HAL
open science

On second order necessary conditions in infinite dimensional optimal control with state constraints

Hélène Frankowska, E M Marchini, M Mazzola

To cite this version:

Hélène Frankowska, E M Marchini, M Mazzola. On second order necessary conditions in infinite dimensional optimal control with state constraints. 58th IEEE Conference on Decision and Control, Dec 2019, Nice, France. pp.2416-2421, 10.1109/CDC40024.2019.9030176 . hal-02343395

HAL Id: hal-02343395

https://hal.science/hal-02343395v1

Submitted on 6 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On second order necessary conditions in infinite dimensional optimal control with state constraints

H. Frankowska, E.M. Marchini and M. Mazzola

***Abstract*— This paper is devoted to second order necessary optimality conditions for control problems in infinite dimensions. The main novelty of our work is the presence of pure state constraints together with end point constraints, quite useful in the applications.**

Second order analysis for control problems involving PDEs has been extensively discussed in the literature. The most usual approach to derive necessary optimality conditions is to rewrite the control problem as an abstract mathematical programming one. Our approach is different, we avoid the reformulation of the optimal control problem and use instead second order variational analysis. The necessary optimality conditions are in the form of a maximum principle and a second order variational inequality. They are first obtained in the form of nonintersection of convex sets. A suitable separation theorem allows to deduce their dual characterization.

I. INTRODUCTION

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

$$\dot{x}(t) = \mathbb{A}x(t) + f(t, x(t), u(t)), \quad \text{a.e. } t \in I, \quad x(0) = x_0 \quad (1)$$

that satisfy an end point constraint

$$x(1) \in Q = \bigcap_{i=1,\dots,k} Q_i = \bigcap_{i=1,\dots,k} \left\{ x \in X : g_i(x) \le 0 \right\}, \quad (2)$$

and a state constraint, for any $t \in I$,

$$x(t) \in K = \bigcap_{j=1,\dots,q} K_j = \bigcap_{j=1,\dots,q} \left\{ x \in X : \varphi_j(x) \le 0 \right\}. \quad (3)$$

Here, u is a measurable control, that is a function from I to 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, while the functions $g_i : X \to \mathbb{R}$, $\varphi_j : X \to \mathbb{R}$ are twice Fréchet differentiable. The trajectories of (1) are understood in the mild sense (see [18]). In this paper we analyze an infinite dimensional Mayer problem: given a twice Fréchet differentiable map $g_0 : X \to \mathbb{R}$, consider the solutions of the problem

$$\text{minimize } \left\{ g_0(x(1)) : x(\cdot) \text{ is a solution of (1)–(3), for some control } u(\cdot) \right\}. \quad (4)$$

The main results of this paper deal with second order necessary optimality conditions. More precisely, let $(\bar{x}, \bar{u})$ be an optimal solution for our constrained problems and take a critical direction $\bar{y}$ that is tangent to the constraints. Then

$$\langle \nabla g_0(\bar{x}(1)), z(1) \rangle + \frac{1}{2} \langle g_0''(\bar{x}(1))\bar{y}(1), \bar{y}(1) \rangle \ge 0, \quad (5)$$

for any suitable second-order tangent z to the constraints, see the exact definition of tangents in Section II.

We work in a quite general infinite dimensional framework, hence our results apply to optimal control problems involving some classes of PDEs, see [17], [18], [21] where reduction of some PDEs to the form (1) is discussed. In many phenomena, such as heat conduction, reaction-diffusion, population dynamics, economics, one seeks to optimize measures of best performances. The optimal control theory involving PDEs represents the natural framework to deal with such models. In this setting, second order analysis has been largely studied, with particular emphasis on sufficient second order conditions, due to their application in numerical analysis. It is impossible to provide here an exhaustive list of papers. Some significant contributions can be found in [3], [4], [5], [6], [7], [16], [19], [20], where evolution equations are analyzed with a particular interest to the parabolic case, see also the bibliographies therein. Second order optimality conditions are usually obtained by rewriting the control problem as an abstract mathematical programming one. However this approach requires Robinson like constraint qualification conditions, implying severe restrictions on the data. Control constraints, mixed (control-state) constraints, and some particular cases of state constraints were already investigated in the literature. Nevertheless, to our knowledge a general theory involving pure state constraints and end-point constraints is still lacking.

The main novelty of our work are variational techniques, quite different from those based on the reduction to a mathematical programming problem. Following the approach developed in the finite dimensional framework in [14], [15] we avoid such reformulation of the optimal control problem. This allows to work directly with the class of measurable controls, as in the well-known and developed theory of first order conditions, see e.g. [22], and to treat quite general

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 and of the support from the Air Force Office of Scientific Research under award number FA9550-18-1-0254.

H. Frankowska is with CNRS, Institut de Mathématiques de Jussieu - Paris Rive gauche, Sorbonne Université, 4 Place Jussieu, 75252 Paris, France, helene.frankowska@imj-prg.fr

E.M. Marchini is with Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy, elsa.marchini@polimi.it

M. Mazzola is with Institut de Mathématiques de Jussieu - Paris Rive gauche, CNRS, Sorbonne Université, 4 Place Jussieu, 75252 Paris, France, marco.mazzola@imj-prg.fr

pure state constraints and end point constraints, useful in applications. Further our abstract approach allows to apply the results of this paper directly to different kinds of control problems involving PDEs. For lack of space, we provide only one classical example involving a parabolic equation, nevertheless applications to wave equations are also possible, as analyzed in a forthcoming paper [13].

The necessary optimality conditions, involving a maximum principle and a second order variational inequality, are first obtained in the form of nonintersection of convex sets, using an approximation result dealing with second-order variations. By applying a suitable separation theorem, we get their dual characterization. We do not need normality of the maximum principle to get our results, so we can impose more general assumptions with respect to the classical ones present in the literature.

II. PRELIMINARIES

We list the notation, the definitions and the main assumptions in use.

A. 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; $\mathscr{C}(I,X)$ 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; $\mathscr{M}(I,Y)$ the space of countably additive regular measures of bounded variation on I with values in Y.
  It is known that $\mathscr{M}(I,X^*)$ is isomorphic to the dual space of $\mathscr{C}(I,X)$, see [9];
- $\langle \cdot,\cdot \rangle$ stands for the duality pairing on $X^* \times X$;
- given a set-valued map $\mathscr{F} : X \rightsquigarrow X$, $x \in X$ and $y \in \mathscr{F}(x)$, the adjacent derivative $d\mathscr{F}(x,y) : X \rightsquigarrow X$ is defined by $v \in d\mathscr{F}(x,y)u$ iff
  $$\lim_{h\to 0+} \operatorname{dist}\Big(v, \frac{\mathscr{F}(x+hu)-y}{h}\Big) = 0;$$
  and, for $v \in d\mathscr{F}(x,y)(u)$, the second order adjacent variation $d^2\mathscr{F}(x,y,u,v) : X \rightsquigarrow X$ by $z \in d^2\mathscr{F}(x,y,u,v)(w)$ iff
  $$\lim_{h\to 0+} \operatorname{dist}\Big(z, \frac{\mathscr{F}(x+hu+h^2w)-y-hv}{h^2}\Big) = 0;$$
- given $K \subset X$ and $x \in \overline{K}$, define the adjacent tangent cone to K at x by
  $$T_K^\flat(x) = \Big\{ y \in X : \lim_{h\to 0^+} \operatorname{dist}\Big(y, \frac{K-x}{h}\Big) = 0 \Big\}$$
  and the second order tangent set to K at (x,y) by
  $$T_K^{\flat(2)}(x,y) = \Big\{ z \in X : \lim_{h\to 0^+} \operatorname{dist}\Big(z, \frac{K-x-hy}{h^2}\Big) = 0 \Big\};$$
- $d_K(x)$ denotes the distance from $x \in X$ to K.

Definition 2.1: Let $x_0 \in X$. A function $x \in \mathscr{C}(I,X)$ is a (mild) solution of (1) with initial datum $x(0) = x_0$ if it satisfies, for any $t \in I$,

$$x(t) = S(t)x_0 + \int_0^t S(t-s) f(s,x(s),u(s))\, ds\,,$$

for some control $u(\cdot)$. We denote $f^x(t) = f(t,x(t),u(t))$ for any $t \in I$. If in addition x satisfies (2)-(3), we say that (x,u) is an admissible pair for problem (4).
Notice that, since $S(\cdot)$ is a strongly continuous semigroup, there exists $M_S > 0$ such that

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

B. Assumptions

The following conditions (H) are imposed in the main results:

(*i*) f is measurable in t, twice Fréchet differentiable w.r.t. (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
$$\|f(t,x,u) - f(t,y,u)\|_X \le k_R(t)\|x-y\|_X\,;$$

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

(*iv*) g_i, for $i = 0,\dots k$, and φ_j, for $j = 1,\dots q$, are twice Fréchet differentiable.

C. Critical directions and second order tangent variations

For a.e. $t \in I$ and any $x \in X$, set

$$F(t,x) = \overline{\text{co}} f(t,x,U)\,,$$

where $\overline{\text{co}} f(t,x,U)$ is the closed convex hull of $f(t,x,U)$.

Given x a solution of (1), we introduce the linearized differential inclusion:

$$\dot y(t) \in \mathbb{A}y(t) + d_xF(t,x(t),f^x(t))y(t)\,, \quad y(0) = 0\,. \tag{6}$$

Definition 2.2: A solution y of (6) is a *first order variation* if there exist $a \in L^1(I,\mathbb{R}^+)$, $h_0 > 0$ s.t. $\forall\, 0 < h \le h_0$,

$$\operatorname{dist}\big(f^x(t) + h\pi^y(t), F(t,x(t)+hy(t))\big) \le a(t)h^2\,, \tag{7}$$

where $\pi^y(t) \in d_xF(t,x(t),f^x(t))y(t)$ is an integrable selection such that, for every $t \in I$,

$$y(t) = \int_0^t S(t-s)\pi^y(s)ds\,. \tag{8}$$

Sufficient conditions ensuring the validity of (7) can be found in [13], see also [14] for the case $Z = \mathbb{R}^n$. We say that a first order variation y is *admissible*, and we write $y \in \mathscr{V}^1(x)$, if

$$\langle \nabla g_i(x(1)), y(1)\rangle \le 0\,,\ \ \forall\, i = 1,\dots,k \text{ s.t. } x(1) \in \partial Q_i\,,$$
$$\langle \nabla \varphi_j(x(t)), y(t)\rangle \le 0\,,\ \ \forall\, j = 1,\dots,q, t \in I \text{ s.t. } x(t) \in \partial K_j\,.$$

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

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

Definition 2.3: Let x solve (1) and y be a first order variation. We say that $z \in \mathscr{C}(I,X)$ is a *second order variation* at (x,y) if it is a solution of the inclusion

$$\dot{z}(t) \in \mathbb{A}z(t) + d_x^2 F[t]z(t)\,, \quad z(0) = 0\,,$$

where $[t] = (t, x(t), f^x(t), y(t), \pi^y(t))$ and π^y is as in (8).
To deduce second order conditions (5) we need to find suitable second order tangents to the constraints. For this aim, we introduce the following sets:

$$\mathscr{S} = \left\{x \in \mathscr{C}(I,X) : x \text{ solves (1) for some control } u(\cdot)\right\},$$
$$\mathscr{Q} = \left\{x \in \mathscr{C}(I,X) : x(1) \in Q\right\},$$
$$\mathscr{K} = \left\{x \in \mathscr{C}(I,X) : x(I) \subset K\right\},$$

Given $(\bar{x}, \bar{u})$ that is optimal for problem (4) and $\bar{y} \in \mathscr{V}^1(\bar{x})$ that is critical, we aim to express some second order tangents to $\mathscr{S} \cap \mathscr{Q} \cap \mathscr{K}$. Since, in general, the inclusion

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

may be strict, we need to work with convex subsets of $T^{\flat(2)}_{\mathscr{S}}(\bar{x},\bar{y})$, $T^{\flat(2)}_{\mathscr{Q}}(\bar{x},\bar{y})$, and $T^{\flat(2)}_{\mathscr{K}}(\bar{x},\bar{y})$. To this aim define the sets:

$$\mathscr{S}^2 = \left\{z \in \mathscr{C}(I,X) : z \text{ is a } \textit{second order variation} \text{ at } (\bar{x},\bar{y})\right\},$$
$$\mathscr{Q}^2_i = \Big\{z \in \mathscr{C}(I,X) : \langle \nabla g_i(\bar{x}(1)), z(1)\rangle + \frac{1}{2}\langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle < 0\Big\}, \quad i = 0, \dots k,$$
$$\mathscr{Q}^2 = \bigcap_{I_g} \mathscr{Q}^2_i\,,$$
$$I_g = \left\{i = 1, ..., k : g_i(\bar{x}(1)) = 0\,, \langle \nabla g_i(\bar{x}(1)), \bar{y}(1)\rangle = 0\right\},$$
$$\mathscr{K}^2_j = \Big\{z \in \mathscr{C}(I,X) : \max_{t\in\mathscr{M}_{j0}} \Big(\langle \nabla\varphi_j(\bar{x}(t)), z(t)\rangle + \frac{1}{2}\langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle\Big) < 0\Big\}, \quad j = 1, \dots q,$$
$$\mathscr{K}^2 = \bigcap_{j=1}^{q} \mathscr{K}^2_j\,, \quad \mathscr{M}_{j0} = \left\{t \in I : \bar{x}(t) \in \partial K_j\right\}.$$

Notice that $\mathscr{M}_{j0}$ is closed and

$$\langle \nabla\varphi_j(\bar{x}(\cdot)), z(\cdot)\rangle + \frac{1}{2}\langle \varphi_j''(\bar{x}(\cdot))\bar{y}(\cdot), \bar{y}(\cdot)\rangle \in \mathscr{C}(I,\mathbb{R})\,.$$

Proposition 3.1 below guarantees the desired properties: $\mathscr{S}^2 \subset T^{\flat(2)}_{\mathscr{S}}(\bar{x},\bar{y})$, $\mathscr{Q}^2 \subset T^{\flat(2)}_{\mathscr{Q}}(\bar{x},\bar{y})$, $\mathscr{K}^2 \subset T^{\flat(2)}_{\mathscr{K}}(\bar{x},\bar{y})$, and

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

III. THE MAIN RESULTS

Let $(\bar{x}, \bar{u})$ be a local minimizer for problem (1)–(4), namely there exists $\varepsilon > 0$ such that

$$g_0(x(1)) \geq g_0(\bar{x}(1))\,,$$

for any admissible x satisfying $\|x - \bar{x}\|_{L^\infty(I,X)} \leq \varepsilon$, and let $\bar{y} \in \mathscr{V}^1(\bar{x})$ be critical. We prove below three results dealing with second order conditions. The first is an abstract result which implies, as a direct consequence, the validity of (5), for any $z \in \mathscr{S}^2 \cap \mathscr{Q}^2 \cap \mathscr{K}^2 \subset T^{\flat(2)}_{\mathscr{S}\cap\mathscr{Q}\cap\mathscr{K}}(\bar{x},\bar{y})$. Then, using a separation theorem and duality arguments we deduce some more concrete second order conditions together with the validity of the Pontryagin minimum principle, see Proposition 3.2 and Theorem 3.1 below.

Proposition 3.1: Assume (H) and that

$$\exists\, \delta > 0 \;:\; \max_{t\in\mathscr{M}_{j\delta}} \langle \nabla\varphi_j(\bar{x}(t)), \bar{y}(t)\rangle \leq 0\,, \quad \forall\, j = 1, \dots q\,, \quad (9)$$

where $\mathscr{M}_{j\delta} = \left\{t \in [0,1] : \varphi_j(\bar{x}(t)) \geq -\delta,\ d_{\partial K_j}(\bar{x}(t)) \leq \delta\right\}$. Then,

$$\mathscr{S}^2 \cap \mathscr{Q}^2 \cap \mathscr{K}^2 \subset T^{\flat(2)}_{\mathscr{S}\cap\mathscr{Q}\cap\mathscr{K}}(\bar{x},\bar{y}). \quad (10)$$

Moreover,

$$\mathscr{S}^2 \cap \mathscr{Q}^2_0 \cap \mathscr{Q}^2 \cap \mathscr{K}^2 = \emptyset\,. \quad (11)$$

Proof: Let $z \in \mathscr{S}^2 \cap \mathscr{Q}^2 \cap \mathscr{K}^2$ and let $h_n \to 0^+$ as $n \to +\infty$. By Proposition 5.1, there exists a sequence $x_n = \bar{x} + h_n\bar{y} + h_n^2 z_n \in \mathscr{S}$, such that $z_n \to z$ uniformly. We need to show that, for any n large enough, x_n satisfies the constraints (2) and (3). Let $j \in \{1, \dots, q\}$. Since $z \in \mathscr{K}^2$, from the regularity of φ_j we deduce the existence of $\delta > 0$ such that

$$\max_{t\in\mathscr{M}_{j\delta}} \Big(\langle \nabla\varphi_j(\bar{x}(t)), z(t))\rangle + \frac{1}{2}\langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle\Big) < 0\,. \quad (12)$$

Hence, by Taylor expansion and (12), for any n large enough and any $t \in \mathscr{M}_{j\delta}$,

$$\begin{aligned}
\varphi_j(x_n(t)) &= \varphi_j(\bar{x}(t)) + h_n\langle \nabla\varphi_j(\bar{x}(t)), \bar{y}(t)\rangle \\
&\quad + h_n^2\langle \nabla\varphi_j(\bar{x}(t)), z_n(t)\rangle + h_n^2\frac{1}{2}\langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle + h_n^2 r_n(t) \\
&\leq \varphi_j(\bar{x}(t)) + h_n^2\Big(\langle \nabla\varphi_j(\bar{x}(t)), z_n(t)\rangle \\
&\quad + \frac{1}{2}\langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle\Big) + h_n^2 r_n(t) < 0\,,
\end{aligned}$$

where $r_n(\cdot) \to 0$ uniformly, as $n \to \infty$. If $t \notin \mathscr{M}_{j\delta}$, we have that either $\varphi_j(\bar{x}(t)) < -\delta$ or $d_{K_j}(\bar{x}(t)) > \delta$, and we obtain again $x_n(t) \in K$. Hence, $x_n(I) \subset K$, for all large n.

We now consider the constraint Q. If $i \notin I_g$, then either $g_i(\bar{x}(1)) < 0$ or

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

The Taylor expansion yields, 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 definition of $\mathscr{Q}^2_i$, applying the Taylor expansion again, we obtain

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

for n large enough, implying $x_n(1) \in Q$. We can conclude that $x_n \in T^{\flat(2)}_{\mathscr{S}\cap\mathscr{Q}\cap\mathscr{K}}(\bar{x},\bar{y})$, showing (10).
In order to prove (11), suppose by contradiction that there exists

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

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)–(3) such that $z_n \to z$ uniformly. Reasoning as above, replacing g_i with g_0 and taking $\tilde{x} = x_n$, for some n sufficiently large, we obtain the existence of a solution $\tilde{x}$ of (1)–(3) such that

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

This contradicts the optimality of $\bar{x}$. ■

As a consequence of Proposition 3.1, we deduce the second order necessary condition in the form of the variational inequality (5).

Now, applying a separation theorem, we can analyze (11) and obtain sharper conditions, as outlined in the results below.

Proposition 3.2: Let $\bar{y} \in \mathscr{V}^1(\bar{x})$ be critical and assume (H), (9) and that

$$\nabla\varphi_j(\bar{x}(t)) \neq 0, \quad \forall t \in \mathscr{M}_{j0},\ \forall j = 1,\dots,q. \tag{13}$$

Then, for every convex nonempty subset $\tilde{\mathscr{S}}^2 \subset \overline{\mathscr{S}^2}$ there exist $\lambda_i \geq 0$, for $i \in I_g$, positive $\psi_j \in \mathscr{M}(I,\mathbb{R})$ with $\operatorname{supp}\psi_j \subset \mathscr{M}_{j0}$, for $j = 1,\dots,q$, not vanishing simultaneously, and $x^* \in \mathscr{C}(I,X)^*$, such that, for any $z \in \mathscr{C}(I,X)$,

$$\sum_{i\in I_g} \lambda_i \langle \nabla g_i(\bar{x}(1)), z(1)\rangle + \sum_{j=1}^{q} \int_I \langle \nabla\varphi_j(\bar{x}(t)), z(t)\rangle d\psi_j(t) = \langle x^*, z\rangle \tag{14}$$

and

$$\inf\langle x^*, \tilde{\mathscr{S}}^2\rangle + \frac{1}{2}\sum_{i\in I_g} \lambda_i \langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle + \frac{1}{2}\sum_{j=1}^{q} \int_I \langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle d\psi_j(t) \geq 0. \tag{15}$$

Moreover, if

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

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

The proof of Proposition 3.2 follows from a separation theorem and density properties. Because of the lack of space we postpone it to [13].

To state our third result involving the second order necessary conditions and the Pontryagin minimum principle, let us consider the linearized system

$$\begin{cases} \dot{y}(t) = \mathbb{A}y(t) + f_x[t]y(t) + f_u[t]v(t) \\ y(0) = 0, \end{cases} \tag{16}$$

where

$v(\cdot)$ is a measurable selection of the set-valued map

$$t \rightsquigarrow T_U^{\flat}(\bar{u}(t)). \tag{17}$$

Assume that, for any $R > 0$ there exists $\psi_R > 0$ such that

$$\|f(t,x,u) - f(t,x,v)\|_X \leq \psi_R \|u - v\|_Z, \tag{18}$$

for a.e. $t \in I$ and any $u, v \in U$ and $x \in RB$. Then, it is not difficult to prove that a solution $(\bar{y},\bar{v})$ of (16)–(17), with $\bar{v} \in L^\infty(I,Z)$, solves (6). Assume further that

$$f''[\cdot](\bar{y}(\cdot),\bar{v}(\cdot))^2 \in L^1(I,X), \tag{19}$$

and consider the second order linearization

$$\begin{cases} \dot{z}(t) = \mathbb{A}z(t) + f_x[t]z(t) + f_u[t]w(t) \\ \qquad\qquad + \frac{1}{2}f''[t](\bar{y}(t),\bar{v}(t))^2 + \eta(t) \\ z(0) = 0, \end{cases} \tag{20}$$

with

$$\eta \in L^1(I,X),\ \eta(t) \in T_{F(t,\bar{x}(t))}(f^{\bar{x}}(t)),\ \text{a.e. in } I. \tag{21}$$

and $w \in \mathscr{W}$, where

$$\mathscr{W} = \{w : I \to Z \text{ measurable} : f_u[\cdot]w(\cdot) \in L^1(I,X), \quad w(t) \in T_U^{\flat(2)}(\bar{u}(t),\bar{v}(t)) \text{ a.e. in } I\}. \tag{22}$$

Theorem 3.1: Assume (H), (13), and (18). Let $(\bar{x},\bar{u})$ be a local minimizer and $(\bar{y},\bar{v})$ be a solution of (16)–(17) satisfying (19) and $\bar{v} \in L^\infty(I,Z)$. Assume that $\bar{y}$ is a critical admissible first order variation satisfying (9) and $\mathscr{W} \neq \emptyset$. Then, there exist $\lambda_i \geq 0$, for $i \in I_g$, and positive $\psi_j \in \mathscr{M}(I,\mathbb{R})$ with $\operatorname{supp}\psi_j \subset \mathscr{M}_{j0}$, for $j = 1,\dots,q$, not all equal to zero such that the function $p : I \to X^*$ defined by

$$p(t) = \mathscr{T}(1,t)^*\Big(\sum_{i\in I_g} \lambda_i \nabla g_i(\bar{x}(1))\Big) + \int_t^1 \mathscr{T}(s,t)^*\Big(\sum_{j=1}^{q} \nabla\varphi_j(\bar{x}(s))d\psi_j(s)\Big), \tag{23}$$

with $\mathscr{T}$ the solution operator associated with

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

satisfies the minimum principle

$$\langle p(t), f(t,\bar{x}(t),\bar{u}(t))\rangle = \min_{u\in U}\langle p(t), f(t,\bar{x}(t),u)\rangle. \tag{24}$$

Further, for any $w \in \mathscr{W}$, the following second order condition holds

$$\frac{1}{2}\sum_{i\in I_g} \lambda_i \langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle + \frac{1}{2}\sum_{j=1}^{q} \int_I \langle \varphi_j''(\bar{x}(s))\bar{y}(s), \bar{y}(s)\rangle d\psi_j(s) + \frac{1}{2}\int_0^1 \langle p(s), f''[s](\bar{y}(s),\bar{v}(s))^2\rangle ds + \int_0^1 \langle p(s), f_u[s]w(s)\rangle ds \geq 0. \tag{25}$$

If X is reflexive, then, by [9], p is the mild solution of the adjoint measure-driven equation

$$\begin{cases} dp(t) = -\big(\mathbb{A}^* + f_x(t,\bar{x}(t),\bar{u}(t))^*\big)p(t)dt \\ \qquad\qquad - \sum_{j=1}^{q} \nabla\varphi_j(\bar{x}(t))d\psi_j(t) \\ p(1) = \sum_{i=0}^{k} \lambda_i \nabla g_i(\bar{x}(1)). \end{cases} \tag{26}$$

Notice that, under the assumptions of Theorem 3.1, the set

$$\tilde{\mathscr{S}}^2 := \overline{\mathscr{L}^2}^{\mathscr{C}(I,X)}, \tag{27}$$

where

$$\mathscr{L}^2 = \big\{z \in \mathscr{C}(I,X) : z \text{ solves (20) for } w, \eta \text{ as in (22),(21)}\big\}$$

is convex and nonempty, the technical details can be found in [13].

Proof: (Proof of Theorem 3.1.) Let $\tilde{\mathscr{S}}^2$ be defined as in (27) and let λ_i, for $i \in I_g$, ψ_j, for $j = 1, \dots, q$, and $x^* \in \mathscr{C}(I,X)^*$ be as in Proposition 3.2. Let $z \in \mathscr{L}^2$. Then,

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

where

$$\beta(t) = f_u[t]w(t) + \frac{1}{2}f''[t](\bar{y}(t), \bar{v}(t))^2$$

for some w and η as in (22) and (21). Applying an integration by parts, see Lemma 4.1. in [8], we obtain from (14):

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

Then, from (15) we get

$$\begin{aligned}
&\frac{1}{2}\sum_{i=0}^k \lambda_i \langle g_i''(\bar{x}(1))\bar{y}(1), \bar{y}(1)\rangle \\
&\quad + \frac{1}{2}\sum_{j=1}^q \int_I \langle \varphi_j''(\bar{x}(t))\bar{y}(t), \bar{y}(t)\rangle d\psi_j(t) \\
&\quad + \frac{1}{2}\int_0^1 \langle p(s), f''[s](\bar{y}(s), \bar{v}(s))^2\rangle ds \\
&\quad + \inf_{\{w \text{ as in } (22)\}} \int_0^1 \langle p(s), f_u(s,\bar{x}(s),\bar{u}(s))w(s)\rangle ds \\
&\quad + \inf_{\{\eta \text{ as in } (21)\}} \int_0^1 \langle p(s), \eta(s)\rangle ds \geq 0,
\end{aligned}$$

yielding (25) and

$$\int_0^1 \langle p(s), \eta(s)\rangle ds \geq 0, \qquad \text{for any } \eta \text{ as in (21)}, \tag{28}$$

because the set of functions η satisfying (21) is a cone. As in [12, Theorem 4.2], (28) implies (24). ■

IV. APPLICATIONS: CONTROL PROBLEMS INVOLVING PDEs

Our analysis is performed in great generality, so a large class of concrete models can be considered. For lack of space, we propose here only one example of optimal control problem involving a heat equation.

Example 4.1 (A controlled heat equation): We analyze a control system describing a heat transfer problem. A similar problem (without state constraints) has been considered in [1], dealing with second order conditions, and it has been studied in [12] to get first order state constrained necessary conditions. Given $\Omega \subset \mathbb{R}^N$, a bounded domain with smooth boundary $\partial\Omega$, we consider a heat equation where the heat supply is represented by a multiplicative control:

$$\partial_t x(t,\mathbf{x}) - \Delta x(t,\mathbf{x}) = \varphi(t,\mathbf{x}) + u(t)b(\mathbf{x})x(t,\mathbf{x}). \tag{29}$$

Here $\varphi \in L^1(I, L^2(\Omega))$, $b \in L^\infty(\Omega)$, $x = x(t,\mathbf{x})$ is the temperature distribution, a function of the time $t \in I = [0,1]$ and the position $\mathbf{x} \in \Omega$, the control u takes values in the closed interval $U = [c,d]$ of $\mathbb{R}$, where $c < d$. Below we omit writing explicitly the dependence on the variable $\mathbf{x}$. Equation (29) is endowed with Dirichlet boundary conditions and initial condition $x(0) = x_0 \in L^2(\Omega)$.

In order to handle (29) as system (1) and implement our abstract machinery, 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 $S(t)$ on $X = L^2(\Omega)$. Thus, (29) can be written as the abstract system (1) with $f(t,x,u) = \varphi(t) + ubx$.

Our aim is to find a temperature x to be close, at the final time $t = 1$, to a reference temperature $x_D \in X$, namely we want 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

$$K = \left\{x \in X : \|x\|_X^2 - 1 \leq 0\right\}$$

and the end point constraint

$$Q = \left\{x \in X : \|x - x_1\|_X^2 - r \leq 0\right\},$$

for some fixed $x_1 \in L^2(\Omega)$ and $r > 0$. It is not difficult to prove that assumption (H) is satisfied.

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

$$\begin{cases} \dot{y}(t) = \mathbb{A}y(t) + b\bar{u}(t)y(t) + bv(t)\bar{x}(t) \\ y(0) = 0\,, \end{cases} \tag{30}$$

satisfying all the assumptions of Theorem 3.1. Since X is a reflexive space, the function p defined as in (23) is a mild solution of the adjoint equation

$$\begin{cases} dp(t) = -\big(\mathbb{A} + b\bar{u}(t)\big)p(t)dt - 2\bar{x}(t)d\psi(t)\,, \quad t \in I \\ p(1) = \lambda_0(\bar{x}(1) - x_D) + \lambda_1 2(\bar{x}(1) - x_1)\,, \end{cases} \tag{31}$$

for some positive $\psi \in \mathscr{M}(I, \mathbb{R}^+)$ and $\lambda_0, \lambda_1 \geq 0$ satisfying the properties stated in Theorem 3.1. Then, the following second order optimality condition holds

$$\begin{aligned}
&(\lambda_0 + 2\lambda_1)\|\bar{y}(1)\|_X^2 \\
&+ \int_I 2\|\bar{y}(t)\|_X^2 d\psi(t) + \int_0^1 \langle p(s), b\bar{y}(s)\bar{v}(s)\rangle ds \\
&+ \inf_{\{w \in L^1(I,X) \text{ satisfying } (22)\}} 2\int_0^1 \langle p(s), b\bar{x}(s)w(s)\rangle ds \geq 0\,,
\end{aligned}$$

together with the minimum principle

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

It is not difficult to prove that, if $\langle p(t), b\bar{x}(t)\rangle > 0$, then $\bar{u}(t) = c$ and, if $\langle p(t), b\bar{x}(t)\rangle < 0$, then $\bar{u}(t) = d$, a.e. in I.

V. APPENDIX

This section contains a technical result dealing with second order variations, needed in the proof of Proposition 3.1.

Proposition 5.1: Assume (H) (i)–(iii). Let $x \in \mathscr{S}$, y be a first order variation and z be a second order variation at (x,y). Then, for any $h_n \to 0^+$, there exist $x_n \in \mathscr{S}$ such that

$$\frac{x_n - x - h_n y}{h_n^2} \to z, \quad \text{uniformly on } I \text{ as } n \to \infty.$$

Proof: Set

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

We prove first that there exists a solution x_n^2 to

$$\dot{x}(t) \in \mathbb{A}x(t) + F(t, x(t)),$$

such that

$$\frac{x_n^1 - x_n^2}{h_n^2} \to 0, \quad \text{uniformly on } I \text{ as } n \to \infty. \tag{32}$$

Let π^y be as in (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 the second order variation, we obtain that, for a.e. $t \in I$,

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

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 obtain that, for a.e. t,

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

Further, using assumption (H) (ii) and (7) we get for a.e. $t \in I$,

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

with $R > 0$ such that, for every $n \ge 1$,

$$\|x(t) + h_n y(t)\|_X + \|x(t) + h_n y(t) + h_n^2 z(t)\|_X \le R.$$

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 [11], we obtain the claimed x_n^2 satisfying (32). A relaxation theorem, see [10], guarantees the existence of $x_n \in \mathscr{S}$ such that

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

Finally, (32) allows to end the proof. ■

VI. ACKNOWLEDGMENTS

The authors gratefully acknowledge remarks of reviewers that helped to improve the preliminary version of this manuscript.

REFERENCES

[1] M.S. Aronna, J.F. Bonnans and A. Kröner, Optimal control of infinite dimensional bilinear systems: application to the heat and wave equations, *Math. Program., Ser. B*, vol. 1-2, 2018, pp 717-757.

[2] J.F. Bonnans and A. Shapiro, *Perturbation analysis of optimization problems*, Springer, New York, 2000.

[3] E. Casas, J.C. De Los Reyes and F. Tröltzsch, Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints, *SIAM J. Control Optim.*, vol. 19, 2008, pp 616-643.

[4] E. Casas and F. Tröltzsch, Second order necessary and sufficient optimality conditions for optimization problems and applications to control theory, *SIAM J. Control Optim.*, vol. 13, 2002, pp 406-431.

[5] E. Casas and F. Tröltzsch, Second order analysis for optimal control problems: improving results expected from abstract theory, *SIAM J. Control Optim.*, vol. 22, 2012, pp 261-279.

[6] E. Casas and F. Tröltzsch, Second order optimality conditions and their role in PDE control, *Jahresber. Dtsch. Math. Ver.*, vol. 117, 2015, pp 3-44.

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

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

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

[10] H. Frankowska, *A priori estimates for operational differential inclusions*, J. Differential Equations, vol. 84, 1990, pp 100-128.

[11] H. Frankowska, E.M. Marchini, M. Mazzola, A relaxation result for state constrained inclusions in infinite dimension, *Math. Control Relat. Fields*, vol. 6, 2016, pp 113-141.

[12] H. Frankowska, E.M. Marchini, M. Mazzola, Necessary optimality conditions for infinite dimensional state constrained control problems, *J. Differential Equations*, vol. 264, 2018, pp 7294-7327.

[13] H. Frankowska, E.M. Marchini and M. Mazzola, A unified approach to first and second order necessary conditions in optimal control for PDEs, preprint, 2019.

[14] H. Frankowska and N.P. Osmolovskii, Second-order necessary optimality conditions for a strong local minimum in a control problem with general control constraints, *Appl. Math. Optim.*, vol. 80, 2019, pp 135-164.

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

[16] K. Krumbiegel and J. Rehberg, Sufficient second-order optimality conditions for parabolic optimal control problems with pointwise state constraints,, *SIAM J. Control Optim.*, vol. 51, 2013, pp 304-331.

[17] P.S. Lomdhal, O.H. Soerensen and P.L. Christiansen, Soliton excitations in Josephson tunnel junctions, *Phys. Rev. B*, vol. 25, 1982, pp 5737-5748.

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

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

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

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

[22] R.B. Vinter, *Optimal control*, Birkhäuser, Boston, Basel, Berlin, 2000.