First-Order Necessary Conditions in Optimal Control

David Mayne¹ · Richard Vinter¹

Received: 10 October 2020 / Accepted: 27 February 2021 / Published online: 26 March 2021
© The Author(s) 2021

Abstract

In an earlier analysis of strong variation algorithms for optimal control problems with endpoint inequality constraints, Mayne and Polak provided conditions under which accumulation points satisfy a condition requiring a certain optimality function, used in the algorithms to generate search directions, to be nonnegative for all controls. The aim of this paper is to clarify the nature of this optimality condition, which we call the first-order minimax condition, and of a related integrated form of the condition, which, also, is implicit in past algorithm convergence analysis. We consider these conditions, separately, when a pathwise state constraint is, and is not, included in the problem formulation. When there are no pathwise state constraints, we show that the integrated first-order minimax condition is equivalent to the minimum principle and that the minimum principle (and equivalent integrated first-order minimax condition) is strictly stronger than the first-order minimax condition. For problems with state constraints, we establish that the integrated first-order minimax condition and the minimum principle are, once again, equivalent. But, in the state constrained context, it is no longer the case that the minimum principle is stronger than the first-order minimax condition, or vice versa. An example confirms the perhaps surprising fact that the first-order minimax condition is a distinct optimality condition that can provide information, for problems with state constraints, in some circumstances when the minimum principle fails to do so.

Keywords Optimal control · Necessary conditions · Minimum principle · State constraints

Communicated by Dean A. Carlson.

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

✉ Richard Vinter
r.vinter@imperial.ac.uk

David Mayne
d.mayne@imperial.ac.uk

¹ Imperial College, London, UK

Mathematics Subject Classification 49K05 · 49K15 · 49K35

1 Introduction

Convergence analysis of optimal control algorithms typically aims to demonstrate that accumulation points of sequences of control functions, generated by the algorithm under consideration, satisfy necessary conditions of optimality. The nature of the necessary conditions is not fixed, but depends on the algorithm and on the analytical techniques employed in the convergence analysis.

In a series of papers [6,9], Mayne and Polak, building on earlier work by Jacobson and Mayne [3], proposed algorithms (with accompanying convergence analysis) for solving optimal control problems with endpoint inequality constraints, based on strong variations. In consequence of the strong variations employed in the algorithms, one might expect that the necessary conditions featuring in the convergence analysis would be the minimum principle. (This is on account of fact that the minimum principle can be proved by consideration of strong variations to the control.) But, in fact, another condition was used in this earlier literature, a necessary condition of optimality that we, in this paper, shall call the first-order minimax condition. Necessary conditions similar to an integral version of the first-order minimax condition, but applicable to optimal control problems with pathwise state constraints (we shall refer to them as ‘state constraint problems’), are implicit in the convergence analysis of feasible directions algorithms due to Pytlak and Vinter [11].

The first-order minimax condition (for state constraint-free problems) originates in a systematic approach to algorithm construction in nonlinear programming and optimal control, due to Polak [8]; the idea, in the case of optimal control problems, is to find, for a given control u' , an ‘optimality function’ $u \rightarrow \theta(u, u')$ with the property that $\arg \min\{u \rightarrow \theta(u, u')\}$ provides, in the event that $\min_u \theta(u, u') < 0$, search directions for the reduction of cost and endpoint constraint violations. The convergence analysis involves showing that accumulation points $\bar{u}$ satisfy $\min_u \theta(u, \bar{u}) = 0$. For the optimal control problems treated by Mayne and Polak in [6,9], ‘ $\min_u \theta(u, \bar{u}) = 0$ ’ can be interpreted as our first-order minimax condition.

Investigating the strength of a necessary condition based on particular computational scheme gives insights into pathological situations when the necessary condition is satisfied at some non-minimizing process and hence into circumstances when the scheme might fail to yield a minimizer. This is the reason why, in this paper, we investigate the strength of two necessary conditions previously encountered in algorithm analysis, by comparing them with the known minimum principle.

We provide a rather complete picture of the relations between the following necessary conditions of optimality, both for problems with and without pathwise state constraints: the first-order minimax condition (on which the algorithms in [6,9] are based), the integrated first-order minimax condition (implicit in the algorithm convergence analysis of [11]), and the minimum principle. For problems without state constraints, our investigations reveal that the minimum principle is a stronger necessary condition than the first-order minimax condition and that the integrated first-order minimax condition and the minimum principle are equivalent. An example is provided

demonstrating that the minimum principle is strictly stronger than the first-order minimax condition.

When we pass to problems with pathwise state constraints, we find, once again, that the integrated first-order minimax condition and the minimum principle are equivalent. But we discover that the first-order minimax condition is neither stronger nor weaker than the minimum principle. Thus, the first-order minimax condition is revealed to be an optimality condition that is distinct from the minimum principle. An example illustrates how it can be used to show that a certain admissible process is not a minimizer, when the minimum principle fails to do so.

It is well known that first necessary conditions for free right endpoint problem can be simply derived, by consideration of needle variations and performance of elementary gradient calculations. When general endpoint constraints are present, the derivation of necessary conditions akin to the minimum principle necessitates the use of a more sophisticated analysis (based on separation of approximating cones of the reachable set [2], or non-smooth perturbation methods [7]). We observe, however, that when the endpoint constraints take the form of *inequality constraints*, the derivation of necessary conditions based on simple gradient calculations becomes once again possible. (The ‘necessary conditions’ referred to here are ‘first-order minimax conditions.’) There is a parallel here with the derivation of Fritz John-type first-order necessary conditions in nonlinear programming where, as is well known, the analysis is greatly simplified, if the constraints are inequality constraints [5], not mixed inequality and equality constraints. A secondary purpose of this paper is to make explicit these simplifications, also in an optimal control context.

2 Problem Formulation

Consider the optimal control problem with endpoint inequality constraints:

$$(P) \left\{ \begin{array}{l} \text{Minimize } g_0(x(1)) \\ \text{over absolutely continuous functions } x \text{ and meas. functions } u \text{ such that} \\ \dot{x}(t) = f(x(t), u(t)), \quad \text{a.e. } t \in [0, 1], \\ u(t) \in \Omega, \quad \text{a.e. } t \in [0, 1], \\ x(0) = x_0 \text{ and } g_j(x(1)) \leq 0, \text{ for } j = 1, 2, \dots, r. \end{array} \right.$$

The data comprise functions $g_j : R^n \rightarrow R$, $j = 0, \dots, r$ and $f : R^n \times R^m \rightarrow R^n$ and a closed set $\Omega \subset R^m$.

A pair of functions (x, u) is called a process if x is absolutely continuous, u is Lebesgue measurable, $\dot{x}(t) = f(x(t), u(t))$ a.e., $u(t) \in \Omega$ a.e. and $x(0) = x_0$. If x also satisfies the right endpoint constraints in (P), we say that (x, u) is an admissible process. The first component x of a process (x, u) is called a state trajectory. The second component is called a control function.

We shall also consider a generalization that includes state constraints

$$h_k(x(t)) \leq 0, \quad \text{for all } t \in [0, 1] \text{ and } k = 1, \dots, N_s,$$

for given functions $h_k : R^n \rightarrow R, k = 1, \dots, N_s$.

$$(S) \begin{cases} \text{Minimize } g_0(x(1)) \\ \text{over absolutely continuous functions } x \text{ and meas. functions } u \text{ such that} \\ \dot{x}(t) = f(x(t), u(t)), \text{ a.e. } t \in [0, 1], \\ h_k(x(t)) \leq 0, \text{ for all index values } j \in \{1, \dots, N_s\} \text{ and } t \in [0, 1], \\ u(t) \in \Omega, \text{ a.e. } t \in [0, 1], \\ x(0) = x_0 \text{ and } g_j(x(1)) \leq 0, \text{ for } j = 1, 2, \dots, r. \end{cases}$$

‘Admissible processes for (S)’ are understood in the obvious sense. The following hypotheses will be invoked:

- (H1): The set Ω is compact, and the functions $g_j, j = 0, \dots, N$ and $h_k, k = 1, \dots, N_s$ are continuously differentiable,
 (H2): $f(., u)$ is continuously differentiable for each $u \in \Omega$, $f, \nabla_x f$ is continuous, and there exists $c > 0$ such that

$$|f(x, u)| \leq c(1 + |x|) \text{ for all } (x, u) \in R^n \times \Omega.$$

Notation: In Euclidean space, the Euclidean length of a vector x is denoted by $|x|$ and the closed unit ball in Euclidean space is written B . Given numbers a and b , we write $a \vee b := \max\{a, b\}$ and $a \wedge b := \min\{a, b\}$.

Given $x \in L^\infty([0, 1]; R^n)$, we write the L^∞ norm of x as $\|x\|_{L^\infty}$. $NBV^+([0, 1]; R^{N_s})$ denotes the space of N_s -tuples of (normalized) increasing functions $v = \{v_k\}$ on $[0, 1]$ such that each v_k is right continuous on $(0, 1)$. For each k , $d\nu_k$ is the Stieltjes measure associated with v_k . R^+ denotes $[0, \infty)$.

For a given nominal admissible process $(\bar{x}, \bar{u})$, $S(t, s)$ is the transition matrix for the linear system

$$\dot{y}(t) = \nabla_x f(\bar{x}(t), \bar{u}(t))y(t), \text{ a.e. } t \in [0, 1], \quad (1)$$

i.e., for each $s \in [0, 1]$, $t \rightarrow S(t, s)$ is the unique solution on $[0, 1]$ of the matrix differential equation

$$\begin{cases} dS/dt(t, s) = \nabla_x f(\bar{x}(t), \bar{u}(t))S(t, s), \text{ a.e. } t \in [0, 1], \\ S(s, s) = I_{n \times n} \quad (I_{n \times n} \text{ denotes the } n \times n \text{ unit matrix}), \end{cases}$$

$$\begin{aligned} \mathcal{U} &:= \{\text{meas. functions } u \text{ s.t. } u(t) \in \Omega \text{ a.e.}\}, \\ \Delta f(t, u) &:= f(\bar{x}(t), u) - f(\bar{x}(t), \bar{u}(t)), \text{ for } t \in [0, 1] \text{ and } u \in \Omega, \\ \Lambda^N &:= \{(\lambda_0, \dots, \lambda_N) \in R^{N+1} : \lambda_j \geq 0 \text{ for all } j \text{ and } \sum_j \lambda_j = 1\}, \\ I(\bar{x}) &:= \{0\} \cup \{j \in \{1, \dots, r\} : g_j(\bar{x}(1)) = 0\}. \end{aligned}$$

3 Analytical Tools

1. *Relaxation.* The following optimal control problem is known as the *relaxation* of (P).

$$(RP) \left\{ \begin{array}{l} \text{Minimize } g_0(x(1)) \\ \text{over abs. continuous functions } x, \text{ meas. functions } \mu : [0, 1] \rightarrow \Lambda^n \\ \text{and meas. functions } \{u_k : [0, 1] \rightarrow \Omega\}_{k=0}^n \text{ such that} \\ \dot{x}(t) = \sum_{k=0}^n \mu_k(t) f(x(t), u_k(t)), \quad \text{a.e. } t \in [0, 1], \\ x(0) = x_0 \text{ and } g_j(x(1)) \leq 0, \text{ for } j = 1, 2, \dots, r \end{array} \right.$$

A process $(x, \{(u_k, \mu_k)\}_{k=0}^n)$ for (R) is called a *relaxed process* for (P). If x satisfies the endpoint constraints, the relaxed process is called a *relaxed admissible process*. A function x (associated with some relaxed process) is called a *relaxed state trajectory*, and the $(n+1)$ -tuple or pairs $\{(\mu_i, u_i)\}$ (likewise associated) are called a *relaxed control function*. If, say, $\mu_k = 0$ for $k = 2, \dots, r$, we write $\{(\mu_k, u_k)\}$ simply as $((\mu_0, u_0), (\mu_1, u_1))$.

The relaxation theorem [1] tells us that a given relaxed state trajectory can be uniformly approximated by a state trajectory.

Theorem 3.1 (The Relaxation Theorem) *Take $(x, \{(\mu_k, u_k)\}_{k=0}^n)$ to be any relaxed process for (P) and any $\delta > 0$. Then, there exists a process (x', u') for (P) such that*

$$\|x - x'\|_{L^\infty} \leq \delta.$$

2. *A Minimax Theorem.* Minimax theorems originate in the game theory literature. We shall make use of their important role also in variational analysis.

Theorem 3.2 (Von Neumann Minimax Theorem) *Consider a function $F : X \times Y \rightarrow \mathbb{R}$ in which X and Y are subsets of topological linear spaces. Assume that*

- (i) X and Y are convex, compact sets,
- (ii) $F(\cdot, y)$ is convex and continuous for every $y \in Y$,
- (x,.) $F(x, \cdot)$ is concave and continuous for every $x \in X$.

Then, there exists $(x^, y^*) \in X \times Y$ which is a saddlepoint for F , i.e.,*

$$\sup_{y \in Y} F(x^*, y) = F(x^*, y^*) = \inf_{x \in X} F(x, y^*) \quad (2)$$

and

$$\inf_{x \in X} \sup_{y \in Y} F(x, y) = \sup_{y \in Y} \inf_{x \in X} F(x, y).$$

For a proof, see, e.g., [10, Thm. 3.4.6].

3. *First-Order Local Approximations*

The following proposition gives error bounds for needle variations associated with a given nominal process $(\bar{x}, \bar{u})$.

Proposition 3.1 *Take a process $(\bar{x}, \bar{u})$ for (P) . Assume hypotheses (H1)–(H2) are satisfied. Fix $t \in [0, 1)$ and $u \in \Omega$. Define, for each $\sigma \in (0, 1 - t]$, the control function*

$$u_\sigma(s) := \begin{cases} u & \text{if } t \leq s \leq t + \sigma \\ \bar{u}(s) & \text{otherwise} \end{cases}. \quad (3)$$

Let x_σ be the corresponding state trajectory on $[0, 1]$ such that $x_\sigma(0) = x_0$. Then, there exist a number K (that does not depend on σ) and a continuity modulus θ (i.e., a function $\theta : (0, \infty) \rightarrow (0, \infty)$ such that $\lim_{s \downarrow 0} \theta(s) = 0$) with the following properties.

(a): $\|x_\sigma - \bar{x}\|_{L^\infty} \leq K\sigma$,

(b): for any index value $j \in \{0, \dots, r\}$ and process (x, u) ,

$$g_j(x_\sigma(1)) - g_j(\bar{x}(1)) \leq \int_t^{t+\sigma} \Delta f(s, u(s)) \cdot S^T(1, s) \nabla g_j(\bar{x}(1)) ds + \sigma \times \theta(\sigma)'$$

(c): for any index value $k = 1, \dots, N_s$, $t' \in (t, 1]$ and $\sigma' := \sigma \wedge (t' - t)$

$$\begin{aligned} & h_k(x_\sigma(t')) - h_k(\bar{x}(t')) \\ & \leq \int_t^{t+\sigma'} \Delta f(s, u(s)) \cdot S^T(t', s) \nabla h_k(\bar{x}(t')) ds + \sigma' \times \theta(\sigma'). \end{aligned}$$

Proof Choose constants $K_1, R > 0$ and k with the following properties:

$$\begin{aligned} & K_1 > e^c |x_0| + (e^c - 1), \quad R > \max\{|f(x, u)| : |x| \leq K_1, u \in \Omega\} \text{ and} \\ & k > \max\{|\nabla_x f(x, u)| : |x| \leq K_1, u \in \Omega\}. \end{aligned}$$

We deduce from hypothesis (H2) and Gronwall's inequality that, for any process (x', u') such that $x'(0) = x_0$,

$$|x'(t)| < K_1 \text{ for all } t \in [0, 1] \text{ and } |f(x'(t), u'(t))| < R \text{ a.e. } t \in [0, 1].$$

Furthermore, $x \rightarrow f(x, v)$ is Lipschitz continuous on $K_1 B$, with Lipschitz constant k , for all $v \in \Omega$.

Using the facts that x_σ and $\bar{x}$ are state trajectories with the same initial state, we can show that

$$\begin{aligned} |x_\sigma(t) - \bar{x}(t)| &= \left| \int_0^t f(x_\sigma(s), u_\sigma(s)) - f(\bar{x}(s), \bar{u}(s)) ds \right| \\ &\leq \int_0^t k |x_\sigma(s) - \bar{x}(s)| ds + \int_{[t, t+\sigma] \cap [0, t]} |\Delta f(s, u)| ds \end{aligned}$$

for each $t \in [0, 1]$. But then, by Gronwall's inequality,

$$\|x_\sigma - \bar{x}\|_{L^\infty} \leq K \times \sigma, \quad (4)$$

in which $K := 2e^k R$. We have confirmed property (a).

To complete the proof, it is required only to establish property (c). (The proof of property (b) follows from (c), when we substitute g_j in place of h_k in the analysis and select $t' = 1$.) Take any index value k and $t' \geq t$. We have

$$\begin{aligned} h_k(x_\sigma(t')) - h_k(\bar{x}(t')) &= h_k(x_\sigma(t')) - h_k(\bar{x}(t')) + \\ &\int_0^{t'} (\dot{x}_\sigma(s) - \dot{\bar{x}}(s) - [f(x_\sigma(s), u_\sigma(s)) - f(\bar{x}(s), \bar{u}(s))]) \cdot S^T(1, s) \nabla h_k(\bar{x}(1)) ds. \end{aligned}$$

The second term on the right can be written

$$\begin{aligned} &+ \int_0^{t'} (\dot{x}_\sigma(s) - \dot{\bar{x}}(s) - [f(x_\sigma(s), \bar{u}(s)) - f(\bar{x}(s), \bar{u}(s))]) \cdot S^T(1, s) \nabla h_k(\bar{x}(1)) ds \\ &+ \int_t^{t+\sigma'} \left(\Delta f(s, u) + [f(x_\sigma(s), u) - f(\bar{x}(s), u)] \right) \cdot S^T(1, s) \nabla h_k(\bar{x}(1)) ds. \end{aligned}$$

Furthermore, since $(d/dt)S^T(1, t) = -f_x^T(\bar{x}(t), \bar{u}(t))S^T(1, t)$,

$$\begin{aligned} \int_0^{t'} (\dot{x}_\sigma(s) - \dot{\bar{x}}(s) \cdot S^T(1, s) \nabla h_k(\bar{x}(t')) ds &= (x_\sigma(t') - \bar{x}(t')) \cdot \\ \nabla h_k(\bar{x}(t')) - \int_0^{t'} [\nabla_x f(\bar{x}(s), \bar{u}(s))(x_\sigma(s) - \bar{x}(s))] \cdot S^T(1, s) \nabla h_k(\bar{x}(t')) ds \end{aligned}$$

(‘integration by parts’). We know, however, that, under the differentiability hypotheses on the g_j 's and $x \rightarrow f(x, u)$, there exists a modulus of continuity θ_1 (that does not depend on k or t') such that

$$\begin{aligned} &|h_k(x_\sigma(t')) - h_k(\bar{x}(t')) - \nabla h_k(\bar{x}(t')) \cdot (x_\sigma(t') - \bar{x}(t'))| \\ &\leq |x_\sigma(t') - \bar{x}(t')| \times \theta_1(|x_\sigma(t') - \bar{x}(t')|), \text{ for each } k \in \{1, \dots, N_s\} \end{aligned}$$

and

$$\begin{aligned} &|f(x_\sigma(s), u) - f(\bar{x}(s), u) - \nabla_x f(\bar{x}(s), u) \cdot (x_\sigma(s) - \bar{x}(s))| \\ &\leq |x_\sigma(s) - \bar{x}(s)| \times \theta_1(|x_\sigma(s) - \bar{x}(s)|), \end{aligned}$$

for each $s \in [t, t + \sigma]$. Combining these relations and noting property (a), we find that, for some $M > 0$ that does not depend on i ,

$$\begin{aligned} & h_k(x_\sigma(t')) - h_k(\bar{x}(t')) \\ & \leq \int_t^{t+\sigma'} \Delta f(s, u) \cdot S^T(1, s) \nabla h_k(\bar{x}(s)) ds + M \times \sigma' \times \left(\theta_1(E \times \sigma') + \sigma' \right). \end{aligned} \quad (5)$$

Here, as before, $\sigma' = \sigma \wedge (t' - t)$. We have confirmed property (c) with modulus of continuity $\theta(s) := M \times (\theta_1(E \times s) + s)$. $\square$

The preceding proposition provides estimates on solutions to the controlled differential equation $\dot{x} = f(x, u)$ induced by needle variations. Similar analysis yields estimates on solutions, induced by another kind of ‘local’ variation:

Proposition 3.2 *Take a process $(\bar{x}, \bar{u})$. Take any control function $u \in \mathcal{U}$. For each $\sigma \in [0, 1]$, define x_σ to be the solution of the differential equation*

$$\begin{cases} \dot{x}_\sigma(t) = \sigma \left(f(x_\sigma(t), u(t)) - f(x_\sigma(t), \bar{u}(t)) \right) + f(x_\sigma(t), \bar{u}(t)), \\ x_\sigma(0) = x_0. \end{cases}$$

(Notice that x_σ is the relaxed state trajectory corresponding to the relaxed control $(\mu_0 = (1 - \sigma), \bar{u}), (\mu_1 = \sigma, u)$.)

Then, there exists a $K > 0$ (independent of σ) and a continuity modulus θ (i.e., a function $\theta : (0, \infty) \rightarrow (0, \infty)$, such that $\lim_{s \downarrow 0} \theta(s) = 0$), and

- (a): $\|x_\sigma - \bar{x}\|_{L^\infty} \leq K\sigma$,
 (b): for any index value $j \in \{0, \dots, r\}$ and process (x, u) ,

$$\sigma^{-1}(g_j(x_\sigma(1)) - g_j(\bar{x}(1))) \leq \int_0^1 \Delta f(s, u(s)) \cdot S^T(1, s) \nabla g_j(\bar{x}(1)) ds + \theta(\sigma)$$

- (c): for any index value $k = 1, \dots, N_s$, $t' \in [t, 1]$

$$\begin{aligned} & \sigma^{-1} \left(h_k(x_\sigma(t')) - h_k(\bar{x}(t')) \right) \\ & \leq \int_0^{t'} \Delta f(s, u(s)) \cdot S^T(t', s) \nabla h_k(\bar{x}(t')) ds + \theta(\sigma). \end{aligned}$$

4 Necessary Conditions of Optimality (No State Constraints)

This section provides three sets of necessary conditions for an admissible process $(\bar{x}, \bar{u})$ to be a minimizer for (P).

Theorem 4.1 (The First-Order Minimax Condition)

Let $(\bar{x}, \bar{u})$ be a minimizer for (P). Assume (H1)-(H2) are satisfied. Then, for a.e. $t \in [0, 1]$

$$\max_{j \in I(\bar{x})} \left\{ \Delta f(t, u) \cdot S^T(1, t) \nabla g_j(\bar{x}(1)) \right\} \geq 0, \text{ for all } u \in \Omega.$$

Proof We can always arrange, by adding a constant to g_0 if necessary, that

$$g_0(\bar{x}(1)) = 0.$$

Since $(\bar{x}, \bar{u})$ is a minimizer, we must have

$$g_0(x(1)) \vee \dots \vee g_r(x(1)) \geq 0 \quad (6)$$

for all processes (x, u) for (P). (Otherwise, there would exist a process for which the cost is reduced and all the endpoint constraints are satisfied, a contradiction.)

Let $\mathcal{T}$ be the subset of points $t \in (0, 1)$ such that t is a Lebesgue point of $s \rightarrow f(\bar{x}(s), \bar{u}(s))$, $\dot{x}(t)$ exists, $\dot{x}(t) = f(\bar{x}(t), \bar{u}(t))$ and $\bar{u}(t) \in \Omega$.

Assume that the assertions of the theorem are false. Since $\mathcal{T}$ is a set of full measure, there exist a time $t \in \mathcal{T}$, $u \in \Omega$ and a number $\gamma > 0$ such that

$$\Delta f(\bar{t}, u) \cdot S^T(1, \bar{t}) \nabla g_j(\bar{x}(1)) \leq -\gamma \quad (7)$$

for all $j \in I(\bar{x})$. Take $\sigma_i \downarrow 0$ and, for each i , let $u_i \in \mathcal{U}$ to be

$$\begin{cases} u_i(s) = u & \text{if } s \in [\bar{t}, \bar{t} + \sigma_i] \\ u_i(s) = \bar{u}(s) & \text{otherwise.} \end{cases}$$

Recalling that $g_0(\bar{x}(1)) = 0$ and that $g_j(\bar{x}(1)) \leq 0$ for $j > 0$, we deduce from Proposition 3.1 that there exists a modulus of continuity $\theta(\cdot)$ such that, for each $j \in \{0, \dots, r\}$ and $i = 1, 2, \dots$,

$$\begin{aligned} \sigma_i^{-1} g_j(x_i(1)) &\leq \sigma_i^{-1} (g_j(x_i(1)) - g_j(\bar{x}(1))) \\ &\leq \sigma_i^{-1} \int_t^{t+\sigma_i} \Delta f(s, u) \cdot S^T(1, s) \nabla g_j(\bar{x}(1)) ds + \theta(\sigma_i). \end{aligned} \quad (8)$$

By properties of Lebesgue points, there exists a modulus of continuity $\psi : (0, \infty) \rightarrow (0, \infty)$ (that does not depend on j) such that, for all i ,

$$\begin{aligned} \sigma_i^{-1} \int_t^{t+\sigma_i} \Delta f(s, u) \cdot S^T(1, s) \nabla g_j(\bar{x}(1)) ds \\ \leq \Delta f(t, u) \cdot S^T(1, t) \nabla g_j(\bar{x}(1)) + \psi(\sigma_i). \end{aligned} \quad (9)$$

It follows from (7), (8) and (9) that, for all i sufficiently large,

$$\sigma_i^{-1} g_j(x_i(1)) \leq -\gamma/2 \text{ for all } j \in I(\bar{x}). \quad (10)$$

Note, on the other hand, that for each $j \notin I(\bar{x})$, $g_j(\bar{x}(1)) < 0$. So there exists $\gamma_1 > 0$ such that

$$g_j(\bar{x}(1)) \leq -\gamma_1 \text{ for each } j \notin I(\bar{x}).$$

It follows that, for all i sufficiently large,

$$g_j(x_i(1)) \leq -(\sigma_i \gamma \wedge \gamma_1)/2 \text{ for all } j \in \{0, \dots, r\}. \quad (11)$$

This contradicts (6) when $x = x_i$. The proof is concluded. $\square$

Theorem 4.2 (The Integral First-Order Minimax Condition)

Let $(\bar{x}, \bar{u})$ be a minimizer for (P). Assume (H1)-(H2) are satisfied. Then,

$$\max_{j \in I(\bar{x})} \left\{ \int_0^1 \Delta f(t, u(t)) \cdot S^T(1, t) \nabla g_j(\bar{x}(1)) dt \right\} \geq 0,$$

for all control functions $u \in \mathcal{U}$.

Proof Assume again, without loss of generality, that $g_0(\bar{x}(1)) = 0$. Take any $u \in \mathcal{U}$ and $\sigma \in [0, 1]$. Let x_σ be the relaxed state trajectory of Proposition 3.2. We claim that

$$g_0(x_\sigma(1)) \vee \dots \vee g_r(x_\sigma(1)) \geq 0, \quad (12)$$

This follows from the fact that $(x_\sigma, \{(\mu_0 = (1 - \sigma), \mu_1 = \sigma), (u_0 = \bar{u}, u_1 = u)\})$ is a relaxed process. Indeed if, to the contrary, there exists $\bar{r} > 0$ such that

$$g_0(x_\sigma(1)) \vee \dots \vee g_r(x_\sigma(1)) < -\bar{r},$$

then, according to the relaxation theorem and in view of the continuity of the g_k 's, we would be able to find an (unrelaxed) state trajectory x' and such that $g_0(x'(1)) \vee \dots \vee g_r(x'(1)) < -\bar{r}/2$, in contradiction of the optimality of $(\bar{x}, \bar{u})$. The claim is therefore correct.

Suppose that the assertions of the theorem are false. In view of the foregoing observations, we are justified in reproducing the analysis in the proof of the earlier theorem, but now based on the perturbation estimates provided by Proposition 3.2 in place of Proposition 3.1, to obtain a contradiction of (12), for $\sigma > 0$ sufficiently small. The assertions of the theorem must therefore be true. $\square$

Finally, we state a special case of the minimum principle, in a form that emphasizes its connection with the preceding two sets of necessary conditions.

Theorem 4.3 (The Minimum Principle)

Let $(\bar{x}, \bar{u})$ be a minimizer for (P). Assume (H1)-(H2) are satisfied. Then, there exists $\lambda = (\lambda_0, \dots, \lambda_r) \in \Lambda^r$ such that $\lambda_j = 0$, for $j \notin I(\bar{x})$ and

$$\Delta f(t, u) \cdot S^T(1, t) \left(\sum_{j \in I(\bar{x})} \lambda_j \nabla g_j(\bar{x}(1)) \right) \geq 0 \quad (13)$$

for all $u \in \Omega$, a.e. $t \in [0, 1]$.

We shall give an independent proof of this well-known optimality condition, as a by-product of our investigations into the relationships between the first-order minimax condition, the integral first-order minimax condition and the minimum principle.

Comments.

1. The first-order minimax condition, originating in earlier algorithm convergence of Mayne and Polak [6,9], has the alternative expression:

$$\max_{j \in I(\bar{x})} \{ \Delta f(t, u) \cdot p_j(t) \} \geq 0 \text{ for all } u \in \Omega, \text{ a.e. } t \in [0, 1],$$

in which, for each j , $p_j(t) := S^T(1, t) \nabla g_j(\bar{x}(1))$; it can be interpreted as the solution to the costate equation:

$$\begin{cases} -\dot{p}_j(t) = f_x^T(\bar{x}(t), \bar{u}(t)) p_j(t) & \text{a.e. } t \in [0, 1], \\ p_j(1) = \nabla g_j(\bar{x}(1)). \end{cases}$$

2. The optimality condition of Theorem 4.3 can be equivalently written in terms of a single costate arc $p(t) := S^T(1, t)(\sum_j \lambda_j \nabla g_j(\bar{x}(1)))$, thus:

$$\Delta f(t, u) \cdot p(t) \geq 0 \text{ for all } u \in \Omega, \text{ a.e. } t \in [0, 1].$$

By the properties of the transition matrix $S(t, s)$, the costate arc p satisfies

$$\begin{cases} -\dot{p}(t) = f_x^T(\bar{x}(t), \bar{u}(t)) p(t) & \text{a.e. } t \in [0, 1], \\ p(1) = \sum_j \lambda_j \nabla g_j(\bar{x}(1)). \end{cases}$$

In this form, it will be recognized as the special case of the minimum principle, applied to problems with endpoint inequality constraints. (Note the Lagrange multipliers satisfy the non-triviality condition $\lambda \neq 0$, since $\lambda \in \Lambda^r$, and $\lambda_j = 0$ if $h_j(\bar{x}(1)) < 0$.)

5 Relations Between the Necessary Conditions

The following theorem relates the necessary conditions of Sect. 4.

Theorem 5.1 *Assume the data for problem (P) satisfy hypotheses (H1)-(H2), for some arbitrary admissible process $(\bar{x}, \bar{u})$. We have:*

- (i): *An admissible process $(\bar{x}, \bar{u})$ satisfies the first-order minimax condition if it satisfies the integral first-order minimax condition, i.e., the integral first-order minimax condition is stronger than the first-order minimax condition.*
- (ii): *An admissible process $(\bar{x}, \bar{u})$ satisfies the integral first-order minimax condition if and only if it satisfies the minimum principle.*

- (iii): *The minimum principle (and therefore also the integral first-order minimax condition) is a strictly stronger necessary condition than the first-order minimax condition, in the sense that data can be chosen for problem (P) such that the integrated first-order minimax condition can be used to exclude a non-minimizer, but the first-order minimax condition fails to do so.*

Proof (i): Suppose the minimax condition is not satisfied. Then, there exists $u \in \Omega$ and $t \in (0, 1)$ such that t is a Lebesgue point of $s \rightarrow f(\bar{x}(s), \bar{u}(s))$, $\dot{\bar{x}}(t)$ exists, $\bar{u}(t) \in \Omega$, $\dot{\bar{x}}(t) \in f(\bar{x}(t), \bar{u}(t))$ and

$$\Delta f(t, u) \cdot S^T(1, t) \nabla g_j(\bar{x}(1)) < 0 \text{ for all } j \in I(\bar{x}).$$

For arbitrary $\sigma \in (0, 1 - t)$, define the ‘needle variation’ $u_\sigma \in \mathcal{U}$ according to (3). From the Lebesgue point property, we have that, for all $j \in I(\bar{x})$,

$$\begin{aligned} & \sigma^{-1} \int_0^1 \Delta f(s, u_\sigma(s)) \cdot S^T(1, s) \nabla g_j(\bar{x}(1)) ds \\ &= \sigma^{-1} \int_t^{t+\sigma} \Delta f(s, u) \cdot S^T(1, s) \nabla g_j(\bar{x}(1)) ds < 0, \end{aligned}$$

for σ sufficiently small. We have shown that the integrated first-order minimax condition is not satisfied. The proof is complete.

- (ii): Take any $u \in \mathcal{U}$. Suppose the minimum principle is satisfied (with Lagrange multipliers $\lambda = (\lambda_0, \dots, \lambda_r) \in \Lambda^r$ such that $\lambda_j = 0$ if $j \notin I(\bar{x})$). Then, for any $u \in \mathcal{U}$,

$$\int_0^1 \Delta f(t, u(t)) \cdot S^T(1, t) \left(\sum_{j \in I(\bar{x})} \lambda_j \nabla g_j(\bar{x}(1)) \right) dt \geq 0.$$

This implies the existence of $\bar{j} \in I(\bar{x})$ such that

$$\int_0^1 \Delta f(t, u(t)) \cdot S^T(1, t) \nabla g_{\bar{j}}(\bar{x}(1)) dt \geq 0.$$

Hence,

$$\max_{j \in I(\bar{x})} \left\{ \int_0^1 \Delta f(t, u(t)) \cdot S^T(1, t) \nabla g_j(\bar{x}(1)) dt \right\} \geq 0.$$

We have shown that the conditions of the minimum principle imply the conditions of the integral first-order minimax theorem.

Now, suppose that the conditions of the integral first-order minimax theorem are satisfied. It is convenient to introduce, at this stage, the notation:

$$E := \{e \in L^1(0, 1) : e(t) \in \Delta f(t, \Omega) \text{ a.e. } t \in [0, 1]\}.$$

Write $\tilde{A}^r := \{\lambda \in A^r : \lambda_j = 0 \text{ if } j \notin I(\bar{x})\}$. Define $J : L^1 \times \tilde{A}^r \rightarrow R$:

$$J(e, \lambda) := \int_0^1 e(t) \cdot S^T(1, t) \left(\sum_{j \in I(\bar{x})} \lambda_j \nabla g_j(\bar{x}(1)) \right) dt.$$

Notice that the integral first-order minimax condition can be expressed as

$$\max_{j \in I(\bar{x})} \int_0^1 e(t) \cdot S^T(1, t) \nabla g_j(\bar{x}(1)) dt \geq 0, \text{ for all } e \in E. \quad (14)$$

Making use of the following representation of the convex hull of E :

$$\begin{aligned} \text{co } E &:= \{e \in L^1(0, 1) : e(t) = \sum_{i=0}^N v_i \Delta f(t, u_i(t)) \text{ for an integer } N, \\ &\text{control functions } \{u_0, \dots, u_N\} \text{ and } v \in A^N, \text{ a.e. } t \in [0, 1]\}, \end{aligned}$$

and noting that $e \rightarrow J(e, \lambda)$ is continuous w.r.t. the weak L^1 topology, we can replace (14) by the strengthened condition

$$\max_{j \in I(\bar{x})} \int_0^1 e(t) \cdot S^T(1, t) \nabla g_j(\bar{x}(1)) dt \geq 0, \text{ for all } e \in \overline{\text{co}} E, \quad (15)$$

in which $\overline{\text{co}}$ denotes ‘convex closure’ w.r.t. the L^1 norm. Note the following properties of J and its domain (which take to be $\overline{\text{co}} E \times \tilde{A}^r$):

- (A): $\text{co } E$ and $\tilde{A}^r$ are convex sets, and they are compact w.r.t. the weak L^1 topology and the Euclidean topology, respectively. (The weak L^1 compactness of E can be deduced from the hypotheses (H1) and (H2), with the help of Ascoli’s theorem),
- (B): $J(e, \lambda)$ is continuous in e , for fixed $\lambda \in \tilde{A}^r$, and continuous in λ , for fixed $e \in \text{co } E$, w.r.t. the above topologies.
- (C): $J(e, \lambda)$ is convex in e for fixed λ and convex in λ for fixed e .

We have checked the hypotheses are satisfied for the application of Von Neumann’s Minimax theorem. (See [10, Thm. 3.4.6]). This tells us that J has a saddlepoint $(e^*, \lambda^*) \in \text{co } E \times \tilde{A}^r$, i.e.,

$$J(e^*, \lambda) \leq J(e^*, \lambda^*) \leq J(e, \lambda^*) \text{ for all } e \in E \text{ and } \lambda \in \tilde{A}^r.$$

Now, notice that (15) implies that, for any $e' \in E$ (and in particular $e' = e^*$)

$$0 \leq \max_{\lambda} J(e', \lambda).$$

But then, by the saddlepoint condition,

$$0 \leq \max_{\lambda} J(e^*, \lambda) = J(e^*, \lambda^*) \leq J(e, \lambda^*),$$

for any $e \in E$. This inequality tells us that

$$\int_0^1 \Delta f(t, u(t)) \cdot S^T(1, t) \left(\sum_{j=0}^r \lambda_j^* \nabla g_j(\bar{x}(1)) \right) dt \geq 0, \quad (16)$$

for any control function $u \in \mathcal{U}$. A standard ‘needle variation’ argument permits us to conclude from this last relation that

$$\Delta f(t, u) \cdot S^T(1, t) \left(\sum_{j \in I(\bar{x})} \lambda_j^* \nabla g_j(\bar{x}(1)) \right) \geq 0 \text{ for all } u \in \Omega, \text{ a.e. } t \in [0, 1].$$

This is the minimum principal condition, in which the endpoint constraint Lagrange multiplier vector is λ^* .

(iii): This assertion is confirmed by the example of Sect. 6. $\square$

6 Example One

Consider the following example of (P), in which the state and control dimensions are both 2 and there is one endpoint constraint.

$$(P) \begin{cases} \text{Minimize } [0.5, -1] x(1) \\ \text{subject to} \\ \dot{x}(t) = u, \quad \text{a.e. } t \in [0, 1], \\ u(t) \in \Omega := \{[0, 0]^T, [1, 0]^T, [0, 1]^T\}, \quad \text{a.e. } t \in [0, 1], \\ x(0) = 0 \text{ and } [-1, 0.5] x(1) \leq 0 \end{cases}$$

Take $(\bar{x}, \bar{u}) \equiv (0, 0)$. (H1) and (H2) are satisfied.

- Proposition 6.1(a):** *The admissible process $(\bar{x}, \bar{u})$ is not a minimizer;*
 (b): *the first-order minimax condition is satisfied at $(\bar{x}, \bar{u})$,*
 (c): *the integral minimax condition (and therefore also the minimum principle condition) is not satisfied at $(\bar{x}, \bar{u})$.*

Proof Consider the process (u, x) such that $u(t) = [1, 0]^T$ on $[0, 1/2]$ and $\tilde{u}(t) = [0, 1]^T$ on $(1/2, 1]$. Then, $x(1) = [1/2, 1/2]^T$. We see that the values of the cost and endpoint constraint functional are, respectively,

$$[0.5, -1] x(1) = -1/2 \text{ and } [-1, 0.5] x(1) = -1/2,$$

i.e., (u, x) is admissible and has strictly negative cost. It follows that $(x \equiv [0, 0]^T, u \equiv [0, 0]^T)$ cannot therefore be a minimizer.

(b): Using our generic notation, we have

$$\begin{aligned} S(s, t) &\equiv I_{2 \times 2} \text{ for all values of } s \text{ and } t, \\ \{\Delta f(t, u) : u \in \Omega\} &= \{[0, 0]^T, [1, 0], [0, 1]^T\}, \\ \nabla g_0(\bar{x}(1)) &= [0.5, -1]^T \text{ and } \nabla g_1(\bar{x}(1)) = [-1, 5]^T \end{aligned}$$

Then, for any time $t \in [0, 1]$,

$$\begin{aligned} \max_{j \in \{0, \dots, r\}} \Delta f(t, u) \cdot S(t, 1) \nabla g_j(\bar{x}(1)) &= \\ \begin{cases} \max \{ [0, 0][0.5, -1]^T, [0, 0][-1, 0.5]^T \} = 0 & \text{if } u = [1, 0]^T \\ \max \{ [1, 0][0.5, -1]^T, [1, 0][-1, 0.5]^T \} = 0.5 & \text{if } u = [1, 0]^T \\ \max \{ [0, 1][0.5, -1]^T, [0, 1][-1, 0.5]^T \} = 0.5 & \text{if } u = [0, 1]^T. \end{cases} \end{aligned}$$

It follows that the first-order minimax condition is satisfied at $(\bar{u}, \bar{x})$.

(c): For $u \in \mathcal{U}$ as in part (a), we calculate

$$\begin{aligned} \max_j \int_0^1 \Delta f(t, u(t)) \cdot S^T(1, t) \nabla g_j(\bar{x}(1)) dt &= \\ \max \{ [1/2, 1/2][0.5, -1]^T, [1/2, 1/2][-1, 0.5]^T \} &= -1/2 < 0. \end{aligned}$$

We have shown that the integrated first-order minimax condition is violated. $\square$

7 State Constraints

Consider the state constrained problem formulated in the introduction

$$(S) \begin{cases} \text{Minimize } g_0(x(1)) \\ \text{over absolutely continuous functions } x \text{ and meas. functions } u \text{ such that} \\ \dot{x}(t) = f(x(t), u(t)), \text{ a.e. } t \in [0, 1], \\ h_k(x(t)) \leq 0, \text{ for all index values } k \in \{1, \dots, N_s\} \text{ and } t \in [0, 1], \\ u(t) \in \Omega, \text{ a.e. } t \in [0, 1], \\ x(0) = x_0 \text{ and } g_j(x(1)) \leq 0, \text{ for } j = 1, 2, \dots, r. \end{cases}$$

For $k = 1, \dots, N_s$, $A_k(\bar{x})$ will denote the set of times at which the k 'th state constraint is active, for the nominal process $(\bar{x}, \bar{u})$, that is

$$A_k(\bar{x}) := \{t \in [0, 1] : h_k(\bar{x}(t)) = 0\}, \text{ for } k = 0, \dots, r.$$

We derive similar necessary conditions of optimality to those of Sect. 4, but now allowing for state constraints.

Theorem 7.1 The First-Order Minimax Condition (State Constraints)

Let $(\bar{x}, \bar{u})$ be a minimizer for (S). Assume (H1)-(H2) are satisfied. Then, for a.e. $t \in [0, 1]$,

$$\begin{aligned} \max_{j \in I(\bar{x})} \{ (\Delta f(t, u) \cdot S^T(1, t) \nabla g_j(\bar{x}(1))) \} \\ \vee \max_{k \in \{1, \dots, N_s\}} \left(\max_{t' \in A_k(\bar{x}) \cap [t, 1]} \{ \Delta f(t', u) \cdot S^T(t', t) \nabla h_k(\bar{x}(t')) \} \right) \geq 0, \end{aligned}$$

for all $u \in \Omega$.

Proof We may assume that g_0 satisfies $g_0(\bar{x}(1)) = 0$. Since $(\bar{x}, \bar{u})$ is a minimizer, we must have

$$g_0(x(1)) \vee \dots \vee g_r(x(1)) \vee \max_{t \in [0,1]} h_j(x(t)) \vee \dots \vee \max_{t \in [0,1]} h_{N_s}(x(t)) \geq 0, \quad (17)$$

for all processes (x, u) for (S).

Let $\mathcal{T}$ be the subset of points $t \in [0, 1]$ with the following properties: t is a Lebesgue point of $s \rightarrow f(\bar{x}(s), \bar{u}(s))$, $\dot{x}(t)$ exists, $\dot{x}(t) = f(\bar{x}(t), \bar{u}(t))$ and $\bar{u}(t) \in \Omega$. Define, for arbitrary $\epsilon > 0$ and $k \in \{1, \dots, N_s\}$,

$$A_k^\epsilon(\bar{x}) := \{t \in [0, 1] : h_k(\bar{x}_k(t)) \geq -\epsilon\}. \quad (18)$$

We claim it suffices to prove a modified version of the theorem, in which we require, for arbitrary $t \in \mathcal{T}$ and $\epsilon > 0$,

$$\begin{aligned} & \max_{j \in I(\bar{x})} \left\{ (\Delta f(t, u(t)) \cdot S^T(1, t) \nabla g_j(\bar{x}(1))) \right\} \\ & \vee \max_{k \in \{0, \dots, N_s\}} \left(\max_{t' \in A_k^\epsilon(\bar{x}) \cap [t, 1]} \left\{ \Delta f(t', u) \cdot S^T(t', t) \nabla h_k(\bar{x}(t)) \right\} \right) \geq 0 \end{aligned}$$

for all $u \in \Omega$ and $t' \geq t$.

Indeed, if this modified condition was true for arbitrary $\epsilon > 0$, then it would be valid for each $\epsilon = \epsilon_i$, $i = 1, 2, \dots$, where $\epsilon_i \downarrow 0$. This means that, for each $t \in \mathcal{T}$ and $u \in \Omega$ and every i , either there exists $j(i) \in I(\bar{x})$ such that

$$\Delta f(t, u(t)) \cdot S^T(1, t) \nabla g_{j(i)}(\bar{x}(1)) \geq 0$$

or there exist $k(i) \in \{1, \dots, N_s\}$ and $t'(i) \in A_{k(i)}^{\epsilon_i} \cap [t, 1]$ such that

$$\Delta f(t'(i), u) \cdot S^T(t'(i), t) \nabla h_{k(i)}(\bar{x}(t)) \geq 0.$$

By extracting subsequences, we can arrange the $i(j) = \bar{i} \in I(\bar{x})$, $k(i) = \bar{k}$ for all i and $t'_i \rightarrow \bar{t}'$ as $i \rightarrow \infty$, for some $\bar{i} \in I(\bar{x})$, $\bar{k} \in \{1, \dots, N_s\}$ and $\bar{t}' \in A_{\bar{k}} \cap [t, 1]$. Since, for each k ,

$$A_k(\bar{x}) \cap [t, 1] = \{\lim_i t_i : t_i \in A_k^{\epsilon_i}(\bar{x}) \cap [t, 1] \text{ for each } i\},$$

we can deduce that either there exists $\bar{j} \in I(\bar{x})$ such that

$$\Delta f(t, u(t)) \cdot S(t, 1) \nabla g_{\bar{j}}(\bar{x}(1)) \geq 0$$

or there exist $\bar{k} \in \{1, \dots, N_s\}$ and $\bar{t}' \in A_{\bar{k}}(\bar{x})$ such

$$\Delta f(\bar{t}', u) \cdot S(\bar{t}', t) \nabla h_{\bar{k}}(\bar{x}(t)) \geq 0.$$

These relations combine to yield the required (stronger) necessary condition of the theorem statement.

Assume, for some $\epsilon > 0$, the modified condition above is false; we show that this leads to a contradiction. Then, for some $t \in \mathcal{T}$, $u \in \Omega$ and $\gamma > 0$,

$$\Delta f(t, u) \cdot S^T(1, t) \nabla g_j(\bar{x}(1)) \leq -\gamma, \text{ for all } j \in I(\bar{x}) \quad (19)$$

and

$$\begin{aligned} \Delta f(t, u) \cdot S^T(t', t) \nabla h_k(\bar{x}(t')) \\ \leq -\gamma, \text{ for all } k \in \{1, N_s\}, t' \in A_k^\epsilon(\bar{x}) \cap [t, 1]. \end{aligned} \quad (20)$$

Take $\sigma_i \downarrow 0$ and, for each i , let $u_i \in \mathcal{U}$ to be control function

$$\begin{cases} u_i(s) = u & \text{if } s \in [\bar{t}, \bar{t} + \sigma_i] \\ u_i(s) = \bar{u}(s) & \text{otherwise.} \end{cases}$$

Since $g_j(\bar{x}(1)) \leq 0$ for all $j \in \{0, \dots, r\}$ and $h_k(\bar{x}(t)) \leq 0$ for all $t \in [0, 1]$ and $k \in \{0, \dots, N_s\}$, we deduce from Proposition 3.1 that there exists a modulus of continuity $\theta(\cdot)$ such that, for each $j \in \{0, \dots, r\}$ and $j = 1, 2, \dots$,

$$\begin{aligned} \sigma_i^{-1} g_j(x_i(1)) &\leq \sigma_i^{-1} (g_j(x_i(1)) - g_j(\bar{x}(1))) \\ &\leq \sigma_i^{-1} \int_t^{t+\sigma_i} \Delta f(s, u) \cdot S^T(1, s) \nabla g_j(\bar{x}(1)) ds + \theta(\sigma_i). \end{aligned} \quad (21)$$

and, for each $k \in \{1, \dots, N_s\}$ and $\sigma'_i := \sigma_i \wedge (t' - t)$,

$$\begin{aligned} h_k(x_i(t')) &\leq h_k(x_i(t')) - h_k(\bar{x}(t')) \\ &\leq \int_t^{t+\sigma'_i} (\Delta f(s, u) \cdot S^T(t', s) \nabla h_k(\bar{x}(t'))) ds + \sigma'_i \times \theta(\sigma'_i). \end{aligned} \quad (22)$$

Since t is a Lebesgue point, there exists a further modulus of continuity $\psi : (0, \infty) \rightarrow (0, \infty)$ (that does not depend on j, k or $t'(\geq \bar{t})$) such that

$$\begin{aligned} \sigma_i^{-1} \int_t^{t+\sigma_i} \Delta f(s, u) \cdot S^T(1, s) \nabla g_j(\bar{x}(1)) ds \\ \leq \Delta f(t, u) \cdot S^T(1, t) \nabla g_j(\bar{x}(1)) + \psi(\sigma_i). \end{aligned} \quad (23)$$

and

$$\begin{aligned} \int_t^{t+\sigma'_i} \Delta f(s, u) \cdot S^T(t', s) \nabla h_k(\bar{x}(t')) ds \\ \leq \sigma'_i \Delta f(t, u) \cdot S^T(1, t) \nabla h_k(\bar{x}(1)) + \sigma'_i \psi(\sigma'_i). \end{aligned} \quad (24)$$

Since $g_j(\bar{x}(1)) \leq 0$ for each j , it follows from (19), (21), (22), (23) and (24) that, for all i sufficiently large,

$$\sigma_i^{-1} g_j(x_i(1)) \leq -\gamma/2 \text{ for all } j \in I(\bar{x}). \quad (25)$$

Note, on the other hand, that for each $j \notin I(\bar{x})$, $g_j(\bar{x}(1)) < 0$. So there exists $\gamma_1 > 0$ such that

$$g_j(\bar{x}(1)) \leq -\gamma_1 \text{ for each } j \notin I(\bar{x}). \quad (26)$$

It follows from Proposition 3.1(a) that, for all i sufficiently large,

$$g_j(x_i(1)) \leq -\gamma_1/2 \text{ for all } j \notin I(\bar{x}). \quad (27)$$

Now take any $k \in \{1, \dots, N_s\}$. Since $h_k(\bar{x}(t)) \leq 0$ for all $t \in [0, 1]$, we can deduce from (20) (22) and (24) that, for all i sufficiently large,

$$(\sigma_1 \vee (t' - t))^{-1} h_k(x_i(t')) \leq -\gamma/2 \text{ for } t' \in A_k^\epsilon(\bar{x}) \cap (t, 1]. \quad (28)$$

Since x_i coincides with $\bar{x}$ on $[0, t]$ and $h_k(\bar{x}(t)) \leq 0$, we know also that

$$h_k(x_i(t')) \leq 0 \text{ for all } t' \in [0, t]. \quad (29)$$

Since $h_k(\bar{x}(t)) < -\epsilon$ for all $t' \in [0, 1]$ such that $t' \notin A_k^\epsilon(\bar{x})$, we deduce from Proposition 3.1 (a) that, for i sufficiently large,

$$h_k(x_i(t')) \leq -\epsilon/2 \text{ for all } t' \in [0, t] \text{ such that } t' \notin A_k^\epsilon. \quad (30)$$

Relations (25), (27), (28), (29) and (30) combine to tell us that, for i sufficiently large, condition (17) is violated by the admissible state trajectory x_i . The validity of the assertions of the theorem follow from this contradiction. $\square$

Theorem 7.2 (The State Constrained Integral First-Order Minimax Condition)
Let $(\bar{x}, \bar{u})$ be a minimizer for (S). Assume (H1)–(H2) are satisfied. Then

$$\begin{aligned} & \max_{j \in I(\bar{x})} \left\{ \int_0^1 \Delta f(t, u(t)) \cdot S^T(t, 1) \nabla g_j(\bar{x}(1)) dt \right\} \\ & \vee \max_{k \in \{1, \dots, N_s\}} \left(\max_{t' \in A_k(\bar{x})} \left\{ \int_0^{t'} \Delta f(t, u(t)) \cdot S^T(t', t) \nabla h_k(\bar{x}(t')) dt \right\} \right) \geq 0, \end{aligned}$$

for all $u \in \mathcal{U}$.

(We interpret $\max_{t' \in A_j(\bar{x})} \{\dots\} := -\infty$, if $A_j(\bar{x}) = \emptyset$.)

Proof We may assume, without loss of generality, that $g_0(\bar{x}(1)) = 0$. We claim it suffices to prove a weaker form of the theorem, in which the inequality is replaced, for arbitrary $\epsilon > 0$, by the weaker condition

$$\max_{j \in I(\bar{x})} \left\{ \int_0^1 (\Delta f(t, u(t)) \cdot S(t, 1) \nabla g_j(\bar{x}(1))) dt \right\} \\ \vee \max_{k \in \{1, \dots, N_s\}} \left(\max_{t' \in A_k^\epsilon(\bar{x})} \left\{ \int_0^{t'} \Delta f(t', u(t)) \cdot S(t', t) \nabla h_k(\bar{x}(t)) dt \right\} \right) \geq 0,$$

for all $u \in \mathcal{U}$. (Here, the set $A_k^\epsilon(\bar{x})$ is as defined by (18).) Indeed if this weaker condition were true for arbitrary $\epsilon > 0$ then it would be valid for $\epsilon = \epsilon_i, i = 1, 2, \dots$, where $\epsilon_i \downarrow 0$. This means that, for any $u \in \mathcal{U}$, either there exists $j(i) \in I(\bar{x})$ such that

$$\int_0^1 \Delta f(t, u(t)) \cdot S(t, 1) \nabla g_{j(i)}(\bar{x}(1)) dt \geq 0, \quad \text{for all } i,$$

or there exists $k(i) \in \{1, \dots, N_s\}$ and $t'(i) \in A_{k(i)}^{\epsilon_i}$ such that

$$\int_0^{t'(i)} \Delta f(t'(i), u(t)) \cdot S(t'(i), t) \nabla h_{k(i)}(\bar{x}(t')) dt \geq 0, \quad \text{for all } i.$$

By extracting subsequences, we can arrange that there exist $\bar{j} \in I(\bar{x})$, $\bar{k} \in \{1, \dots, N_s\}$ and $\bar{t}' \in A_{\bar{k}}(\bar{x})$ such that $j(i) = \bar{j}$, $k(i) = \bar{k}$, for all i ; furthermore $t'_i \rightarrow \bar{t}'$.

We deduce in the limit that either

$$\int_0^1 \Delta f(t, u(t)) \cdot S(t, 1) \nabla g_{\bar{j}}(\bar{x}(1)) dt \geq 0$$

or

$$\int_0^{\bar{t}'} \Delta f(\bar{t}', u) \cdot S(\bar{t}', t) \nabla h_{\bar{k}}(\bar{x}(t)) dt \geq 0,$$

in which, we recall, $\bar{j} \in I(\bar{x})$, $\bar{k} \in \{1, \dots, N_s\}$ and $\bar{t}' \in A_{\bar{k}}(\bar{x})$. These relations combine to yield the required (stronger) necessary condition of the theorem.

So assume that the assertions of this weaker version of the theorem, in which $A_k^\epsilon(\bar{x})$ replaces $A_k(\bar{x})$ for some arbitrary $\epsilon > 0$, are false; we shall show that this leads to a contradiction. Take any $u \in \mathcal{U}$. Under the contraposition hypothesis, there exists $\gamma > 0$ such that, for all $\sigma \in (0, 1)$,

$$\max_{j \in I(\bar{x})} \left\{ \int_0^1 \Delta f(t, u(t)) \cdot S(t, 1) \nabla g_j(\bar{x}(1)) dt \right\} \\ \vee \max_{k \in \{1, \dots, N_s\}} \left(\max_{t' \in A_k^\epsilon(\bar{x})} \left\{ \int_0^{t'} \Delta f(t, u(t)) \cdot S(t', t) \nabla h_k(\bar{x}(t)) dt \right\} \right) < -\gamma. \quad (31)$$

Take $\sigma_i \downarrow 0$. For each i , let x_i be the relaxed state trajectory corresponding to the relaxed control $((\sigma_i, \bar{u}), (1 - \sigma_i), u)$. We can deduce from the relaxation theorem and the optimality of $(\bar{x}, \bar{u})$ that, for each i ,

$$g_0(x_i(1)) \vee \dots \vee g_r(x_i(1)) \vee \max_{t \in [0,1]} h_1(x_i(t)) \vee \dots \vee \max_{t \in [0,1]} h_{N_s}(x_i(t)) \geq 0. \quad (32)$$

Since $g_j(\bar{x}(1)) \leq 0$ for all $j \in \{0, \dots, r\}$ and $h_k(\bar{x}(t)) \leq 0$ for all $t \in [0, 1]$ and $k \in \{0, \dots, N_s\}$, we deduce from Proposition 3.2 that there exists a modulus of continuity $\theta(\cdot)$ with the following properties: for any i ,

$$\sigma_i^{-1} g_j(x_i(1)) \leq \int_0^1 \Delta f(s, u(s)) \cdot S^T(1, s) \nabla g_j(\bar{x}(1)) ds + \theta(\sigma_i),$$

for $j = 0, \dots, r$ and

$$\sigma_i^{-1} h_k(x_i(t')) \leq \int_0^{t'} \Delta f(s, u(s)) \cdot S^T(t', s) \nabla h_k(\bar{x}(t')) ds + \theta(\sigma_i),$$

for all $t' \in A_k^\epsilon(\bar{x})$, $k = 1, \dots, N_s$.

We conclude from (31) that, for all i sufficiently large,

$$\sigma_i^{-1} g_j(x_i(1)) \leq -\gamma/2 \text{ for all } j \in I(\bar{x})$$

and

$$\sigma_i^{-1} h_k(x_i(t')) \leq -\gamma/2, \text{ for } k \in 1, \dots, N_s \text{ and } t' \in A_k^\epsilon(\bar{x}).$$

Note, on the other hand, that for each $j \notin I(\bar{x})$, $g_j(\bar{x}(1)) < 0$. It follows that there exists $\gamma_1 > 0$ such that

$$g_j(\bar{x}(1)) \leq -\gamma_1 \text{ for each } j \notin I(\bar{x}). \quad (33)$$

Proposition 3.1 now tells us that, for all i sufficiently large,

$$g_j(\bar{x}_i(1)) \leq -\gamma_1/2 \text{ for each } j \notin I(\bar{x}). \quad (34)$$

By the definition of $A_k^\epsilon(\bar{x})$,

$$h_k(\bar{x}(1)) \leq -\epsilon \text{ for each } k \in \{1, \dots, N_s\} \text{ and } t \notin A_k^\epsilon(\bar{x}), \quad (35)$$

and consequently,

$$h_k(x_i(1)) \leq -\epsilon/2 \text{ for each } k \in \{1, \dots, N_s\} \text{ and } t \notin A_k^\epsilon(\bar{x}). \quad (36)$$

We conclude from the preceding four relations that

$$\begin{aligned} & g_0(x_i(1)) \vee \dots \vee g_r(x_i(1)) \vee \max_{t \in [0,1]} h_1(x_i(t)) \vee \dots \vee \max_{t \in [0,1]} h_{N_s}(x_i(t)) \\ & \leq -\min\{\epsilon, \sigma_i \times \gamma, \gamma_1\}/2, \end{aligned}$$

for i sufficiently large. This contradicts (32). The proof is concluded. $\square$

The following necessary condition is a version of the minimum principle for the state constrained problem (S), which makes its relation with the preceding first-order minimax theorem explicit.

Theorem 7.3 (The State Constrained Minimum Principle)

Let $(\bar{x}, \bar{u})$ be a minimizer for (S). Assume (H1)–(H2) are satisfied. Then, there exist $\lambda = \{\lambda_j\} \in (R^+)^{r+1}$ and non-decreasing functions of bounded variation $v_k \in BV(0, 1)$, $k = 1, \dots, N_s$, with the following properties:

- (a): $\lambda_j = 0$ if $j \notin I(\bar{x})$, $\text{supp}\{v_k\} \subset A_k(\bar{x})$ for $k = 1, \dots, N_s$,
- (b) $(\lambda, \{v_k\}) \neq (0, 0)$ and
- (c): for a.e. $t \in [0, 1]$,

$$\begin{aligned} & \Delta f(t, u) \cdot S^T(t, 1) \sum_{j \in I(\bar{x})} \lambda_j \nabla g_j(\bar{x}(1)) \\ & + \Delta f(t, u) \cdot \sum_{k=1}^{N_s} \left(\int_t^1 S^T(t', t) \nabla h_k(\bar{x}(t')) dv_k(t') \right) \geq 0, \text{ for all } u \in \Omega. \end{aligned}$$

Comments.

1. The necessary condition of Theorem 7.1 (the state constrained first-order minimax condition) can also be expressed: for a.e. $t \in [0, 1]$,

$$\begin{aligned} & \max_{j \in I(\bar{x})} \{(\Delta f(t, u(t)) \cdot p_j(t))\} \\ & \vee \max_{k \in \{1, \dots, N_s\}} \left(\max_{t' \in A_k(\bar{x}) \cap [t, 1]} \{\Delta f(t', u) \cdot p_k(t', t)\} \right) \geq 0 \text{ for all } u \in \Omega. \end{aligned} \quad (37)$$

Here, $p_j(t) := S^T(1, t) \nabla g_j(\bar{x}(1))$, for each $j \in I(\bar{x})$ and $p_k(t', t) := S^T(t', t) \nabla h_k(\bar{x}(t'))$, for each $k \in \{1, \dots, N_s\}$. For each $j \in I(\bar{x})$, p_j (the ‘costate function corresponding to g_j ’) satisfies

$$\begin{cases} -\dot{p}_j(t) = f_x^T(\bar{x}(t), \bar{u}(t)) p_j(t) & \text{a.e. } t \in [0, 1], \\ p_j(1) = \nabla g_j(\bar{x}(1)). \end{cases}$$

For each $k \in \{1, \dots, N_s\}$ and $t' \in [0, 1]$, $p_k(t', \cdot)$ (the ‘costate corresponding to h_k at time t' ’) is the solution to the differential equation

$$\begin{cases} -\dot{p}_k(t', t) = f_x^T(\bar{x}(t), \bar{u}(t)) p_k(t', t) & \text{a.e. } t \in [0, t'], \\ p_k(t', t') = \nabla h_k(\bar{x}(t')). \end{cases}$$

2. The necessary condition provided by Theorem 7.2 is implicit in convergence analysis of [11]. It can be expressed as

$$\max_{j \in I(\bar{x})} \left\{ \int_0^1 \Delta f(t, u(t)) \cdot p_j(t) dt \right\} \\ \vee \max_{k \in \{1, \dots, N_s\}} \left(\max_{t' \in A_k(\bar{x})} \left\{ \int_0^{t'} \Delta f(t', u(t)) \cdot p_k(t', t) \right\} \right) \geq 0, \text{ for all } u \in \mathcal{U}.$$

Here, $p_j(t)$ and $p_k(t, s)$ are as defined in the preceding comment.

3. An equivalent version of the necessary condition of Theorem 7.3 is: There exist $\lambda = \{\lambda_j\} \in (R^+)^{r+1}$ and non-decreasing functions of bounded variation $v_k \in BV(0, 1)$, $k = 1, \dots, N_s$, such that $\lambda_j = 0$ if $j \notin I(\bar{x})$, $\text{supp}\{v_k\} \subset A_k(\bar{x})$ for $k = 1, \dots, N_s$, $(\lambda, \{v_k\}) \neq (0, 0)$ and

$$\Delta f(t, u) \cdot p(t) \geq 0 \text{ for all } u \in \Omega, \text{ a.e. } t \in [0, 1],$$

Here, p satisfies the ‘measure driven’ differential equation

$$\begin{cases} -dp(t) = f_x^T(\bar{x}(t), \bar{u}(t))p(t)dt + \sum_{k=1}^{N_s} \nabla h_k(\bar{x}(t))dv_k(t) & \text{a.e. } t \in [0, 1], \\ p(T) = \sum_k \lambda_k \nabla g_k(\bar{x}(1)). \end{cases}$$

In this form, it will be recognized as a special case of the standard state constrained minimum principle ([10, Thm. 9.3.1]).

The following theorem relates the necessary conditions of Theorem 7.1 (the state constrained first-order minimax condition), Theorem 7.2 (the state constrained integral first-order minimax condition) and Theorem 7.3 (the state constrained minimum principle).

Theorem 7.4 *Assume the data for problem (S) satisfy hypotheses (H1)–(H2), for some arbitrary admissible process $(\bar{x}, \bar{u})$. We have:*

- (i): *An admissible process $(\bar{x}, \bar{u})$ satisfies the state constrained integral first-order minimax condition if and only if it satisfies the (state constrained) minimum principle.*
- (ii): *Data can be chosen for problem (S) and an admissible process $(\bar{x}, \bar{u})$ such that the first-order minimax condition confirms that $(\bar{x}, \bar{u})$ is not a minimizer, but the minimum principle fails to do so. The converse is also true: Data can be chosen for problem (S) and an admissible process $(\bar{x}, \bar{u})$ such that the minimum principle confirms that $(\bar{x}, \bar{u})$ is not a minimizer, but the first-order minimax condition fails to do so.*

Comment:

The assertions of Theorem 7.4 remain true even if we replace the classical state constrained minimum principle by Arutyunov and Aseev’s strengthened ‘non-degenerate’ state minimum principle [4]. This is because in Example Two,

establishing that, for some (non-optimal) admissible process $(\bar{x}, \bar{u})$, the first-order minimax condition is not satisfied, while the minimum principle is satisfied, both endpoints of $\bar{x}$ are interior to the state constraint set. For such problems, the controllability hypotheses of [4] are automatically satisfied because there are no nonzero normal vectors to the state constraint set at $\bar{x}(0)$ or $\bar{x}(1)$ and the assertions of both the original minimum principle and its non-degenerate form are the same.

- Proof** (i): The first assertion in (i) is already confirmed by Example One of Sect. 6, in the special case of problem (P) in which none of the state constraints are active. The second assertion is confirmed by Example Two in Sect. 8.
- (ii) First, suppose that necessary condition of Theorem 7.3 (the minimum principle) is satisfied. Take any $u \in \mathcal{U}$. Inserting $u = u(t)$ into the inequality in part (b) of the theorem statement, integrating w.r.t. t and changing the order of integration give

$$\begin{aligned} & \int_0^1 \Delta f(t, u(t)) \cdot S^T(1, t) \sum_{j \in I(\bar{x})} \lambda_j \nabla g_j(\bar{x}(1)) dt \\ & + \sum_{j=1}^{N_s} \int_0^1 \left(\int_0^{t'} \Delta f(t, u(t)) \cdot S^T(t', t) \nabla h_j(\bar{x}(t')) dt \right) dv_k(t') \geq 0. \quad (38) \end{aligned}$$

If the necessary condition of Theorem 7.2 is violated, we know that $u \in \mathcal{U}$ can be chosen such that, for some $\gamma > 0$,

$$\begin{aligned} & \max_{j \in I(\bar{x})} \left\{ \int_0^1 \Delta f(t, u(t)) \cdot S^T(t, 1) \nabla g_j(\bar{x}(1)) dt \right\} \\ & \vee \max_{k \in \{1, \dots, N_s\}} \left(\max_{t' \in A_k(\bar{x})} \left\{ \int_0^{t'} \Delta f(t, u(t)) \cdot S^T(t', t) \nabla h_k(\bar{x}(t')) dt \right\} \right) \leq -\gamma. \end{aligned}$$

We deduce from the non-triviality of the Lagrange multipliers, together with the facts that the elements $\{\lambda_j\}$ have support in $I(\bar{x})$ and the measures $d\mu_k$ have support in $A_k(\bar{x})$ for each k , that there exists $\rho > 0$ such that

$$\sum_{j \in I(\bar{x})} \lambda_j + \sum_{\{k : A_k(\bar{x}) \neq \emptyset\}} \int_{A_k(\bar{x})} d\mu_k(t) = \rho. \quad (39)$$

It follows from (39) and (39) that

$$\begin{aligned} & \int_0^1 \Delta f(t, u(t)) \cdot S^T(1, t) \sum_{j \in I(\bar{x})} \lambda_j \nabla g_j(\bar{x}(1)) dt \\ & + \sum_{j=1}^{N_s} \int_0^1 \left(\int_0^{t'} \Delta f(t, u(t)) \cdot S^T(t', t) \nabla h_j(\bar{x}(t')) dt \right) dv_k(t') \\ & \leq -\gamma \times \left(\sum_{j \in I(\bar{x})} + \sum_{\{k: A_k(\bar{x}) \neq \emptyset\}} \int_{A_k(\bar{x})} dv_k(t) \right) = \gamma \times \rho < 0. \end{aligned}$$

This contradicts (38). We have shown that the integral first-order minimax condition is satisfied.

Now, suppose that the conditions of the integral first-order minimax theorem are satisfied. Define:

$$\begin{aligned} E &:= \{e \in L^1(0, 1) : e(t) \in \Delta f(t, \Omega) \text{ a.e. } t \in [0, 1]\} \text{ and write} \\ D &:= \{(\lambda, v) \in (R^+)^{(r+1)} \times NBV^+([0, 1]; R^{N_s}) : \\ & \lambda_j = 0 \text{ if } j \notin I(\bar{x}), \text{ supp } \{v_k\} \in A_k(\bar{x}) \text{ and } \sum_j \lambda_j + \sum_k \int_{A_k(\bar{x})} dv_k = 1\}. \end{aligned}$$

Define also the function $J : L^1 \times D \rightarrow R$ to be

$$\begin{aligned} J(e, (\lambda, v)) &:= \int_0^1 e(t) \cdot \left(S^T(1, t) \left(\sum_{j \in I(\bar{x})} \lambda_j \nabla g_j(\bar{x}(1)) \right) \right. \\ & \left. + \sum_{j=1}^{N_s} \left(\int_t^1 S^T(t', t) \nabla h_j(\bar{x}(t')) dv_k(t') \right) \right) dt. \end{aligned}$$

The integral first-order minimax condition can be expressed as

$$\max_{(\lambda, v) \in D} J(e, (\lambda, v)) \geq 0 \text{ for all } e \in E.$$

Now, equip $\overline{\text{co}} E$ with its relative weak L^1 topology. We also take the topology on D to be the relative topology induced on this set by the product topology $E^r \times \mathcal{D}$, where E^r is the Euclidean topology and $\mathcal{D}$ is the weak* topology on $C^*([0, 1]; R^{N_s})$. Since (with respect to these topologies) $e \rightarrow J(e, (\lambda, v))$ is continuous and linear for fixed (λ, v) , we can deduce from the preceding relation that

$$\max_{(\lambda, v) \in D} J(e, (\lambda, v)) \geq 0 \text{ for all } e \in \overline{\text{co}} E. \quad (40)$$

Now, apply the Von Neumann Minimax theorem, taking the pay-off function to be J with domain $\overline{\text{co}} E \times D$ and identifying the topologies as above. We thereby obtain

an element $(e^*, (\lambda^*, v^*)) \in \overline{\text{co}} E \times D$ such that

$$J(e^*, (\lambda, v)) \leq J(e^*, (\lambda^*, v^*)) \leq J(e, (\lambda^*, v^*)) \text{ for all } (e, (\lambda, v)) \in \overline{\text{co}} E \times D.$$

It follows from (40) that $0 \leq \max_{(\lambda, v) \in D} J(e^*, (\lambda, v))$. From the saddlepoint condition then

$$0 \leq \max_{(\lambda, v) \in D} J(e^*, (\lambda, v)) = J(e^*, (\lambda^*, v^*)) \leq J(e, (\lambda^*, v^*)) \text{ for all } e \in E.$$

We conclude that, for all $u \in \mathcal{U}$,

$$\begin{aligned} \int_0^1 \Delta f(t, u(t)) \cdot \left(S^T(1, t) \left(\sum_{j \in I(\bar{x})} \lambda_j \nabla g_j(\bar{x}(1)) \right) \right. \\ \left. + \sum_{j=1}^{N_s} \left(\int_t^1 S^T(t', t) \nabla h_j(\bar{x}(t')) dv_k(t') \right) \right) dt \geq 0. \end{aligned}$$

This relation will be recognized as the state constrained minimum principle condition. The proof is complete. $\square$

8 Example Two

Consider the following example of Problem (S), the data which satisfy hypotheses (H1) and (H2):

$$\begin{cases} \text{Minimize } x_1(1) \text{ subject to} \\ \dot{x}_1(t), \dot{x}_2(t) = (x_2(t), u(t)), \text{ a.e. } t \in [0, 1], \\ u \in [-1, +1], \text{ a.e. } t \in [0, 1], \\ x_1(t) \leq 0, \text{ for all } t \in [0, 1], \\ x_1(0) = -\frac{1}{8} \text{ and } x_2(0) = \frac{1}{2}. \end{cases}$$

Take as nominal admissible process $(\bar{x}, \bar{u})$, in which $\bar{u}(t) = \begin{cases} -1 & \text{if } 0 \leq t \leq \frac{1}{2} \\ -\frac{1}{2} & \text{if } \frac{1}{2} < t \leq 1 \end{cases}$ and

$$(\bar{x}_1(t), \bar{x}_2(t)) = \begin{cases} (-\frac{1}{8} + \frac{1}{2}t - \frac{1}{2}t^2, \frac{1}{2} - t) & \text{if } 0 \leq t \leq \frac{1}{2} \\ (-\frac{1}{4} \times (t - \frac{1}{2})^2, -\frac{1}{2}(t - \frac{1}{2})) & \text{if } \frac{1}{2} < t \leq 1 \end{cases}.$$

Proposition 8.1(a): $(\bar{x}, \bar{u})$ is not a minimizer for (S),
 (b): the first-order minimax condition is not satisfied at $(\bar{x}, \bar{u})$,
 (c): the minimum principle condition is satisfied at $(\bar{x}, \bar{u})$.

Proof(a): $(\bar{x}, \bar{u})$ is an admissible process and has cost $-\frac{1}{16}$. Since the cost of the admissible process (x', u') in which

$$u'(t) = 1 \text{ and } (x'_1(t), x'_2(t)) = (-\frac{1}{8} + \frac{1}{2}t - \frac{1}{2}t^2, \frac{1}{2} - t) \text{ for } t \in [0, 1]$$

has cost $-\frac{1}{8}$, the nominal process $(\bar{x}, \bar{u})$ is not a minimizer.

(b): For any $t \in (\frac{1}{2}, 1)$, the left side of the first-order minimax condition at time t , for control value $u = -1$, is

$$\Delta f(t, u) \cdot S^T(1, t) \nabla g(\bar{x}(1)) = (-1 + \frac{1}{2}) \times (1 - t) = -\frac{1}{2} \times (1 - t) < 0.$$

(Notice that $A_1(\bar{x}) \cap [t, 1]$ is empty, so the state constraint makes no contribution to the relation.) We see that the first-order minimax condition is violated, confirming that $(\bar{x}, \bar{u})$ is not a minimizer.

(c): The minimum principle (in the traditional form described in Comment 2 after the statement of Theorem 7.3) is satisfied for $(\bar{x}, \bar{u})$, when we choose

$$\lambda_0 = 0, \quad dv(t) = \frac{1}{2} \times \delta(t - \frac{1}{2}) dt$$

$(\delta(t - \frac{1}{2}))$ denotes the Dirac delta function concentrated at time $t = \frac{1}{2}$ and

$$(p_1(t), p_2(t)) = \begin{cases} (1, \frac{1}{2} - t) & \text{for } 0 \leq t \leq \frac{1}{2} \\ (0, 0) & \text{for } \frac{1}{2} < t \leq 1. \end{cases}$$

□

Comment: To clarify the ideas behind Example Two, we note that it is an optimal control problem in which the velocity function $f(x, u)$ is linear in x and u , the state constraint function $h(x)$ and right endpoint cost function $g(x)$ are linear in x , and there is no right endpoint state constraint. For such problems, the minimum principle is not satisfied at an admissible nominal process $(\bar{x}, \bar{u})$ if and only if there is exists an admissible process (x, u) such that

$$g(x(1)) < g(\bar{x}(1)) \text{ and } \max_{s \in [0, 1]} h(x(s)) < 0.$$

For the problem of Example Two, we have

$$\max_{t \in [0, 1]} h(x(t)) = 0 \text{ for all admissible processes } (x, u).$$

So the minimum principle is satisfied at all admissible processes. On the other hand, it is easy to check that the first-order minimax condition is not satisfied at the specified

admissible process $(\bar{x}, \bar{u})$ if there exists an admissible process (x', u') and $t \in (0, 1)$ such that

$$g(x'(1)) < g(\bar{x}'(1)) \text{ and } \max_{s \in [t, 1]} h(x(s)) < 0.$$

Since, in Example Two, we exhibit such a process (x', u') , the first-order minimax condition is not satisfied at $(\bar{x}, \bar{u})$.

Conclusions

It is important to investigate the strength of the optimality condition associated with a particular optimal control algorithm, because this gives insights into anomalous situations where the algorithm fails to generate minimizing processes. We have investigated two optimality conditions that have arisen in earlier convergence analysis, the first-order minimax conditions and the integrated first-order minimax conditions, and have compared them with the minimum principle. For problems without pathwise state constraints, we find that the minimum principle is strictly stronger than the first-order minimax and the minimum principle and the integrated first-order minimax condition are equivalent. For problems with state constraints, we have found once again that the minimum principle and the integrated first-order minimax condition are equivalent, but the (non-integrated) first-order minimax condition is neither strictly stronger nor strictly weaker than the minimum principle. We provide an example in which the first-order minimax condition can be used to exclude a non-minimizer, but the minimum principle fails to do so. This example establishes that the first-order minimax condition for state constrained problems is an independent necessary condition that can, in certain circumstances, supply more information than the minimum principle.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Arutyunov, A.V., Aseev, S.M.: Investigation of the degeneracy phenomenon of the maximum principle for optimal control problems with state constraints. *SIAM J. Control Optim.* **35**, 930–952 (1977)
2. Clarke, F.H.: *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983; reprinted as vol.5 of *Classics in Appl. Math.*, SIAM, Philadelphia, PA (1990)
3. Jacobson, D.H., Mayne, D.Q.: *Differential Dynamic Programming*. Elsevier Press, New York (1970)
4. Mayne, D.Q., Polak, E.: First-order strong variation algorithms for optimal control. *J. Optim. Theory Appl.* **16**, 277–301 (1975)

5. Polak, E.: Computational Methods in Optimization: A Unified Approach. Academic Press, New York (1970)
6. Polak, E.: Optimization: algorithms and consistent approximations, Springer Verlag, Applied Math. Sciences, Vol. 124, New York (1997)
7. Polak, E., Mayne, D.Q.: First-order strong variation algorithms for optimal control problems with terminal inequality constraints. *J. Optim. Theory Appl.* **16**, 303–325 (1975)
8. Pontryagin, L.S., Boltyanskii, V.G., Gramkelidze, R.V., Mischenko, E.F.: The Mathematical Theory of Optimal Processes, Tririgoff, K. N. , Transl., L. W. Neustadt, Ed., Wiley, New York (1962)
9. Pytlak, R., Vinter, R.B.: A feasible directions type algorithm for optimal control problems with state and control constraints. *SIAM J. Control Optim.* **36**, 1999–2019 (1998)
10. Vinter, R.B.: Optimal Control. Birkhäuser, Boston (2000)
11. Warga, J.: Optimal Control of Differential and Functional Equations. Academic Press, New York (1972)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.