



Free End-Time Optimal Control Problems: Conditions for the Absence of an Infimum Gap

Richard B. Vinter¹

Received: 26 November 2018 / Accepted: 2 June 2019 / Published online: 1 August 2019
© The Author(s) 2019

Abstract

This paper concerns free end-time optimal control problems, in which the dynamic constraint takes the form of a controlled differential inclusion. Such problems may fail to have a minimizer. Relaxation is a procedure for enlarging the domain of an optimization problem to guarantee existence of a minimizer. In the context of problems studied here, the standard relaxation procedure involves replacing the velocity sets in the original problem by their convex hulls. It is desirable that the original and relaxed versions of the problem have the same infimum cost. For then we can obtain a sub-optimal state trajectory, by obtaining a solution to the relaxed problem and approximating it. It is important, therefore, to investigate when the infimum costs of the two problems are the same; for otherwise the above strategy for generating sub-optimal state trajectories breaks down. We explore the relation between the existence of an infimum gap and abnormality of necessary conditions for the free-time problem. Such relations can translate into verifiable hypotheses excluding the existence of an infimum gap. Links between existence of an infimum gap and normality have previously been explored for fixed end-time problems. This paper establishes, for the first time, such links for free end-time problems.

Keywords Necessary conditions · Optimal control · Ordinary differential equations · State constraints

Mathematics Subject Classification (2010) Primary 49N · 49K

Dedicated to Professor Alexander Ioffe in honor of his 80th birthday.

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

¹ Department of Electrical and Electronic Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

1 Introduction

Consider the optimal control problem

$$(P) \quad \begin{cases} \text{Minimize } g(T, x(0), x(T)) \\ \text{over } T \geq 0, \text{ absolutely continuous functions } x(\cdot) : [0, T] \rightarrow \mathbb{R}^n \\ \text{and measurable functions } u(\cdot) \text{ satisfying} \\ \dot{x}(t) = f(t, x(t), u(t)) \text{ a.e. } t \in [0, T], \\ u(t) \in U(t) \text{ a.e. } t \in [0, T], \\ (T, x(0), x(T)) \in C, \end{cases}$$

the data for which comprise: functions $g : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ and $f : [0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$, a closed set $C \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ and a multifunction $U(\cdot) : [0, +\infty) \rightsquigarrow \mathbb{R}^m$.

A process $(T, x(\cdot), u(\cdot))$ comprises a number $T \geq 0$ and a pair of functions, of which the first $x(\cdot) : [0, T] \rightarrow \mathbb{R}^n$ is an absolutely continuous function, and the second $u(\cdot) : [0, T] \rightarrow \mathbb{R}^m$ is a measurable function satisfying

$$\dot{x}(t) = f(t, x(t), u(t)) \quad \text{and} \quad u(t) \in U(t) \quad \text{a.e. } t \in [0, T].$$

A process $(T, x(\cdot), u(\cdot))$ is said to be admissible if $(T, x(0), x(T)) \in C$. The second component $x(\cdot)$ of an (admissible) process $(T, x(\cdot), u(\cdot))$ is called an (admissible) state trajectory and the third an (admissible) control function.

We say that a process $(\bar{T}, \bar{x}(\cdot), \bar{u}(\cdot))$ is a minimizer if it achieves the minimum of $g(T, x(0), x(T))$ over all admissible processes $(T, x(\cdot), u(\cdot))$. It is called a strong local minimizer if, for some $\varepsilon > 0$,

$$g(T, x(0), x(T)) \geq g(\bar{T}, \bar{x}(0), \bar{x}(T))$$

for all processes $(T, x(\cdot), u(\cdot))$ such that $|T - \bar{T}| + \|x^e(\cdot) - \bar{x}^e(\cdot)\|_{L^\infty} \leq \varepsilon$.

Here, $x^e(\cdot) : [0, \infty) \rightarrow \mathbb{R}^n$ denotes the extension, by constant extrapolation from the right, of the function $x(\cdot) : [0, T] \rightarrow \mathbb{R}^n$.

We write $\inf(P)$, ‘the infimum cost’, for the infimum value of the cost function $g(T, x(0), x(T))$ over the set of admissible processes $(T, x(\cdot), u(\cdot))$.

Hypotheses ensuring the existence of a unique state trajectory for every control function and initial state and the continuous dependence of this state trajectory on these quantities, the continuity of the cost function and closedness of C and the values of $U(\cdot)$, and the non-emptiness and boundedness of the set of admissible processes, are not, alone, sufficient to guarantee existence of solutions to (P). Existence of solutions is assured only if an additional ‘geometric’ hypothesis is satisfied; this typically places a convexity requirement on the velocity sets $f(t, x, U(t)) := \{f(t, x, u) \mid u \in U(t)\}$, namely:

$$(C): \quad f(t, x, U(t)) \text{ is convex for all } t \in [0, \infty), x \in \mathbb{R}^n.$$

Relaxation is a procedure for enlarging the domain of an optimization problem, but not too much, to guarantee existence of minimizers. Here ‘not too much’ means, that elements in the enlarged domain can be closely approximated by elements in the original domain. Typically, the enlarged domain is the closure, in some sense, of the original domain. Relaxation opens a way, if not to solving the original problem (which might not be possible), but to generating a sub-optimal minimizer of the original problem, i.e an element whose cost is arbitrarily close to the infimum cost of the original problem. The idea is to solve the relaxed problem and approximate the solution by elements in the original domain.

The concept of relaxation makes sense for any optimization problem. In the context of optimal control, however, relaxation is usually taken to mean modifying the optimal control

problem in such a manner that the original velocity sets are replaced by their convex hulls. It will be clear from the preceding discussion that the relaxed problem satisfies the extra geometric condition, and therefore has solutions. On the other hand, the Relaxation Theorem, which tells us that relaxed state trajectories can be approximated arbitrarily closely by original state trajectories, w.r.t. the supremum norm, means that relaxation via convexification meets the other ‘not too much’ requirement of the procedure.

There are a number of ways of convexifying the velocity sets. Following Gamkrelidze [3], we adopt the following formulation of the relaxed problem (R):

$$(R) \quad \begin{cases} \text{Minimize } g(T, x(0), x(T)) \\ \text{over } T \geq 0, \text{ absolutely continuous functions } x(\cdot) : [0, T] \rightarrow \mathbb{R}^n \\ \text{and measurable functions } (u^0(\cdot), \dots, u^n(\cdot)), (\mu^0(\cdot), \dots, \mu^n(\cdot)) \text{ satisfying} \\ \dot{x}(t) = \sum_{j=0}^n \mu^j(t) f(t, x(t), u^j(t)) \text{ a.e. } t \in [0, T], \\ (u^0(t), \dots, u^n(t)) \in U(t) \times \dots \times U(t) \text{ a.e. } t \in [0, T], \\ (\mu^0(t), \dots, \mu^n(t)) \in \Sigma \text{ a.e. } t \in [0, T], \\ (T, x(0), x(T)) \in C. \end{cases}$$

Here, Σ is the set of simplicial index values in n -dimensional space:

$$\Sigma := \left\{ (\mu^0 \geq 0, \dots, \mu^n \geq 0) \mid \sum_{j=0}^n \mu^j = 1 \right\}.$$

Processes, state trajectories and strong local minimizers for (R) are referred to as ‘relaxed processes’, ‘relaxed state trajectories’ and ‘relaxed strong local minimizers’, respectively. A (non-relaxed) process $(T, x(\cdot), u(\cdot))$ is a special case of a relaxed process $(x(\cdot), \{(\mu^k(\cdot), u^k(\cdot))\}_{k=0}^n)$, in which $\mu^0 = 1, \mu^1 = 0, \dots, \mu^n = 0$, and $u^0(\cdot) = u^1(\cdot) = \dots = u^n(\cdot) = u(\cdot)$.

Notice that the velocity sets for the relaxed problem are

$$\left\{ \sum_{j=0}^n \mu^j f(t, x, u^j) \mid (u^0, \dots, u^n) \in U(t) \times \dots \times U(t), (\mu^0, \dots, \mu^n) \in \Sigma \right\} \\ = \text{co}\{f(t, x, U(t))\} \quad \text{for } t \in [0, \infty), x \in \mathbb{R}^n.$$

So the convexity hypothesis (C) is satisfied and, under unrestrictive additional hypotheses, the relaxed problem has a minimizer. Denote the minimum cost by $\min(R)$.

While relaxed state trajectories can be approximated arbitrarily closely in the supremum norm by non-relaxed state trajectories, it is nonetheless possible that there is an ‘infimum gap’, i.e.,

$$\min(R) < \inf(P).$$

The pathology here is that, while according to the Relaxation Theorem, feasible relaxed state trajectories can be approximated arbitrary closely by non-relaxed state trajectories, doing so may be incompatible with satisfying all the constraints in the original problem. It is important to identify situations where an infimum gap arises because, then, the earlier-discussed justification for studying the relaxed problem is no longer valid.

The aim of this paper is to derive sufficient conditions for absence of an infimum gap, for free end-time optimal control problems. These conditions require necessary conditions of optimality, expressed in terms of the free end-time Pontryagin Maximum Principle, to be satisfied in normal form, i.e. for any valid multiplier set, the cost multiplier component must be non-zero.

Two kinds of relationship between occurrence of an infimum gap and abnormality of associated necessary conditions of optimality are explored. In the first we focus attention on a strong local minimizer which cannot also be interpreted as a strong relaxed minimizer; in the second, on a relaxed minimizer, whose cost is strictly less than the infimum cost over admissible (non-relaxed) processes. The following terminology, relating these properties was introduced in [5]:

Type A Relations: A strong local minimizer satisfies the Pontryagin Maximum Principle in abnormal form (i.e., with cost multiplier zero) if, when regarded as a relaxed admissible process, it is not also a relaxed strong local minimizer.

Type B Relations: A relaxed strong local minimizer satisfies the relaxed Pontryagin Maximum Principle in abnormal form if its cost is strictly less than the infimum cost over all admissible processes, whose state trajectories are close (in the L^∞ sense) to that of the relaxed strong local minimizer.

The main results of this paper can be summarized as

- (A): ' $(\bar{T}, \bar{x}(\cdot), \bar{u}(\cdot))$ is a strong local minimizer (for the free end-time problem) but not a relaxed strong local minimizer' implies ' $(\bar{T}, \bar{x}(\cdot), \bar{u}(\cdot))$ satisfies an 'averaged' version of the Pontryagin Maximum Principle in abnormal form', and
- (B): ' $(\bar{T}, \bar{x}(\cdot), \bar{u}(\cdot))$ is a relaxed strong local minimizer with cost strictly less than that of any admissible (non-relaxed) process, whose state trajectory is close (in the L^∞ sense) to that of the relaxed strong local minimizer' implies ' $(\bar{T}, \bar{x}(\cdot), \bar{u}(\cdot))$ satisfies the relaxed Pontryagin Maximum Principle in abnormal form'.

(A) and (B) will be recognized as forms of Type A and Type B relations, respectively. The relations are proved here, for the first time, for general, non-smooth free end-time optimal control problems.

When we limit attention to optimal control problems with fixed initial state x_0 , there is a connection with the conditions excluding an infimum gap, considered in this paper, and sensitivity relations in optimal control. Here, by 'sensitivity relations' we mean information about generalized derivatives of the value function $V(\cdot, \cdot)$ for the relaxed problem at x_0 . It can be deduced from results in [1] that, under a normality hypothesis on the data for the relaxed problem, the convexified limiting subdifferential of $V(0, \cdot)$ at x_0 is bounded. This, in turn, implies Lipschitz continuity of $V(0, \cdot)$ near x_0 . On the other hand, continuity of $V(0, \cdot)$ near x_0 implies 'no infimum gap', in consequence of the Relaxation Theorem (see, e.g., [8, Theorem 2.7.2]). Combining these facts, we arrive at an independent proof of 'normality of the relaxed problem' implies 'no infimum gap'; this is a Type B property. By contrast, our analysis covers problems for which the initial state is not fixed and also supplements conditions for Type B properties by conditions for Type A properties.

Previous work has addressed, exclusively, fixed end-time optimal control problems. Warga was the first to investigate the relation between the existence of an infimum gap and validity of the Pontryagin Maximum Principle in abnormal form. Warga announced a *Type B* relation for optimal control problems with smooth data in his early paper [9]. In his monograph [10] he proved a *Type B* relation for optimal control problems with state constraints. In a subsequent paper [11], Warga generalized his earlier *Type B* results to allow for nonsmooth data, making use of local approximations based on 'derivative containers', developed in [12]. Type A relations were proved, for the first time, by Palladino and Vinter in [5] and [6], for problems in which the dynamic constraint takes the form of a differential inclusion and a controlled differential inclusion, respectively. Related work by Ioffe is reported in [4].

We point out that, when the time dependence of the dynamic constraint is Lipschitz continuous and the control constraint set is independent of time, free end-time problems can be reduced to fixed end-time problems, of standard form, by a change of independent variable. Thus free time optimal control problems, in this special case, are covered by earlier work. However this reduction is not possible in the setting of this paper, because we allow a measurably time dependent dynamic constraint and a measurably time dependent control constraint set.

The following notation will be employed: for vectors $x \in \mathbb{R}^n$, $|x|$ denotes the Euclidean length. \mathbb{B} denotes the closed unit ball in \mathbb{R}^n . Given a set $A \subset \mathbb{R}^n$ and a point $x \in \mathbb{R}^n$, we denote by $d_A(x)$ the Euclidean distance of a point $x \in \mathbb{R}^n$ from A :

$$d_A(x) := \inf\{|x - y| \mid y \in A\}.$$

$W^{1,1}([0, T]; \mathbb{R}^n)$ is the space of absolutely continuous \mathbb{R}^n -valued functions $x(\cdot)$ on $[0, T]$ with norm $|x(0)| + \|\dot{x}(\cdot)\|_{L^1}$. Given $a, b \in \mathbb{R}$, $a \wedge b$ and $a \vee b$ denote $\min\{a, b\}$ and $\max\{a, b\}$, respectively.

We write $W^{1,1}$ or $W^{1,1}(0, T)$ in place of $W^{1,1}([0, T]; \mathbb{R}^n)$, L^∞ or $L^\infty(0, T)$ in place of $L^\infty([0, T]; \mathbb{R}^n)$, etc. when the meaning is clear.

We shall use several constructs of nonsmooth analysis. Given a closed set $D \subset \mathbb{R}^k$ and a point $\bar{x} \in D$, the *limiting normal cone* $N_D(\bar{x})$ of D at \bar{x} is defined to be

$$N_D(\bar{x}) := \left\{ p \mid \exists x_i \xrightarrow{D} \bar{x}, p_i \rightarrow p \text{ s.t. } \limsup_{x \rightarrow x_i} \frac{p_i \cdot (x - x_i)}{|x - x_i|} \leq 0 \text{ for each } i \in \mathbb{N} \right\}.$$

Here, $y_i \xrightarrow{D} y$ indicates that all points in the convergent sequence $\{y_i\}$ lie in D .

Given a lower semicontinuous function $f : \mathbb{R}^k \rightarrow \mathbb{R} \cup \{+\infty\}$ and a point $\bar{x} \in \text{dom } f := \{x \in \mathbb{R}^k \mid f(x) < +\infty\}$, the limiting subdifferential of f at \bar{x} is denoted $\partial f(\bar{x})$:

$$\partial f(\bar{x}) := \left\{ \xi \mid \exists \xi_i \rightarrow \xi \text{ and } x_i \xrightarrow{\text{dom } f} \bar{x} \text{ such that } \limsup_{x \rightarrow x_i} \frac{\xi_i \cdot (x - x_i) - f(x) + f(x_i)}{|x - x_i|} \leq 0 \text{ for all } i \in \mathbb{N} \right\}.$$

For details of definition and properties of these objects, we refer the reader to [2, 7, 8].

2 Conditions for Non-Existence of an Infimum Gap

In this section we state two theorems relating the existence of a gap between the infimum costs for the optimal control problem (P) and its relaxed counterpart (R), and the validity of a Pontryagin Maximum Principle in abnormal form. The following hypotheses, in which $(\bar{T}, \bar{x}(\cdot), \bar{u}(\cdot))$ is a given process and $\bar{\varepsilon} > 0$, will be invoked.

(H1): $f(\cdot, x, u)$ is \mathcal{L} -measurable for each (x, u) and $f(t, \cdot, \cdot)$ is continuous for each t . $U(\cdot)$ is a Borel measurable multifunction taking values compact sets.

(H2): There exist $\bar{\varepsilon} > 0, k \geq 0$ and $c \geq 0$ such that

$$|f(t, x, u) - f(t, x', u)| \leq k|x - x'| \quad \text{and} \quad |f(t, x, u)| \leq c$$

for all $x, x' \in \bar{x}(t) + \bar{\varepsilon}\mathbb{B}, u \in U(t)$, a.e. $t \in [0, \bar{T} + \bar{\varepsilon}]$.

Define the Hamiltonian function

$$H(t, x, u, p) := p \cdot f(t, x, u).$$

Given an essentially bounded function $h(\cdot) : [0, \infty) \rightarrow \mathbb{R}$ and a point $t > 0$, the *essential value* of $h(\cdot)$ at t is the set

$$\text{ess } h(s) := \left[\lim_{\delta \downarrow 0} \left(\text{ess inf}_{s \in [t-\delta, t+\delta]} h(s) \right), \lim_{\delta \downarrow 0} \left(\text{ess sup}_{s \in [t-\delta, t+\delta]} h(s) \right) \right].$$

See [8, Chapter 8], where the main properties of the essential value are stated and proved.

Theorem 1 *Let $(\bar{T}, \bar{x}(\cdot), \bar{u}(\cdot))$ be a strong local minimizer for problem (P) such that $\bar{T} > 0$. Assume that hypotheses (H1) and (H2) are satisfied and that $g(\cdot, \cdot, \cdot)$ is Lipschitz continuous on a neighborhood of $(\bar{T}, \bar{x}(0), \bar{x}(\bar{T}))$.*

(a) *Then there exist an arc $p(\cdot) \in W^{1,1}([0, \bar{T}]; \mathbb{R}^n)$, and $\lambda \geq 0$ such that*

- (i) $(p(\cdot), \lambda) \neq (0, 0)$,
- (ii) $-\dot{p}(t) \in \text{co } \partial_x H(t, \bar{x}(t), p(t), \bar{u}(t))$ a.e. $t \in [0, \bar{T}]$,
- (iii) $(h, p(0), -p(\bar{T})) \in \lambda \partial g(\bar{T}, \bar{x}(0), \bar{x}(\bar{T})) + N_C(\bar{T}, \bar{x}(0), \bar{x}(\bar{T}))$,
- (iv) $H(t, \bar{x}(t), p(t), \bar{u}(t)) \geq H(t, \bar{x}(t), p(t), u)$, for all $u \in U(t)$, a.e. $t \in [0, \bar{T}]$,
- (v) $h \in \text{ess}_{s \rightarrow \bar{T}} \left(\sup_{u \in U(s)} H(s, \bar{x}(\bar{T}), u, p(\bar{T})) \right)$,

in which $\partial_x H(t, x, p, u)$ denotes the limiting subdifferential of $H(t, \cdot, p, u)$.

(b) *Suppose that, for every $\varepsilon > 0$, there exists an admissible relaxed process $(T, x(\cdot), \{(\mu^k(\cdot), u^k(\cdot))\}_{k=0}^n)$ such that*

$$g(\bar{T}, \bar{x}(0), \bar{x}(\bar{T})) > g(T, x(0), x(T))$$

and $|T - \bar{T}| + \|x^\varepsilon(\cdot) - \bar{x}^\varepsilon(\cdot)\|_{L^\infty} \leq \varepsilon$, (i.e., $(\bar{T}, \bar{x}(\cdot), \bar{u}(\cdot))$ is not also a local relaxed minimizer).

Then conditions (i)–(v) above are satisfied for some choice of multipliers $(p(\cdot) : [0, \bar{T}] \rightarrow \mathbb{R}^n, \lambda)$ such that $\lambda = 0$ and in which (ii) is replaced by

$$(ii)' \quad -\dot{p}(t) \in \text{co } \bigcup_{u \in U(t)} \partial_x H(t, \bar{x}(t), p(t), u) \quad \text{a.e. } t \in [0, T].$$

Comments

- (1) Part (a) is a known version of the free end-time Pontryagin Maximum Principle for problems with measurably time dependent data (see, e.g. [8, Chapter 8]).
- (2) Observe that the contrapositive statement of part (b) is a sufficient condition for the absence of an infimum gap (in a local sense): if $(\bar{T}, \bar{x}(\cdot), \bar{u}(\cdot))$ is a strong local minimizer such that, given any multipliers $p(\cdot) \in W^{1,1}$, and $\lambda \geq 0$ satisfying conditions (i), (ii)', (iii)–(iv), we have $\lambda \neq 0$, then $(\bar{T}, \bar{x}(\cdot), \bar{u}(\cdot))$ is also a strong local relaxed minimizer. This is a Type A relation.
- (3) Part (b) can be regarded as a pathway to providing sufficient conditions, of a verifiable nature, for the absence of an infimum gap. The idea is to give verifiable sufficient conditions for normality, which translate, via part (b), into verifiable sufficient conditions for absence of an infimum gap.
- (4) It is not known whether, in part (b), the ‘averaged’ costate inclusion (ii)', in part (b), can be replaced by the stronger pointwise condition (ii).

Theorem 2 Let $(\bar{T}, \bar{x}(\cdot), \{(\bar{\mu}^k(\cdot), \bar{u}^k(\cdot))\}_{k=0}^n)$ be a relaxed admissible process related to problem (R) such that $\bar{T} > 0$. Assume that hypotheses (H1) and (H2) are satisfied.

(a) Assume that $g(\cdot, \cdot, \cdot)$ is Lipschitz continuous on a neighborhood of $(\bar{T}, \bar{x}(0), \bar{x}(\bar{T}))$. Now suppose that $(\bar{x}(\cdot), \{(\bar{\mu}^k(\cdot), \bar{u}^k(\cdot))\}_{k=0}^n)$ is a strong local relaxed minimizer. Then there exist an arc $p(\cdot) \in W^{1,1}([0, \bar{T}]; \mathbb{R}^n)$, $h \in \mathbb{R}$ and $\lambda \geq 0$ such that

- (i) $(p(\cdot), \lambda) \neq (0, 0)$,
- (ii) $-\dot{p}(t) \in \sum_{k=0}^n \bar{\mu}^k(t) \text{co}\{\partial_x H(t, \bar{x}(t), p(t), \bar{u}^k(t))\}$ a.e. $t \in [0, \bar{T}]$,
- (iii) $(h, p(0), -p(\bar{T})) \in \lambda \partial g(\bar{T}, \bar{x}(0), \bar{x}(\bar{T})) + N_C(\bar{T}, \bar{x}(0), \bar{x}(\bar{T}))$,
- (iv) for $k = 0, \dots, n$
 $H(t, \bar{x}(t), p(t), \bar{u}^k(t)) \geq H(t, \bar{x}(t), p(t), u)$, for all $u \in U(t)$, a.e. $t \in [0, \bar{T}]$,
- (v) $h \in \text{ess}_{s \rightarrow \bar{T}} (\sup_{u \in U(s)} H(s, \bar{x}(\bar{T}), u, p(\bar{T}))$.

(b) Assume also that the following condition is satisfied:

(S) $g(\cdot, \cdot, \cdot)$ is continuous on a neighborhood of $(\bar{T}, \bar{x}(0), \bar{x}(\bar{T}))$, and numbers $\varepsilon > 0$ and $\delta > 0$ can be chosen such that

$$g(T, x(0), x(T)) \geq g(\bar{T}, \bar{x}(0), \bar{x}(\bar{T})) + \delta$$

for all admissible processes $(T, x(\cdot), u(\cdot))$ such that $|T - \bar{T}| + \|x^e(\cdot) - \bar{x}^e(\cdot)\|_{L^\infty} \leq \varepsilon$.

Then relations (i)–(v) above are satisfied for some set of multipliers $p(\cdot) \in W^{1,1}([0, \bar{T}]; \mathbb{R}^n)$ and $\lambda \geq 0$, such that $\lambda = 0$.

Comments

- (1) Part (a) of Theorem 2 merely reproduces a known version of the free-time Pontryagin Maximum Principle for the relaxed problem [8], for problems with measurably time dependent data. Interest resides then in Part (b), which asserts properties of feasible relaxed controls, with cost strictly less than that of neighboring (non-relaxed) feasible trajectories.
- (2) Directing attention to the case when $(\bar{x}(\cdot), \{(\bar{\mu}^k(\cdot), \bar{u}^k(\cdot))\}_{k=0}^n)$ is a strong local relaxed minimizer, we deduce from the contrapositive statement of Theorem 2 part (b) another sufficient condition for the non-existence of an infimum gap: if $(\bar{x}(\cdot), \{(\bar{\mu}^k(\cdot), \bar{u}^k(\cdot))\}_{k=0}^n)$ is a relaxed strong local minimizer and, given any multipliers $p(\cdot) \in W^{1,1}$ and $\lambda \geq 0$ satisfying conditions (i)–(v), we have $\lambda \neq 0$, then the cost of $(\bar{T}, \bar{x}(\cdot), \{(\bar{\mu}^k(\cdot), \bar{u}^k(\cdot))\}_{k=0}^n)$ is the infimum cost of admissible processes $(T, x(\cdot), u(\cdot))$ satisfying $|T - \bar{T}| + \|x(\cdot) - \bar{x}(\cdot)\|_{L^\infty} \leq \varepsilon$, for some $\varepsilon > 0$. This is a Type B relation.

3 Proofs of Theorems 1 and 2

We first prove Theorem 2, as the proof of Theorem 1 uses that of Theorem 2.

We need only prove part (b), since a proof of part (a) is available in [8]. Let the relaxed process $(\bar{T}, \bar{x}(\cdot), \{(\bar{\mu}^k(\cdot), \bar{u}^k(\cdot))\}_{k=0}^n)$ and $\varepsilon > 0$ be as in the theorem statement. Define the

function $\bar{\xi}(\cdot) : [0, \bar{T}] \rightarrow \mathbb{R}^n$ to be

$$\bar{\xi}(t) = \int_0^t \sum_{j=0}^n \bar{\mu}^j(s) e^j ds,$$

where $e^0 = (0, \dots, 0)$, $e^1 = (1, 0, \dots, 0)$, \dots , $e^n = (0, \dots, 0, 1)$ are the canonical basis vectors in \mathbb{R}^n . We observe that $(\bar{x}(\cdot), \bar{\xi}(\cdot))$ is a solution of the differential inclusion

$$(\dot{x}(t), \dot{\xi}(t)) \in \text{co } F(t, x(t)) \text{ a.e. } t \in [0, \bar{T}],$$

in which

$$F(t, x) := \bigcup_{k=0}^n \left\{ \left(f\left(t, x, \bar{u}^k(t)\right), e^k \right) \right\}.$$

Take $\rho'_i \downarrow 0$. Invoking the Relaxation Theorem (see, e.g. [8, Theorem 2.7.2]), we can find, for each i , an element $(\bar{T}, x_i(\cdot), \xi_i(\cdot))$ such that $\xi_i(0) = 0$,

$$(\dot{x}_i(t), \dot{\xi}_i(t)) \in F(t, x_i(t)) \text{ a.e. } t \in [0, \bar{T}]$$

and

$$\|(x_i(\cdot), \xi_i(\cdot)) - (\bar{x}(\cdot), \bar{\xi}(\cdot))\|_{L^\infty(0, \bar{T})} \leq \rho'_i. \tag{1}$$

Consider now the set \mathcal{S} of elements $(T, (x(\cdot), \xi(\cdot)), (u(\cdot), \omega(\cdot)))$ comprising absolutely continuous functions $x(\cdot)$ and $\xi(\cdot)$ and measurable functions $u(\cdot)$ and $\omega(\cdot)$ with domain $[0, T]$:

$$\mathcal{S} := \{(T, (x(\cdot), \xi(\cdot)), (u(\cdot), \omega(\cdot))) \mid T \in [0, \bar{T} + \bar{\epsilon}], (\dot{x}(t), \dot{\xi}(t)) = (f(t, x(t), u(t)), \omega(t)), (u(t), \omega(t)) \in U(t) \times V \text{ a.e., } \xi(0) = 0, \|x^e(\cdot) - \bar{x}^e(\cdot)\|_{L^\infty(0, \bar{T} + \bar{\epsilon})} \leq \epsilon\},$$

where $V := \bigcup_{k=0}^n \{e^k\}$. Here, as usual, $x^e(\cdot)$ denotes ‘extension by constant extrapolation from the right’, etc.

By reducing $\epsilon > 0$, if necessary, we can arrange that $(\mathcal{S}, d_{\mathcal{S}}(\cdot, \cdot))$ is a complete metric space, with metric

$$d_{\mathcal{S}}((T, (x(\cdot), \xi(\cdot)), (u(\cdot), \omega(\cdot))), (T', (x'(\cdot), \xi'(\cdot)), (u'(\cdot), \omega'(\cdot)))) := |T - T'| + |x(0) - x'(0)| + \text{meas}\{t \in [0, (T \wedge T')]: (u(t), \omega(t)) \neq (u'(t), \omega'(t))\}.$$

Define the function $J(\cdot) : \mathcal{S} \rightarrow R$

$$J(T, (x(\cdot), \xi(\cdot)), (u(\cdot), \omega(\cdot))) := d_C(T, x(0), x(T)).$$

$J(\cdot)$ is continuous on $(\mathcal{S}, d_{\mathcal{S}}(\cdot, \cdot))$. For each i , $(\bar{T}, x_i(\cdot), \xi_i(\cdot))$ can be interpreted as the state trajectory component of a process $(\bar{T}, (x_i(\cdot), \xi_i(\cdot)), (u_i(\cdot), \omega_i(\cdot)))$ for the dynamical system:

$$\begin{cases} (\dot{x}(t), \dot{\xi}(t)) = (f(t, x(t), u(t)), \omega(t)) \text{ a.e. } t \in [0, \bar{T}], \\ (u(t), \omega(t)) \in U(t) \times V \text{ a.e. } t \in [0, \bar{T}], \end{cases}$$

in which, for each index value i and a.e. $t \in [0, \bar{T}]$,

$$(u_i(t), \omega_i(t)) = \left(\bar{u}^{j(i,t)}(t), e^{j(i,t)} \right).$$

Here, $j(i, t)$ is the unique index value j such that $\frac{d}{dt} \xi_i(t) = e^j$. Notice that

$$(u_i(t), \omega_i(t)) \in \bigcup_{j=0}^n \{(\bar{u}^j(t), e^j)\} \text{ a.e.} \tag{2}$$

For some constant K independent of i ,

$$0 \leq d_C(\bar{T}, x_i(0), x_i(\bar{T})) \leq K\rho_i'.$$

Let $\rho_i := K\rho_i'$. Noting that $J(\cdot) \geq 0$, we deduce that, for each i , $(\bar{T}, (x_i(\cdot), \xi_i(\cdot)), (u_i(\cdot), \beta_i(\cdot)))$ is a ρ_i -minimizer for the optimization problem:

$$\begin{cases} \text{Minimize } J(T, (x(\cdot), \xi(\cdot)), (u(\cdot), \omega(\cdot))) \\ \text{over } (T, (x(\cdot), \xi(\cdot)), (u(\cdot), \omega(\cdot))) \in \mathcal{S}. \end{cases}$$

By Ekeland’s Theorem, there exists, for i sufficiently large, a minimizer $(T_i, (y_i(\cdot), \eta_i(\cdot)), (v_i(\cdot), \beta_i(\cdot))) \in \mathcal{S}$ for the perturbed optimization problem:

$$\begin{cases} \text{Minimize } d_C(T, x(0), x(T)) + \rho_i^{1/2} (|x(0) - y_i(0)| + (T_i - T) \vee 0 + \int_0^T m_i(t, (u(t), \omega(t))) dt) \\ \text{over } (T, (x(\cdot), \xi(\cdot)), (u(\cdot), \omega(\cdot))) \in \mathcal{S}. \end{cases}$$

Furthermore,

$$|x_i(0) - y_i(0)| + (T_i - \bar{T}) \vee 0 + \int_0^{\bar{T}} m_i(t, (u(t), \omega(t))) dt \leq \rho_i^{1/2}. \tag{3}$$

Here $m_i(\cdot, \cdot, \cdot) : [0, \bar{T} + \bar{\varepsilon}] \times \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ is the function

$$m_i(t, u, \omega) := \begin{cases} 1 & \text{if } (u, \omega) \neq (v_i(t), \beta_i(t)) \text{ or } t \in (T_i, T], \\ 0 & \text{otherwise.} \end{cases}$$

Under the stated hypotheses, (1) and (3) imply

$$\|(y_i^e(\cdot), \eta_i^e(\cdot)) - (\bar{x}^e(\cdot), \bar{\xi}^e(\cdot))\|_{L^\infty(0, \bar{T} + \bar{\varepsilon})} \rightarrow 0 \quad \text{as } i \rightarrow \infty. \tag{4}$$

Recall, once again, that $y_i^e : [0, \bar{T} + \bar{\varepsilon}] \rightarrow \mathbb{R}^n$ is the extension of $y_i : [0, T_i] \rightarrow \mathbb{R}^n$ by constant extrapolation to the right. Also, along some subsequence (we do not relabel),

$$(\dot{y}_i^e(\cdot), \dot{\eta}_i^e(\cdot)) \rightarrow (\dot{\bar{x}}^e(\cdot), \dot{\bar{\xi}}^e(\cdot)) \quad \text{weakly in } L^1(0, \bar{T} + \bar{\varepsilon}). \tag{5}$$

$(T_i, (y_i(\cdot), \eta_i(\cdot)), (v_i(\cdot), \beta_i(\cdot)))$ can be interpreted as a strong local minimizer for the optimal control problem:

$$(Q_i) \begin{cases} \text{Minimize } d_C(T, x(0), x(T)) \\ \quad + \rho_i^{1/2} (|x(0) - y_i(0)| + (T_i - T) \vee 0 + \int_0^T m_i(t, (u(t), \omega(t))) dt) \\ \text{over } (T, x(\cdot), \xi(\cdot), u(\cdot), w(\cdot)) \text{ satisfying} \\ \dot{x}(t), \dot{\xi}(t) = (f(t, x(t), u(t)), \omega(t)) \text{ a.e. } t \in [0, T], \\ (u(t), \omega(t)) \in U(t) \times V \text{ a.e. } t \in [0, T], \\ \xi(0) = 0. \end{cases}$$

In view (4), (5) and (S),

$$d_C(T, y_i(0), y_i(T)) > 0 \text{ for } i \text{ sufficiently large.} \tag{6}$$

Now apply the free end-time Pontryagin Maximum Principle [8, Chapter 8] to a reformulation of (Q_i) in which the integral cost term has been eliminated by state augmentation. Since right endpoints of state trajectories in (Q_i) are unconstrained, we are justified in setting the cost multiplier to 1. We conclude: for each i there exist an arc $p_i(\cdot) \in W^{1,1}(0, T_i)$, $h_i \in \mathbb{R}$, and $\lambda_i \geq 0$ such that

- (i) $-\dot{p}_i(t) \in \text{co}\{\partial_x H(t, y_i(t), (p_i(t), v_i(t)))\}$ a.e. $t \in [0, T_i]$,
- (ii) $(h_i, p_i(0), -p_i(T_i)) \in \partial d_C(T_i, y_i(0), y_i(T_i)) + \rho_i^{1/2} \mathbb{B}$,
- (iii) $p_i(t) \cdot \dot{y}_i(t) \geq \max_{u \in U(t)} p_i(t) \cdot f(t, y_i(t), u) - \rho_i^{1/2}$ a.e. $t \in [0, T_i]$,

$$(iv) \quad h_i \in \text{ess}_{s \rightarrow T_i} \left(\sup_{u \in U(s)} H(s, y_i(T_i), u, p_i(T_i)) \right) + \rho_i^{1/2} \mathbb{B}.$$

(Notice that the costate arc associated with the state component $\eta_i(\cdot)$ is the zero function, and so does not appear in the above conditions.)

Since ‘ $\xi \in \partial d_C(z)$ and $z \notin C$ implies $|\xi| = 1$ ’, we deduce from (6) that

$$|(h_i, p_i(0), -p_i(T_i))| \in \left[1 - \rho_i^{1/2}, 1 + \rho_i^{1/2} \right]. \tag{7}$$

Now the (extended) functions $p_i^e(\cdot) \in W^{1,1}(0, \bar{T} + \bar{\varepsilon})$ are uniformly bounded and have uniformly integrably bounded derivatives. It follows that, along some subsequence (we do not relabel), $p_i^e(\cdot) \rightarrow p^e(\cdot)$ strongly in $L^\infty(0, \bar{T} + \bar{\varepsilon})$ and $\dot{p}_i^e(\cdot) \rightarrow \dot{p}^e(\cdot)$ weakly in $L^1(0, \bar{T} + \bar{\varepsilon})$ for some $p^e(\cdot) \in W^{1,1}$. We can also arrange that $h_i \rightarrow h$, for some $h \in \mathbb{R}$.

For any selector $u(t) \in U(t)$ we have, from (iii),

$$\int_{[0, T_i]} p_i(t) \dot{y}_i(t) dt \geq \int_{[0, T_i]} p_i(t) \cdot f(t, y_i(t), u(t)) dt - (K\rho_i)^{1/2}.$$

In the limit

$$\int_{[0, \bar{T}]} p(t) \dot{\bar{x}}(t) dt \geq \int_{[0, \bar{T}]} p(t) \cdot f(t, \bar{x}(t), u(t)) dt.$$

Since this relation is valid for arbitrary selectors $u(\cdot)$ we can conclude that

$$p(t) \cdot \dot{\bar{x}}(t) = \max_{u \in U(t)} p(t) \cdot f(t, \bar{x}(t), u) \quad \text{a.e. } t \in [0, \bar{T}]. \tag{8}$$

Condition (ii) yields, in the limit as $i \rightarrow \infty$,

$$(h, p(0), -p(\bar{T})) \in \partial d_C(\bar{T}, \bar{x}(0), \bar{x}(\bar{T})) \subset N_C(\bar{T}, \bar{x}(0), \bar{x}(\bar{T})). \tag{9}$$

From (7), $|(h, p(0), -p(\bar{T}))| = 1$. Since ‘ $p(\cdot) = 0$ ’ implies $h = 0$, we conclude that

$$p(\cdot) \neq 0. \tag{10}$$

From the stability properties of essential values (see [8, Chapter 8]),

$$h \in \text{ess}_{s \rightarrow \bar{T}} \left(\sup \{ H(s, \bar{x}(\bar{T}), u, p(\bar{T})) : u \in U(t) \} \right). \tag{11}$$

Now define

$$\mathcal{A}_i := \{ t \in [0, \bar{T} + \bar{\varepsilon}] : (v_i(t), \beta_i(t)) \neq (u_i(t), \omega_i(t)) \text{ or } t > T_i \}.$$

By (2),

$$(v_i(t), \beta_i(t)) \in \bigcup_{j=0}^n \left\{ (\bar{u}^j(t), e^j) \right\} \quad \text{a.e. } t \in [0, \bar{T} + \bar{\varepsilon}] \setminus \mathcal{A}_i.$$

It follows from condition (i) that, for a.e. $t \in [0, \bar{T} + \bar{\varepsilon}] \setminus \mathcal{A}_i$,

$$(-\dot{p}_i(t), \dot{\eta}_i(t)) \in \bigcup_{j=0}^n \left(\text{co } \partial_x H(t, y_i(t), p_i(t), \bar{u}^j(t)) \times \{ (f(t, y_i(t), \bar{u}^j(t)), e^j) \} \right).$$

But (along some subsequence) $(p_i^e(\cdot), y_i^e(\cdot), \eta_i^e(\cdot)) \rightarrow (p^e(\cdot), \bar{x}^e(\cdot), \bar{\xi}^e(\cdot))$ uniformly, $(\dot{p}_i^e(\cdot), \dot{y}_i^e(\cdot), \dot{\eta}_i^e(\cdot)) \rightarrow (\dot{p}^e(\cdot), \dot{\bar{x}}^e(\cdot), \dot{\bar{\xi}}^e(\cdot))$ weakly in $L^1(0, \bar{T} + \bar{\varepsilon})$. A convergence analysis similar to in ([8, Chapter 9]) yields, in the limit as $i \rightarrow \infty$,

$$\left(-\dot{p}(t), \dot{\bar{\xi}}(t) \right) \in \text{co} \left(\bigcup_{j=0}^n \left(\text{co } \partial_x H(t, \bar{x}(t), p(t), \bar{u}^j(t)) \times \{ (f(t, \bar{x}(t), \bar{u}^j(t)), e^j) \} \right) \right)$$

for a.e. $t \in [0, \bar{T}]$. By the Carathéodory Representation Theorem (see [10]), there exist measurable functions $\mu^j(\cdot)$, $j = 0, 1, \dots, n$ such that $\{\mu^j(t)\} \in \Lambda$ a.e. and

$$(-\dot{p}(t), \dot{\bar{x}}(t), \dot{\bar{\xi}}(t)) \in \sum_{j=0}^n \mu^j(t) \text{co} \left(\partial_x H(t, \bar{x}(t), r(t), \bar{u}^j(t)) \times \{(f(t, \bar{x}(t), \bar{u}^j(t)), e^j)\} \right) \text{ a.e.}$$

But then

$$\dot{\bar{\xi}}(t) = \sum_{j=0}^n \mu^j(t) e_j = \sum_{j=0}^n \bar{\mu}^j(t) e_j \text{ a.e.}$$

Since the vectors e_0, \dots, e_n are in ‘general position’, it follows that $\mu_j(t) = \bar{\mu}_j(t)$ a.e., $j = 0, \dots, n$. But then

$$\left(-\dot{p}(t), \dot{\bar{x}}(t), \dot{\bar{\xi}}(t) \right) \in \sum_{j=0}^n \bar{\mu}^j(t) \text{co} \left(\partial_x H(t, \bar{x}(t), p(t), \bar{u}^j(t)) \times \left\{ \left(f(t, \bar{x}(t), \bar{u}^j(t)), e^j \right) \right\} \right) \text{ a.e.}$$

This relation implies that:

$$-\dot{p}(t) \in \sum_{j=0}^n \bar{\mu}^j(t) \text{co} \partial_x H(t, \bar{x}(t), p(t), \bar{u}^j(t)) \text{ a.e.} \tag{12}$$

(We observe that if, in the preceding analysis, we had not augmented the state trajectories $x(\cdot)$ with the arcs $\xi(\cdot)$, we would have obtained a weaker form of the costate differential inclusion (12), in which the $\bar{\mu}^j(\cdot)$ ’s was replaced by some other, possibly different, weight functions $\mu^j(\cdot)$ ’s. The reason for augmenting the state trajectories $x(\cdot)$ was precisely to avoid this eventuality.)

Reviewing (8), (9), (10), (11) and (12), we see that all the assertions of Theorem 2 have been proved.

3.1 Proof of Theorem 1

Part (b) merely states a known version of the free end-time Pontryagin Maximum Principle for the relaxed problem (R). (See [8, Chapter 8]). We need to attend, then, only to part (b). Take a sequence $\varepsilon_i \downarrow 0$. Under the assumed conditions, there exists a sequence of admissible relaxed trajectories $(T_i, x_i(\cdot), \{\mu_i^k(\cdot), u_i^k(\cdot)\}_{k=0}^n)$ such that

$$g(T_i, x_i(0), x_i(T_i)) < g(\bar{T}, \bar{x}(0), \bar{x}(\bar{T}))$$

and

$$|T_i - \bar{T}| + \|x_i^e(\cdot) - \bar{x}^e(\cdot)\|_{L^\infty(0, \bar{T} + \varepsilon)} \leq \varepsilon_i.$$

The preceding relation implies that $x_i^e(\cdot) \rightarrow \bar{x}^e(\cdot)$ uniformly. Now apply Theorem 2, with reference to the process $(T_i, x_i(\cdot), \{\mu_i^k(\cdot), u_i^k(\cdot)\}_{k=0}^n)$. Notice that, since $\lambda = 0$ in the theorem statement, the nontriviality condition on the co-state multiplier becomes $p(\cdot) \neq 0$, which can be strengthened to $\|p(\cdot)\|_{L^\infty} = 1$, by scaling. We deduce, then, existence of $p_i(\cdot) \in W^{1,1}([0, T_i]; \mathbb{R}^n)$, and $h_i \in \mathbb{R}$ such that:

- (i)' $\|p_i(\cdot)\|_{L^\infty} = 1$,
- (ii)' $-\dot{p}_i(t) \in \sum_{k=0}^n \lambda_i^k \text{co} \partial_x H(t, x_i(t), u_i^k(t), p_i(t)) \text{ a.e. } t \in [0, T_i]$,
- (iii)' $(h_i, p_i(0), -p_i(T_i)) \in N_C(T_i, x_i(0), x_i(T_i))$,
- (iv)' $\forall k = 0, \dots, n,$
 $p_i(t) \cdot f(t, x_i(t), u_i^k(t)) = \max_{u \in U(t)} p_i(t) \cdot f(t, x_i(t), u) \text{ a.e. } t \in [0, T_i]$,
- (v)' $h_i \in \text{ess}_{s \rightarrow T_i} (\sup_{u \in U(t)} H(s, x_i(T_i), u, p_i(T_i)))$.

From (ii)' it follows that

$$-\dot{p}_i(t) \in \bigcup_{u \in U(t)} \text{co } \partial_x H(t, x_i(t), u, p_i(t)) \quad \text{a.e. } t \in [0, T_i].$$

The sequences $\{x_i^e(\cdot)\}$ and $\{p_i^e(\cdot)\}$ are uniformly bounded. Furthermore $\{\dot{x}_i^e(\cdot)\}$ and $\{\dot{p}_i^e(\cdot)\}$ are equi-integrable. From the Compactness of Trajectories Theorem [8, Theorem 2.5.3], it follows that $x_i^e(\cdot) \rightarrow \bar{x}^e(\cdot)$ uniformly and $\dot{x}_i^e(\cdot) \rightarrow \dot{\bar{x}}^e(\cdot)$ weakly in $L^1(0, \bar{T} + \bar{\varepsilon})$, as well as $p_i^e(\cdot) \rightarrow p^e(\cdot)$ uniformly and $\dot{p}_i^e(\cdot) \rightarrow \dot{p}^e(\cdot)$ weakly in $L^1(0, \bar{T} + \bar{\varepsilon})$. A standard convergence analysis (see, e.g. [8, Chapter 9]) permits us to pass to the limit in (i)'–(v)', and thereby show:

- (i) $p(\cdot) \neq 0$,
- (ii) $-\dot{p}(t) \in \text{co } \bigcup_{u \in U(t)} \partial_x H(t, \bar{x}(t), p(t))$ a.e. $t \in [0, \bar{T}]$,
- (iii) $(h, p(0), -p(\bar{T})) \in N_C(\bar{T}, \bar{x}(0), \bar{x}(\bar{T}))$,
- (iv) $p(t) \cdot f(t, \bar{x}(t), \bar{u}(t)) = \max_{u \in U(t)} p(t) \cdot f(t, \bar{x}(t), u)$ a.e. $t \in [0, \bar{T}]$,
- (v) $h \in \text{ess}_{s \rightarrow \bar{T}} \bar{p}(\sup_{u \in U(t)} H(s, \bar{x}(\bar{T}), u, p(\bar{T})))$.

These are the desired necessary conditions, in which the cost multiplier is zero. The proof is complete.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Clarke, F.H., Loewen, P.D.: The value function in optimal control: sensitivity, controllability, and time-optimality. *SIAM J. Control Optim.* **24**, 243–263 (1986)
- Clarke, F.H., Ledyaev, Y.S., Stern, R.J., Wolenski, P.R.: *Nonsmooth Analysis and Control Theory*. Graduate Texts in Mathematics, vol. 178. Springer, New York (1998)
- Gamkrelidze, R.V.: On some extremal problems in the theory of differential equations with applications to the theory of optimal control. *J. SIAM Control* **3**, 106–128 (1965)
- Ioffe, A.D.: Euler-Lagrange and Hamiltonian formalisms in dynamic optimization. *Trans. Am. Math. Soc.* **349**, 2871–2900 (1997)
- Palladino, M., Vinter, R.B.: Minimizers that are not also relaxed minimizers. *SIAM J. Control Optim.* **52**, 2164–2179 (2014)
- Palladino, M., Vinter, R.B.: When are minimizing controls also minimizing relaxed controls? *Discret. Contin. Dyn. Syst. Ser. A* **35**, 4573–4592 (2015)
- Rockafellar, R.T., Wets, J.-B.R.: *Variational Analysis*. Grundlehren Der Mathematischen Wissenschaften, vol. 317. Springer, Berlin (1998)
- Vinter, R.B.: *Optimal Control*. Birkhäuser, Boston (2000)
- Warga, J.: Normal control problems have no minimizing strictly original solutions. *Bull. Am. Math. Soc.* **77**, 625–628 (1971)
- Warga, J.: *Optimal Control of Differential and Functional Equations*. Academic Press, New York (1972)
- Warga, J.: Controllability, extremality, and abnormality in nonsmooth optimal control. *J. Optim. Theory Appl.* **41**, 239–260 (1983)
- Warga, J.: Optimization and controllability without differentiability assumptions. *SIAM J. Control Optim.* **21**, 837–855 (1983)

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