



Minimax optimal control problem with state constraints



Dmitry Karamzin^{a,b,c,*}, Valeriano de Oliveira^d, Fernando Pereira^e, Geraldo Silva^d

^a Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Vavilova Street, 40, Moscow 119333, Russia

^b RUDN University, Miklukho-Maklaya Street, 6, Moscow 117198, Russia

^c SYSTEC/Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal (Visiting researcher)

^d Instituto de Biociências, Letras e Ciências Exatas, UNESP – Universidade Estadual Paulista, Rua Cristóvão Colombo, N. 2265, CEP 15054-000, São José do Rio Preto – SP, Brazil

^e SYSTEC/Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, Portugal

ARTICLE INFO

Article history:

Received 31 March 2016

Received in revised form

30 May 2016

Accepted 20 June 2016

Recommended by A. Zemouche

Available online 9 July 2016

MSC:

49K15

49K21

49K35

Keywords:

Minimax control

State constraints

Maximum principle

ABSTRACT

In this article, nondegenerate necessary conditions of optimality are derived and discussed for the so-called “minimax” optimal control problems with state constraints. In this class of problems, the data depends on an unknown parameter that takes values in a given compact set called “uncertainty set”. The solution of the problem is considered for the worst case value of the unknown parameter. The result is obtained in the two stages: after the basic necessary conditions of optimality are derived, additional conditions ensuring their nondegeneracy in the minimax context are obtained.

© 2016 European Control Association. Published by Elsevier Ltd. All rights reserved.

1. Introduction

In this article, we study a minimax optimal control problem with state constraints. As a starting point of our investigation, we consider the article [33] by R.B. Vinter, where a maximum principle of the Pontryagin type for an optimal control problem involving a set of unknown parameters is derived. We follow the path traversed by the author of [33] by extending his methodology and results to the case in which state constraints are present. The consideration of this case involves some challenges and adds a novel flavor to the arguments.

From a mathematical point of view, the investigation of Necessary Conditions of Optimality (NCO) for “minimax” control processes is complicated by a certain degeneracy effect associated with the endpoint constraints. This phenomenon has been described in detail in [33]. Therein, a number of examples are also

given showing that the NCO do not generally hold true for certain types of endpoint constraints, such as, for example, equality constraints imposed on the both endpoints. In the proof, this is reflected by the fact that the set of normalized Lagrange multipliers becomes non-convex which does not allow the application of standard techniques to find measurable (w.r.t. the unknown parameter) Lagrange multipliers.

Nevertheless, from an engineering point of view, the uncertainty parameter implies that it does not make a lot of sense to consider fixed endpoints or even some equality constraints imposed on the both endpoints in a “minimax” problem. Moreover, the endpoint constraint set S (see the problem statement below) should generally be “sufficiently good” and large enough, in order to exhibit at least a nonzero Lebesgue measure. Indeed, a relevant perturbation in the uncertainty parameter should lead to a perturbation of the final state trajectory value (otherwise we do not have uncertainty in a practical sense). But, if the terminal point, or some of its coordinates, is fixed, then we may easily incur in constraint violation and a lack, or even total absence, of feasibility. Then, the problem statement becomes ill-posed. So, it is natural to consider that the set S , or the so-called “aiming area”, is given by only the inequality type of constraints. This is one of the reasons why we do not address the above issue of

* Corresponding author at: Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Vavilova Street, 40, Moscow 119333, Russia.

E-mail addresses: dmitry_karamzin@mail.ru (D. Karamzin), antunes@ibilce.unesp.br (V.d. Oliveira), flp@fe.up.pt (F. Pereira), gsilva@ibilce.unesp.br (G. Silva).

degeneracy related to the endpoint constraint in this particular work (which is surely a rather complex and interesting issue in general). On the other hand, another type of degeneracy associated with the state constraints (which are highly relevant in various classes of engineering applications) is addressed.

One of the distinguishing features of NCO for minimax problems is that the maximum condition and the Hamiltonian become, distributed over the set of unknown parameter values. In what follows, we consider the case when the uncertainty parameter w does not depend on time. It is worth remarking that, from the engineering perspective, these two cases, the time-dependent and the time-constant w , are inherently different. While the time-constant case can usually be found in systems whose design is such that perturbations are in general not affected by natural processes (and thus are fully human “determined”) during the life cycle of the system, the time-varying case appears due to natural causes, such as wind, and turbulence, etc. Then, the role of the control is to ensure some kind of robust optimality in the natural environment. The time-varying case can also be addressed in the context of a generalization of differential games [35,37], notably, in the case of two players controlling the system with conflicting objectives. However, it is important to emphasize that the minimax control problem that we consider here and the class of problems designated by differential games are inherently different in their formulations.

Regarding the study of state constraints, the authors are not aware whether this type of constraints had been investigated before in the context of NCO for minimax control problems. However, the state constraints have been appearing in the literature in the context of robust control as a whole (see, e.g., [25] and the bibliography cited therein).

The article is organized as follows. In Section 2 we present and discuss the problem formulation. The NCO of the Pontryagin type in the framework of the Gamkrelidze approach are presented, discussed, and proved for different instances of increasing complexity in Section 3. The issues concerning the nondegeneracy of the optimality conditions are addressed in Section 4. An example illustrating the usefulness of the controllability conditions in the minimax framework is provided. Finally, in the last section, a brief outline of the key conclusions is exposed.

2. Problem statement

Consider the minimax optimal control problem

$$\begin{aligned} \text{Minimize } & \max_{w \in C} \{\varphi(x_1, w)\}, \\ \text{subject to } & \dot{x} = f(x, u, w), \\ & x(0) = x_0, \quad x(1) = x_1 \in S, \\ & u(t) \in U \quad \text{a.a. } t \in [0, 1], \\ & g(x(t), w) \leq 0 \quad \forall t \in [0, 1]. \end{aligned} \quad (1)$$

Hereinafter, $\dot{x} = \frac{dx}{dt}$, where $t \in [0, 1]$ denotes time, x is the state variable with values in the n -dimensional Euclidean space \mathbb{R}^n , $u \in U \subset \mathbb{R}^m$ is the value of the control function, and $w \in \mathbb{R}^l$ is the so-called unknown parameter. The vector x_1 is termed endpoint. The maps

$$\varphi : \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}, \quad f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \rightarrow \mathbb{R}^n, \quad g : \mathbb{R}^n \times \mathbb{R}^l \rightarrow \mathbb{R}^k,$$

satisfy the following general hypothesis (which are, from now on, adopted without any reference to them):

Hypothesis 1. Maps φ, f are continuously differentiable w.r.t. x , map g is twice continuously differentiable w.r.t. x , and maps φ, f ,

and g together with all their partial derivatives w.r.t. x are continuous.

The sets C, S , and U are compact. The control function is a measurable function $u : [0, 1] \rightarrow U$. Note that the unknown parameter $w \in C$ does not depend on time. We denote the set of control functions by \mathcal{U} .

Let $u \in \mathcal{U}$ be a control function. A control process (x, u) comprises u and the family $\{x(t; w)\}_{w \in C}$ of arcs, satisfying for each $w \in C$, the equation

$$\dot{x}(t; w) = f(x(t; w), u(t), w) \quad \text{a.a. } t \in [0, 1].$$

A control process (x, u) is said to be feasible if the endpoint constraint $x(1; w) \in S$, with $x(0) = x_0$, and the state constraints

$$g(x(t; w), w) \leq 0 \quad \forall t \in [0, 1] \text{ are satisfied for each } w \in C$$

A control process (\bar{x}, \bar{u}) is said to be optimal if

$$\max_{w \in C} \{\varphi(x(1; w), w)\} \geq \max_{w \in C} \{\varphi(\bar{x}(1; w), w)\}$$

for all feasible control processes (x, u) .

Therefore, we first maximize the cost over the set of all unknown parameters which yields the worst possible performance case for the given control strategy $u(\cdot)$. Then, the result is minimized over all control strategies $u(\cdot)$ in order to obtain the solution or the so-called robust optimal control. The robust optimal control guarantees a certain minimal cost which is considered fair or acceptable for all the values of the uncertainty parameter w because it is the cost of the worst performance case.

3. The necessary conditions of optimality

In this section, we state and derive the NCO in the form of a Pontryagin maximum principle for the following three cases of increasing complexity w.r.t. the set C : singleton, finite, and infinite valued. The NCO considered here are derived in the framework of Gamkrelidze approach (for details, see Chapter 6 of [29] and also [22,30,13]).

Let us define some notation.

Consider the extended Hamilton–Pontryagin function

$$\bar{H}(x, u, w, \psi, \mu) := \langle \psi, f(x, u, w) \rangle - \langle \mu, \Gamma(x, u, w) \rangle,$$

where $\Gamma(x, u, w) := g'_x(x, w)f(x, u, w)$, and by h'_y we denote either the gradient or the Jacobian of the function h w.r.t. the variable y .

Throughout the article, $N_S(x)$ designates the limiting normal cone to the set S at point x in the sense proposed by Mordukhovich in [26]. The limiting subdifferential is denoted as usual by symbol ∂ [27].

3.1. The case of singleton set C

Since this case corresponds to the classical problem statement, throughout this subsection we omit the dependence of functions on the unknown parameter w . The following NCO in the form of a Pontryagin Maximum Principle can be derived with the help of the Variational Analysis technique [27,32] relying essentially on Ekeland's variational principle [19].

Theorem 1. Suppose that (\bar{x}, \bar{u}) is a solution to problem (1), in which the set C is a singleton, and φ is a merely Lipschitz continuous function.

Then, there exist Lagrange multipliers $\lambda \in [0, 1]$, $\psi \in W_{1,\infty}([0, 1])$, and $\mu \in BV([0, 1])$ such that

$$\lambda + \max_{t \in [0, 1]} |\psi(t)| + |\mu(0)| = 1,$$

$$\dot{\psi}(t) = -\bar{H}'_x(\bar{x}(t), \bar{u}(t), \psi(t), \mu(t)) \quad \text{a.a. } t \in [0, 1],$$

$$-\psi(1) \in \lambda \partial \varphi(\bar{x}_1) + N_S(\bar{x}_1),$$

$$\max_{u \in U} \{\bar{H}(\bar{x}(t), u, \psi(t), \mu(t))\} = \bar{H}(\bar{x}(t), \bar{u}(t), \psi(t), \mu(t)) \quad \text{a.a. } t \in [0, 1],$$

where the map $\mu = (\mu^1, \mu^2, \dots, \mu^k)$ satisfies the following properties:

- (a) for a given arbitrary subinterval $[a, b] \subseteq [0, 1]$, the function μ^j is constant on $[a, b]$, when $g^j(\bar{x}(t)) < 0 \forall t \in [a, b]$;
- (b) each function μ^j is nonincreasing;
- (c) for each j , $\mu^j(1) = 0$.

Another approach to derive this theorem is to use a certain variable change in the multipliers. Let us demonstrate how [Theorem 1](#) can be easily deduced from the corresponding result of Chapter 9 in [\[32\]](#). Let the multipliers (λ, p, q, η) , where $q(0) = p(0)$, satisfy Theorem 9.3.1 of [\[32\]](#). Consider the new multipliers:

$$\mu(t) := \int_{[t, 1]} d\eta \quad \forall t \in [0, 1], \quad \mu(1) = 0,$$

$$\psi(t) := q(t) + \mu(t)g'_x(\bar{x}(t)), \quad t \in [0, 1].$$

It is a straightforward task to verify from the conditions of Theorem 9.3.1 of [\[32\]](#) that the multipliers (λ, ψ, μ) satisfy [Theorem 1](#) (see [\[13, 6\]](#) for details).

Despite the fact that the NCO in Gamkrelidze form [\[22, 30\]](#) are closely related (due to the above change of the conjugate variable) to the ones given by A.Ya. Dubovitskii and A.A. Milyutin in [\[18\]](#), independently, by L. Neustadt in [\[28\]](#), and in its nonsmooth versions in [\[32, 15\]](#), the consideration of the extended Hamilton–Pontryagin function \bar{H} has some advantages in what concerns their application. A good example is the geodesic equation, see, e.g., [\[29, 12, 16\]](#). The form of the NCO in [Theorem 1](#) also turns out to be convenient to study the continuity of the Lagrange multiplier $\mu(t)$ under various regularity assumptions, see [\[30, 3, 14, 36, 4\]](#). Such a property is important in some applications, see, e.g., [\[1\]](#). Another interesting outcome follows from the study of conditions for nondegeneracy and the conservation law in the presence of state constraints. Moreover, the extended maximum condition often helps to relate ψ and μ , to be more precise, it helps to express μ through ψ under a natural regularity condition. Note that the NCO involving the extended Hamilton–Pontryagin function require the extra hypothesis of the existence of the second order derivative for the state constraint function. This is the price to pay for an additional useful information that we may expect from these NCO.

The idea to derive the Gamkrelidze-like NCO (like in [Theorem 1](#)) from the NCO in the Dubovitskii–Milyutin–Neustadt form was first proposed in [\[13\]](#). The direct proof of Gamkrelidze-like NCO involving a penalization technique and the application of Ekeland's variational principle can be found in, e.g., [\[12, 7\]](#). The approach based on the variational principles has been extensively used in Variational Analysis, see books [\[27, 32, 15\]](#). Note that, along with a commonly used non-smooth penalization, one could also employ smooth penalty functions. Smooth penalization is highly relevant to derive second-order NCO for optimal control problems [\[31, 2, 8, 23\]](#).

3.2. The case of finite set C

Let $C = \{w_1, w_2, \dots, w_N\}$. The problem [\(1\)](#) with finite set C turns out easy to treat because it can be reduced to a conventional optimal control problem, through a bookkeeping scheme yielding a certain non-smooth control problem [\[33\]](#). Consider N state variable components $x^{(i)}$, $i = 1, \dots, N$, each of them in \mathbb{R}^n .

Consider the following reduced problem (P_{aux}) equivalent to [\(1\)](#):

$$\text{Minimize } \max\{\varphi(x_1^{(1)}, w_1), \varphi(x_1^{(2)}, w_2), \dots, \varphi(x_1^{(N)}, w_N)\},$$

$$\text{subject to } \dot{x}^{(i)} = f(x^{(i)}, u, w_i),$$

$$x^{(i)}(0) = x_0, \quad x^{(i)}(1) = x_1^{(i)} \in S,$$

$$u(t) \in U \quad \text{a.a. } t \in [0, 1],$$

$$g(x^{(i)}(t), w_i) \leq 0 \quad \forall t \in [0, 1], \quad i = 1, \dots, N.$$

Let us apply [Theorem 1](#) to the reduced problem (P_{aux}) . By using non-smooth calculus, the subgradient of the max function yields the existence of N non-negative numbers α_i such that $\sum_{i=1}^N \alpha_i = 1$, a number $\lambda \geq 0$, and functions $\psi^{(i)} \in W_{1,\infty}([0, 1])$, $\mu^{(i)} \in BV([0, 1])$, for $i = 1, \dots, N$, such that

$$\lambda + \sum_{i=1}^N \max_{t \in [0, 1]} |\psi^{(i)}(t)| + \sum_{i=1}^N |\mu^{(i)}(0)| = 1,$$

$$\begin{aligned} \dot{\psi}^{(i)}(t) &= -\bar{H}'_x(\bar{x}^{(i)}(t), \bar{u}(t), w_i, \psi^{(i)}(t), \mu^{(i)}(t)) \quad \text{a.a. } t \in [0, 1], \\ -\psi^{(i)}(1) &\in \lambda \alpha_i \varphi'_x(\bar{x}_1^{(i)}, w_i) + N_S(\bar{x}_1^{(i)}), \end{aligned}$$

and

$$\begin{aligned} \max_{u \in U} \left\{ \sum_{i=1}^N \bar{H}(\bar{x}^{(i)}(t), u, w_i, \psi^{(i)}(t), \mu^{(i)}(t)) \right\} \\ = \sum_{i=1}^N \bar{H}(\bar{x}^{(i)}(t), \bar{u}(t), w_i, \psi^{(i)}(t), \mu^{(i)}(t)). \end{aligned}$$

Here, the control process $((\bar{x}^{(1)}, \bar{x}^{(2)}, \dots, \bar{x}^{(N)}), \bar{u})$ is optimal.

Let us designate $\lambda_i = \lambda \alpha_i$, and let

$$c_i = \lambda_i + \max_{t \in [0, 1]} |\psi^{(i)}(t)| + |\mu^{(i)}(0)|.$$

By introducing the probability measure

$$\Lambda := \sum_{i=1}^N c_i \delta(w_i),$$

and by dividing $\lambda_i, \psi^{(i)}$ and $\mu^{(i)}$ by c_i for i such that $c_i > 0$, we rewrite the above conditions in an equivalent form in order to obtain the following result.

Theorem 2. Suppose that (\bar{x}, \bar{u}) is a solution to [\(1\)](#), in which the set C is finite. Then, there exist a probability measure Λ on C , and, for each $w \in C$, Lagrange multipliers $\lambda(w) \in [0, 1]$, $\psi(\cdot; w) \in W_{1,\infty}([0, 1])$, and $\mu(\cdot; w) \in BV([0, 1])$ such that

$$\lambda(w) + \max_{t \in [0, 1]} |\psi(t; w)| + |\mu(0; w)| = 1, \quad \Lambda\text{-a.e.},$$

$$\begin{aligned} \dot{\psi}(t; w) &= -\bar{H}'_x(\bar{x}(t; w), \bar{u}(t), w, \psi(t; w), \mu(t; w)) \quad \text{a.a. } t \in [0, 1], \\ -\psi(1; w) &\in \lambda(w) \varphi'_x(\bar{x}(1; w), w) + N_S(\bar{x}(1; w)), \quad w \in C, \end{aligned}$$

$$\begin{aligned} \max_{u \in U} \int_C \bar{H}(\bar{x}(t; w), u, w, \psi(t; w), \mu(t; w)) d\Lambda \\ = \int_C \bar{H}(\bar{x}(t; w), \bar{u}(t), w, \psi(t; w), \mu(t; w)) d\Lambda \quad \text{a.a. } t \in [0, 1], \end{aligned}$$

where the map $\mu(t; w)$ satisfies, for each w , the following properties:

- (a) for a given arbitrary subinterval $[a, b] \subseteq [0, 1]$, the function $\mu^j(\cdot; w)$ is constant on $[a, b]$, if $g^j(\bar{x}(t; w), w) < 0 \forall t \in [a, b]$;
- (b) each function $\mu^j(\cdot; w)$ is non-increasing;
- (c) for each j , $\mu^j(1; w) = 0$.

Remark 1. In the absence of endpoint and state constraints, one can state in addition that

$$\text{supp}(\Lambda) = \arg \max_{w \in C} \varphi(\bar{x}_1, w).$$

However, this formula holds only for the simplest problem formulation and fails in the presence of constraints.

Remark 2. A mere verbatim repetition of the above arguments enables the extension of the above NCO for the case in which f

depends measurably on the time variable t . See [13] for the detailed hypothesis on the data.

3.3. The case of infinite set C

The study of this case is significantly complicated by a degeneracy effect in the NCO associated with the endpoint constraint set S [33]. This issue was mentioned in the Introduction. In order to avoid this degeneracy effect, we simplify the problem statement by additionally assuming that

$$S = \{x \in \mathbb{R}^n : e_j(x) \leq 0, j = 1, \dots, r\},$$

where $e_j : \mathbb{R}^n \rightarrow \mathbb{R}$ are given smooth functions. Thereby, from now on, we shall deal with inequality right endpoint constraints only.

Our main result is as follows.

Theorem 3. Suppose that (\bar{x}, \bar{u}) is a solution to (1), in which the set S is replaced by inequality constraints. Then, there exist a Radon probability measure Λ on C and, for each $w \in C$, Lagrange multipliers $\lambda(w) \in [0, 1]$, $\nu(w) \in [0, 1]^r : \langle \nu(w), e(\bar{x}(1; w)) \rangle = 0$, $\psi(\cdot; w) \in W_{1,\infty}([0, 1])$, and $\mu(\cdot; w) \in BV([0, 1])$ such that

$$\lambda(w) + \sum_{j=1}^r \nu_j(w) + \sum_{j=1}^k \mu_j^0(0; w) = 1, \quad \Lambda\text{-a.e.},$$

$$\begin{aligned} \dot{\psi}(t; w) &= -\bar{H}'_x(\bar{x}(t; w), \bar{u}(t; w), \psi(t; w), \mu(t; w)) \quad \text{a.a. } t \in [0, 1], \\ -\psi(1; w) &= \lambda(w)\varphi'_x(\bar{x}(1; w), w) \\ &\quad + \sum_{j=1}^r \nu_j(w)e'_{jx}(\bar{x}(1; w)), \quad w \in C, \end{aligned}$$

$$\begin{aligned} \max_{u \in U} \int_C \bar{H}(\bar{x}(t; w), u, w, \psi(t; w), \mu(t; w)) d\Lambda \\ = \int_C \bar{H}(\bar{x}(t; w), \bar{u}(t; w), w, \psi(t; w), \mu(t; w)) d\Lambda \quad \text{a.a. } t \in [0, 1], \end{aligned}$$

where the function μ satisfies the properties (a)–(c) of Theorem 2, the Lagrange multipliers λ and ν are Λ -measurable, ψ is Λ -measurable for all $t \in [0, 1]$, and μ is Λ -measurable for a.a. $t \in (0, 1]$, and for $t=0$.

In the proof, we need the following lemma (Proposition 6.1 in [33]). Consider some closed set E in a Euclidean space.

Lemma 1. Take a sequence of Radon probability measures Δ_i defined on the Borel subsets of E , such that $\Delta_i \xrightarrow{*} \Delta$ for some Radon probability measure Δ , and a sequence of set-valued maps $D_i : E \rightarrow 2^{\mathbb{R}^n}$ such that

$$\limsup_{i \rightarrow \infty} \text{Gr} D_i \subseteq \text{Gr} D$$

for some set-valued map $D : E \rightarrow 2^{\mathbb{R}^n}$ with compact graph $\text{Gr} D$. Take a sequence of Δ_i -measurable Borel functions $\gamma_i : E \rightarrow \mathbb{R}^n$ such that $\gamma_i(x) \in D_i(x)$ for a.a. $x \in E$ with respect to Δ_i .

If $D(x)$ is convex for all x , then there exists a Δ -measurable Borel function $\gamma : E \rightarrow \mathbb{R}^n$ such that $\gamma(x) \in D(x)$ for a.a. $x \in E$ w.r.t. Δ , and, along a subsequence,

$$\int_E \gamma_i(x) d\Delta_i \rightarrow \int_E \gamma(x) d\Delta.$$

The proof of this key lemma can be found in [32,34]. Note that, here, its formulation is simplified and adapted for the proof given below.

Proof of Theorem 3. First, let us observe that the NCO stated in the theorem degenerate whenever there exists $w \in C$ s.t. $g^j(x_0, w)$

$= 0$ for some j .¹ Therefore, it is not restrictive to assume that, $g^j(x_0, w) < 0 \forall w \in C, j=1, \dots, k$.

Pick any sufficiently large number c such that

$$c > \max_{w \in C} \max_{t \in [0,1]} |\bar{x}(t; w)|.$$

Let $\mathcal{F} \subseteq \mathcal{U}$ be the subset of all feasible controls, such that $u \in \mathcal{F}$ if and only if the control process (x, u) is feasible and

$$\max_{w \in C} \max_{t \in [0,1]} |x(t; w)| \leq c.$$

It is a straightforward task to verify that the set \mathcal{F} is closed with respect to the L_1 -topology and, therefore, it is a complete metric space (w.r.t. L_1 -metric).

Take a sequence $\varepsilon_i \rightarrow 0^+$. There is no loss of generality in assuming that

$$\max_{w \in C} \varphi(\bar{x}_1, w) = 0.$$

Let $\{w_i\}_{i=1}^\infty$ be a countable and everywhere dense in C set of points. In view of the optimality of \bar{x} , the compactness of C , and properties of continuity w.r.t. w , for each i , there is a number $N = N(i)$ such that

$$J_i(u) := \max_{w \in C_N} \varphi(x(1; w), w) + \varepsilon_i^2 > 0 \quad \forall u \in \mathcal{F},$$

where $C_N := \{w_1, w_2, \dots, w_N\}$.

It is clear that $J_i(u)$ is continuous with respect to the L_1 -metric, and $J_i(\bar{u}) \leq \varepsilon_i^2$. Then, from Ekeland's variational principle, it follows that, for each i , there exists a control function $u_i \in \mathcal{F}$ satisfying

$$\int_0^1 |u_i(t) - \bar{u}(t)| dt \leq \varepsilon_i,$$

and

$$J_i(u_i) = \min_{u \in \mathcal{F}} \left\{ J_i(u) + \varepsilon_i \int_0^1 |u(t) - u_i(t)| dt \right\}.$$

Thus, we obtain the following auxiliary optimal control problem (which is referred to as i -problem in what follows):

$$\text{Minimize} \quad \max_{w \in C_N} \varphi(x_1, w) + \varepsilon_i \int_0^1 |u(t) - u_i(t)| dt,$$

subject to $\dot{x} = f(x, u, w)$,

$$x(0; w) = x_0, \quad x(1; w) = x_1 \in S,$$

$$u(t) \in U \quad \text{a.a. } t \in [0, 1],$$

$$g(x(t; w), w) \leq 0, \quad |x(t; w)| \leq c \quad \forall t \in [0, 1], \quad w \in C_N,$$

whose minimizer is the control process (x_i, u_i) . It is clear that $u_i(t) \rightarrow \bar{u}(t)$ in L_1 , and $x_i(t; w) \rightarrow \bar{x}(t; w) \forall w \in C_N$.

The next step consists in applying the NCO from Theorem 2 to the i -problem. By taking into account Remark 2, the fact that the constraint $|x| \leq c$ is not active for i sufficiently large, for each $w \in C$, there exist Lagrange multipliers $\lambda_i(w)$, $\nu_i(w)$, $\psi_i(t; w)$, and $\mu_i(t; w)$ satisfying²:

$$\lambda_i(w) + \sum_{j=1}^r \nu_j^i(w) + \sum_{j=1}^k \mu_j^i(0; w) = 1, \quad \Lambda_i\text{-a.e.},$$

$$\dot{\psi}_i(t; w) = -\bar{H}'_x(x_i(t; w), u_i(t; w), \psi_i(t; w), \mu_i(t; w)) \quad \text{a.a. } t \in [0, 1],$$

¹ For details, see Section 4.

² The non-triviality condition is not the one of Theorem 2. However, it is achieved by rearranging Λ_i over C_N in a simple way (see the arguments of Theorem 2).

$$-\psi_i(1; w) = \lambda_i(w) \varphi'_x(x_i(1; w), w) + \sum_{j=1}^r \nu_j^i(w) e'_{jx}(x_i(1; w)),$$

$$\max_{u \in U} \int_C (\bar{H}(x_i(t; w), u, w, \psi_i(t; w), \mu_i(t; w)) - \lambda_i(w) \varepsilon_i |u - u_i(t)|) d\Lambda_i \\ = \int_C \bar{H}(x_i(t; w), u_i(t), w, \psi_i(t; w), \mu_i(t; w)) d\Lambda_i, \quad \text{a.a. } t,$$

where Λ_i is some probability measure on C with $\text{supp} \Lambda_i \subseteq C_N$, and μ_i satisfies the (a)–(c) conditions in which \bar{x} is replaced by x_i .

Above, the Lagrange multipliers are trivial for those w which are not in $\text{supp} \Lambda_i$.

Denote by $L(w)$ the set of the Lagrange multipliers with respect to w . By definition, the four-tuple

$$q = (\lambda, \nu, \mu(\cdot), \psi(\cdot)),$$

where $\lambda \in \mathbb{R}$, $\nu \in \mathbb{R}^r$, $\mu \in BV([0, 1])$ and $\psi \in W_{1,\infty}([0, 1])$, belongs to the set $L(w)$ if and only if the following conditions are satisfied:

$$\lambda + \sum_{j=1}^r \nu_j + \sum_{j=1}^k \mu^j(0) = 1, \quad \lambda \geq 0, \quad \nu \geq 0, \quad \langle \nu, e(\bar{x}(1; w)) \rangle = 0,$$

$$\dot{\psi}(t) = -\bar{H}'_x(\bar{x}(t; w), \bar{u}(t), w, \psi(t), \mu(t)) \quad \text{a.a. } t \in [0, 1],$$

$$-\psi(1) = \lambda \varphi'_x(\bar{x}(1; w), w) + \sum_{j=1}^r \nu_j^i e'_{jx}(\bar{x}(1; w)),$$

and $\mu(t)$ satisfies the conditions (a)–(c).

We may regard the set $L(w)$ as a set of functions

$$q(t) : [0, 1] \rightarrow \mathbb{R} \times \mathbb{R}^r \times \mathbb{R}^k \times \mathbb{R}^n,$$

being the above multipliers λ , and ν considered as constant functions with respect to t . Denote by \mathcal{L} the set of functions

$$q(t; w) : [0, 1] \times C \rightarrow \mathbb{R}^{1+r+k+n},$$

such that $q(t; w) \in L(w) \forall w \in C$, being

$$q(t; w) = (\lambda(w), \nu(w), \mu(t; w), \psi(t; w)).$$

By definition, the set \mathcal{L} is convex. (Here, we essentially used the fact that the endpoint constraints are given by inequalities only, and thereby, the corresponding multipliers ν^j are non-negative.)

Consider a countable and everywhere dense in \mathcal{U} family of control functions v_j , $j = 1, 2, \dots$. Take a positive integer N and the linear operators A_j , $j = 1, 2, \dots, N$, such that they map the functions of \mathcal{L} into scalar functions defined on C as follows:

$$A_j(q(\cdot; w)) = \int_0^1 (\bar{H}(\bar{x}(t; w), v_j(t), w, \psi(t; w), \mu(t; w)) \\ - \bar{H}(\bar{x}(t; w), \bar{u}(t), w, \psi(t; w), \mu(t; w))) dt.$$

Let us put $B_N = (A_1, A_2, \dots, A_N)$.

Now, define the map $D_N(t; w)$

$$D_N(t; w) := \left\{ (a, b) \in \mathbb{R}^{1+r+k+n} \times \mathbb{R}^N : \right. \\ \left. a = q(t; w), \quad b = B_N(q(\cdot; w)) \text{ for some } q \in \mathcal{L} \right\}.$$

Note that the set $D_N(t; w)$ is convex because \mathcal{L} is convex, while the operator B_N is linear. Thanks to Helly and Arzela–Ascoli theorems, and also due to compactness arguments and continuity of the data with respect to w , the set-valued map D_N is upper semicontinuous. Moreover it is bounded and thereby its graph $\text{Gr} D_N$ is a compact set.

Similarly, define the set $L_i(w)$ of the Lagrange multipliers for the i -problem. It consists of four-tuples

$$q = (\lambda, \nu, \mu(\cdot), \psi(\cdot)),$$

which satisfy the conditions

$$\lambda + \sum_{j=1}^r \nu_j + \sum_{j=1}^k \mu^j(0) = 1,$$

$$\lambda \geq 0, \quad \nu \geq 0, \quad \langle \nu, e(\bar{x}_i(1; w)) \rangle = 0,$$

$$\dot{\psi}(t) = -\bar{H}'_x(x_i(t; w), u_i(t), w, \psi(t), \mu(t)) \quad \text{a.a. } t \in [0, 1],$$

$$-\psi(1) = \lambda \varphi'_x(x_i(1; w), w) + \sum_{j=1}^r \nu_j^i e'_{jx}(x_i(1; w)),$$

and $\mu(t)$ satisfies the (a)–(c) conditions, in which \bar{x} is replaced by x_i .

Similarly, define the set \mathcal{L}_i . The linear operator $A_j^{(i)}$ is given by

$$A_j^{(i)}(q(\cdot; w)) = \int_0^1 (\bar{H}(x_i(t; w), v_j(t), w, \psi(t; w), \mu(t; w)) \\ - \bar{H}(x_i(t; w), u_i(t), w, \psi(t; w), \mu(t; w))) dt.$$

The map $D_N^{(i)}(t; w)$ is defined as follows:

$$D_N^{(i)}(t; w) := \left\{ (a, b) \in \mathbb{R}^{1+r+k+n} \times \mathbb{R}^N : \right. \\ \left. a = q(t; w), \quad b = B_N^{(i)}(q(\cdot; w)) \text{ for some } q \in \mathcal{L}_i \right\},$$

where $B_N^{(i)} = (A_1^{(i)}, A_2^{(i)}, \dots, A_N^{(i)})$.

Note that due to the NCO for the i -problem obtained above, the set-valued map $D_N^{(i)}(t; w)$ contains a selector $(a_i(t; w), b_i(t; w))$ such that

$$a_i = q_i, \quad \text{where } q_i(t; w) = (\lambda_i(w), \nu_i(w), \mu_i(t; w), \psi_i(t; w)).$$

Moreover, it is a straightforward exercise to check that

$$\limsup_{i \rightarrow \infty} \text{Gr} D_N^{(i)} \subseteq \text{Gr} D_N.$$

Now, we are ready to apply [Lemma 1](#). Let ℓ denotes the Lebesgue measure on \mathbb{R} . Denote by ℓ_B the corresponding Borel length-measure (its unique Lebesgue–Stieltjes extension equals ℓ).

In [Lemma 1](#), let us consider $E := [0, 1] \times C$, $\Delta := \ell_B \otimes \Lambda$, $\Delta_i := \ell_B \otimes \Lambda_i$, $D(x) := D_N(t; w)$, $D_i(x) := D_N^{(i)}(t; w)$ respectively, and, let us take as γ_i the above-mentioned selector (a_i, b_i) which is surely measurable with respect to Δ_i .

Then [Lemma 1](#) yields the existence of a $\ell_B \otimes \Lambda$ measurable selector of D_N . Therefore, there exist $\ell_B \otimes \Lambda$ -measurable Lagrange multipliers $q_N \in \mathcal{L}$, where

$$q_N(t; w) = (\lambda_N(w), \nu_N(w), \mu_N(t; w), \psi_N(t; w)).$$

Moreover, the passage to limit ensures that

$$\int_{[0, 1] \times C} B_N^{(i)}(q_i) d(\ell_B \otimes \Lambda_i) \rightarrow \int_{[0, 1] \times C} B_N(q) d(\ell_B \otimes \Lambda) \leq 0,$$

where the inequality “ ≤ 0 ” follows from the NCO for the i -problem. By definition of B_N and also in view of Fubini’s theorem [\[24, p. 359\]](#), the last inequality means that, for $j = 1, 2, \dots, N$,

$$\int_C \int_0^1 (\bar{H}(\bar{x}(t; w), v_j(t), w, \psi_N(t; w), \mu_N(t; w)) \\ - \bar{H}(\bar{x}(t; w), \bar{u}(t), w, \psi_N(t; w), \mu_N(t; w))) d\ell d\Lambda \leq 0. \quad (2)$$

Let $\mathcal{L}^m \subset \mathcal{L}$ be the subset of $\ell \otimes \Lambda$ measurable functions. Note that we have just proved that $q_N \in \mathcal{L}^m$, and thus $\mathcal{L}^m \neq \emptyset$. To be more accurate in the arguments, we recall the condition (a), and that, by assumption, $g^j(x_0, w) < 0$. Thus, in the definition of \mathcal{L} , we can replace $\mu^j(0)$ by $\text{ess sup}_{t \in [0, 1]} \mu^j(t)$.

Now consider the Hilbert space

$$\mathcal{X} = L_2([0, 1] \times C; \ell \otimes \Lambda)$$

of the $\ell \otimes \Lambda$ measurable functions $q(t, w)$ defined on $[0, 1] \times C$ with values in $\mathbb{R}^{1+r+k+n}$.

Note that $\mathcal{L}^m \subset \mathcal{X}$, and \mathcal{L}^m is closed which is a straightforward task to verify. The key point is to ensure the monotonicity – of course, up to the equivalence class – of the limit $\mu(\cdot; w)$ for Λ -a.a. w . However, this can be deduced from Egorov's theorem [24, p. 290]. Then, by taking into account the convexity of \mathcal{L}^m , Mazur's theorem [24] yields that the set \mathcal{L}^m is weakly closed. Since \mathcal{L}^m is bounded, and L_2 is separable, the set \mathcal{L}^m is weakly sequentially compact. Then, along a subsequence, $q_N \xrightarrow{w} q \in \mathcal{L}^m$ weakly in L_2 , for some

$$q(t; w) = (\lambda(w), \nu(w), \mu(t; w), \psi(t; w)).$$

Moreover, by taking the limit as $N \rightarrow \infty$ in the integral inequality (2), and by bearing in mind that $\{v_j\}_{j=1}^\infty$ is everywhere dense in \mathcal{U} , we obtain the maximum condition in the integral form

$$\begin{aligned} \max_{u \in \mathcal{U}} \int_C \int_0^1 \bar{H}(\bar{x}(t; w), u(t), w, \psi(t; w), \mu(t; w)) dt d\Lambda \\ = \int_C \int_0^1 \bar{H}(\bar{x}(t; w), \bar{u}(t), w, \psi(t; w), \mu(t; w)) dt d\Lambda, \end{aligned}$$

which is equivalent to the corresponding pointwise condition stated in the theorem.

In order to complete the proof it remains to notice that the set \mathcal{L}^m corresponds exactly to the set of Lagrange multipliers in the theorem. \square

4. Non-degeneracy

The NCO obtained in Theorems 1–3 degenerate whenever there exists $w \in C$ s.t. $g^j(x_0, w) = 0$ for some j . Indeed, in order to see this, we just need to consider the Dirac measure $\Lambda = \delta_w$ and the Lagrange multipliers

$$\lambda(w) = 0, \quad \nu(w) = 0, \quad \psi(\cdot; w) = 0, \quad \mu^i(\cdot; w) = 0 \quad \text{if } i \neq j,$$

and

$$\mu^j(t; w) = \begin{cases} 1, & t = 0, \\ 0, & t > 0. \end{cases}$$

Thus, in the case when the starting point lies on the boundary of the state constraint set, the above results hold trivially true.³ Then, a natural question arises: How to overcome the degeneracy phenomenon? Several approaches have been developed in the literature for the classic problem statement, that is, for the case in which C is a singleton, see, e.g., [9,17,21,11,20,10,5]. The same approaches can be applied to the study of minimax state constrained control problem as well. Below, we employ the method proposed in [10].

The key-idea is based on the time-transversality conditions. Suppose, for a while, that the time interval $[0, 1]$ in problem (1) is not fixed, and now it is $[t_0, 1]$, where the left time-endpoint t_0 is free. At that, let the optimal process be still the same, that is, (\bar{x}, \bar{u}) defined on $[0, 1]$, so $\bar{t}_0 = 0$. Then, the NCO from Theorem 3 can be supplemented by a new condition (the so-called time-transversality condition):

$$\max_{u \in \mathcal{U}} \int_C \bar{H}(x_0, u, w, \psi(0; w), \mu(0; w)) d\Lambda = 0. \quad (3)$$

³ The NCO may also degenerate at the terminal point, see [13]. Herein, we shall consider only the left endpoint x_0 .

We omit the proof which is lengthy but is along the same lines as the preceding analysis. Its details can be found in [10]. The important distinguished feature with regard to the minimax control context is that, in the proof of Theorem 3, we consider not a four-tuple, but the five-tuple:

$$q(t; w) = (\lambda(w), \nu(w), \xi(w), \mu(t; w), \psi(t; w)),$$

where $\xi = (\xi^1, \xi^2, \dots, \xi^k)$, and $\xi(w) \geq \mu(t; w) \forall t \in [0, 1]$, while $\xi^j(w) = \mu^j(0+; w)$ whenever $g^j(x_0, w) < 0$, and

$$\lambda + \sum_{j=1}^r \nu^j + \sum_{j=1}^k \xi^j = 1.$$

Note that, in this way, by setting $\mu(0; w) = \xi(w)$, we ensure that the map $\mu(0; \cdot)$ is Λ -measurable.

Let us associate with the problem (1) the minimax ν -problem:

$$\begin{aligned} \text{Minimize } \max_{w \in C} \varphi(x_1, w) \\ \text{subject to } \dot{x} = (v+1)f(x, u, w), \\ \dot{\chi} = v+1, \\ x(t_0) = x_0, \quad x(1) = x_1 \in S, \\ \chi(t_0) = 0, \quad \chi(1) = 1, \\ u(t) \in U, \quad |v(t)| \leq \frac{1}{2} \quad \text{a.a. } t \in [t_0, 1], \\ g(x(t), w) \leq 0 \quad \forall t \in [t_0, 1]. \end{aligned} \quad (4)$$

Problem (4) is autonomous with free left time-endpoint t_0 . There is a one-to-one correspondence between the feasible processes of (4) and the feasible processes of the original problem (1). Moreover, the values of cost on the processes corresponding to each other coincide. This is true due to a simple time-variable change $t \rightarrow \chi(t)$. Therefore, the two problems (1) and (4) are equivalent.

It is easy to see that the feasible process $x(t; w) = \bar{x}(t; w)$, $\chi(t) = t$, $u(t) = \bar{u}(t)$, $v(t) = 0$, $t_0 = 0$ corresponds to the solution of (1) with the same value of cost. Therefore, it is a solution to (4) as well.

By applying to (4) the NCO from Theorem 3, supplemented by (3), and also by considering the fact that χ does not depend on w , we deduce

$$\max_{u \in \mathcal{U}} \int_C \bar{H}(\bar{x}(t; w), u, w, \psi(t; w), \mu(t; w)) d\Lambda = \kappa \quad \text{a.a. } t \in [0, 1], \quad (5)$$

$$\max_{u \in \mathcal{U}} \int_C \bar{H}(x_0, u, w, \psi(0; w), \mu(0; w)) d\Lambda = \kappa, \quad (6)$$

for some number $\kappa \in \mathbb{R}$. This is just due to a verbatim repetition of the arguments from [10, Chapter 2, Section 2.12].

Introduce the controllability condition imposed on the left endpoint.

Definition 1. The controllability condition is said to be satisfied on the left endpoint x_0 provided that there exists a vector $u_0 \in U$, such that

$$\langle g_x^j(x_0, w), f(x_0, u_0, w) \rangle < 0 \quad \forall j: g^j(x_0, w) = 0 \quad \forall w \in C. \quad (7)$$

Let us verify now, whether under the just imposed assumption, the above-mentioned degenerate Lagrange multipliers combination may appear or not. Due to the conditions (5) and (6), the answer is obviously “no”. Indeed, let $(\lambda, \nu, \mu, \psi, \Lambda)$ be degenerate. Then, due to (6), it is clear that the controllability condition (7) implies that $\kappa > 0$. However, in view of (5), $\kappa = 0$. This contradiction prevents the existence of degenerate multipliers.

Similar arguments lead us to the following more general result.

Theorem 4. Suppose that (\bar{x}, \bar{u}) is a solution to (1), in which the set S is given by inequality constraints, and the controllability assumptions

are in force. Then, there exist a Radon probability measure Λ on C and, for each $w \in C$, Lagrange multipliers $\lambda(w) \in [0, 1]$, $\nu(w) \in [0, 1]^r$, $\psi(\cdot; w) \in W_{1,\infty}([0, 1])$, and $\mu(\cdot; w) \in BV([0, 1])$ such that the assertion of [Theorem 3](#) is valid, and, moreover, there exists a Borel set $\tilde{C} \subseteq C$, such that $\Lambda(\tilde{C}) > 0$, and, for a.a. $w \in \tilde{C}$

$$\lambda(w) + \sum_{j=1}^r \nu^j(w) + \sum_{j=1}^k \mu^j(0+; w) > 0.$$

Note that the function $\mu(0+; w)$ above is Λ -measurable as the pointwise (w.r.t. Λ) limit of Λ -measurable functions. The last expression is a stronger non-triviality condition which ensures non-degeneracy.

The following example clearly demonstrates the essential need of the controllability condition. (This example is an extension into the minimax context of the example given in [\[17\]](#).)

Example 1. Consider the minimax problem

$$\begin{aligned} \text{Minimize} \quad & \max_{w \in \{0,1\}} \int_0^1 (1+w)u(t)dt \\ \text{subject to} \quad & \dot{x} = wtu + (1-w)u, \\ & x(0; w) = 0, \quad x(1; w) = 0, \\ & x(t; w) \geq 0, \\ \text{and} \quad & u \in U = [-1, 1]. \end{aligned}$$

It is easy to check that the solution is $\bar{u} = 0$. Degeneration occurs at the value $w=1$ with $\Lambda = \delta(1)$, where $\delta(1)$ is the corresponding Dirac's measure. Then, the nontriviality condition of [Theorem 4](#) fails to hold. This happens because the controllability condition (7) is violated at $w=1$.

Remark 3. The same arguments, theorems and proofs, as in this article, remain valid also for the NCO in the Dubovitskii–Milyutin–Neustadt form. It follows directly from the preceding analysis. In this context, the existence of the second order derivative of g w.r.t. x is not required.

5. Conclusions

In this article, we presented, discussed and proved NCO in the Gamkrelidze-like form for optimal control problems with state constraints. Besides the basic ones, an additional “time transversality” condition was considered in order to ensure that the obtained conditions do not degenerate.

As we mentioned in [Remark 3](#), the above proofs also allow to obtain the NCO in the Dubovitskii–Milyutin–Neustadt form. Thus, the requirement of second-order differentiability of g w.r.t. x appears to be a redundant assumption. However, this is not quite true. As we noted earlier, the NCO in the Gamkrelidze-like form stand out much closer to applications than the Dubovitskii–Milyutin–Neustadt alternative. Many examples such as the geodesic equation, continuity of the measure multiplier, and conservation laws, which require the existence of the second-order derivative of g , confirm this thesis as well (for details, see [Section 3.1](#)).

This issue is even more pertinent, when taking into account a rather high cumbersomeness of the NCO for “minimax” optimal control problems with state constraints, for which the Gamkrelidze-like form seems to be more preferable than the Dubovitskii–Milyutin–Neustadt alternative. It is also worth adding, that, by all means, in most engineering applications the second order derivative of the state constraint function exist.

Acknowledgments

The authors thank a valuable support of the Russian Foundation for Basic Research with grants 16-31-60005 and 16-01-00283; São Paulo Research Foundation (FAPESP) with grants 2013/07375-0 and 2016/03540-4, CNPq (Brazil) with grants 309335/2012-4, 479109/2013-3, 457785/2014-4 and 310955/2015-7; and FCT (Portugal) research funding granted to the SYSTEC R&D Unit under project UID/EEA/00147/2013, and project NORTE-01-0145-FEDER-000033 – STRIDE (COMPETE 2020). The authors are also thankful to anonymous reviewers for useful comments.

References

- [1] V. Alexandrov, M. Budninskiy, On kinematic control extremals, in: European Control Conference (ECC), Zurich, Switzerland, 2013, pp. 210–214.
- [2] A. Arutyunov, Perturbations of extremal problems with constraints and necessary optimality conditions, *J. Sov. Math.* 54 (6) (1991) 1342–1400.
- [3] A. Arutyunov, Properties of the Lagrange multipliers in the Pontryagin maximum principle for optimal control problems with state constraints, *Differ. Equ.* 48 (12) (2012) 1586–1595.
- [4] A. Arutyunov, D. Karamzin, On some continuity properties of the measure Lagrange multiplier from the maximum principle for state constrained problems, *SIAM J. Control Optim.* 53 (4) (2015) 2514–2540.
- [5] A. Arutyunov, D. Karamzin, F. Pereira, A nondegenerate maximum principle for the impulse control problem with state constraints, *SIAM J. Control Optim.* 43 (5) (2005) 1812–1843.
- [6] A. Arutyunov, D. Karamzin, F. Pereira, R.V. Gamkrelidze's maximum principle for optimal control problems with bounded phase coordinates and its relation to other optimality conditions, *Dokl. Math.* 83 (1) (2011) 131–135.
- [7] A. Arutyunov, D. Karamzin, F. Pereira, State constraints in impulsive control problems: Gamkrelidze-like conditions of optimality, *J. Optim. Theory Appl.* 166 (2) (2015) 440–459.
- [8] A. Arutyunov, D. Silin, N. Zerkalov, Maximum principle and second-order conditions for minimax problems of optimal control, *J. Optim. Theory Appl.* 75 (3) (1992) 521–533.
- [9] A. Arutyunov, N. Tynianskij, The maximum principle in a problem with phase constraints, *Sov. J. Comput. Syst. Sci.* 23 (1) (1984) 28–35.
- [10] A.V. Arutyunov, Optimality Conditions. Abnormal and Degenerate Problems Mathematics and its Applications, Kluwer Academic Publishers, Dordrecht, 2000.
- [11] A.V. Arutyunov, S.M. Aseev, Investigation of the degeneracy phenomenon of the maximum principle for optimal control problems with state constraints, *SIAM J. Control Optim.* 35 (3) (1997) 930–952.
- [12] A.V. Arutyunov, D.Y. Karamzin, Non-degenerate necessary optimality conditions for the optimal control problem with equality-type state constraints, *J. Global Optim.* 64 (4) (2016) 623–647.
- [13] A.V. Arutyunov, D.Y. Karamzin, F.L. Pereira, The maximum principle for optimal control problems with state constraints by R.V. Gamkrelidze: revisited, *J. Optim. Theory Appl.* 149 (3) (2011) 474–493.
- [14] A.V. Arutyunov, D.Y. Karamzin, F.L. Pereira, Conditions for the absence of jumps of the solution to the adjoint system of the maximum principle for optimal control problems with state constraints, *Proc. Steklov Inst. Math.* 292 (1) (2016) 27–35.
- [15] F. Clarke, Optimization and Nonsmooth Analysis, Wiley-Interscience, New York, 1983.
- [16] A.V. Davydova, D.Y. Karamzin, On some properties of the shortest curve in a compound domain, *Differ. Equ.* 51 (12) (2015) 1626–1636.
- [17] A. Dubovitskij, V. Dubovitskij, Necessary conditions for a strong minimum in optimal control problems with degeneration of final and phase constraints, *Uspekhi Mat. Nauk.* 40 (2(242)) (1985) 175–176.
- [18] A. Dubovitskij, A. Milyutin, Extremum problems with constraints, *Sov. Math. Dokl.* 4 (1963) 452–455.
- [19] I. Ekeland, On the variational principle, *J. Math. Anal. Appl.* 47 (1974) 324–353.
- [20] M.M.A. Ferreira, F.A.C.C. Fontes, R.B. Vinter, Nondegenerate necessary conditions for nonconvex optimal control problems with state constraints, *J. Math. Anal. Appl.* 233 (1) (1999) 116–129.
- [21] M.M.A. Ferreira, R.B. Vinter, When is the maximum principle for state constrained problems nondegenerate? *J. Math. Anal. Appl.* 187 (2) (1994) 438–467.
- [22] R.V. Gamkrelidze, Optimal control processes for bounded phase coordinates, *Izv. Akad. Nauk. SSSR Ser. Mat.* 24 (1960) 315–356.
- [23] D.Y. Karamzin, Necessary conditions for an extremum in a control problem with phase constraints, *Zh. Vychisl. Mat. Mat. Fiz.* 47 (7) (2007) 1123–1150.
- [24] A. Kolmogorov, S. Fomin, Introductory Real Analysis, Dover Books on Mathematics, Dover Publications, 1975.
- [25] A.B. Kurzhanski, I.M. Mitchell, P. Varaiya, Optimization techniques for state-constrained control and obstacle problems, *J. Optim. Theory Appl.* 128 (3) (2006) 499–521.

- [26] B. Mordukhovich, Maximum principle in the problem of time optimal response with nonsmooth constraints, *J. Appl. Math. Mech.* 40 (1976) 960–969.
- [27] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation I. Basic Theory, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 330, Springer-Verlag, Berlin, 2006.
- [28] L. Neustadt, An abstract variational theory with applications to a broad class of optimization problems. ii. Applications, *SIAM J. Control Optim.* 4 (1966) 505–527.
- [29] L. Pontryagin, V. Boltyanskij, R. Gamkrelidze, E. Mishchenko, Selected works. Vol. 4. The Mathematical Theory of Optimal Processes. Ed. and with a preface by R.V. Gamkrelidze. Transl. from the Russian by K.N. Trirogoff. Transl. ed. by L. W. Neustadt. With a preface by L.W. Neustadt and K.N. Trirogoff. Reprint of the 1962 Engl. translation, *Classics of Soviet Mathematics*, xxiv, Gordon and Breach Science Publishers, New York, NY, 1986, 360 p.
- [30] I.B. Russak, On general problems with bounded state variables, *J. Optim. Theory Appl.* 6 (6) (1970) 424–452.
- [31] I.B. Russak, Second-order necessary conditions for general problems with state inequality constraints, *J. Optim. Theory Appl.* 17 (112) (1975) 43–92.
- [32] R.B. Vinter, *Optimal Control*, Birkhäuser, Boston, 2000.
- [33] R.B. Vinter, Minimax optimal control, *SIAM J. Control Optim.* 44 (3) (2005) 939–968 (electronic).
- [34] R.B. Vinter, G. Pappas, A maximum principle for nonsmooth optimal-control problems with state constraints, *J. Math. Anal. Appl.* 89 (1) (1982) 212–232.
- [35] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, London, 1972.
- [36] E.V. Zakharov, D.Y. Karamzin, On the study of conditions for the continuity of the lagrange multiplier measure in problems with state constraints, *Differ. Equ.* 51 (3) (2015) 399–405.
- [37] M.I. Zelikin, N.T. Tynyanskii, Determined differential games, *Uspekhi Mat. Nauk* 20 (4) (1965) 151–157.