SECOND ORDER NECESSARY CONDITIONS IN OPTIMAL CONTROL OF EVOLUTION SYSTEMS

H. FRANKOWSKA, E.M. MARCHINI, AND M. MAZZOLA

Abstract. This paper is devoted to second order necessary optimality conditions for a class of infinite dimensional optimal control problems under *functional* pure state constraints together with end point constraints. Using tools of second order variational analysis, we derive necessary optimality conditions in the form of a minimum principle and a second order variational inequality. We further propose sufficient conditions guaranteeing normality of the minimum principle. Finally, applications to optimal control models involving PDEs are provided.

Keywords. semilinear control system, state constraint, optimal control, second order necessary conditions, normality

1. Introduction

In many phenomena, such as heat conduction, reaction-diffusion, population dynamics, economics, one seeks to optimize measures of best performances. Dealing with an optimal control model involving PDEs represents then the natural framework. For this reason optimal control theory of PDEs is a very active field of research since the sixties.

In the present paper we consider a quite general optimal control problem governed by an abstract semilinear system in the infinite dimensional setting. Our aim is to provide tools and results useful when working with some classes of controlled PDEs. Reduction of evolution PDEs to an abstract system can be found in classical books, see e.g. [35, 40]. Thereby our analysis, although performed in an abstract infinite dimensional setting, can be applied to a number of concrete problems.

The main goal of our paper concerns second order necessary optimality conditions. Second order analysis for PDEs has been extensively discussed (with particular emphasis on second order sufficient conditions) since it plays a crucial role in getting stability properties that are fundamental for the numerical investigations. Concerning the evolution systems, there is a rich literature on optimal control of semilinear parabolic equations. Significant contributions in the last two decades can be found in [3, 4, 5, 8, 9, 10, 11, 12, 13, 14, 17, 18, 31, 33, 36, 37, 38], see also the bibliographies therein. State constrained problems have been analyzed too because their presence is compulsory in many applied models. The usual approach to derive second order optimality conditions in the presence of state constraints consists in rewriting the control problem as an abstract mathematical programming one. To obtain the Lagrange multipliers and second order optimality conditions in such a context, some constraint qualifications, like Robinson's condition,

2010 *Mathematics Subject Classification.* 34G25, 35Q93, 46N10, 47J22, 47N70, 49J20, 93C20.

This research benefited from the support of the FJMH Program PGMO and from the support to this program from EDF-THALES-ORANGE-CRITEO' under grant PGMO 2018-0047H..

are usually imposed. However going back to the original problem and to translate the obtained necessary conditions in terms of the problem data is, in general, difficult, the abstract constraint qualifications being not simple to verify in concrete cases. For this reason, in the literature, some additional structural requirements like continuity (or piecewise continuity) of optimal controls, their bang-bang or singular nature and so on, are imposed to derive optimality conditions. Let us underline that piecewise continuity of optimal controls can not be guaranteed (especially for state-constrained nonlinear systems) by the available existence theorems even in a much simpler finite dimensional setting. Also often structural assumptions on the set of controls and the underlying control system are imposed, like being affine in the control variable with the set of controls having a polyhedral structure. Naturally, it is preferable to have results avoiding structural assumptions on optimal controls and to develop an approach valid for general nonlinear control systems. To our knowledge, a general theory for the pure state constrained and end-point constrained case is lacking.

In the present work we consider state and end point constraints described by inequalities on a Banach space X. Neither convexity assumptions nor Robinson's type constraints qualifications are imposed, but we request the non triviality of gradients of maps describing pure state constraints at their boundary points. This assumption guarantees the non-emptiness of the interior of the second order tangent (see the definition in subsection 2.1) to the set of continuous functions from an interval I to X satisfying the pure state constraints for every $t \in I$.

Instead of a reduction to a mathematical programming problem, we use second order variational analysis and derivatives of set-valued maps to handle systems depending nonlinearly on controls. Derivatives of set-valued maps allow us to express larger linearizations of nonlinear systems than the familiar ones. More precisely, we consider second order tangents to functional constraints and solutions to a second order linearization of control system at a locally optimal solution and derive a variational inequality. This inequality and separation theorems lead to both first and second order necessary optimality conditions.

Here, on one hand, we deal directly with arbitrary measurable optimal controls and general control sets avoiding structural assumptions mentioned above. On the other hand, using semilinear differential inclusions, we linearize twice the nonlinear evolution system and profit from a relaxation theorem and from some additive properties of the set-valued derivatives, to get a larger linearization. The necessary optimality conditions are first obtained in the form of nonintersection of convex sets related to the second order linearization of constraints and of the cost function. Its dual characterization implies second order necessary optimality conditions. Strikingly, it leads also to first order optimality conditions in the form of the minimum principle, allowing to have a unified light on both.

More to the point, in an infinite dimensional separable Banach space X, we consider the solutions $x : I = [0,1] \to X$ of the control system

$$\dot{x}(t) = \mathbb{A}\, x(t) + f(t, x(t), u(t))\,, \quad x(0) = x_0\,, \tag{1.1}$$

that satisfy an end point constraint

$$x(1) \in Q = \bigcap_{i=1,\dots,k} Q_i = \bigcap_{i=1,\dots,k} \{x \in X : g_i(x) \le 0\} \tag{1.2}$$

and a state constraint

$$x(t) \in K = \bigcap_{j=1,\dots,q} K_j = \bigcap_{j=1,\dots,q} \{x \in X : \varphi_j(x) \le 0\}, \quad \text{for any } t \in I\,. \tag{1.3}$$

Here, u is a (Lebesgue) measurable function taking values in a given closed non-empty bounded set $U \subset Z$, and Z is a separable Banach space. The densely defined unbounded linear operator $\mathbb{A}$ is the infinitesimal generator of a strongly continuous semigroup $S(t) : X \to X$, the map $f : I \times X \times Z \to X$ is twice Fréchet differentiable with respect to the second variable x and the third variable u. The initial condition $x_0 \in X$ is given. The functions $g_i : X \to \mathbb{R}$ are twice Fréchet differentiable while $\varphi_j : X \to \mathbb{R}$ are twice continuously Fréchet differentiable. The trajectories of (1.1) are understood in the mild sense (see [35] or Definition 2.1 below). In this paper we investigate an infinite dimensional Mayer problem: given a twice Fréchet differentiable map $g_0 : X \to \mathbb{R}$, consider the problem

$$\text{minimize} \quad \left\{ g_0(x(1)) : x(\cdot) \text{ is a solution of (1.1)–(1.3) for some control } u(\cdot) \right\}. \tag{1.4}$$

When the semigroup $S(\cdot)$ is compact, sufficient conditions for the existence of optimal solutions to the above problem can be deduced for instance from [6], provided there exists at least one trajectory of (1.1)–(1.3). The main additional requirement being the assumption that $f(t, x, U)$ is closed and convex for all (t, x).

In this paper U is assumed to be an arbitrary bounded and closed subset of a separable Banach space. In particular it may be a finite union of bounded closed sets with smooth boundaries, the case that is not handled by the classical mathematical programming, where one rather works with intersections. To define second order linearizations of control systems we shall use second order tangent sets to U, in particular, assuming their nonemptiness. That is for the second order theory more regularity from U is requested, than for the first order conditions. To compensate the generality of the set U, when making first order variations of controls, we shall also impose a geometric assumption (2.12) below.

Concrete realizations of our very general problem are provided in section 4 where we apply our analysis to a Mayer problem governed by the hyperbolic equation:

$$\partial_{tt}\xi(t, \mathbf{x}) + \partial_t\xi(t, \mathbf{x}) - \Delta\xi(t, \mathbf{x}) + \sin(\xi(t, \mathbf{x})) = u(t, \mathbf{x}), \tag{1.5}$$

with $\xi = \xi(t, \mathbf{x})$, $t \in I$, $\mathbf{x} \in \Omega$, a smooth and bounded domain in $\mathbb{R}^3$, $u = u(t, \cdot) \in U$, a given bounded closed non empty subset of $L^2(\Omega)$.

A further example, involving a parabolic equation is analyzed in detail in section 4:

$$\partial_t x(t, \mathbf{x}) - \Delta x(t, \mathbf{x}) = \zeta(t, \mathbf{x}) + b(\mathbf{x})^T u(t) x(t, \mathbf{x}). \tag{1.6}$$

Here the state space is $X = L^2(\Omega)$, $x = x(t, \mathbf{x})$, $t \in I$, and $\mathbf{x} \in \Omega$, a bounded domain of $\mathbb{R}^N$ with smooth boundary $\partial\Omega$, $\zeta \in L^1(I, L^2(\Omega))$, $b \in L^\infty(\Omega, \mathbb{R}^m)$, u takes values in U, a convex polytope of $\mathbb{R}^m$. When state constraints are of a functional type as above, the spatial domain Ω may have any dimension N to derive necessary optimality conditions.

We would like to underline already here that for problems involving controlled PDEs, continuity (or boundedness) of solutions do allow to extend the proposed approach also to optimal control problems involving pointwise state constraints, see section 5 for an illustration of this remark. Such continuity requirement, however, does not diminish the developed approach. Indeed, in the literature devoted to necessary optimality conditions, usually authors do impose assumptions implying, thanks to embedding theorems, that solutions of controlled PDEs are continuous in the state variable $\mathbf{x}$, or are at least essentially bounded. We have privileged to deal with an abstract case, to propose a blue print technique that may be applied to various frameworks.

To simplify the exposition, in the examples we consider quadratic costs and simple constraints, nevertheless non quadratic costs and the presence of a finite number of regular inequality constraints (as in (1.2) and (1.3)) can be investigated as well.

In this paper, the necessary optimality conditions are stated in a very general abstract setting, including both hyperbolic and parabolic problems. This choice entails the assumptions that turn out to be restrictive in some applications. For instance, concerning the forcing term f, while in (1.5) we could consider a larger class of nonlinear functions, twice Fréchet differentiability in the state space $L^2(\Omega)$ is a strong assumption. For this reason in (1.6) we restrict our analysis to the bilinear case. Let us underline that many papers dealing with second order conditions for control problems involving parabolic PDEs rely on smoothing properties of the control-to-state mapping. This regularity allows to treat more general forcing term f and to weaken the strong differentiability assumptions required in our paper. Nevertheless, the developed abstract tools can be applied also to some parabolic PDEs after minor modifications of proofs that allow to profit of such smoothing effect and to request the twice differentiability of the forcing term only on a subspace of $L^2(\Omega)$. In particular, this was illustrated in [25, Example 5.2], where an optimal control problem not involving state constraints was considered and second order optimality conditions were derived for weak minima. For this reason we believe that the techniques developed in our paper could be adapted and refined to suit a larger class of problems involving PDEs, by using this further regularity information on the solutions. For instance, in section 5 we discuss the case of pointwise state constraints based on the abstract ideas of section 3.

Let us compare our results with some relevant contributions on second order conditions in the pure state constrained case. The main novelty of this work is the approach based on second order variational analysis, dealing with abstract evolution PDEs. Concerning the applications presented here, notice that the great generality of our setting allows to treat problems involving state constrained hyperbolic equations, that is new in this field, to our knowledge.

On the other hand, the parabolic case has been frequently analyzed. The results have been obtained under various regularity assumptions on the data and imposing some constraints qualifications, leading to qualified (normal) necessary conditions.

In the approach developed here the only constraint qualification assumed in order to get (possibly abnormal) first and second order optimality conditions is that $\nabla\varphi_j$ do not vanish on the boundary of K_j, for all j.

Concerning the literature devoted to parabolic optimal control problems with pure state constraints let us mention [3, 38], where the authors obtain, among other results, second order necessary conditions for problems with integral constraints on the final state. Several assumptions on the data are imposed (such as monotonicity), together with Robinson's type constraint qualifications. Another contribution in the same direction can be found in [4], where no end point constraints are present and a simple functional state constraint is considered. Constraints qualifications are so that they imply normality of the obtained necessary conditions. This

In this paper we are mostly interested by the functional constraints (1.2), (1.3). In particular cases when X is a Sobolev space of functions defined on an open set Ω, one might be also interested by the pointwise state constraint for $x(t) \in X$, in the form $g(t, x(t, \mathbf{x})) \leq 0$ for all (or almost all) $\mathbf{x} \in \Omega$. This case can not be reduced to functional state constraints. However the developed techniques still apply by a suitable adaptation of the proofs of section 3, using the embedding theorems that guarantee continuity of $x(t, \cdot)$. This is an ingredient that figures also in many publications on second order optimality conditions because it is important to insure the openness of some sets involved in the Robinson or Slater type constraint qualifications. In [8] for a one-dimensional, linear in the control variable parabolic problem involving pointwise state constraints, first order necessary conditions in normal form and second order sufficient ones are proposed. To insure normality, linearized Slater's conditions are imposed. In [33] this analysis is extended in dimensions two and three when the state appears linearly in the constraint. Such restriction on the dimension of the space variable is needed to guarantee continuity of solutions. The results of [8] are further generalized in [18] to an arbitrary space dimension, the finite dimensional control space and again Slater's conditions are assumed to get normal first order necessary conditions. Other regular constraints qualifications and some restrictions on the dimension of the domains are required also in [36].

In our paper, we do not need to have normal first order necessary conditions to deduce (possibly abnormal) second order ones. Nevertheless, for the sake of completeness, a sufficient condition for normality is provided. It consists in a constraint qualification involving an inward pointing condition that can be linked to the Slater condition when state constraints are convex. The interested reader can find many details on this condition in [26, 27, 28].

We would like to mention that a further step in our analysis will be derivation of second order sufficient conditions applying methods similar to those developed here.

The paper is organized as follows. In section 2 we recall notations, definitions and main assumptions in use. Section 3 contains our main results and their proofs, while some technical proofs are postponed to the appendix. Applications to optimal control problems involving heat and wave equations are proposed in section 4. In section 5 we briefly discuss how the same approach works for controlled PDEs in the presence of pointwise state constraints.

2. Preliminaries

In this section we list the notation and the main assumptions in use throughout the paper.

2.1. **Notation.**

- X is a separable Banach space;
- $B(x,r)$ denotes the closed ball of center $x \in X$ and radius $r > 0$; B is the closed unit ball in X centered at 0;
- given a Banach space Y, $\mathbb{L}(X,Y)$ denotes the Banach space of bounded linear operators from X into Y (if $Y = X$, we write $L(X)$ instead of $\mathbb{L}(X,X)$), $\mathcal{C}(I,X)$ denotes the space of continuous functions from I to X, $L^p(I,X)$ the space of Bochner L^p integrable functions from I to X, and $L^\infty(I,X)$ the space of measurable essentially bounded functions from I to X, $\mathcal{M}(I,Y)$ the space of countably additive regular measures of bounded variation on I with values in Y; recall that $\mathcal{M}(I,X^*)$ is isomorphic to the dual space of $\mathcal{C}(I,X)$, see [22, Theorem 10.2.14];
- $\langle \cdot, \cdot \rangle$ stands for the duality pairing on $X^* \times X$;
- μ is the Lebesgue measure on the real line;
- given $K \subset X$, denote by ∂K its boundary, by $\mathrm{int}(K)$ its interior, by $\overline{K}$ its closure, by K^c its complement, by $\mathrm{co}K$ its convex hull, by $\overline{\mathrm{co}}K$ its closed convex hull and by $\mathrm{cone}K$ its conical hull;
- given $x \in X$, a locally Lipschitz around x set-valued map[1] $\mathcal{F} : X \rightsquigarrow X$ and $y \in \mathcal{F}(x)$, the adjacent derivative $d\mathcal{F}(x,y) : X \rightsquigarrow X$ is defined by
$$v \in d\mathcal{F}(x,y)u \quad \Longleftrightarrow \quad \lim_{h \to 0+} \mathrm{dist}\Big(v, \frac{\mathcal{F}(x+hu) - y}{h}\Big) = 0\,;$$
and, for $v \in d\mathcal{F}(x,y)u$, the second order adjacent variation $d^2\mathcal{F}(x,y,u,v) : X \rightsquigarrow X$ by
$$z \in d^2\mathcal{F}(x,y,u,v)w \quad \Longleftrightarrow \quad \lim_{h \to 0+} \mathrm{dist}\Big(z, \frac{\mathcal{F}(x+hu+h^2w) - y - hv}{h^2}\Big) = 0\,.$$
- given $K \subset X$, denote by $T_K^\flat(x)$ and $T_K^{\flat(2)}(x,y)$ respectively the adjacent tangent cone to K at $x \in \overline{K}$ and the second order tangent set to K at $(x,y) \in \overline{K} \times X$, defined as
$$T_K^\flat(x) = \Big\{ y \in X : \lim_{h \to 0^+} \mathrm{dist}\Big(y, \frac{K - x}{h}\Big) = 0 \Big\}$$
and
$$T_K^{\flat(2)}(x,y) = \Big\{ z \in X : \lim_{h \to 0^+} \mathrm{dist}\Big(z, \frac{K - x - hy}{h^2}\Big) = 0 \Big\}\,,$$
where dist stands for the distance. When K is convex, we write $T_K(x)$ instead of $T_K^\flat(x)$ to follow the usual notations of convex analysis. $C_K(x)$ denotes the Clarke tangent cone to K at x:
$$C_K(x) = \Big\{ y \in X : \lim_{h \to 0^+, x' \to_K x} \mathrm{dist}\Big(y, \frac{K - x'}{h}\Big) = 0 \Big\}\,;$$

[1] a set-valued map $\mathcal{F} : X \rightsquigarrow X$ is called locally Lipschitz around $x \in X$, if there exists a neighborhood $\mathcal{U}$ of x and $l > 0$ satisfying
$$\mathcal{F}(x_1) \subset \mathcal{F}(x_2) + l\|x_1 - x_2\|_X B\,, \quad \forall\, x_1, x_2 \in \mathcal{U}\,;$$

Notice that $z \in T_K^{\flat(2)}(x,y)$ if and only if for every $h>0$, there exists $r(h) \in X$ such that $\|r(h)\|_X = o(h^2)$ and $x + hy + h^2 z + r(h) \in K$, implying that if $T_K^{\flat(2)}(x,y) \neq \emptyset$, then $y \in T_K^\flat(x)$. See [2, Chapter 4] for some calculus of tangent cones.
Below, every Lebesgue measurable function $u : I \to U$ is called a control.

Definition 2.1. Let $x_0 \in X$. A function $x \in \mathcal{C}(I,X)$ is a (mild) solution of (1.1) for the initial datum $x(0) = x_0$ and a control $u : I \to U$ if it satisfies

$$x(t) = S(t)x_0 + \int_0^t S(t-s) f(s, x(s), u(s))\, ds, \quad \text{for any } t \in I\,, \tag{2.1}$$

see [35]. We denote by $\mathcal{S}$ the set of all (mild) solutions of (1.1), namely

$$\mathcal{S} = \big\{ x \in \mathcal{C}(I,X) : x \text{ solves (2.1) for some control } u \big\}\,.$$

If $x \in \mathcal{S}$ corresponds to a control $u(\cdot)$, then we use the notation $f^x(t) = f(t, x(t), u(t))$ for a.e. $t \in I$. Let us remark that even though $u(\cdot)$ associated to $x \in \mathcal{S}$ is not necessarily unique, $f^x(t)$ is well defined for a.e. $t \in I$. If in addition x satisfies (1.2)–(1.3), we say that (x,u) is an admissible pair for problem (1.4).

Since $S(\cdot)$ is a strongly continuous semigroup, there exists $M_S > 0$ such that

$$\|S(t)\|_{L(X)} \le M_S, \qquad \text{for any } t \in I\,. \tag{2.2}$$

Definition 2.2. We say that $(\bar{x}, \bar{u})$ is a strong local minimizer for problem (1.4) if there exists $\varepsilon > 0$ such that, for any admissible pair (x,u) for problem (1.4) satisfying $\|x - \bar{x}\|_{L^\infty(I,X)} < \varepsilon$, we have $g_0(x(1)) \ge g_0(\bar{x}(1))$.

2.2. **Assumptions.** The following conditions (H) on the functions f, g_i and φ_j are imposed in the main results:

(i) f is measurable in t, twice Fréchet differentiable with respect to (x,u);

(ii) for any $R > 0$, there exists $k_R \in L^1(I, \mathbb{R}^+)$ such that, for a.e. $t \in I$ and any $u \in U$, $f(t,\cdot,u)$ is $k_R(t)$-Lipschitz on RB, namely, for a.e. $t \in I$, any $u \in U$, and, any $x, y \in RB$,

$$\|f(t,x,u) - f(t,y,u)\|_X \le k_R(t) \|x-y\|_X\,; \tag{2.3}$$

(iii) there exists $\phi \in L^1(I, \mathbb{R}^+)$ such that, for a.e. $t \in I$, any $x \in X$ and any $u \in U$,

$$\|f(t,x,u)\|_X \le \phi(t)\big(1 + \|x\|_X\big)\,; \tag{2.4}$$

(iv) for every $R > 0$ there exist $\psi_R > 0$ and $\delta > 0$ such that for a.e. $t \in I$

$$\|f(t,x,u) - f(t,x,v)\|_X \le \psi_R \|u - v\|_Z\,, \quad \forall\, u,v \in U + \delta B\,,\ \forall\, x \in RB\,; \tag{2.5}$$

(v) for every $R > 0$ there exists $\phi_R \in L^1(I, \mathbb{R}^+)$ and $\delta > 0$ such that for a.e. $t \in I$

$$\begin{aligned} &\|f_x(t,x,u) - f_x(t,y,v)\|_{L(X)} + \|f_u(t,x,u) - f_u(t,y,v)\|_{\mathbb{L}(Z,X)} \\ &\qquad \le \phi_R(t)(\|x-y\|_X + \|u-v\|_Z)\,, \quad \forall\, u,v \in U + \delta B\,,\ \forall\, x,y \in RB\,; \end{aligned} \tag{2.6}$$

(vi) g_i, for $i = 0, \dots k$, are twice Fréchet differentiable; φ_j, for $j = 1, \dots q$, are twice continuously Fréchet differentiable;

(vii) $\nabla \varphi_j(x) \neq 0\,, \quad \forall\, x \in \partial K_j\,,\ j = 1, \dots, q\,.$

Remark 2.3. The hypotheses (H)$(i)-(iii)$ on the dynamics ensure that to every control $u(\cdot)$ corresponds a unique mild solution defined on the whole interval I. Assumption (2.4) prevents blow-up phenomena, while Lipschitz condition (2.3) implies that to each control corresponds the unique mild solution. Assumption (iii) is quite strong. It is needed to guarantee the existence and apriori bounds on solutions corresponding to each control u. For many semilinear evolution equations of mathematical physics, existence and uniqueness can be shown for arbitrary long time intervals in the framework of weak solutions. Also the above smooth differentiability assumptions (i), (v), (vi) are very strong. For some PDEs enjoying smoothing effects such differentiability can be required on smaller subspaces of X, see [25, Example 5.2].

2.3. **First and second order variations of mild solutions.** Let $G : I \times X \rightsquigarrow X$ be a set-valued map. Recall that $x \in \mathcal{C}(I, X)$ is called a (mild) solution to the semilinear differential inclusion

$$\dot{x}(t) \in \mathbb{A}x(t) + G(t, x(t))\,, \quad x(0) = x_0$$

if $x(0) = x_0$ and for an integrable selection $v(t) \in G(t, x(t))$ it holds

$$x(t) = S(t)x_0 + \int_0^t S(t-s)v(s)ds\,, \quad \forall\, t \in I\,.$$

We consider here first and second order linearizations of a semilinear differential inclusion associated to our control system. For $t \in I$ and $x \in X$, set

$$y(t) = \int_0^t S(t-s)\pi^y(s)ds\,, \quad \forall\, t \in I\,; \qquad F(t,x) = \overline{\mathrm{co}} f(t,x,U)\,.$$

Let $x \in \mathcal{S}$ and consider the following linearized differential inclusion:

$$\dot{y}(t) \in \mathbb{A}y(t) + d_x F(t, x(t), f^x(t))y(t)\,, \quad y(0) = 0\,, \tag{2.7}$$

where $d_x F(t,\alpha,\beta)$ stands for the adjacent derivative of $F(t,\cdot)$ at (α,β). A function $y \in \mathcal{C}(I,X)$ is called a mild solution of (2.7) if for some integrable selection $\pi^y(t) \in d_x F(t, x(t), f^x(t))y(t)$,

$$y(t) = \int_0^t S(t-s)\pi^y(s)ds\,, \quad \forall\, t \in I\,. \tag{2.8}$$

Definition 2.4. A mild solution y of (2.7) is called a *first order variation* if there exist $a \in L^1(I, \mathbb{R}^+)$ and $h_0 > 0$ such that

$$\mathrm{dist}\big(f^x(t) + h\pi^y(t), F(t, x(t) + hy(t))\big) \le a(t)h^2\,, \quad \forall\, 0 < h \le h_0\,. \tag{2.9}$$

We say that a first order variation y is *admissible* if

$$\langle \nabla g_i(x(1)), y(1)\rangle \le 0\,, \text{ for any } i \in \{1,\dots,k\} \text{ such that } x(1) \in \partial Q_i$$

and

$$\langle \nabla \varphi_j(x(t)), y(t)\rangle \le 0\,, \text{ for any } j \in \{1,\dots q\} \text{ and any } t \in I \text{ such that } x(t) \in \partial K_j\,.$$

We will denote by $\mathcal{V}^1(x)$ the set of all admissible first order variations associated to x.

A function $y \in \mathcal{V}^1(x)$ is called *critical* if

$$\langle \nabla g_0(x(1)), y(1)\rangle = 0\,.$$

Fix an admissible pair $(\bar{x}, \bar{u})$ for problem (1.4). In order to obtain first order variations associated to $\bar{x}$, we will consider the linearized control system

$$\dot{y}(t) = \mathbb{A}y(t) + f_x[t]y(t) + f_u[t]v(t)\,, \quad y(0) = 0\,, \tag{2.10}$$

with

$$v(\cdot) \in L^\infty(I, X) \text{ and } v(t) \in T_U^\flat(\bar{u}(t)) \text{ for a.e. } t \in I. \tag{2.11}$$

Here, $[t]$ stands for $(t, \bar{x}(t), \bar{u}(t))$. We shall consider trajectory-control pairs $(\bar{y}, \bar{v})$ that are solutions of (2.10)–(2.11) and satisfy the following:

$$\begin{cases} \text{There exist } c \in L^1(I, \mathbb{R}^+) \text{ and } h_0 > 0 \text{ such that} \\ \operatorname{dist}_U(\bar{u}(t) + h\bar{v}(t)) \le c(t)h^2\,, \quad \forall\, 0 < h \le h_0\,, \text{ a.e. in } I\,. \end{cases} \tag{2.12}$$

Remark 2.5. Condition (2.12) plays an important role to deduce the conclusion of Lemma 2.9 below from the Lebesgue dominated convergence theorem.

In order to illustrate a case where the estimate (2.12) can be easily verified, let us suppose that the oriented distance $\tilde{d}_U : Z \to \mathbb{R}$ to the set U, defined by

$$\tilde{d}_U(z) = \begin{cases} \operatorname{dist}(z, U) & \text{if } z \notin U \\ -\operatorname{dist}(z, U^c) & \text{if } z \in U\,, \end{cases}$$

is Fréchet differentiable on $\partial U + \varepsilon B$ for some $\varepsilon > 0$ with Lipschitz $\nabla\tilde{d}_U(\cdot)$ on $\partial U + \varepsilon B$. Remark that in this case if $u \in \partial U$, then $v \in T_U^\flat(u)$ implies $\langle \nabla\tilde{d}_U(u), v\rangle \le 0$. Let $(\bar{y}, \bar{v})$ be a solution of (2.10)–(2.11) and assume that there exists $\delta > 0$ such that for a.e. $t \in I$ satisfying $\tilde{d}_U(\bar{u}(t)) > -\delta$, it holds

$$\langle \nabla\tilde{d}_U(\bar{u}(t)), \bar{v}(t)\rangle \le 0\,.$$

By the Taylor's formula, for all small $h > 0$ and for all $t \in I$ as above we have

$$\tilde{d}_U(\bar{u}(t) + h\bar{v}(t)) = \tilde{d}_U(\bar{u}(t)) + h\langle \nabla\tilde{d}_U(\bar{u}(t)), \bar{v}(t)\rangle + O_t(h^2),$$

where $|O_t(h^2)| \le Ch^2$ for some $C > 0$ independent from time. If h is sufficiently small, then for all t, satisfying $\tilde{d}_U(\bar{u}(t)) \le -\delta$, we obtain $\tilde{d}_U(\bar{u}(t) + h\bar{v}(t)) \le 0$. Otherwise, whenever $h > 0$ is sufficiently small, for a.e. $t \in I$ satisfying $\tilde{d}_U(\bar{u}(t)) > -\delta$, our assumptions imply $\tilde{d}_U(\bar{u}(t) + h\bar{v}(t)) \le h^2 C$. Therefore, there exist $c, h_0 > 0$ such that

$$\operatorname{dist}_U(\bar{u}(t) + h\bar{v}(t)) \le ch^2\,, \quad \forall\, 0 < h \le h_0\,, \text{ a.e. } t \in I\,.$$

For instance if $Z = L^2(\Omega; \mathbb{R}^n)$ with $\Omega \subset \mathbb{R}^n$ open, the unit ball in Z has the oriented distance function $\tilde{d}_U(u) = \|u\| - 1$ if $\|u\| \ge 1$ and $\tilde{d}_U(u) = 1 - \|u\|$ if $\|u\| \le 1$. Hence $\tilde{d}_U(\cdot)$ is Fréchet differentiable with Lipschitz $\nabla\tilde{d}_U$ on $\partial U + \varepsilon B$ for every $\varepsilon \in (0, 1)$.

The above example can be generalized to the case where the set U is defined by a system of inequalities (see [30, section 7]). For instance in the case of box constraints U, (2.12) is verified for $\bar{v} = u - \bar{u}$, where u is any control satisfying box constraints. The same is valid for convex constraints U.

Lemma 2.6. *Assume (H) and let $(\bar{x}, \bar{u})$ be an admissible pair for problem* (1.4) *and $(\bar{y}, \bar{v})$ be a solution of* (2.10)–(2.11) *satisfying* (2.12). *Then, $\bar{y}$ is a first order variation.*

The proof of Lemma 2.6 is given in the Appendix.

Definition 2.7. Let $x \in \mathcal{S}$ and y be a first order variation. Then $z \in \mathcal{C}(I, X)$ is called a *second order variation* if it is a solution of the semilinear inclusion

$$\dot{z}(t) \in \mathbb{A}z(t) + d_x^2 F[t]z(t)\,, \quad \text{for a.e. } t \in I\,, \quad z(0) = 0\,,$$

with $[t] = (t, x(t), f^x(t), y(t), \pi^y(t))$ and $\pi^y(t) \in d_x F(t, x(t), f^x(t))y(t)$ as in (2.8). We will denote by $\mathcal{S}^2(x, y)$ the set of second-order variations associated to the pair (x, y).

Remark 2.8. In general, the set $\mathcal{S}^2(x, y)$ may be empty. In subsection 3.2 we provide conditions implying its nonemptiness.

Lemma 2.9. *Assume (H)(i)-(iii) and let $x \in \mathcal{S}$, y be a first order variation and $z \in \mathcal{S}^2(x, y)$. Then, for any sequence $h_n \to 0^+$, there exist $x_n \in \mathcal{S}$, $n = 1, 2, \dots,$ such that*

$$\frac{x_n - x - h_n y}{h_n^2} \to z \qquad \text{in } \mathcal{C}(I, X), \text{ as } n \to \infty\,.$$

The proof can be found in the Appendix.

Define the trace operator $\Gamma : \mathcal{C}(I, X) \to X$ by $\Gamma(x) := x(1)$. We conclude this section with two technical lemmas whose proofs are again postponed to the Appendix.

Lemma 2.10. *Let $0 \neq \xi \in X^*$ and consider the set*

$$\mathcal{A} = \big\{ w \in \mathcal{C}(I, X) : \langle \xi, w(1) \rangle \leq 0 \big\}\,.$$

Then, the negative polar cone $\mathcal{A}^-$ is given by

$$\mathcal{A}^- = \big\{ \lambda \Gamma^* \xi : \lambda \geq 0 \big\}\,.$$

Lemma 2.11. *Let $\xi \in \mathcal{C}(I, X^*)$, and $M \subset I$ be closed non empty and such that $\xi(t) \neq 0$, for any $t \in M$. Consider the set*

$$\mathcal{A} = \big\{ w \in \mathcal{C}(I, X) : \sup_{t \in M} \langle \xi(t), w(t) \rangle \leq 0 \big\}.$$

Then

$$\begin{aligned} \mathcal{A}^- = \Big\{ \gamma \in \mathcal{M}(I, X^*) : \exists\, \psi \in \mathcal{M}(I, \mathbb{R}) \ &\textit{positive such that } \operatorname{supp} \psi \subset M \ \textit{ and} \\ \int_I \langle w(t), \gamma(dt) \rangle = \int_I \langle \xi(t), w(t) \rangle \psi(dt),\ &\forall\, w \in \mathcal{C}(I, X) \Big\}\,. \end{aligned} \tag{2.13}$$

3. Main results

To state the main result on second order necessary conditions, see Theorem 3.1 below, fix a local minimizer $(\bar{x}, \bar{u})$ for problem (1.4) (see definition 2.2) and a trajectory-control pair $(\bar{y}, \bar{v})$ of (2.10)–(2.11) satisfying (2.12). By Lemma 2.6, we know that if assumptions (H) hold true, then such $\bar{y}$ is a first order variation. We call such $(\bar{y}, \bar{v})$ a *critical* pair whenever $\bar{y}$ is critical in the sense of Definition 2.4. Define the Hamiltonian $\mathcal{H} : I \times X \times Z \times X^* \to \mathbb{R}$ as

$$\mathcal{H}(t, x, u, p) = \langle p, f(t, x, u) \rangle\,.$$

For all $0 \le t_0 \le t \le 1$, denote by $\mathcal{T}(t_0, t) : X \to X$ the solution operator that associates to any $y_0 \in X$ the value $y(t)$ at time t of the solution to the linear system

$$\begin{cases} \dot{y}(s) = \mathbb{A}y(s) + f_x[s]y(s) \\ y(t_0) = y_0 \,. \end{cases}$$

Assumption 2.3 ensures that $\mathcal{T}$ is strongly continuous, see [21, Theorem 5.5.6]. Finally, introduce the sets

$$\mathcal{W} = \{w : I \to Z \ \text{ measurable} : f_u[\cdot]w(\cdot) \in L^1(I, X)\,,\ w(t) \in T_U^{\flat(2)}(\bar{u}(t), \bar{v}(t)) \text{ a.e. in } I\}\,,$$

$$I_g = \{i \in \{1, \ldots, k\} : \bar{x}(1) \in \partial Q_i\,,\ \langle \nabla g_i(\bar{x}(1)), \bar{y}(1)\rangle = 0\}\,,$$

$$\mathcal{M}_{j0} = \{t \in I : \bar{x}(t) \in \partial K_j\}\,, \quad j = 1, \ldots q\,,$$

and, for all $\delta > 0$,

$$\mathcal{M}_{j\delta} = \left\{t \in I : \operatorname{dist}(\bar{x}(t), \partial K_j) \le \delta\right\}, \quad j = 1, \ldots q.$$

Below we shall use the convention $\max \emptyset = -\infty$.

Theorem 3.1. *Assume (H). Let $(\bar{x}, \bar{u})$ be a local minimizer of (1.4) and $(\bar{y}, \bar{v})$ be a solution of (2.10)–(2.11) satisfying (2.12) and such that $\mathcal{W} \neq \emptyset$. If $\bar{y}$ is critical and*

$$\exists\, \delta > 0 \ \ \textit{satisfying} \ \ \max_{t \in \mathcal{M}_{j\delta}} \langle \nabla\varphi_j(\bar{x}(t)), \bar{y}(t)\rangle \le 0\,, \quad \forall\, j = 1, \ldots q\,, \tag{3.1}$$

then, there exist $\lambda_0 \in \{0, 1\}$, $\lambda_i \ge 0$, for $i \in I_g$, and positive $\psi_j \in \mathcal{M}(I, \mathbb{R})$ with $\operatorname{supp} \psi_j \subset \mathcal{M}_{j0}$, for $j = 1, \ldots, q$, not all equal to zero, such that the function $p : I \to X^*$ defined by

$$p(t) = \mathcal{T}(1, t)^*\Big(\sum_{i \in I_g \cup \{0\}} \lambda_i \nabla g_i(\bar{x}(1))\Big) + \int_t^1 \mathcal{T}(s, t)^*\Big(\sum_{j=1}^q \nabla\varphi_j(\bar{x}(s))\Big)\psi_j(ds) \tag{3.2}$$

satisfies the Pontryagin minimum principle

$$\mathcal{H}(t, \bar{x}(t), \bar{u}(t), p(t)) = \min_{u \in U} \mathcal{H}(t, \bar{x}(t), u, p(t))\,, \quad a.e.\ t \in I. \tag{3.3}$$

and the second order integral condition

$$\begin{aligned} &\frac{1}{2}\sum_{i \in I_g \cup \{0\}} \lambda_i \langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle + \frac{1}{2}\sum_{j=1}^q \int_I \langle \varphi_j''(\bar{x}(s))\bar{y}(s), \bar{y}(s)\rangle \psi_j(ds) \\ &+ \frac{1}{2}\int_0^1 \langle \mathcal{H}''_{(x,u)}(s, \bar{x}(s), \bar{u}(s), p(s))(\bar{y}(s), \bar{v}(s)), (\bar{y}(s), \bar{v}(s))\rangle ds + \int_0^1 \mathcal{H}_u[s]w(s)ds \ge 0 \end{aligned} \tag{3.4}$$

for any $w \in \mathcal{W}$, where

$$\mathcal{H}_u[t] := \mathcal{H}_u(t, \bar{x}(t), \bar{u}(t), p(t)).$$

Remark 3.2. The function p defined in (3.2) is a solution in the sense of [22, p. 486] of the adjoint equation

$$\begin{cases} dp(t) = -\big(\mathbb{A}^* + f_x(t, \bar{x}(t), \bar{u}(t))^*\big)p(t)\mathrm{d}t - \sum_{j=1}^q \nabla\varphi_j(\bar{x}(t))\psi_j(dt)\,, \quad t \in I \\ p(1) = \sum_{i \in I_g \cup \{0\}} \lambda_i \nabla g_i(\bar{x}(1))\,. \end{cases} \tag{3.5}$$

Let us underline that the integrand in (3.2) is continuous on I, see [22, pp. 217, 218]. Therefore its Bochner integral is well defined even though X^*, in general, is not separable, see [19, Theorems 2, pp. 42, 45]. It is well known that, in general, multipliers λ_i and ψ_j appearing in (3.2) are not uniquely defined and for this reason the adjoint variable $p(\cdot)$ may be not unique. Clearly to each set of multipliers λ_i and ψ_j corresponds at most one $p(\cdot)$ defined by (3.2).

Remark 3.3. If X is reflexive, then A^* generates a strongly continuous semigroup. Hence p is the mild solution of the adjoint measure-driven equation (3.5), see [22, Theorems 5.2.6, 5.5.10]. In section 4, we provide some examples of adjoint equations derived from concrete models involving PDEs.

Remark 3.4. The expression in the left-hand side of (3.4) is the sum of a quadratic form involving second order derivatives of Hamiltonian and functions describing the cost and the constraints and a linear term involving the first order derivatives of Hamiltonian applied to second order tangents. In the finite dimensional framework, when U is described by inequalities, this expression was rewritten as a quadratic form with the Hamiltonian replaced by the augmented Hamiltonian, while skipping the linear term, see [29, Section 5.2].

In general, the obtained minimum principle can be abnormal, namely it may happen that it does not depend on the cost functional g_0. Conditions guaranteeing the validity of the above necessary optimality conditions in normal form, i.e. with $\lambda_0 = 1$, will be provided in subsection 3.3.

Remark 3.5. We would like to underline that Theorem 3.1 contains the celebrated minimum principle. Indeed let $(\bar{x}, \bar{u})$ be a local minimizer of (1.4). Then $(\bar{y}, \bar{v}) = 0$ is critical and $T_U^{\flat(2)}(\bar{u}(t), 0) = T_U^{\flat}(\bar{u}(t)) \ni 0$, $I_g = \{i \in \{1, \dots, k\} : \bar{x}(1) \in \partial Q_i\}$ and (3.1) is satisfied. By Theorem 3.1 the minimum principle (3.3) holds true with p as in (3.2). In this case the second order condition (3.4) does not bring an additional information because

$$f_u[t] T_U^{\flat}(\bar{u}(t)) \subset \overline{\bigcup_{\lambda > 0} \lambda \left(\overline{co}\, f(t, \bar{x}(t), U) - f[t] \right)}.$$

Hence for any measurable selection $w(t) \in T_U^{\flat(2)}(\bar{u}(t), 0)$ a.e. such that $f_u[\cdot]w(\cdot) \in L^1(I, X)$ we have

$$\int_0^1 \langle p(s), f_u[s]w(s) \rangle ds \geq \int_0^1 \inf_{u \in U,\, \lambda > 0} \lambda \langle p(s), f(s, \bar{x}(s), u) - f[s] \rangle ds.$$

This and (3.3) yield

$$\int_0^1 \langle p(s), f_u[s]w(s) \rangle ds \geq 0$$

implying (3.4).

The proof of Theorem 3.1 is performed in three main steps. At first in Proposition 3.6 we prove a second order optimality condition (3.7) in the form of nonintersection of second order variations of constraints with second order negative variations of the cost functional. In Lemma 3.7, using a separation theorem, we get an abstract second order condition involving multipliers λ_i and ψ_j (see (3.10)). Finally, in subsection 3.2, we introduce a suitable family of second order variations $\tilde{\mathcal{S}}^2 \subset \overline{\mathcal{S}^2}$ that yields the minimum principle (3.3) and the second order optimality condition.

3.1. **An abstract second order condition.** Fix a local minimizer $(\bar{x}, \bar{u})$ for problem (1.4) and a critical $\bar{y} \in \mathcal{V}^1(\bar{x})$. Consider the sets:

$$\mathcal{S}^2 = \mathcal{S}^2(\bar{x}, \bar{y}),$$

$$\mathcal{Q}_i^2 = \Big\{ z \in \mathcal{C}(I,X) : \langle \nabla g_i(\bar{x}(1)), z(1) \rangle + \frac{1}{2} \langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1) \rangle < 0 \Big\}, \quad i \in I_g \cup \{0\},$$

$$\mathcal{K}_j^2 = \Big\{ z \in \mathcal{C}(I,X) : \max_{t \in \mathcal{M}_{j0}} \Big(\langle \nabla \varphi_j(\bar{x}(t)), z(t) \rangle + \frac{1}{2} \langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t) \rangle \Big) < 0 \Big\}, \quad j = 1, \dots q.$$

Since we use the convention $\max \emptyset = -\infty$, if $\mathcal{M}_{j0} = \emptyset$ then $\mathcal{K}_j^2 = \mathcal{C}(I,X)$. Remark that $\mathcal{Q}_i^2$ and $\mathcal{K}_j^2$ are open convex subsets of $\mathcal{C}(I,X)$. Moreover, $\mathcal{K}_j^2$ is empty if and only if $\nabla \varphi_j(\bar{x}(t)) = 0$ and $\langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t) \rangle \geq 0$ for some $t \in \mathcal{M}_{j0}$. In particular, if (H)(vii) holds true, then $\mathcal{K}_j^2$ is not empty for all $j = 1, \dots q$.

Proposition 3.6. *Assume (H)(i),(ii),(iii),(vi) and* (3.1)*. Then,*

$$\mathcal{S}^2 \cap \mathcal{Q}^2 \cap \mathcal{K}^2 \subset T^{\flat(2)}_{\mathcal{S} \cap \mathcal{Q} \cap \mathcal{K}}(\bar{x}, \bar{y}), \tag{3.6}$$

where

$$\mathcal{Q}^2 = \bigcap_{i \in I_g} \mathcal{Q}_i^2, \quad \mathcal{K}^2 = \bigcap_{j=1}^{q} \mathcal{K}_j^2,$$

$\mathcal{S}$ *is as in definition 2.1, and* $\mathcal{Q}, \mathcal{K}$ *are defined by*

$$\mathcal{Q} = \big\{ x \in \mathcal{C}(I,X) : x(1) \in Q \big\}, \qquad \mathcal{K} = \big\{ x \in \mathcal{C}(I,X) : x(t) \in K, \, \forall t \in I \big\}.$$

Moreover,

$$\mathcal{S}^2 \cap \mathcal{Q}_0^2 \cap \mathcal{Q}^2 \cap \mathcal{K}^2 = \emptyset. \tag{3.7}$$

Observe that (3.7) is equivalent to a more familiar second order necessary condition in the form of a variational inequality: for any $z \in \mathcal{S}^2 \cap \mathcal{Q}^2 \cap \mathcal{K}^2$

$$\langle \nabla g_0(\bar{x}(1)), z(1) \rangle + \frac{1}{2} \langle g_0''(\bar{x}(1))\bar{y}(1), \bar{y}(1) \rangle \geq 0.$$

Proof. Let $z \in \mathcal{S}^2 \cap \mathcal{Q}^2 \cap \mathcal{K}^2$ and let $h_n \to 0^+$ as $n \to +\infty$. By Lemma 2.9, there exists a sequence $x_n = \bar{x} + h_n \bar{y} + h_n^2 z_n \in \mathcal{S}$, such that $z_n \to z$ uniformly. We need to show that, for any n large enough, x_n satisfies the constraints (1.2) and (1.3). Let $\delta > 0$ be as in (3.1) and fix $j \in \{1, \dots, q\}$. Since $z \in \mathcal{K}^2$, by the regularity of φ_j we deduce the existence of $0 < \tilde{\delta} < \delta$ such that

$$\max_{t \in \mathcal{M}_{j\tilde{\delta}}} \Big(\langle \nabla \varphi_j(\bar{x}(t)), z(t) \rangle + \frac{1}{2} \langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t) \rangle \Big) < 0. \tag{3.8}$$

Applying the Taylor's formula with integral rest (see for instance [1, p. 29]), we have

$$\varphi_j(x_n(t)) = \varphi_j(\bar{x}(t)) + h_n \langle \nabla \varphi_j(\bar{x}(t)), \bar{y}(t) \rangle + h_n^2 \langle \nabla \varphi_j(\bar{x}(t)), z(t) \rangle$$
$$+ h_n^2 \frac{1}{2} \langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t) \rangle + h_n^2 r_n(t),$$

where

$$r_n(t) = \langle \nabla\varphi_j(\bar{x}(t)), z_n(t) - z(t)\rangle + h_n\langle \varphi_j''(\bar{x}(t))\bar{y}(t), z_n(t)\rangle + h_n^2\frac{1}{2}\langle \varphi_j''(\bar{x}(t))z_n(t), z_n(t)\rangle$$

$$+ \int_0^1 (1-\lambda)\left\langle \left[\varphi_j''(\bar{x}(t) + \lambda(h_n\bar{y}(t) + h_n^2 z_n(t))) - \varphi_j''(\bar{x}(t))\right](\bar{y}(t) + h_n z_n(t)), \bar{y}(t) + h_n z_n(t)\right\rangle d\lambda\,.$$

Since $z_n \to z$ uniformly, φ_j'' is continuous and $\bar{x}(I)$ is compact, it follows that $r_n(\cdot) \to 0$ uniformly as $n \to \infty$. Hence, by (3.1) and (3.8), for any n large enough and any $t \in \mathcal{M}_{j\tilde{\delta}}$, we have

$$\varphi_j(x_n(t)) \le h_n^2\Big(\langle \nabla\varphi_j(\bar{x}(t)), z(t)\rangle + \frac{1}{2}\langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle + r_n(t)\Big) < 0\,.$$

Thus, for sufficiently large n, $x_n(t) \in K_j$, for all $t \in \mathcal{M}_{j\tilde{\delta}}$. If $t \notin \mathcal{M}_{j\tilde{\delta}}$, we have $d_{\partial K_j}(\bar{x}(t)) > \tilde{\delta}$, and we obtain again, for all large n and all $t \notin \mathcal{M}_{j\tilde{\delta}}$, $x_n(t) \in K_j$.

We now consider the end-point constraint Q. If $i \in \{1,\dots,k\} \setminus I_g$, then either $\bar{x}(1) \in \mathrm{int}(Q_i)$ or

$$\bar{x}(1) \in \partial Q_i \quad \text{and} \quad \langle \nabla g_i(\bar{x}(1)), \bar{y}(1)\rangle < 0\,.$$

Thus, for n large enough,

$$g_i(x_n(1)) = g_i(\bar{x}(1)) + h_n\langle \nabla g_i(\bar{x}(1)), \bar{y}(1)\rangle + o(h_n) < 0\,.$$

On the other hand, if $i \in I_g$, by the Taylor's formula for g_i at $\bar{x}(1)$ (see for instance [7, Theorem 5.6.3]) and the definition of $\mathcal{Q}_i^2$, we obtain

$$\begin{aligned} g_i(x_n(1)) &= g_i(\bar{x}(1)) + h_n\langle \nabla g_i(\bar{x}(1)), \bar{y}(1)\rangle + h_n^2\langle \nabla g_i(\bar{x}(1)), z_n(1)\rangle \\ &\quad + h_n^2\frac{1}{2}\langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle + o(h_n^2) \\ &= h_n^2\Big(\langle \nabla g_i(\bar{x}(1)), z_n(1)\rangle + \frac{1}{2}\langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle + o(1)\Big) < 0 \end{aligned}$$

for n large enough, implying that $x_n(1) \in Q$. We can conclude that $x_n \in T^{\flat(2)}_{\mathcal{S}\cap\mathcal{Q}\cap\mathcal{K}}(\bar{x},\bar{y})$ and that (3.6) follows.

To prove (3.7), suppose by contradiction that there exists

$$z \in \mathcal{S}^2 \cap \mathcal{Q}_0^2 \cap \mathcal{Q}^2 \cap \mathcal{K}^2\,.$$

By the first part of the proof, given $h_n \to 0^+$ there exists a sequence $x_n = \bar{x} + h_n\bar{y} + h_n^2 z_n$ of solutions of (1.1)–(1.3) such that $z_n \to z$ uniformly. Reasoning as above, replacing g_i by g_0, for n sufficiently large, we obtain the existence of a solution $\tilde{x} := x_n$ of (1.1)–(1.3) such that

$$g_0(\tilde{x}(1)) < g_0(\bar{x}(1))\,,$$

in contradiction with the optimality of $\bar{x}$. □

The second part of Proposition 3.6 provides a formulation of the second order optimality condition as the nonintersection of the convex set $\mathcal{Q}_0^2 \cap \mathcal{Q}^2 \cap \mathcal{K}^2$ with $\mathcal{S}^2$. We shall analyze (3.7) using the duality arguments for convex subsets of $\mathcal{S}^2$. In what follows, for any $z^* \in \mathcal{C}(I,X)^*$ and $Q \subset \mathcal{C}(I,X)$ we use the notation $\inf\langle z^*, Q\rangle = \inf_{z\in Q}\langle z^*, z\rangle$.

Lemma 3.7. *Assume (H)(i),(ii),(iii),(vi),(vii) and let $\bar{y} \in \mathcal{V}^1(\bar{x})$ be critical and satisfy (3.1). Then, for every convex nonempty subset $\tilde{\mathcal{S}}^2 \subset \overline{\mathcal{S}^2}$ there exist $\lambda_0 \in \{0,1\}$, $\lambda_i \geq 0$, for $i \in I_g$, positive $\psi_j \in \mathcal{M}(I,\mathbb{R})$ with $\operatorname{supp}\psi_j \subset \mathcal{M}_{j0}$, for $j = 1,\dots,q$, not vanishing simultaneously, and $x^* \in \mathcal{C}(I,X)^*$, such that*

(3.9)

$$\sum_{i\in I_g\cup\{0\}} \lambda_i \langle \nabla g_i(\bar{x}(1)), z(1)\rangle + \sum_{j=1}^{q} \int_I \langle \nabla \varphi_j(\bar{x}(t)), z(t)\rangle \psi_j(dt) = \langle x^*, z\rangle\,, \quad \forall\, z \in \mathcal{C}(I,X)$$

and

(3.10)

$$\inf\langle x^*, \tilde{\mathcal{S}}^2\rangle + \frac{1}{2}\sum_{i\in I_g\cup\{0\}} \lambda_i \langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle + \frac{1}{2}\sum_{j=1}^{q} \int_I \langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle \psi_j(dt) \geq 0\,.$$

Moreover, if

$$\tilde{\mathcal{S}}^2 \cap \mathcal{Q}^2 \cap \mathcal{K}^2 \neq \emptyset\,,$$

then the above necessary optimality conditions hold in normal form, i.e. with $\lambda_0 = 1$.

Proof. Notice that assumption (H)(vii) implies that $\mathcal{K}_j^2 \neq \emptyset$ for any j. Consider two different cases:

Case 1: $\mathcal{Q}_i^2 \neq \emptyset$, for any $i \in I_g \cup \{0\}$. By Proposition 3.6, we have $\mathcal{S}^2 \cap \mathcal{Q}_0^2 \cap \mathcal{Q}^2 \cap \mathcal{K}^2 = \emptyset$. As in [29, Lemma 2.2], define the open convex set

$$\mathcal{A} = \bigcup_{x\in\mathcal{S}^2} \Big[\prod_{i\in I_g\cup\{0\}} \big(\mathcal{Q}_i^2 - x\big) \times \prod_{j=1}^{q} \big(\mathcal{K}_i^2 - x\big)\Big]\,.$$

Since $0 \notin \mathcal{A}$, by a separation theorem, there exist $x^*, x_i^*, y_j^* \in \mathcal{C}(I,X)^*, i \in I_g \cup \{0\}, j = 1,\dots,q$ not vanishing simultaneously such that

$$\sum_{i\in I_g\cup\{0\}} \langle x_i^*, x_i - x\rangle + \sum_{j=1}^{q} \langle y_j^*, y_j - x\rangle \geq 0\,, \quad \forall\, x \in \mathcal{S}^2, x_i \in \mathcal{Q}_i^2, y_j \in \mathcal{K}_i^2, i \in I_g\cup\{0\}, j = 1,\dots,q\,.$$

Setting $x_i^* = 0$ for any $i \notin I_g \cup \{0\}$ and

$$-x^* = x_0^* + x_1^* + \cdots + x_k^* + y_1^* + \cdots + y_q^*\,,$$

we obtain

$$\inf\langle x^*, \tilde{\mathcal{S}}^2\rangle + \inf\langle x_0^*, \mathcal{Q}_0^2\rangle + \cdots + \inf\langle x_k^*, \mathcal{Q}_k^2\rangle + \inf\langle y_1^*, \mathcal{K}_1^2\rangle + \cdots + \inf\langle y_q^*, \mathcal{K}_q^2\rangle \geq 0\,. \tag{3.11}$$

Let $i \in I_g \cup \{0\}$ be such that $x_i^* \neq 0$. Inequality (3.11) yields

$$\inf\big\{\langle x_i^*, z\rangle : z \in \mathcal{Q}_i^2\big\} > -\infty\,, \tag{3.12}$$

implying that $\nabla g_i(\bar{x}(1)) \neq 0$. Otherwise, since $\mathcal{Q}_i^2 \neq \emptyset$, we get $\langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle < 0$, yielding $\mathcal{Q}_i^2 = X$, in contradiction with (3.12). We claim that

$$x_i^* = -\lambda_i \Gamma^* \nabla g_i(\bar{x}(1)) \quad \text{for some positive } \lambda_i\,. \tag{3.13}$$

Let $\mathcal{A} = \{z \in \mathcal{C}(I,X) : \langle \nabla g_i(\bar{x}(1)), z(1)\rangle \le 0\}$, we notice that $\mathcal{Q}_i^2 + \mathcal{A} = \mathcal{Q}_i^2$ and that $-x_i^* \in \mathcal{A}^-$. Indeed, by contradiction, suppose that there exists $\bar{q} \in \mathcal{A}$ satisfying $\langle x_i^*, \bar{q}\rangle < 0$. Taking any $\bar{z} \in \mathcal{Q}_i^2$, we obtain that

$$\inf_{z\in\mathcal{Q}_i^2} \langle x_i^*, z\rangle = \inf_{z\in\mathcal{Q}_i^2, q\in\mathcal{A}} \langle x_i^*, z+q\rangle \le \langle x_i^*, \bar{z}\rangle + \langle x_i^*, \lambda\bar{q}\rangle , \qquad \text{for any } \lambda \ge 0 .$$

Letting $\lambda \to +\infty$, we get $\inf_{z\in\mathcal{Q}_i^2}\langle x_i^*, z\rangle = -\infty$. Lemma 2.10 allows to deduce (3.13).

Further, by the definition of $\mathcal{Q}_i^2$,

$$\inf\langle -\lambda_i \Gamma^* \nabla g_i(\bar{x}(1)), \mathcal{Q}_i^2\rangle = \frac{1}{2}\lambda_i \langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle .$$

Concerning $\mathcal{K}_j^2$, again, we can consider the case $y_j^* \neq 0$, that implies $\mathcal{M}_{j0} \neq \emptyset$. By assumption, $\nabla\varphi_j(\bar{x}(t)) \neq 0$ for any $t \in \mathcal{M}_{j0}$. Define

$$\mathcal{A} = \left\{ z \in \mathcal{C}(I,X) : \max_{t\in\mathcal{M}_{j0}} \langle \nabla\varphi_j(\bar{x}(t)), z(t)\rangle \le 0 \right\} .$$

Then $\mathcal{K}_j^2 + \mathcal{A} = \mathcal{K}_j^2$ and, as above, we deduce that, $-y_j^* \in \mathcal{A}^-$. By Lemma 2.11 applied with $\xi(\cdot) = \nabla\varphi_j(\bar{x}(\cdot))$, there exists a positive $\psi_j \in \mathcal{M}(I,\mathbb{R})$ such that $\operatorname{supp}\psi_j \subset \mathcal{M}_{j0}$ and

$$\int_I \langle z(t), y_j^*(dt)\rangle = -\int_I \langle \nabla\varphi_j(\bar{x}(t)), z(t)\rangle \psi_j(dt), \;\; \text{for any } z \in \mathcal{C}(I,X) .$$

By the definition of $\mathcal{K}_j^2$, we obtain immediately

$$\inf_{z\in\mathcal{K}_j^2} \langle y_j^*, z\rangle = \inf_{z\in\mathcal{K}_j^2} -\int_I \langle \nabla\varphi_j(\bar{x}(t)), z(t)\rangle\psi_j(dt) \ge \frac{1}{2}\int_I \langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle \psi_j(dt) . \tag{3.14}$$

We claim that the inequality in (3.14) is an equality. Indeed, assume by contradiction that, for some $\varepsilon > 0$,

$$\inf_{z\in\mathcal{K}_j^2} \int_I -\langle \nabla\varphi_j(\bar{x}(t)), z(t)\rangle\psi_j(dt) = \frac{1}{2}\int_I \langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle \psi_j(dt) + \varepsilon . \tag{3.15}$$

Consider the set-valued map $R : M_{j0} \rightsquigarrow X$ defined by

$$R(t) = \left\{ \alpha \in X \;:\; \langle \nabla\varphi_j(\bar{x}(t)), \alpha\rangle + \frac{1}{2}\langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle \in \left[-\frac{\varepsilon}{2\int_I d\psi_j(s)}, -\frac{\varepsilon}{3\int_I d\psi_j(s)} \right] \right\} .$$

It is not difficult to show that R is lower semicontinuous with nonempty closed convex values. By Michael's Theorem (see [2, Theorem 9.1.2]) there exists a continuous selection $\bar{z}$ of R. Therefore $\bar{z} \in \mathcal{K}_j^2$ and

$$-\int_I \langle \nabla\varphi_j(\bar{x}(t)), \bar{z}(t)\rangle\psi_j(dt) \le \frac{1}{2}\int_I \langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle \psi_j(dt) + \frac{\varepsilon}{2} ,$$

in contradiction with (3.15). Hence, (3.9) and (3.10) follow.

Case 2: $\mathcal{Q}_i^2 = \emptyset$, for some i, implying $\nabla g_i(\bar{x}(1)) = 0$ and $\frac{1}{2}\langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle \ge 0$. Then, taking $\lambda_i = 1$, $\lambda_s = 0$ for any $s \neq i$, $\psi_j = 0$, for $j = 1, \ldots, q$, and $x^* = 0$, we deduce again (3.9) and (3.10).

If $\lambda_0 > 0$, then dividing all the multipliers λ_i, ψ_j by λ_0 we may assume that $\lambda_0 = 1$.

Normality. If $\lambda_0 = 0$, then $x_0^* = 0$ and $-x^* = x_1^* + \cdots + x_k^* + y_1^* + \cdots + y_q^*$. By (3.11), we have

$$\inf\langle x^*, \tilde{\mathcal{S}}^2\rangle + \sum_{i\in I_g} \inf\langle x_i^*, \mathcal{Q}_i^2\rangle + \sum_{j=1}^{q} \inf\langle y_j^*, \mathcal{K}_j^2\rangle \geq 0\,.$$

Applying again [29, Lemma 2.2], (a necessary and sufficient condition for separation of non empty convex sets), we obtain the contradictory

$$\tilde{\mathcal{S}}^2 \cap \Big(\bigcap_{i\in I_g} \mathcal{Q}_i^2\Big) \cap \mathcal{K}^2 = \emptyset\,,$$

ending the proof. □

3.2. **Proof of the minimum principle.** In this subsection we introduce a suitable non-empty convex set $\tilde{\mathcal{S}}^2 \subset \overline{\mathcal{S}^2}$ that yields the minimum principle.

As in the statement of Theorem 3.1, assume (H) and let $(\bar{x}, \bar{u})$ be a local minimizer of (1.4) and $(\bar{y}, \bar{v})$ be a solution of (2.10)–(2.11) satisfying (2.12). From [2, Chapter 8] it can be easily deduced that the set-valued maps $t \rightsquigarrow T_U^\flat(\bar{u}(t))$ and $t \rightsquigarrow T_U^{\flat(2)}(\bar{u}(t), \bar{v}(t))$ are measurable. Consider the second-order linearization

$$\dot{z}(t) = \mathbb{A}z(t) + f_x[t]z(t) + f_u[t]w(t) + \frac{1}{2}f''[t](\bar{y}(t), \bar{v}(t))^2 + \eta(t)\,, \quad z(0) = 0\,, \tag{3.16}$$

where $w : I \to Z$ is a Lebesgue measurable function satisfying

$$f_u[\cdot]w(\cdot) \in L^1(I, X)\,, \qquad w(t) \in T_U^{\flat(2)}(\bar{u}(t), \bar{v}(t))\,, \quad \text{a.e. in } I\,, \tag{3.17}$$

and

$$\eta \in L^1(I, X)\,, \qquad \eta(t) \in T_{F(t,\bar{x}(t))}(f^{\bar{x}}(t))\,, \quad \text{a.e. in } I\,. \tag{3.18}$$

Assume that there exists w satisfying (3.17). Then the set

$$\mathcal{L}^2 = \big\{z \in \mathcal{C}(I, X) : z \text{ solves (3.16) for some } w, \eta \text{ as in (3.17), (3.18)}\big\}$$

is nonempty. The following Lemma 3.8 describes a convenient convex set $\tilde{\mathcal{S}}^2$.

Lemma 3.8. *Assume (H) and that there exists w satisfying* (3.17)*. Then, the set*

$$\tilde{\mathcal{S}}^2 := \overline{\mathcal{L}^2} \tag{3.19}$$

is a non-empty convex subset of $\overline{\mathcal{S}^2}$.

Proof. We first prove that

$$\mathrm{co}\mathcal{L}^2 \subset \tilde{\mathcal{S}}^2\,,$$

implying $\mathrm{co}\tilde{\mathcal{S}}^2 = \tilde{\mathcal{S}}^2$. Pick $z_1, z_2 \in \mathcal{L}^2$, corresponding to the controls w_1, η_1 and w_2, η_2 respectively, and $\lambda \in (0,1)$. We aim to prove that $z_\lambda = \lambda z_1 + (1-\lambda)z_2 \in \tilde{\mathcal{S}}^2$. The function z_λ is a mild solution of the differential inclusion

$$\dot{z}(t) \in \mathbb{A}z(t) + f_x[t]z(t) + f_u[t][w_1(t), w_2(t)] + \frac{1}{2}f''[t](\bar{y}(t), \bar{v}(t))^2 + \eta_\lambda(t)\,,$$

where $[w_1(t), w_2(t)]$ is the segment joining $w_1(t)$ and $w_2(t)$, and $\eta_\lambda = \lambda\eta_1 + (1-\lambda)\eta_2$. Since $T_{F(t,\bar{x}(t))}(f^{\bar{x}}(t))$ is convex, $\eta_\lambda(t) \in T_{F(t,\bar{x}(t))}(f^{\bar{x}}(t))$ a.e. in I. By a relaxation theorem [23, Theorem 2.5], there exists a sequence of solutions z_n to

$$\dot{z}(t) \in \mathbb{A}z(t) + f_x[t]z(t) + f_u[t]\{w_1(t), w_2(t)\} + \frac{1}{2}f''[t](\bar{y}(t), \bar{v}(t))^2 + \eta_\lambda(t)\,,$$

such that $z_n \to z_\lambda$ uniformly in I. Since the set-valued mapping $t \rightsquigarrow T_U^{\flat(2)}(\bar{u}(t), \bar{v}(t))$ is measurable, applying a measurable selection theorem, we deduce that, for any n, there exists a measurable selection $w_n(t) \in \{w_1(t), w_2(t)\} \subset T_U^{\flat(2)}(\bar{u}(t), \bar{v}(t))$ such that $f_u(\cdot, \bar{x}(\cdot), \bar{u}(\cdot))w_n(\cdot) \in L^1(I, X)$ and z_n solves

$$\dot{z}(t) = \mathbb{A}z(t) + f_x[t]z(t) + f_u[t]w_n(t) + \frac{1}{2}f''[t](\bar{y}(t), \bar{v}(t))^2 + \eta_\lambda(t)\,, \qquad z(0) = 0,$$

implying $z_n \in \mathcal{L}^2$ and $z_\lambda \in \tilde{\mathcal{S}}^2$.

To conclude the proof, it is sufficient to verify that $\mathcal{L}^2 \subset \mathcal{S}^2$. Let $z \in \mathcal{C}(I, X)$ satisfy (3.16) for some w and η as in (3.17), (3.18). In order to show that $z \in \mathcal{S}^2$, we shall prove that

$$f_x[t]z(t) + f_u[t]w(t) + \frac{1}{2}f''[t](\bar{y}(t), \bar{v}(t))^2 + \eta(t) \in d_x^2F[t]z(t)\,, \quad \text{for a.e. } t \in I\,, \tag{3.20}$$

where $F[t] = F(t, \bar{x}(t), f^{\bar{x}}(t), \bar{y}(t), f_x[t]\bar{y}(t) + f_u[t]\bar{v}(t))$. The set-valued map $F(t,\cdot)$ being locally Lipschitz continuous with convex values, with arguments analogous to [24, Proposition 2.11] it is possible to verify that

$$d_x^2F[t]z(t) + T_{F(t,\bar{x}(t))}(f^{\bar{x}}(t)) \subset d_x^2F[t]z(t)\,.$$

Since $\eta(t) \in T_{F(t,\bar{x}(t))}(f^{\bar{x}}(t))$ for a.e. $t \in I$, to get (3.20) it is enough to show that

$$b(t) := f_x[t]z(t) + f_u[t]w(t) + \frac{1}{2}f''[t](\bar{y}(t), \bar{v}(t))^2 \in d_x^2F[t]z(t)\,, \quad \text{for a.e. } t \in I\,. \tag{3.21}$$

Let $t \in I$ be such that $w(t) \in T_U^{\flat(2)}(\bar{u}(t), \bar{v}(t))$. Then for every $h > 0$, there exists $r(h) \in Z$ such that $\|r(h)\|_Z = o(h^2)$ and $\bar{u}(t) + h\bar{v}(t) + h^2w(t) + r(h) \in U$. In particular,

$$\begin{aligned} &\operatorname{dist}\Big(b(t), \frac{F(t, (\bar{x} + h\bar{y} + h^2z)(t)) - (f^{\bar{x}} + hf_x[\cdot]\bar{y} + hf_u[\cdot]\bar{v})(t)}{h^2}\Big) \le \\ &\Big\|b(t) - \frac{f(t, (\bar{x} + h\bar{y} + h^2z)(t), (\bar{u} + h\bar{v} + h^2w)(t) + r(h)) - (f^{\bar{x}} + hf_x[\cdot]\bar{y} + hf_u[\cdot]\bar{v})(t)}{h^2}\Big\|_X. \end{aligned} \tag{3.22}$$

Applying the second order Taylor's formula for f at $[t]$ (see for instance [7, Theorem 5.6.3]), we have that the right hand side of the above inequality converges to zero as $h \to 0$. Therefore we obtain (3.21) and conclude the proof. $\square$

Proof of Theorem 3.1. Let $\tilde{\mathcal{S}}^2$ be defined as in (3.19) and let λ_i, for $i \in I_g \cup \{0\}$, ψ_j, for $j = 1, \ldots, q$, and $x^* \in \mathcal{C}(I, X)^*$ be as in Lemma 3.7. Let z be a mild solution to (3.16) corresponding to some controls w and η satisfying (3.17) and (3.18). Then,

$$z(t) = \int_0^t \mathcal{T}(t, s)(\beta(s) + \eta(s))ds\,,$$

where

$$\beta(s) = f_u[s]w(s) + \frac{1}{2}f''[s](\bar{y}(s), \bar{v}(s))^2.$$

Then, from (3.9) and [21, Lemma 4.1], we obtain that

$$\begin{aligned}
\langle x^*, z\rangle &= \int_0^1 \Big\langle \sum_{i\in I_g\cup\{0\}} \lambda_i \nabla g_i(\bar{x}(1)), \mathcal{T}(1,s)(\beta(s)+\eta(s))\Big\rangle ds \\
&\qquad + \int_0^1 \sum_{j=1}^q \Big\langle \nabla\varphi_j(\bar{x}(t)), \int_0^t \mathcal{T}(t,s)(\beta(s)+\eta(s))ds\Big\rangle \psi_j(dt) \\
&= \int_0^1 \Big\langle \mathcal{T}(1,s)^*\Big(\sum_{i\in I_g\cup\{0\}} \lambda_i \nabla g_i(\bar{x}(1))\Big), \beta(s)+\eta(s)\Big\rangle ds \\
&\qquad + \int_0^1 \Big\langle \int_s^1 \mathcal{T}(t,s)^*\Big(\sum_{j=1}^q \nabla\varphi_j(\bar{x}(t))\psi_j(dt)\Big), \beta(s)+\eta(s)\Big\rangle ds \\
&= \int_0^1 \langle p(s), \beta(s)+\eta(s)\rangle ds,
\end{aligned}$$

see [28] for details. Plugging this equality in (3.10) we derive the second order condition

$$\begin{aligned}
&\frac{1}{2}\sum_{i\in I_g\cup\{0\}} \lambda_i \langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle + \frac{1}{2}\sum_{j=1}^q \int_I \langle \varphi_j''(\bar{x}(s))\bar{y}(s), \bar{y}(s)\rangle \psi_j(ds) \\
&\quad + \frac{1}{2}\int_0^1 \langle p(s), f''[s](\bar{y}(s), \bar{v}(s))^2\rangle ds + \inf_{\{w \text{ as in } (3.17)\}} \int_0^1 \langle p(s), f_u[s]w(s)\rangle ds \\
&\quad + \inf_{\{\eta \text{ as in } (3.18)\}} \int_0^1 \langle p(s), \eta(s)\rangle ds \geq 0,
\end{aligned}$$

yielding (3.4) and

$$\int_0^1 \langle p(s), \eta(s)\rangle ds \geq 0, \qquad \text{for any } \eta \text{ as in } (3.18)\,, \tag{3.23}$$

because the set of functions η satisfying (3.18) is a cone. From (3.23) we obtain the strong minimum principle (3.3), in the same way as in [28, Theorem 4.2]. □

3.3. **Normality.** In this subsection, we show that an appropriate inward pointing condition for the optimal pair $(\bar{x}, \bar{u})$ allows to formulate the optimality conditions of Theorem

3.1 in normal form. We will assume the following inward pointing condition

$$\exists\, \eta, \rho, M > 0,\ \exists J \subset I \text{ such that } \mu(J) = 1 \text{ and :} \tag{3.24}$$

$$\begin{aligned}
&(i) \text{ if } \sup_{z \in B(S(\tau_0)\bar{x}(s),\eta)} \langle \nabla\varphi_{j_0}(z), S(\tau_0) f^{\bar{x}}(s)\rangle \geq 0 \text{ for some } s \in J,\ j_0 = 1,\dots,q,\ \tau_0 \leq \eta,\\
&\quad \text{with } S(\tau_0)\bar{x}(s) \in \partial K_{j_0} + \eta B, \text{ then } \exists\, v \in T_{F(s,\bar{x}(s))}(f^{\bar{x}}(s)) \text{ satisfying } \|v\|_X \leq M \text{ and}\\
&\quad \langle \nabla\varphi_j(\bar{x}(t)), S(t-s)v\rangle \leq -\rho, \forall\, j = 1,\dots,q, \forall\, t \in [s,(s+\eta)\wedge 1] \cap \mathcal{M}_{j0}\\
&\quad \text{such that } \bar{x}(t) \in B(S(t-s)\bar{x}(s),\eta)\,;\\
&(ii) \text{ if } s \in J \cap [1-\eta, 1]\,,\ \exists\, v \in T_{F(s,\bar{x}(s))}(f^{\bar{x}}(s)) \text{ satisfying the same conditions as in } (i)\\
&\quad (\text{whenever } \sup_{z \in B(S(\tau_0)\bar{x}(s),\eta)} \langle \nabla\varphi_{j_0}(z), S(\tau_0) f^{\bar{x}}(s)\rangle \geq 0 \text{ for some } j_0 = 1,\dots,q,\ \tau_0 \leq \eta)\\
&\quad \text{and} \quad \langle \nabla g_i(\bar{x}(1)), S(1-s)v\rangle \leq -\rho\,,\ \forall\, i \in I_g\,.
\end{aligned}$$

This is an infinite dimensional version of the classical inward pointing condition proposed by Soner in [39]. As shown in the following lemma, such condition allows to construct first order variations pointwise belonging to the interior of the tangent cone to the set of active constraints. This, in turn, allows to verify the condition for normality provided in Lemma 3.7. In the infinite dimensional setting, some uniformity on a neighborhood of the boundary of the set of constraints has to be imposed and the semigroup has to be involved in the formulation of the inward pointing condition (3.24). However, if the data of the problem enjoy some compactness, the condition can be simplified in a similar way as done in [27].

Lemma 3.9. *Let $t_1 \in I$ and $y_1 \in X$. Assume that K is positively invariant by the semigroup $S(\cdot)$, i.e.*

$$S(t)K \subset K, \qquad \forall\, t \in I\,, \tag{3.25}$$

and (3.24). If $t_1 = 0$ suppose that $\bar{x}(0) \notin \partial K$. If $t_1 > 0$ and $t_1 \in \mathcal{M}_{j0}$, suppose in addition that $\langle \nabla\varphi_j(\bar{x}(t_1)), y_1\rangle < 0$. Then there exist $\delta > 0$, independent from t_1 and y_1, and a solution y of

$$\dot{y}(t) = \mathbb{A}y(t) + f_x[t]y(t) + v(t)\,, \qquad y(t_1) = y_1, \tag{3.26}$$

for some

$$\textit{integrable selection } v(\cdot) \textit{ of the set-valued map } T_{F(\cdot,\bar{x}(\cdot))}(f^{\bar{x}}(\cdot))\,, \tag{3.27}$$

satisfying

$$\langle \nabla\varphi_j(\bar{x}(t)), y(t)\rangle < 0 \quad \forall\, j = 1,\dots,q,\ \forall\, t \in [t_1,(t_1+\delta)\wedge 1] \cap \mathcal{M}_{j0} \tag{3.28}$$

and, if $1 - t_1 \leq \delta$,

$$\langle \nabla g_i(\bar{x}(1)), y(1)\rangle < 0 \quad \forall\, i \in I_g\,. \tag{3.29}$$

Proof. Let η, ρ, M, J, M_S be as in (3.24), (2.2) and define, in light of (2.3), positive reals

$$M_x := \int_I \|f_x[t]\|_{\mathbb{L}(X)}\,\mathrm{d}t < \infty\,, \qquad M_K > \max_{j=1,\dots,q\,,\,t \in \mathcal{M}_{j0}} \|\nabla\varphi_j(\bar{x}(t))\|_{X^*}\,,$$

$$M_Q > \max_{i \in I_g} \|\nabla g_i(\bar{x}(1))\|_{X^*}\,.$$

By (2.4), there exists $0<\delta<\eta$, independent from t_1 and y_1, satisfying

$$\|\bar{x}(t)-S(t-s)\bar{x}(s)\|_X \leq \eta/2\,,\ \forall\, 0\leq s<t\leq 1 \text{ such that } t-s\leq\delta \tag{3.30}$$

and

$$\int_E \|f_x[t]\|_{\mathbb{L}(X,X)}\,\mathrm{d}t \leq \frac{\rho}{2M(M_K+M_Q)M_S^2 \mathrm{e}^{M_S M_x}} \quad \forall\, E\subset I \text{ with } \ \mu(E)\leq\delta\,. \tag{3.31}$$

Set

$$\Gamma=\Big\{s\in J:\ \text{ for some } \tau_0\leq\eta \text{ and } j_0=1,\dots,q\,,\ S(\tau_0)\,\bar{x}(s)\in\partial K_{j_0}+\eta B \ \text{ and}$$
$$\sup_{\{z\in X:\|z-S(\tau_0)\bar{x}(s)\|\leq\eta\}}\langle\nabla\varphi_{j_0}(z),S(\tau_0)f[s]\rangle\geq 0\Big\}.$$

Then, by (3.24) and a measurable selection theorem, there exists a measurable function $v:I\to X$ such that $v(s)\in T_{F(s,\bar{x}(s))}(f^{\bar{x}}(s))$ a.e. $s\in I$, $v\equiv 0$ on $I\setminus\Gamma$, and, for any $s\in\Gamma$,

$$\|v(s)\|_X\leq M\,,$$

$$\langle\nabla\varphi_j(\bar{x}(t)),S(t-s)v(s)\rangle\leq-\rho \tag{3.32}$$

for every $j=1,\dots,q$ and $t\in[s,(s+\eta)\wedge 1]\cap\mathcal{M}_{j0}$ such that $\bar{x}(t)\in B(S(t-s)\bar{x}(s),\eta)$. Moreover,

$$\langle\nabla g_i(\bar{x}(1)),S(1-s)v(s)\rangle\leq-\rho\,,\ \forall\, i\in I_g\,,\ \forall\, s\in J\cap[1-\eta,1]\,.$$

For any $c\geq 0$, consider the mild solution y_c of

$$\begin{cases}\dot{y}(t)=\mathbb{A}y(t)+f_x[t]y(t)+cv(t)\\ y(t_1)=y_1\,.\end{cases} \tag{3.33}$$

By applying the Gronwall lemma, it is possible to deduce that for all $t\in[t_1,1]$,

$$\|y_c(t)\|_X\leq M_S\mathrm{e}^{M_S M_x}\|y_1\|_X+MM_S\mathrm{e}^{M_S M_x}c\mu(\Gamma\cap[t_1,t])\,. \tag{3.34}$$

We will show that, for a suitable $c\geq 0$, the solution y_c of (3.33) satisfies (3.28) and, if $1-t_1\leq\delta$, also (3.29).

In order to prove (3.28), fix $j\in\{1,\dots,q\}$.

Case 1: $t_1\notin\mathcal{M}_{j0}$. If $[t_1,(t_1+\delta)\wedge 1]\cap\mathcal{M}_{j0}=\emptyset$, then there is nothing to prove. Otherwise, let $t_2=\min\big([t_1,(t_1+\delta)\wedge 1]\cap\mathcal{M}_{j0}\big)$.

Observe that $\mu(\Gamma\cap[t_1,t_2])>0$. Indeed, suppose by contradiction that for a.e. $t\in[t_1,t_2]$ and every $0\leq\tau\leq\eta$ we have either $S(\tau)\,\bar{x}(t)\notin\partial K_j+\eta B$ or

$$\langle\nabla\varphi_j(z),S(\tau)f[t]\rangle<0\,,\quad\forall\, z\in X \text{ such that } \|z-S(\tau)\bar{x}(t)\|\leq\eta\,. \tag{3.35}$$

By the mean value theorem (see [15, Theorem 2.3.7]), there exists $z\in[\bar{x}(t_2),S(t_2-t_1)\bar{x}(t_1)]$, satisfying

$$\varphi_j(\bar{x}(t_2))=\varphi_j(S(t_2-t_1)\bar{x}(t_1))+\langle\nabla\varphi_j(z),\bar{x}(t_2)-S(t_2-t_1)\bar{x}(t_1)\rangle\,.$$

Hence, from (2.1) and (3.25) we deduce

$$0\leq\int_{t_1}^{t_2}\langle\nabla\varphi_j(z),S(t_2-t)f[t]\rangle\mathrm{d}t\,. \tag{3.36}$$

Let $t \in [t_1, t_2]$. By (3.30) we have that $S(t_2-t)\bar{x}(t) \in \partial K+\eta B$ and $\|z-S(t_2-t)\bar{x}(t)\|_X \leq \eta$. Therefore, (3.35) and (3.36) are in contradiction.

Let $t \in [t_1, (t_1+\delta) \wedge 1] \cap \mathcal{M}_{j0}$. We have $t \geq t_2$ and, by (3.31), (3.32), (3.34),

$$\begin{aligned}
\langle \nabla\varphi_j(\bar{x}(t)), y_c(t)\rangle &= \langle \nabla\varphi_j(\bar{x}(t)), S(t-t_1)y_1\rangle + \int_{t_1}^{t} \langle \nabla\varphi_j(\bar{x}(t)), S(t-s)f_x[s]y_c(s)\rangle \mathrm{d}s \\
&\quad + c\int_{\Gamma\cap[t_1,t]} \langle \nabla\varphi_j(\bar{x}(t)), S(t-s)v(s)\rangle \mathrm{d}s \\
&\leq M_K M_S \|y_1\|_X \big(1+M_S M_x \mathrm{e}^{M_S M_x}\big) \\
&\quad - c\,\mu(\Gamma\cap[t_1,t])\big(\rho - M M_K M_S^2 \mathrm{e}^{M_S M_x}\int_{t_1}^{t} \|f_x[s]\|_{L(X)}\mathrm{d}s\big) \\
&\leq M_K M_S \|y_1\|_X \big(1+M_S M_x \mathrm{e}^{M_S M_x}\big) - c\,\mu(\Gamma\cap[t_1,t_2])\,\frac{\rho}{2}\,.
\end{aligned} \tag{3.37}$$

This quantity is strictly negative for all c greater than some $c_1 > 0$, independent on t. Increasing c_1 we may assume that it is independent from j.

Case 2: $t_1 \in \mathcal{M}_{j0}$. In this case, by the hypothesis of the lemma, $t_1 > 0$ and $\langle \nabla\varphi_j(\bar{x}(t_1)), y_1\rangle < 0$. Let $t_3 \in (t_1, (t_1+\delta)\wedge 1]$ be such that

$$\langle \nabla\varphi_j(\bar{x}(t)), S(t-t_1)y_1\rangle + M_K M_S^2 \mathrm{e}^{M_S M_x}\|y_1\|_X \int_{t_1}^{t} \|f_x[s]\|_{L(X)}\mathrm{d}s < 0\,, \quad \forall\, t \in (t_1, t_3]\,.$$

If $(t_1, t_3] \cap \mathcal{M}_{j0} = \emptyset$, we can find, as in Case 1, $c_1 \geq 0$ independent from j such that for any $c \geq c_1$

$$\langle \nabla\varphi_j(\bar{x}(t)), y_c(t)\rangle < 0 \quad \forall\, t \in [t_1, (t_1+\delta)\wedge 1] \cap \mathcal{M}_{j0}\,.$$

Otherwise, let $t_2 \in (t_1, t_3] \cap \mathcal{M}_{j0}$. As in Case 1, we have $\mu(\Gamma \cap (t_1, t_2]) > 0$. For $t \in (t_1, t_2] \cap \mathcal{M}_{j0}$, we obtain

$$\begin{aligned}
\langle \nabla\varphi_j(\bar{x}(t)), y_c(t)\rangle &= \langle \nabla\varphi_j(\bar{x}(t)), S(t-t_1)y_1\rangle + \int_{t_1}^{t} \langle \nabla\varphi_j(\bar{x}(t)), S(t-s)f_x[s]y_c(s)\rangle \mathrm{d}s \\
&\quad + c\int_{\Gamma\cap[t_1,t]} \langle \nabla\varphi_j(\bar{x}(t)), S(t-s)v(s)\rangle \mathrm{d}s \\
&\leq \langle \nabla\varphi_j(\bar{x}(t)), S(t-t_1)y_1\rangle + M_K M_S^2 \mathrm{e}^{M_S M_x}\|y_1\|_X \int_{t_1}^{t} \|f_x[s]\|_{L(X)}\mathrm{d}s \\
&\quad - c\,\mu(\Gamma\cap[t_1,t])\big(\rho - M M_K M_S^2 \mathrm{e}^{M_S M_x}\int_{t_1}^{t} \|f_x[s]\|_{L(X)}\mathrm{d}s\big) < 0
\end{aligned}$$

for every $c > 0$. If $t \in (t_2, (t_1+\delta)\wedge 1] \cap \mathcal{M}_{j0}$, the same computation as in (3.37) yields $\langle \nabla\varphi_j(\bar{x}(t)), y_c(t)\rangle < 0$ for all c greater than some $c_1 > 0$, independent on t. Increasing c_1 we may assume that it is independent from j.

If $1-t_1>\delta$ the proof is complete. Otherwise, let $i\in I_g$. We have

$$\begin{aligned}\langle\nabla g_i(\bar{x}(1)),y_c(1)\rangle &= \langle\nabla g_i(\bar{x}(1)),S(1-t_1)y_1\rangle+\int_{t_1}^1\langle\nabla g_i(\bar{x}(1)),S(1-s)f_x[s]y_c(s)\rangle\mathrm{d}s\\ &\quad+c\int_{t_1}^1\langle\nabla g_i(\bar{x}(1)),S(1-s)v(s)\rangle\mathrm{d}s\\ &\le M_QM_S\|y_1\|_X\big(1+M_SM_x\mathrm{e}^{M_SM_x}\big)\\ &\quad-c\,(1-t_1)\Big(\rho-MM_QM_S^2\mathrm{e}^{M_SM_x}\int_{t_1}^1\|f_x[s]\|_{L(X)}\mathrm{d}s\Big)\\ &\le M_QM_S\|y_1\|_X\big(1+M_SM_x\mathrm{e}^{M_SM_x}\big)-c\,(1-t_1)\,\frac{\rho}{2}\,.\end{aligned}$$

This quantity is strictly negative for all c greater than some $c_2>0$, independent on t and i. To conclude, it will be sufficient to choose $c=\max\{c_1,c_2\}$. □

Theorem 3.10. *Let $(\bar{x},\bar{u})$ be an optimal pair such that $\bar{x}(0)\notin\partial K$. In addition to the hypotheses of Theorem 3.1, assume* (3.24) *and* (3.25)*. Then the optimality conditions of Theorem 3.1 hold in the normal form.*

Proof. Let $\tilde{\mathcal{S}}^2$ be defined as in Lemma 3.8. Let $\bar{z}$ be a solution of (3.16) for some $\bar{w}$ and $\bar{\eta}$ satisfying (3.17) and (3.18). By Lemma 3.9, we can construct by iteration a solution y of (3.26) with $t_1=0$ and $y_1=0$, for some integrable selection v of the set-valued map $T_{F(\cdot,\bar{x}(\cdot))}(f^{\bar{x}}(\cdot))$, satisfying

$$\langle\nabla\varphi_j(\bar{x}(t)),y(t)\rangle<0\quad\forall\,j=1,\dots,q,\ \forall\,t\in\mathcal{M}_{j0}$$

and

$$\langle\nabla g_i(\bar{x}(1)),y(1)\rangle<0\quad\forall\,i\in I_g\,.$$

For any $\alpha>0$, define $z_\alpha=\bar{z}+\alpha y$. Then z_α is the mild solution of

$$\dot{z}_\alpha(t)=\mathbb{A}z_\alpha(t)+f_x[t]z_\alpha(t)+f_u[t]\bar{w}(t)+\frac{1}{2}f''[t](\bar{y}(t),\bar{v}(t))^2+\bar{\eta}(t)+\alpha v(t)\,,\quad z_\alpha(0)=0\,.$$

Consequently, $z_\alpha\in\tilde{\mathcal{S}}^2$. Finally, since the set $\mathcal{M}_{j0}$ is compact and the function $\langle\nabla\varphi_j(\bar{x}(\cdot)),y(\cdot)\rangle$ is continuous, we can find α sufficiently large such that

$$\langle\nabla\varphi_j(\bar{x}(t)),z_\alpha(t))\rangle+\frac{1}{2}\langle\varphi_j''(\bar{x}(t))\bar{y}(t),\bar{y}(t)\rangle<0\quad\forall\,t\in\mathcal{M}_{j0}\,,\ j=1,\dots q,$$

$$\langle\nabla g_i(\bar{x}(1)),z_\alpha(1)\rangle+\frac{1}{2}\langle g_i''(\bar{x}(1))\bar{y}(1),\bar{y}(1)\rangle<0\quad\forall\,i\in I_g\,.$$

Hence $z_\alpha\in\tilde{\mathcal{S}}^2\cap\mathcal{K}^2\cap\mathcal{Q}^2$. By Lemma 3.7, this inclusion is sufficient to deduce the second order optimality condition in normal form. □

4. Applications to control problems involving PDEs

In this section we apply our results to some concrete models described by PDEs. Both a wave and a heat equation will be considered.

Example 4.1 (*The Sine-Gordon equation*)**.** We investigate an optimal control problem involving the Sine-Gordon equation, a model for the dynamics of a Josephson junction driven by a current source, see e.g. [28, 34, 40] for further details. Let Ω be a smooth and bounded domain in $\mathbb{R}^3$, $U \subset L^2(\Omega)$ be a given bounded closed non empty set, and consider the equation

$$\partial_{tt}\xi(t,\mathbf{x}) + \partial_t\xi(t,\mathbf{x}) - \Delta\xi(t,\mathbf{x}) + \sin\xi(t,\mathbf{x}) = u(t,\mathbf{x}). \tag{4.1}$$

where $\xi = \xi(t,\mathbf{x})$, $t \in I$, $\mathbf{x} \in \Omega$, is a scalar function, and $\xi(0,\mathbf{x}) = \xi_0(\mathbf{x}) \in H_0^1(\Omega)$, $\partial_t\xi(0,\mathbf{x}) = \xi_1(\mathbf{x}) \in L^2(\Omega)$ are given. The parameter $u = u(t,\cdot) \in U$ controls the external current. We impose Dirichlet boundary conditions. Below the explicit dependence on the variable $\mathbf{x}$ will be omitted. To write (4.1) as an abstract system (1.1), and implement our abstract machinery, let $X = H_0^1(\Omega) \times L^2(\Omega)$, with associated norm $\|x\|_X = \|(x_1,x_2)\|_X = \left(\|\nabla x_1\|^2_{L^2(\Omega)} + \|x_2\|^2_{L^2(\Omega)}\right)^{1/2}$, and

$$x = (x_1, x_2) = (\xi, \xi + \partial_t\xi),$$

yielding

$$\begin{cases} \dot{x}_1(t) = x_2(t) - x_1(t) \\ \dot{x}_2(t) = \Delta x_1(t) - \sin(x_1(t)) + u(t)\,. \end{cases}$$

Define the operator

$$\mathbb{A} = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix}$$

with domain $D(\mathbb{A}) = \left(H^2(\Omega) \cap H_0^1(\Omega)\right) \times H_0^1(\Omega)$, and

$$f(t,x,u) = -(x_1, \sin(x_1)) + (0,u)\,.$$

Then, (4.1) can be seen as the abstract system (1.1). Notice that, since $x_1 \in H_0^1(\Omega)$ and f takes values in $X = H_0^1(\Omega) \times L^2(\Omega)$, the required differentiability of f holds true, and $\partial_x f(t,x,u)(y_1,y_2) : \mathbf{x} \mapsto -(y_1(\mathbf{x}), \cos(x_1(\mathbf{x}))y_1(\mathbf{x}))$ for any $(y_1,y_2) \in X$ (see for instance [22, section 6.8]). The operator $\mathbb{A}$ is skew adjoint, see [22, p. 229], and generates a strongly continuous semigroup of contractions $S(t)$ in the Hilbert space X. Concerning existence of solutions, for any measurable control function u, we notice that our assumptions (2.3) and (2.4) are satisfied, yielding the validity of Hypotheses I and II of [22, section 5.5], see section 5.7 for the details. This implies local existence of a solution. Finally, (2.4) guarantees the validity of [22, Corollary 5.5.5], proving the existence of a mild solution (see definition 2.1) for our system. Given $R > 0$, we take $K = \{x \in X : \varphi(x) = \|x\|^2_X - R \leq 0\}$ as state constraint, complying with the request that the energy associated to the system is bounded, and no end-point constraint is imposed. Assumption (3.25) holds, further $\nabla\varphi(x) = 2x \neq 0$, for any $x \in \partial K$.

We take as a cost a functional $g_0 : X \to \mathbb{R}$ defined by

$$g_0(z) = \frac{1}{2}\|z - x_D\|^2_X,$$

for some fixed $x_D \in X$. Observe that assumptions (H) are satisfied by the data of the problem. Let $(\bar{x},\bar{u})$ be optimal and $(\bar{y},\bar{v})$ be a solution of the linearized system

$$\begin{cases} \dot{y}(t) = \mathbb{A}y(t) + (-y_1(t), -\cos(\bar{x}_1(t))y_1(t)) + (0, v(t)) \\ y(0) = 0\,, \end{cases}$$

where $y = (y_1, y_2) \in X$, $v(\cdot)$ is an essentially bounded measurable selection of $T_U^\flat(\bar{u}(\cdot))$. If $(\bar{y}, \bar{v})$ satisfies the assumptions of Theorem 3.1, then the solution $p = (p_1, p_2) : I \to H^{-1}(\Omega) \times L^2(\Omega)$ to the measure-driven adjoint equation

$$\begin{cases} dp(t) = \big(\mathbb{A} - \partial_x f(t, \bar{x}(t), \bar{u}(t))^*\big)p(t)dt - 2\bar{x}(t)\psi(dt)\,, \quad t \in I \\ p(1) = \lambda_0(\bar{x}(1) - x_D) \end{cases}$$

for a positive $\psi \in \mathcal{M}(I, \mathbb{R})$ and $\lambda_0 \geq 0$ not both equal to zero and with $\operatorname{supp}\psi \subset \{t \in I : \|\bar{x}(t)\|_X^2 = R\}$, satisfies the following minimality condition

$$\langle p_2(t), \bar{u}(t))\rangle = \min_{u \in U} \langle p_2(t), u\rangle\,, \qquad \text{for a.e. } t \in I\,,$$

and the second order condition:

$$\lambda_0\|\bar{y}(1)\|_X^2 + \int_I 2\|\bar{y}(s)\|_X^2\psi(ds) + \int_0^1 \big\langle p_2(s), \sin(\bar{x}_1(s))\big(\bar{y}_1(t)\big)^2\big\rangle_{L^2(\Omega)}ds$$
$$+ \inf_w 2\int_0^1 \langle p_2(s), w(s)\rangle_{L^2(\Omega)}ds \geq 0\,.$$

where the infimum is taken over functions $w(\cdot) \in L^1(I, L^2(\Omega))$ satisfying

$$w(t) \in T_U^{\flat(2)}(\bar{u}(t), \bar{v}(t))\,, \quad \text{a.e. in } I\,.$$

Further, by Theorem 3.10 the normal necessary conditions hold if $\bar{x}(0) \notin \partial K$, U is compact, and the following simplified inward pointing condition is satisfied:

$$\exists\, \rho > 0 \text{ such that, if } \|\bar{x}(s)\|_X^2 = R \text{ and}$$
$$-\|\bar{x}_1(s)\|_{H_0^1(\Omega)}^2 + \langle \bar{x}_2(s), -\sin(\bar{x}_1(s)) + w\rangle_{L^2(\Omega)} \geq 0\,, \text{ for some } s \in I \text{ and } w \in U,$$
$$\text{then, } \exists\, v \in C_U(\bar{u}(t)) \text{ satisfying } \quad \langle \bar{x}_2(s), v\rangle_{L^2(\Omega)} \leq -\rho. \tag{4.2}$$

Indeed we check next by a contradiction argument that (4.2) implies (3.24). Notice that, if $v \in C_U(\bar{u}(t))$, then $f_u[t]v \in T_{F(t,\bar{x}(t))}\big(f^{\bar{x}}(t)\big)$.

If (3.24) does not hold, then, in particular, for any $i \in \mathbb{N}$ there exist $s_i \in I$, $\tau_i \in (0, 1/i)$, $z_i \in B\Big(S(\tau_i)\bar{x}(s_i), \frac{1}{i}\Big)$ such that:

$$S(\tau_i)\bar{x}(s_i) \in \partial K + \frac{1}{i}B \qquad \text{and} \qquad \Big\langle z_i, S(\tau_i)\big(-\bar{x}_1(s_i), -\sin(\bar{x}_1(s_i)) + \bar{u}(s_i)\big)\Big\rangle \geq 0$$

but, for any $v_i \in C_U(\bar{u}(s_i))$ we can find $t_i \in [s_i, (s_i + 1/i) \wedge 1]$ such that $\bar{x}(t_i) \in \partial K$, $\bar{x}(t_i) \in B\Big(S(t_i - s_i)\bar{x}(s_i), \frac{1}{i}\Big)$ and

$$\langle \bar{x}(t_i), S(t_i - s_i)(0, v_i)\rangle \geq -\frac{1}{i}. \tag{4.3}$$

By compactness assumptions we get, taking a subsequence and keeping the same notation,

$$s_i \to s \in I, \quad \tau_i \to 0, \quad z_i \to \bar{x}(s) \in \partial K, \quad \bar{u}(s_i) \to w \in U, \qquad \text{yielding}$$
$$-\|\bar{x}_1(s)\|_{H_0^1(\Omega)}^2 + \langle \bar{x}_2(s), -\sin(\bar{x}_1(s)) + w\rangle_{L^2(\Omega)} \geq 0.$$

Hence, by (4.2), there exists $v \in C_U(w)$ such that

$$\langle \bar{x}_2(s), v\rangle_{L^2(\Omega)} \leq -\rho.$$

From the lower semicontinuity of Clarke tangent cone, there exist $v_i \in C_U(\bar{u}(s_i))$ such that $v_i \to v$. Let t_i be as in (4.3). Taking the limit and keeping the same notation, we obtain that

$$t_i \to s, \quad \bar{x}(t_i) \to \bar{x}(s), \quad \text{ and } \quad \langle \bar{x}_2(s), v\rangle_{L^2(\Omega)} \geq 0,$$

a contradiction.

Example 4.2 (*A controlled heat equation)***.** The second example involves a controlled heat equation, analyzed in [28] to get first order necessary optimality conditions in the presence of state constraints. Given $\Omega \subset \mathbb{R}^N$, a bounded domain with smooth boundary $\partial\Omega$, consider a heat equation

$$\partial_t x(t, \mathbf{x}) - \Delta x(t, \mathbf{x}) = \zeta(t, \mathbf{x}) + b(\mathbf{x})^T u(t) x(t, \mathbf{x}). \tag{4.4}$$

Here, $\zeta \in L^1(I, L^2(\Omega))$, $b \in L^\infty(\Omega, \mathbb{R}^m)$ and $x = x(t, \mathbf{x})$ is the temperature distribution, a function of the time $t \in I$ and the position $\mathbf{x} \in \Omega$. $x(0, \mathbf{x}) = x_0(\mathbf{x}) \in L^2(\Omega)$ is given and the heat supply is represented by a multiplicative control u taking values in a convex polytope U of $\mathbb{R}^m$, i.e. a compact intersection of a finite family of affine half-spaces.

Finite dimensionality of the control variable seems to be frequent for many real applications of control theory.

In this case, we have

$$T_U^\flat(u) = \begin{cases} \mathbb{R}^m & \text{if } u \in \operatorname{int} U \\ \operatorname{cone}(U - u) & \text{if } u \in \partial U \end{cases} \tag{4.5}$$

and

$$T_U^{\flat(2)}(u, v) = \begin{cases} \mathbb{R}^m & \text{if } u \in \operatorname{int} U \\ \mathbb{R}^m & \text{if } u \in \partial U \text{ and } v \in \operatorname{int} T_U^\flat(u) \\ \operatorname{cone}(T_U^\flat(u) - v) = T_U^\flat(u) + \mathbb{R}v & \text{if } u \in \partial U \text{ and } v \in \partial T_U^\flat(u)\,. \end{cases} \tag{4.6}$$

As in the previous example, below we omit writing explicitly the dependence on the variable $\mathbf{x}$. Equation (4.4) is endowed with Dirichlet boundary conditions.

To handle (4.4) as an abstract system (1.1), we define the operator $\mathbb{A} = \Delta$ with domain $D(\mathbb{A}) = H^2(\Omega) \cap H_0^1(\Omega)$. $\mathbb{A}$ generates a strongly continuous semigroup $S(t)$ of contractions on $X = L^2(\Omega)$. The forcing term $f : I \times X \times \mathbb{R}^m \to X$ is defined as $f(t, x, u) = \zeta(t) + b^T u x$.

Given a reference temperature $x_D \in X$, our aim is to minimize the functional

$$g_0(x(1)) = \frac{1}{2}\|x(1) - x_D\|_X^2$$

among all the trajectory/control pairs (x, u) satisfying the energy state constraint $K = \{x \in X : \varphi(x) = \|x\|_X^2 - 1 \leq 0\}$ and the end-point constraint $Q = \{x \in X : g_1(x) = \|x - x_1\|_X^2 - r \leq 0\}$, for some fixed $x_1 \in L^2(\Omega)$ and $r > 0$. Then assumptions (H) and (3.25) hold.

Let $(\bar{x}, \bar{u})$ be optimal and let $(\bar{y}, \bar{v})$ be a solution of the linearized system

$$\begin{cases} \dot{y}(t) = \mathbb{A}y(t) + b^T \bar{u}(t) y(t) + b^T v(t) \bar{x}(t) \\ y(0) = 0\,, \end{cases}$$

such that $\bar{v}(\cdot)$ is an essentially bounded measurable selection of $T_U^\flat(\bar{u}(\cdot))$. Remark that (2.12) holds. Assume that $\bar{y}$ satisfies the following conditions:

$$\exists\,\delta > 0\,:\,\langle \bar{x}(t), \bar{y}(t)\rangle \le 0\,, \quad \forall\, t \in I \text{ s.t. } \big|\|\bar{x}(t)\|_X - 1\big| < \delta\,,$$

$$\langle \bar{x}(1) - x_1, \bar{y}(1)\rangle \le 0 \quad \text{if } \|\bar{x}(1) - x_1\|_X^2 = r$$

and

$$\langle \bar{x}(1) - x_D, \bar{y}(1)\rangle = 0\,.$$

Then, there exists a function $p : I \to X$ such that $p, \bar{x}, \bar{u}$ satisfy the necessary conditions of Theorem 3.1. To be precise, there exist a positive $\psi \in \mathcal{M}(I, \mathbb{R}^+)$, $\lambda_0 \in \{0, 1\}$, $\lambda_1 \ge 0$ not all equal to zero such that the solution of the adjoint equation

$$\begin{cases} dp(t) = -\big(\mathbb{A} + b^T\bar{u}(t)\big)p(t)dt - 2\bar{x}(t)\psi(dt)\,, & t \in I \\ p(1) = \lambda_0(\bar{x}(1) - x_D) + \lambda_1 2(\bar{x}(1) - x_1) \end{cases}$$

satisfies the minimum principle

$$\langle p(t), b^T\bar{u}(t)\,\bar{x}(t)\rangle = \min_{u \in U}\langle p(t), b^T u\,\bar{x}(t)\rangle\,, \qquad \text{for a.e. } t \in I \tag{4.7}$$

and the second order necessary condition

$$(\lambda_0 + 2\lambda_1)\|\bar{y}(1)\|_X^2 + \int_I 2\|\bar{y}(s)\|_X^2\psi(ds) + \int_0^1 \langle p(s), b^T\bar{v}(s)\bar{y}(s)\rangle ds$$
$$+ \inf_{\{w \in L^1(I,X)\,:\,w(t) \in T_U^{\flat(2)}(\bar{u}(t),\bar{v}(t)) \text{ a.e. in } I\}} 2\int_0^1 \langle p(s), b^T w(s)\bar{x}(s)\rangle ds \ge 0\,.$$

Since for every $u \in U$ and $v \in T_U^\flat(u)$ the second order tangent set $T_U^{\flat(2)}(u, v)$ is a cone containing the origin, the second order necessary condition implies

$$\inf_{w \in T_U^{\flat(2)}(\bar{u}(t),\bar{v}(t))} \langle p(t), b^T w\bar{x}(t)\rangle = 0\,, \qquad \text{for a.e. } t \in I\,, \tag{4.8}$$

yielding

$$(\lambda_0 + 2\lambda_1)\|\bar{y}(1)\|_X^2 + \int_I 2\|\bar{y}(s)\|_X^2\psi(ds) + \int_0^1 \langle p(s), b^T\bar{v}(s)\bar{y}(s)\rangle ds \ge 0\,.$$

Observe that the first order condition (4.7) can be formulated as a variational inequality: for a.e. $t \in I$

$$\langle p(t), b^T(u - \bar{u}(t))\,\bar{x}(t)\rangle \ge 0\,, \qquad \text{for any } u \in U\,\,,$$

or via the polar cone,

$$-\int_\Omega p(t, \mathbf{x})b(\mathbf{x})\,\bar{x}(t, \mathbf{x})\,d\mathbf{x} \in (U - \bar{u}(t))^-\,, \qquad \text{for a.e. } t \in I\,,$$

where $()^-$ denotes the negative polar cone. The second order condition (4.8) contains an additional information. Indeed, by (4.5), (4.6) and (4.8) we obtain that for a.e. $t \in I$ either

$$\int_\Omega p(t, \mathbf{x})b(\mathbf{x})\,\bar{x}(t, \mathbf{x})\,d\mathbf{x} = 0$$

or

$$\bar{u}(t) \in \partial U\,, \quad \bar{v}(t) \in \partial\big(\mathrm{cone}(U - \bar{u}(t))\big)$$
$$\text{and} \quad -\textstyle\int_\Omega p(t,\mathbf{x}) b(\mathbf{x})\, \bar{x}(t,\mathbf{x})\, d\mathbf{x} \in \big(\mathrm{cone}(U - \bar{u}(t)) + \mathbb{R}\bar{v}(t)\big)^-\,.$$

We have in particular

$$\Big(\int_\Omega p(t,\mathbf{x}) b(\mathbf{x})\, \bar{x}(t,\mathbf{x})\, d\mathbf{x}\Big) \cdot \bar{v}(t) = 0 \quad \text{a.e. } t \in I\,.$$

Finally, as in Example 4.1 sufficient conditions for normality (3.24) can be simplified, namely, if $\zeta \in \mathcal{C}(I, L^2(\Omega))$, then the assumption

$$\exists\, \rho > 0 \text{ such that, if } \bar{x}(s) \in \partial K \text{ and } \langle \bar{x}(s), \zeta(s) + b^T w \bar{x}(s)\rangle \geq 0\,,$$

for some $s \in I$ and $w \in U$, then, $\exists\, v \in \mathrm{cone}(U - w)$ satisfying $\langle \bar{x}(s), b^T v \bar{x}(s)\rangle \leq -\rho$

implies (i) in (3.24).

5. The case of pointwise constraints

The arguments developed in the previous sections for optimal control problems under functional inequality constraints (1.2)–(1.3) can be adapted to problems involving constraints of a different type. In the present section we show how to adapt the results in the case of pointwise state constraints.

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary and $X = W^{1,p}(\Omega)$, with $p > n$. Then, by the embedding theorem, any $x \in X$ is continuous and has a continuous extension to $\overline{\Omega}$. Given a bounded closed subset U of a separable Banach space Z, the infinitesimal generator $\mathbb{A}$ of a strongly continuous semigroup $S(t) : X \to X$ and maps $g_0 : X \to \mathbb{R}$, $f : I \times X \times Z \to X$ satisfying the same assumptions as in the previous sections, we consider the solutions $x : I \to X : t \mapsto x(t, \cdot)$ of the control system (1.1) that satisfy running pointwise constraints of the form

$$x(t,\mathbf{x}) \leq a(\mathbf{x})\,, \quad \forall\, (t,\mathbf{x}) \in I \times \Omega \tag{5.1}$$

and the terminal pointwise constraint

$$x(1,\mathbf{x}) \leq b(\mathbf{x})\,, \quad \forall\, \mathbf{x} \in \Omega\,, \tag{5.2}$$

where a and b are continuous real functions on $\mathbb{R}^n$. Given a trajectory $\bar{x}$ of (1.1) satisfying (5.1) and (5.2), we say that a solution $(\bar{y}, \bar{v})$ of the linearized control system (2.10)–(2.11) is *pointwise strongly tangent* if there exists $\delta' > 0$ such that

$$\bar{y}(t,\mathbf{x}) \leq 0 \text{ for every } (t,\mathbf{x}) \in I \times \Omega \text{ satisfying } \bar{x}(t,\mathbf{x}) \geq a(\mathbf{x}) - \delta'\,,$$
$$\bar{y}(1,\mathbf{x}) \leq 0 \text{ for every } \mathbf{x} \in \Omega \text{ satisfying } \bar{x}(1,\mathbf{x}) \geq b(\mathbf{x}) - \delta'\,.$$

We then denote for any $0 \leq \delta \leq \delta'$

$$\Sigma_\delta = \{(t,\mathbf{x}) \in I \times \Omega : \bar{x}(t,\mathbf{x}) \geq a(\mathbf{x}) - \delta \text{ and } \bar{y}(t,\mathbf{x}) \geq -\delta\}\,,$$
$$\Omega_\delta = \{\mathbf{x} \in \overline{\Omega} : \bar{x}(1,\mathbf{x}) \geq b(\mathbf{x}) - \delta \text{ and } \bar{y}(1,\mathbf{x}) \geq -\delta\}$$

and introduce the convex cones

$$\mathcal{A} = \Big\{z \in \mathcal{C}(I, X) : \exists\, 0 < \delta \leq \delta' \text{ such that } z(t,\mathbf{x}) \leq 0\,,\ \forall\, (t,\mathbf{x}) \in \Sigma_\delta\Big\}\,, \tag{5.3}$$

$$\mathcal{B} = \Big\{ x \in X : x(\mathbf{x}) \le 0 \,,\ \forall \mathbf{x} \in \Omega_0 \Big\} \,.$$

Theorem 5.1. *Let $x_0 \in X$ and $(\bar{x}, \bar{u})$ be locally optimal for the problem of minimizing $g_0(x(1,\cdot))$ among all the trajectory/control pairs (x,u) satisfying* (2.1), (5.1) *and* (5.2). *Let $(\bar{y}, \bar{v})$ be a solution of* (2.10)–(2.11) *satisfying* (2.12). *Assume (H)(i)–(v), that g_0 is twice continuously Fréchet differentiable and that the set*

$$\mathcal{W} = \{ w : I \to Z \ \ \textit{measurable} : f_u[\cdot]w(\cdot) \in L^1(I, X) \,,\ w(t) \in T_U^{\flat(2)}(\bar{u}(t), \bar{v}(t)) \ \textit{a.e. in } I \}$$

is nonempty. If $\bar{y}$ is critical and pointwise strongly tangent, then there exist $\lambda_0 \in \{0,1\}$, $x^* \in \mathcal{B}^-$ and $y^* \in \mathcal{A}^-$, not all equal to zero, such that the solution $p : I \to X^*$ of the measure-driven adjoint system

$$\begin{cases} dp(t) = -\big(\mathbb{A}^* + f_x[t]^*\big)p(t)\mathrm{d}t + y^*(dt) \,, & t \in I \\ p(1) = \lambda_0 \nabla g_0(\bar{x}(1,\cdot)) + x^* \end{cases} \tag{5.4}$$

satisfies the minimum principle (3.3) *and the second order condition*

$$\lambda_0 \langle g_0''(\bar{x}(1,\cdot))\bar{y}(1,\cdot), \bar{y}(1,\cdot)\rangle + \int_0^1 \big\langle p(s), f''[s](\bar{y}(s,\cdot), \bar{v}(s))^2 + 2f_u[s]w(s)\big\rangle ds \ge 0 \,, \tag{5.5}$$

for any $w \in \mathcal{W}$.

Proof. The proof follows the same arguments as those developed in the previous sections. For this reason we detail here only the parts where some changes are made necessary by the different nature of the constraints. Consider the sets

$$\mathcal{S}^2 = \mathcal{S}^2(\bar{x}, \bar{y}) \,,$$

$$\mathcal{Q}_0^2 = \Big\{ z \in \mathcal{C}(I,X) : \langle \nabla g_0(\bar{x}(1,\cdot)), z(1,\cdot)\rangle + \frac{1}{2}\langle g_0''(\bar{x}(1,\cdot))\bar{y}(1,\cdot), \bar{y}(1,\cdot)\rangle < 0 \Big\} \,,$$

$$\mathcal{Q}_1^2 = \Big\{ z \in \mathcal{C}(I,X) : \exists\, \delta > 0 \text{ such that } z(1,\mathbf{x}) < 0 \,,\ \forall \mathbf{x} \in \Omega_\delta \Big\} \,,$$

$$\mathcal{K}^2 = \Big\{ z \in \mathcal{C}(I,X) : \exists\, \delta, \eta > 0 \text{ such that } z(t,\mathbf{x}) < -\eta \,,\ \forall\, (t,\mathbf{x}) \in \Sigma_\delta \Big\} \,.$$

Observe that $\mathcal{Q}_1^2$ and $\mathcal{K}^2$ are nonempty open convex cones. We have

$$\mathcal{S}^2 \cap \mathcal{Q}_0^2 \cap \mathcal{Q}_1^2 \cap \mathcal{K}^2 = \emptyset \,. \tag{5.6}$$

Indeed, suppose by contradiction that $z \in \mathcal{S}^2 \cap \mathcal{Q}_0^2 \cap \mathcal{Q}_1^2 \cap \mathcal{K}^2$ and let $h_n \to 0^+$ as $n \to +\infty$. By Lemma 2.9, there exists a sequence $x_n = \bar{x} + h_n\bar{y} + h_n^2 z_n \in \mathcal{S}$, such that $z_n \to z$ uniformly. As in Proposition 3.6, using the fact that $X \subset L^\infty(\Omega)$ with continuous injection and $z \in \mathcal{K}^2 \cap \mathcal{Q}_1^2$, one can prove that, for any n large enough, x_n satisfies the constraints (5.1) and (5.2). Reasoning as in the proof of Proposition 3.6, we obtain

$$g_0(x_n(1)) < g_0(\bar{x}(1)) \,,$$

for n sufficiently large, in contradiction with the optimality of $\bar{x}$.

Let $\tilde{\mathcal{S}^2} \subset \overline{\mathcal{S}^2}$ be defined as in (3.19) and suppose $\mathcal{Q}_0^2 \neq \emptyset$. By (5.6) and [29, Lemma 2.2], there exist $z^*, z_0^*, z_1^*, y^* \in \mathcal{C}(I,X)^*$ not vanishing simultaneously, such that

$$-z^* = z_0^* + z_1^* - y^* \tag{5.7}$$

and

$$\inf\langle z^*, \tilde{\mathcal{S}}^2\rangle + \inf\langle z_0^*, \mathcal{Q}_0^2\rangle + \inf\langle z_1^*, \mathcal{Q}_1^2\rangle + \inf\langle -y^*, \mathcal{K}^2\rangle \geq 0\,. \tag{5.8}$$

As in the proof of Lemma 3.7, we have that

$$z_0^* = -\lambda_0 \Gamma^* \nabla g_0(\bar{x}(1,\cdot)) \quad \text{for some positive } \lambda_0\,. \tag{5.9}$$

Consequently, we have

$$\inf\langle z_0^*, \mathcal{Q}_0^2\rangle = \inf\langle -\lambda_0 \Gamma^* \nabla g_0(\bar{x}(1,\cdot)), \mathcal{Q}_0^2\rangle = \frac{1}{2}\lambda_0 \langle g_0''(\bar{x}(1,\cdot))\bar{y}(1,\cdot), \bar{y}(1,\cdot)\rangle\,. \tag{5.10}$$

Further, since $\mathcal{Q}_1^2$ and $\mathcal{K}^2$ are nonempty cones, inequality (5.8) implies that

$$\inf\langle z_1^*, \mathcal{Q}_1^2\rangle = \inf\langle -y^*, \mathcal{K}^2\rangle = 0\,. \tag{5.11}$$

Consider the sets $\mathcal{A}$, defined by (5.3), and

$$\mathcal{A}_1 = \Big\{ z \in \mathcal{C}(I, X) : z(1, \mathbf{x}) \leq 0\,,\ \forall\, \mathbf{x} \in \Omega_0 \Big\}\,.$$

Reasoning as in the proof of Lemma 3.7, we have that $\mathcal{K}^2 + \mathcal{A} = \mathcal{K}^2$ and $\mathcal{Q}_1^2 + \mathcal{A}_1 = \mathcal{Q}_1^2$, implying that $y^* \in \mathcal{A}^-$ and $-z_1^* \in \mathcal{A}_1^-$. For any $z \in \mathcal{C}(I, X)$ such that $z(1,\cdot) \equiv 0$ we have that $\langle -z_1^*, z\rangle = \langle -z_1^*, -z\rangle = 0$, implying that

$$\langle -z_1^*, z_1\rangle = \langle -z_1^*, z_2\rangle, \qquad \text{for any } z_1, z_2 \in \mathcal{C}(I, X) \text{ with } z_1(1,\cdot) \equiv z_2(1,\cdot).$$

In other words, $\langle -z_1^*, z\rangle$ depends only on the terminal value $z(1,\cdot)$ and not on $z(t,\cdot)$, $t < 1$. For any $x \in X$, define $z_x \in \mathcal{C}(I, X)$ by $z_x(t,\cdot) \equiv x(\cdot)$ for every $t \in I$, and define $x^* \in X^*$ by $\langle x^*, x\rangle = \langle -z_1^*, z_x\rangle$. Then we obtain

$$\langle -z_1^*, z\rangle = \langle -z_1^*, z_{z(1,\cdot)}\rangle = \langle x^*, z(1,\cdot)\rangle = \langle \Gamma^* x^*, z\rangle \quad \forall\, z \in \mathcal{C}(I, X)\,. \tag{5.12}$$

Since $-z_1^* \in \mathcal{A}_1^-$ and $z_x \in \mathcal{A}_1$ for any $x \in \mathcal{B}$, we deduce that $x^* \in \mathcal{B}^-$. Combining (5.7), (5.8), (5.9), (5.10), (5.11) and (5.12), we obtain

$$\langle z^*, z\rangle = \lambda_0 \langle \nabla g_0(\bar{x}(1,\cdot)), z(1,\cdot)\rangle + \langle x^*, z(1,\cdot)\rangle + \langle y^*, z\rangle\,, \quad \forall\, z \in \mathcal{C}(I, X) \tag{5.13}$$

and

$$\inf\langle z^*, \tilde{\mathcal{S}}^2\rangle + \frac{1}{2}\lambda_0 \langle g_0''(\bar{x}(1,\cdot))\bar{y}(1,\cdot), \bar{y}(1,\cdot)\rangle \geq 0\,. \tag{5.14}$$

Let $p : I \to X^*$ be the solution of (5.4). We have

$$p(t) = \mathcal{T}(1,t)^* \Big(\lambda_0 \nabla g_0(\bar{x}(1,\cdot)) + x^*\Big) + \int_t^1 \mathcal{T}(s,t)^* y^*(ds)\,, \tag{5.15}$$

where $\mathcal{T}$ is the solution operator associated with

$$\dot{y}(t) = \mathbb{A}y(t) + f_x[t]y(t)\,.$$

Reasoning as in the proof of Theorem 3.1, we can deduce from (5.13) and (5.14) the minimum principle (3.3) and the second order condition (5.5) for any $w \in \mathcal{W}$.

To conclude the proof, it is sufficient to observe that if $\mathcal{Q}_0^2 = \emptyset$, then $\nabla g_0(\bar{x}(1,\cdot)) = 0$ and $\frac{1}{2}\langle g_0''(\bar{x}(1,\cdot))\bar{y}(1,\cdot), \bar{y}(1,\cdot)\rangle \geq 0$. Thus, the theorem holds with $\lambda_0 = 1$, $x^* = 0$ and $y^* = 0$. If $\lambda_0 > 0$, then dividing all the multipliers by λ_0 we may assume that $\lambda_0 = 1$. $\square$

Example 5.2. Consider the one-dimensional heat equation

$$\partial_t x(t,s) = \partial_{ss} x(t,s) - x(t,s) + u(t)\zeta(x(t,s)) \quad (t,s) \in [0,1]\times[0,1]\,, \tag{5.16}$$

with Neumann boundary conditions. The control function u takes values in the interval $[-1,1]$. The operator $\tilde{\mathbb{A}} = \partial_{ss} - I$ with domain $D(\tilde{\mathbb{A}}) = \big\{x \in H^2(0,1) : x'(0) = x'(1) = 0\big\}$ generates a strongly continuous semigroup $\tilde{S}(t)$ on $Y = L^2(0,1)$. The space $X = H^1(0,1)$ is an invariant subspace of Y by the semigroup $\tilde{S}$, and the restriction of $\tilde{S}(t)$ to X (called $S(t)$) is a strongly continuous semigroup on X, see [40]. Calling $\mathbb{A}$ its infinitesimal generator, we can write (5.16) in the form (1.1). Here, we assume that $f : (t,x,u) \mapsto u\,\zeta(x)$ satisfies (H)(i)–(iii). Given $x_0, x_D \in X$, let $(\bar{x},\bar{u})$ be locally optimal for the problem of minimizing the functional

$$g_0(x(1,\cdot)) = \|x(1,\cdot) - x_D\|_X^2\,,$$

among all the admissible trajectory/control pairs (x,u) satisfying the initial condition $x(0,\cdot) = x_0$, the running pointwise constraint (5.1) and the terminal pointwise constraint (5.2), with a and b continuous real functions on $[0,1]$.

Let $(\bar{y},\bar{v})$ be a solution of (2.10)–(2.11), such that $\bar{y}$ is pointwise strongly tangent and critical, and $\mathcal{W} \neq \emptyset$. By Theorem 5.1, there exist $\lambda_0 \in \{0,1\}$, $x^* \in \mathcal{B}^-$ and $y^* \in -\mathcal{A}^-$, not all equal to zero, such that the solution $p : I \to X^*$ of the measure-driven equation

$$\begin{cases} dp(t) = -\big(\mathbb{A}^* + \bar{u}(t) D\zeta(\bar{x}(t,\cdot))^*\big)p(t)\mathrm{d}t - y^*(dt)\,, & t \in I \\ p(1) = \lambda_0 2(\bar{x}(1,\cdot) - x_D) + x^* \end{cases} \tag{5.17}$$

satisfies the minimum principle

$$\bar{u}(t)\langle p(t), \zeta(\bar{x}(t,\cdot))\rangle = \min_{u\in[-1,1]} u\langle p(t), \zeta(\bar{x}(t,\cdot))\rangle\,, \quad \text{a.e. } t \in I \tag{5.18}$$

and the second order conditions

$$\lambda_0\|\bar{y}(1,\cdot)\|_X^2 + \int_0^1 \Big\langle p(t), \frac{1}{2}\bar{u}(t)\zeta''(\bar{x}(t,\cdot))(\bar{y}(t,\cdot))^2 + \bar{v}(t)D\zeta(\bar{x}(t,\cdot))\bar{y}(t,\cdot)\Big\rangle dt \geq 0\,,$$

$$\inf_w \int_0^1 w(t)\langle p(t), \zeta(\bar{x}(t,\cdot))\rangle dt = 0\,, \tag{5.19}$$

where the infimum is taken over measurable essentially bounded selections $w(t) \in T^{\flat(2)}_{[-1,1]}(\bar{u}(t),\bar{v}(t))$, a.e. in I. The minimum principle (5.18) implies that for a.e. $t \in I$

$$\langle p(t), \zeta(\bar{x}(t,\cdot))\rangle > 0 \quad \Rightarrow \quad \bar{u}(t) = -1$$

$$\langle p(t), \zeta(\bar{x}(t,\cdot))\rangle < 0 \quad \Rightarrow \quad \bar{u}(t) = 1\,,$$

while (5.19) implies that $\langle p(t), \zeta(\bar{x}(t,\cdot))\rangle = 0$ for a.e. $t \in I$ satisfying $\bar{v}(t) \neq 0$.

Appendix

We provide here proofs of Lemmas 2.6, 2.9, 2.10 and 2.11.

Proof of Lemma 2.6. At first we prove that $\bar{y}$ solves (2.7). Indeed, as in [23], for a.e. $t \in I$,

$$d_x F(t, \bar{x}(t), f^{\bar{x}}(t))\bar{y}(t) + T_{F(t,\bar{x}(t))}(f^{\bar{x}}(t)) \subset d_x F(t, \bar{x}(t), f^{\bar{x}}(t))\bar{y}(t),$$

and

$$f_x[t]\bar{y}(t) \in d_x F(t, \bar{x}(t), f^{\bar{x}}(t))\bar{y}(t)\,, \tag{A.1}$$

$$f_u[t]\bar{v}(t) \in T_{F(t,\bar{x}(t))}(f^{\bar{x}}(t))\,. \tag{A.2}$$

(A.1) follows from the very definition of $d_x F$. To prove (A.2), consider $h_n \to 0^+$ and $v_n \to \bar{v}(t)$ in Z, for $n \to \infty$, such that $\bar{u}(t) + h_n v_n \in U$ for any $n \in \mathbb{N}$. Taking

$$w_n := \frac{f(t, \bar{x}(t), \bar{u}(t) + h_n v_n) - f(t, \bar{x}(t), \bar{u}(t))}{h_n},$$

we get $f^{\bar{x}}(t) + h_n w_n \in F(t, \bar{x}(t))$ and, by (2.5), that, for some $R > 0$ and a.e. $t \in I$,

$$\begin{aligned}\|w_n - f_u[t]\bar{v}(t)\|_X &= \left\| \frac{f(t, \bar{x}(t), \bar{u}(t) + h_n v_n) - f(t, \bar{x}(t), \bar{u}(t))}{h_n} - f_u[t]\bar{v}(t) \right\|_X \\ &\le \psi_R \|v_n - \bar{v}(t)\|_Z + \left\| \frac{f(t, \bar{x}(t), \bar{u}(t) + h_n \bar{v}(t)) - f(t, \bar{x}(t), \bar{u}(t))}{h_n} - f_u[t]\bar{v}(t) \right\|_X \to 0\,,\end{aligned}$$

for $n \to \infty$, yielding (A.2).

To prove (2.9) for $\pi^{\bar{y}}(t) = f_x[t]\bar{y}(t) + f_u[t]\bar{v}(t)$, notice that, by (2.12), for any $0 < h \le h_0$ there exists a measurable $u_h(\cdot)$ satisfying, for a.e. $t \in I$, $u_h(t) \in U$ and $\|u_h(t) - \bar{u}(t) - h\bar{v}(t)\|_Z \le (c(t) + 1)h^2$. Then, $f(t, \bar{x}(t) + h\bar{y}(t), u_h(t)) \in F(t, \bar{x}(t) + h\bar{y}(t))$ and, by (2.5) and (2.6),

$$\begin{aligned}&\operatorname{dist}\big(f^{\bar{x}}(t) + hf_x[t]\bar{y}(t) + hf_u[t]\bar{v}(t), F(t, \bar{x}(t) + h\bar{y}(t))\big) \\ &\le \big\| f(t, \bar{x}(t) + h\bar{y}(t), u_h(t)) - f^{\bar{x}}(t) - hf_x[t]\bar{y}(t) - hf_u[t]\bar{v}(t)) \big\|_X \\ &\le \psi_R \|\bar{u}(t) + h\bar{v}(t) - u_h(t)\|_Z \\ &+ \int_0^1 \big[f_x(t, \bar{x}(t) + \lambda h\bar{y}(t), \bar{u}(t) + \lambda h\bar{v}(t)) - f_x[t]\big] h\bar{y}(t)\, d\lambda \\ &+ \int_0^1 \big[f_u(t, \bar{x}(t) + \lambda h\bar{y}(t), \bar{u}(t) + \lambda h\bar{v}(t)) - f_u[t]\big] h\bar{v}(t)\, d\lambda \\ &\le \big(\psi_R(c(t) + 1) + b(t)\big)h^2\end{aligned}$$

for some $R > 0$ and $b \in L^1(I, \mathbb{R}^+)$. Condition (2.9) follows, ending the proof. $\square$

Proof of Lemma 2.9. Set

$$x_n^1(t) = x(t) + h_n y(t) + h_n^2 z(t)\,.$$

We prove first that there exist solutions x_n^2 to

$$\dot{x}(t) \in \mathbb{A}x(t) + F(t, x(t))\,, \quad x(0) = x_0$$

such that

$$\frac{x_n^1 - x_n^2}{h_n^2} \to 0\,, \qquad \text{uniformly on } I \text{ as } n \to \infty\,. \tag{A.3}$$

Let π^y be as in (2.8) and $\alpha^z(t) \in d_x^2 F[t]z(t)$ be an integrable selection such that

$$z(t) = \int_0^t S(t-s)\alpha^z(s)ds\,, \quad \forall\, t \in I\,.$$

By the definition of set-valued derivative, we obtain that

$$\lim_{h\to 0+} \operatorname{dist}\Big(\alpha^z(t), \frac{F(t, x(t) + hy(t) + h^2 z(t)) - f^x(t) - h\pi^y(t)}{h^2}\Big) = 0\,, \qquad \text{for a.e. } t \in I\,.$$

Hence, setting

$$\gamma_n(t) = \operatorname{dist}\big(f^x(t) + h_n\pi^y(t) + h_n^2\alpha^z(t), F(t, x_n^1(t))\big)$$

we deduce that, for a.e. $t \in I$,

$$\lim_{n\to\infty} \frac{\gamma_n(t)}{h_n^2} = 0\,.$$

Further, using (2.3) and (2.9) we get

$$\begin{aligned}
\gamma_n(t) &\le \operatorname{dist}\big(f^x(t) + h_n\pi^y(t) + h_n^2\alpha^z(t), F(t, x(t) + h_n y(t))\big) + h_n^2 k_R(t)\|z(t)\|_X \\
&\le \operatorname{dist}\big(f^x(t) + h_n\pi^y(t), F(t, x(t) + h_n y(t))\big) + h_n^2\|\alpha^z(t)\|_X + h_n^2 k_R(t)\|z(t)\|_X \\
&\le h_n^2\big(a(t) + \|\alpha^z(t)\|_X + k_R(t)\|z\|_{L^\infty(I,X)}\big)
\end{aligned}$$

with $R > 0$ such that $\|x(t) + h_n y(t)\|_X + \|x(t) + h_n y(t) + h_n^2 z(t)\|_X \le R$, for every $t \in I, n \ge 1$. The dominated convergence theorem ensures that

$$\lim_{n\to\infty} \frac{1}{h_n^2}\int_0^1 \gamma_n(t)dt = 0\,.$$

By Lemma A.1 from [26], we obtain x_n^2 as required. A relaxation theorem, [23, Theorem 2.5], guarantees the existence of $\tilde{x}_n$ solving

$$\dot{x}(t) \in \mathbb{A}\, x(t) + \overline{f(t, x(t), U)}\,, \quad x(0) = x_0\,,$$

such that

$$\frac{\tilde{x}_n - x_n^2}{h_n^2} \to 0\,, \qquad \text{uniformly on } I \text{ as } n \to \infty. \tag{A.4}$$

Now, for some measurable $v_n(\cdot) \in \overline{f(\cdot, \tilde{x}_n(\cdot), U)}$, we have that

$$\tilde{x}_n(t) = S(t)x_0 + \int_0^t S(t-s)v_n(s)ds\,.$$

By [2, Theorem 8.2.9] there exist controls $u_n(\cdot)$ such that $\|f(t, \tilde{x}_n(t), u_n(t)) - v_n(t)\|_X \le h_n^3$ for every $t \in I$. Define

$$y_n(t) = S(t)x_0 + \int_0^t S(t-s)f(s, \tilde{x}_n(s), u_n(s))\,ds \quad \text{for any } t \in I\,.$$

Then $\|\tilde{x}_n(t) - y_n(t)\|_X \leq M_S h_n^3$ for every $t \in I$ and therefore for some $R > 0$ and all n we have

$$\|f(t, \tilde{x}_n(t), u_n(t)) - f(t, y_n(t), u_n(t))\|_X \leq k_R(t) M_S h_n^3 \qquad \text{for a.e. } \ t \in I\,.$$

Applying again [26, Lemma A.1], we finally deduce the existence of $x_n \in \mathcal{S}$ satisfying

$$\frac{x_n - \tilde{x}_n}{h_n^2} \to 0\,, \qquad \text{uniformly on } I \text{ as } n \to \infty.$$

This, (A.3) and (A.4) complete the proof. □

Proof of Lemma 2.10. It is clear that for every $\lambda \geq 0$ and $w \in \mathcal{A}$ we have $\langle \lambda \Gamma^* \xi, w\rangle = \lambda \langle \xi, w(1)\rangle \leq 0$, so that $\lambda \Gamma^* \xi \in \mathcal{A}^-$. On the other hand, let $\alpha \in \mathcal{A}^-$. For any $w \in \mathcal{C}(I, X)$ such that $w(1) = 0$ we have that $\langle \alpha, w\rangle = \langle \alpha, -w\rangle = 0$, implying that

$$\langle \alpha, w_1\rangle = \langle \alpha, w_2\rangle, \qquad \text{for any } w_1, w_2 \in \mathcal{C}(I, X) \text{ with } w_1(1) = w_2(1).$$

In other words, $\langle \alpha, w\rangle$ depends only on the terminal value $w(1)$ and not on $w(t)$, $t < 1$. For any $x \in X$, define $w_x \in \mathcal{C}(I, X)$ by $w_x(t) \equiv x$, and $\beta \in X^*$ as $\langle \beta, x\rangle = \langle \alpha, w_x\rangle$. Then, for any $w \in \mathcal{C}(I, X)$, $\langle \alpha, w\rangle = \langle \alpha, w_{w(1)}\rangle = \langle \beta, w(1)\rangle = \langle \Gamma^* \beta, w\rangle$. To prove the claim, we show that $\beta = \lambda \xi$, for some $\lambda \geq 0$. Indeed, assume by contradiction that $\beta \notin \mathcal{B} = \{\lambda \xi : \lambda \geq 0\}$. We claim that the set $\mathcal{B}$ is weakly-star closed. Indeed, let $\lambda_n \xi \overset{*}{\rightharpoonup} \eta$, weakly-star in X^*, then λ_n is bounded. Taking a subsequence and using the same notation, $\lambda_n \to \lambda \geq 0$. So, $\lambda_n \xi \to \lambda \xi$ in X^*, implying $\eta = \lambda \xi \in \mathcal{B}$. Hence $\mathcal{B}$ is sequentially weakly-star closed in X^*. Since X is separable and $\mathcal{B} \subset X^*$ is convex, by [16, Corollary V.12.7] $\mathcal{B}$ is also weakly-star closed.

By the separation theorem, there exist $x \in X$ and $\varepsilon > 0$ such that, for any $\lambda \geq 0$,

$$\langle \lambda \xi, x\rangle + \varepsilon \leq \langle \beta, x\rangle.$$

Since $\mathcal{B}$ is a cone, $\sup_{\lambda \geq 0} \langle \lambda \xi, x\rangle = 0$. Hence, the continuous function w_x defined above satisfies $\langle \alpha, w_x\rangle = \langle \beta, x\rangle > 0$. On the other hand, $\langle \xi, x\rangle \leq 0$ and therefore $w_x \in \mathcal{A}$. Hence $\langle \alpha, w_x\rangle \leq 0$. The obtained contradiction ends the proof. □

Proof of Lemma 2.11. Denote by $\mathcal{B}$ the right hand side of (2.13) and observe that it is a convex cone. Let $\gamma \in \mathcal{B}$ and let ψ be the associated positive measure in $\mathcal{M}(I, \mathbb{R})$. We obtain

$$\int_I \langle w(t), \gamma(dt)\rangle = \int_I \langle \xi(t), w(t)\rangle \psi(dt) = \int_M \langle \xi(t), w(t)\rangle \psi(dt) \leq 0\,, \ \text{ for any } w \in \mathcal{A}\,,$$

proving that $\gamma \in \mathcal{A}^-$. In order to show the inclusion $\mathcal{A}^- \subset \mathcal{B}$, let us first remark that

$$\|\gamma\|_{\mathcal{M}(I, X^*)} = \int_M \|\xi(t)\|_{X^*} \psi(dt)\,, \quad \forall\, \gamma \in \mathcal{B}\,. \tag{A.5}$$

Indeed, by definition

$$\|\gamma\|_{\mathcal{M}(I, X^*)} = \sup_{w \in \mathcal{C}(I, X) : \|w\| \leq 1} \int_M \langle \xi(t), w(t)\rangle \psi(dt) \leq \int_M \|\xi(t)\|_{X^*} \psi(dt)\,, \tag{A.6}$$

where ψ is as in (2.13). If $\psi(I) = 0$, inequality (A.6) implies (A.5). On the other hand, suppose $\psi(I) > 0$ and

$$\sup_{w \in \mathcal{C}(I,X): \|w\| \le 1} \int_M \langle \xi(t), w(t) \rangle \psi(dt) = \int_M \|\xi(t)\|_{X^*} \psi(dt) - \varepsilon \tag{A.7}$$

for some $\varepsilon > 0$. Consider the set-valued map $R : M \rightsquigarrow X$ defined by

$$R(t) = \left\{ \alpha \in B \; : \; \langle \xi(t), \alpha \rangle - \|\xi(t)\|_{X^*} \in \left[-\frac{\varepsilon}{2\psi(I)}, 0 \right] \right\} .$$

R is lower semicontinuous (see for instance [2, Proposition 1.5.1]) with nonempty closed convex values. Then by Michael's Theorem (see [2, Theorem 9.1.2]) there exists a continuous selection $\bar{w}$ of R, satisfying

$$\int_M \langle \xi(t), \bar{w}(t) \rangle \psi(dt) \ge \int_M \|\xi(t)\|_{X^*} \psi(dt) - \frac{\varepsilon}{2} .$$

By the Dugundji theorem, $\bar{w}$ can be extended on I to a continuous function denoted again by $\bar{w}$ and taking values in B, see [20, Theorem 4.1]. This contradicts (A.7) and proves (A.5). Further, remark that the set $\mathcal{B}$ is convex, non empty ($0 \in \mathcal{B}$) and weakly-star closed. Indeed, let $\gamma_n \in \mathcal{B}$ be such that $\gamma_n \overset{*}{\rightharpoonup} \tilde{\gamma}$, weakly-star in $\mathcal{M}(I, X^*)$, namely,

$$\int_I \langle w(t), \gamma_n(dt) \rangle \to \int_I \langle w(t), \tilde{\gamma}(dt) \rangle , \qquad \text{for any } w \in \mathcal{C}(I, X) .$$

We claim that $\tilde{\gamma} \in \mathcal{B}$. If $\tilde{\gamma} = 0$, then the claim is clearly true. Assume that $\tilde{\gamma} \neq 0$. It is not restrictive to suppose that $\gamma_n \neq 0$, for any n. For any n, let ψ_n be as in (2.13), i.e.

$$\int_I \langle w(t), \gamma_n(dt) \rangle = \int_I \langle \xi(t), w(t) \rangle \psi_n(dt) , \qquad \text{for any } w \in \mathcal{C}(I, X) ,$$

and $C > 0$ be such that $\sup_{n \in \mathbb{N}} \|\gamma_n\|_{\mathcal{M}(I,X^*)} \le C$. Since there exists $\varepsilon_0 > 0$ such that $\|\xi(t)\|_{X^*} \ge \varepsilon_0$, for any $t \in M$, we have

$$\varepsilon_0 \sup_{n \in \mathbb{N}} \int_I \psi_n(dt) = \varepsilon_0 \sup_{n \in \mathbb{N}} \int_M \psi_n(dt) \le \sup_{n \in \mathbb{N}} \int_M \|\xi(t)\|_{X^*} \psi_n(dt) = \sup_{n \in \mathbb{N}} \|\gamma_n\|_{\mathcal{M}(I,X^*)} \le C .$$

Hence, taking a subsequence and keeping the same notation, $\psi_n \overset{*}{\rightharpoonup} \tilde{\psi}$ weakly-star in $\mathcal{M}(I, \mathbb{R})$, namely

$$\int_I \nu(t) \psi_n(dt) \to \int_I \nu(t) \tilde{\psi}(dt) , \qquad \text{for any } \nu \in \mathcal{C}(I, \mathbb{R}) .$$

The measure $\tilde{\psi} \in \mathcal{M}(I, \mathbb{R})$ is clearly positive and $\operatorname{supp} \tilde{\psi} \subset M$. Let $w \in \mathcal{C}(I, X)$, then $\langle \xi(\cdot), w(\cdot) \rangle \in \mathcal{C}(I, \mathbb{R})$, and

$$\int_I \langle w(t), \gamma_n(dt) \rangle = \int_I \langle \xi(t), w(t) \rangle \psi_n(dt) .$$

Taking the limit as $n \to +\infty$, we obtain

$$\int_I \langle \xi(t), w(t) \rangle \tilde{\psi}(dt) = \int_I \langle w(t), \tilde{\gamma}(dt) \rangle,$$

proving that $\tilde{\gamma} \in \mathcal{B}$. Hence $\mathcal{B}$ is sequentially weakly-star closed in $\mathcal{M}(I, X^*)$. Since X is separable, then the space $\mathcal{C}(I, X)$ is also separable (see [32, Theorem I.4.19]). Applying again [16, Corollary V.12.7], we obtain that $\mathcal{B}$ is weakly-star closed.

Let $\gamma \in \mathcal{A}^-$. We can suppose that $\gamma \neq 0$, otherwise γ is obviously in $\mathcal{B}$. Assume by contradiction that $\gamma \notin \mathcal{B}$, then by a separation theorem, there exists $w \in \mathcal{C}(I, X)$ such that

$$0 = \sup_{\tilde{\gamma} \in \mathcal{B}} \int_I \langle w(t), \tilde{\gamma}(dt) \rangle < \int_I \langle w(t), \gamma(dt) \rangle \,. \tag{A.8}$$

This implies that

$$\int_I \langle \xi(t), w(t) \rangle \tilde{\psi}(dt) \leq 0 \,, \qquad \text{for any positive } \tilde{\psi} \in \mathcal{M}(I, \mathbb{R}) \text{ with } \operatorname{supp} \tilde{\psi} \subset \mathrm{M} \,.$$

yielding

$$\langle \xi(t), w(t) \rangle \leq 0 \,, \qquad \text{for any } t \in M.$$

Hence $w \in \mathcal{A}$ and

$$\int_I \langle w(t), \gamma(dt) \rangle \leq 0,$$

in contradiction with (A.8). The proof is complete. $\square$

Conflict of interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Data availability

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

[1] A. Ambrosetti, G. Prodi, *A Primer of Nonlinear Analysis*, Cambridge University Press, Cambridge, 1993.
[2] J.-P. Aubin, H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston, Basel, Berlin, 1990.
[3] T. Bayen, F.J. Silva, *Second order analysis for strong solutions in the optimal control of parabolic equations*, SIAM J. Control Optim. **54** (2016), 819–844.
[4] J.F. Bonnans, P. Jaisson, *Optimal control of a parabolic equation with time-dependent state constraints*, SIAM J. Control Optim. **48** (2010), 4550–4571.
[5] J.F. Bonnans, A. Shapiro, *Perturbation analysis of optimization problems*, Springer, New York, 2000.
[6] P. Cannarsa, H. Frankowska, *Value function and optimality conditions for semilinear control problems*, Applied Mathematics and Optimization **26** (1992), 139-169.
[7] H. Cartan, *Differential calculus*, Houghton Mifflin, Boston, 1971.
[8] E. Casas, J.C. De Los Reyes, F. Tröltzsch, *Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints*, SIAM J. Control Optim. **19** (2008), 616–643.
[9] E. Casas, F. Kruse, K. Kunish, *Optimal control of semilinear parabolic equations by BV-functions*, SIAM J. Control Optim. **55** (2017), 1752–1788.
[10] E. Casas, C. Ryll, F. Tröltzsch, *Optimal control of a class of reaction-diffusion systems*, Comput. Optim. Appl. **70** (2018), 677–707.

[11] E. Casas, F. Tröltzsch, *Second order necessary and sufficient optimality conditions for optimization problems and applications to control theory*, SIAM J. Control Optim. **13** (2002), 406–431.

[12] E. Casas, F. Tröltzsch, *Second order analysis for optimal control problems: improving results expected from abstract theory*, SIAM J. Control Optim. **22** (2012), 261–279.

[13] E. Casas, F. Tröltzsch, *Second order optimality conditions and their role in PDE control*, Jahresber. Dtsch. Math. Ver. **117** (2015), 3–44.

[14] E. Casas, F. Tröltzsch, *Second-order optimality conditions for weak and strong local solutions of parabolic optimal control problems*, Vietnam J. Math. **44** (2016), 181–202.

[15] F.H. Clarke, *Optimization and nonsmooth analysis*, SIAM, Philadelphia, 1990.

[16] J.B. Conway, *A course in functional analysis*, Springer-Verlag, New York, 1985.

[17] S. Court, K. Kunish, L. Pfeiffer, *Hybrid optimal control problems for a class of semilinear parabolic equations*, Discrete Contin. Dyn. Syst. **11** (2018), 1031–1060.

[18] J.C. de Los Reyes, P. Merino, J. Rehberg, F.S. Tröltzsch, *Optimality conditions for state-constrained PDE control problems with time-dependent controls*, Control Cybernet. **37** (2008), 5–38.

[19] J. Diestel, J.J. Uhl Jr., *Vector measures*, AMS, Providence, RI, 1977.

[20] J. Dugundji, *An extension of Tietze's theorem*, Pacific J. Math. **1** (1951), 353–367.

[21] H. O. Fattorini, *Optimal control problems with state constraints for semilinear distributed-parameter system*, J. Optim. Theory Appl. **88** (1996), 25–59.

[22] H. O. Fattorini, *Infinite-dimensional optimization and control theory*, Cambridge University Press, Cambridge, 1999.

[23] H. Frankowska, *A priori estimates for operational differential inclusions*, J. Differential Equations **84** (1990), 100–128.

[24] H. Frankowska, D. Hoehener and D. Tonon, A second-order maximum principle in optimal control under state constraints, *Serdica Math. J.*, vol. 39, 2013, pp 233-270.

[25] H. Frankowska, Q. Lü *Second order necessary conditions for optimal control problems of evolution equations with mixed final point constraints*, ESAIM: COCV **27** (2020), 71.

[26] H. Frankowska, E.M. Marchini, M. Mazzola, *A relaxation result for state constrained inclusions in infinite dimension*, Math. Control Relat. Fields **6** (2016), 113–141.

[27] H. Frankowska, E.M. Marchini, M. Mazzola, *Distance estimates for state constrained trajectories of infinite dimensional differential inclusions*, ESAIM: COCV **24** (2017), 1207–1229.

[28] H. Frankowska, E.M. Marchini, M. Mazzola, *Necessary optimality conditions for infinite dimensional state constrained control problems*, J. Differential Equations **264** (2018), 7294–7327.

[29] H. Frankowska, N.P. Osmolovskii, *Strong local minimizers in optimal control problems with state constraints: second-order necessary conditions*, SIAM J. Control Optim. **56** (2018), 2353–2376.

[30] H. Frankowska, N.P. Osmolovskii, *Second-order necessary conditions for a strong local minimum in a control problem with general control constraints*, Appl. Math. Optim. **80** (2019), 135–164.

[31] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich, *Optimization with PDE constraints, volume 23 of mathematical modelling: theory and applications*, Springer, New York, 2009.

[32] A. Kechris, *Classical Descriptive Set Theory*, Springer-Verlag, New York, 1995.

[33] K. Krumbiegel, J. Rehberg, *Second order sufficient optimality conditions for parabolic optimal control problems with pointwise state constraints*, SIAM J. Control Optim. **51** (2013), 304–331.

[34] P.S. Lomdhal, O.H. Soerensen, P.L. Christiansen, *Soliton excitations in Josephson tunnel junctions*, Phys. Rev. B **25** (1982), 5737–5748.

[35] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, New York, 1983.

[36] J.-P. Raymond, F. Tröltzsch, *Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints*, Discrete Contin. Dyn. Syst. **6** (2000), 431–450.

[37] A. Rösch, F. Tröltzsch, *Sufficient second-order optimality conditions for a parabolic optimal control problem with pointwise control-state constraints*, SIAM J. Control Optim. **42** (2003), 138–154.

[38] F.J. Silva, *Second order analysis for the optimal control of parabolic equations under control and final state constraints*, Set Valued Var. Anal **24** (2016), 57–81.

[39] H. M. Soner, *Optimal control with state-space constraints*, SIAM J. Control Optim. **24** (1986), 552–561.

[40] R. Temam, *Infinite-dimensional dynamical systems in mechanics and physics*, Springer, New York, 1997.

CNRS, Institut de Mathématiques de Jussieu - Paris Rive Gauche, Sorbonne Université, Case 247, 4 Place Jussieu, 75252 Paris, France
Email address: helene.frankowska@imj-prg.fr

Dipartimento di Matematica, Politecnico di Milano
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
Email address: elsa.marchini@polimi.it

Sorbonne Université, Institut de Mathématiques de Jussieu - Paris Rive Gauche, CNRS, Case 247, 4 Place Jussieu, 75252 Paris, France
Email address: marco.mazzola@imj-prg.fr