then exists ρ_k such that, for every $0 < \rho < \rho_k$, we have

$$d((v^k_{1,\rho}, w^k_{1,\rho}), (\bar{v}, \bar{w})) \leq \rho\mathcal{L}^N(\Sigma) + \varepsilon_{n(k)} \leq \lambda(k).$$

Therefore, for every $k > 0$ and every $0 < \rho < \rho_k$, $(v^k_{1,\rho}, w^k_{1,\rho})$ belongs to $(V_{ad}(\bar{v}, k) \times W_{ad}) \cap \mathcal{B}^d_{\lambda(k)}(\bar{v}, \bar{w})$.

If we set $(v, w) = (v^k_{1,\rho}, w^k_{1,\rho})$ in (33), it follows that

$$\limsup_{\rho\to 0} \frac{J_{r,n(k)}(y_k, v_k, w_k) - J_{r,n(k)}(y^k_{1,\rho}, v^k_{1,\rho}, w^k_{1,\rho})}{\rho} \leq \varepsilon_{n(k)}\mathcal{L}^N(\Sigma). \tag{38}$$

Taking (30), (36) and the definition of $J_{r,n}$ into account, we get

$$-\Delta J^1_k - \langle \mu_k, \phi'(y_k)z^1_k\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})} \leq \varepsilon_{n(k)}\mathcal{L}^N(\Sigma) \tag{39}$$

where

$$\mu_k = \begin{cases} \frac{r d_{\mathcal{C}}(\phi(y_k))\nabla d_{\mathcal{C}}(\phi(y_k))}{[d_{\mathcal{C}}(\phi(y_k))^2 + n(k)^{-2}]^{\frac{1}{2}}} & \text{if } d_{\mathcal{C}}(\phi(y_k)) > 0, \\ 0 & \text{if not.} \end{cases} \tag{40}$$

For every $k > 0$, we consider the weak solution p_k of

$$\begin{aligned} &-\frac{\partial p_k}{\partial t} + Ap_k + f'_y(x, t, y_k)p_k = F'_y(x, t, y_k) + [\phi'(y_k)^*\mu_k]|_Q && \text{in } Q, \\ &\frac{\partial p_k}{\partial n_A} + g'_y(s, t, y_k, v_k)p_k = G'_y(s, t, y_k, v_k) + [\phi'(y_k)^*\mu_k]|_\Sigma && \text{on } \Sigma, \\ &p_k(T) = L'_y(x, y_k(T), w_k) + [\phi'(y_k)^*\mu_k]|_{\overline{\Omega}_T} && \text{in } \Omega, \end{aligned} \tag{41}$$

where $[\phi'(y_k)^*\mu_k]|_Q$ is the restriction of $[\phi'(y_k)^*\mu_k]$ to Q, $[\phi'(y_k)^*\mu_k]|_\Sigma$ is the restriction of $[\phi'(y_k)^*\mu_k]$ to Σ and $[\phi'(y_k)^*\mu_k]|_{\overline{\Omega}_T}$ is the restriction of $[\phi'(y_k)^*\mu_k]$ to $\overline{\Omega}_T$. By using the Green formula of Theorem 3.2, we obtain

$$\int_Q F'_y(x, t, y_k)z^1_k\, dxdt + \int_\Sigma G'_y(s, t, y_k, v_k)z^1_k\, dsdt + \int_\Omega L'_y(x, y_k(T), w_k)z^1_k(T)\, dx +$$

$$+\langle \mu_k, \phi'(y_k)z^1_k\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})} =$$

$$= \int_Q p_k \left(\frac{\partial z^1_k}{\partial t} + Az^1_k + f'_y(x, t, y_k)z^1_k\right) dxdt + \int_\Sigma p_k \left(\frac{\partial z^1_k}{\partial n_A} + g'_y(s, t, y_k, v_k)z^1_k\right) dsdt$$

$$= \int_\Sigma p_k\, [g(s, t, y_k, v_k) - g(s, t, y_k, v_{0k})]\, dsdt.$$

With this equality, with (39) and the definition of ΔJ^1_k, we get

$$\int_\Sigma [G(s, t, y_k, v_k) - p_k g(s, t, y_k, v_k)]\, dsdt \leq$$

$$\leq \int_\Sigma [G(s, t, y_k, v_{0k}) - p_k g(s, t, y_k, v_{0k})]\, dsdt + \varepsilon_{n(k)}\mathcal{L}^N(\Sigma)$$

$$\leq \int_\Sigma [G(s,t,y_k,v_{0k}) - p_k g(s,t,y_k,v_{0k})]\,dsdt + \frac{1}{k^{2\sigma}}\mathcal{L}^N(\Sigma) \tag{42}$$

for every $k > 0$ and every $v_0 \in V_{ad}$ (where v_{0k} is defined in function of v_0 in (34)).

Step 2. Approximate optimality conditions for the initial control w_k satisfying (32), (33).
Let w_0 be in W_{ad}, we consider the sequence $(v^k_{2,\rho}, w^k_{2,\rho})$ defined by

$$v^k_{2,\rho} = v_k, \quad w^k_{2,\rho} = w_k + \rho(w_0 - w_k)$$

and we denote by $y^k_{2,\rho}$ the state corresponding to $(v^k_{2,\rho}, w^k_{2,\rho})$. Since W_{ad} is convex, we see that $\{(v^k_{2,\rho}, w^k_{2,\rho}), k > 0\} \subset V_{ad}(\bar{v}, k) \times W_{ad}$. Moreover, we have

$$\begin{aligned}
d((v^k_{2,\rho}, w^k_{2,\rho}), (\bar{v}, \bar{w})) &\leq \rho\|w_0 - w_k\|_{\infty,\Omega} + \varepsilon_{n(k)} \\
&\leq \varepsilon_{n(k)} + \rho(\|w_0 - \bar{w}\|_{\infty,\Omega} + \|w_k - \bar{w}\|_{\infty,\Omega}) \\
&\leq \varepsilon_{n(k)} + \rho(\varepsilon_{n(k)} + \|w_0 - \bar{w}\|_{\infty,\Omega}).
\end{aligned}$$

As in Step 1, if ρ_k is small enough, for every $0 < \rho < \rho_k$, we have:

$$d((v^k_{2,\rho}, w^k_{2,\rho}), (\bar{v}, \bar{w})) \leq \lambda(k).$$

Thus $(v^k_{2,\rho}, w^k_{2,\rho})$ belongs to $(V_{ad}(\bar{v}, k) \times W_{ad}) \cap \mathcal{B}^d_{\lambda(k)}(\bar{v}, \bar{w})$, for every $k > 0$ and for every $0 < \rho < \rho_k$.
Theorem 4.1 then yields

$$y^k_{2,\rho} = y_k + \rho z^2_k + r^k_\rho, \quad \lim_{\rho\to 0} \frac{1}{\rho}\|r^k_\rho\|_{C(\overline{Q})} = 0, \tag{43}$$

$$J(y^k_{2,\rho}, v^k_{2,\rho}, w^k_{2,\rho}) = J(y_k, v_k, w_k) + \rho\Delta J^2_k + o(\rho), \tag{44}$$

where z^2_k is the weak solution of

$$\begin{aligned}
&\frac{\partial z^2_k}{\partial t} + Az^2_k + f'_y(x,t,y_k)z^2_k = 0 && \text{in } Q, \\
&\frac{\partial z^2_k}{\partial n_A} + g'_y(s,t,y_k,v_k)z^2_k = 0 && \text{on } \Sigma, \\
&z^2_k(0) = w_0 - w_k && \text{in } \Omega,
\end{aligned}$$

and

$$\begin{aligned}
\Delta J^2_k = &\int_Q F'_y(x,t,y_k)z^2_k\,dxdt + \int_\Sigma G'_y(s,t,y_k,v_k)z^2_k\,dsdt + \\
&+ \int_\Omega L'_y(x,y_k(T),w_k)z^2_k(T)\,dx + \int_\Omega L'_w(x,y_k(T),w_k)(w_0 - w_k)\,dx.
\end{aligned}$$

As in Step 1, from (33), (44) we deduce that

$$-\Delta J^2_k - \langle \mu_k, \phi'(y_k)z^2_k\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})} = \limsup_{\rho\to 0} \frac{J_{r,n(k)}(y_k,v_k,w_k) - J_{r,n(k)}(y^k_{2,\rho}, v^k_{2,\rho}, w^k_{2,\rho})}{\rho}$$

$$\leq \varepsilon_{n(k)} \|w_0 - w_k\|_{\infty,\Omega} \leq \varepsilon_{n(k)} (\varepsilon_{n(k)} + \|w_0 - \bar{w}\|_{\infty,\Omega}), \tag{45}$$

where μ_k is defined in (40). If we consider the weak solution p_k of (41), still by using the Green formula of Theorem 3.2, we obtain

$$\int_Q F'_y(x,t,y_k) z_k^2 \, dxdt + \int_\Sigma G'_y(s,t,y_k,v_k) z_k^2 \, dsdt + \int_\Omega L'_y(x, y_k(T), w_k) z_k^2(T) \, dx +$$

$$+ \langle [\phi'(y_k)^* \mu_k]|_{\overline{Q} \setminus \overline{\Omega}_0}, z_k^2 \rangle_{\mathcal{M}_b(\overline{Q}\setminus\overline{\Omega}_0) \times C_b(\overline{Q}\setminus\overline{\Omega}_0)} =$$

$$= \int_Q p_k \left(\frac{\partial z_k^2}{\partial t} + A z_k^2 + f'_y(x,t,y_k) z_k^2 \right) dxdt + \int_\Sigma p_k \left(\frac{\partial z_k^2}{\partial n_A} + g'_y(s,t,y_k,v_k) z_k^2 \right) dsdt +$$

$$+ \langle p_k(0), z_k^2(0) \rangle_{\mathcal{M}(\overline{\Omega}) \times C(\overline{\Omega})}$$

$$= \langle p_k(0), w_0 - w_k \rangle_{\mathcal{M}(\overline{\Omega}) \times C(\overline{\Omega})}.$$

Taking (45) and the definition of ΔJ_k^2 into account, we get

$$- \int_\Omega L'_w(x, y_k(T), w_k)(w_0 - w_k) \, dx - \langle p_k(0), w_0 - w_k \rangle_{\mathcal{M}(\overline{\Omega}) \times C(\overline{\Omega})}$$

$$- \langle [\phi'_y(y_k)^* \mu_k]|_{\overline{\Omega}_0}, w_0 - w_k \rangle_{\mathcal{M}(\overline{\Omega}) \times C(\overline{\Omega})} \leq \varepsilon_{n(k)} (\varepsilon_{n(k)} + \|w_0 - \bar{w}\|_{\infty,\Omega}) \tag{46}$$

for every $w_0 \in W_{ad}$.

Step 3. Convergence of sequences $(\mu_k)_k$ and $(p_k)_k$.
We observe that

$$|\mu_k|_{\mathcal{M}(\overline{D})} \leq r d_{\mathcal{C}}(\phi(y_k)) \left[(d_{\mathcal{C}}(\phi(y_k)))^2 + n(k)^{-2} \right]^{-1/2} \leq r.$$

The sequence $(\mu_k)_k$ is bounded in $\mathcal{M}(\overline{D})$, so there exist $\bar{\mu} \in \mathcal{M}(\overline{D})$ and a subsequence, still denoted by $(\mu_k)_k$, such that

$$\mu_k \rightharpoonup \bar{\mu} \quad \text{weak}^* \text{ in } \mathcal{M}(\overline{D}). \tag{47}$$

Let (δ, d) be a pair verifying $(C_{q\sigma})$. From Theorem 3.2, we have

$$\|p_k\|_{L^{\delta'}(0,T;W^{1,d'}(\Omega))} \leq C_4 \{ \|F'_y(\cdot, y_k)\|_{1,Q} + \|G'_y(\cdot, y_k, v_k)\|_{1,\Sigma} + \|L'_y(\cdot, y_k(T), w_k)\|_{1,\Omega} +$$

$$+ |\mu_k|_{\mathcal{M}(\overline{D})} \|\phi'(y_k)\|_{\mathcal{L}(C(\overline{Q}); C(\overline{D}))} \}.$$

(Here $\mathcal{L}(C(\overline{Q}); C(\overline{D}))$ denotes the space of linear continuous mappings from $C(\overline{Q})$ to $C(\overline{D})$.)
Since the sequences $(\mu_k)_k$, $(y_k)_k$, $(v_k)_k$ and $(w_k)_k$ are bounded respectively in $\mathcal{M}(\overline{D})$, $C(\overline{Q})$, $L^\sigma(\Sigma)$ and in $C(\overline{\Omega})$, the sequence $(p_k)_k$ is bounded in $L^{\delta'}(0,T;W^{1,d'}(\Omega))$. There then exist $\bar{p} \in L^{\delta'}(0,T;W^{1,d'}(\Omega))$ and a subsequence, still denoted by $(p_k)_k$, such that $(p_k)_k$ weakly converges to $\bar{p}$ in $L^{\delta'}(0,T;W^{1,d'}(\Omega))$.
Let us prove that $\bar{p}$ is the weak solution of equation (7).
Let φ be in $C^1(\overline{Q})$ satisfying $\varphi(\cdot, 0) = 0$ in $\overline{\Omega}$, for every $k > 0$, we have

$$\int_Q \{ p_k \frac{\partial \varphi}{\partial t} + \sum_{i,j=1}^N a_{ij} D_j p_k D_i \varphi + f'_y(x,t,y_k) p_k \varphi \} \, dxdt + \int_\Sigma \{ p_k \frac{\partial \varphi}{\partial n_A} + g'_y(s,t,y_k,v_k) p_k \varphi \} \, dsdt =$$

$$\int_Q F'_y(x,t,y_k)\varphi\,dxdt + \int_\Sigma G'_y(s,t,y_k,v_k)\varphi\,dsdt + \int_\Omega L'_y(x,y_k(T),w_k)\varphi(T)\,dx$$
$$+\langle \mu_k, \phi'(y_k)\varphi\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})}. \tag{48}$$

Since (δ,d) satisfies $(C_{q\sigma})$, the following imbeddings are continuous:

$$L^{\delta'}(0,T;W^{1,d'}(\Omega)) \hookrightarrow L^{q'}(Q), \quad L^{\delta'}(0,T;W^{1-\frac{1}{d'},d'}(\Gamma)) \hookrightarrow L^{\sigma'}(\Sigma).$$

Therefore, $(p_k)_k$ weakly converges to $\bar p$ in $L^{q'}(Q)$ and the sequence of traces $(p_k|_\Sigma)_k$ weakly converges to the trace $\bar p|_\Sigma$ in $L^{\sigma'}(\Sigma)$. Moreover, since $(v_k)_k$ converges to $\bar v$ in $L^\sigma(\Sigma)$ (indeed, since $d_E(v_k,\bar v)\le \varepsilon_{n(k)} \le \frac{1}{k^{2\sigma}}$ and $|v_k-\bar v|\le k$ a.e. on Σ, we have $\int_\Sigma |v_k-\bar v|^\sigma\,dsdt \le \frac{1}{k^\sigma}$), $(w_k)_k$ converges to $\bar w$ in $C(\overline{\Omega})$ and $(y_k)_k$ converges to $\bar y$ in $C(\overline{Q})$, thanks to assumptions on f,g,F,G,L,ϕ, we have

$$\lim_k \|f'_y(\cdot,y_k)-f'_y(\cdot,\bar y)\|_{q,Q}=0, \quad \lim_k \|F'_y(\cdot,y_k)-F'_y(\cdot,\bar y)\|_{1,Q}=0,$$

$$\lim_k \|g'_y(\cdot,y_k,v_k)-g'_y(\cdot,\bar y,\bar v)\|_{\sigma,\Sigma}=0, \quad \lim_k \|G'_y(\cdot,y_k,v_k)-G'_y(\cdot,\bar y,\bar v)\|_{1,\Sigma}=0,$$

$$\lim_k \|L'_y(\cdot,y_k(T),w_k)-L'_y(\cdot,\bar y(T),\bar w)\|_{1,\Omega}=0, \quad \lim_k \|\phi'(y_k)-\phi'(\bar y)\|_{\mathcal{L}(C(\overline{Q});C(\overline{D}))}=0.$$

Thus, by passing to the limit in (48), it follows that

$$\int_Q \{\bar p\frac{\partial\varphi}{\partial t} + \sum_{i,j=1}^N a_{ij}D_j\bar p D_i\varphi + f'_y(x,t,\bar y)\bar p\varphi\}\,dxdt + \int_\Sigma \{\bar p\frac{\partial\varphi}{\partial n_A} + g'_y(s,t,\bar y,\bar v)\bar p\varphi\}\,dsdt =$$

$$= \int_Q F'_y(x,t,\bar y)\varphi\,dxdt + \int_\Sigma G'_y(s,t,\bar y,\bar v)\varphi\,dsdt + \int_\Omega L'_y(x,\bar y(T),\bar w)\varphi(T)\,dx +$$
$$+\langle \bar\mu, \phi'(\bar y)\varphi\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})} \tag{49}$$

for every $\varphi\in C^1(\overline{Q})$ satisfying $\varphi(\cdot,0)=0$ in $\overline{\Omega}$. Therefore $\bar p$ is the unique weak solution for (7). Since the weak solution of (7) is unique in the sense of definition 3.1, we can deduce by classical arguments that $\bar p$ is independent of the pair (δ,d) (chosen after (47)) and that the original sequence $(p_k)_k$ converges weakly to $\bar p$ in $L^{\delta'}(0,T;W^{1,d'}(\Omega))$ for every (δ,d) satisfying $(C_{q\sigma})$.
To pass to the limit in (46), we prove that

$$(p_k(0)+[\phi'_y(y_k)^*\mu_k]|_{\overline{\Omega}_0})_k \rightharpoonup \bar p(0)+[\phi'_y(\bar y)^*\bar\mu]|_{\overline{\Omega}_0} \quad \text{weakly star in } \mathcal{M}(\overline{\Omega}). \tag{50}$$

For this, let φ be in $C(\overline{\Omega})$ and let y be the solution of

$$\frac{\partial y}{\partial t}+Ay=0 \text{ in } Q, \quad \frac{\partial y}{\partial n_A}=0 \text{ on } \Sigma, \quad y(0)=\varphi \text{ in } \Omega.$$

Thanks to the Green formula of Theorem 3.2, we get

$$\langle p_k(0)+[\phi'_y(y_k)^*\mu_k]|_{\overline{\Omega}_0},\varphi\rangle_{\mathcal{M}(\overline{\Omega})\times C(\overline{\Omega})} - \langle \bar p(0)+[\phi'_y(\bar y)^*\bar\mu]|_{\overline{\Omega}_0},\varphi\rangle_{\mathcal{M}(\overline{\Omega})\times C(\overline{\Omega})} =$$

$$\int_Q \left[\bar p f'_y(x,t,\bar y)y - p_k f'_y(x,t,y_k)y\right]\,dxdt + \int_\Sigma \left[\bar p g'_y(s,t,\bar y,\bar v)y - p_k g'_y(s,t,y_k,v_k)y\right]\,dsdt+$$

$$\langle \mu_k, \phi'_y(y_k)y\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})} - \langle \bar{\mu}, \phi'_y(\bar{y})y\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})}.$$

Now (50) follows from the previous convergence results.

Step 4. Integral Pontryagin's principle.

Notice that $(v_{0k})_k$ tends to v_0 in $L^\sigma(\Sigma)$ and $(v_k)_k$ tends to $\bar{v}$ in $L^\sigma(\Sigma)$. By passing to the limit when k tends to infinity in (42) and (46), and by using the convergence results stated in Step 3, we obtain

$$\int_\Sigma H_\Sigma(s,t,\bar{y}(s,t),\bar{v}(s,t),\bar{p}(s,t))\,dsdt \le \int_\Sigma H_\Sigma(s,t,\bar{y}(s,t),v_0(s,t),\bar{p}(s,t))\,dsdt \tag{51}$$

for every $v_0 \in V_{ad}$, and

$$\int_\Omega L'_w(x,\bar{y}(T),\bar{w}).(\bar{w}-w_0)\,dx + \langle \bar{p}(0) + [\phi'_y(\bar{y})^*\bar{\mu}]|_{\overline{\Omega}_0}, \bar{w}-w_0\rangle_{\mathcal{M}(\overline{\Omega})\times C(\overline{\Omega})} \le 0 \tag{52}$$

for every $w_0 \in W_{ad}$. On the other hand, from the definition of μ_k and from (31), we deduce

$$\langle \mu_k, z-\phi(y_k)\rangle_{\mathcal{M}(\overline{D})\times C(\overline{D})} \le 0 \qquad \forall z \in \mathcal{C}.$$

By passing to the limit in this expression, we obtain (6).

Step 5. Pointwise Pontryagin's principle. Now we prove the pointwise boundary Pontryagin principle. The functions

$$(s,t) \longmapsto \bar{v}(s,t), \qquad (s,t) \longmapsto H_\Sigma(s,t,\bar{y}(s,t),\bar{v}(s,t),\bar{p}(s,t)),$$

are measurable on Σ and the function

$$(s,t,v) \longmapsto H_\Sigma(s,t,\bar{y}(s,t),v,\bar{p}(s,t)),$$

is a Carathéodory function from $\Sigma \times R$ into R. Thanks to Lusin's Theorem and Scorza-Dragoni's Theorem, for every $\epsilon > 0$, there exist a compact subset $\Sigma_\epsilon \subset \Sigma$, continuous mappings $\varphi_0^\epsilon, \varphi_1^\epsilon$ from Σ_ϵ into R and a continuous mapping φ_2^ϵ from $\Sigma_\epsilon \times R$ into R such that

$$\mathcal{L}^N(\Sigma \setminus \Sigma_\epsilon) \le \epsilon,$$

$$\varphi_0^\epsilon(s,t) = \bar{v}(s,t) \ \text{ on } \Sigma_\epsilon,$$

$$\varphi_1^\epsilon(s,t) = H_\Sigma(s,t,\bar{y}(s,t),\bar{v}(s,t),\bar{p}(s,t)) \ \text{ on } \Sigma_\epsilon,$$

$$\varphi_2^\epsilon(s,t,v) = H_\Sigma(s,t,\bar{y}(s,t),v,\bar{p}(s,t)) \ \text{ on } \Sigma_\epsilon \times R.$$

Since $\bar{v}$ is continuous on Σ_ϵ, $\bar{v}$ is bounded on Σ_ϵ and, for $M > \|\bar{v}\|_{\infty,\Sigma_\epsilon}$, the multimapping

$$(s,t) \longmapsto K_M(s,t) := K_V(s,t) \cap [-M,M]$$

has nonempty compact values. From a Lusin type Theorem for measurable multimappings with compact values (see for example [2], Theorem 1.4.1), for every integer $M > \|\bar{v}\|_{\infty,\Sigma_\epsilon}$, there exists a measurable subset $\Sigma_{\epsilon M} \subset \Sigma_\epsilon$ such that $\mathcal{L}^N(\Sigma_\epsilon \setminus \Sigma_{\epsilon M}) \leq \frac{\epsilon}{2^M}$ and the restriction of K_M to $\Sigma_{\epsilon M}$ is continuous. Let us denote by $\widetilde{\Sigma}_{\epsilon M}$ the set of Lebesgue points in $\Sigma_{\epsilon M}$, of the characteristic function of $\Sigma_{\epsilon M}$.

Now let (s_0, t_0) be in $\widetilde{\Sigma}_{\epsilon M}$ and let $v \in K_M(s_0, t_0)$. Since the multimapping K_M is continuous on $\Sigma_{\epsilon M}$ (in fact we only use the lower semicontinuity of K_M), for every integer $n > 0$, the multimapping K_M admits a measurable selection v_n and there exists an increasing function γ from R^+ into R^+ such that

$$|v_n(s,t) - v| \leq \frac{1}{n} \qquad \text{on } B((s_0,t_0), \gamma(\frac{1}{n})) \cap \Sigma_{\epsilon M} \text{ and } \lim_{n\to\infty} \gamma(\frac{1}{n}) = 0,$$

where $B((s_0,t_0),\gamma(\frac{1}{n}))$ is the ball in R^N centered on (s_0,t_0) and of radius $\gamma(\frac{1}{n})$. We now set

$$S_{\epsilon,M,n} = \widetilde{\Sigma}_{\epsilon M} \cap B((s_0,t_0),\gamma(\frac{1}{n})),$$

and we consider the variation

$$v_M(s,t) = \begin{cases} \bar{v}(s,t) & \text{on } \Sigma \setminus S_{\epsilon,M,1}, \\ v_n(s,t) & \text{on } S_{\epsilon,M,n} \setminus S_{\epsilon,M,n+1} \text{ for every } n > 0, \\ v & \text{if } (s,t) = (s_0,t_0). \end{cases}$$

It is clear that $v_M \in V_{ad}$ and that

$$\lim_{(s,t)\to(s_0,t_0)} v_M(s,t) = v.$$

If we take $v_0 = \chi_{S_{\epsilon,M,n}} v_M + (1 - \chi_{S_{\epsilon,M,n}})\bar{v}$ in (51) (where $\chi_{S_{\epsilon,M,n}}$ is the characteristic function of $S_{\epsilon,M,n}$), it follows

$$\frac{1}{\mathcal{L}^N(S_{\epsilon,M,n})} \int_{S_{\epsilon,M,n}} \varphi_1^\varepsilon(s,t)\, dsdt \leq \frac{1}{\mathcal{L}^N(S_{\epsilon,M,n})} \int_{S_{\epsilon,M,n}} \varphi_2^\varepsilon(s,t,v_M(s,t))\, dsdt.$$

(Note that for every $n \geq 1$, $\mathcal{L}^N(S_{\epsilon,M,n}) \neq 0$ because $(s_0,t_0) \in \widetilde{\Sigma}_{\epsilon M}$.) By passing to the limit in the above inequality when n tends to infinity and using the continuity of φ_1^ε and φ_2^ε, we obtain

$$\varphi_1^\varepsilon(s_0,t_0) = H_\Sigma(s_0,t_0,\bar{y}(s_0,t_0),\bar{v}(s_0,t_0),\bar{p}(s_0,t_0)) \leq \varphi_2^\varepsilon(s_0,t_0,v) = H_\Sigma(s_0,t_0,\bar{y}(s_0,t_0),v,\bar{p}(s_0,t_0))$$

for every $(s_0,t_0) \in \widetilde{\Sigma}_{\epsilon M}$ and for every $v \in K_V(s_0,t_0)$ such that $|v| \leq M$. We set

$$\widetilde{\Sigma}_\epsilon = \bigcap_{\substack{M\in N^* \\ M > \|\bar{v}\|_{\infty,\Sigma_\epsilon}}} \widetilde{\Sigma}_{\epsilon M}$$

and we observe that $\mathcal{L}^N(\Sigma \setminus \widetilde{\Sigma}_\epsilon) \leq 2\epsilon$. For every $(s,t) \in \widetilde{\Sigma}_\epsilon$ and every $v \in K_V(s,t)$ we have

$$H_\Sigma(s,t,\bar{y}(s,t),\bar{v}(s,t),\bar{p}(s,t)) \leq H_\Sigma(s,t,\bar{y}(s,t),v,\bar{p}(s,t)).$$

Upon setting $\widetilde{\Sigma} = \bigcup_{\epsilon>0} \widetilde{\Sigma}_\epsilon$, we have $\mathcal{L}^N(\widetilde{\Sigma}) = \mathcal{L}^N(\Sigma)$, the pointwise Pontryagin's principle is satisfied on $\widetilde{\Sigma}$ and the proof is complete. $\square$

5.3 Proof of Pontryagin principle in nonqualified form.

In this case, as in [20], [29], [40], [24], [12], we can consider the penalized functional

$$J_n(y,v,w) = \left\{ \left[(J(y,v,w) - J(\bar{y},\bar{v},\bar{w}) + \frac{1}{n^2})^+ \right]^2 + (d_{\mathcal{C}}(\phi(y)))^2 \right\}^{1/2}.$$

With such a choice, for every $k > 0$, $(\bar{y},\bar{v},\bar{w})$ is a $\frac{1}{n^2}$-solution of the penalized problem

$$(P_n^k) \qquad \inf\{J_n(y,v,w) \mid y \in W(0,T) \cap C(\overline{Q}), (v,w) \in V_{ad}(\bar{v},k) \times W_{ad},\ (y,v,w) \text{ satisfies } (1)\}.$$

As in Section 5.2, for every $k > 0$, we choose $n(k)$ such that

$$\frac{1}{n(k)} \leq \frac{1}{k^{2\sigma}}.$$

Thanks to the Ekeland's principle, there exists $(v_k, w_k) \in V_{ad}(\bar{v},k) \times W_{ad}$ such that

$$d((v_k,w_k),(\bar{v},\bar{w})) \leq \frac{1}{n(k)},$$

$$J_{n(k)}(y_k,v_k,w_k) \leq J_{n(k)}(y_{vw},v,w) + \frac{1}{n(k)} d((v_k,w_k),(v,w))$$

for every $(v,w) \in V_{ad}(\bar{v},k) \times W_{ad}$ (y_k is the solution of (1) corresponding to (v_k,w_k)). With calculations similar to those in [24], [40], [12] by using diffuse perturbations, we get

$$\int_\Sigma H(s,t,y_k(s,t),v_k(s,t),p_k(s,t),\nu_k)\,dsdt \leq \int_\Sigma H(s,t,y_k(s,t),v_{0k}(s,t),p_k(s,t),\nu_k) + \frac{1}{n(k)}\mathcal{L}^N(\Sigma)$$

for every $k > 0$ and every $v_0 \in V_{ad}$ (v_{0k} is defined in function of v_0 in (34)) and

$$\int_\Omega \nu_k L'_w(x,y_k(T),w_k)(w_0 - w_k)\,dx - \langle p_k(0) + [\phi'_y(y_k)^*\mu_k]|_{\overline{\Omega}_0}, w_0 - w_k\rangle_{\mathcal{M}(\overline{\Omega})\times C(\overline{\Omega})} \leq$$

$$\frac{1}{n(k)}\left(\frac{1}{n(k)} + \|w_0 - \bar{w}\|_{\infty,\Omega}\right)$$

for every $w_0 \in W_{ad}$, where

$$\nu_k = \frac{(J(y_k,v_k,w_k) - J(\bar{y},\bar{v},\bar{w}) + \frac{1}{n(k)^2})^+}{J_{n(k)}(y_k,v_k,w_k)},$$

$$\mu_k = \begin{cases} \frac{d_{\mathcal{C}}(\phi(y_k))\nabla d_{\mathcal{C}}(\phi(y_k))}{J_{n(k)}(y_k,v_k,w_k)} & \text{if } \phi(y_k)) \notin \mathcal{C}, \\ 0 & \text{otherwise,} \end{cases}$$

and p_k is the weak solution of

$$-\frac{\partial p_k}{\partial t} + Ap_k + f'_y(x,t,y_k)p_k = \nu_k F'_y(x,t,y_k) + [\phi'_y(y_k)^*\mu_k]|_Q \qquad \text{in } Q,$$

$$\frac{\partial p_k}{\partial n_A} + g'_y(s,t,y_k,v_k)p_k = \nu_k G'_y(s,t,y_k,v_k) + [\phi'_y(y_k)^*\mu_k]|_\Sigma \qquad \text{on } \Sigma,$$

$$p_k(T) = \nu_k L'_y(x, y_k(T), w_k) + [\phi'_y(y_k)^* \mu_k]|_{\overline{\Omega}_T} \qquad \text{in } \Omega.$$

By passing to the limit when k tends to infinity, as in Section 5.2, we finally get the Pontryagin principle in nonqualified form with $\bar{\nu} = \lim_k \nu_k$ and $\bar{\mu}$ the weak star limit of μ_k. To prove that $(\bar{\nu}, \bar{\mu})$ is nonzero, we remark that $\nu_k^2 + |\mu_k|^2_{\mathcal{M}(\overline{D})} = 1$. If $\bar{\nu} > 0$, the proof is complete. If $\bar{\nu} = 0$, we can prove that $|\bar{\mu}|_{\mathcal{M}(\overline{D})} > 0$ by using $\lim_k |\mu_k|_{\mathcal{M}(\overline{D})} = 1$ and $int_{C(\overline{D})}\mathcal{C} \neq \emptyset$. Indeed if $int_{C(\overline{D})}\mathcal{C}$ is nonempty, there exists a ball $B(z;\rho) \subset \mathcal{C}$ with $\rho > 0$ (where $B(z;\rho)$ is the ball in $C(\overline{D})$, centered at z and with radius ρ). We can choose $z_k \in B(0;\rho)$ such that $\langle \mu_k, z_k \rangle_{\mathcal{M}(\overline{D}) \times C(\overline{D})} = \frac{1}{2}\rho |\mu_k|_{\mathcal{M}(\overline{D})}$. Since $z + z_k \in \mathcal{C}$, from the definition of μ_k and from (31), we have

$$\langle \mu_k, z + z_k - \phi(y_k) \rangle_{\mathcal{M}(\overline{D}) \times C(\overline{D})} \leq 0.$$

By passing to the limit, we get

$$\frac{1}{2}\rho + \langle \bar{\mu}, z - \phi(\bar{y}) \rangle_{\mathcal{M}(\overline{D}) \times C(\overline{D})} \leq 0,$$

thus $\bar{\mu} \neq 0$.

References

[1] F. Abergel, R. Temam, Optimality Conditions for Some Nonqualified Problems of Distributed Control, SIAM J. Control Optim. Vol. 27 (1989), p. 1-12.

[2] N. U. Ahmed, K. L. Teo, "Optimal Control of Distributed Parameter Systems", Elsevier North Holland, New York, 1981.

[3] J. J. Alibert, J. P. Raymond, Optimal Control Problems Governed by Semilinear Elliptic Equations with Pointwise State Constraints, preprint 1994.

[4] D. Azé, Some Remarks on Duality in State-Constrained Optimal Control Problems Governed by Elliptic Equations, preprint 1994.

[5] D. Azé, S. Bolintinéanu, A Boundary Control Problem with Parabolic Dynamic and Final Boundary Observation, Proceedings of the 2nd Catalan Days on Applied Mathematics, Presses Universitaires de Perpignan, M. Sofonea and J. N. Corvellec Eds., 1995, p. 12-25.

[6] V. Barbu, "Analysis and Control of Nonlinear Infinite Dimensional Systems", Academic Press, New York, 1993.

[7] M. Bergounioux, Optimal Control of Parabolic Problems with State Constraints: a Penalization Method for Optimality Conditions, Appl. Math. Optim., Vol. 29 (1994), 285-307.

[8] J. F. Bonnans, E. Casas, An Extension of Pontryagin's Principle for State Constrained Optimal Control of Semilinear Elliptic Equations and Variational Inequalities, SIAM J. Cont. and Optim., Vol. 33 (1995).

[9] J. V. Burke, Calmness and Exact Penalization, SIAM J. Control Optim. Vol. 29 (1991), 493-497.

[10] E. Casas, Control of an Elliptic Problem with Pointwise State Constraints, SIAM J. Cont. and Optim., Vol. 24 (1986), p. 1309-1318.

[11] E. Casas, Boundary Control of Semilinear Elliptic Equations with Pointwise State Constraints, SIAM J. Control Optim. Vol. 31 (1993), 993-1006.

[12] E. Casas, Pontryagin's Principle for State-Constrained Boundary Control Problems of Semilinear Parabolic Equations, preprint, 1995.

[13] E. Casas, J. Yong, Maximum Principle for State Constrained Optimal Control Problems Governed by Quasilinear Elliptic Equations, Diff. and Integ. Equat., Vol 8 (1995), p. 1-18.

[14] F. H. Clarke, The Maximum Principle under Minimal Hypotheses, SIAM J. Cont. and Optim., Vol 14 (1976), 1078-1091.

[15] F. H. Clarke, Optimization and Nonsmooth Analysis, John Wiley, New York, 1983.

[16] E. DiBenedetto, "Degenerate Parabolic Equations", Springer-Verlag, New York, 1993.

[17] J. Diestel, "Geometry of Banach Spaces- Selected Topics", Lecture Notes in Mathematics No 485, Springer-Verlag/Berlin/Heidelberg/New York, 1975.

[18] I. Ekeland, On the Variational Principle, J. Math. Anal. Appl., Vol. 47 (1974), p. 324-353.

[19] I. Ekeland, Nonconvex Minimization Problems, Bull. Amer. Math. Soc. N.S., Vol. 1 (1979), p. 443-474.

[20] H. O. Fattorini, A Unified Theory of Necessary Conditions for Nonlinear Nonconvex Control Systems, Appl. Math. Optim., Vol. 15(1987), p. 141-185.

[21] H. O. Fattorini, T. Murphy, Optimal Control Problems for Nonlinear Parabolic Boundary Control Systems: The Dirichlet Boundary Condition, Diff. and Integ. Equat., Vol 7 (1994), p. 1367-1388.

[22] H. O. Fattorini, T. Murphy, Optimal Problems for Nonlinear Parabolic Boundary Control Systems, SIAM J. Control and Optim., Vol. 32 (1994), p. 1577-1596.

[23] H. O. Fattorini, S. Sritharan, Necessary and Sufficient Conditions for Optimal Controls in Viscous Flow Problems, Proceedings of the Royal Society of Edinburgh, 124A (1994), p. 211-251.

[24] B. Hu, J. Yong, Pontryagin Maximum Principle for Semilinear and Quasilinear Parabolic Equations with Pointwise State Constraints, SIAM J. Cont. and Optim. Vol. 33 (1995), p. 1857-1880.

[25] O. A. Ladyzenskaja, V. A. Solonnikov, and N. N. Ural'ceva, "Linear and Quasilinear Equations of Parabolic Type," Translations of Mathematical Monographs, Vol. 23, Amer. Math. Soc., Providence, R. I., 1968.

[26] X. J. Li, Vector Measure and the Necessary Conditions for the Optimal Control Problems of Linear Systems, Proceedings of the Third IFAC Symposium on the Control of Distributed Parameter Systems, Toulouse, France 1982.

[27] X. J. Li, S. N. Chow, Maximum Principle of Optimal Control for Functional Differential Systems, J. Optim. Theory and Applic., Vol. 54 (1987), p. 335-360.

[28] X. J. Li, Y. Yao, Maximum Principle of Distributed Parameter Systems with Time Lags, Proceedings on the Conference on Control Theory of Distributed Parameter Systems and Applications, F. Kappel and K. Kunish Eds., Springer-Verlag, New York, 1985, p.410-427.

[29] X. J. Li, J. Yong, Necessary Conditions for Optimal Control of Distributed Parameter Systems, SIAM J. Control and Optim., Vol. 29 (1991), p. 895-908.

[30] X. J. Li, J. Yong, "Optimal Control Theory for Infinite Dimensional Systems", Birkhäuser, Boston/Basel/Berlin, 1995.

[31] U. Mackenroth, Convex Parabolic Boundary Control Problems with Pointwise State Constraints, J. Math. Anal. Appli., Vol. 87 (1982), p. 256-277.

[32] J. P. Raymond, Nonlinear Boundary Control of Semilinear Parabolic Equations with Pointwise State Constraints, to appear in Disc. Cont. Dyn. Syst.

[33] J. P. Raymond, Optimal Control Problem Governed by a Semilinear Parabolic Equation with Pointwise State Constraints, to appear in Proceedings of IFIP Conference on Modelling and Optimization of Distributed Parameter Systems with Applications to Engineering, July 1995, Warsaw, Poland.

[34] J. P. Raymond, Pontryagin's Principle for State-Constrained Control Problems with Unbounded Controls, Proceedings of the 2nd Catalan Days on Applied Mathematics, Presses Universitaires de Perpignan, M. Sofonea and J. N. Corvellec Eds., 1995, p. 227-237.

[35] J. P. Raymond, H. Zidani, Hamiltonian Pontryagin's Principles for Control Problems Governed by Semilinear Parabolic Equations, preprint 1995.

[36] J. P. Raymond, H. Zidani, Optimal Control Problem Governed by a Semilinear Parabolic Equation, in System Modelling and Optimization, J. Dolezal and J. Fidler Eds., Chapman and Hall, 1996, p. 211-217.

[37] F. Tröltzsch, Optimality Conditions for Parabolic Control Problems and Applications, Teubner-Texte zür Mathematik, 62, B.G. Teubner Verlagsgesellshaft, Leipzig, 1984.

[38] Y. Yao, Vector Measure and Maximum Principle of Distributed Parameter Systems, Sci. Sinica Ser., Vol. 26 (1983), p. 102-112.

[39] Y. Yao, Maximum principle of Semi-linear Distributed Systems, Proceedings of the Third IFAC Symposium on the Control of Distributed Parameter Systems, Toulouse, France 1982.

[40] J. Yong, Pontryagin Maximum Principle for Semilinear Second Order Elliptic Partial Differential Equations and Variational Inequalities with State Constraints, Diff. and Integ. Equat., Vol. 5 (1992), p. 1307-1334.