

## PONTRYAGIN'S PRINCIPLE FOR STATE-CONSTRAINED BOUNDARY CONTROL PROBLEMS OF SEMILINEAR PARABOLIC EQUATIONS\*

EDUARDO CASAS†

**Abstract.** This paper deals with state-constrained optimal control problems governed by semilinear parabolic equations. We establish a minimum principle of Pontryagin's type. To deal with the state constraints, we introduce a penalty problem by using Ekeland's principle. The key tool for the proof is the use of a special kind of spike perturbations distributed in the domain where the controls are defined. Conditions for normality of optimality conditions are given.

**Key words.** Pontryagin principle, boundary control, semilinear parabolic equations, optimality conditions, state constraints

**AMS subject classifications.** 49K20, 35K20

**PII.** S0363012995283637

**1. Introduction.** In the last years, some proofs of minimum principles of Pontryagin's type have appeared. For long time, the optimality conditions for control problems governed by partial differential equations (PDEs) have been given in an integral form, assuming the convexity of the control set and the differentiability with respect to the control and state of all functions involved in the problem. This makes a big difference with the control theory for problems governed by ordinary differential equations (ODEs), where a Pontryagin principle is derived without the previous assumptions. In my opinion, the reason for this difference is the difficulty of extending the methods used for ODEs to infinite-dimensional systems. In particular, the classical spike perturbations of the controls localized around a point do not work properly for PDEs because they lead to some equations with Dirac measures as data, which produce noncontinuous solutions. This makes it difficult to treat the state constraints, especially the pointwise state constraints.

A new type of spike perturbation was developed by a group of mathematicians from Fudan University; see Li [25], Li and Yao [26], and Li and Yong [27]. They used these perturbations to study control problems of evolution equations. The spike perturbations were defined by using the representation of the state given by the corresponding semigroup. This idea was also followed by Fattorini [17], [18]; Fattorini and Frankowska [19]; and Fattorini and Murphy [20], [21]. Later Yong [33] and Casas and Yong [14] built a similar kind of spike perturbations for elliptic equations by using the representation of the solution with the aid of the Green function. Afterwards, Casas suggested a new construction of the set where the perturbations were localized; see Casas [11] and Bei Hu and Yong [22]. This construction was independent of the equation. For a different viewpoint explaining the true nature of this new type of spike perturbations, the reader is referred to Casas [12], where the boundary control of a quasi-linear elliptic equation was considered.

Bonnans and Casas [5], [6] followed a different approach to derive Pontryagin's principle that did not use this type of spike perturbations. However, it was necessary

---

\*Received by the editors March 24, 1995; accepted for publication (in revised form) May 13, 1996. This research was partially supported by Dirección General de Investigación Científica y Técnica (Madrid).

<http://www.siam.org/journals/sicon/35-4/28363.html>

†Departamento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, 39071 Santander, Spain (casas@etsiso.macc.unican.es).

to assume a stability condition of the optimal cost functional with respect to small perturbation of the feasible state set.

In this paper, we consider a boundary control problem governed by a parabolic semilinear equation. General state constraints are included in the formulation of the problem. The idea developed in [12] is used here. To deal with the state constraints we penalize them. The lack of convexity of the control set and the noncontinuity with respect to the control of the functions involved in the control problem make it difficult to formulate a penalty problem having a solution converging to the optimal control of the original problem, however. Ekeland's variational principle is the key to obtaining the suitable penalization.

Pontryagin's principle is often established in a nonqualified form, which implies that the cost functional does not appear in the conditions for optimality. In the absence of equality state constraints, we give a condition that leads to a qualified optimality system. This condition was introduced by Bonnans [4] and Bonnans and Casas [6]. It consists of assuming a certain kind of Lipschitz dependence of the optimal cost functional with respect to small perturbations of the state constraint. It is proved that this condition is satisfied "almost everywhere (a.e.)." We will distinguish strong and weak Pontryagin principles, depending on whether the optimality system is qualified or not. To prove the strong principle we make an exact penalization of the state constraints.

One of the difficulties found in the optimality system is the adjoint state equation. This equation can have measures as data in the domain, on the boundary, and as a final condition. There are not many papers written about parabolic equations involving measures. For these equations the reader is referred to Barbu and Precupanu [1], Lasiecka [24], Tröltzsch [32], and Boccardo and Gallouët [3], the last one dealing with quasi-linear equations. Here we use the transposition method to derive a general result of existence and "uniqueness" of solution. Since we do not assume continuity of the coefficients of the state equation, we need to be precise in which sense the solution is unique; see Serrin [30] for a nonuniqueness result in  $W_0^{1,p}(\Omega)$  ( $p < 2$ ) of an elliptic problem well posed in  $H^1(\Omega)$ .

The paper is organized as follows. In the next section, the control problem is formulated. The state constraints are presented in an abstract framework. We show through some examples how the usual state constraints are included in the abstract formulation. The weak and strong Pontryagin principles are formulated in sections 3 and 4, respectively. In section 5, the state equation is studied and the spike perturbations are defined. The linear parabolic equations involving measures are analyzed in section 6. All the mentioned papers dealing with control of evolution equations, except [22], followed the semigroup approach to analyze the state and adjoint state equations. Here we will follow the variational approach, which allows us to obtain some pointwise information of the solutions of the PDEs. This information is very important for studying the control problems with pointwise state constraints. Finally, the proofs of weak and strong principles are given in section 7.

**2. Setting of the control problem.** Let  $\Omega \subset \mathbb{R}^n$ ,  $n \geq 1$ , be an open and bounded set, with Lipschitz boundary  $\Gamma$ . Given  $0 < T < +\infty$ , we set  $\Omega_T = \Omega \times (0, T)$  and  $\Sigma_T = \Gamma \times (0, T)$ . Let  $(\mathcal{K}, d)$  be a metric space and let us consider a function  $f : \Sigma_T \times \mathbb{R} \times \mathcal{K} \rightarrow \mathbb{R}$  of class  $C^1$  with respect to the second variable and satisfying the following assumptions:

$$(2.1) \quad \frac{\partial f}{\partial y}(x, t, y, u) \leq 0 \quad \forall (x, t, y, u) \in \Sigma_T \times \mathbb{R} \times \mathcal{K};$$

$$(2.2) \quad \begin{cases} \forall M > 0 \exists C_M > 0 \text{ such that } \forall(x, t, u) \in \Sigma_T \times \mathcal{K} \text{ and } |y| \leq M, \\ |f(x, t, 0, u)| + \left| \frac{\partial f}{\partial y}(x, t, y, u) \right| \leq C_M. \end{cases}$$

The state equation is as follows:

$$(2.3) \quad \begin{cases} \frac{\partial y}{\partial t}(x, t) + Ay(x, t) + a_0(x, t, y(x, t)) = 0 & \text{in } \Omega_T, \\ \partial_{\nu_A} y(x, t) = f(x, t, y(x, t), u(x, t)) & \text{on } \Sigma_T, \\ y(x, 0) = y_0(x) & \text{in } \Omega, \end{cases}$$

where  $y_0 \in C(\bar{\Omega})$ ,  $A$  is the linear operator

$$(2.4) \quad \begin{aligned} Ay = & - \sum_{j=1}^n \partial_{x_j} \left\{ \sum_{i=1}^n [a_{ij}(x, t) \partial_{x_i} y(x, t)] + b_j(x, t) y(x, t) \right\} \\ & + \sum_{j=1}^n d_j(x, t) \partial_{x_j} y(x, t) + c(x, t) y(x, t), \end{aligned}$$

and

$$(2.5) \quad \partial_{\nu_A} y(x, t) = \sum_{j=1}^n \left\{ \sum_{i=1}^n [a_{ij}(x, t) \partial_{x_i} y(x, t)] + b_j(x, t) y(x, t) \right\} \nu_j(x),$$

$\nu(x)$  being the outward unit normal vector to  $\Gamma$  at the point  $x$ ; see Casas [9] or Casas and Fernández [13] for an interpretation of this Neumann condition in a trace sense. Function  $a_0 : \Omega_T \times \mathbb{R} \rightarrow \mathbb{R}$  is a Carathéodory function of class  $C^1$  with respect to the second variable and satisfies the following assumptions:

$$(2.6) \quad \begin{cases} \exists \psi_0 \in L^{\hat{p}}([0, T], L^{\hat{q}}(\Omega)) \text{ and } C_1 > 0 \text{ such that} \\ a_0(x, t, y) y \geq \psi_0(x, t) - C_1 y^2 \quad \forall(x, t, y) \in \Omega_T \times \mathbb{R}; \end{cases}$$

$$(2.7) \quad \begin{cases} a_0(\cdot, \cdot, 0) \in L^{\hat{p}}([0, T], L^{\hat{q}}(\Omega)) \text{ and } \forall M > 0 \exists C_M > 0 \text{ such that} \\ \left| \frac{\partial a_0}{\partial y}(x, t, y) \right| \leq C_M \quad \forall(x, t) \in \Omega_T, |y| \leq M; \end{cases}$$

where  $\hat{q}, \hat{p} \in [1, +\infty]$  and  $1/\hat{p} + n/2\hat{q} < 1$ .

As usual, we assume the following hypotheses on  $A$ :

$$(2.8) \quad \begin{cases} a_{ij}, b_j, d_j, c \in L^\infty(\Omega_T) \quad \forall i, j = 1, \dots, n; \\ \sum_{i,j=1}^n a_{ij}(x, t) \xi_i \xi_j \geq \Lambda |\xi|^2 \quad \forall \xi \in \mathbb{R}^n \text{ a.e. } (x, t) \in \Omega_T, \text{ with } \Lambda > 0. \end{cases}$$

Once given the state equation, we introduce the cost functional

$$J(u) = \int_{\Omega_T} L(x, t, y_u(x, t)) dx dt + \int_{\Sigma_T} l(x, t, y_u(x, t), u(x, t)) d\sigma(x) dt,$$

where  $y_u$  is the solution of (2.3) associated with  $u$ ;  $\sigma$  denotes the usual  $(n - 1)$ -dimensional measure on  $\Gamma$  induced by the parametrization (remember that  $\Gamma$  is a Lipschitz manifold); and  $L : \Omega_T \times \mathbb{R} \rightarrow \mathbb{R}$  and  $l : \Sigma_T \times \mathbb{R} \times \mathcal{K} \rightarrow \mathbb{R}$  are of class  $C^1$  with respect to the second variable,  $L$  being measurable with respect to the first one, satisfying

$$(2.9) \quad \begin{cases} \forall M > 0 \exists \psi_{dM} \in L^1(\Omega_T) \text{ such that } \forall (x, t) \in \Omega_T, |y| \leq M, \\ |L(x, t, 0)| + \left| \frac{\partial L}{\partial y}(x, t, y) \right| \leq \psi_{dM}(x, t) \end{cases}$$

and

$$(2.10) \quad \begin{cases} \forall M > 0 \exists \psi_{bM} \in L^1(\Sigma_T) \text{ such that } \forall (x, t, u) \in \Sigma_T \times \mathcal{K}, |y| \leq M, \\ |l(x, t, 0, u)| + \left| \frac{\partial l}{\partial y}(x, t, y, u) \right| \leq \psi_{bM}(x, t). \end{cases}$$

The space of controls  $\mathcal{U}$  is formed by the measurable functions  $u : \Sigma_T \rightarrow \mathcal{K}$  such that the mapping

$$(x, t) \in \Sigma_T \rightarrow (f(x, t, y, u(x, t)), l(x, t, y, u(x, t))) \in \mathbb{R}^2$$

is measurable for every  $y \in \mathbb{R}$ . In section 5 we will prove that there exists a unique solution of (2.3) in the space  $Y = C(\bar{\Omega}_T) \cap L^2([0, T], H^1(\Omega))$  for every  $u \in \mathcal{U}$ , so that functional  $J : \mathcal{U} \rightarrow \mathbb{R}$  is well defined.

Finally we introduce the state constraints. Let  $Z$  be a separable Banach space and  $Q \subset Z$  a closed convex subset with nonempty interior. Given two mappings of class  $C^1$ ,  $G : Y \rightarrow Z$  and  $F : C(\bar{\Omega}_T) \rightarrow \mathbb{R}^s$ ,  $s \geq 1$ , we formulate the optimal control problem as follows:

$$(P) \quad \text{Minimize } \{J(u) : u \in \mathcal{U}, G(y_u) \in Q, F(y_u) = 0\}.$$

Let us show how the usual examples of state constraints can be handled with this formulation.

*Example 2.1.* Given a continuous function  $g : \bar{\Omega}_T \times \mathbb{R} \rightarrow \mathbb{R}$  of class  $C^1$  in respect to the second variable, the constraint  $g(x, t, y_u(x, t)) \leq \delta$  for all  $(x, t) \in \bar{\Omega}_T$ , with  $\delta > 0$  being a given number, can be written in the above framework by putting  $Z = C(\bar{\Omega}_T)$ ,  $G : Y \rightarrow C(\bar{\Omega}_T)$ , defined by  $G(y) = g(\cdot, y(\cdot))$ , and

$$Q = \{z \in C(\bar{\Omega}_T) : z(x, t) \leq \delta \quad \forall (x, t) \in \bar{\Omega}_T\}.$$

*Example 2.2.* Let  $\{(x_j, t_j)\}_{j=1}^s \subset \bar{\Omega}_T$ ; then we can include the equality constraints  $y_u(x_j, t_j) = \delta_j$ ,  $1 \leq j \leq s$ , in the above formulation. Indeed, it is enough to define the functions  $F_j : C(\bar{\Omega}_T) \rightarrow \mathbb{R}$  given by  $F_j(y) = y(x_j) - \delta_j$  and to take  $F = (F_1, \dots, F_s)^T$ . Then  $F$  is of class  $C^1$ .

*Example 2.3.* Let  $g : \Omega \times [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  be a function measurable with respect to the first variable, continuous with respect to the second, of class  $C^1$  with respect to the third, and such that  $\partial g / \partial y$  is also continuous in the last two variables. Moreover, it is assumed that for every  $M > 0$  there exists a function  $\psi_M \in L^1(\Omega)$  such that

$$|g(x, t, 0)| + \left| \frac{\partial g}{\partial y}(x, t, y) \right| \leq \psi_M(x) \quad \text{a.e. } x \in \Omega \quad \forall t \in [0, T] \text{ and } |y| \leq M.$$

Then the constraint

$$\int_{\Omega} g(x, t, y_u(x, t)) dx \leq \delta \quad \forall t \in [0, T]$$

is included in the above formulation by taking  $Z = C[0, T]$ ,

$$Q = \{z \in C[0, T] : z(t) \leq \delta \quad \forall t \in [0, T]\},$$

and  $G : Y \rightarrow C[0, T]$  given by

$$G(y) = \int_{\Omega} g(x, \cdot, y(x, \cdot)) dx.$$

*Example 2.4.* The constraint

$$\int_{\Omega_T} |y_u(x, t)| dx dt \leq \delta$$

is considered by taking  $Z = L^1(\Omega_T)$ ,  $G : Y \rightarrow L^1(\Omega)$ , with  $G(y) = y$ , and  $Q$  the closed ball in  $L^1(\Omega)$  of center at 0 and radius  $\delta$ .

*Example 2.5.* For every  $1 \leq j \leq k$  let  $g_j : \Omega_T \times \mathbb{R} \rightarrow \mathbb{R}$  be a measurable function of class  $C^1$  with respect to the second variable such that for each  $M > 0$  there exists a function  $\eta_M^j \in L^1(\Omega_T)$  satisfying

$$|g_j(x, t, 0)| + \left| \frac{\partial g_j}{\partial y}(x, t, y) \right| \leq \eta_M^j(x, t) \quad \text{a.e. } (x, t) \in \Omega_T \quad \forall |y| \leq M.$$

Then the constraints

$$\int_{\Omega} g_j(x, t, y_u(x, t)) dx dt \leq \delta_j, \quad 1 \leq j \leq k,$$

are included in the formulation of (P) by choosing  $G = (G_1, \dots, G_k)^T$ , with

$$G_j(y) = \int_{\Omega} g_j(x, t, y(x, t)) dx dt,$$

$Z = \mathbb{R}^k$ , and  $Q = (-\infty, \delta_1] \times \dots \times (-\infty, \delta_k]$ .

*Example 2.6.* The equality constraints

$$\int_{\Omega} f_j(x, t, y_u(x, t)) dx = \delta_j, \quad 1 \leq j \leq l,$$

can also be included in problem (P) in the obvious way by assuming the same hypotheses as in Example 2.5.

*Example 2.7.* Integral constraints on the gradient of the state can be considered within our formulation of problem (P):

$$G(y_u) = \int_0^T \int_{\Omega} |\nabla_x y_u(x, t)|^2 dx dt \leq \delta.$$

In this case we can take  $Z = \mathbb{R}$  and  $Q = (-\infty, \delta]$ .

**3. The weak Pontryagin principle.** Before formulating the weak Pontryagin principle, we introduce some notation. Given  $\alpha \geq 0$ , we define the Hamiltonian  $H_\alpha : \Sigma_T \times \mathbb{R} \times \mathcal{K} \times \mathbb{R} \rightarrow \mathbb{R}$  as follows:

$$H_\alpha(x, t, y, u, \varphi) = \alpha l(x, t, y, u) + \varphi f(x, t, y, u).$$

Now we can establish Pontryagin's principle.

**THEOREM 3.1.** *If  $\bar{u} \in \mathcal{U}$  is a solution of (P), then there exist  $\bar{\alpha} \geq 0$ ,  $\bar{y} \in C(\bar{\Omega}_T) \cap L^2([0, T], H^1(\Omega))$ , and  $\bar{\varphi} \in L^r([0, T], W^{1,p}(\Omega))$  for all  $p, r \in [1, 2)$  with  $(2/r) + (n/p) > n + 1$ ,  $\bar{\mu} \in Z'$  and  $\bar{\lambda} \in R^s$  such that*

$$(3.1) \quad \bar{\alpha} + \|\bar{\mu}\|_{Z'} + |\bar{\lambda}| > 0;$$

$$(3.2) \quad \begin{cases} \frac{\partial \bar{y}}{\partial t} + A\bar{y} + a_0(x, t, \bar{y}(x, t)) = 0 & \text{in } \Omega_T, \\ \partial_{\nu_A} \bar{y}(x, t) = f(x, t, \bar{y}(x, t), \bar{u}(x, t)) & \text{on } \Sigma_T, \\ \bar{y}(0) = y_0 & \text{in } \Omega; \end{cases}$$

$$(3.3) \quad \begin{cases} -\frac{\partial \bar{\varphi}}{\partial t} + A^* \bar{\varphi} + \frac{\partial a_0}{\partial y}(x, t, \bar{y}) \bar{\varphi} = \bar{\alpha} \frac{\partial L}{\partial y}(x, t, \bar{y}) \\ \quad + [DG(\bar{y})^* \bar{\mu}]|_{\Omega_T} + [DF(\bar{y})^* \bar{\lambda}]|_{\Omega_T} & \text{in } \Omega_T, \\ \partial_{\nu_{A^*}} \bar{\varphi} = \frac{\partial f}{\partial y}(x, t, \bar{y}, \bar{u}) \bar{\varphi} + \bar{\alpha} \frac{\partial l}{\partial y}(x, t, \bar{y}, \bar{u}) \\ \quad + [DG(\bar{y})^* \bar{\mu}]|_{\Sigma_T} + [DF(\bar{y})^* \bar{\lambda}]|_{\Sigma_T} & \text{on } \Sigma_T, \\ \bar{\varphi}(T) = [DG(\bar{y})^* \bar{\mu}]|_{\bar{\Omega} \times \{T\}} + [DF(\bar{y})^* \bar{\lambda}]|_{\bar{\Omega} \times \{T\}} & \text{in } \bar{\Omega}; \end{cases}$$

$$(3.4) \quad \langle \bar{\mu}, z - G(\bar{y}) \rangle \leq 0 \quad \forall z \in Q;$$

$$(3.5) \quad \begin{aligned} & \int_{\Sigma_T} H_{\bar{\alpha}}(x, t, \bar{y}(x, t), \bar{u}(x, t), \bar{\varphi}(x, t)) d\sigma(x) dt \\ & = \min_{u \in \mathcal{U}} \int_{\Sigma_T} H_{\bar{\alpha}}(x, t, \bar{y}(x, t), u(x, t), \bar{\varphi}(x, t)) d\sigma(x) dt; \end{aligned}$$

where  $A^*$  denotes the formal adjoint operator of  $A$ . Moreover, if one of the following assumptions is satisfied,

(A1) Functions  $f$  and  $l$  are continuous with respect to the third variable on  $(\mathcal{K}, d)$  and this space is separable;

(A2) There exists a set  $\Sigma_T^0 \subset \Sigma_T$ , with  $m_{\Sigma_T}(\Sigma_T^0) = m_{\Sigma_T}(\Sigma_T)$ , such that the function

$$(x, t) \in \Sigma_T \rightarrow (f(x, t, y, u), l(x, t, y, u)) \in \mathbb{R}^2$$

is continuous in  $\Sigma_T^0$  for every  $(y, u) \in \mathbb{R} \times \mathcal{K}$ , then the following pointwise relation holds:

$$(3.6) \quad \begin{aligned} & H_{\bar{\alpha}}(x, t, \bar{y}(x, t), \bar{u}(x, t), \bar{\varphi}(x, t)) \\ & = \min_{u \in \mathcal{K}} H_{\bar{\alpha}}(x, t, \bar{y}(x, t), u, \bar{\varphi}(x, t)) \quad \text{a.e.}[\sigma] \quad x \in \Gamma \quad \text{and a.e. } t \in [0, T]. \end{aligned}$$

*Remark 3.2.* In the previous theorem,  $[DG(\bar{y})]^* \bar{\mu}$  and  $[DF(\bar{y})]^* \bar{\lambda}$  are elements of

$$Y' = C(\bar{\Omega}_T)' + L^2([0, T], H^1(\Omega))' = M(\bar{\Omega}_T) + L^2([0, T], H^1(\Omega)'),$$

where  $M(\bar{\Omega}_T)$  is the space of the real and regular Borel measures in  $\bar{\Omega}_T$ . Let us assume that  $[DG(\bar{y})]^* \bar{\mu} = \bar{\phi} + \bar{\nu}$ , with  $\bar{\phi} \in L^2([0, T], H^1(\Omega)')$  and  $\bar{\nu} \in M(\bar{\Omega}_T)$ , then we can write

$$[DG(\bar{y})]^* \bar{\mu}|_{\Omega_T} = \bar{\phi} + \bar{\nu}|_{\Omega_T}, \quad [DG(\bar{y})]^* \bar{\mu}|_{\Sigma_T} = \bar{\nu}|_{\Sigma_T}, \quad \text{and} \quad [DG(\bar{y})]^* \bar{\mu}|_{\bar{\Omega} \times \{T\}} = \bar{\nu}|_{\bar{\Omega} \times \{T\}}.$$

Analogous considerations can be made for  $[DF(\bar{y})]^* \bar{\lambda}$ .

Let us apply the above principle to the examples given in section 2.

*Example 3.3.* In Example 2.1,  $Z = C(\bar{\Omega}_T)$ ; therefore, the Lagrange multiplier  $\bar{\mu}$  whose existence is established in Theorem 3.1 is a measure in  $\bar{\Omega}_T$ . In this case the transversality condition (3.4) is written as follows:

$$\int_{\bar{\Omega}_T} (z(x, t) - g(x, t, \bar{y}(x, t))) d\bar{\mu}(x, t) \leq 0 \quad \forall z \in C(\bar{\Omega}_T) \text{ with } z(x, t) \leq \delta.$$

From this relation we can deduce that  $\bar{\mu}$  is a positive measure concentrated in the set of points  $(x, t) \in \bar{\Omega}_T$ , where  $g(x, t, \bar{y}(x, t)) = \delta$ . In particular, it could be a Dirac measure or a combination of Dirac measures; see Casas [7].

The adjoint state equation (3.2) now becomes

$$\begin{cases} -\frac{\partial \bar{\varphi}}{\partial t} + A^* \bar{\varphi} + \frac{\partial a_0}{\partial y}(x, t, \bar{y}) \bar{\varphi} = \bar{\alpha} \frac{\partial L}{\partial y}(x, t, \bar{y}) + \frac{\partial g}{\partial y}(x, t, \bar{y}) \bar{\mu}|_{\Omega_T} & \text{in } \Omega_T, \\ \partial_{\nu_{A^*}} \bar{\varphi} = \frac{\partial f}{\partial y}(x, t, \bar{y}, \bar{u}) \bar{\varphi} + \bar{\alpha} \frac{\partial l}{\partial y}(x, t, \bar{y}, \bar{u}) + \frac{\partial g}{\partial y}(x, t, \bar{y}) \bar{\mu}|_{\Sigma_T} & \text{on } \Sigma_T, \\ \bar{\varphi}(T) = \frac{\partial g}{\partial y}(x, T, \bar{y}(x, T)) \bar{\mu}|_{\bar{\Omega} \times \{T\}} & \text{in } \bar{\Omega}. \end{cases}$$

Since  $\partial g / \partial y$  is a continuous function in  $\bar{\Omega}_T$ , then the product  $(\partial g / \partial y) \bar{\mu}$  is well defined and can be identified again with a measure.

*Example 3.4.* In Example 2.2

$$[DF(\bar{y})]^* \bar{\lambda} = \sum_{j=1}^l \bar{\lambda}_j \delta_{(x_j, t_j)}.$$

If the points  $(x_j, t_j)$  are all of them included in  $\Omega_T$ , then the adjoint state equation is

$$\begin{cases} -\frac{\partial \bar{\varphi}}{\partial t} + A^* \bar{\varphi} + \frac{\partial a_0}{\partial y}(x, \bar{y}(x)) \bar{\varphi} = \bar{\alpha} \frac{\partial L}{\partial y}(x, \bar{y}(x)) + \sum_{j=1}^l \bar{\lambda}_j \delta_{(x_j, t_j)} & \text{in } \Omega_T, \\ \partial_{\nu_{A^*}} \bar{\varphi} = \frac{\partial f}{\partial y}(x, t, \bar{y}, \bar{u}) \bar{\varphi} + \bar{\alpha} \frac{\partial l}{\partial y}(x, t, \bar{y}, \bar{u}) & \text{on } \Sigma_T, \\ \bar{\varphi}(T) = 0 & \text{in } \Omega. \end{cases}$$

If some points  $x_j$  are in  $\Gamma$ , then the corresponding term  $\bar{\lambda}_j \delta_{(x_j, t_j)}$  should appear on the Neumann condition. Analogously, if  $t_j = T$  for some index  $j$ , then  $\bar{\lambda} \delta_{(x_j, T)}$  should be included in the final condition.

*Example 3.5.* In Example 2.3, the Lagrange multiplier  $\bar{\mu}$  is a positive Borel measure in  $[0, T]$  concentrated in the set of points  $t$  where the state constraint is active and

$$DG(\bar{y})^* \bar{\mu} = \frac{\partial g}{\partial y}(x, t, \bar{y}(x, t)) \bar{\mu}(t).$$

Then we have the following equation for  $\bar{\varphi}$ :

$$\begin{cases} -\frac{\partial \bar{\varphi}}{\partial t} + A^* \bar{\varphi} + \frac{\partial a_0}{\partial y}(x, t, \bar{y}) \bar{\varphi} = \bar{\alpha} \frac{\partial L}{\partial y}(x, t, \bar{y}) + \frac{\partial g}{\partial y}(x, t, \bar{y}) \bar{\mu}|_{(0, T)} & \text{in } \Omega_T, \\ \partial_{\nu_{A^*}} \bar{\varphi} = \frac{\partial f}{\partial y}(x, t, \bar{y}, \bar{u}) \bar{\varphi} + \bar{\alpha} \frac{\partial l}{\partial y}(x, t, \bar{y}, \bar{u}) & \text{on } \Sigma_T, \\ \bar{\varphi}(T) = \frac{\partial g}{\partial y}(x, T, \bar{y}(x, T)) \bar{\mu}(\{T\}) & \text{in } \Omega. \end{cases}$$

So, in particular, we have that  $\bar{\varphi}(T) = 0$  if the state constraint is not active in  $T$ . This type of state constraints has been studied by many authors; see Barbu and Precupanu [1], Lasiecka [24], and Tröltzsch [32]. All of them consider the semigroup theory approach to deal with the state and adjoint state equations. They prove some regularity of the adjoint state  $\bar{\varphi}$ ; see section 6.

*Example 3.6.* In Example 2.4, the Lagrange multiplier  $\bar{\mu}$  is an element of  $Z' = L^\infty(\Omega_T)$ ; therefore, (3.2) reduces in this case to

$$\begin{cases} -\frac{\partial \bar{\varphi}}{\partial t} + A^* \bar{\varphi} + \frac{\partial a_0}{\partial y}(x, t, \bar{y}) \bar{\varphi} = \bar{\alpha} \frac{\partial L}{\partial y}(x, t, \bar{y}) + \bar{\mu} & \text{in } \Omega_T, \\ \partial_{\nu_{A^*}} \bar{\varphi} = \frac{\partial f}{\partial y}(x, t, \bar{y}, \bar{u}) \bar{\varphi} + \bar{\alpha} \frac{\partial l}{\partial y}(x, t, \bar{y}, \bar{u}) & \text{on } \Sigma_T, \\ \bar{\varphi}(T) = 0 & \text{in } \Omega. \end{cases}$$

In this case, assuming more regularity for the functions  $\psi_{dM}$  and  $\psi_{bM}$  given in (2.8)–(2.9), we can obtain additional regularity for  $\bar{\varphi}$ . For instance, if we take function  $\psi_{bM} \in L^{\hat{p}}([0, T], L^{\hat{q}}(\Omega))$ , then  $\bar{\varphi} \in Y$ .  $H^{2,1}(\Omega)$ -regularity is also obtained provided that  $\Gamma$  is of class  $C^2$  and the coefficients  $a_{ij}$  of  $A$  are Lipschitz in the variable  $x$ .

*Example 3.7.* The Lagrange multipliers in Example 2.5 are positive real numbers  $\{\bar{\mu}_j\}_{j=1}^k$ . The positivity is a consequence of the transversality condition (3.3). The adjoint state equation can be written as follows:

$$\begin{cases} -\frac{\partial \bar{\varphi}}{\partial t} + A^* \bar{\varphi} + \frac{\partial a_0}{\partial y}(x, t, \bar{y}) \bar{\varphi} = \bar{\alpha} \frac{\partial L}{\partial y}(x, t, \bar{y}) + \sum_{j=1}^k \bar{\mu}_j \frac{\partial g_j}{\partial y}(x, t, \bar{y}) & \text{in } \Omega_T, \\ \partial_{\nu_{A^*}} \bar{\varphi} = \frac{\partial f}{\partial y}(x, t, \bar{y}, \bar{u}) \bar{\varphi} + \bar{\alpha} \frac{\partial l}{\partial y}(x, t, \bar{y}, \bar{u}) & \text{on } \Sigma_T, \\ \bar{\varphi}(T) = 0 & \text{in } \Omega. \end{cases}$$

By increasing the regularity of functions  $\eta_j$ , we can improve the regularity of  $\bar{\varphi}$  such as it was described in Example 3.6.

For the equality constraints considered in Example 2.6 the adjoint state equation is similar to the above one. The only difference is that the Lagrange multipliers can be negative.



*Example 3.8.* In Example 2.7, the Lagrange multiplier  $\bar{\mu}$  is a nonnegative real number,  $\bar{\varphi} \in Y$ , and the adjoint state equation is

$$\begin{cases} -\frac{\partial \bar{\varphi}}{\partial t} + A^* \bar{\varphi} + \frac{\partial a_0}{\partial y}(x, t, \bar{y}) \bar{\varphi} = \bar{\alpha} \frac{\partial L}{\partial y}(x, t, \bar{y}) + 2\bar{\mu} \nabla^* \nabla_x y(x, t) & \text{in } \Omega_T, \\ \partial_{\nu_{A^*}} \bar{\varphi} = \frac{\partial f}{\partial y}(x, t, \bar{y}, \bar{u}) \bar{\varphi} + \bar{\alpha} \frac{\partial l}{\partial y}(x, t, \bar{y}, \bar{u}) & \text{on } \Sigma_T, \\ \bar{\varphi}(T) = 0 & \text{in } \Omega, \end{cases}$$

where  $\nabla^* \nabla_x y \in L^2([0, T], H^1(\Omega)')$  is given by

$$\langle \nabla^* \nabla_x y, z \rangle = \int_{\Omega_T} \nabla_x y(x, t) \nabla_x z(x, t) dx dt.$$

The restriction of  $\nabla^* \nabla_x y$  to  $L^2([0, T], H_0^1(\Omega))$  is equal to  $-\Delta_x y$ .

**4. The strong Pontryagin principle.** In this section we will prove that, in the absence of equality constraints, Theorem 3.1 holds with  $\bar{\alpha} = 1$  for “almost all” control problems. We will be precise about this term later. The key to achieving this result is the introduction of a stability assumption of the optimal cost functional with respect to small perturbations of the set of feasible controls. This stability allows us to accomplish an exact penalization of the state constraints. First of all let us formulate the following control problem:

$$(P_\delta) \begin{cases} \text{Minimize } J(u), \\ u \in \mathcal{U}, G(y_u) \in Q_\delta \end{cases}$$

with the same notation and assumptions of section 2 and setting  $Q_\delta = Q + \bar{B}_\delta(0)$  for every  $\delta > 0$ .

DEFINITION 4.1. *We say that  $(P_\delta)$  is strongly stable if there exist  $\epsilon > 0$  and  $C > 0$  such that*

$$(4.1) \quad \inf (P_\delta) - \inf (P_{\delta'}) \leq C(\delta' - \delta) \quad \forall \delta' \in [\delta, \delta + \epsilon].$$

This concept was first introduced in relation with optimal control problems by Bonnans [4]; see also Bonnans and Casas [6]. A weaker stability concept was used by Casas [8] to analyze the convergence of the numerical discretizations of optimal control problems. The following proposition states that almost all problems  $(P_\delta)$  are strongly stable.

PROPOSITION 4.2. *Let  $\delta_0 \geq 0$  be the smallest number such that  $(P_\delta)$  has feasible controls for every  $\delta > \delta_0$ . Then  $(P_\delta)$  is strongly stable for all  $\delta > \delta_0$  except at most a zero Lebesgue measure set.*

*Proof.* It is enough to consider the function  $h : (\delta_0, +\infty) \rightarrow \mathbb{R}$  defined by

$$h(\delta) = \inf (P_\delta)$$

and remark that it is a nonincreasing monotone function and, consequently, differentiable at every point of  $(\delta_0, +\infty)$  except at a zero measure set. Now it is obvious to check that  $(P_\delta)$  is strongly stable at every point where  $h$  is differentiable.  $\square$

Now we state the strong Pontryagin principle.

THEOREM 4.3. *If  $(P_\delta)$  is strongly stable and  $\bar{u}$  is a solution of this problem, then Theorem 3.1 remains to be true with  $\bar{\alpha} = 1$ .*

The proof of this theorem is postponed until section 7.

**5. Analysis of the state equation.** In this section we will see that (2.3) is well posed in  $Y = C(\bar{\Omega}_T) \cap L^2([0, T], H^1(\Omega))$  for every control  $u \in \mathcal{U}$ . Also we will study the variations of the state with respect to some pointwise perturbations of the control which are the crucial point in the proof of Pontryagin's principle. In  $\mathcal{U}$  we consider Ekeland's distance

$$(5.1) \quad d_E(u, v) = m_{\Sigma_T}(\{(x, t) \in \Sigma_T : u(x, t) \neq v(x, t)\}),$$

where  $m_{\Sigma_T}$  is the measure on  $\Sigma_T$  obtained as the product of  $\sigma$  and the Lebesgue measure in the interval  $(0, T)$ . It is easy to check that  $(\mathcal{U}, d_E)$  is a complete metric space. Indeed the proof given by Ekeland [16] can be repeated in this framework.

**THEOREM 5.1.** *Under assumptions (2.1)–(2.8), problem (2.3) has a unique solution in  $Y = C(\bar{\Omega}_T) \cap L^2([0, T], H^1(\Omega))$  for every control  $u \in \mathcal{U}$ . Moreover, there exists a constant  $M > 0$  such that*

$$(5.2) \quad \|y_u\|_\infty + \|y_u\|_{L^2([0, T], H^1(\Omega))} \leq M \quad \forall u \in \mathcal{U}.$$

Finally, if  $\{u_k\}_{k=1}^\infty \subset \mathcal{U}$  is a sequence converging to  $u$  in  $\mathcal{U}$ , i.e.  $d_E(u_k, u) \rightarrow 0$ , then  $\{y_{u_k}\}_{k=1}^\infty$  converges to  $y_u$  strongly in  $Y$ .

*Proof.* The uniqueness of the solution in  $Y$  can be proved by using the Gronwall inequality in the standard way along with the monotonicity of the nonlinear terms. Let us prove the existence.

If  $a_0$  and  $f$  are bounded functions, then the existence and uniqueness of a solution in  $L^\infty([0, T], L^2(\Omega)) \cap L^2([0, T], H^1(\Omega))$  is a consequence of the monotonicity of  $f$  imposed in (2.1) and the condition on  $a_0$  given in (2.6); see Lions [29] or Ladyzhenskaya, Solonnikov, and Ural'tseva [23] for a proof based in Galerkin's approximation of the problem. If  $f$  is not bounded, we can consider the usual truncation of the function

$$f_m(x, t, y, u) = \begin{cases} f(x, t, y, u) & \text{if } |y| \leq m, \\ f(x, t, m, u) & \text{if } y > m, \\ f(x, t, -m, u) & \text{if } y < -m. \end{cases}$$

Thus hypothesis (2.2) implies the boundedness of  $f_m$ .

An analogous modification can be made on  $a_0$ . Then we deduce the existence and uniqueness of a solution  $y_m \in L^\infty([0, T], L^2(\Omega)) \cap L^2([0, T], H^1(\Omega))$  for problem (2.3) with  $a_0$  and  $f$  replaced by  $a_{0m}$  and  $f_m$ , respectively. Now thanks to the assumptions (2.1)–(2.8), we can apply the procedure of Ladyzhenskaya, Solonnikov, and Ural'tseva [23] to deduce the existence of a constant  $M > 0$  independent of  $m$  and  $u \in \mathcal{U}$  such that (5.2) holds for  $y_u$  replaced by  $y_m$ . This implies that

$$a_m(x, t, y_m(x, t)) = a(x, t, y_m(x, t)) \quad \forall m \geq M$$

and

$$f_m(x, t, y_m(x, t), u(x, t)) = f(x, t, y_m(x, t), u(x, t)) \quad \forall m \geq M.$$

Consequently, the uniqueness of a solution of (2.3) lets us obtain the identity  $y_m = y_u$  and the inequality (5.2).

In order to prove the continuity of  $y_u$ , we first suppose that  $y_0 \in C^\theta(\bar{\Omega}_T)$  for some constant  $\theta \in (0, 1]$ . Then, by applying the results of di Benedetto [2], we deduce

that  $y_u \in C^{\beta, \beta/2}(\bar{\Omega}_T)$  for some  $\beta \in (0, \theta]$ . When  $y_0$  is not a Hölder function, we can take a sequence  $\{y_{0k}\}_{k=1}^\infty \subset C^\theta(\bar{\Omega}_T)$  converging uniformly to  $y_0$  in  $\bar{\Omega}_T$ . Then the corresponding solutions of (2.3), denoted by  $y_k$ , are Hölder functions. Now, by applying the methods of [23] is easy to deduce the convergence  $y_k \rightarrow y_u$  in  $L^\infty(\Omega_T)$ , which proves the continuity of  $y_u$ .

Finally, the convergence  $y_{u_k} \rightarrow y_u$  in  $L^2([0, T], H^1(\Omega))$  when  $d_E(u_k, u) \rightarrow 0$  is easily derived. The uniform convergence is obtained again by using the arguments of [23].  $\square$

The rest of the section is devoted to the proof of the following theorem

**THEOREM 5.2.** *Let  $u, v \in \mathcal{U}$ . Given  $\rho \in (0, 1)$ , there exist  $m_{\Sigma_T}$ -measurable sets  $E_\rho \subset \Sigma_T$ , with  $m_{\Sigma_T}(E_\rho) = \rho m_{\Sigma_T}(\Sigma_T)$ , such that if we define*

$$u_\rho(x, t) = \begin{cases} u(x, t) & \text{if } (x, t) \in \Sigma_T \setminus E_\rho, \\ v(x, t) & \text{if } (x, t) \in E_\rho, \end{cases}$$

and if we denote by  $y_\rho$  and  $y$  the states corresponding to  $u_\rho$  and  $u$ , respectively, then the following equalities hold:

$$(5.3) \quad y_\rho = y + \rho z + r_\rho, \quad \lim_{\rho \rightarrow 0} \frac{1}{\rho} \|r_\rho\|_Y = 0,$$

and

$$(5.4) \quad J(u_\rho) = J(u) + \rho z^0 + r_\rho^0, \quad \lim_{\rho \rightarrow 0} \frac{1}{\rho} r_\rho^0 = 0,$$

where  $z \in Y$  satisfies

$$(5.5) \quad \begin{cases} \frac{\partial z}{\partial t} + Az + \frac{\partial a_0}{\partial y}(x, t, y(x, t))z = 0 & \text{in } \Omega_T, \\ \partial_{\nu_A} z = \frac{\partial f}{\partial y}(x, t, y(x, t), u(x, t))z \\ + f(x, t, y(x, t), v(x, t)) - f(x, t, y(x, t), u(x, t)) & \text{on } \Sigma_T, \\ z(x, 0) = 0 & \text{in } \Omega \end{cases}$$

and

$$(5.6) \quad \begin{aligned} z^0 &= \int_{\Omega_T} \frac{\partial L}{\partial y}(x, t, y(x, t))z(x, t) dx dt + \int_{\Sigma_T} \frac{\partial l}{\partial y}(x, t, y(x, t), u(x, t))z(x, t) d\sigma(x) dt \\ &+ \int_{\Sigma_T} [l(x, t, y(x, t), v(x, t)) - l(x, t, y(x, t), u(x, t))] d\sigma(x) dt. \end{aligned}$$

The first step is the proof of the following result

**PROPOSITION 5.3.** *For every  $0 < \rho < 1$  there exists a sequence of  $m_{\Sigma_T}$ -measurable sets  $\{E_k\}_{k=1}^\infty$  satisfying*

- (1)  $E_k = E_\Gamma^k \times J^k$ , with  $E_k \subset \Gamma$  and  $J^k \subset (0, T)$ ,  $\sigma(E_\Gamma^k) = \sqrt{\rho}\sigma(\Gamma)$ , and  $|J^k| = \sqrt{\rho}T$ .
- (2)  $(1/\sqrt{\rho})\chi_{E_\Gamma^k} \rightarrow 1$  \*weakly in  $L^\infty(\Gamma)$ ;  $(1/\sqrt{\rho})\chi_{J^k} \rightarrow 1$  \*weakly in  $L^\infty(0, T)$ ; and  $(1/\rho)\chi_{E_k} \rightarrow 1$  \*weakly in  $L^\infty(\Sigma_T)$ .

*Proof.* We divide the proof into several steps.

*Step 1.* The sets  $E_\Gamma^k$ .

Let us construct the sets  $E_\Gamma^k$ . Since  $\Omega$  is bounded and  $\Gamma$  is a Lipschitz manifold, we can obtain a finite collection of  $\sigma$ -measurable sets  $\{\Gamma_r\}_{r=1}^d$  and functions  $\{a_r\}_{r=1}^d$  satisfying

(i)  $\bigcup_{r=1}^d \Gamma_r = \Gamma$ ,  $\overset{\circ}{\Gamma}_i \cap \overset{\circ}{\Gamma}_j = \emptyset$  if  $i \neq j$  and  $\sigma(\Gamma) = \sum_{r=1}^d \sigma(\overset{\circ}{\Gamma}_r)$ .

(ii) The functions  $a_r : (-\Lambda_\Gamma, +\Lambda_\Gamma)^{n-1} \rightarrow \mathbb{R}$  are Lipschitz, and for some coordinate system  $(x'_r, x_{r,n}) = (x_{r,1}, \dots, x_{r,n})$  in  $\mathbb{R}^n$  we have that

$$\overset{\circ}{\Gamma}_r = \{(x'_r, a_r(x'_r)) : x'_r \in (-\Lambda_\Gamma, +\Lambda_\Gamma)^{n-1}\}$$

and for every set  $E = \{(x'_r, a_r(x'_r)) : x'_r \in F\}$ , with  $F \subset (-\Lambda_\Gamma, +\Lambda_\Gamma)^{n-1}$  Lebesgue measurable, the following identity holds:

$$\sigma(E) = \int_F \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r.$$

For every  $k \in \mathbb{N}$  we decompose the interval  $[-\Lambda_\Gamma, +\Lambda_\Gamma]$  into  $k$  closed subintervals of length  $2\Lambda_\Gamma/k$  and disjoint interiors. Now we make all possible Cartesian products of these subintervals and obtain a family of cubes  $\{Q_{k,i}\}_{i=1}^{k^{n-1}}$  of equal Lebesgue measure, covering  $[-\Lambda_\Gamma, +\Lambda_\Gamma]^{n-1}$  and with disjoint interiors. For every  $r = 1, \dots, d$  and every cube we take a measurable set  $F_{k,j}^r \subset \overset{\circ}{Q}_{k,j}$  such that

$$\int_{F_{k,j}^r} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r = \sqrt{\rho} \int_{Q_{k,j}} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r.$$

Let us see that such an  $F_{k,j}^r$  exists. For every  $t \in [0, 1]$  we define  $Q_{k,j}(t)$  as the cube with the same center as  $Q_{k,j}$  and the length of each side being equal to  $t$  times the length of the sides of  $Q_{k,j}$ . So  $Q_{k,j}(1) = Q_{k,j}$  and  $Q_{k,j}(0)$  is reduced to one point: the center of  $Q_{k,j}$ . Let us consider the function  $g : [0, 1] \rightarrow \mathbb{R}$  defined by

$$g(t) = \int_{Q_{k,j}(t)} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r.$$

Then it is obvious that  $g$  is continuous and

$$0 = g(0) < \sqrt{\rho} \int_{Q_{k,j}} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r < g(1).$$

Therefore there exists  $0 < t_0 < 1$  such that

$$g(t_0) = \sqrt{\rho} \int_{Q_{k,j}} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r.$$

Thus we can choose  $F_{k,j}^r = Q_{k,j}(t_0)$ .

Now we set

$$F_k^r = \bigcup_{i=1}^{k^{n-1}} F_{k,i}^r, \quad E_k^r = \{(x'_r, a_r(x'_r)) : x'_r \in F_k^r\} \subset \overset{\circ}{\Gamma}_r, \quad E_\Gamma^k = \bigcup_{r=1}^d E_k^r.$$

Then

$$\begin{aligned} \sigma(E_\Gamma^k) &= \sum_{r=1}^d \sigma(E_k^r) = \sum_{r=1}^d \int_{F_k^r} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r \\ &= \sqrt{\rho} \sum_{r=1}^d \int_{[-\Lambda_\Gamma, +\Lambda_\Gamma]^{n-1}} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r = \sqrt{\rho} \sum_{r=1}^d \sigma(\Gamma_r^o) = \sqrt{\rho} \sigma(\Gamma). \end{aligned}$$

We are going to prove that

$$(5.7) \quad \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \sigma(A \cap E_\Gamma^k) = \sigma(A) \quad \forall A \subset \Gamma \text{ } \sigma \text{ measurable.}$$

Once this is proved, the convergence  $(1/\sqrt{\rho})\chi_{E_\Gamma^k} \rightarrow 1$  \*weakly in  $L^\infty(\Gamma)$  follows from the density of the simple functions in  $L^1(\Gamma)$ .

First, let us assume that  $A \subset \overset{o}{\Gamma}_r$  is an open set. Let us take the open set  $B \subset (-\Lambda_\Gamma, +\Lambda_\Gamma)^{n-1}$  such that  $A = \{(x'_r, a_r(x'_r)) : x'_r \in B\}$ . Then, from Lemma 5.4 proved below, we deduce

$$\begin{aligned} \sigma(A) &= \int_B \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r \\ &= \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \int_{B \cap F_k^r} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r \\ &= \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \sigma(A \cap E_k^r) = \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \sigma(A \cap E_\Gamma^k). \end{aligned}$$

If  $A \subset \Gamma$  is an open set, then

$$\sigma(A) = \sum_{r=1}^d \sigma(A \cap \overset{o}{\Gamma}_r) = \sum_{r=1}^d \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \sigma(A \cap \overset{o}{\Gamma}_r \cap E_\Gamma^k) = \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \sigma(A \cap E_\Gamma^k).$$

Thus (5.7) holds for every open subset of  $\Gamma$ . Let us take a closed set  $K \subset \Gamma$ ,

$$\begin{aligned} \sigma(K) &= \sigma(\Gamma) - \sigma(\Gamma \setminus K) = \sigma(\Gamma) - \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \sigma([\Gamma \setminus K] \cap E_\Gamma^k) \\ &= \sigma(\Gamma) - \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \{ \sigma(E_\Gamma^k) - \sigma(K \cap E_\Gamma^k) \} = \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \sigma(K \cap E_\Gamma^k). \end{aligned}$$

Finally, let  $A \subset \Gamma$  be a  $\sigma$ -measurable set. Given  $\epsilon > 0$  arbitrary, we can take  $K \subset \Gamma$  closed and  $V \subset \Gamma$  open such that  $K \subset A \subset V$  and

$$\sigma(A) - \epsilon \leq \sigma(K) \leq \sigma(V) \leq \sigma(A) + \epsilon.$$

Then

$$\begin{aligned} \sigma(A) - \epsilon &\leq \sigma(K) \leq \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \sigma(K \cap E_\Gamma^k) \leq \frac{1}{\sqrt{\rho}} \liminf_{k \rightarrow \infty} \sigma(A \cap E_\Gamma^k) \\ &\leq \frac{1}{\sqrt{\rho}} \limsup_{k \rightarrow \infty} \sigma(A \cap E_\Gamma^k) \leq \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \sigma(V \cap E_\Gamma^k) = \sigma(V) \leq \sigma(A) + \epsilon, \end{aligned}$$

which concludes the proof of (5.7).

*Step 2. The sets  $J^k$ .*

To construct the sets  $J^k$ , we decompose the interval  $[0, T]$  into  $k$  closed intervals  $I_j^k$  of length  $T/k$  and disjoint interiors. For each  $j = 1, \dots, k$  we take a subinterval  $J_j^k \subset I_j^k$  of length  $\sqrt{\rho}T/k$  and the same center as  $I_j^k$ . Finally, we define  $J^k$  as the union of the intervals  $\{J_j^k\}_{j=1}^k$ . Then  $|J^k| = \sqrt{\rho}T$  and the convergence  $(1/\sqrt{\rho})\chi_{J^k} \rightarrow 1$  \*weakly in  $L^\infty(0, T)$  can be proved following the same ideas as in the previous step.

*Step 3. The sets  $E_k$ .*

Taking  $E_k = E_\Gamma^k \times J^k$ , it remains to prove the convergence  $(1/\rho)\chi_{E_k} \rightarrow 1$  \*weakly in  $L^\infty(\Sigma_T)$ . Given  $f \in L^1(\Gamma)$  and  $h \in L^1(0, T)$ , we get from Steps 1 and 2 that

$$\begin{aligned} &\lim_{k \rightarrow \infty} \int_{\Sigma_T} \frac{1}{\rho} \chi_{E_k}(x, t) f(x) h(t) dm_{\Sigma_T}(x, t) \\ &= \left( \lim_{k \rightarrow \infty} \int_\Gamma \frac{1}{\sqrt{\rho}} \chi_{E_\Gamma^k}(x) f(x) d\sigma(x) \right) \left( \lim_{k \rightarrow \infty} \int_0^T \frac{1}{\sqrt{\rho}} \chi_{J^k}(t) h(t) dt \right) \\ &= \int_{\Sigma_T} f(x) h(t) dm_{\Sigma_T}(x, t). \end{aligned}$$

Since the functions  $f(x)h(t)$ , with  $f \in L^1(\Gamma)$  and  $h \in L^1(0, T)$ , expand a subspace dense in  $L^1(\Sigma_T)$ , we conclude the proof.  $\square$

LEMMA 5.4. *With the notations of the above proof, the following identity holds for all open sets  $B \subset (-\Lambda_\Gamma, +\Lambda_\Gamma)^{n-1}$ :*

$$(5.8) \quad \int_B \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r = \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \int_{B \cap F_k^r} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r$$

for every  $r = 1, \dots, d$ .

*Proof.* Let us take a sequence  $\{C_k\}_{k=1}^\infty$  of closed cubes with sides parallel to the axes and  $\overset{\circ}{C}_k \cap \overset{\circ}{C}_i = \emptyset$  if  $i \neq k$ , so that  $B = \bigcup_{k=1}^\infty C_k$ ; see Stein [31, pp. 167–170].

Fixed  $r$ , for each cube  $C_l$ , it is obvious that

$$\begin{aligned} \int_{C_l} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r &= \lim_{k \rightarrow \infty} \sum_{Q_{k,j} \subset C_l} \int_{Q_{k,j}} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r \\ &= \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \sum_{Q_{k,j} \subset C_l} \int_{F_{k,j}^r} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r \\ &= \frac{1}{\sqrt{\rho}} \lim_{k \rightarrow \infty} \int_{C_l \cap F_k^r} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r. \end{aligned}$$

Now, given  $\epsilon > 0$  there exists  $k_\epsilon \in \mathbb{N}$  such that

$$\left| \int_B \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r - \sum_{l=1}^{k_\epsilon} \int_{C_l} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r \right| < \epsilon.$$

From here it follows

$$\begin{aligned} & \int_B \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r - \epsilon \\ & \leq \sum_{l=1}^{k_\epsilon} \int_{C_l} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r \\ & = \lim_{k \rightarrow \infty} \frac{1}{\sqrt{\rho}} \sum_{l=1}^{k_\epsilon} \int_{C_l \cap F_k^r} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r \\ & \leq \liminf_{k \rightarrow \infty} \frac{1}{\sqrt{\rho}} \int_{B \cap F_k^r} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r \leq \\ & \limsup_{k \rightarrow \infty} \frac{1}{\sqrt{\rho}} \int_{B \cap F_k^r} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r \\ & \leq \limsup_{k \rightarrow \infty} \frac{1}{\sqrt{\rho}} \sum_{l=1}^{k_\epsilon} \int_{C_l \cap F_k^r} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r + \frac{\epsilon}{\sqrt{\rho}} \\ & = \sum_{l=1}^{k_\epsilon} \int_{C_l} \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r + \frac{\epsilon}{\sqrt{\rho}} \\ & \leq \int_B \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_{r,i}}(x'_r) \right|^2} dx'_r + \left(1 + \frac{1}{\sqrt{\rho}}\right) \epsilon. \end{aligned}$$

Since  $\epsilon > 0$  is arbitrary, the previous relations conclude the proof.  $\square$

Finally, we are ready to prove Theorem 5.2.

*Proof of Theorem 5.2.* Let  $\rho \in (0, 1)$  be fixed. Applying Proposition 5.3, we deduce the existence of measurable sets  $\{E_k\}_{k=1}^\infty$  such that  $m_{\Sigma_T}(E_k) = \rho m_{\Sigma_T}(\Sigma_T)$  and  $(1/\rho)\chi_{E_k} \rightarrow 1$  \*weakly in  $L^\infty(\Sigma_T)$ . For every  $k \in \mathbb{N}$ , we set

$$u_k(x, t) = \begin{cases} u(x, t) & \text{if } (x, t) \in \Sigma_T \setminus E_k, \\ v(x, t) & \text{if } (x, t) \in E_k, \end{cases}$$

and we denote by  $y_k$  and  $y$  the states corresponding to  $u_k$  and  $u$ , respectively. Now, subtracting the equations satisfied by  $y_k$  and  $y$ , and putting  $z_k = (y_k - y)/\rho$  we obtain

$$(5.9) \quad \begin{cases} \frac{\partial z_k}{\partial t} + Az_k + c_k(x, t)z_k = 0 & \text{in } \Omega_T, \\ \partial_{\nu_A} z_k = b_k(x, t)z_k + h(x, t)\frac{1}{\rho}\chi_{E_k} & \text{on } \Sigma_T, \\ z(x, 0) = 0 & \text{in } \Omega, \end{cases}$$

where

$$c_k(x, t) = \int_0^1 \frac{\partial a_0}{\partial y}(x, t, y(x, t) + \tau[y_k(x, t) - y(x, t)])d\tau,$$

$$b_k(x, t) = \int_0^1 \frac{\partial f}{\partial y}(x, t, y(x, t) + \tau[y_k(x, t) - y(x, t)], u_k(x, t))d\tau,$$

and

$$h(x, t) = f(x, t, y(x, t), v(x, t)) - f(x, t, y(x, t), u(x, t)).$$

By subtracting (5.9) and (5.5) and writing  $\zeta_k = z_k - z$ , we deduce

$$(5.10) \quad \begin{cases} \frac{\partial \zeta_k}{\partial t} + A\zeta_k + c_k(x, t)\zeta_k = \left[ \frac{\partial a_0}{\partial y}(x, t, y(x, t)) - c_k(x, t) \right] z & \text{in } \Omega_T, \\ \partial_{\nu_A} \zeta_k = b_k(x, t)\zeta_k + \left[ b_k(x, t) - \frac{\partial f}{\partial y}(x, t, y(x, t), u(x, t)) \right] z & \\ + h(x, t) \left( \frac{1}{\rho}\chi_{E_k} - 1 \right) & \text{on } \Sigma_T, \\ \zeta_k(x, 0) = 0 & \text{in } \Omega. \end{cases}$$

Now we decompose  $\zeta_k = \zeta_k^1 + \zeta_k^2$ , with

$$(5.11) \quad \begin{cases} \frac{\partial \zeta_k^1}{\partial t} + A\zeta_k^1 + c_k(x, t)\zeta_k^1 = \left[ \frac{\partial a_0}{\partial y}(x, t, y(x, t)) - c_k(x, t) \right] z & \text{in } \Omega_T, \\ \partial_{\nu_A} \zeta_k^1 = b_k(x, t)\zeta_k^1 + \left[ b_k(x, t) - \frac{\partial f}{\partial y}(x, t, y(x, t), u(x, t)) \right] z & \text{on } \Sigma_T, \\ \zeta_k^1(x, 0) = 0 & \text{in } \Omega \end{cases}$$

and

$$(5.12) \quad \begin{cases} \frac{\partial \zeta_k^2}{\partial t} + A\zeta_k^2 + c_k(x, t)\zeta_k^2 = 0 & \text{in } \Omega_T, \\ \partial_{\nu_A} \zeta_k^2 = b_k(x, t)\zeta_k^2 + h(x, t) \left( \frac{1}{\rho}\chi_{E_k} - 1 \right) & \text{on } \Sigma_T, \\ \zeta_k^2(x, 0) = 0 & \text{in } \Omega. \end{cases}$$



Taking into account (5.2) and (2.1)–(2.8), multiplying equation (5.12) by the function  $\exp(-\omega t)\zeta_k^2$ , with  $\omega > 0$  large enough, and integrating by parts, we deduce

$$\begin{aligned}
 & C \left( \|\zeta_k^2\|_{L^2(\Omega_T)}^2 + \|\zeta_k^2\|_{L^2([0,T],H^1(\Omega))}^2 \right) \\
 & \leq \frac{\exp(-\omega T)}{2} \|\zeta_k^2(T)\|_{L^2(\Omega)}^2 + \frac{\omega}{2} \int_0^T \exp(-\omega t) \int_{\Omega} |\zeta_k^2(x,t)|^2 dx dt \\
 & + \int_0^T \exp(-\omega t) \langle A\zeta_k^2, \zeta_k^2 \rangle dt + \int_0^T \exp(-\omega t) \int_{\Omega} c_k(x,t) |\zeta_k^2(x,t)|^2 dx dt \\
 & = \int_0^T \int_{\Gamma} \exp(-\omega t) b_k(x,t) |\zeta_k^2(x,t)|^2 d\sigma(x) dt \\
 & + \int_0^T \int_{\Gamma} \exp(-\omega t) h(x,t) \left( \frac{1}{\rho} \chi_{E_k}(x,t) - 1 \right) \zeta_k^2(x,t) d\sigma(x) dt \\
 (5.13) \quad & \leq \int_0^T \int_{\Gamma} \exp(-\omega t) h(x,t) \left( \frac{1}{\rho} \chi_{E_k}(x,t) - 1 \right) \zeta_k^2(x,t) d\sigma(x) dt.
 \end{aligned}$$

From here it follows that

$$(5.14) \quad \|\zeta_k^2\|_{L^2(\Omega_T)}^2 \leq C' \left\| h \left( \frac{1}{\rho} \chi_{E_k} - 1 \right) \right\|_{C^{\beta,\beta/2}(\bar{\Omega}_T)'} \|\zeta_k^2\|_{C^{\beta,\beta/2}(\bar{\Omega}_T)}$$

for some  $\beta \in (0, 1]$ . The Hölder regularity of  $\zeta_k^2$  follows from the assumptions (2.1)–(2.8) and the results of di Benedetto [2].

On the other hand, for  $\theta \in (0, \beta)$ , the inclusions

$$C^{\beta,\beta/2}(\bar{\Omega}_T) \subset C^{\theta,\theta/2}(\bar{\Omega}_T) \subset L^2(\Omega_T)$$

are compact. Then we can apply the Lions lemma [28] to obtain

$$(5.15) \quad \|\zeta_k^2\|_{C^{\theta,\theta/2}(\bar{\Omega}_T)} \leq \epsilon \|\zeta_k^2\|_{C^{\beta,\beta/2}(\bar{\Omega}_T)} + C_{\epsilon} \|\zeta_k^2\|_{L^2(\Omega_T)}.$$

Since  $y$ ,  $y_k$ , and  $h$  are uniformly bounded, the Hölder estimate of  $\zeta_k^2$  can be chosen depending only on  $\rho$ :

$$(5.16) \quad \|\zeta_k^2\|_{C^{\beta,\beta/2}(\bar{\Omega}_T)} \leq C_{\rho} \quad \forall k \in \mathbb{N}.$$

Taking  $\epsilon = \rho/(2[1 + C_{\rho}])$  in (5.15) and using (5.14) and (5.16), it follows

$$\begin{aligned}
 (5.17) \quad \|\zeta_k^2\|_{C^{\theta,\theta/2}(\bar{\Omega}_T)} & \leq \frac{\rho}{2} + C_{\epsilon} \left\{ C' \left\| h \left( \frac{1}{\rho} \chi_{E_k} - 1 \right) \right\|_{C^{\beta,\beta/2}(\bar{\Omega}_T)'} C_{\rho} \right\}^{1/2} \\
 & = \frac{\rho}{2} + C'_{\rho} \left\| h \left( \frac{1}{\rho} \chi_{E_k} - 1 \right) \right\|_{C^{\beta,\beta/2}(\bar{\Omega}_T)'}^{1/2}.
 \end{aligned}$$

Then, for  $\rho$  fixed, the convergence  $(1/\rho)\chi_{E_k} \rightarrow 1$  \*weakly in  $L^\infty(\Sigma_T)$ , the boundedness of  $h$ , and the compactness of the inclusion  $L^\infty(\Sigma_T) \subset C^{\beta,\beta/2}(\bar{\Omega}_T)'$  implies strong convergence  $(1/\rho)h\chi_{E_k} \rightarrow h$  in  $C^{\beta,\beta/2}(\bar{\Omega}_T)'$ . Therefore we can take  $k_\rho \in \mathbb{N}$  large enough in such a way that

$$(5.18) \quad \left| \int_{\Sigma_T} h_0(x, t) \left( \frac{1}{\rho} \chi_{E_k} - 1 \right) d\sigma(x) dt \right| + \left\| h(x, t) \left( \frac{1}{\rho} \chi_{E_k} - 1 \right) \right\|_{C^{\beta,\beta/2}(\bar{\Omega}_T)'} < \frac{\rho^2}{4(1 + C'_\rho)^2} \quad \forall k \geq k_\rho,$$

where

$$h_0(x, t) = l(x, t, y(x, t), v(x, t)) - l(x, t, y(x, t), u(x, t)).$$

Let us set  $E_\rho = E_{k_\rho}$ ,  $u_\rho = u_{k_\rho}$ , and the analogous changes for  $y_\rho, \zeta_\rho, \zeta_\rho^i, i = 1, 2$ . It is obvious that  $d_E(u_\rho, u) \rightarrow 0$  when  $\rho \rightarrow 0$ . Hence Theorem 5.1 implies that  $y_\rho \rightarrow y$  in  $Y$ . This convergence along with the estimates of di Benedetto [2] allow us to deduce from (5.11) the strong convergence  $\zeta_\rho^1 \rightarrow 0$  in  $Y$  when  $\rho \rightarrow 0$ . Combining this with (5.13), (5.17), and (5.18), it is easy to derive the strong convergence  $\zeta_\rho \rightarrow 0$  in  $Y$ , which proves (5.3).

To conclude the proof it is enough to note that

$$\begin{aligned} & \frac{J(u_\rho) - J(u)}{\rho} - z^0 \\ &= \int_{\Omega_T} \left\{ \frac{L(x, t, y_\rho(x, t)) - L(x, t, y(x, t))}{\rho} - \frac{\partial L}{\partial y}(x, t, y(x, t))z(x, t) \right\} dx dt \\ & \quad \int_{\Sigma_T} \left\{ \frac{l(x, t, y_\rho(x, t), u_\rho(x, t)) - l(x, t, y(x, t), u_\rho(x, t))}{\rho} \right. \\ & \quad \left. - \frac{\partial l}{\partial y}(x, t, y(x, t), u(x, t))z(x, t) \right\} d\sigma(x) dt \\ & \quad + \int_{\Sigma_T} h_0(x, t) \left( \frac{1}{\rho} \chi_{E_\rho}(x, t) - 1 \right) d\sigma(x) dt \end{aligned}$$

and to take into account the convergences previously established and (5.18).  $\square$

**6. Linear parabolic equations involving measure data.** Let  $\mu$  be a regular Borel measure in  $\bar{\Omega}_T$ . We can write  $\mu = \mu_{\Omega_T} + \mu_{\Sigma_T} + \mu_T + \mu_0$ , where  $\mu_{\Omega_T} = \mu|_{\Omega_T}$ ,  $\mu_{\Sigma_T} = \mu|_{\Sigma_T}$ ,  $\mu_T = \mu|_{\bar{\Omega} \times \{T\}}$ , and  $\mu_0 = \mu|_{\bar{\Omega} \times \{0\}}$ . The aim of this section is the study of the following problem:

$$(6.1) \quad \begin{cases} -\frac{\partial \varphi}{\partial t} + A^* \varphi = \mu_{\Omega_T} & \text{in } \Omega_T, \\ \partial_{\nu_{A^*}} \varphi = \mu_{\Sigma_T} & \text{on } \Sigma_T, \\ \varphi(T) = \mu_T & \text{in } \bar{\Omega}. \end{cases}$$

The reader is referred to Boccardo and Gallouët [3] for the study of a quasi-linear parabolic equation with a measure in  $\Omega_T$  as a datum. Here we improve the results of [3] by exploiting the linearity of the equation.

Let us denote

$$Y_0 = \{y \in Y : y(x, 0) = 0 \ \forall x \in \Omega\}.$$

DEFINITION 6.1. *Given  $p, r \in [1, 2)$ , with  $(2/r) + (n/p) > n + 1$ , we will say that a function  $\varphi \in L^r([0, T], W^{1,p}(\Omega))$  is a solution of (6.1) if for every  $y \in Y_0 \cap C^1(\bar{\Omega}_T)$*

$$\begin{aligned} & \int_{\Omega_T} \left\{ \frac{\partial y}{\partial t} \varphi + \sum_{j=1}^n \left[ \sum_{i=1}^n a_{ij} \partial_{x_i} y \partial_{x_j} \varphi + b_j y \partial_{x_j} \varphi + d_j \partial_{x_j} y \varphi \right] + cy \varphi \right\} dxdt \\ (6.2) \quad & = \int_{\bar{\Omega}_T} y d\mu(x, t) = \int_{\Omega_T} y d\mu_{\Omega_T}(x, t) + \int_{\Sigma_T} y d\mu_{\Sigma_T}(x, t) + \int_{\bar{\Omega}} y(x, T) d\mu_T(x). \end{aligned}$$

Let us note that (6.2) implies that  $-(\partial\varphi/\partial t) + A^*\varphi = \mu_{\Omega_T}$  in the distribution sense in  $\Omega_T$ . Let us take  $\vec{w} = (w_1, \dots, w_{n+1})$ , with

$$w_i = \sum_{j=1}^n a_{ij} \partial_{x_j} \varphi + d_i \varphi, \quad 1 \leq i \leq n, \quad \text{and} \quad w_{n+1} = \varphi.$$

Then  $\vec{w} \in L^q(\Omega_T)^{n+1}$ ,  $q = \min\{r, p\} < (n + 1)/n$ , and

$$\begin{aligned} \operatorname{div}_{(x,t)} \vec{w} &= \frac{\partial \varphi}{\partial t} + \sum_{i=1}^n \partial_{x_i} \left[ \sum_{j=1}^n a_{ij} \partial_{x_j} \varphi + d_i \varphi \right] = \frac{\partial \varphi}{\partial t} - A^* \varphi + \sum_{i=1}^n b_i \partial_{x_i} \varphi + c \varphi \\ (6.3) \quad &= -\mu_{\Omega_T} + \sum_{i=1}^n b_i \partial_{x_i} \varphi + c \varphi \in M(\Omega_T). \end{aligned}$$

Thus we have  $\vec{w} \in V^q(\Omega_T)$ ,

$$V^q(\Omega_T) = \{\vec{w} \in L^q(\Omega_T)^{n+1} : \operatorname{div}_{(x,t)} \vec{w} \in M(\Omega_T)\}.$$

This space, endowed with the graph norm, is a Banach space. We have the following result.

THEOREM 6.2 (see Casas [10]). *Given  $q \in (1, (n + 1)/n)$ , there exists a unique continuous linear mapping  $\gamma_{\nu_T} : V^q(\Omega_T) \rightarrow W^{-1/q,q}(\partial\Omega_T)$  satisfying*

$$(6.4) \quad \gamma_{\nu_T}(\vec{w}) = \vec{w} \cdot \vec{\nu}_T \quad \forall \vec{w} \in C^1(\bar{\Omega}_T)$$

and

$$\begin{aligned} & \int_{\Omega_T} \vec{w} \cdot \nabla_{(x,t)} \phi dxdt + \langle \operatorname{div}_{(x,t)} \vec{w}, \phi \rangle_{M(\Omega_T), C_b(\Omega_T)} \\ (6.5) \quad & = \langle \gamma_{\nu_T}(\vec{w}), \gamma(\phi) \rangle_{W^{-1/q,q}(\partial\Omega_T), W^{1/q,q'}(\partial\Omega_T)} \quad \forall \phi \in W^{1,q'}(\Omega_T), \end{aligned}$$

where  $C_b(\Omega_T)$  is the space of bounded and continuous functions in  $\Omega_T$  and  $\vec{\nu}_T(x, t)$  is the outward unit normal vector to  $\partial\Omega_T$  at the point  $(x, t)$ .

By applying this theorem to the function  $\vec{w}$  defined above and using (6.2) and (6.3), we have for all  $y \in Y_0 \cap C^1(\bar{\Omega}_T)$

$$\begin{aligned} & \langle \gamma_{\nu_T}(\vec{w}), \gamma(y) \rangle_{W^{-1/q, q}(\partial\Omega_T), W^{1/q, q'}(\partial\Omega_T)} \\ &= \int_{\Omega_T} \vec{w} \cdot \nabla_{(x,t)} y dx dt + \langle \operatorname{div}_{(x,t)} \vec{w}, y \rangle_{M(\Omega_T), C_b(\Omega_T)} \\ &= \int_{\Omega_T} \left\{ \frac{\partial y}{\partial t} \varphi + \sum_{i=1}^n \left[ \sum_{j=1}^n a_{ij} \partial_{x_i} y \partial_{x_j} \varphi + b_i y \partial_{x_i} \varphi + d_i \partial_{x_i} y \varphi \right] + cy \varphi \right\} dx dt \\ & \quad - \int_{\Omega_T} y d\mu_{\Omega_T} = \int_0^T \int_{\Gamma} y d\mu_{\Sigma_T}(x, t) + \int_{\Omega} y(x, T) d\mu_T(x). \end{aligned}$$

From the identity

$$\langle \gamma_{\nu_T}(\vec{w}), \gamma(y) \rangle_{W^{-1/q, q}(\partial\Omega_T), W^{1/q, q'}(\partial\Omega_T)} = \int_0^T \int_{\Gamma} y d\mu_{\Sigma_T}(x, t) + \int_{\Omega} y(x, T) d\mu_T(x)$$

and taking into account that

$$\vec{\nu}_T(x, t) = \begin{pmatrix} \vec{\nu}(x) \\ 0 \end{pmatrix} \quad \forall (x, t) \in \Sigma_T \quad \text{and} \quad \vec{\nu}_T(x, T) = \begin{pmatrix} \vec{0} \\ 1 \end{pmatrix} \quad \forall x \in \Omega,$$

we can identify

$$\partial_{\nu_{A^*}} \varphi = \gamma_{\nu_T}(\vec{w})|_{\Sigma_T} = \mu_{\Sigma_T} \quad \text{and} \quad \varphi(x, T) = \gamma_{\nu_T}(\vec{w})|_{\bar{\Omega} \times \{T\}} = \mu_T.$$

Now we have the following result of existence and uniqueness of solution for problem (6.1).

**THEOREM 6.3.** *There exists a unique function  $\varphi \in L^r([0, T], W^{1,p}(\Omega)) \forall r, p \in [1, 2)$  with  $(2/r) + (n/p) > n + 1$  such that it is a solution of (6.1) and*

$$(6.6) \quad \int_{\Omega_T} \left( \frac{\partial y}{\partial t} + Ay \right) \varphi dx dt + \int_{\Sigma_T} \partial_{\nu_A} y \varphi d\sigma(x) dt = \int_{\bar{\Omega}_T} y d\mu(x, t) \quad \forall y \in Y_0^\infty,$$

with

$$Y_0^\infty = \left\{ y \in Y_0 : \frac{\partial y}{\partial t} + Ay \in L^\infty(\Omega_T) \quad \text{and} \quad \partial_{\nu_A} y \in L^\infty(\Sigma_T) \right\}.$$

Moreover, there exists a constant  $C_{r,p} > 0$  independent of  $\mu$  such that

$$(6.7) \quad \|\varphi\|_{L^r([0,T], W^{1,p}(\Omega))} \leq C_{r,p} \|\mu\|_{M(\bar{\Omega}_T)}.$$

*Proof.* Let  $\{f_k\}_k \subset C(\bar{\Omega}_T)$ ,  $\{g_k\}_k \subset C(\Gamma \times [0, T])$  and  $\{h_k\}_k \subset C(\bar{\Omega})$  such that  $f_k \rightarrow \mu_{\Omega_T}$ ,  $g_k \rightarrow \mu_{\Sigma_T}$ , and  $h_k \rightarrow \mu_T$  \*weakly in  $M(\Omega_T)$ ,  $M(\Sigma_T)$ , and  $M(\bar{\Omega})$ , respectively. Moreover, we can assume that

$$\|f_k\|_{L^1(\Omega_T)} \leq \|\mu_{\Omega_T}\|_{M(\Omega_T)}, \quad \|g_k\|_{L^1(\Sigma_T)} \leq \|\mu_{\Sigma_T}\|_{M(\Sigma_T)}, \quad \text{and} \quad \|h_k\|_{L^1(\Omega)} \leq \|\mu_T\|_{M(\bar{\Omega})}.$$

Let us take  $\varphi_k \in Y$  such that

$$(6.8) \quad \begin{cases} -\frac{\partial \varphi_k}{\partial t} + A^* \varphi_k = f_k \text{ in } \Omega_T, \\ \partial_{\nu_{A^*}} \varphi_k = g_k \text{ on } \Sigma_T, \\ \varphi_k(T) = h_k \text{ in } \Omega. \end{cases}$$

Now for every  $\psi = (\psi_0, \psi_1, \dots, \psi_n) \in \mathcal{D}(\Omega_T)^{n+1}$ , we denote by  $y_\psi$  the solution in  $Y$  of

$$(6.9) \quad \begin{cases} \frac{\partial y}{\partial t} + Ay = \psi_0 - \sum_{j=1}^n \partial_{x_j} \psi_j \text{ in } \Omega_T, \\ \partial_{\nu_A} y = 0 \text{ on } \Sigma_T, \\ y(0) = 0 \text{ in } \Omega. \end{cases}$$

Then

$$(6.10) \quad \begin{aligned} & \int_{\Omega_T} \left( \psi_0 \varphi_k + \sum_{j=1}^n \psi_j \partial_{x_j} \varphi_k \right) dxdt = \int_{\Omega_T} \left( \frac{\partial y_\psi}{\partial t} + Ay_\psi \right) \varphi_k dxdt \\ & = \int_{\Omega_T} \left( -\frac{\partial \varphi_k}{\partial t} + A^* \varphi_k \right) y_\psi dxdt + \int_{\Sigma_T} \partial_{\nu_{A^*}} \varphi_k y_\psi d\sigma(x)dt + \int_{\Omega} \varphi_k(T) y_\psi(T) dx. \end{aligned}$$

Using (6.8) and the properties of  $f_k$ ,  $g_k$ , and  $h_k$ , we deduce from (6.8)

$$(6.11) \quad \begin{aligned} & \int_{\Omega_T} \left( \psi_0 \varphi_k + \sum_{j=1}^n \psi_j \partial_{x_j} \varphi_k \right) dxdt \\ & \leq \|\mu\|_{M(\bar{\Omega}_T)} \|y_\psi\|_{C(\bar{\Omega}_T)} \leq C_{r,p} \|\mu\|_{M(\bar{\Omega}_T)} \sum_{j=0}^n \|\psi_j\|_{L^r([0,T], L^{p'}(\Omega))}, \end{aligned}$$

the last inequality being a consequence of the estimates for the solution of (6.9); see di Benedetto [2] and Ladyzhenskaya, Solonnikov, and Ural'tseva [23]. From the density of the space  $\{\psi_0 - \sum_{j=1}^n \partial_{x_j} \psi_j : \psi \in \mathcal{D}(\Omega_T)^{n+1}\}$  in  $L^r([0, T], W^{1,p}(\Omega)')$  and estimate (6.11) follows the boundedness of  $\{\varphi_k\}_k$  in the space  $L^r([0, T], W^{1,p}(\Omega))$ . Moreover, by taking a subsequence if necessary, we can assume that  $\varphi_k \rightharpoonup \varphi$  weakly in  $L^r([0, T], W^{1,p}(\Omega))$  and (6.7) is satisfied.

Let us prove that  $\varphi$  does not depend on  $r$  and  $p$ . Indeed, passing to the limit in (6.10) and remembering that  $y_\psi(0) = 0$ , we get

$$(6.12) \quad \int_{\Omega_T} \left( \psi_0 \varphi + \sum_{j=1}^n \psi_j \partial_{x_j} \varphi \right) dxdt = \int_{\bar{\Omega}_T} y_\psi d\mu \quad \forall \psi \in \mathcal{D}(\Omega_T)^{n+1}.$$

It is obvious that there is at most one function  $\varphi$  in  $L^1([0, T], W^{1,1}(\Omega))$  satisfying (6.12), which proves that  $\varphi$  is independent of  $r$  and  $p$ .

Given  $y \in Y_0 \cap C^1(\bar{\Omega}_T)$ , multiplying (6.8) by  $y$  and integrating by parts, it follows that

$$\int_{\Omega_T} \left\{ \frac{\partial y}{\partial t} \varphi_k + \sum_{j=1}^n \left[ \sum_{i=1}^n a_{ij} \partial_{x_i} y \partial_{x_j} \varphi_k + b_j y \partial_{x_j} \varphi_k + d_j \partial_{x_j} y \varphi_k \right] + c y \varphi_k \right\} dx dt$$

$$= \int_{\Omega_T} f_k y dx dt + \int_{\Sigma_T} g_k y d\sigma(x) dt + \int_{\Omega} h_k y(T) dx.$$

Now passing to the limit we deduce (6.2) and consequently  $\varphi$  is a solution of (6.1).

Let us prove (6.6). Given  $y \in Y_0^\infty$ , multiplying (6.8) by  $y$  and integrating by parts, we deduce

$$\int_{\Omega_T} f_k y dx dt + \int_{\Sigma_T} g_k y d\sigma(x) dt + \int_{\Omega} h_k y(T) dx$$

$$= \int_{\Omega_T} \left( \frac{\partial y}{\partial t} + A y \right) \varphi_k dx dt + \int_{\Sigma_T} \partial_{\nu_A} y \varphi_k d\sigma(x) dt.$$

Now (6.6) is obtained by passing to the limit.

Finally, the uniqueness of  $\varphi$  follows from (6.6). Indeed, the regularity results for the Neumann problem associated with the operator  $(\partial/\partial t) + A$  (see [2] or [23]) prove the surjectivity of the mapping

$$y \in Y_0^\infty \longrightarrow \left( \frac{\partial y}{\partial t} + A y, \partial_{\nu_A} y \right) \in L^\infty(\Omega_T) \times L^\infty(\Sigma_T).$$

This along with (6.6) implies that the zero function of  $L^r([0, T], W^{1,p}(\Omega))$  is the only one satisfying

$$\int_{\Omega_T} \left( \frac{\partial y}{\partial t} + A y \right) \varphi dx dt + \int_{\Sigma_T} \partial_{\nu_A} y \varphi d\sigma(x) dt = 0 \quad \forall y \in Y_0^\infty.$$

This shows the uniqueness of  $\varphi$ . □

An interesting case arises when  $\mu = g\omega$ , with  $g \in C([0, T], L^2(\Omega))$  and  $\omega \in M[0, T]$

$$\int_{\bar{\Omega}_T} z d\mu = \int_0^T \left( \int_{\Omega} z(x, t) g(x, t) dx \right) d\omega(t) \quad \forall z \in C([0, T], L^2(\Omega));$$

see Example 3.5. In this particular case we have the following result.

**THEOREM 6.4.** *With the above notation, there exists a unique function  $\varphi$  in the space  $L^2([0, T], H^1(\Omega)) \cap L^\infty([0, T], L^2(\Omega))$  solution of the problem*

$$(6.13) \quad \begin{cases} -\frac{\partial \varphi}{\partial t} + A^* \varphi = g\omega & \text{in } \Omega_T, \\ \partial_{\nu_{A^*}} \varphi = 0 & \text{on } \Sigma_T, \\ \varphi(T) = g(T)\omega(\{T\}) & \text{in } \Omega. \end{cases}$$

*Proof.* Uniqueness can be obtained in the standard way. For the proof of the existence we take a sequence  $\{\omega_k\}_k \subset C[0, T]$  converging  $*$ weakly to  $\omega$  in  $M[0, T]$  and satisfying

$$\|\omega_k\|_{L^1(\Omega_T)} \leq \|\omega\|_{M[0, T]}.$$

Let us take  $\varphi_k \in Y$  such that

$$(6.14) \quad \begin{cases} -\frac{\partial \varphi_k}{\partial t} + A^* \varphi_k = g\omega_k \text{ in } \Omega_T, \\ \partial_{\nu_{A^*}} \varphi_k = 0 \text{ on } \Sigma_T, \\ \varphi_k(T) = g(T)\omega(\{T\}) \text{ in } \Omega. \end{cases}$$

Given  $f \in \mathcal{D}(\Omega_T)$ , let us denote by  $y_f$  the solution in  $Y$  of the problem

$$(6.15) \quad \begin{cases} \frac{\partial y}{\partial t} + Ay = f \text{ in } \Omega_T, \\ \partial_{\nu_A} y = 0 \text{ on } \Sigma_T, \\ y(0) = 0 \text{ in } \Omega. \end{cases}$$

Then

$$(6.16) \quad \begin{aligned} \int_{\Omega_T} f\varphi_k dxdt &= \int_{\Omega_T} \left( \frac{\partial y}{\partial t} + Ay \right) \varphi_k dxdt = \int_{\Omega_T} g\omega_k y dxdt + \int_{\Omega} \omega(\{T\})g(T)y(T)dx \\ &\leq \|g\|_{C([0,T],L^2(\Omega))} \|\omega\|_{M[0,T]} \|y\|_{C([0,T],L^2(\Omega))}. \end{aligned}$$

From (6.15) it follows by using the classical arguments that

$$\|y\|_{C([0,T],L^2(\Omega))} \leq C_1 \|f\|_{L^1([0,T],L^2(\Omega))} \quad \text{and} \quad \|y\|_{C([0,T],L^2(\Omega))} \leq C_2 \|f\|_{L^2([0,T],H^1(\Omega)')}.$$

From the first inequality and (6.16) we deduce the boundedness of the sequence  $\{\varphi_k\}_k$  in the space  $L^\infty([0, T], L^2(\Omega))$ . The second inequality leads to the boundedness of the same sequence in  $L^2([0, T], H^1(\Omega))$ . The rest of the proof is easy.  $\square$

As mentioned in section 3, problems of type (6.13) have been studied by Barbu and Precupanu [1], Lasiecka [24], and Tröltzsch [32].

In the case of a measure  $\mu = g\omega$ , with  $g \in L^1[0, T]$  and  $\omega \in \mathcal{M}(\bar{\Omega})$ , we deduce from Theorem 6.3 and the inclusion  $W^{1,p}(\Omega) \subset \mathcal{M}(\bar{\Omega}) \subset W^{1,p'}(\Omega)'$  the existence of a solution  $\varphi \in L^1([0, T], W^{1,p}(\Omega))$  for all  $p \in [1, n/(n - 1))$  and such that  $\partial\varphi/\partial t \in L^1([0, T], W^{1,p'}(\Omega)')$ . Hence we deduce that  $\varphi \in C([0, T], W^{1,p'}(\Omega)')$  after a modification on a set of zero measure.

**7. Proof of Pontryagin principle.** In this section we prove Theorems 3.1 and 4.3. A crucial point in the proofs is the use of Ekeland's variational principle that we state now.

LEMMA 7.1 (see Ekeland [16]). *Let  $(E, d)$  be a complete metric space and  $F : E \rightarrow \mathbb{R} \cup \{+\infty\}$  a lower semicontinuous function, and let  $e_\epsilon \in E$  satisfy*

$$F(e_\epsilon) \leq \inf_{e \in E} F(e) + \epsilon.$$

*Then there exists an element  $\bar{e}_\epsilon \in E$  such that*

$$F(\bar{e}_\epsilon) \leq F(e_\epsilon), \quad d(\bar{e}_\epsilon, e_\epsilon) \leq \sqrt{\epsilon},$$

*and*

$$F(\bar{e}_\epsilon) \leq F(e) + \sqrt{\epsilon}d(e, \bar{e}_\epsilon) \quad \forall e \in E.$$

*Proof of Theorem 3.1.* Since  $Z$  is separable, we can take in  $Z$  a norm  $\|\cdot\|_Z$  such that  $Z'$  endowed with the dual norm  $\|\cdot\|_{Z'}$  is strictly convex. Then the function

$$d_Q : (Z, \|\cdot\|_Z) \longrightarrow \mathbb{R},$$

$$d_Q(z) = \inf_{y \in Q} \|y - z\|_Z$$

is convex, Lipschitz and Gâteaux differentiable at every point  $z \notin Q$ , with  $\partial d_Q(z) = \{\nabla d_Q(z)\}$ , where the Clarke's generalized gradient and the subdifferential in the sense of the convex analysis coincide for this function. Therefore, given  $\xi \in \partial d_Q(y)$ , we have that

$$(7.1) \quad \langle \xi, z - y \rangle + d_Q(y) \leq d_Q(z) \quad \forall z \in Z.$$

Moreover,  $\|\nabla d_Q(z)\|_{Z'} = 1$  for every  $z \notin Q$ ; see Clarke [15] and Casas and Yong [14].

Let us take  $J_\epsilon : \mathcal{U} \longrightarrow \mathbb{R}$  defined by

$$J_\epsilon(u) = \{[(J(u) - J(\bar{u}) + \epsilon)^+]^2 + d_Q(G(y_u))^2 + |F(y_u)|^2\}^{1/2}.$$

It is obvious that  $J_\epsilon(u) > 0$  for every  $u \in \mathcal{U}$  and  $J_\epsilon(\bar{u}) = \epsilon$ . On the other hand, thanks to Theorem 5.1 we have that  $J_\epsilon$  is continuous in  $(\mathcal{U}, d_E)$ , with  $d_E$  defined by (5.1). Therefore we can apply Ekeland's variational principle and deduce the existence of  $u^\epsilon \in \mathcal{U}$  such that

$$(7.2) \quad d_E(u^\epsilon, \bar{u}) \leq \sqrt{\epsilon} \quad \text{and} \quad 0 < J_\epsilon(u^\epsilon) \leq J_\epsilon(u) + \sqrt{\epsilon} d_E(u^\epsilon, u) \quad \forall u \in \mathcal{U}.$$

Given  $v \in \mathcal{U}$  arbitrary, let us take  $E_\rho$  and  $u_\rho^\epsilon$  as in Theorem 5.2,

$$u_\rho^\epsilon(x) = \begin{cases} u^\epsilon(x) & \text{if } x \in \Sigma_T \setminus E_\rho, \\ v(x) & \text{if } x \in E_\rho. \end{cases}$$

Then with the help of (5.3) and (5.4) we get

$$(7.3) \quad -\sqrt{\epsilon} m_{\Sigma_T}(\Sigma) \leq \frac{J_\epsilon(u_\rho^\epsilon) - J_\epsilon(u^\epsilon)}{\rho} = \frac{[(J(u_\rho^\epsilon) - J(\bar{u}) + \epsilon)^+]^2 - [(J(u^\epsilon) - J(\bar{u}) + \epsilon)^+]^2}{\rho[J_\epsilon(u_\rho^\epsilon) + J_\epsilon(u^\epsilon)]} + \frac{d_Q(G(y_\rho^\epsilon))^2 - d_Q(G(y^\epsilon))^2 + |F(y_\rho^\epsilon)|^2 - |F(y^\epsilon)|^2}{\rho[J_\epsilon(u_\rho^\epsilon) + J_\epsilon(u^\epsilon)]}$$

$$\xrightarrow{\rho \rightarrow 0} \{(J(u^\epsilon) - J(\bar{u}) + \epsilon)^+ z^{0,\epsilon} + \langle \xi^\epsilon, DG(y^\epsilon)z^\epsilon \rangle + \langle F(y^\epsilon), DF(y^\epsilon)z^\epsilon \rangle\} / J_\epsilon(u^\epsilon)$$

$$= \alpha_\epsilon z^{0,\epsilon} + \langle [DG(y^\epsilon)]^* \mu^\epsilon, z^\epsilon \rangle + \langle [DF(y^\epsilon)]^* \lambda^\epsilon, z^\epsilon \rangle,$$

where  $y^\epsilon$  and  $y_\rho^\epsilon$  are the states associated with  $u^\epsilon$  and  $u_\rho^\epsilon$ , respectively, and  $z^\epsilon \in Y$  satisfies

$$(7.4) \quad \begin{cases} \frac{\partial z^\epsilon}{\partial t} + Az^\epsilon + \frac{\partial a_0}{\partial y}(x, t, y^\epsilon(x))z^\epsilon = 0 & \text{in } \Omega_T, \\ \partial_{\nu_A} z^\epsilon = \frac{\partial f}{\partial y}(x, t, y^\epsilon(x, t), u^\epsilon(x, t))z^\epsilon \\ + f(x, t, y^\epsilon(x, t), v(x, t)) - f(x, t, y^\epsilon(x, t), u^\epsilon(x, t)) & \text{on } \Sigma_T, \\ z^\epsilon(x, 0) = 0 & \text{in } \Omega, \end{cases}$$



$$z^{0,\epsilon} = \int_{\Omega_T} \frac{\partial L}{\partial y}(x, t, y^\epsilon(x, t))z^\epsilon(x, t)dxdt + \int_{\Sigma_T} \frac{\partial l}{\partial y}(x, t, y^\epsilon(x, t), u^\epsilon(x, t))z^\epsilon(x, t)d\sigma(x)dt$$

$$(7.5) \quad + \int_{\Sigma_T} [l(x, t, y^\epsilon(x, t), v(x, t)) - l(x, t, y^\epsilon(x, t), u(x, t))]d\sigma(x)dt,$$

$$(7.6) \quad \alpha_\epsilon = \frac{(J(u^\epsilon) - J(\bar{u}) + \epsilon)^+}{J_\epsilon(u^\epsilon)}, \quad \mu^\epsilon = \frac{\xi^\epsilon}{J_\epsilon(u^\epsilon)}, \quad \lambda^\epsilon = \frac{F(y^\epsilon)}{J_\epsilon(u^\epsilon)},$$

$$(7.7) \quad \xi^\epsilon = \begin{cases} d_Q(G(y^\epsilon))\nabla d_Q G(y^\epsilon) & \text{if } G(y^\epsilon) \notin Q, \\ 0 & \text{otherwise.} \end{cases}$$

By using Theorem 6.3, we can take a function  $\varphi^\epsilon \in L^r([0, T], W^{1,p}(\Omega)) \forall r, p \in [1, 2)$  with  $(2/r) + (n/p) > n + 1$  such that

$$(7.8) \quad \begin{cases} -\frac{\partial \varphi^\epsilon}{\partial t} + A^* \varphi^\epsilon + \frac{\partial a_0}{\partial y}(x, t, y^\epsilon)\varphi^\epsilon = \alpha_\epsilon \frac{\partial L}{\partial y}(x, t, y^\epsilon) \\ \quad + [DG(y^\epsilon)^* \mu_\epsilon]_{|\Omega_T} + [DF(y^\epsilon)^* \lambda_\epsilon]_{|\Omega_T} \text{ in } \Omega_T, \\ \partial_{\nu_{A^*}} \varphi^\epsilon = \frac{\partial f}{\partial y}(x, t, y^\epsilon, u^\epsilon)\varphi^\epsilon + \alpha_\epsilon \frac{\partial l}{\partial y}(x, t, y^\epsilon, u^\epsilon) \\ \quad + [DG(y^\epsilon)^* \mu_\epsilon]_{|\Sigma_T} + [DF(y^\epsilon)^* \lambda_\epsilon]_{|\Sigma_T} \text{ on } \Sigma_T, \\ \varphi^\epsilon(T) = [DG(y^\epsilon)^* \mu_\epsilon]_{|\Omega \times \{T\}} + [DF(y^\epsilon)^* \lambda_\epsilon]_{|\Omega \times \{T\}} \text{ in } \Omega. \end{cases}$$

Thanks to the assumptions (2.2) and (2.7), we have that  $z^\epsilon \in Y_0^\infty$ . Then we can apply (6.6) with  $y = z^\epsilon$  and deduce from (7.3)–(7.5) and the definition of  $H_\alpha$  given in section 3 the inequality

$$(7.9) \quad \int_{\Sigma_T} H_{\alpha_\epsilon}(x, t, y^\epsilon(x, t), u^\epsilon(x, t), \varphi^\epsilon(x, t))d\sigma(x)dt \leq \int_{\Sigma_T} H_{\alpha_\epsilon}(x, t, y^\epsilon(x, t), v(x, t), \varphi^\epsilon(x, t))d\sigma(x)dt + \sqrt{\epsilon}m_{\Sigma_T}(\Sigma_T) \quad \forall v \in \mathcal{U}.$$

Now we pass to the limit when  $\epsilon \rightarrow 0$ . To do this, let us remark that

$$(7.10) \quad \alpha_\epsilon^2 + \|\mu^\epsilon\|_{Z'}^2 + |\lambda^\epsilon|^2 = 1.$$

Then we take subsequences, denoted in the same way, satisfying

$$(7.11) \quad \begin{cases} \alpha_\epsilon \rightarrow \bar{\alpha} \text{ in } \mathbb{R}, \quad \lambda^\epsilon \rightarrow \bar{\lambda} \text{ in } \mathbb{R}^n, \\ \mu^\epsilon \rightarrow \bar{\mu} \text{ in the *weak topology of } Z'. \end{cases}$$

On the other hand, the convergence  $y^\epsilon \rightarrow \bar{y}$  in  $Y$  follows from Theorem 5.1. The boundedness of  $\{\varphi^\epsilon\}$  in  $L^r([0, T], W^{1,p}(\Omega))$  follows from (6.7) and (7.10). Then, using (7.11), it is easy to pass to the limit in (7.8) and (7.9) and to deduce (3.3) and (3.5). Now remembering the definition of  $\mu^\epsilon$  and  $\xi^\epsilon$  and (7.1), we deduce

$$(7.12) \quad \langle \mu^\epsilon, z - G(y^\epsilon) \rangle \leq 0 \quad \forall z \in Q.$$

Passing to the limit in this expression we obtain (3.4). Let us prove (3.1). To do this, let us suppose that  $\bar{\alpha} = |\bar{\lambda}| = 0$ ; then from (7.10) it follows  $\|\mu^\epsilon\|_{Z'} \rightarrow 1$  as  $\epsilon \rightarrow 0$ . Let us take  $z_0 \in \overset{\circ}{Q}$  and  $\rho > 0$  such that  $\bar{B}_\rho(z_0) \subset \overset{\circ}{Q}$ . Then (7.12) implies that

$$\langle \mu^\epsilon, z + z_0 - G(y^\epsilon) \rangle \leq 0 \quad \forall z \in \bar{B}_\rho(0).$$

Hence

$$\rho \|\mu^\epsilon\|_{Z'} = \sup_{z \in B_\rho(0)} \langle \mu^\epsilon, z \rangle \leq \langle \mu^\epsilon, G(y^\epsilon) - z_0 \rangle.$$

Passing to the limit

$$0 < \rho \leq \lim_{\epsilon \rightarrow 0} \langle \mu^\epsilon, G(y^\epsilon) - z_0 \rangle = \langle \bar{\mu}, G(\bar{y}) - z_0 \rangle,$$

which proves that  $\bar{\mu} \neq 0$ .

It remains to prove (3.6); see Bonnans and Casas [5] or Casas [11] for the study of analogous situations. To do this we consider the coordinate system  $\{(\Gamma_r, a_r)\}_{r=1}^d$  of  $\Gamma$  introduced in the proof of Proposition 5.3. Given a point  $x_0 \in \overset{\circ}{\Gamma}_r$  for some  $1 \leq r \leq d$  we denote for each  $\epsilon > 0$  small enough

$$\Gamma_\epsilon(x_0) = \{x = (x'_r, a_r(x'_r)) : x'_r \in B_\epsilon(x'_{0r}) \subset (0, 1)^{n-1}\},$$

where  $B_\epsilon(x'_{0r})$  is the ball in  $\mathbb{R}^{n-1}$  centered at  $x'_{0r}$  and having radius  $\epsilon$ . Now given  $0 < t_0 < T$ , we set

$$\Sigma_T^\epsilon(x_0, t_0) = \Gamma_\epsilon(x_0) \times (t_0 - \epsilon, t_0 + \epsilon).$$

The following lemma is used in this proof.

LEMMA 7.2. *Given  $f \in L^1(\Sigma_T)$ , there exists a  $m_{\Sigma_T}$ -measurable set  $S \subset \bigcup_{r=1}^d \overset{\circ}{\Gamma}_r \times (0, T)$ , with  $m_{\Sigma_T}(S) = m_{\Sigma_T}(\Sigma_T)$ , such that for every  $(x_0, t_0) \in S$  we have*

$$(7.13) \quad \lim_{\epsilon \rightarrow 0} \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} |f(x, t) - f(x_0, t_0)| dm_{\Sigma_T}(x, t) = 0.$$

*Proof.* Let us denote for all  $(x'_r, t) \in (0, 1)^{n-1} \times (0, T)$

$$\omega_r(x'_r) = \sqrt{1 + \sum_{i=1}^{n-1} \left| \frac{\partial a_r}{\partial x_i}(x'_r) \right|^2} \quad \text{and} \quad f_r(x'_r, t) = \omega_j(x'_r) f(x'_r, a_r(x'_r), t).$$

Since  $\omega_r$  and  $f_r$  are Lebesgue integrable functions in  $(0, 1)^{n-1}$  and  $(0, 1)^{n-1} \times (0, T)$ , respectively, we know that the set of Lebesgue points of these functions  $U_r$  and  $V_r$ , respectively, have measure equal to 1 and  $T$ , respectively. Let us define

$$S_r = \{(x, t) \in V_r = (x'_r, a_r(x'_r), t) : x'_r \in U_r\} \quad \text{and} \quad S = \bigcap_{r=1}^d S_r.$$

Then  $m_{\Sigma_T}(S) = m_{\Sigma_T}(\Sigma_T)$  and  $S_r \subset \overset{\circ}{\Gamma}_r \times (0, T)$ ,  $1 \leq r \leq d$ .

Let us take  $(x_0, t_0) = (x'_{0r}, a(x'_{0r}), t_0) \in S_r$ . Then  $x'_{0j}$  and  $(x_0, t_0)$  are Lebesgue points of  $\omega_r$  and  $f_r$ ; consequently,

$$\begin{aligned} & \lim_{\epsilon \rightarrow 0} \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} |f(x, t) - f(x_0, t_0)| dm_{\Sigma_T}(x, t) \\ &= \lim_{\epsilon \rightarrow 0} \left( \frac{1}{2\epsilon |B_\epsilon(x'_{0r})|} \int_{t_0-\epsilon}^{t_0+\epsilon} \int_{B_\epsilon(x'_{0r})} |f_r(x'_r, t) - f_r(x'_{0r}, t)| dx'_r dt \right), \end{aligned}$$

$$\begin{aligned} & \lim_{\epsilon \rightarrow 0} \left( \frac{1}{|B_\epsilon(x'_{0r})|} \int_{B_\epsilon(x'_{0r})} \omega_r(x'_r) dx'_r \right)^{-1} \\ & = f_r(x'_{0r}, t_0) / \omega_r(x'_{0r}) = f(x_0, t_0), \end{aligned}$$

where  $|B_\epsilon(x'_{0r})|$  denotes the  $(n - 1)$ -measure of  $B_\epsilon(x'_{0r})$ .  $\square$

The set points of  $S$  will be called the Lebesgue points of  $f$ . This set depends on the system of coordinates  $\{(\Gamma_r, a_r)\}_{r=1}^d$ , but this dependence only affects a set of  $\sigma$ -measure equal to zero.

We return to the proof of (3.6). Assume first that (A1) holds. Let us take a numerable dense subset  $\{v_r\}_{j=1}^\infty$  of  $\mathcal{K}$ . Let  $F$  and  $\{F_r\}_{j=1}^\infty$  be measurable subsets of  $\Omega$ , with  $m_{\Sigma_T}(F) = m_{\Sigma_T}(\Sigma_T) = m_{\Sigma_T}(F_r)$  for every  $j$ , such that the Lebesgue point sets of functions  $(x, t) \in \Sigma_T \rightarrow H_{\bar{\alpha}}(x, t, \bar{y}(x, t), \bar{u}(x, t), \bar{\varphi}(x, t))$  and  $(x, t) \in \Omega \rightarrow H_{\bar{\alpha}}(x, t, \bar{y}(x, t), v_j, \bar{\varphi}(x, t))$  are  $F$  and  $F_j$ , respectively. Let us set  $F_0 = F \cap [\cap_{j=1}^\infty F_j]$ . Then we have  $m_{\Sigma_T}(F_0) = m_{\Sigma_T}(\Sigma_T)$ . Now given  $(x_0, t_0) \in F_0$  arbitrary, for every  $\epsilon > 0$  small enough and  $j \geq 1$  we define the admissible controls

$$u_j^\epsilon(x, t) = \begin{cases} \bar{u}(x, t) & \text{if } (x, t) \notin \Sigma_T^\epsilon(x_0, t_0), \\ v_j & \text{otherwise.} \end{cases}$$

Then from (3.5) we deduce

$$\begin{aligned} & \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} H_{\bar{\alpha}}(x, t, \bar{y}(x, t), \bar{u}(x, t), \bar{\varphi}(x, t)) d\sigma(x) dt \\ & \leq \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} H_{\bar{\alpha}}(x, t, \bar{y}(x, t), v_j, \bar{\varphi}(x, t)) d\sigma(x) dt, \quad 1 \leq j. \end{aligned}$$

Passing to the limit where  $\epsilon \rightarrow 0$ , with the help of Lemma 7.2 we get

$$H_{\bar{\alpha}}(x_0, t_0, \bar{y}(x_0, t_0), \bar{u}(x_0, t_0), \bar{\varphi}(x_0, t_0)) \leq H_{\bar{\alpha}}(x_0, t_0, \bar{y}(x_0, t_0), v_j, \bar{\varphi}(x_0, t_0))$$

for every  $(x_0, t_0) \in F_0$  and  $j \geq 1$ . Taking into account that function

$$v \rightarrow H_{\bar{\alpha}}(x_0, t_0, \bar{y}(x_0, t_0), v, \bar{\varphi}(x_0, t_0))$$

is continuous and that  $\{v_j\}_{j=1}^\infty$  is dense in  $\mathcal{K}$ , (3.6) follows from the above inequality.

Now let us suppose that assumption (A2) holds. Let  $F_{\bar{\varphi}}$  be a measurable subset of  $\Sigma_T$  such that for every  $(x_0, t_0) \in F_{\bar{\varphi}}$

$$(7.14) \quad \lim_{\epsilon \rightarrow 0} \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} |\bar{\varphi}(x, t) - \bar{\varphi}(x_0, t_0)| d\sigma(x) dt = 0.$$

Let  $F_0 = F_{\bar{\varphi}} \cap \Sigma_T^0 \cap F$ , where  $F$  is taken as above. Thus we have that  $m_{\Sigma_T}(F_0) = m_{\Sigma_T}(\Sigma_T)$ , and taking spike perturbations as before we deduce

$$\begin{aligned} & \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} H_{\bar{\alpha}}(x, t, \bar{y}(x, t), \bar{u}(x, t), \bar{\varphi}(x, t)) d\sigma(x) dt \\ & \leq \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} H_{\bar{\alpha}}(x, t, \bar{y}(x, t), v, \bar{\varphi}(x, t)) d\sigma(x) dt \end{aligned}$$

for every  $(x_0, t_0) \in F_0$  and  $v \in \mathcal{K}$ . Since  $(x_0, t_0) \in F$ , we can pass to the limit on the left-hand side of the inequality. Let us study the right-hand side:

$$\begin{aligned} & \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} H_{\bar{\alpha}}(x, t, \bar{y}(x, t), v, \bar{\varphi}(x, t)) d\sigma(x) dt \\ &= \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} \bar{\alpha}l(x, t, \bar{y}(x, t), v) d\sigma(x) dt \\ &+ \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} f(x, t, \bar{y}(x, t), v) d\sigma(x) dt \bar{\varphi}(x_0, t_0) \\ &+ \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} [\bar{\varphi}(x, t) - \bar{\varphi}(x_0, t_0)] f(x, t, \bar{y}(x, t), v) d\sigma(x) dt. \end{aligned}$$

The first two terms converge to  $H_{\bar{\alpha}}(x_0, t_0, \bar{y}(x_0, t_0), v, \bar{\varphi}(x_0, t_0))$  because of the continuity of the integrands in  $(x_0, t_0) \in \Sigma_T^0$ . Let us prove that the last term goes to zero.

$$\begin{aligned} & \left| \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} [\bar{\varphi}(x, t) - \bar{\varphi}(x_0, t_0)] f(x, t, \bar{y}(x, t), v) d\sigma(x) dt \right| \\ & \leq C \frac{1}{m_{\Sigma_T}(\Sigma_T^\epsilon(x_0, t_0))} \int_{\Sigma_T^\epsilon(x_0, t_0)} |\bar{\varphi}(x, t) - \bar{\varphi}(x_0, t_0)| d\sigma(x) dt \longrightarrow 0, \end{aligned}$$

thanks to (7.14) and the fact that  $(x, t) \rightarrow f(x, t, \bar{y}(x, t), v)$  is bounded in  $\Sigma_T$  because of the assumption (2.2) and the boundedness of  $\bar{y}$ .  $\square$

Now we will prove Theorem 4.3. The key to achieving this result is to carry out an exact penalization of the state constraint. To do this, we will use the distance function  $d_{Q_\delta}$  associated with the set  $Q_\delta$  and defined in the same way as in the proof of Theorem 4.3.

PROPOSITION 7.3. *If  $(P_\delta)$  is strongly stable and  $\bar{u}$  is a solution of this problem, then there exists  $q_0 > 0$  such that  $\bar{u}$  is also a solution of*

$$(7.15) \quad \inf_{u \in \mathcal{U}} J_q(u) = J(u) + qd_{Q_\delta}(G(y_u))$$

for every  $q \geq q_0$ .

*Proof.* Let us suppose that it is false. Then there exists a sequence  $\{q_k\}_{k=1}^\infty$  of real numbers, with  $q_k \rightarrow +\infty$  and elements  $\{u_k\}_{k=1}^\infty \subset \mathcal{U}$  such that

$$J(u_k) + q_k d_{Q_\delta}(G(y_k)) < J(\bar{u}) \quad \forall k \geq 1,$$

where  $y_k$  is the state corresponding to  $u_k$ . From here we obtain that

$$d_{Q_\delta}(G(y_k)) < \frac{J(\bar{u}) - J(u_k)}{q_k} \longrightarrow 0 \quad \text{when } k \rightarrow +\infty$$

and  $G(y_k) \notin Q_\delta$ . Let  $\delta_k > \delta$  be the smallest number such that  $G(y_k) \in Q_{\delta_k}$ . Since  $\delta_k \rightarrow \delta$ , we can use (4.1) to deduce

$$\begin{aligned} C(\delta_k - \delta) &\geq \inf(P_\delta) - \inf(P_{\delta_k}) \geq J(\bar{u}) - J(u_k) \\ &> q_k d_{Q_\delta}(G(y_k)) = q_k(\delta_k - \delta) \quad \forall k \geq k_\epsilon, \end{aligned}$$

which is not possible.  $\square$

Since  $J_q$  is not Gâteaux differentiable on  $Q_\delta$ , we are going to modify slightly this functional to attain the differentiability necessary for the proof.

PROPOSITION 7.4. *Let us take  $q \geq q_0$  and for every  $\epsilon > 0$  let us consider the problem*

$$(P_{\delta,\epsilon}) \inf_{u \in \mathcal{U}} J_{q,\epsilon}(u) = J(u) + q \{d_{Q_\delta}(G(y_u))^2 + \epsilon^2\}^{1/2}.$$

Then  $\inf(P_{\delta,\epsilon}) \rightarrow \inf(P_\delta)$  when  $\epsilon \rightarrow 0$ .

*Proof.* It is an immediate consequence of the inequality

$$J_q(u) \leq J_{q,\epsilon}(u) \leq J_q(u) + q\epsilon \quad \forall u \in \mathcal{U}. \quad \square$$

Finally we are ready to prove the strong Pontryagin principle.

*Proof of Theorem 4.3.* Propositions 7.3 and 7.4 imply that  $\bar{u}$  is a  $\sigma_\epsilon^2$ -solution of  $(P_{\delta,\epsilon})$ , with  $\sigma_\epsilon \rightarrow 0$  when  $\epsilon \rightarrow 0$ ; i.e.

$$J_{q,\epsilon}(\bar{u}) \leq \inf(P_{\delta,\epsilon}) + \sigma_\epsilon^2.$$

Then we can apply again Ekeland's principle and deduce the existence of an element  $u^\epsilon \in \mathcal{U}$  such that

$$d(u^\epsilon, \bar{u}) \leq \sigma_\epsilon, \quad J_{q,\epsilon}(u^\epsilon) \leq J_{q,\epsilon}(\bar{u}),$$

and

$$J_{q,\epsilon}(u^\epsilon) \leq J_{q,\epsilon}(u) + \sigma_\epsilon d_E(u^\epsilon, u) \quad \forall u \in \mathcal{U}.$$

Now we argue as in the proof of Theorem 3.1 and replace (7.3) by

$$-\sigma_\epsilon m_{\Sigma_T}(\Sigma_T) \leq \lim_{\rho \rightarrow 0} \frac{J_{q,\epsilon}(u_\rho^\epsilon) - J_{q,\epsilon}(u^\epsilon)}{\rho} = z^{0,\epsilon} + \langle \mu^\epsilon, DG(y^\epsilon)z^\epsilon \rangle,$$

where  $\mu^\epsilon \in Z'$  is given by

$$\mu^\epsilon = \begin{cases} \frac{q d_{Q_\delta}(G(y^\epsilon))}{\{d_{Q_\delta}(G(y^\epsilon))^2 + \epsilon^2\}^{1/2}} \nabla d_{Q_\delta}(G(y^\epsilon)) & \text{if } G(y^\epsilon) \notin Q_\delta, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore we have  $\|\mu^\epsilon\|_{Z'} \leq q$  for every  $\epsilon > 0$ . Now we can take a subsequence that converges weakly\* to an element  $\bar{\mu} \in Z'$ . The rest is as in the proof of Theorem 3.1, taking  $\alpha_\epsilon = 1$ .  $\square$

REFERENCES

- [1] V. BARBU AND T. PRECUPANU, *Convexity and Optimization in Banach Spaces*, Editura Academiei, Sijthoff & Noordhoff, Bucharest, 1978.
- [2] E. DI BENEDETTO, *On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 13 (1986), pp. 487–535.
- [3] L. BOCCARDO AND T. GALLOUËT, *Non-linear elliptic and parabolic equations involving measure data*, J. Funct. Anal., 87 (1989), pp. 149–169.
- [4] J. BONNANS, *Pontryagin's principle for the optimal control of semilinear elliptic systems with state constraints*, in 30th IEEE Conference on Control and Decision, Brighton, England, 1991, pp. 1976–1979.

- [5] J. BONNANS AND E. CASAS, *Un principe de Pontryagine pour le contrôle des systèmes elliptiques*, J. Differential Equations, 90 (1991), pp. 288–303.
- [6] J. BONNANS AND E. CASAS, *An extension of Pontryagin's principle for state-constrained optimal control of semilinear elliptic equations and variational inequalities*, SIAM J. Control Optim., 33 (1995), pp. 274–298.
- [7] E. CASAS, *Control of an elliptic problem with pointwise state constraints*, SIAM J. Control Optim., 24 (1986), pp. 1309–1318.
- [8] E. CASAS, *Finite element approximations for some state-constrained optimal control problems*, in Mathematics of the Analysis and Design in Process Control, P. Borne, S. Tzafestas, and N. Radhy, eds., North Holland, Amsterdam, 1992, pp. 293–301.
- [9] E. CASAS, *Introducción a las Ecuaciones en Derivadas Parciales*, University of Cantabria, Santander, 1992.
- [10] E. CASAS, *Boundary control of semilinear elliptic equations with pointwise state constraints*, SIAM J. Control Optim., 31 (1993), pp. 993–1006.
- [11] E. CASAS, *Pontryagin's principle for optimal control problems governed by semilinear elliptic equations*, in International Conference on Control and Estimation of Distributed Parameter Systems: Nonlinear Phenomena, F. Kappel and K. Kunisch, eds., Birkhäuser, Basel, 1994, pp. 97–114.
- [12] E. CASAS, *Boundary control problems of quasilinear elliptic equations: A Pontryagin's principle*, Appl. Math. Optim., 33 (1996), pp. 265–291.
- [13] E. CASAS AND L. FERNÁNDEZ, *A Green's formula for quasilinear parabolic operators*, in Equadiff 91. International Conference on Differential Equations, C. Perelló, C. Simó, and J. Solá-Morales, eds., World Scientific Publishing, Singapore, 1993, pp. 363–367.
- [14] E. CASAS AND J. YONG, *Maximum principle for state-constrained optimal control problems governed by quasilinear elliptic equations*, Differential Integral Equations, 8 (1995), pp. 1–18.
- [15] F. CLARKE, *Optimization and Nonsmooth Analysis*, John Wiley, Toronto, 1983.
- [16] I. EKELAND, *Nonconvex minimization problems*, Bull. Amer. Math. Soc., 1 (1979), pp. 76–91.
- [17] H. FATTORINI, *Optimal control problems for distributed parameter systems governed by semilinear parabolic equations in  $L^1$  and  $L^\infty$  spaces*, in Optimal Control of Partial Differential Equations, Lecture Notes in Control and Information Sciences 149, K. Hoffmann and W. Krabs, eds., Springer-Verlag, Berlin, 1991, pp. 60–80.
- [18] H. FATTORINI, *Optimal control problems for distributed parameter systems in Banach spaces*, Appl. Math. Optim., 4 (1993), pp. 225–257.
- [19] H. FATTORINI AND H. FRANKOWSKA, *Infinite dimensional control problems with state constraints*, in Proceedings of IFIP-IIASA Conference on Modelling and Inverse Problems of Control for Distributed Parameter Systems, Lecture Notes in Control and Information Sciences 154, Springer-Verlag, Berlin, 1991, pp. 52–62.
- [20] H. FATTORINI AND T. MURPHY, *Optimal controls problems for nonlinear parabolic boundary control systems: the Dirichlet boundary condition*, Differential Integral Equations, 6 (1994), pp. 1367–1388.
- [21] H. FATTORINI AND T. MURPHY, *Optimal controls problems for nonlinear parabolic boundary control systems*, SIAM J. Control Optim., 32 (1994), pp. 1577–1596.
- [22] B. HU AND J. YONG, *Pontryagin Maximum Principle for Semilinear and Quasilinear Parabolic Equations with Pointwise State Constraints*, Tech. Report 1141, Institute for Mathematics and Its Applications, University of Minnesota, IMA Preprint Series, June 1993.
- [23] O. LADYZHENSKAYA, V. SOLONNIKOV, AND N. URAL'TSEVA, *Linear and Quasilinear Equations of Parabolic Type*, AMS, Providence, RI, 1968.
- [24] I. LASIECKA, *State constrained control problems for parabolic systems: Regularity of optimal solutions*, Appl. Math. Optim., 6 (1980), pp. 1–29.
- [25] X. LI, *Vector-valued measure and the necessary conditions for the optimal control problems of linear systems*, in Proc. IFAC 3rd Symposium on Control of Distributed Parameter Systems, Toulouse, France, 1982.
- [26] X. LI AND Y. YAO, *Maximum principle of distributed parameter systems with time lags*, in Distributed Parameter Systems, Lecture Notes in Control and Information Sciences 75, Springer-Verlag, New York, 1985, pp. 410–427.
- [27] X. LI AND J. YONG, *Necessary conditions of optimal control for distributed parameter systems*, SIAM J. Control Optim., 29 (1991), pp. 895–908.
- [28] J. LIONS, *Problèmes aux limites non homogènes IV*, Ann. Scuola Norm. Sup. Pisa, 15 (1961), pp. 311–236.
- [29] J. LIONS, *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires*, Dunod, Paris, 1969.

- [30] J. SERRIN, *Pathological solutions of elliptic differential equations*, Ann. Scuola Norm. Sup. Pisa, 18 (1964), pp. 385–387.
- [31] E. STEIN, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, NJ, 1970.
- [32] F. TRÖLTZSCH, *On some parabolic boundary control problems with constraints on the control and functional-constraints on the state*, Z. Anal. Anwendungen, 1 (1982), pp. 1–13.
- [33] J. YONG, *Pontryagin maximum principle for semilinear second order elliptic partial differential equations and variational inequalities with state constraints*, Differential Integral Equations, 5 (1992), pp. 1307–1334.