

11-1-2003

Neumann Boundary Control of Hyperbolic Equations with Pointwise State Constraints

Boris S. Mordukhovich

Wayne State University, boris@math.wayne.edu

Jean-Pierre Raymond

Wayne State University

Recommended Citation

Mordukhovich, Boris S. and Raymond, Jean-Pierre, "Neumann Boundary Control of Hyperbolic Equations with Pointwise State Constraints" (2003). *Mathematics Research Reports*. Paper 16.

http://digitalcommons.wayne.edu/math_reports/16

This Technical Report is brought to you for free and open access by the Mathematics at DigitalCommons@WayneState. It has been accepted for inclusion in Mathematics Research Reports by an authorized administrator of DigitalCommons@WayneState.

**NEUMANN BOUNDARY CONTROL OF HYPERBOLIC
EQUATIONS WITH POINTWISE STATE
CONSTRAINTS**

BORIS S. MORDUKHOVICH and JEAN-PIERRE RAYMOND

**WAYNE STATE
UNIVERSITY**

Detroit, MI 48202

**Department of Mathematics
Research Report**

**2003 Series
#11**

This research was partly supported by the National Science Foundation.

NEUMANN BOUNDARY CONTROL OF HYPERBOLIC EQUATIONS WITH POINTWISE STATE CONSTRAINTS

BORIS S. MORDUKHOVICH¹ and JEAN-PIERRE RAYMOND²

Abstract. We consider optimal control problems for hyperbolic systems with controls in Neumann boundary conditions with pointwise (hard) constraints on control and state functions. Focusing on hyperbolic dynamics governed by the multidimensional wave equation with a nonlinear term, we derive new necessary optimality conditions in the pointwise form of the Pontryagin Maximum Principle for the state-constrained problem under consideration. Our approach is based on modern methods of variational analysis that allows us to obtain refined necessary optimality conditions with no convexity assumptions on integrands in the minimizing cost functional.

Key words. optimal control, wave equation, Neumann boundary controls, state constraints, necessary optimality conditions

AMS subject classifications. 49K20, 93C20, 35L20

1 Introduction

This paper concerns optimal control problems for hyperbolic systems with controls in Neumann boundary conditions in the presence of *pointwise/hard* constraints on control and state functions. It is well known that *state-constrained* control problems are among the most challenging and difficult in dynamic optimization. To the best of our knowledge, they have not been studied yet for *hyperbolic systems* with controls in *Neumann boundary conditions*, which is the objective of this paper.

We pay the main attention to the following optimal control problem governed by the *semilinear wave equation*: minimize

$$J(y, u) = \int_{\Omega} f(x, y(T)) dx + \int_Q g(x, t, y) dx dt + \int_{\Sigma} h(s, t, u) ds dt$$

over admissible pairs (y, u) satisfying

$$(1.1) \quad \begin{cases} y_{tt} - \Delta y + \Phi(x, t, y) = 0 & \text{in } Q := \Omega \times]0, T[, \\ \partial_{\nu} y = u & \text{on } \Sigma := \Gamma \times]0, T[, \\ y(0) = y_0, \quad y_t(0) = y_1 & \text{in } \Omega \end{cases}$$

¹Department of Mathematics, Wayne State University, Detroit, MI 48202 (boris@math.wayne.edu). Research of the first author was partly supported by the National Science Foundation under grants DMS-0072179 and DMS-0304989.

²Laboratoire MIP, Université Paul Sabatier, 31062 Toulouse Cedex 4, France (raymond@mip.ups-tlse.fr).

under the pointwise constraints on control and state functions

$$u \in U_{ad} \subset L^2(\Sigma), \quad y \in \mathcal{C} \subset C([0, T]; L^2(\Omega)).$$

We denote this problem by (P) and shortly write as

$$(P) \quad \inf\{J(y, u) \mid (y, u) \text{ satisfies (1.1), } u \in U_{ad}, y \in \mathcal{C}\}.$$

The hyperbolic system (1.1) is given by a Neumann boundary control problem for the semilinear wave equation with a nonlinear term given by the function Φ . Assumptions on this function as well as on the integrands f , g , and h are presented and discussed in Section 2. Note that $\Omega \subset \mathbb{R}^N$ is an open bounded domain with the boundary Γ of class C^2 , the time $T > 0$ and the initial state $(y_0, y_1) \in H^1(\Omega) \times L^2(\Omega)$ are fixed. Note that the main constructions and results of the paper can be extended to hyperbolic equations governed by more general *strongly elliptic operators* in (1.1), not just the Laplacian Δ , with *time-independent* coefficients under certain regularity assumptions that are not considered in this paper.

As mentioned, we are not familiar with any publications concerning state-constrained Neumann boundary control problems for hyperbolic equations. Some results for *distributed controls* in state-constrained hyperbolic systems are obtained in [4, 6, 23, 24]. Our preceding paper [16] concerns necessary optimality conditions for state-constrained problems governed by the wave equation with *Dirichlet* boundary controls.

It has been well recognized that Neumann and Dirichlet boundary conditions are essentially different in both parabolic and hyperbolic dynamical settings. While for parabolic equations the Dirichlet boundary value problem is considerably more difficult than the Neumann one, it is not the case for the hyperbolic dynamics; see, e.g., [8, 9, 10, 11, 12, 17, 19] and the references therein. Quite opposite, the fundamental *regularity theory* for hyperbolic equations with Dirichlet boundary conditions has come first; cf. [8] and [10]. The sharp regularity results developed by Lasiecka and Triggiani for hyperbolic equations of the Neumann type [10] play a crucial role in this paper.

The main goal of this paper is to establish *necessary optimality conditions* for the state-constrained Neumann boundary control problem (P) , which will be derived under rather mild and natural assumptions. In contrast to the Dirichlet case [16], we concern the *nonlinear dynamics* in (1.1) imposing *stronger* regularity requirements on the initial state: $(y_0, y_1) \in H^1(\Omega) \times L^2(\Omega)$ in (1.1) instead of $(y_0, y_1) \in L^2(\Omega) \times H^{-1}(\Omega)$ in [16]. On the other hand, the Neumann case provides *more regularity* of the corresponding solutions to the boundary value problem in (1.1), which eventually allows us to entirely *avoid the convexity assumptions* on the integrands in the cost functional $J(y, u)$ that play a crucial role in the Dirichlet control problem considered in [16]. Moreover, in this paper we are able to establish necessary optimality condition for (P) in the *pointwise form* of the Pontryagin Maximum Principle in contrast to the weaker integral form of [16].

Our approach to deriving necessary optimality conditions in the Neumann problem (P) is completely different from the one in [16] developed for the Dirichlet case. Instead of reducing the original problem to an abstract optimization problem and then using a suitable version of the Lagrange Multiplier Rule in [16], we now employ *perturbation methods of modern variational analysis* involving *penalizing* state constraints and then passing to the limit from necessary optimality conditions in unconstrained approximating problems. In the case of optimal control problems governed by ordinary differential systems with even nonsmooth data this approach has been developed in the seventies; see, e.g., [2, 15, 22]. For problems governed by partial differential equations the situation is more complicated and the first results have been obtained in the nineties only for bounded controls; see [12] and the references therein. As mentioned in [6, p. 595], versions of the maximum principle for unbounded control operators in the case of problems governed by partial differential equations were discovered by Fattorini [6] and independently by Raymond-Zidani [20]. Based on Ekeland's variational principle [3], and the approach developed in [20], we derive necessary optimality conditions for the original state-constrained hyperbolic problem (P) in the strong/pointwise form of the maximum (actually minimum) principle of Pontryagin's type. Note that the approach developed in this paper allows us to obtain necessary optimality conditions for a more general version of problem (P), where the integrand h depends also on the *state variable* y . We are not going to pursue this issue here for simplicity.

The rest of the paper is organized as follows. In Section 2 we present and discuss the *basic assumptions* used throughout the paper and then formulate the *main result* giving necessary conditions for optimal solutions to (P). Section 3 is devoted to the proper definitions of solutions and the subsequent analysis of the Neumann boundary control problem in the *state system* (1.1) and in the corresponding *adjoint system* appearing in the necessary optimality conditions.

Section 4 contains a preparatory material allowing us to derive in the concluding section necessary optimality condition in the pointwise maximum principle form for approximating problems with no state constraints. Namely, we obtain the so-called *increment formula* for the minimizing functional with respect to *diffuse/needle variations* of controls. This technique, which is well known for the case of ordinary differential equations (see, e.g., [7, 15]), requires a more delicate analysis in the case of partial differential equations. Following the constructions of [20] developed in somewhat different parabolic setting, we obtain an increment formula for approximating hyperbolic problems based on suitable *Taylor expansions* of the problem data with respect to diffuse perturbations of reference controls.

In the final Section 5 we give the proof of the main result of this paper that involves the three major steps of variational analysis: (a) *perturbation* of the original state-constrained problem by a family of approximating problems with *no state constraints* by using Ekeland's variational principle in an appropriate metric space, (b) deriving necessary optimality conditions in the *approximating problems* that provide *suboptimality conditions* to the original problem, and finally (c) *passing to the*

limit from the approximating problems to obtain the desired necessary conditions for the reference optimal solution to the state-constrained problem (P).

2 Basic Assumptions and Statement of the Main Theorem

Throughout the paper we use standard notation. For the reader convenience we recall that $\mathcal{M}([0, T]; L^2(\Omega))$ is the space of measures on $[0, T]$ with values in $L^2(\Omega)$, which is the topological dual of $C([0, T]; L^2(\Omega))$. The topological dual of

$$C_0([0, T]; L^2(\Omega)) := \{y \in C([0, T]; L^2(\Omega)) \mid y(0) = 0\}$$

is denoted by $\mathcal{M}_b(]0, T[; L^2(\Omega))$. It is well known that $\mathcal{M}_b(]0, T[; L^2(\Omega))$ can be identified with the subspace of $\mathcal{M}([0, T]; L^2(\Omega))$ of measures $\mu \in \mathcal{M}([0, T]; L^2(\Omega))$ such that $\mu|_{\Omega \times \{0\}} = 0$, where $\mu|_{\Omega \times \{0\}}$ denotes the restriction of μ to $\Omega \times \{0\}$. The same kind of notation is used in the paper in similar settings. For $z \in L^2(Q)$ we denote by z_t (respectively by z_{tt}) the derivative (respectively the second derivative) of z in t in the sense of distributions on Q .

For a Banach space Z the duality pairing between Z and Z^* is denoted by $\langle \cdot, \cdot \rangle_{Z \times Z^*}$. When there is no ambiguity, we sometimes write $\langle \cdot, \cdot \rangle$ instead of $\langle \cdot, \cdot \rangle_{Z \times Z^*}$. If $y \in C([0, T]; L^2(\Omega))$ and $\mu \in \mathcal{M}_b(]0, T[; L^2(\Omega))$, we still use the notation

$$\langle y, \mu \rangle_{C([0, T]; L^2(\Omega)) \times \mathcal{M}_b(]0, T[; L^2(\Omega))} \quad \text{for} \quad \langle y, \tilde{\mu} \rangle_{C([0, T]; L^2(\Omega)) \times \mathcal{M}([0, T]; L^2(\Omega))},$$

where $\tilde{\mu}$ is the extension of μ by zero to $\Omega \times \{0\}$. The same is used for similar spaces of measures. To emphasize a specific kind of regularity of solutions to the hyperbolic equations under considerations, we may write, e.g., that $(y, y_t) \in C([0, T]; X) \times C([0, T]; Y)$ is a solution to (1.1) instead of just indicating that y is a solution to this system.

For the definition of the space $BV([0, T]; (H^1(\Omega))^*)$ —the space of functions of bounded variation on $[0, T]$ with values in $(H^1(\Omega))^*$ —we refer to [1, 14]. If $p \in BV([0, T]; (H^1(\Omega))^*)$, one can define $p(t^-)$ and $p(t^+)$ for every $t \in]0, T[$ and also $p(0^+)$ and $p(T^-)$, while the values $p(0)$ and $p(T)$ may be generally different from $p(0^+)$ and $p(T^-)$. There is a unique Radon measure on $[0, T]$ with values in $(H^1(\Omega))^*$, denoted by $d_t p$, such that the restriction of $d_t p$ to $]0, T[$ is the vector-valued distributional derivative of p in $]0, T[$ with $d_t p(\{0\}) = (p(0^+) - p(0))$ and $d_t p(\{T\}) = (p(T) - p(T^-))$. Moreover, identifying p with its representative right-hand side continuous in $]0, T[$, one has

$$p(0^+) = p(0) + d_t p(\{0\}) \quad \text{and} \quad p(t) = p(0) + d_t p([0, t]) \quad \text{for every } t \in]0, T[.$$

Recall that if (p_n) is a bounded sequence in $BV([0, T]; (H^1(\Omega))^*)$, then there is a subsequence (p_{n_k}) and a function $p \in BV([0, T]; (H^1(\Omega))^*)$ such that

$$p_{n_k}(t) \rightarrow p(t) \quad \text{weakly in } (H^1(\Omega))^* \quad \text{for almost every } t \in [0, T].$$

Note that this convergence may hold for every $t \in [0, T]$ if the above representative right-hand side continuous in $]0, T[$ is not specified; see [1, Theorem 3.5] and [14, Proposition 16.1]. In particular,

$$p_{n_k}(T) \rightarrow p(T) \quad \text{weakly in } (H^1(\Omega))^*.$$

Now let us formulate the standing **Basic Assumptions** on the initial data of problem (P) that are needed throughout this paper.

(A1) For every $y \in \mathbb{R}$, $\Phi(\cdot, \cdot, y)$ is measurable in Q . For almost every (a.e.) pair $(x, t) \in Q$, $\Phi(x, t, \cdot)$ is of class C^1 on \mathbb{R} . Moreover, one has

$$(2.1) \quad \Phi(\cdot, 0) \in L^1(0, T; L^2(\Omega)), \quad |\Phi'_y(x, t, y)| \leq M \quad \text{in } Q \times \mathbb{R},$$

where M is a positive constant.

(A2) For every $y \in \mathbb{R}$, $f(\cdot, y)$ is measurable in Ω with $f(\cdot, 0)$ belonging to $L^1(\Omega)$. For a.e. $x \in \Omega$, $f(x, \cdot)$ is a function of class C^1 on \mathbb{R} . Moreover, there is a constant $C > 0$ such that

$$|f'_y(x, y)| \leq C(1 + |y|) \quad \text{whenever } (x, y) \in \Omega \times \mathbb{R}.$$

(A3) For every $y \in \mathbb{R}$, $g(\cdot, \cdot, y)$ is measurable in Q with $g(\cdot, 0)$ belonging to $L^1(Q)$. For a.e. $(x, t) \in Q$, $g(x, t, \cdot)$ is of class C^1 . Moreover, there is a constant $C > 0$ such that

$$|g'_y(x, t, y)| \leq C(1 + |y|) \quad \text{whenever } (x, t, y) \in Q \times \mathbb{R}.$$

(A4) For every $u \in \mathbb{R}$, $h(\cdot, \cdot, u)$ is measurable on Σ with $h(\cdot, 0)$ belonging to $L^1(\Sigma)$. For a.e. $(s, t) \in \Sigma$, $h(s, t, \cdot)$ is of class C^1 . Moreover, there is a constant $C > 0$ such that

$$|h'_u(s, t, u)| \leq C(1 + |u|) \quad \text{whenever } (s, t, u) \in \Sigma \times \mathbb{R}.$$

(A5) The state constraint set $\mathcal{C} \subset C([0, T]; L^2(\Omega))$ is a closed and convex with $\text{int} \mathcal{C} \neq \emptyset$. We suppose that the initial state function $(x, t) \mapsto y_0(x)$ belongs to the interior of \mathcal{C} .

(A6) The control set U_{ad} is given in the form

$$U_{ad} := \{u \in L^2(\Sigma) \mid u(s, t) \in K(s, t) \text{ a.e. on } \Sigma\},$$

where K is a measurable multifunction whose values are nonempty and closed subsets of \mathbb{R} .

Of course, we suppose as usual that the set of *feasible pairs* (y, u) to (P) is *nonempty*, i.e., there is $u \in U_{ad}$ such that $J(y_u, u) < \infty$ and $y_u \in \mathcal{C}$ for the corresponding solution to system (1.1) rigorously defined and discussed in Section 3.

Observe that the above basic assumptions do *not* impose any *convexity* requirements on the integrands in the cost functional with respect to either state or control variables, as well as on the

control set U_{ad} . This is different from the setting of [16] for the corresponding Dirichlet problem. The reason for such an essential relaxation is that the Neumann boundary value problem offers *more regularity* in comparison with the Dirichlet one and allows us to employ powerful variational methods to prove necessary optimality conditions that do *not rely on weak convergences*. The latter definitely requires convexity for the limiting procedures to end up with pointwise results. On the other hand, in this paper we do not establish any existence theorems for optimal solutions, in contrast to [16]. In fact, in the Neumann setting under consideration it would be enough to assume convexity only with *respect to control* variables to justify the existence of optimal solutions by direct methods. The stronger convexity assumptions imposed in [16] with respect to *both state and control* variables are due to the lack of regularity in the Dirichlet setting and are needed not only for the existence of optimal solutions but also for the proof of necessary optimality conditions as given in [16].

To formulate our main result, let us define the *Hamiltonian* function

$$\mathcal{H}(s, t, u, p, \lambda) := pu + \lambda h(s, t, u)$$

for the control problem (P). The following theorem gives necessary conditions for optimal solutions to (P) that are counterparts of the *Pontryagin Maximum Principle* in the pointwise form for the Neumann boundary control problem under consideration. Note that it is more convenient in our case to formulate this result with the *minimum* (not maximum) condition.

Theorem 2.1 (pointwise necessary optimality conditions). *Let (\bar{y}, \bar{u}) be a optimal solution to problem (P) satisfying assumptions (A1)–(A6). Then there $\lambda \in \mathbb{R}^+$, $\mu \in \mathcal{M}_b([0, T]; L^2(\Omega))$, and a measurable subset $\tilde{\Sigma} \subset \Sigma$ of full Lebesgue measure such that*

$$(2.2) \quad (\lambda, \mu) \neq 0, \quad \langle \mu, z - \bar{y} \rangle \leq 0 \quad \text{for all } z \in \mathcal{C}, \quad \text{and}$$

$$(2.3) \quad \mathcal{H}(s, t, \bar{u}(s, t), p(s, t), \lambda) = \min_{u \in K(s, t)} \mathcal{H}(s, t, u, p(s, t), \lambda) \quad \text{for all } (s, t) \in \tilde{\Sigma},$$

where p is the corresponding solution to the adjoint system

$$(2.4) \quad \begin{cases} p_{tt} - \Delta p + \Phi'_y(\cdot, \bar{y})p = \lambda g'_y(x, t, \bar{y}) + \mu|_Q & \text{in } Q, \\ \partial_\nu p = 0 & \text{on } \Sigma = \Gamma \times]0, T[, \\ p(T) = y_0, \quad p_t(T) = -\lambda f'_y(x, \bar{y}(T)) - \mu|_{\Omega \times \{T\}} & \text{in } \Omega. \end{cases}$$

The proof of Theorem 2.1 is conducted in Section 5. The definitions of solutions to the state and adjoint systems in this theorem are given and discussed in the next section.

3 Analysis of the State and Adjoint Systems

Let us start with the Neumann boundary value problem for the *linear wave equation*

$$(3.1) \quad \begin{cases} y_{tt} - \Delta y = \phi & \text{in } Q, \\ \partial_\nu y = u & \text{on } \Sigma, \\ y(0) = y_0, \quad y_t(0) = y_1 & \text{in } \Omega. \end{cases}$$

The following fundamental *regularity* result are established by Lasiecka and Triggiani [10].

Lemma 3.1 (basic regularity). *Let $y(\phi, u, y_0, y_1)$ be a solution to the linear Neumann boundary value problem (3.1). Then the mapping $u \mapsto y(0, u, 0, 0)$ is bounded from $L^2(\Sigma)$ to $C([0, T]; H^{1/2}(\Omega)) \cap C^1([0, T]; H^{-1/2}(\Omega))$, and it is also bounded from $L^2(\Sigma)$ to $H^{3/5-\varepsilon}(Q)$ for all $\varepsilon > 0$. Furthermore, the mapping $(\phi, 0, y_0, y_1) \mapsto y(\phi, 0, y_0, y_1)$ is bounded from $L^1(0, T; L^2(\Omega)) \times H^1(\Omega) \times L^2(\Omega)$ to $C([0, T]; H^1(\Omega)) \cap C([0, T]; L^2(\Omega))$.*

Next we consider the Neumann boundary value problem for the linear wave equation with an *irregular coefficient*:

$$(3.2) \quad \begin{cases} y_{tt} - \Delta y + ay = \phi & \text{in } Q, \\ \partial_\nu y = u & \text{on } \Sigma, \\ y(0) = y_0, \quad y_t(0) = y_1 & \text{in } \Omega, \end{cases}$$

where the irregular coefficient $a(x, t)$ belongs to $L^\infty(Q)$. The following upper estimate of solutions to the homogeneous problem in (3.2) is needed in the sequel.

Lemma 3.2 (solution estimate for the homogeneous Neumann problem). *Let y be a solution to (3.2) with $u = 0$. Then one has the estimate*

$$\|y\|_{C([0, T]; H^1(\Omega))} + \|y_t\|_{C([0, T]; L^2(\Omega))} \leq C(\|\phi\|_{L^1(0, T; L^2(\Omega))} + \|y_0\|_{H^1(\Omega)} + \|y_1\|_{L^2(\Omega)}),$$

where the constant C may depend on $\|a\|_{L^\infty(Q)}$ and $\|\phi\|_{L^1(0, T; L^2(\Omega))}$, but it is invariant with respect to all $a(x, t)$ having the same $L^\infty(Q)$ -norm.

Proof. The proof is classical. It is sufficient to multiply the first equation in (3.2) by y_t , to integrate it over Ω , and then to use Gronwall's Lemma (see for example [13, page 184]). \triangle

Lemma 3.3 (compactness of the solution operator). *Let $y(u)$ be the unique solution to (3.2) corresponding to $(\phi, y_0, y_1) = (0, 0, 0)$ and u . Then the mapping $u \mapsto y(u)$ is a compact operator from $L^2(\Sigma)$ into $C([0, T]; L^2(\Omega))$.*

Proof. The existence and uniqueness of the corresponding solution to (3.2) can be deduced from the well-known result for (3.1) by using a fixed point method in $L^1(0, \bar{t}; L^2(\Omega))$ as \bar{t} is sufficiently small and then by iterating the process n times with $n\bar{t} > T$; cf. [18] for more details. Moreover, in this way we get the estimate

$$\|y\|_{C([0,T]; H^{1/2}(\Omega))} + \|y_t\|_{C([0,T]; H^{-1/2}(\Omega))} \leq C\|u\|_{L^2(\Sigma)},$$

where C depends on an upper bound for $\|a\|_{L^\infty(Q)}$ being independent of a . Now the compactness result follows from [21, Corollary 5]. \triangle

Our next goal is to study the Neumann boundary value problem (1.1), which is labelled as the *state system* for convenience. We first recall the notion of *weak solutions* to the Neumann problem in (1.1) that is appropriate for the purposes of this paper.

Definition 3.4 (weak solutions to the state system). A function $(y, y_t) \in C([0, T]; L^2(\Omega)) \times C([0, T]; (H^1(\Omega))^*)$ is a WEAK SOLUTION to system (1.1) if

$$(3.3) \quad \begin{aligned} \int_Q -\Phi(x, t, y)z \, dxdt &= \int_Q y\varphi \, dxdt + \langle y_t(T), z^0 \rangle_{(H^1(\Omega))^* \times H^1(\Omega)} \\ &- \int_\Omega y_t(0)z(0) \, dx - \int_\Omega y(T)z^1 \, dx + \int_\Omega y(0)z_t(0) \, dx + \int_\Sigma zu \, dsdt \end{aligned}$$

for all $(\varphi, z^0, z^1) \in L^1(0, T; L^2(\Omega)) \times H^1(\Omega) \times L^2(\Omega)$, where z solves the homogeneous Neumann boundary value problem

$$(3.4) \quad \begin{cases} z_{tt} - \Delta z = \varphi & \text{in } Q, \\ \partial_\nu z = 0 & \text{on } \Sigma, \\ z(T) = z^0, \quad z_t(T) = z^1 & \text{in } \Omega. \end{cases}$$

The advantage of the above definition is that it allows to establish the existence, uniqueness, and regularity of weak solutions to the original state system under the standing assumptions made in Section 2.

Theorem 3.5 (existence, uniqueness, and regularity of weak solutions to the state system). For every $(u, y_0, y_1) \in L^2(\Sigma) \times H^1(\Omega) \times L^2(\Omega)$ the state system (1.1) admits a unique weak solution (y, y_t) in $C([0, T]; H^{1/2}(\Omega)) \times C([0, T]; H^{-1/2}(\Omega))$ satisfying the estimate

$$\|y\|_{C([0,T]; H^{1/2}(\Omega))} + \|y_t\|_{C([0,T]; H^{-1/2}(\Omega))} \leq C(\|u\|_{L^2(\Sigma)} + \|y_0\|_{H^1(\Omega)} + \|y_1\|_{L^2(\Omega)} + 1)$$

with some constant $C > 0$. Moreover, the mapping $(u, y_0, y_1) \mapsto y$ is continuous from $(u, y_0, y_1) \in L^2(\Sigma) \times H^1(\Omega) \times L^2(\Omega)$ into $C([0, T]; H^{1/2}(\Omega)) \cap C^1([0, T]; H^{-1/2}(\Omega))$.

Proof. The existence of solutions in the space $C([0, \bar{t}]; L^2(\Omega)) \cap C^1([0, \bar{t}]; (H^1(\Omega))^*)$ with \bar{t} sufficiently small can be obtained by a standard fixed point method. Then assumption (A1) and the estimates in Lemmas 3.2 and 3.3 allow to ensure the existence of solutions in the space given in the theorem. The proof of uniqueness is also standard and is omitted for brevity. The estimate of (y, y_t) in $C([0, T]; H^{1/2}(\Omega)) \cap C^1([0, T]; H^{-1/2}(\Omega))$ follows from the estimate of y in $C([0, T]; L^2(\Omega))$ due to the basic Lemma 3.1. To justify the continuity of the mapping $(u, y_0, y_1) \mapsto y$ from $(u, y_0, y_1) \in L^2(\Sigma) \times H^1(\Omega) \times L^2(\Omega)$ into $C([0, T]; H^{1/2}(\Omega)) \cap C^1([0, T]; H^{-1/2}(\Omega))$, we use again assumption (A1) and the corresponding estimates for the linearized system (3.2) presented in Lemmas 3.2 and 3.3. \triangle

Next we consider the *adjoint system* given by

$$(3.5) \quad \begin{cases} p_{tt} - \Delta p + ap = \mu|_Q & \text{in } Q, \\ \partial_\nu p = 0 & \text{on } \Sigma, \\ p(T) = 0, \quad p_t(T) = -\mu|_{\Omega \times \{T\}} & \text{in } \Omega, \end{cases}$$

where $\mu \in \mathcal{M}_b([0, T]; L^2(\Omega))$, where $\mu|_Q$ and $\mu|_{\Omega \times \{T\}}$ denote the restrictions of μ to Q and to $\Omega \times \{T\}$, respectively, and where $a \in L^\infty(Q)$.

Take $(p, p_t) \in L^2(0, T; H^1(\Omega)) \times L^2(0, T; L^2(\Omega))$ and assume that the combination $p_{tt} - \Delta p$, calculated in the sense of distributions on Q , belongs to $\mathcal{M}_b(Q)$. Following the construction in [19, Lemma 4.3] and using the divergence theorem, we define the *normal trace* on ∂Q of the vectorfield $(-\nabla p, p_t)$ as an element of $H^{-1/2}(\partial Q)$. Denoting this normal trace by $\gamma_{\nu Q}(-\nabla p, p_t)$, one has

$$\|\gamma_{\nu Q}(-\nabla p, p_t)\|_{H^{-1/2}(\partial Q)} \leq C(\|p\|_{L^2(0, T; H^1(\Omega))} + \|p_t\|_{L^2(Q)} + \|p_{tt} - \Delta p\|_{\mathcal{M}_b(Q)}),$$

where the constant $C > 0$ is independent of p . In this way we get

$$\gamma_{\nu Q}(-\nabla p, p_t)|_{\Omega \times \{0\}} = p_t(0) \in H^{-1/2}(\Omega),$$

which allows us to justify the following definition of *weak solutions* to the adjoint hyperbolic system with the Neumann boundary conditions.

Definition 3.6 (weak solutions to the adjoint system). A function $(p, p_t) \in L^\infty(0, T; H^1(\Omega)) \times L^\infty(0, T; L^2(\Omega))$ is a WEAK SOLUTION to (3.5) if

$$(3.6) \quad \begin{cases} - \int_{\Omega} p(0) y_1 \, dx + \langle p_t(0), y_0 \rangle_{(H^1(\Omega))^* \times H^1(\Omega)} \\ + \langle y(\varphi, y_0, y_1), \mu \rangle_{C([0, T]; L^2(\Omega)) \times \mathcal{M}_b([0, T]; L^2(\Omega))} - \int_Q p \phi \, dx dt = 0 \end{cases}$$

for all $(\varphi, y_0, y_1) \in L^2(Q) \times H^1(\Omega) \times L^2(\Omega)$, where $y(\varphi, y_0, y_1)$ is the solution to

$$(3.7) \quad \begin{cases} y_{tt} - \Delta y + ay = \varphi & \text{in } Q, \\ \partial_\nu y = 0 & \text{on } \Sigma, \\ y(0) = y_0, \quad y_t(0) = y_1 & \text{in } \Omega. \end{cases}$$

The next theorem establishes the existence, uniqueness, and regularity of weak solutions to the adjoint system under the standing assumptions made.

Theorem 3.7 (existence, uniqueness and regularity of weak solutions to the adjoint system). *The adjoint system (3.5) admits a unique weak solution $(p, p_t) \in L^\infty(0, T; H^1(\Omega)) \times L^\infty(0, T; L^2(\Omega))$ such that $p_t \in BV([0, T]; (H^1(\Omega))^*)$, $p \in C_w([0, T]; H^1(\Omega))$, and*

$$p_t(\tau) \in L^2(\Omega) \quad \text{whenever } \tau \in \{t \in [0, T] \mid \mu(\Omega \times \{t\}) = 0\},$$

which implies that $p_t(0) \in L^2(\Omega)$ ($C_w([0, T]; H^1(\Omega))$ denotes the space of continuous functions from $[0, T]$ into $H^1(\Omega)$ endowed with its weak topology). Moreover, one has the estimate

$$(3.8) \quad \|p\|_{L^\infty(0, T; H^1(\Omega))} + \|p_t\|_{L^\infty(0, T; L^2(\Omega))} \leq C \|\mu\|_{\mathcal{M}_b([0, T]; L^2(\Omega))},$$

where C depends on $\|a\|_{L^\infty(Q)}$ but is invariant with respect to the functions $a(x, t)$ having the same norm in the space $L^\infty(Q)$.

Proof. Observe that $p = 0$ when the pair $(p, p_t) \in L^\infty(0, T; H^1(\Omega)) \times L^\infty(0, T; L^2(\Omega))$ satisfies (3.6) with $\mu = 0$. This implies that the adjoint system (3.5) cannot admit more than one weak solution. To prove the existence of a weak solution, we pick a sequence $(\mu_n) \subset L^1(0, T; L^2(\Omega))$ satisfying

$$\lim_{n \rightarrow \infty} \int_Q y \mu_n \, dx dt = \langle y, \mu \rangle_{C([0, T]; L^2(\Omega)) \times \mathcal{M}_b([0, T]; L^2(\Omega))} \quad \text{for all } y \in C([0, T]; L^2(\Omega)),$$

$$\int_Q \mu_n \, dx dt = \|\mu\|_{\mathcal{M}_b([0, T]; L^2(\Omega))}.$$

Taking the unique solution p_n to

$$(3.9) \quad \begin{cases} p_{tt} - \Delta p + ap = \mu_n & \text{in } Q, \\ \partial_\nu p = 0 & \text{on } \Sigma, \\ p(T) = 0, \quad p_t(T) = -\mu|_{\Omega \times \{T\}} & \text{in } \Omega \end{cases}$$

and using [8, Theorem 2.1], we have the estimate

$$(3.10) \quad \|p_n\|_{L^\infty(0, T; H^1(\Omega))} + \|p_{nt}\|_{L^\infty(0, T; L^2(\Omega))} + \|p_n(0)\|_{H^1(\Omega)} + \|p_{nt}(0)\|_{L^2(\Omega)} \leq C \|\mu\|_{\mathcal{M}_b([0, T]; L^2(\Omega))}$$

with a constant $C > 0$ independent of n , where p_{nt} stands for the derivative of p_n with respect to t in $]0, T[$ in the sense of vector valued distributions. Denoting by p_{ntt} the derivative of p_{nt} with respect to t in $]0, T[$ in the sense of vector valued distributions, we get from (3.9) the representation

$$p_{ntt} = \pi_n + \mu_n \in L^\infty(0, T; (H^1(\Omega))^*) + \mathcal{M}_b(]0, T[; L^2(\Omega)) \subset \mathcal{M}_b(]0, T[; (H^1(\Omega))^*),$$

where the operator π_n is defined by

$$\langle \pi_n, y \rangle_{L^\infty(0, T; (H^1(\Omega))^*) \times L^1(0, T; H^1(\Omega))} = \int_Q \nabla p_n \nabla y \, dx dt.$$

Therefore, in addition to estimates (3.10), the sequences (p_{ntt}) and (p_{nt}) are bounded in the spaces $\mathcal{M}_b(]0, T[; (H^1(\Omega))^*)$ and $BV([0, T]; (H^1(\Omega))^*)$, respectively. Observing that $\mathcal{M}_b(]0, T[; (H^1(\Omega))^*)$ is the dual of a *separable* Banach space, we select weak* convergent *subsequences* of the above sequences. The same *sequential compactness* property holds also for the space $BV([0, T]; (H^1(\Omega))^*)$; see Section 2. In this way we find $p \in L^\infty(0, T; H^1(\Omega))$ with $p_t \in L^\infty(0, T; L^2(\Omega)) \cap BV([0, T]; (H^1(\Omega))^*)$ and a subsequence of (p_n) converging to p in the weak* topology of $L^\infty(0, T; H^1(\Omega))$ and such that the corresponding subsequence of (p_{nt}) converges weak* in $L^\infty(0, T; L^2(\Omega))$ to p_t . Furthermore, we can also suppose, since $\gamma_{\nu_Q}(-\nabla p_n, p_{nt})$ is bounded in $L^2(\partial Q)$, that $\gamma_{\nu_Q}(-\nabla p_n, p_{nt})$ converges to $\gamma_{\nu_Q}(-\nabla p, p_t)$ in the weak topology of $L^2(\partial Q)$. Taking into account the relations

$$\gamma_{\nu_Q}(-\nabla p_n, p_{nt})|_{\Omega \times \{T\}} = \mu|_{\Omega \times \{T\}} \quad \text{and} \quad \gamma_{\nu_Q}(-\nabla p_n, p_{nt})|_{\Sigma} = 0,$$

one gets that $\gamma_{\nu_Q}(-\nabla p, p_t)|_{\Sigma} = -\partial_\nu p = 0$ and that

$$\gamma_{\nu_Q}(-\nabla p_n, p_{nt})|_{\Omega \times \{0\}} = p_{nt}(0) \rightarrow \gamma_{\nu_Q}(-\nabla p, p_t)|_{\Omega \times \{0\}} = p_t(0)$$

in the weak topology of $L^2(\Omega)$. Finally, by passing to the limit in the equality

$$\begin{aligned} & -\langle p_n(0), y_1 \rangle_{L^2(\Omega)} + \langle p_{nt}(0), y_0 \rangle_{(H^1(\Omega))^* \times H^1(\Omega)} \\ & + \langle y(f, y_0, y_1), \mu \rangle_{C([0, T]; L^2(\Omega)) \times \mathcal{M}_b(]0, T[; L^2(\Omega))} - \langle p_n, f \rangle_{L^2(Q)} = 0, \end{aligned}$$

we conclude that (p, p_t) is the desired weak solution to the adjoint system (3.5) and complete the proof of the theorem. \triangle

The last result of this section gives a useful Green-type relationship between the corresponding solutions of the state and adjoint systems.

Theorem 3.8 (Green formula). *Let y and p satisfy the state (3.2) and adjoint (3.5) systems, respectively. Then*

$$\begin{aligned} (3.11) \quad & \langle y, \mu \rangle_{C([0, T]; L^2(\Omega)) \times \mathcal{M}_b(]0, T[; L^2(\Omega))} - \langle p, y_{tt} - \Delta y \rangle_{L^\infty(0, T; H^1(\Omega)) \times L^1(0, T; (H^1(\Omega))^*)} \\ & = - \int_\Omega y(0) p_t(0) \, dx + \int_\Omega y_t(0) p(0) \, dx + \int_\Sigma y p \, ds dt. \end{aligned}$$

Proof. This is proved in Theorem 3.7 for the sequence (p_n) of solutions to the approximating adjoint systems (3.9). Passing there to the limit as $n \rightarrow \infty$, we obtain the desired Green formula (3.11) as formulated in the theorem. \triangle

4 Diffuse Perturbations and Increment Formula

As mentioned in Section 1, our approach to deriving necessary optimality conditions in the original state-constrained problem (P) includes an approximation procedure to penalize the state constraints. In this way we arrive at a family of Neumann boundary control problems for hyperbolic equations with hard/pointwise constraints on controls but with *no state constraints*. Although the latter approximating problems are essentially easier than the initial state-constrained problem (P) , they still require a delicate variational analysis. As well known in the control theory for ordinary differential equations, a key element in obtaining maximum-type conditions for problems with hard constraints on control but not on state variables is the so-called *increment formula* for the minimizing cost functional with respect to *needle variations* of reference controls; see, e.g., [7, 15]. In this section we obtain some counterparts of such results for the hyperbolic control problems under consideration, where analogues of needle variations are known as “diffuse perturbations” introduced in [20].

Given a reference control $\bar{u} \in U_{ad}$, an admissible control $u \in U_{ad}$, and a number $\rho \in]0, 1[$, a *diffuse perturbation* of \bar{u} is defined by

$$(4.1) \quad u_\rho(s, t) := \begin{cases} \bar{u}(s, t) & \text{on } \Sigma \setminus E_\rho, \\ u(s, t) & \text{on } E_\rho, \end{cases}$$

where E_ρ is a measurable subset of Σ . The next theorem can be viewed as an increment formula for the cost functional $J(y, u)$ with respect to diffuse perturbations of the reference control. Note that it also contains the corresponding Taylor expansion for state trajectory of (1.1), which is an essential ingredient of the increment formula. By \mathcal{L}^N we denote the *Lebesgue measure* on \mathbb{R}^N .

Theorem 4.1 (increment formula). *Given arbitrary controls $\bar{u}, u \in U_{ad}$ and a number $\rho \in]0, 1[$, we consider the diffuse perturbation defined in (4.1) and the weak solutions \bar{y} and y_ρ of system (1.1) corresponding to \bar{u} and u_ρ , respectively. Then there exists a measurable subset $E_\rho \subset \Sigma$ such that the following hold:*

$$(4.2) \quad \mathcal{L}^N(E_\rho) = \rho \mathcal{L}^N(\Sigma),$$

$$(4.3) \quad \int_{E_\rho} (h(s, t, \bar{u}) - h(s, t, u)) \, ds dt = \rho \int_{\Sigma} (h(s, t, \bar{u}) - h(s, t, u)) \, ds dt,$$

$$(4.4) \quad y_\rho = \bar{y} + \rho z + \rho r_\rho \quad \text{with} \quad \lim_{\rho \rightarrow 0} \|r_\rho\|_{C([0, T]; L^2(\Omega))} = 0, \quad \text{and}$$

$$(4.5) \quad J(y_\rho, u_\rho) = J(\bar{y}, \bar{u}) + \rho \Delta J + o(\rho) \quad \text{with} \quad \Delta J := J'_y(\bar{y}, \bar{u})z + J(\bar{y}, u) - J(\bar{y}, \bar{u}),$$

where z is the weak solution to the system

$$(4.6) \quad \begin{cases} z_{tt} - \Delta z + \Phi'_y(\cdot, \bar{y})z = 0 & \text{in } Q, \\ \partial_\nu z = \bar{u} - u & \text{on } \Sigma, \\ z(0) = 0, \quad z_t(0) = 0 & \text{in } \Omega. \end{cases}$$

The proof of the theorem given below relies on the following technical lemma, which follows from [20, Lemma 4.1].

Lemma 4.2 (domains of diffuse perturbations). *Let $\bar{u}, u \in U_{ad}$. For every $\rho \in]0, 1[$ there is a sequence of measurable subsets E_ρ^n in Σ such that:*

$$(4.7) \quad \mathcal{L}^N(E_\rho^n) = \rho \mathcal{L}^N(\Sigma),$$

$$(4.8) \quad \int_{E_\rho^n} (h(s, t, \bar{u}) - h(s, t, u)) ds dt = \rho \int_\Sigma (h(s, t, \bar{u}) - h(s, t, u)) ds dt,$$

$$(4.9) \quad \frac{1}{\rho} \chi_{E_\rho^n} \rightarrow 0 \quad \text{weak}^* \text{ in } L^\infty(\Sigma) \quad \text{as } n \rightarrow \infty,$$

where χ_Ω stands for the characteristic function of the set Ω .

Proof of Theorem 4.1. The existence of the domain sets E_ρ satisfying (4.2) and (4.3) is an easy consequence of Lemma 4.2. The main issue is to justify the Taylor expansion (4.4) for the trajectories y_ρ of (1.1) corresponding to the diffuse control perturbations. One clearly sees that (4.4) and (4.3) imply the increment formula (4.5) due to the constructions of diffuse perturbations.

To prove (4.4), we pick a number $\rho \in]0, 1[$, take the sets E_ρ^n from Lemma 4.2, and build the diffuse control perturbations

$$u_\rho^n(s, t) := \begin{cases} \bar{u}(s, t) & \text{on } \Sigma \setminus E_\rho^n, \\ u(s, t) & \text{on } E_\rho^n. \end{cases}$$

Let y_ρ^n be the solution of (1.1) corresponding to u_ρ^n , and let z be the (unique) weak solution of (4.6). It is easy to see that for all n the function $\xi_\rho^n = (y_\rho^n - \bar{y})/\rho - z$ is the unique weak solution to the system

$$\begin{cases} \xi_{tt} - \Delta \xi + a_\rho^n \xi = f_\rho^n & \text{in } Q, \\ \partial_\nu \xi = w_\rho^n & \text{on } \Sigma, \\ \xi(0) = 0, \quad \xi_t(0) = 0 & \text{in } \Omega \end{cases}$$

with the following data:

$$a_\rho^n := \int_0^1 \Phi'_y(\cdot, (\bar{y} + \theta(y_\rho^n - \bar{y}))) d\theta, \quad f_\rho^n := (\Phi'_y(\cdot, \bar{y}) - a_\rho^n)z, \quad w_\rho^n := (1 - \frac{1}{\rho} \chi_{E_\rho^n}) \cdot (u - \bar{u}),$$

Denote by $\xi_\rho^{n,1}$ the solution to

$$\begin{cases} \xi_{tt} - \Delta \xi + a_\rho^n \xi = f_\rho^n & \text{in } Q, \\ \partial_\nu \xi = 0 & \text{on } \Sigma, \\ \xi(0) = 0, \quad \xi_t(0) = 0 & \text{in } \Omega, \end{cases}$$

by $\xi_\rho^{n,2}$ the solution to

$$\begin{cases} \xi_{tt} - \Delta \xi + a_\rho^n \xi = 0 & \text{in } Q, \\ \partial_\nu \xi = w_\rho^n & \text{on } \Sigma, \\ \xi(0) = 0, \quad \xi_t(0) = 0 & \text{in } \Omega, \end{cases}$$

and by ζ_ρ^n the solution to

$$\begin{cases} \zeta_{tt} - \Delta \zeta + a \zeta = 0 & \text{in } Q, \\ \partial_\nu \zeta = w_\rho^n & \text{on } \Sigma, \\ \zeta(0) = 0, \quad \zeta_t(0) = 0 & \text{in } \Omega, \end{cases}$$

where $a(x, t) := \Phi'_y(x, t, \bar{y}(x, t))$. One clearly has

$$\begin{aligned} (\xi_\rho^{n,2} - \zeta_\rho^n)_{tt} - \Delta(\xi_\rho^{n,2} - \zeta_\rho^n) + a_\rho^n(\xi_\rho^{n,2} - \zeta_\rho^n) &= (a - a_\rho^n)\zeta_\rho^n & \text{in } Q, \\ \partial_\nu(\xi_\rho^{n,2} - \zeta_\rho^n) &= 0 & \text{on } \Sigma, \\ (\xi_\rho^{n,2} - \zeta_\rho^n)(0) = 0, \quad (\xi_\rho^{n,2} - \zeta_\rho^n)_t(0) &= 0 & \text{in } \Omega. \end{aligned}$$

Employing Lemmas 3.1 and 3.3, we find a constant $C > 0$, independent of n and ρ , ensuring the following estimates for all $n = 1, 2, \dots$ and $0 < \rho < 1$:

$$(4.10) \quad \begin{aligned} \|\xi_\rho^{n,2} - \zeta_\rho^n\|_{C([0,T];L^2(\Omega))} &\leq C \|a - a_\rho^n\|_{L^1(0,T;L^{2N}(\Omega))} \|\zeta_\rho^n\|_{L^\infty(0,T;L^{2N/(N-1)}(\Omega))}, \\ &\leq C \|a - a_\rho^n\|_{L^1(0,T;L^{2N}(\Omega))} \|\zeta_\rho^n\|_{L^\infty(0,T;H^{1/2}(\Omega))}, \end{aligned}$$

$$(4.11) \quad \|\xi_\rho^{n,1}\|_{C([0,T];L^2(\Omega))} \leq C \|f_\rho^n\|_{L^1(0,T;L^2(\Omega))}.$$

Taking (4.9) into account, we conclude that the sequence w_ρ^n converges to zero in the weak* topology of $L^\infty(\Sigma)$ for all $0 < \rho < 1$. Therefore, by Lemma 3.3, the sequence ζ_ρ^n converges to zero in $C([0, T]; L^2(\Omega))$. Thus there is an integer $n(\rho)$ such that

$$(4.12) \quad \|\zeta_\rho^{n(\rho)}\|_{C([0,T];L^2(\Omega))} \leq \rho \quad \text{for all } 0 < \rho < 1.$$

Observe further that $u_\rho^{n(\rho)}$ converges to \bar{u} in $L^2(\Sigma)$ as $\rho \downarrow 0$. It follows now from Theorem 3.5 that $y_\rho^{n(\rho)}$ converges to \bar{y} in $C([0, T]; L^2(\Omega))$ as $\rho \downarrow 0$. Invoking assumption (A1), one has that $\phi_\rho^{n(\rho)}$

converges to zero in $L^1(0, T; L^2(\Omega))$ and that $(a - a_\rho^{n(\rho)})$ converges to zero in $L^1(0, T; L^{2N}(\Omega))$ as $\rho \downarrow 0$. This imply together with (4.10)-(4.12) that

$$\begin{aligned} \lim_{\rho \rightarrow 0} \|\xi_\rho^{n(\rho)}\|_{C([0, T]; L^2(\Omega))} &\leq \lim_{\rho \rightarrow 0} (\|\xi_\rho^{n(\rho), 1}\|_{C([0, T]; L^2(\Omega))} \\ &\quad + \|\xi_\rho^{n(\rho), 2} - \zeta_\rho^{n(\rho)}\|_{C([0, T]; L^2(\Omega))} + \|\zeta_\rho^{n(\rho)}\|_{C([0, T]; L^2(\Omega))}) = 0. \end{aligned}$$

Setting finally $E_\rho := E_\rho^{n(\rho)}$, $u_\rho := u_\rho^{n(\rho)}$, and $\frac{1}{\rho}r_\rho := \xi_\rho^{n(\rho)}$, we end the proof of the theorem. \triangle

5 Proof of Necessary Optimality Conditions

As mentioned, in the proof of our main theorem we are going to use *Ekeland's variational principle* [3], which is one of the most powerful tools of nonlinear analysis especially important in applications of variational methods. In the framework of deriving necessary optimality conditions for the state-constrained problem (P) , Ekeland's variational principle allows us to perform an efficient *strong approximation* of the given optimal solution to the original problem by some functions that happen to be optimal solutions to *perturbed* optimal control problems with *no state constraints*. To furnish this procedure, we first describe a complete metric space and a lower semicontinuous functional, which are suitable for the application of Ekeland's principle to our problem.

Given $\bar{u} \in U_{ad}$ and a fixed positive number k , we define the set

$$U_{ad}(\bar{u}, k) := \{u \in U_{ad} \mid |u(s, t) - \bar{u}(s, t)| \leq k \text{ for a.e. } (s, t) \in \Sigma\}$$

and endow this set with the metric, which goes back to Ekeland's seminal paper [3],

$$d(v, u) := \mathcal{L}^N(\{(s, t) \mid \bar{u}(s, t) \neq u(s, t)\}),$$

where $\mathcal{L}^N(\Omega)$ denotes as before the Lebesgue measure of $\Omega \subset \mathbb{R}^N$. Observe that if $(u_n) \subset U_{ad}(\bar{u}, k)$ and $u \in U_{ad}(\bar{u}, k)$ are such that $\lim_{n \rightarrow \infty} d(u_n, u) = 0$, then the sequence (u_n) *strongly* converges to u in the norm of $L^2(\Sigma)$. It is easy to see that the space $(U_{ad}(\bar{u}, k), d)$ is metric. The next result provides more information about this space and about the cost functional of (P) on it, where y_u stands for the weak solution of (1.1) corresponding to u .

Lemma 5.1 (proper setting for Ekeland's principle). *The metric space $(U_{ad}(\bar{u}, k), d)$ is complete, and the mapping $u \mapsto (y_u, J(y_u, u))$ is continuous from $(U_{ad}(\bar{u}, k), d)$ into $C([0, T]; L^2(\Omega)) \times \mathbb{R}$.*

Proof. The completeness of the space $(U_{ad}(\bar{u}, k), d)$ is a well-known fact; cf. [3, 20]). Let us prove the continuity statement of the lemma based on the regularity of weak solutions to the state system (1.1) established in Section 3.

Take $(u_n) \subset U_{ad}(\bar{u}, k)$ and $u \in U_{ad}(\bar{u}, k)$ such that the control sequence u_n converges to u in the above d -metric as $n \rightarrow \infty$. Denote by y and by y_n the weak solutions of (1.1) corresponding

to u and to u_n , respectively. Since $u_n \rightarrow u$ strongly in $L^2(\Sigma)$, the corresponding arcs y_n converge to y strongly in $C([0, T]; L^2(\Omega))$ by Theorem 3.5. Furthermore, it follows from the estimates in assumptions (A2)–(A4) that the value sequence $J(y_n, u_n)$ converges to $J(y, u)$ as $n \rightarrow \infty$, which ensures the desired continuity. \triangle

Now following [12, Theorem 2.18] and using the classical results in the geometry of Banach spaces, we conclude that the separable space $C([0, T]; L^2(\Omega))$ admits an equivalent norm $|\cdot|_{C([0, T]; L^2(\Omega))}$, which is *strictly convex* on $C([0, T]; L^2(\Omega))$. Moreover, the space $\mathcal{M}([0, T]; L^2(\Omega))$ endowed with the dual norm of $|\cdot|_{C([0, T]; L^2(\Omega))}$, denoted by $|\cdot|_{\mathcal{M}([0, T]; L^2(\Omega))}$, is also strictly convex. Then we define the *distance function*

$$(5.1) \quad d_{\mathcal{C}}(\varphi) := \inf_{z \in \mathcal{C}} |\varphi - z|_{C([0, T]; L^2(\Omega))}$$

to the set \mathcal{C} in the new norm $|\cdot|$ on $C([0, T]; L^2(\Omega))$. Since \mathcal{C} is convex, the distance function (5.1) is convex as well. It is always globally Lipschitzian of rank one satisfying

$$\limsup_{\rho \downarrow 0, \theta \rightarrow \varphi} \frac{d_{\mathcal{C}}(\theta + \rho z) - d_{\mathcal{C}}(\theta)}{\rho} = \max \{ \langle \xi, z \rangle_{\mathcal{M}([0, T]; L^2(\Omega)) \times C([0, T]; L^2(\Omega))} \mid \xi \in \partial d_{\mathcal{C}}(\varphi) \}$$

for every $\varphi, z \in C([0, T]; L^2(\Omega))$, where $\partial d_{\mathcal{C}}$ is the *subdifferential* in the sense of *convex analysis*. Therefore, for a given $\varphi \in C([0, T]; L^2(\Omega))$ one has

$$\langle \xi, z - \varphi \rangle + d_{\mathcal{C}}(\varphi) \leq d_{\mathcal{C}}(z) \quad \text{whenever } \xi \in \partial d_{\mathcal{C}}(\varphi) \text{ and } z \in C([0, T]; L^2(\Omega)),$$

$$|\xi|_{\mathcal{M}([0, T]; L^2(\Omega))} \leq 1 \quad \text{whenever } \xi \in \partial d_{\mathcal{C}}(\varphi).$$

It follows from standard facts of convex analysis that

$$|\xi|_{\mathcal{M}([0, T]; L^2(\Omega))} = 1 \quad \text{for every } \xi \in \partial d_{\mathcal{C}}(\varphi) \text{ and } \varphi \notin \mathcal{C}.$$

Moreover, since $\partial d_{\mathcal{C}}(\varphi)$ is convex and the norm $|\cdot|_{\mathcal{M}([0, T]; L^2(\Omega))}$ is strictly convex on $\mathcal{M}([0, T]; L^2(\Omega))$, the subdifferential $\partial d_{\mathcal{C}}(\varphi)$ is a *singleton* and $d_{\mathcal{C}}$ is *Gâteaux-differentiable* at φ for $\varphi \notin \mathcal{C}$.

Let (\bar{y}, \bar{u}) be an optimal solution to the original problem (P). For each $k = 1, 2, \dots$ we define the *penalized functional* by

$$J_k(y, u) := \left[\left(J(y, u) - J(\bar{y}, \bar{u}) + \frac{1}{k^2} \right)^+ \right]^2 + d_{\mathcal{C}}^2(y),$$

where J is the cost functional in (P) while the distance function $d_{\mathcal{C}}(\cdot)$ is given in (5.1). Since $J_k(\bar{y}, \bar{u}) = k^{-4}$, one has that

$$J_k(\bar{y}, \bar{u}) < \inf \{ J_k(y, u) \mid u \in U_{ad}(\bar{u}, k^{1/3}), (y, u) \text{ satisfies (1.1)} \} + \frac{1}{k^2},$$

for all k , i.e., (\bar{y}, \bar{u}) is a $\frac{1}{k^2}$ -*optimal solution* to the penalized problem.

Note that the functional J_k is *smooth* at points where it *does not vanish*, in the sense that it is Gâteaux-differentiable at such points; cf. [15] in the case of control systems governed by ordinary differential equations. This follows from the construction of J_k , assumptions (A2)–(A4), and the above property of the function (5.1). Ekeland’s principle allows us to *strongly* approximate (\bar{y}, \bar{u}) by a pair (y_k, u_k) satisfying (1.1) in such a way that (y_k, u_k) is an *exact solution* to some *perturbed* optimal control problem for system (1.1) with the same control constraints and no state constraints. The new functional may be *nonsmooth* due to distance-type perturbations, but we can handle this by the choice of metric d in Lemma 5.1.

Proof of Theorem 2.1. We divide the proof of this theorem into the three major steps.

Step 1: Approximating problems via Ekeland’s principle. Given an optimal solution (\bar{y}, \bar{u}) to the original problem (P) , we fix a natural number $k = 1, 2, \dots$ and get from Lemma 5.1 that the metric space $(U_{ad}(\bar{u}, k^{1/3}), d)$ is complete, and that the functional $u \mapsto J_k(y_u, u)$ is lower semicontinuous (even continuous) on this space. By the strong form of Ekeland’s variational principle [3] we find an admissible control u_k satisfying

$$(5.2) \quad \begin{aligned} u_k &\in U_{ad}(\bar{u}, k^{1/3}), \quad d(u_k, \bar{u}) \leq \frac{1}{k}, \quad \text{and} \\ J_k(y_k, u_k) &\leq J_k(y_u, u) + \frac{1}{k}d(u_k, u) \quad \text{for all } u \in U_{ad}(\bar{u}, k^{1/3}), \end{aligned}$$

where y_k and y_u are the weak solutions of (1.1) corresponding to u_k and u , respectively. The latter means that u_k is an *optimal solution* to the *perturbed problem*

$$(P_k) \quad \inf \left\{ J_k(y, u) + \frac{1}{k} \mid u \in U_{ad}(\bar{u}, k^{1/3}), (y, u) \text{ satisfies (1.1)} \right\}$$

for all natural numbers k .

Step 2: Necessary conditions in approximating problems. First take an arbitrary $u_0 \in U_{ad}$ and construct the following modification of \bar{u} feasible to (P_k) by

$$(5.3) \quad u_{0k}(s, t) := \begin{cases} u_0(s, t) & \text{if } |u_0(s, t) - \bar{u}(s, t)| \leq k^{1/3}, \\ \bar{u}(s, t) & \text{otherwise.} \end{cases}$$

Then, given any $0 \leq \rho < 1$, we define *diffuse perturbations* of the optimal control u_k in (P_k) as

$$(5.4) \quad u_\rho^k(s, t) := \begin{cases} u_k(s, t) & \text{on } \Sigma \setminus E_\rho^k, \\ u_{0k}(s, t) & \text{on } E_\rho^k. \end{cases}$$

Theorem 4.1 ensures the existence of measurable sets $E_\rho^k \subset \Sigma$ for which $\mathcal{L}^N(E_\rho^k) = \rho \mathcal{L}^N(\Sigma)$,

$$(5.5) \quad y_\rho^k = y_k + \rho z_k + \rho r_\rho^k, \quad \lim_{\rho \rightarrow 0} \|r_\rho^k\|_{C([0, T]; L^2(\Omega))} = 0, \quad \text{and}$$

$$(5.6) \quad J(y_\rho^k, u_\rho^k) = J(y_k, u_k) + \rho \Delta J^k + o(\rho),$$

where y_ρ^k is the weak solution of (1.1) corresponding to u_ρ^k , where z_k is the weak solution of

$$\begin{cases} z_{tt} - \Delta z + \Phi'_y(\cdot, y_k)z = 0 & \text{in } Q, \\ \partial_\nu z = u_k - u_{0k} & \text{on } \Sigma, \\ z(0) = 0, \quad z_t(0) = 0 & \text{in } \Omega, \end{cases}$$

and where ΔJ^k is defined by

$$\Delta J^k := \int_Q g'_y(\cdot, y_k)z_k \, dxdt + \int_\Omega f'_y(\cdot, y_k(T))z_k \, dx + \int_\Sigma (h(\cdot, u_{0k}) - h(\cdot, u_k)) \, dsdt.$$

Since each u_ρ^k is clearly feasible to (P_k) , we have from (5.2) and the definition of the metric d that

$$(5.7) \quad \lim_{\rho \rightarrow 0} \frac{J_k(y_k, u_k) - J_k(y_\rho^k, u_\rho^k)}{\rho} \leq \frac{1}{k} \mathcal{L}^N(\Sigma).$$

Observe that $J_k(y_k, u_k) \neq 0$ for all k due the optimality of u_k in (P_k) and the structure of J_k . Hence J_k is *Gâteaux differentiable* at (y_k, u_k) by the discussions above. Then it easily follows from (5.6) and (5.7) that

$$(5.8) \quad -\lambda_k \Delta J^k - \langle \mu_k, z_k \rangle \leq \frac{1}{k} \mathcal{L}^N(\Sigma),$$

where the multipliers λ_k and μ_k are computed by

$$\lambda_k := \frac{(J(y_k, u_k) - J(\bar{y}, \bar{u}) + \frac{1}{k^2})^+}{J_k(y_k, u_k)}, \quad \mu_k := \begin{cases} \frac{d_C(\Phi(y_k)) \nabla d_C(\Phi(y_k))}{J_k(y_k, u_k)} & \text{if } \Phi(y_k) \notin \mathcal{C}, \\ 0 & \text{otherwise.} \end{cases}$$

Now let p_k be the (unique) weak solution to the *adjoint system*

$$(5.9) \quad \begin{cases} p_{tt} - \Delta p + \Phi'_y(\cdot, y_k)p = \lambda_k g'_y(\cdot, y_k) + \mu_k|_Q & \text{in } Q, \\ \partial_\nu p = 0 & \text{on } \Sigma = \Gamma \times]0, T[, \\ p(T) = 0, \quad p_t(T) = -\lambda_k f'_y(\cdot, y_k(T)) - \mu_k|_{\Omega \times \{T\}} & \text{in } \Omega, \end{cases}$$

where $\mu_k|_Q$ and $\mu_k|_{\Omega \times \{T\}}$ are the restrictions of μ_k to Q and $\Omega \times \{T\}$, respectively. Employing the Green formula of Theorem 3.8, we have

$$\begin{aligned} & \lambda_k \int_Q g'_y(x, t, y_k)z_k \, dxdt + \lambda_k \int_\Omega f'_y(x, y_k(T))z_k(T) \, dx + \langle \mu_k, z_k \rangle \\ &= \int_Q p_k(z_{ktt} - \Delta z_k + \Phi'_y(\cdot, y_k)z_k) \, dxdt + \int_\Sigma p_k \partial_\nu z_k \, dsdt \\ &= \int_\Sigma p_k(u_k - u_{0k}) \, dsdt. \end{aligned}$$

The latter implies, by (5.8) and the definition of ΔJ^k , that

$$(5.10) \quad \int_{\Sigma} (\lambda_k h(s, t, u_k) + p_k u_k) ds dt \leq \int_{\Sigma} (\lambda_k h(s, t, u_{0k}) + p_k u_{0k}) ds dt + \frac{1}{k} \mathcal{L}^N(\Sigma)$$

for every $k = 1, 2, \dots$, which gives necessary optimality conditions for the solutions u_k to the approximating problems (P_k) .

Step 3: Passing to the limit. To conclude the proof of the theorem, we need to pass to the limit in the above relations for the optimal solutions u_k to (P_k) as $k \rightarrow \infty$. First observe that

$$\lambda_k^2 + |\mu_k|_{\mathcal{M}([0, T]; L^2(\Omega))}^2 = 1 \quad \text{for all } k = 1, 2, \dots$$

Invoking basic functional analysis, we find an element $(\lambda, \mu) \in \mathbb{R} \times \mathcal{M}([0, T]; L^2(\Omega))$, with $\lambda \geq 0$, and a subsequence of (λ_k, μ_k) such that

$$\lambda_k \rightarrow \lambda \text{ in } \mathbb{R} \quad \text{and} \quad \mu_k \rightarrow \bar{\mu} \text{ weak* in } \mathcal{M}([0, T]; L^2(\Omega)).$$

Furthermore, Theorem 3.7 ensures the estimate

$$\begin{aligned} & \|p_k\|_{L^\infty(0, T; H^1(\Omega))} + \|p_{kt}\|_{L^\infty(0, T; L^2(\Omega))} \\ & \leq \|\mu\|_{\mathcal{M}([0, T]; L^2(\Omega))} + \|g'_y(\cdot, y_k)\|_{L^1(0, T; L^2(\Omega))} + \|f'_y(\cdot, y_k(T))\|_{L^2(\Omega)}. \end{aligned}$$

Since the sequence $(\lambda_k) \subset \mathbb{R}$, $(\mu_k) \subset \mathcal{M}([0, T]; L^2(\Omega))$, $(y_k) \subset C([0, T]; L^2(\Omega))$, and $(u_k) \subset L^2(\Sigma)$ are bounded, the sequence (p_k, p_{kt}) is bounded in $L^\infty(0, T; H^1(\Omega)) \times L^\infty(0, T; L^2(\Omega))$. Then there is a subsequence of (p_k, p_{kt}) , which converges weak* in $L^\infty(0, T; H^1(\Omega)) \times L^\infty(0, T; L^2(\Omega))$ to some (p, p_t) belonging to this space. Employing standard arguments, we prove that (p, p_t) is the (unique) weak solution to (2.4).

Taking into account assumption (A5) on the convexity and nonempty interiority of the set \mathcal{C} , one has the necessary condition (2.2) for the limiting multipliers (λ, μ) . It remains to justify the minimum condition (2.3). To furnish this, recall that $u_k \rightarrow \bar{u}$ strongly in $L^2(\Sigma)$. Passing to the limit as $k \rightarrow \infty$ in (5.10), we get

$$(5.11) \quad \int_{\Sigma} (\lambda h(s, t, \bar{u}) + p \bar{u}) ds dt \leq \int_{\Sigma} (\lambda h(s, t, u_0) + p u_0) ds dt \quad \text{for every } u_0 \in U_{ad}.$$

Finally, taking into account the structure of U_{ad} in (A6) and employing the standard arguments (see, e.g., [20, Section 5.2]), we derive the pointwise condition (2.3) from the integral one in (5.11).

References

- [1] V. BARBU, T. PRECUPANU, *Convexity and Optimization in Banach Spaces*, 2nd ed., D. Reidel, Dordrecht, 1986.
- [2] F. H. CLARKE, *The maximum principle under minimal hypotheses*, SIAM J. Control Optim., 14 (1976), pp. 1908–1091.

- [3] I. EKELAND, *On the variational principle*, J. Math. Anal. Appl., 47 (1974), pp. 324–353.
- [4] H. O. FATTORINI, *Optimal control problems with state constraints for semilinear distributed-parameter systems*, J. Optim. Theory Appl., 88 (1996), pp. 25–59.
- [5] H. O. FATTORINI, *Nonlinear infinite-dimensional optimal control problems with state constraints and unbounded control sets*, Rend. Istit. Mat. Univ. Trieste, 28 (1996), suppl. 127-146 (1997).
- [6] H. O. FATTORINI, *Infinite-Dimensional Optimization and Control Theory*, Cambridge University Press, Cambridge, UK, 1999.
- [7] R. GABASOV and F. M. KIRILLOVA, *Qualitative Theory of Optimal Processes*, Marcel Dekker, New York, 1976.
- [8] I. LASIECKA, J.-L. LIONS and R. TRIGGIANI, *Nonhomogeneous boundary value problems for second order hyperbolic operators*, J. Math. Pures Appl., 65 (1986), pp. 149–192.
- [9] I. LASIECKA and R. TRIGGIANI, *Dirichlet boundary control problem for parabolic equations with quadratic cost: analyticity and Riccati's feedback synthesis*, SIAM J. Control Optim., 21 (1983), pp. 41–67.
- [10] I. LASIECKA and R. TRIGGIANI, *Sharp regularity theory for second order hyperbolic equations of Neumann type. Part I: L_2 nonhomogeneous data*, Ann. Mat. Pura Appl., 157 (1990), pp. 285–367.
- [11] I. LASIECKA and R. TRIGGIANI, *Control Theory for Partial Differential Equations*, Cambridge University Press, Cambridge, UK, 2000.
- [12] X. J. LI and J. M. YONG, *Optimal Control Theory for Infinite-Dimensional Systems*, Birkhäuser, Boston, 1995.
- [13] J.-L. LIONS, *Contrôle des Systèmes Distribués Singuliers*, Gauthier-Villars, Paris, 1983.
- [14] J. J. MOREAU, *Bounded variation in time*, in Topics in Nonsmooth Mechanics, J.J. Moreau, P.D. Panagiotopoulos, G. Strang (eds.), pp. 1–74, Birkhäuser, Boston, 1988.
- [15] B. S. MORDUKHOVICH, *Maximum principle in problems of time optimal control with nonsmooth constraints*, J. Appl. Math. Mech., 40 (1976), pp. 960–969.
- [16] B. S. MORDUKHOVICH and J. P. RAYMOND, *Dirichlet boundary control of hyperbolic equations in the presence of state constraints*, Research Report #10, Dept. of Mathematics, Wayne State University, Detroit, MI, 2003.
- [17] B. S. MORDUKHOVICH and K. ZHANG, *Minimax control of parabolic systems with Dirichlet boundary conditions and state constraints*, Appl. Math. Optim., 36 (1997), pp. 323-360.
- [18] P. A. NGUYEN and J. P. RAYMOND, *Control problems for convection-diffusion equations with a control localized on manifolds*, ESAIM COCV, 6 (2001), pp. 417–448.

- [19] J. P. RAYMOND, *Nonlinear boundary control of semilinear parabolic problems with pointwise state constraints*, *Discrete Contin. Dynam. Systems*, 3 (1997), pp. 341–370.
- [20] J. P. RAYMOND and H. ZIDANI, *Pontryagin's principle for state-constrained control problems governed by parabolic equations with unbounded controls*, *SIAM J. Control Optim.*, 38 (1998), pp. 1853–1879.
- [21] J. SIMON, *Compact sets in the space $L^p(0, T; B)$* , *Ann. Mat. Pura Appl.*, 146 (1987), pp. 65–96.
- [22] R. VINTER, *Optimal Control*, Birkhäuser, Boston, 2000.
- [23] L. W. WHITE, *Control of a hyperbolic problem with pointwise stress constraints*, *J. Optim. Theory Appl.*, 41 (1983), pp. 359–369.
- [24] L. W. WHITE, *Distributed control of a hyperbolic problem with control and stress constraints*, *J. Math. Anal. Appl.*, 106 (1985), pp. 41–53.