Journal of Convex Analysis
Volume 14 (2007), No. 2, 361–393

Convex Cones and Conjugacy for Inequality Control Constraints

Javier F. Rosenblueth
*IIMAS–UNAM, Apartado Postal 20-726,
México DF 01000, México*
jfrl@servidor.unam.mx

Received: May 24, 2005
Revised manuscript received: November 21, 2005

Three approaches to conjugacy available in the literature, applicable to optimal control problems with inequality and equality constraints in the control, are studied. A clear and unified review of the second order conditions obtained in each case is given, and the three definitions of "generalized conjugate points" are improved in the sense that, either the range of applicability is enlarged, or the conditions defining membership of the different sets of points are simplified.

Keywords: Calculus of variations, optimal control, conjugate points, first and second order necessary conditions, convex cones, normality

1991 Mathematics Subject Classification: 49K15

The underlying principles in optimization are simple.
Hestenes, *Optimization Theory* (1975)

1. Introduction

In this paper we study possible definitions of conjugate points for certain optimal control problems involving equality and inequality constraints in the control. The theory of conjugacy in optimal control is rather controversial, even for problems without constraints, but we intend to show the reader that the underlying principles are simpler than what they (apparently) seem to be.

We shall concentrate on three approaches to conjugacy which are available in the literature and are widely quoted. The main contribution of this paper corresponds to a clear and unified review of the second order conditions obtained in each case, as well as an improvement of the three definitions, introduced in those papers, of "generalized conjugate points" in the sense that, either the range of applicability is enlarged, or the conditions defining membership of the different sets of points are simplified.

Let us begin by stating the optimal control problem we shall be concerned with. Suppose we are given an interval $T := [t_0, t_1]$ in $\mathbf{R}$, a point $\xi_0 \in \mathbf{R}^n$, open sets $O \subset \mathbf{R}^n$ and $V \subset \mathbf{R}^m$, and functions

$$(g, h)\colon O \to \mathbf{R} \times \mathbf{R}^k \ (k \le n), \quad (L, f)\colon T \times O \times V \to \mathbf{R} \times \mathbf{R}^n, \quad \varphi\colon V \to \mathbf{R}^q \ (q \le m).$$

Denote by $X(T, O)$ the space of piecewise C^1 functions mapping T to O, by $\mathcal{U}(T, V)$ the space of piecewise continuous functions mapping T to V, and set $Z := X(T, O) \times \mathcal{U}(T, V)$.

ISSN 0944-6532 / $ 2.50

The sets of constraints will be expressed in terms of $C := \{x \in O \mid h(x) = 0\}$ and

$$U := \{u \in V \mid \varphi_i(u) \le 0 \ (i \in R), \ \varphi_j(u) = 0 \ (j \in Q)\}$$

where $R = \{1, \ldots, r\}$ and $Q = \{r+1, \ldots, q\}$. Let $\mathcal{A} := T \times O \times U$, define

$$\begin{aligned} D &:= \{(x,u) \in Z \mid \dot{x}(t) = f(t, x(t), u(t)) \ (t \in T)\}, \\ Z_e &:= \{(x,u) \in D \mid (t, x(t), u(t)) \in \mathcal{A} \ (t \in T), \ x(t_0) = \xi_0, \ x(t_1) \in C\} \end{aligned}$$

and let $I \colon Z \to \mathbf{R}$ be given by

$$I(x,u) := g(x(t_1)) + \int_{t_0}^{t_1} L(t, x(t), u(t))dt \quad ((x,u) \in Z).$$

The problem we shall be concerned with, which we label (P), is that of minimizing I over Z_e.

A common and concise way of formulating this problem is as follows:

Minimize $I(x,u) = g(x(t_1)) + \int_{t_0}^{t_1} L(t, x(t), u(t))dt$ subject to

(a) $(x,u) \in X(T,O) \times \mathcal{U}(T,V)$;

(b) $\dot{x}(t) = f(t, x(t), u(t)) \ (t \in T)$;

(c) $x(t_0) = \xi_0$, $h(x(t_1)) = 0$;

(d) $\varphi_i(u(t)) \le 0 \ (i \in R)$, $\varphi_j(u(t)) = 0 \ (j \in Q) \ (t \in T)$.

For problem (P) the theory of conjugacy presents two fundamental aspects. One deals with the derivation of second order conditions, while the other treats the question of how to characterize them. Now, even for simpler problems such as the basic fixed-endpoint problem in the calculus of variations, where the former aspect is easily and successfully solved, the latter is rather controversial. In [26] the reader can find several references of different (not equivalent) ways of stating and proving "the" necessary condition of Jacobi where the assumptions on the sets and functions of the problem, and the definition itself of conjugate points, may vary considerably.

Several attempts to generalize the notion of conjugate points to optimal control have been made. In particular, special attention deserve those of Bernhard [4], Breakwell and Ho [5], Caroff and Frankowska [6], Dmitruk [9, 10], Hestenes [12, 13, 15–18], Loewen and Zheng [19], Popescu [23], Stefani and Zezza [28, 29], Zeidan [31, 32], Zeidan and Zezza [33–36]. It might be extremely complicated to compare these, or more, approaches to conjugacy and, in this paper, we shall concentrate on one line of research which has been widely quoted.

It corresponds to the one initiated in 1988 by Zeidan and Zezza [33] where, for problem (P), a second order necessary condition is derived and a set $\mathcal{G}_0$ of "generalized (classical) conjugate points" is proposed. This approach generalizes the classical theory (represented by, for example, Hestenes [14]) in the sense that, when reduced to a calculus of variations context, both theories coincide. Its main result, stating that if a process satisfies the second order condition then the set $\mathcal{G}_0$ is empty in (t_0, t_1) (the underlying open time interval), is strongly based on the so-called "accessory problem" and it requires nonsingularity and two-sided normality assumptions.

In 1994, Loewen and Zheng [19] questioned not only these two assumptions but also the second order necessary condition, which states that the second variation with respect to the process under consideration is nonnegative along certain convex cone. For certain cases (in particular, assuming convexity of the control set U), the convex cone of [33] is enlarged in [19]. Moreover, they introduce a new set $\mathcal{G}_1$ of "generalized conjugate points" and, without making use of the accessory problem and independently of nonsingularity or sided normality assumptions, prove that if a process satisfies a second order condition (in terms of a certain convex cone) then $\mathcal{G}_1$ (depending on that convex cone) is empty in (t_0, t_1). For the basic problem in the calculus of variations, this set contains (but in general does not coincide with) that of classical conjugate points in (t_0, t_1) under nonsingularity assumptions. Thus, their approach generalizes not only the classical theory to optimal control, but also to singular trajectories in calculus of variations.

In 1996, Zeidan [32] considered a problem where both endpoints vary and obtained second order conditions without convexity assumptions on the control set. Essentially, it is shown that the convex cone proposed in [19] holds for the general case. Also in [32] a new set $\mathcal{G}_2$ of "generalized coupled points" is introduced. It contains that of [19] and, for the normal case, it is proved, again through the accessory problem, that a necessary condition for optimality is the nonexistence of such points in (t_0, t_1).

Now, the sets of conjugate points defined in [33], [19] and [32] do solve certain fundamental questions of conjugacy, but they present several undesirable features. For the set $\mathcal{G}_0$, the nonsingularity and the sided normality assumptions reduce its range of applicability. Though this aspect is no longer present in $\mathcal{G}_1$ and $\mathcal{G}_2$, the nonemptiness of these two sets in the open interval (t_0, t_1) has been established merely as a sufficient condition for the existence of negative second variations. In other words, it is not clear if there are problems with negative second variations for which these sets are empty. Also, the main idea of characterizing a condition is, in general, to obtain a simpler way of verifying it. However, even in the calculus of variations, simple examples show that to solve the question of nonemptiness of these sets may be much more difficult than verifying directly if that condition holds.

The purpose of this paper is twofold. First, to give a clear and unified summary of the different second order necessary conditions for problem (P) obtained in each of these three papers. Second, to introduce a new approach to conjugacy for which the undesirable features inherent in the sets previously defined do not occur. We shall study the sets $\mathcal{G}_0$, $\mathcal{G}_1$ and $\mathcal{G}_2$ and improve them in several respects. To begin with, we modify the assumptions on the functions of the problem which are simply "recalled" in [33], [19] and [32]. This modification enlarges the range of problems to which the first and second order conditions of those references can be applied. We then introduce three sets of points $\mathcal{S}_0$, $\mathcal{S}_1$ and $\mathcal{S}_2$ with the following properties. The first contains, under nonsingularity and normality assumptions, the set $\mathcal{G}_0 \cap (t_0, t_1)$, while the other two contain $\mathcal{G}_1$ and $\mathcal{G}_2$ respectively. Moreover, the corresponding second order conditions are shown to be *equivalent* to the emptiness of $\mathcal{S}_1$ in general, and that of $\mathcal{S}_0 \cup \mathcal{S}_2$ for the normal case. Finally, we provide some examples which illustrate the fact that verifying membership of $\mathcal{S}_0$, $\mathcal{S}_1$ or $\mathcal{S}_2$ may be trivial while that of $\mathcal{G}_0$, $\mathcal{G}_1$ or $\mathcal{G}_2$ may be extremely difficult, in some cases perhaps even a hopeless task.

2. First and second order necessary conditions

The purpose of this section is to state, in a clear and unified way, the first and second order necessary conditions for optimality for problem (P) as stated (in the former case) and obtained (in the latter case) in [33], [19] and [32]. These papers differ not only on the assumptions imposed on the functions and sets delimiting the problem, but also on the convex cones on which the second order conditions are based. For comparison reasons, and to avoid any confusion, we shall quote all the assumptions and the necessary conditions exactly as they are written in each reference (except for the unifying notation we shall use, and silently correcting the misprints, which are plentiful, that appear specially in [32]).

We begin by defining several sets and functions which allow us to write the necessary conditions in a succinct way. We have chosen this notation also to avoid any misleading interpretations which easily occur by assigning well-known names or phrases used with completely different meanings in different texts, sometimes even by the same author. For example, words such as *normal*, *regular* or *extremal*, which are constantly used in [33], [19] and [32], have a different meaning in each of these references. Only a few concepts will receive a certain name when we find convenient to do so and no confusion should arise.

Elements of Z will be called *processes*, and a process (x, u) *solves* (P) if $(x, u) \in Z_e$ and $I(x, u) \leq I(y, v)$ for all $(y, v) \in Z_e$. Without loss of generality, the theory to follow will be applied to global solutions of the problem instead of, as in [33] and [32], to weak local minima. The reason is simply that the first and second order necessary conditions we state hold for any open sets $O \subset \mathbf{R}^n$ and $V \subset \mathbf{R}^m$. Thus, shrinking these sets if necessary, the same conditions remain valid for local minima.

Notation and a general result

• For any $S \subset T$ and $r \in \mathbf{N}$, let $X(S, \mathbf{R}^r)$ (resp. $\mathcal{U}(S, \mathbf{R}^r)$) be the space of piecewise C^1 (resp. piecewise continuous) functions mapping S to $\mathbf{R}^r$. For simplicity, set $X := X(T, \mathbf{R}^n)$, $\mathcal{U} := \mathcal{U}(T, \mathbf{R}^m)$. Given $(x, u) \in Z$ we shall use the notation $(\tilde{x}(t))$ to represent $(t, x(t), u(t))$, and the symbol '$*$' will denote transpose.

• For any $x \in O$ let

$$\mathcal{N}(x) := \{p \in \mathbf{R}^n \mid p = h'(x)^* \gamma \text{ for some } \gamma \in \mathbf{R}^k\}.$$

It should be noted (see [19] for details) that, assuming the matrix $h'(x)$ has full row rank, the set $\mathcal{N}(x)$ corresponds to the normal cone associated to the (adjacent) tangent cone to C at x.

• For all (t, x, u, p, μ) in $T \times O \times V \times \mathbf{R}^n \times \mathbf{R}^q$ let

$$H(t, x, u, p, \mu) = \langle p, f(t, x, u) \rangle - L(t, x, u) - \langle \mu, \varphi(u) \rangle$$

and, for all $(x, u) \in Z$, let $M(x, u)$ be the set of all $(p, \mu) \in X \times \mathcal{U}(T, \mathbf{R}^q)$ satisfying

(i) $\dot{p}(t) + H_x^*(\tilde{x}(t), p(t), \mu(t)) = 0$ $(t \in T)$;
(ii) $H_u(\tilde{x}(t), p(t), \mu(t)) = 0$ $(t \in T)$;
(iii) $\mu_i(t) \geq 0$ $(i \in R)$ with $\mu_i(t) = 0$ whenever $\varphi_i(u(t)) < 0$;
(iv) $-[p(t_1) + g'(x(t_1))^*] \in \mathcal{N}(x(t_1))$.

The letter "M" stands for "multipliers" and, as we shall see presently, this set includes first order conditions which are satisfied, under certain assumptions, by any solution to the problem. In particular, condition (iv), known as the *transversality condition*, is stated in a succinct way above in terms of the normal cone to C at $x(t_1)$. The way this condition is expressed in [19] is as follows. For any $(x,\gamma) \in O \times \mathbf{R}^k$, let $\kappa_\gamma(x) := g(x) + \gamma^* h(x)$ and note that, if

$$\mathcal{K}(x,p) := \{\gamma \in \mathbf{R}^k \mid p^* = -\kappa'_\gamma(x)\} \quad ((x,p) \in O \times \mathbf{R}^n),$$

then (iv) is equivalent to $\mathcal{K}(x(t_1), p(t_1)) \neq \emptyset$.

- For any $\mathcal{V} \subset \mathcal{U}$ and $(x,u) \in Z$ let $A(t) := f_x(\tilde{x}(t))$, $B(t) := f_u(\tilde{x}(t))$ $(t \in T)$,

$$D(\mathcal{V}; x, u) := \{(y,v) \in X \times \mathcal{V} \mid \dot{y}(t) = A(t)y(t) + B(t)v(t) \ (t \in T)\},$$

and define the set of *$\mathcal{V}$-admissible variations along* (x,u) by

$$Y(\mathcal{V}; x, u) := \{(y,v) \in D(\mathcal{V}; x, u) \mid y(t_0) = 0, \ h'(x(t_1))y(t_1) = 0\}.$$

- Let $\mathcal{X} := Z \times X \times \mathcal{U}(T, \mathbf{R}^q)$ and consider the sets

$$\mathcal{E} := \{(x,u,p,\mu) \in \mathcal{X} \mid (x,u) \in D \text{ and } (p,\mu) \in M(x,u)\},$$
$$\mathcal{H}(\mathcal{V}) := \{(x,u,p,\mu,\gamma) \in \mathcal{X} \times \mathbf{R}^k \mid J((x,u);(y,v)) \geq 0 \text{ for all } (y,v) \in Y(\mathcal{V}; x,u)\}$$

where

$$J((x,u);(y,v)) = \langle y(t_1), \Lambda_\gamma y(t_1)\rangle + \int_{t_0}^{t_1} 2\Omega(t, y(t), v(t))dt \quad ((y,v) \in X \times \mathcal{U}),$$

$\Lambda_\gamma := \kappa''_\gamma(x(t_1))$ and, for all $(t,y,v) \in T \times \mathbf{R}^n \times \mathbf{R}^m$,

$$2\Omega(t,y,v) := -\big[\langle y, H_{xx}(t)y\rangle + 2\langle y, H_{xu}(t)v\rangle + \langle v, H_{uu}(t)v\rangle\big]$$

where $H(t)$ denotes $H(\tilde{x}(t), p(t), \mu(t))$.

The following result gives first and second order necessary conditions for problem (P). They hold under two assumptions which we label (H1) and (H2). As we shall see, depending on the different assumptions imposed, the corresponding conditions given in [33], [19] and [32] can be restated in this way.

Theorem 2.1. *Suppose* (x,u) *solves* (P) *and* (H1) *holds. Then there exists* $(p,\mu,\gamma) \in X \times \mathcal{U}(T, \mathbf{R}^q) \times \mathbf{R}^k$ *with* $(x,u,p,\mu) \in \mathcal{E}$ *and* $\gamma \in \mathcal{K}(x(t_1), p(t_1))$*. If also* (H2) *holds, then* $(x,u,p,\mu,\gamma) \in \mathcal{H}(\mathcal{V})$*.*

For clarity of exposition, let us "unfold" this result in terms of the sets and functions involved in the statement. This theorem states that, if (x,u) is a solution to the problem and hypothesis (H1) holds, then there exist $p\colon T \to \mathbf{R}^n$ piecewise C^1, $\mu\colon T \to \mathbf{R}^q$ piecewise continuous, and $\gamma \in \mathbf{R}^k$ such that

(i) $-\dot{p}(t) = f_x^*(t, x(t), u(t))p(t) - L^*(t, x(t), u(t))$ $(t \in T)$;
(ii) $p^*(t) f_u(t, x(t), u(t)) - L_u(t, x(t), u(t)) - \mu^*(t)\varphi'(u(t)) = 0$ $(t \in T)$;
(iii) $\mu_i(t) \geq 0$ $(i = 1, \ldots, r)$ with $\mu_i(t)\varphi_i(u(t)) = 0$ $(t \in T)$;

(iv) $-p(t_1) = g'(x(t_1))^* + h'(x(t_1))^*\gamma$.

If also hypothesis (H2) holds then

$$J((x,u);(y,v)) = \langle y(t_1), \kappa''_\gamma(x(t_1))y(t_1)\rangle + \int_{t_0}^{t_1} 2\Omega(t,y(t),v(t))dt \geq 0$$

for all $(y,v) \in X \times \mathcal{U}$ satisfying

(v) $v \in \mathcal{V}$;
(vi) $\dot{y}(t) = f_x(t,x(t),u(t))y(t) + f_u(t,x(t),u(t))v(t)$ $(t \in T)$;
(vii) $y(t_0) = 0$, $h'(x(t_1))y(t_1) = 0$.

The first statement in this theorem, assuming (H1), is a consequence of some version of the maximum principle. A few remarks on this fact will be discussed in the following section. The second statement, under assumption (H2), corresponds to a second order condition depending on the set of $\mathcal{V}$-admissible variations. The set $\mathcal{V}$, in each of the three references, is a convex cone (containing, implicitly, the zero function) of piecewise continuous functions. Let us write explicitly the different convex cones we shall be concerned with.

• For any $u \in U$ let $I(u) := \{i \in R \mid \varphi_i(u) = 0\}$ denote the set of *active inequality indices* and set

$$\mathcal{T}_0(u) := \{v \in \mathbf{R}^m \mid \varphi'_i(u)v = 0 \ (i \in I(u) \cup Q)\}.$$

• If $(x,u,p,\mu) \in \mathcal{E}$, let $\Gamma_\mu(t) := \{i \in R \mid \mu_i(t) > 0\}$ and

$$\mathcal{T}_1(u(t)) := \{v \in \mathbf{R}^m \mid \varphi'_i(u(t))v \leq 0 \ (i \in I(u(t)) - \Gamma_\mu(t)), \ \varphi'_j(u(t))v = 0 \ (j \in \Gamma_\mu(t) \cup Q)\}.$$

• For any $u \in U$, let

$$\mathcal{T}_2(u) := \{v \in \mathbf{R}^m \mid \varphi'_i(u)v \leq 0 \ (i \in I(u)), \ \varphi'_j(u)v = 0 \ (j \in Q)\}.$$

• For any $u \in \mathcal{U}$ let $\mathcal{W}_i(u) := \{v \in \mathcal{U} \mid v(t) \in \mathcal{T}_i(u(t)) \ (t \in T)\}$ $(i = 0,1,2)$.

Note that, if $(x,u,p,\mu) \in \mathcal{E}$ with $(x,u) \in Z_e$, then $\mathcal{T}_0(u(t)) \subset \mathcal{T}_1((u(t)) \subset \mathcal{T}_2((u(t))$ $(t \in T)$ and, therefore, $\mathcal{W}_0(u) \subset \mathcal{W}_1(u) \subset \mathcal{W}_2(u)$. Clearly, each $\mathcal{T}_i$ (resp. $\mathcal{W}_i$) is a convex cone in $\mathbf{R}^m$ (resp. $\mathcal{U}$).

Let us turn now to assumptions (H1) and (H2) as stated in [33], [19] and [32]. Let $F := (L,f)$.

Zeidan & Zezza

The problem considered in [33] deals with equality constraints in the control, that is, $r = 0$. However, in the "important" Remark 6.4, it is said that the case involving also inequality constraints is within the scope of the paper. The reason given is that, by considering only active inequality constraints, and assuming the active index function to be piecewise constant on T, all the results of the paper apply. Based on this remark, let us state the assumptions imposed in [33] for problem (P) which yield first and second order necessary conditions. In A1(b), A1(c) and A2(a) below it is assumed implicitly that we are given a process $(x,u) \in Z_e$.

Assumption A1

(a) $F(t,\cdot,\cdot)$ is C^1 for all $t \in T$ and g, h, φ are C^1; $F(\cdot,x,u)$, $F_x(\cdot,x,u)$ and $F_u(\cdot,x,u)$ are piecewise continuous for all $(x,u) \in O \times V$; there exists an integrable function $\alpha\colon T \to \mathbf{R}$ such that, at any point $(t,x,u) \in T \times O \times V$,

$$|F(t,x,u)| + |F_x(t,x,u)| + |F_u(t,x,u)| \leq \alpha(t).$$

(b) $h'(x(t_1))$ is of full rank; $(\varphi_i'(u(t)))$ $(i \in I(u(t)) \cup Q)$ is of full rank for all $t \in T$.

(c) There is no nonnull solution p of the system
   (i) $\dot{p}(t) + A^*(t)p(t) = 0$ $(t \in T)$;
   (ii) $\langle B^*(t)p(t), \zeta\rangle = 0$ $(\zeta \in \mathcal{T}_0(u(t)),\ t \in T)$;
   (iii) $-p(t_1) \in \mathcal{N}(x(t_1))$.

Assumption A2

(a) $I(u(\cdot))$ is piecewise constant on T.

(b) $F(t,\cdot,\cdot)$ is C^2 for all $t \in T$ and g, h, φ are C^2; F and its partial derivatives of second order in (x,u) are piecewise continuous in t for all $(x,u) \in O \times V$; there exists an integrable function $\beta\colon T \to \mathbf{R}$ such that, at any point $(t,x,u) \in T \times O \times V$,

$$|F_{xx}(t,x,u)| + |F_{xu}(t,x,u)| + |F_{uu}(t,x,u)| \leq \beta(t).$$

Let us briefly mention the names assigned in [33] to some of these assumptions. A process $(x,u) \in Z_e$ is called *regular* if the second part of A1(b) holds, *strongly normal on* T if A1(c) and the first part of A1(b) hold, and a *regular extremal* if $M(x,u) \neq \emptyset$ for some cost multiplier $\lambda_0 \geq 0$ in the definition of $I(x,u)$ with $\lambda_0 + |\gamma| \neq 0$ where $-p^*(t_1) = \lambda_0 g'(x(t_1)) + \gamma^* h'(x(t_1))$. Theorem 5.1 of this paper corresponds to the following result.

Theorem 2.2. *Theorem* 2.1 *holds replacing* (H) *with* (A) *and* $\mathcal{V} = \mathcal{W}_0(u)$.

Loewen & Zheng

The optimal control problem studied in [19] is precisely (P). The main point of the paper, according to the authors in Remark 5.8, is to give a rigorous treatment of second order necessary conditions for control sets involving *inequality constraints* along with equalities. In the same remark one reads: "Although Zeidan and Zezza claim that the inequality constrained case can be studied by their methods, their justification is neither clear nor convincing. The problem arises because control sets defined by a mixed constraint system may have corners, whereas the methods of [33] are valid exclusively for smooth control sets."

An important contribution of this paper is to show that, for certain cases, the set $\mathcal{V} = \mathcal{W}_0(u)$ on which the second order condition of Theorem 2.2 is based can be replaced with a larger one. This new convex cone satisfies certain properties summarized in the following definition.

• Given $(x,u,p,\mu) \in \mathcal{E}$ with $(x,u) \in Z_e$, $\mathcal{V}$ is an "admissible direction set" if $\mathcal{V} \subset \mathcal{U}$ is a convex cone such that, for any $v \in \mathcal{V}$, there exist constants $\rho, \tau > 0$ and a function $w\colon [t_0,t_1] \times [-\tau,\tau] \to \mathbf{R}^m$ satisfying:

(i) $w(\cdot,\theta)$ is piecewise continuous and $w(t,\cdot)$ is C^2;

(ii) $w(t,0) = u(t)$ and $w_\theta(t,0) = v(t)$;
(iii) $\varphi_i(w(t,\theta)) = 0$ for all $i \in \Gamma_\mu(t) \cup Q$;
(iv) $\varphi_i(w(t,\theta)) \le 0$ for all $i \in R - \Gamma_\mu(t)$, if $0 < \theta < \tau$;
(v) $\|w(t,\cdot)\|_\infty$, $\|w_\theta(t,\cdot)\|_\infty$, $\|w_{\theta\theta}(t,\cdot)\|_\infty \le \rho$.

The assumptions imposed in [19] are as follows. In B1(b) and B1(c) below it is assumed implicitly that we are given $(x,u) \in Z_e$ and, in B2(a), $(x,u,p,\mu) \in \mathcal{E}$.

Assumption B1

(a) $V = \mathbf{R}^m$, $t_0 = 0$; F, g, h, φ are C^2; there exists an integrable function β such that, at any point $(t,x,u) \in T \times O \times U$, the first and second order derivatives of $F(t,\cdot,\cdot)$ are bounded by $\beta(t)$; the set $F(t,x,U)$ is closed for every $(t,x) \in T \times O$.
(b) $h'(x(t_1))$ has full row rank; $\{\varphi_i'(u(t)) \mid i \in I(u(t)) \cup Q\}$ is linearly independent for a.a. $t \in T$.
(c) There is no nonnull solution p of the system
  (i) $\dot p(t) + A^*(t)p(t) = 0$ $(t \in T)$;
  (ii) $\langle B^*(t)p(t), \zeta\rangle \le 0$ $(\zeta \in \mathcal{T}_2(u(t)),\ t \in T)$;
  (iii) $-p(t_1) \in \mathcal{N}(x(t_1))$.

Assumption B2

(a) B1(c) holds with $\mathcal{T}_2(u(t))$ replaced with $\mathcal{T}_1(u(t))$;
(b) U is convex.

According to [19], a process $(x,u) \in Z_e$ is called and *extremal* if $M(x,u) \ne \emptyset$, *normal* if B1(c) holds, and *regular* if B2(a) holds. Combining the first order necessary conditions, Theorem 3.3, and the remark following Definition 4.5 of that paper, we obtain the following result.

Theorem 2.3. *Theorem* 2.1 *holds replacing* (H) *with* (B) *and* $\mathcal{V}$ *an admissible direction set.*

The question of how to produce an admissible direction set $\mathcal{V}$ for problem (P) is a fundamental feature in the theory developed in [19]. Assuming that $(x,u,p,\mu) \in \mathcal{E}$ with $(x,u) \in Z_e$, and $I(u(t))$ is piecewise constant $(t \in T)$, it is proved that

$$\mathcal{V}_K := \{kv \mid k \ge 0,\ v \in \mathcal{U},\ v(t) \in V_K(t)\ (t \in T)\}$$

is an admissible direction set, where $K\colon T \to R$ is a certain function with $I(u(t)) \subset K(t)$ and

$$V_K(t) = \{v \in \mathbf{R}^m \mid \varphi_i(u(t)) + \varphi_i'(u(t))v \le 0\ (i \in K(t) - \Gamma_\mu(t)),\\ \varphi_j'(u(t))v = 0\ (j \in \Gamma_\mu(t) \cup Q)\}.$$

The function K introduced in [19] depends on the set $I_\epsilon(u) = \{i \in R \mid -\epsilon \le \varphi_i(u) \le 0\}$ of "ϵ-active indices at u," together with a partition $s_0 = t_0 < s_1 < \cdots < s_N = t_1$ such that, on each interval (s_i, s_{i+1}), the control u is continuous and the index function $I(u(t))$ is constant. Clearly, if $K(t) = I(u(t))$, then $V_K(t) = \mathcal{T}_1(u(t))$ and so $\mathcal{V}_K = \mathcal{W}_1(u)$. In particular, it is shown in [19] that $\mathcal{V}_K$ with $K(t) = I(u(t))$ is an admissible direction set in any of the following three cases: if u is piecewise constant on T, if the control set involves only equality constraints, or in the free-endpoint case.

Zeidan

The question of whether, for the general case, we can choose $\mathcal{V} = \mathcal{W}_1(u)$ in Theorem 2.1, is left open in [19]. It is answered in [32] not only for a problem more general than (P) (both state endpoints are allowed to vary so that, in the statement of the problem, $g(x(t_1))$ is replaced with $g(x(t_0), x(t_1))$ and condition (c) with $h(x(t_0), x(t_1)) = 0$) but removing the convexity assumption on the control set U. When reduced to the problem we are considering, the assumptions are as follows. In C1(b) and C1(c) below we assume $(x, u) \in Z_e$ given and, in C2, $(x, u, p, \mu) \in \mathcal{E}$.

Assumption C1

(a) $F(t,\cdot,\cdot)$ is C^2 for all $t \in T$ and g, h, φ are C^2; $F(\cdot, x, u)$ and its derivatives in (x, u) are piecewise continuous; there exists an integrable function $\beta\colon T \to \mathbf{R}$ such that

$$|F(t,x,u)| + |\nabla_{(x,u)}F(t,x,u)| + |\nabla^2_{(x,u)}F(t,x,u)| \le \beta(t).$$

(b) $h'(x(t_1))$ is of full rank; there exists $k > 0$ such that $|\sum_{i\in\gamma(t)} \lambda_i \varphi_i'(u(t))| \ge k|\lambda|$ for all λ and $t \in T$, where $\gamma(t) = I(u(t)) \cup Q$.

(c) There is no nonnull solution p of the system
   (i) $\dot p(t) + A^*(t)p(t) = 0$ $(t \in T)$;
   (ii) $\langle B^*(t)p(t), w(t)\rangle \le 0$ $(w \in \mathcal{W}_2(u),\ t \in T)$;
   (iii) $-p(t_1) \in \mathcal{N}(x(t_1))$.

Assumption C2

(a) C1(c) holds with $\mathcal{W}_2(u)$ replaced with $\mathcal{W}_1(u)$;

(b) $I(u(\cdot))$ and Γ_μ are piecewise constant on T.

According to [32], a process $(x, u) \in Z_e$ is an *extremal* if it is a regular extremal in the sense of [33], *weakly normal* if it is an extremal in the sense of [19], and *normal* if $M(x, u) = \emptyset$ for a zero cost multiplier in the definition of $I(x, u)$ with $|p| + |\mu| + |\gamma| \ne 0$ and $-p^*(t_1) = \gamma^* h'(x(t_1))$. Lemma 2.2 of [32] states that, if the second part of C1(b) holds, then (x, u) is normal if and only if the system $\dot y(t) = A(t)y(t) + B(t)v(t)$ $(t \in T)$ is $h'(x(t_1))$-controllable over $\mathcal{W}_2(u)$ which, by Lemma 2.1 of that paper, is equivalent to assumption C1(c). Theorems 2.1 and 4.1 of [32] yield the following result.

Theorem 2.4. *Theorem* 2.1 *holds replacing* (H) *with* (C) *and* $\mathcal{V} = \mathcal{W}_1(u)$.

3. Auxiliary results

A change on the assumptions

In [33], the "regularity" assumptions A1(a) and A2(b) are simply "recalled" and Pontryagin's minimum principle is stated under assumption A1(a). No reference is mentioned at all except for [20] which does not deal with that principle. In [32], on the other hand, the paper starts making the "customary" assumptions C1(a) and C1(b) and Pontryagin's minimum principle is said to be taken from [7], [8] or [14]. However, the assumptions used in those three references differ considerably from C1. Finally, in [19], where Pontryagin's maximum principle is quoted under normality assumptions, the "standing hypotheses" are similar (but stronger) to those of [33] and [32].

Apparently (from the list of references in [33] and [32]) the assumptions and the principle are taken from [11] where first and second order conditions are obtained (for certain Mayer optimal control problem) by using some techniques of [21]. However, in [21], first order necessary conditions are derived (in one of several cases) replacing the last part of assumption A1 with

For each compact subset O_c of O and each compact subset V_c of V, there exists an integrable function $\alpha\colon T \to \mathbf{R}$ such that, for all (t, x, u) in $T \times O_c \times V_c$,

$$|F(t, x, u)| + |F_x(t, x, u)| + |F_u(t, x, u)| \leq \alpha(t).$$

It is worth mentioning that (see [21, p. 283]), for the problem we are dealing with, if F is continuous and continuously differentiable with respect to (x, u), then the function α above may even be taken to be constant. This is certainly not the case with respect to A1(a), B1(a) or C1(a). More importantly, when dealing with accessory problems, these assumptions with respect to the integrand Ω may not hold in general, implying that the first order conditions of [11], [32] or [33] may not be applied.

In the remaining of this paper we replace A1(a) (and in the obvious way also A2(b), B1(a) and C1(a)) as above. With this change in the assumptions, as one can verify from the proofs, Theorems 2.2, 2.3 and 2.4 (based on [11], [19], [21], [32] and [33]) still hold.

The first accessory problem

The approach to conjugacy given in [33] is strongly based on the so-called *accessory problem* (with respect to $\mathcal{H}(\mathcal{W}_0(u))$). Given $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$, this problem, which we label $(\tilde{\mathrm{P}}_0)$, corresponds to that of minimizing $J((x, u); \cdot)/2$ over $Y(\mathcal{W}_0(u); x, u)$, that is,

minimize $K(y, v) := \frac{1}{2}\langle y(t_1), \Lambda_\gamma y(t_1)\rangle + \int_{t_0}^{t_1} \Omega(t, y(t), v(t))dt$ subject to

(a) $(y, v) \in X \times \mathcal{W}_0(u)$;
(b) $\dot{y}(t) = A(t)y(t) + B(t)v(t)$ $(t \in T)$;
(c) $y(t_0) = 0$, $h'(x(t_1))y(t_1) = 0$.

Clearly, necessary conditions for this problem may not follow from Theorem 2.2. To begin with, though no constraints of the form $\varphi_i(v(t)) \leq 0$ and $\varphi_j(v(t)) = 0$ are present, as in problem (P), a process (y, v) for $(\tilde{\mathrm{P}}_0)$ is admissible only if v belongs to $\mathcal{W}_0(u) = \{v \in \mathcal{U} \mid v(t) \in \mathcal{T}_0(u(t))\ (t \in T)\}$. Note that $(\tilde{\mathrm{P}}_0)$ can be reformulated by setting

(a) $(y, v) \in X \times \mathcal{U}$;

leaving (b) and (c) as above, and adding the constraint

(d) $\psi_i(t, v(t)) = 0$ $(i \in I(u(t)) \cup Q,\ t \in T)$, where $\psi_i(t, v) = \varphi_i'(u(t))v$.

For such (equivalent) problems, we could try to apply the results of [11] but, as mentioned above, the assumption A1(a) with respect to Ω may not hold. Under the modified assumptions, let us now define the corresponding sets and functions for $(\tilde{\mathrm{P}}_0)$ which yield the first order necessary conditions obtained from [21].

- Given $(x, u, p, \mu) \in \mathcal{X}$, for all $(t, y, v, q) \in T \times \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^n$ let

$$\tilde{H}(t, y, v, q) := \langle q, A(t)y + B(t)v\rangle - \Omega(t, y, v).$$

• Given $(x,u,p,\mu,\gamma) \in \mathcal{X} \times \mathbf{R}^k$, let $\tilde{M}_0(y,v)$ be the set of all $q \in X$ satisfying

(i) $\dot{q}(t) + \tilde{H}_y^*(\tilde{y}(t), q(t)) = 0 \ (t \in T)$;
(ii) $\langle \tilde{H}_v^*(\tilde{y}(t), q(t)), \zeta \rangle = 0 \ (\zeta \in \mathcal{T}_0(u(t)),\ t \in T)$;
(iii) $-[q(t_1) + \Lambda_\gamma y(t_1)] \in \mathcal{N}(x(t_1))$,

and define the set (depending on (x,u,p,μ,γ))

$$\tilde{\mathcal{E}}_0 := \{(y,v,q) \in X \times \mathcal{U} \times X \mid (y,v) \in D(\mathcal{W}_0(u); x,u) \text{ and } q \in \tilde{M}_0(y,v)\}.$$

Note that $(y,v,q) \in X \times \mathcal{U} \times X$ belongs to $\tilde{\mathcal{E}}_0$ if and only if

(i) $\dot{y}(t) = A(t)y(t) + B(t)v(t),\ \ v(t) \in \mathcal{T}_0(u(t)) \ (t \in T)$;
(ii) $\dot{q}(t) + A^*(t)q(t) = -H_{xx}(t)y(t) - H_{xu}(t)v(t) \ (t \in T)$;
(iii) $\langle B^*(t)q(t) + H_{ux}(t)y(t) + H_{uu}(t)v(t), \zeta \rangle = 0 \ (\zeta \in \mathcal{T}_0(u(t)),\ t \in T)$;
(iv) $q^*(t_1) = -y^*(t_1)\Lambda_\gamma - l^*h'(x(t_1))$ for some $l \in \mathbf{R}^k$.

The following result gives first order conditions for a solution of the accessory problem. It can be easily derived from the necessary conditions given in [21] by transforming the problem into a Mayer one.

Lemma 3.1. *Let $(x,u,p,\mu,\gamma) \in \mathcal{X} \times \mathbf{R}^k$ and suppose* A1 *and* A2 *hold. If (y,v) solves $(\tilde{\mathrm{P}}_0)$ then there exists $q \in X$ such that $(y,v,q) \in \tilde{\mathcal{E}}_0$.*

As we did before, let us write explicitly this set of first order necessary conditions. The result states that, if (y,v) is a solution to the accessory problem $(\tilde{\mathrm{P}}_0)$ then there exists $q\colon T \to \mathbf{R}^n$ piecewise C^1 such that

(i) $\varphi_i'(u(t))v(t) = 0$ for all $i = 1,\dots,r$ with $\varphi_i(u(t)) = 0$ and $i = r+1,\dots,q$ $(t \in T)$;
(ii) $\dot{y}(t) = f_x(t,x(t),u(t))y(t) + f_u(t,x(t),u(t))v(t) \ (t \in T)$;
(iii) $\dot{q}(t) + f_x^*(t,x(t),u(t))q(t) = -H_{xx}(t)y(t) - H_{xu}(t)v(t) \ (t \in T)$;
(iv) $\langle f_u^*(t,x(t),u(t))q(t) + H_{ux}(t)y(t) + H_{uu}(t)v(t), \zeta \rangle = 0 \ (\zeta \in \mathcal{T}_0(u(t)),\ t \in T)$;
(v) $q^*(t_1) = -y^*(t_1)\kappa_\gamma''(x(t_1)) - l^*h'(x(t_1))$ for some $l \in \mathbf{R}^k$.

It should be mentioned that an application of the maximum principle yields the inequality

$$\langle f_u^*(t,x(t),u(t))q(t) + H_{ux}(t)y(t) + H_{uu}(t)v(t), \zeta - v(t) \rangle \le 0 \quad \text{for all } \zeta \in \mathcal{T}_0(u(t)),\ t \in T$$

but the equality in (iv) follows since $\mathcal{T}_0(u(t))$ is a subspace.

The second accessory problem

Let us turn now to the approach to conjugacy given in [32], based not on the accessory problem with respect to $\mathcal{H}(\mathcal{W}_0(u))$ but $\mathcal{H}(\mathcal{W}_1(u))$. Given $(x,u,p,\mu,\gamma) \in \mathcal{X} \times \mathbf{R}^k$, this problem, which we label $(\tilde{\mathrm{P}}_1)$, corresponds to that of minimizing $J((x,u);\cdot)/2$ over $Y(\mathcal{W}_1(u); x,u)$, that is,

Minimize $K(y,v) := \frac{1}{2}\langle y(t_1), \Lambda_\gamma y(t_1) \rangle + \int_{t_0}^{t_1} \Omega(t,y(t),v(t))dt$ subject to

(a) $(y,v) \in X \times \mathcal{W}_1(u)$;
(b) $\dot{y}(t) = A(t)y(t) + B(t)v(t) \ (t \in T)$;
(c) $y(t_0) = 0,\ h'(x(t_1))y(t_1) = 0$.

• Given $(x,u,p,\mu,\gamma) \in \mathcal{X} \times \mathbf{R}^k$, let $\tilde{M}_1(y,v)$ be the set of all $q \in X$ satisfying

(i) $\dot{q}(t) + \tilde{H}_y^*(\tilde{y}(t), q(t)) = 0$ $(t \in T)$;
(ii) $\langle \tilde{H}_v^*(\tilde{y}(t), q(t)), w(t)\rangle \leq 0$ $(w \in \mathcal{W}_1(u),\ t \in T)$;
(iii) $\langle \tilde{H}_v^*(\tilde{y}(t), q(t)), v(t)\rangle = 0$ for all $t \in T$;
(iv) $-[q(t_1) + \Lambda_\gamma y(t_1)] \in \mathcal{N}(x(t_1))$,

and define the set (depending on (x,u,p,μ,γ))

$$\tilde{\mathcal{E}}_1 := \{(y,v,q) \in X \times \mathcal{U} \times X \mid (y,v) \in D(\mathcal{W}_1(u); x,u) \text{ and } q \in \tilde{M}_1(y,v)\}.$$

As before, note that $(y,v,q) \in X \times \mathcal{U} \times X$ belongs to $\tilde{\mathcal{E}}_1$ if and only if

(i) $\dot{y}(t) = A(t)y(t) + B(t)v(t),\ \ v \in \mathcal{W}_1(u)$ $(t \in T)$;
(ii) $\dot{q}(t) + A^*(t)q(t) = -H_{xx}(t)y(t) - H_{xu}(t)v(t)$ $(t \in T)$;
(iii) $\langle B^*(t)q(t) + H_{ux}(t)y(t) + H_{uu}(t)v(t), w(t)\rangle \leq 0$ $(w \in \mathcal{W}_1(u),\ t \in T)$;
(iv) $\langle B^*(t)q(t) + H_{ux}(t)y(t) + H_{uu}(t)v(t), v(t)\rangle = 0$ $(t \in T)$;
(v) $q^*(t_1) = -y^*(t_1)\Lambda_\gamma - l^*h'(x(t_1))$ for some $l \in \mathbf{R}^k$.

From [21], by transforming the problem into a Mayer one and using the special structure of $\mathcal{T}_1(u(t))$, we obtain the following result. The "unfolding" of this lemma is analogous to that of Lemma 3.1.

Lemma 3.2. *Let $(x,u,p,\mu,\gamma) \in \mathcal{X} \times \mathbf{R}^k$ and suppose* C1 *and* C2 *hold. If (y,v) solves* $(\tilde{\mathrm{P}}_1)$ *then there exists $q \in X$ such that $(y,v,q) \in \tilde{\mathcal{E}}_1$.*

Notation

In the following sections we shall make use of several sets and functions satisfying certain properties in subintervals of $T = [t_0, t_1]$. To simplify their constant reference, we shall find convenient to introduce the following notation.

• For any $s \in [t_0, t_1)$ let $T_s := [s, t_1]$, $X_s := X(T_s, \mathbf{R}^n)$, $\mathcal{U}_s := \mathcal{U}(T_s, \mathbf{R}^m)$, and set $Z_s := X_s \times \mathcal{U}_s$.

• Given $(x,u,p,\mu,\gamma) \in \mathcal{X} \times \mathbf{R}^k$, let $\tilde{\mathcal{E}}_i(s)$ denote the set of all $(y,v,q) \in Z_s \times X_s$ satisfying the same properties as $\tilde{\mathcal{E}}_i$ $(i = 0,1)$ but replacing the interval T with T_s. For any $s < t_0$ we set $\tilde{\mathcal{E}}_0(s) := \tilde{\mathcal{E}}_0$.

• Given $S \subset T$ and $\emptyset \neq \mathcal{V} \subset \mathcal{U}$, define $\mathcal{V}(S)$ as the set of all $v \in \mathcal{U}(S, \mathbf{R}^m)$ such that, for some $v_0 \in \mathcal{V}$, $\bar{v} \in \mathcal{V}$ where $\bar{v}(t) = v(t)$ for $t \in S$ and $\bar{v}(t) = v_0(t)$ otherwise. For simplicity we set $\mathcal{V}_s := \mathcal{V}(T_s)$. Given $(x,u) \in Z$, let as before $A(t) := f_x(\tilde{x}(t))$, $B(t) := f_u(\tilde{x}(t))$ $(t \in T)$ and define

$$\begin{aligned} D_s(\mathcal{V}; x,u) &:= \{(y,v) \in X_s \times \mathcal{V}_s \mid \dot{y}(t) = A(t)y(t) + B(t)v(t)\ (t \in T_s)\}, \\ Y_s(\mathcal{V}; x,u) &:= \{(y,v) \in D_s(\mathcal{V}; x,u) \mid y(s) = 0,\ h'(x(t_1))y(t_1) = 0\}. \end{aligned}$$

• Whenever we are given $(x,u,p,\mu) \in Z \times X \times \mathcal{U}(T, \mathbf{R}^q)$ and $(y,v) \in Y_s(\mathcal{V}; x,u)$, we shall consider the functions $\sigma\colon T_s \to \mathbf{R}^n$ and $\rho\colon T_s \to \mathbf{R}^m$ defined by

$$\sigma(t) := -H_{xx}(t)y(t) - H_{xu}(t)v(t), \quad \rho(t) := -H_{ux}(t)y(t) - H_{uu}(t)v(t)$$

where $H(t)$ denotes $H(\tilde{x}(t), p(t), \mu(t))$.

• Given $s \in [t_0, t_1)$ and $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$, define the bilinear form $\mathcal{F}_s \colon Z_s \times Z_s \to \mathbf{R}$ by

$$\mathcal{F}_s((z, w), (y, v)) := \langle z(t_1), \Lambda_\gamma y(t_1) \rangle + \int_s^{t_1} \{\langle z(t), \sigma(t) \rangle + \langle w(t), \rho(t) \rangle\} dt.$$

4. Zeidan & Zezza

This section is devoted to an explanation and improvement in several directions of the approach to conjugacy given in [33]. This approach is, in some sense, a "natural" generalization of that concept in the context of calculus of variations, where a point $s \in [t_0, t_1)$ is conjugate to t_1 if there exists a nontrivial secondary extremal (that is, an extremal for the accessory problem) vanishing at s and t_1.

We assume throughout this section that A1(a), A1(b) and A2 (modified) hold. The assumption A1(c) which, in the terminology of [33], means that (x, u) is (strongly) normal on T, will play a fundamental role in the theory to follow and we shall find convenient to write it explicitly every time it is assumed. Let us remove in the definition the interval T and the adjective "strong" but add explicitly its dependence on $\mathcal{T}_0$ (to distinguish it from the notion given in Section 6) and say that (x, u) is $\mathcal{T}_0$-*normal* if A1(c) holds.

Let us begin with the definition of conjugacy given in [33]. There, the elements of $\mathcal{G}_0(x, u, p, \mu, \gamma)$ below, assuming that $(x, u, p, \mu) \in \mathcal{E}$, $\gamma \in \mathcal{K}(x(t_1), p(t_1))$ and (x, u) is $\mathcal{T}_0$-normal, are called *points conjugate to* t_1.

Definition 4.1. For any $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$ let $\mathcal{G}_0(x, u, p, \mu, \gamma)$ be the set of points $s \in [t_0, t_1)$ for which there exists $(y, v, q) \in \tilde{\mathcal{E}}_0$ with $y(s) = 0$, $h'(x(t_1))y(t_1) = 0$ and $(y, v, q) \not\equiv (0, 0, 0)$.

As it is well-known, the necessary condition of Jacobi for optimality in the classical calculus of variations (in terms of the final point, "there are no conjugate points to t_1 in the open interval (t_0, t_1)") holds if the trajectory under consideration satisfies Legendre's strengthened condition (and therefore, in the terminology of Hestenes [14], it is *nonsingular*). This condition can be generalized by defining

$$\mathcal{L}' := \{(x, u, p, \mu) \in \mathcal{X} \mid G^*(t) H_{uu}(\tilde{x}(t), p(t), \mu(t)) G(t) < 0 \text{ for all } t \in T\}$$

where G is piecewise continuous on T and $G(t)$ is a matrix whose columns form an orthonormal basis for $\mathcal{T}_0(u(t))$ $(t \in T)$. However, as trivial examples show (see [35]), even if (x, u, p, μ) belongs to $\mathcal{L}'$ and (x, u) is $\mathcal{T}_0$-normal, the emptiness of $\mathcal{G}_0 \cap (t_0, t_1)$ may not be necessary for optimality. The approach given in [33] requires also the notion of $\mathcal{T}_0$-normality on subintervals of T.

Definition 4.2. For any $[a, b] \subset T$ we say that $(x, u) \in Z$ is $\mathcal{T}_0$-*normal on* $[a, b]$ if the system

(a) $\dot{p}(t) + A^*(t)p(t) = 0$ $(t \in [a, b])$;
(b) $\langle B^*(t)p(t), \zeta \rangle = 0$ $(\zeta \in \mathcal{T}_0(u(t)),\ t \in [a, b])$;
(c) $-p(t_1) \in \mathcal{N}(x(t_1))$ (in case $b = t_1$)

has no nonnull solution p on $[a, b]$. The set of all $(x, u) \in Z$ which are $\mathcal{T}_0$-normal on $[a, b]$ will be denoted by $N[a, b]$.

The necessary condition for optimality given in [33, Theorem 6.1] in terms of the set $\mathcal{G}_0$ of points conjugate to t_1 corresponds to the following result.

Theorem 4.3. *Suppose (x,u) solves (P) with $(x,u) \in N[a,t_1] \cap N[t_0,b]$ for all $(a,b) \in [t_0,t_1) \times (t_0,t_1)$. If $(x,u,p,\mu) \in \mathcal{E} \cap \mathcal{L}'$ and $\gamma \in \mathcal{K}(x(t_1),p(t_1))$, then $\mathcal{G}_0(x,u,p,\mu,\gamma) \cap (t_0,t_1) = \emptyset$.*

To understand what lies behind this condition, let us first characterize $\tilde{\mathcal{E}}_0$, for the nonsingular case, in terms of a linear system. The following result is obtained by a direct substitution.

Proposition 4.4. *Let $(x,u,p,\mu,\gamma) \in \mathcal{L}' \times \mathbf{R}^k$ and $(y,q) \in X \times X$ with $-[q(t_1)+\Lambda_\gamma y(t_1)] \in \mathcal{N}(x(t_1))$. Then the following are equivalent:*

(a) There exists $v \in \mathcal{U}$ such that $(y,v,q) \in \tilde{\mathcal{E}}_0$.

(b) (y,q) satisfies the linear system

$$\dot{y}(t) = a(t)y(t) + b(t)q(t), \quad \dot{q}(t) = c(t)y(t) - a^*(t)q(t) \quad (t \in T)$$

which we label (J), where

$$\begin{aligned} a(t) &= A(t) - B(t)\Psi(t)H_{ux}(t) \\ b(t) &= -B(t)\Psi(t)B^*(t) \\ c(t) &= H_{xu}(t)\Psi(t)H_{ux}(t) - H_{xx}(t), \end{aligned}$$

$\Psi(t) = G(t)[G^*(t)H_{uu}(t)G(t)]^{-1}G^*(t)$ $(t \in T)$, and $H(t)$ denotes $H(\tilde{x}(t),p(t),\mu(t))$.

Let us now prove an auxiliary result which gives sufficient conditions, in terms of one or two-sided normality conditions, for the vanishing of processes (y,v) on different subintervals of T where, for some $q \in X$, (y,v,q) belongs to $\tilde{\mathcal{E}}_0$.

Lemma 4.5. *Suppose $s \in (t_0,t_1)$ and (x,u,p,μ,γ) and (y,v,q) are such that*

(i) $(x,u,p,\mu) \in \mathcal{L}'$, $(x,u,p,\mu,\gamma) \in \mathcal{H}(\mathcal{W}_0(u))$, and (x,u) is $\mathcal{T}_0$-normal.

(ii) $(y,v,q) \in \tilde{\mathcal{E}}_0$ with $y(s) = 0$ and $h'(x(t_1))y(t_1) = 0$.

Then the following holds:

(a) If $(x,u) \in N[t_0,s]$ then $(y(t),v(t)) = (0,0)$ for all $t \in [s,t_1]$.

(b) If $(x,u) \in N[s,t_1]$ then $(y(t),v(t)) = (0,0)$ for all $t \in [t_0,s]$.

(c) If $(x,u) \in N[t_0,s] \cap N[s,t_1]$, then $(y,v,q) \equiv (0,0,0)$ on T.

Proof. Let $(z(t),w(t)) := (0,0)$ if $t \in [t_0,s]$ and $(z(t),w(t)) := (y(t),v(t))$ if $t \in [s,t_1]$. We have

$$J((x,u);(z,w)) = \langle y(t_1), \Lambda_\gamma y(t_1)\rangle + \int_s^{t_1} 2\Omega(t,y(t),v(t))dt = 0$$

and so (z,w) solves $(\tilde{\mathrm{P}}_0)$. By Lemma 3.1, there exists $r \in X$ such that $(z,w,r) \in \tilde{\mathcal{E}}_0$. In particular, this implies that

$$\dot{r}(t) + A^*(t)r(t) = 0, \quad \langle B^*(t)r(t), \zeta\rangle = 0 \quad (\zeta \in \mathcal{T}_0(u(t)),\ t \in [t_0,s]) \tag{1}$$

$$[\dot{q}(t)-\dot{r}(t)]+A^*(t)[q(t)-r(t)] = 0, \quad \langle B^*(t)[q(t)-r(t)], \zeta\rangle = 0 \quad (\zeta \in \mathcal{T}_0(u(t)),\ t \in [s,t_1]). \tag{2}$$

(a): Suppose $(x,u) \in N[t_0,s]$. By (1), $r(t)=0$ for all $t \in [t_0,s]$. Hence (z,r) satisfies (J) with, in particular, $(z(t_0),r(t_0))=(0,0)$, and so $(z,r)\equiv(0,0)$ on T. Thus $y(t)=0$ for all $t\in[s,t_1]$. Now, let $v_1(t)$ be such that $v(t)=G(t)v_1(t)$ $(t\in T)$, so that

$$v_1(t) = -[G^*(t)H_{uu}(t)G(t)]^{-1}G^*(t)B^*(t)q(t).$$

We have $0=B(t)v(t)$ $(t\in[s,t_1])$ and, therefore,

$$0=\langle v_1(t),G^*(t)B^*(t)q(t)\rangle = \langle v_1(t),[G^*(t)H_{uu}(t)G(t)]v_1(t)\rangle \quad (t\in[s,t_1]).$$

Since $(x,u,p,\mu)\in\mathcal{L}'$ it follows that $v_1(t)=0$ on $[s,t_1]$. Hence also $v(t)=0$ on $[s,t_1]$.

(b): Suppose $(x,u)\in N[s,t_1]$. By (2), $q(t)=r(t)$ for all $t\in[s,t_1]$, and so we have two solutions (y,q) and (z,r) of (J) satisfying, in particular, $(y(t_1),q(t_1))=(z(t_1),r(t_1))$. This implies that $(y,q)\equiv(z,r)$ on T, and so $y(t)=0$ for all $t\in[t_0,s]$. Proceeding as above replacing $[s,t_1]$ with $[t_0,s]$, we conclude that $v(t)=0$ on $[t_0,s]$.

(c): Suppose $(x,u)\in N[t_0,s]\cap N[s,t_1]$. By ($a$) and ($b$), $(y,v)\equiv(0,0)$ on T, and the $\mathcal{T}_0$-normality of (x,u) on T implies that also $q\equiv 0$. □

Note that Theorem 4.3 follows straightforwardly from Lemma 4.5(c). Now, Theorem 2.2 (including a second order necessary condition in terms of $\mathcal{W}_0(u)$), Lemma 3.1 (a first order necessary condition for the accessory problem $(\tilde{\mathrm{P}}_0)$), Definition 4.1 (the set of "generalized conjugate points") and Theorem 4.3 (a necessary condition in terms of the nonexistence of generalized conjugate points in the open time interval) correspond to the main elements in the theory of conjugacy introduced in [33]. This theory is successful in several aspects but, as we shall show next, it can be improved in various directions. In the remaining of this section we shall concentrate on the approach itself to conjugacy (Definition 4.1 and Theorem 4.3) leaving unchanged the other aspects of the theory.

To begin with let us mention that, according to Remark 6.3 of [33], "a one-sided strong normality in the theorem is not enough for the result to hold true. In other words, the strong normality assumption on either side is indispensable and cannot be weakened." The reason, as explained in the introduction of [19], is simply that their definition of conjugate points forces them to impose this condition. This clearly follows from Lemma 4.5. Note that, in the definition given in [33], the corresponding $(y,v,q)\in\tilde{\mathcal{E}}_0$ with $y(s)=0$ and $h'(x(t_1))y(t_1)=0$ *should satisfy* $(y,v,q)\not\equiv(0,0,0)$ on T. But, by 4.5(a), we can impose $y\not\equiv 0$ on $[s,t_1]$, and the nonexistence of such points in the open interval will follow, as in Theorem 4.3, if $(x,u)\in N[t_0,b]$ for all $b\in(t_0,t_1]$. The same occurs, by 4.5(b), if one imposes the condition $y\not\equiv 0$ on $[t_0,s]$ and $(x,u)\in N[a,t_1]$ for all $a\in[t_0,t_1)$. These facts are formalized in the next result.

Definition 4.6. For any $(x,u,p,\mu,\gamma)\in\mathcal{X}\times\mathbf{R}^k$ let $\mathcal{C}_0(x,u,p,\mu,\gamma)$ be the set of points $s\in[t_0,t_1)$ for which there exists $(y,v,q)\in\tilde{\mathcal{E}}_0$ with $y(s)=0$, $h'(x(t_1))y(t_1)=0$ and $y\not\equiv 0$ on $[t_0,s]$. Let $\mathcal{C}_1$ be as $\mathcal{C}_0$ but replacing the interval $[t_0,s]$ with $[s,t_1]$.

Clearly, under the assumptions of Theorem 4.3, $\mathcal{G}_0\subset\mathcal{C}_0\cup\mathcal{C}_1$ (all depending on (x,u,p,μ,γ)). Hence, this theorem is a corollary of the following result.

Theorem 4.7. *Suppose* (x,u) *solves* (P), $(x,u,p,\mu)\in\mathcal{E}\cap\mathcal{L}'$ *and* $\gamma\in\mathcal{K}(x(t_1),p(t_1))$. *Then*

(a) $\quad (x,u) \in N[a,t_1]$ *for all* $a \in [t_0,t_1) \Rightarrow \mathcal{C}_0(x,u,p,\mu,\gamma) \cap (t_0,t_1) = \emptyset$.
(b) $\quad (x,u) \in N[t_0,b]$ *for all* $b \in (t_0,t_1] \Rightarrow \mathcal{C}_1(x,u,p,\mu,\gamma) \cap (t_0,t_1) = \emptyset$.

The sets of points $\mathcal{C}_0$ and $\mathcal{C}_1$ improve the set $\mathcal{G}_0$ of conjugate points provided in [33] because there are problems for which the assumptions of Theorem 4.7(a) or 4.7(b) hold but not both. For such problems, the results of [33] give no information while Theorem 4.7 may be useful to detect nonoptimality. The following simple example illustrates this fact.

Example 4.8. Consider the problem of minimizing

$$I(x,u) = \frac{1}{2}\int_{-\pi}^{\pi}\{u^2(t) - x^2(t)\}dt$$

subject to $\dot{x}(t) = \lambda(t)u(t)$ $(t \in [-\pi,\pi])$ and $x(-\pi) = x(\pi) = 0$, where $\lambda(t) = 0$ if $t \in [-\pi,-1]$ and $\lambda(t) = 1$ if $t \in [-1,\pi]$.

Here, $T = [-\pi,\pi]$, $\xi_0 = 0$, $h(x) = x$, $L(t,x,u) = (u^2 - x^2)/2$ and $f(t,x,u) = \lambda(t)u$.

Let $(x,u) \in Z$. Note first that (x,u) is $\mathcal{T}_0$-normal on $[a,b]$ if the equations $\dot{p}(t) = 0$, $\lambda(t)p(t) = 0$ have no nonnull solution p on $[a,b]$. Thus $(x,u) \in N[a,\pi]$ for all $a \in [-\pi,\pi)$. Also, $H_{uu} \equiv -1$ and so, given any $(x,u,p,\mu) \in \mathcal{E}$ and $\gamma \in \mathcal{K}(x(t_1),p(t_1))$, we have $(x,u,p,\mu) \in \mathcal{L}'$ and the assumptions of Theorem 4.7(a) hold. On the other hand, $(x,u) \notin N[-\pi,b]$ if $b \in (-\pi,-1]$ and, therefore, we cannot apply Theorem 4.7(b) (nor, in consequence, Theorem 4.3, the main result of [33]). Now, as one readily verifies, $\tilde{\mathcal{E}}_0$ is given by those $(y,v,q) \in X \times \mathcal{U} \times X$ satisfying $\dot{y}(t) = \lambda(t)v(t)$, $\dot{q}(t) = -y(t)$, $\lambda(t)q(t) = v(t)$ $(t \in T)$. Setting

$$(y(t),v(t),q(t)) := \begin{cases} (a,0,b-a(t+1)) & \text{if } t \in [-\pi,-1] \\ (\sin t, \cos t, \cos t) & \text{if } t \in [-1,\pi] \end{cases}$$

where $a = \sin(-1)$ and $b = \cos(-1)$, we have $(y,v,q) \in \tilde{\mathcal{E}}_0$, $y(0) = y(\pi) = 0$ and $y \not\equiv 0$ on $[-\pi,0]$. Hence $s = 0$ belongs to $\mathcal{C}_0(x,u,p,\mu,\gamma) \cap (-\pi,\pi)$ and, by Theorem 4.7(a), the problem has no solution.

Before proceeding, let us make a few comments on subsequent papers of Zeidan and Zezza.

• In [34], the optimal control problem considered is slightly different to that of [33]. The problem is as before, except that X corresponds to the space of absolutely continuous functions mapping T to O, $\mathcal{U}$ is the space of measurable functions mapping T to V, and $\mathcal{A} = T \times O \times U$ with $U \subset V$ convex. Second order conditions, according to [34], are derived from [11] or [30] upon writing the problem in Mayer form, and the approach to conjugacy is similar to that of [33]. As with $\mathcal{G}_0$, an example is given "which illustrates the indispensability of the normality assumption on either side." The remark we made just before Definition 4.6 also applies to this case.

• In [36], they consider the problem of [34] with a general boundary condition and U open. Second order conditions are again, according to [36], derived from [11] or [30]. In this paper, they make one of the modifications of $\mathcal{G}_0$ explained above. Their definition of "points coupled with t_1" is based explicitly on the linear system (J), the modification of their previous definition of conjugacy corresponds to $\mathcal{C}_0$, and their main result (applied to our problem) is precisely Theorem 4.7(a).

• The problem considered in [31] is that of [34] with $\mathcal{U} = L^\infty(T, \mathbf{R}^m)$, $O = \mathbf{R}^n$, $V = \mathbf{R}^m$, and

$$\mathcal{A} = \{(t, x, u) \in T \times \mathbf{R}^n \times \mathbf{R}^m \mid \varphi_i(t, x(t), u(t)) \le 0 \ (i \in \{1, \ldots, q\},\ t \in T)\}.$$

Second order conditions are quoted from [21], and conjugacy is defined as in Definition 4.6 in terms of $\mathcal{C}_0$. There, a point belonging to $\mathcal{G}_0$ is called "classically conjugate to t_1." Again, the main theorem (applied to our problem) corresponds to Theorem 4.7(a).

Now, as one easily deduces from the proof of Lemma 4.5, the properties required for an element $(y, v, q) \in \tilde{\mathcal{E}}_0$ have actually only a local character. In particular, there is no need to have these functions defined on the whole interval T, and the properties of normality, nonsingularity, and the nonvanishing of y may be imposed only on certain subintervals of T. With this in mind we introduce the following definition.

Definition 4.9. Given $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$ let $\mathcal{P}(x, u, p, \mu, \gamma)$ be the set of points $s \in [t_0, t_1)$ for which there exists $\epsilon > 0$ such that $G^*(t)H_{uu}(t)G(t) < 0$ for all $t \in [s-\epsilon, s+\epsilon] \cap T$, and either (a) or (b) holds:

(a) $(x, u) \in N([s, s+\epsilon] \cap T)$ and there exists $(y, v, q) \in \tilde{\mathcal{E}}_0(s-\epsilon)$ with $y(s) = 0$, $h'(x(t_1))y(t_1) = 0$ and $y \not\equiv 0$ on $[s-\epsilon, s] \cap T$.

(b) $(x, u) \in N([s-\epsilon, s] \cap T)$ and there exists $(y, v, q) \in \tilde{\mathcal{E}}_0(s)$ with $y(s) = 0$, $h'(x(t_1))y(t_1) = 0$ and $y \not\equiv 0$ on $[s, s+\epsilon] \cap T$.

Note that, under the assumptions of Theorem 4.7(a), $\mathcal{C}_0 \subset \mathcal{P}$ and, under the assumptions of Theorem 4.7(b), $\mathcal{C}_1 \subset \mathcal{P}$. Given $s \in \mathcal{C}_0 \cup \mathcal{C}_1$, the result follows simply by choosing $\epsilon = \max\{s - t_0, t_1 - s\}$.

Theorem 4.10. *If* $(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{W}_0(u))$ *and* (x, u) *is* $\mathcal{T}_0$*-normal, then* $\mathcal{P}(x, u, p, \mu, \gamma) \cap (t_0, t_1) = \emptyset$.

Proof. We proceed basically as in Lemma 4.5. Suppose there exists $s \in \mathcal{P}(x, u, p, \mu, \gamma) \cap (t_0, t_1)$. Let $\epsilon > 0$ be as in 4.9 and suppose that 4.9(a) holds. Since $s \in (t_0, t_1)$ we can assume, without loss of generality, that $\epsilon < \min\{s - t_0, t_1 - s\}$. Let $(y, v, q) \in \tilde{\mathcal{E}}_0(s-\epsilon)$ with $y(s) = 0$, $h'(x(t_1))y(t_1) = 0$ and $y \not\equiv 0$ on $[s-\epsilon, s]$. Let $(z(t), w(t)) := (0, 0)$ if $t \in [t_0, s]$ and $(z(t), w(t)) := (y(t), v(t))$ if $t \in [s, t_1]$. Since (z, w) solves $(\tilde{\mathrm{P}}_0)$, by Lemma 3.1 there exists $r \in X$ such that $(z, w, r) \in \tilde{\mathcal{E}}_0$. In particular, this implies that

$$[\dot{q}(t) - \dot{r}(t)] + A^*(t)[q(t) - r(t)] = 0, \quad \langle B^*(t)[q(t) - r(t)], \zeta\rangle = 0 \quad (\zeta \in \mathcal{T}_0(u(t)),\ t \in [s, t_1]).$$

Since $(x, u) \in N[s, s+\epsilon]$, $q(t) = r(t)$ for all $t \in [s, s+\epsilon]$ and so we have two solutions (y, q) and (z, r) of (J) on $[s-\epsilon, s+\epsilon]$ satisfying $(y(t), q(t)) = (z(t), r(t))$ for any $t \in (s, s+\epsilon)$. This implies that $(y, q) \equiv (z, r)$ on $[s-\epsilon, s+\epsilon]$, and so $y(t) = 0$ for all $t \in [s-\epsilon, s]$, contradicting (a).

Similarly, if 4.9(b) holds and $(y, v, q) \in \tilde{\mathcal{E}}_0(s)$ is such that $y(s) = 0$, $h'(x(t_1))y(t_1) = 0$ and $y \not\equiv 0$ on $[s, s+\epsilon]$, there exists $r \in X$ such that $(z, w, r) \in \tilde{\mathcal{E}}_0$ where (z, w) is defined as above. Thus

$$\dot{r}(t) + A^*(t)r(t) = 0, \quad \langle B^*(t)r(t), \zeta\rangle = 0 \quad (\zeta \in \mathcal{T}_0(u(t)),\ t \in [t_0, s]).$$

Since $(x, u) \in N[s-\epsilon, s]$, $r(t) = 0$ for all $t \in [s-\epsilon, s]$. Hence (z, r) satisfies (J) on $[s-\epsilon, s+\epsilon]$ with, in particular, $(z(t), r(t)) = (0, 0)$ for any $t \in (s-\epsilon, s)$, and so $(z, r) \equiv (0, 0)$ on $[s-\epsilon, s+\epsilon]$. Thus $y(t) = 0$ for all $t \in [s, s+\epsilon]$, contradicting (b). □

In Example 4.8 we provided a simple problem for which the nonemptiness of $\mathcal{C}_0$ implies nonoptimality, but the sets $\mathcal{C}_1$ and $\mathcal{G}_0$ cannot be applied. A slight modification of this example shows some of the difficulties that may occur in trying to use the set $\mathcal{C}_0$. They are, however, trivially solved by using $\mathcal{P}$.

Example 4.11. Consider the problem of Example 4.8 except for λ given now by

$$\lambda(t) = \begin{cases} 0 & \text{if } t \in [-\pi, -1] \\ \psi(t) & \text{if } t \in [-1, 0] \\ 1 & \text{if } t \in [0, \pi] \end{cases}$$

and $\psi\colon [-1, 0] \to \mathbf{R}$ is any positive, continuous function.

Given $(x, u) \in Z$, the same arguments used before imply that the assumptions of Theorem 4.7(a) hold but not those of Theorem 4.7(b) or Theorem 4.3. As before, $\tilde{\mathcal{E}}_0$ is given by those $(y, v, q) \in X \times \mathcal{U} \times X$ satisfying $\dot{y}(t) = \lambda(t)v(t)$, $\dot{q}(t) = -y(t)$, $\lambda(t)q(t) = v(t)$ $(t \in T)$. If we try to apply Theorem 4.7(a), we need to find $s \in (-\pi, \pi)$ and $(y, v, q) \in \tilde{\mathcal{E}}_0$ with $y(s) = y(\pi) = 0$ and $y \not\equiv 0$ on $[-\pi, s]$. Note that, in particular, we require $(y, q) \in X \times X$ to satisfy

$$\dot{q}(t) = -y(t) \ (t \in [-\pi, \pi]) \quad \text{and} \quad \dot{y}(t) = \begin{cases} 0 & \text{if } t \in [-\pi, -1] \\ \psi^2(t)q(t) & \text{if } t \in [-1, 0] \\ q(t) & \text{if } t \in [0, \pi]. \end{cases}$$

Thus we face the problem of finding $(y, q) \in X \times X$ satisfying this differential equation in the entire interval $[-\pi, \pi]$ with ψ an arbitrary positive, continuous function. That this is no longer required follows by an application of Theorem 4.10. Indeed, let $(y(t), v(t), q(t)) := (\sin t, \cos t, \cos t)$ for all $t \in [0, \pi]$. Clearly $(y, v, q) \in \tilde{\mathcal{E}}_0(s)$ for $s = 0$. Let $0 < \epsilon \leq \pi$. Since $(x, u) \in N[-\epsilon, 0]$, $y \not\equiv 0$ on $[0, \epsilon]$ and $H_{uu}(\tilde{x}(t), p(t), 1) = -1$ for all $t \in [-\epsilon, \epsilon]$, the point $s = 0$ satisfies the conditions of 4.9(b) and so $0 \in \mathcal{P}(x, u, p, \mu, \gamma)$. By Theorem 4.10, the problem has no solution.

The set $\mathcal{P}$ certainly may give more information than the sets of points previously defined. This is obviously the case when the nonsingularity assumption of Theorems 4.3 or 4.7 fails and, as the previous example illustrates, this may occur even in the nonsingular case. However, as the following trivial example shows, one may face problems for which the theory related to $\mathcal{P}$ cannot be applied.

Example 4.12. Consider the problem of minimizing

$$I(x, u) = \frac{1}{2} \int_{-\pi}^{\pi} \lambda(t)\{u^2(t) - x^2(t)\} dt$$

subject to $\dot{x}(t) = u(t)$ $(t \in [-\pi, \pi])$ and $x(-\pi) = x(\pi) = 0$, where $\lambda(t) = 0$ $(t \in [-\pi, 0))$, $\lambda(t) = 1$ $(t \in [0, \pi])$.

Clearly we cannot apply Theorems 4.3 or 4.7 since $H_{uu}(t) = -\lambda(t)$ ($t \in [-\pi, \pi]$) and so $\mathcal{L}' = \emptyset$ (that is, any (x, u, p, μ) is singular). Now, if (y, v, q) belongs to $\tilde{\mathcal{E}}_0(s)$ then $q(t) = \lambda(t)\dot{y}(t)$ and $\dot{q}(t) = -\lambda(t)y(t)$ on $[s, \pi]$. Thus, if $s \in \mathcal{P}(x, u, p, \mu, \gamma)$, the condition $y(\pi) = 0$ implies that, for some $\alpha \in \mathbf{R}$, $(y(t), q(t)) = (\alpha \sin t, \alpha \cos t)$ ($t \in [0, \pi]$), and the condition $y(s) = 0$ implies that $s \leq 0$. However, there is no $\epsilon > 0$ for which $H_{uu}(t) < 0$ for all $t \in [s - \epsilon, s + \epsilon]$.

Let us introduce now a set of points which may be used in situations like the above. For this set, its emptiness in the open time interval is necessary for normal minimizers without even one-sided normality assumptions.

Definition 4.13. Given $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$ let $\mathcal{Q}(x, u, p, \mu, \gamma)$ be the set of points $s \in [t_0, t_1)$ for which there exists $(y, v, q) \in \tilde{\mathcal{E}}_0(s)$ with $y(s) = 0$, $h'(x(t_1))y(t_1) = 0$, $\langle H_{uu}(s+)v(s), \zeta\rangle \neq 0$ for some $\zeta \in \mathcal{T}_0(u(s))$, and B is continuous at s.

Theorem 4.14. *If $(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{W}_0(u))$ and (x, u) is $\mathcal{T}_0$-normal, then $\mathcal{Q}(x, u, p, \mu, \gamma) \cap (t_0, t_1) = \emptyset$.*

Proof. Suppose there exists $s \in \mathcal{Q}(x, u, p, \mu, \gamma) \cap (t_0, t_1)$. Let $(y, v, q) \in \tilde{\mathcal{E}}_0(s)$ and $\zeta \in \mathcal{T}_0(u(s))$ be as in 4.13. Define $(z(t), w(t)) := (0, 0)$ if $t \in [t_0, s]$ and $(z(t), w(t)) := (y(t), v(t))$ if $t \in (s, t_1]$. Since (z, w) solves $(\tilde{\mathrm{P}}_0)$, by Lemma 3.1 there exists $r \in X$ such that $(z, w, r) \in \tilde{\mathcal{E}}_0$. In particular, this implies that

$$\langle B^*(t)r(t) + H_{ux}(t)z(t) + H_{uu}(t)w(t), \eta\rangle = 0 \quad (\eta \in \mathcal{T}_0(u(t)),\ t \in T).$$

Therefore $\langle B^*(s-)r(s-), \zeta\rangle = 0$ and $\langle B^*(s+)r(s+) + H_{uu}(s+)v(s), \zeta\rangle = 0$. Since $t \mapsto B^*(t)r(t)$ is continuous at s, we have $\langle H_{uu}(s+)v(s), \zeta\rangle = 0$ and we reach a contradiction. □

Returning to Example 4.12, set $(y(t), v(t), q(t)) := (\sin t, \cos t, \cos t)$ ($t \in [0, \pi]$). Then $(y, v, q) \in \tilde{\mathcal{E}}_0(s)$ for $s = 0$, $H_{uu}(0+)v(0) = -\lambda(0) \neq 0$ and $B \equiv 1$ is continuous at $s = 0$. Therefore $0 \in \mathcal{Q}(x, u, p, \mu, \gamma)$ and, by Theorem 4.14, this problem has no solution.

The sets $\mathcal{P}$ and $\mathcal{Q}$ were first introduced in [27] for the case when only equality constraints are present. For completeness and clarity of exposition we have included the proofs of the necessary conditions they provide.

Let us now give a brief summary of the underlying ideas in their definitions. Suppose that (x, u, p, μ, γ) belongs to $\mathcal{H}(\mathcal{W}_0(u))$, (x, u) is $\mathcal{T}_0$-normal, and there exists a nontrivial "secondary extremal" (y, v, q) (satisfying the conditions of $\tilde{\mathcal{E}}_0$ in $[s, t_1]$) with $y(s) = 0$ and $h'(x(t_1))y(t_1) = 0$. Then, by extending (y, v) to zero outside $[s, t_1]$, one obtains a solution (z, w) of the accessory problem $(\tilde{\mathrm{P}}_0)$, and an application of the maximum principle yields the existence of a piecewise C^1 function r such that $(z, w, r) \in \tilde{\mathcal{E}}_0$. The definition of $\mathcal{P}$ relies on the uniqueness of solutions of the linear system (J) which is satisfied (at least locally) by (z, w, r). The set $\mathcal{Q}$ depends basically on the continuity of r.

We end this section by making use of the bilinear form introduced in Section 3. The set of points we now propose can enlarge and even simplify the applicability of the sets defined so far. It is based on the following simple observation. Let $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$, suppose (x, u) is $\mathcal{T}_0$-normal and, for some $s \in [t_0, t_1)$ and $(y, v) \in Y_s(\mathcal{W}_0(u); x, u)$, we have $\mathcal{F}_s((y, v), (y, v)) = 0$. Then, if there does not exist $q \in X_s$ such that $(y, v, q) \in \tilde{\mathcal{E}}_0(s)$, Lemma 3.1 implies that $(x, u, p, \mu, \gamma) \notin \mathcal{H}(\mathcal{W}_0(u))$. Thus, for certain problems, it suffices to find (y, v) as above to conclude nonoptimality of the process under consideration. This idea can incorporate the sets $\mathcal{P}$ and $\mathcal{Q}$ as follows.

Definition 4.15. For any $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$ let $\mathcal{S}_0(x, u, p, \mu, \gamma)$ be the set of points $s \in [t_0, t_1)$ for which there exists $(y, v) \in Y_s(\mathcal{W}_0(u); x, u)$ such that

(i) $\mathcal{F}_s((y, v), (y, v)) \leq 0$.

(ii) If there exists $q \in X_s$ such that $(y, v, q) \in \tilde{\mathcal{E}}_0(s)$, then $s > t_0$ and $s \in \mathcal{P}(x, u, p, \mu, \gamma) \cup \mathcal{Q}(x, u, p, \mu, \gamma)$.

Theorem 4.16. *If $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$ and (x, u) is $\mathcal{T}_0$-normal, then $(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{W}_0(u)) \Leftrightarrow \mathcal{S}_0(x, u, p, \mu, \gamma) = \emptyset$.*

Proof. "$\Rightarrow$": Suppose there exists $s \in \mathcal{S}_0(x, u, p, \mu, \gamma)$. Let $(y, v) \in Y_s(\mathcal{W}_0(u); x, u)$ be as in 4.15 and define $(\zeta(t), \eta(t)) := (0, 0)$ if $t \in [t_0, s]$, $(\zeta(t), \eta(t)) := (y(t), v(t))$ if $t \in [s, t_1]$. Note that, by 4.15(i),

$$J((x, u); (\zeta, \eta)) = \mathcal{F}_s((y, v), (y, v)) \leq 0.$$

Strict inequality contradicts the assumption $(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{W}_0(u))$ and, therefore, (ζ, η) solves $(\tilde{\mathrm{P}}_0)$. By Lemma 3.1, there exists $q_1 \in X$ such that $(\zeta, \eta, q_1) \in \tilde{\mathcal{E}}_0$. Let q be the restriction of q_1 to $[s, t_1]$, so that $(y, v, q) \in \tilde{\mathcal{E}}_0(s)$. By 4.15(ii), $s > t_0$ and $s \in \mathcal{P}(x, u, p, \mu, \gamma) \cup \mathcal{Q}(x, u, p, \mu, \gamma)$. This contradicts Theorems 4.10 or 4.14.

"$\Leftarrow$": Suppose $(x, u, p, \mu, \gamma) \notin \mathcal{H}(\mathcal{W}_0(u))$. Let $(y, v) \in Y(\mathcal{W}_0(u); x, u)$ be such that $J((x, u); (y, v)) < 0$. Clearly, 4.15(ii) does not apply since, otherwise, $J((x, u); (y, v))$, which coincides with $\mathcal{F}_{t_0}((y, v), (y, v))$, would vanish. Thus $t_0 \in \mathcal{S}_0(x, u, p, \mu, \gamma)$. $\square$

Remark 4.17. The second order condition on which the theory of this section has been based is expressed in terms of the set $\mathcal{W}_0(u)$ defined in Section 2. Theorem 2.2 gives conditions under which, if (x, u) solves (P), then $(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{W}_0(u))$ for certain (p, μ, γ). Combining this result with Theorem 4.16, we obtain a necessary condition for optimality in terms of the set $\mathcal{S}_0(x, u, p, \mu, \gamma)$. Let us just point out that this theory could be modified in terms of the second order conditions given in Theorems 2.3 and 2.4, and we would then deal with either admissible direction sets or $\mathcal{W}_1(u)$.

5. Loewen & Zheng

Throughout [19] the authors compare their results with those of [33] (and occasionally with [34]). In particular, they state that "Some examples will show how our generalized conjugate point gives much more information than that of [33]," and "Our generalized conjugate point is sharper and more general than that of [34]." Basically, there are two main differences between the theories developed in those papers: the second order conditions obtained (as explained in Section 2, the convex cone on which those of [33] are based is, for certain cases, enlarged in [19]) and the approaches to conjugacy. In this section we shall concentrate on this second aspect, assuming that B1(a) (modified), B1(b) and B2(b) hold. No reference to B1(c) or B2(a) (that is, in the terminology of [19], *normality* or *regularity* respectively) will be made since, in the approach to conjugacy given here, there is no use at all of accessory problems, and those assumptions play a role only in the derivation of the second order condition.

From the previous section one concludes, as mentioned in [19], that "when the strengthened Legendre condition and the strong normality conditions of [33] fail to hold, the results of [33] and [34] give little information." Clearly, one of the contributions of the approach we introduced above, in terms of the set $\mathcal{S}_0$, is that those assumptions can be weakened and, in some cases, even removed. This is also the case with respect to the set of conjugate points defined in [19], whose nonemptiness implies the existence of a negative second variation along certain admissible variation. This set is defined as follows (recall the notation given at the end of Section 3).

Definition 5.1. Let $\mathcal{V} \subset \mathcal{U}$. For any $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$ let $\mathcal{G}_1(\mathcal{V}; x, u, p, \mu, \gamma)$ be the set of points $s \in [t_0, t_1)$ for which there exist $(y, v) \in Y_s(\mathcal{V}; x, u)$ and $q \in X_s$ such that, if $\lambda(t) := B^*(t)q(t) - \rho(t)$ $(t \in T_s)$, then

(i) $\dot{q}(t) + A^*(t)q(t) = \sigma(t)$ $(t \in T_s)$.
(ii) $q(s) \neq 0$, $-[q(t_1) + \Lambda_\gamma y(t_1)] \in \mathcal{N}(x(t_1))$.
(iii) $\langle v(t), \lambda(t)\rangle \geq 0$ $(t \in T_s)$.

and either (a) or (b) holds:

(a) $\langle v(t), \lambda(t)\rangle > 0$ on a set of positive measure.
(b) There exists $(z, w) \in Y(\mathcal{V}; x, u)$ such that $\langle z(s), q(s)\rangle > 0$ and $\langle w(t), \lambda(t)\rangle \geq 0$ $(t \in T_s)$.

The main result in [19, Theorem 4.3] relating this set to the second order condition "$(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{V})$" is the following.

Theorem 5.2. *If $\mathcal{V}$ is an admissible direction set and $(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{V})$, then $\mathcal{G}_1(\mathcal{V}; x, u, p, \mu, \gamma) \cap (t_0, t_1) = \emptyset$.*

Let us briefly explain the main difference between $\mathcal{G}_1$ and any of the sets of conjugate points previously defined. Let us consider first the case $\mathcal{V} = \mathcal{W}_0(u)$. In this event, observe that it is required the existence of (y, v, q) in $Z_s \times X_s$ with $y(s) = 0$ and $h'(x(t_1))y(t_1) = 0$, satisfying the same conditions defining membership of $\tilde{\mathcal{E}}_0(s)$ except for $\langle \lambda(t), \zeta\rangle = 0$ $(\zeta \in \mathcal{T}_0(u(t)),\ t \in T_s)$. Instead, it is required that $\langle \lambda(t), v(t)\rangle \geq 0$ $(t \in T_s)$. Of course, if the first relation holds, in which case (y, v, q) is a secondary extremal with respect to $(\tilde{\mathrm{P}}_0)$, then the second is satisfied, but the converse may not occur. In this sense this approach is more general than the previous ones. Now, if $(\zeta(t), \eta(t)) := (0, 0)$ if $t \in [t_0, s]$ and $(\zeta(t), \eta(t)) := (y(t), v(t))$ if $t \in [s, t_1]$, then $J((x, u); (\zeta, \eta)) \leq 0$. If 5.1(a) holds then $(x, u, p, \mu, \gamma) \notin \mathcal{H}(\mathcal{W}_0(u))$. Otherwise, 5.1(b) holds and, setting $(y_\alpha, v_\alpha) := (z + \alpha\zeta, w + \alpha\eta)$ it follows that, for some appropriate α, $J((x, u); (y_\alpha, v_\alpha)) < 0$ and so, once again, $(x, u, p, \mu, \gamma) \notin \mathcal{H}(\mathcal{W}_0(u))$. The generality of this approach is also a consequence of the fact that, if $\mathcal{W}_0(u)$ is a proper subset of $\mathcal{V}$, Definition 5.1 includes admissible variations which are not considered in any of the previous definitions of conjugacy.

Though this approach seems to be successful for certain optimal control problems (note that the emptiness of $\mathcal{G}_1$ in (t_0, t_1) results necessary for optimality even if the process under consideration is singular, and no (strong) normality assumptions are required), let us point out two undesirable features mentioned in the introduction. First, its nonemptiness is established in [19] merely as a sufficient condition for the existence of negative second variations. Second, one can easily find examples for which to solve the question of nonemptiness of this set may be much more difficult than verifying directly if that condition holds (see, for example, [1–3, 24–27]).

These two features can be solved by defining a new set $\mathcal{S}_1$ which generalizes one first introduced in [1], where it is applicable to the fixed-endpoint problem in the calculus of variations, and generalized for optimal control problems with equality constraints in [27]. We refer the reader to [3] for a wide range of problems for which verifying membership of $\mathcal{S}_1$ (in the calculus of variations context) is trivial but that of $\mathcal{G}_1$ may be extremely difficult or perhaps even a hopeless task.

Definition 5.3. Let $\mathcal{V} \subset \mathcal{U}$. For any $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$ let $\mathcal{S}_1(\mathcal{V}; x, u, p, \mu, \gamma)$ be the set of points $s \in [t_0, t_1)$ for which there exists $(y, v) \in Y_s(\mathcal{V}; x, u)$ such that

(i) $\mathcal{F}_s((y, v), (y, v)) \leq 0$.
(ii) There exists $(z, w) \in Y(\mathcal{V}; x, u)$ such that $\mathcal{F}_s((z, w), (y, v)) < 0$.

Theorem 5.4. *Let $\mathcal{V} \subset \mathcal{U}$ and $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$. Then the following holds:*

(*a*) $\mathcal{G}_1(\mathcal{V}; x, u, p, \mu, \gamma) \subset \mathcal{S}_1(\mathcal{V}; x, u, p, \mu, \gamma)$.
(*b*) *If $\mathcal{V}$ is a convex cone, then $(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{V}) \Leftrightarrow \mathcal{S}_1(\mathcal{V}; x, u, p, \mu, \gamma) = \emptyset$.*

Proof. (*a*): Let $s \in \mathcal{G}_1(\mathcal{V}; x, u, p, \mu, \gamma)$ and let $(y, v) \in Y_s(\mathcal{V}; x, u)$ and $q \in X_s$ be as in 5.1. Condition 5.3(i) follows since

$$\begin{aligned} &\mathcal{F}_s((y, v), (y, v)) \\ &= \langle y(t_1), \Lambda_\gamma y(t_1)\rangle + \int_s^{t_1} \{\langle y(t), \dot{q}(t) + A^*(t)q(t)\rangle + \langle v(t), B^*(t)q(t) - \lambda(t)\rangle\}dt \\ &= \langle y(t_1), \Lambda_\gamma y(t_1) + q(t_1)\rangle - \int_s^{t_1} \langle v(t), \lambda(t)\rangle dt \leq 0. \end{aligned}$$

If 5.1(a) holds, then the inequality in 5.3(i) is strict and, therefore, 5.3(ii) holds by setting $(z(t), w(t)) := (0, 0)$ if $t \in [t_0, s]$ and $(z(t), w(t)) := (y(t), v(t))$ if $t \in [s, t_1]$. If 5.1(b) holds, let $(z, w) \in Y(\mathcal{V}; x, u)$ be such that $\langle z(s), q(s)\rangle > 0$ and $\langle w(t), \lambda(t)\rangle \geq 0$ $(t \in T_s)$. Then

$$\mathcal{F}_s((z, w), (y, v)) \leq \langle z(t_1), \Lambda_\gamma y(t_1) + q(t_1)\rangle - \langle z(s), q(s)\rangle - \int_s^{t_1} \langle w(t), \lambda(t)\rangle dt < 0.$$

(*b*) "$\Rightarrow$": Suppose there exists $s \in \mathcal{S}_1(\mathcal{V}; x, u, p, \mu, \gamma)$. Let (y, v) and (z, w) be as in 5.3, and define $(\zeta(t), \eta(t)) := (0, 0)$ if $t \in [t_0, s]$, $(\zeta(t), \eta(t)) := (y(t), v(t))$ if $t \in [s, t_1]$. By 5.3(i),

$$J((x, u); (\zeta, \eta)) = \langle \zeta(t_1), \Lambda_\gamma \zeta(t_1)\rangle + \int_{t_0}^{t_1} 2\Omega(t, \zeta(t), \eta(t))dt = \mathcal{F}_s((y, v), (y, v)) \ \leq \ 0.$$

Set $k := J((x, u); (z, w))$, $\beta := \mathcal{F}_s((z, w), (y, v))$, and $\alpha := -(\beta + k/2\beta)$. Note that $\alpha > 0$ since $k \geq 0$ and $\beta < 0$. Therefore $(y_\alpha, v_\alpha) := (z + \alpha\zeta, w + \alpha\eta)$ belongs to $Y(\mathcal{V}; x, u)$ and

$$\begin{aligned} J((x, u); (y_\alpha, v_\alpha)) &= \langle y_\alpha(t_1), \Lambda_\gamma y_\alpha(t_1)\rangle + \int_{t_0}^{t_1} 2\Omega(t, y_\alpha(t), v_\alpha(t))dt \\ &= k + \alpha^2 J((x, u); (\zeta, \eta)) + 2\alpha\mathcal{F}_s((z, w), (y, v)) \\ &\leq k + 2\alpha\beta = -2\beta^2 < 0. \end{aligned}$$

(*b*) "$\Leftarrow$": Suppose $(x, u, p, \mu, \gamma) \notin \mathcal{H}(\mathcal{V})$. Let $(y, v) \in Y(\mathcal{V}; x, u)$ be such that $J((x, u); (y, v)) < 0$ and let $(z, w) \equiv (y, v)$. Then $t_0 \in \mathcal{S}_1(\mathcal{V}; x, u, p, \mu, \gamma)$. □

Remark 5.5. In view of this result, note first that Theorem 5.2 holds assuming that $\mathcal{V} \subset \mathcal{U}$ is any convex cone and not necessarily an admissible direction set. Also in Theorem 5.2, if $(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{V})$ then $\mathcal{G}_1(\mathcal{V}; x, u, p, \mu, \gamma)$ is empty not only in (t_0, t_1) but in the half-open interval $[t_0, t_1)$. Apart from this, observe that in the proof of Theorem 5.4(*a*) the condition "$q(s) \neq 0$" is never used so that, in Definition 5.1, it can be removed. The reason is that, if 5.1(a) holds, then it is unnecessary and, if 5.1(b) holds, it is redundant. Finally, it is clear that a necessary condition for optimality follows by Theorems 2.3 and 5.4 but, in view of Theorem 2.4, we can replace $\mathcal{V}$ with $\mathcal{W}_1(u)$ in Theorem 5.4 assuming C1 and C2 hold. In this event, in particular, the convexity assumption B2(b) on U is no longer required.

6. Zeidan

The second order necessary condition obtained by Zeidan in [32] states that, assuming C1 and C2, $\mathcal{V}$ can be replaced with $\mathcal{W}_1(u)$ in Theorem 2.1. This contribution enlarges the applicability of the conditions given both in [33] and [19]. But this paper deals also with an approach to conjugacy different from the previous ones. Throughout this section we assume that C1(a) (modified), C1(b) and C2(b) hold. As in Section 4, we shall say that (x, u) is $\mathcal{T}_1$-normal if it satisfies C2(a). When reduced to problem (P), the set of "generalized coupled points" given in [32] is defined as follows.

Definition 6.1. Given $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$ set $\mathcal{V} := \mathcal{W}_1(u)$ and let $\mathcal{G}_2(x, u, p, \mu, \gamma)$ be the set of points $s \in [t_0, t_1)$ for which there exist $(y, v) \in Y_s(\mathcal{V}; x, u)$ and $q \in X_s$ such that, if $\lambda(t) := B^*(t)q(t) - \rho(t)$ $(t \in T_s)$, then

(i) $\dot{q}(t) + A^*(t)q(t) = \sigma(t)$ $(t \in T_s)$.

(ii) $-[q(t_1) + \Lambda_\gamma y(t_1)] \in \mathcal{N}(x(t_1))$.

(iii) $\langle v(t), \lambda(t)\rangle \geq 0$ $(t \in T_s)$

(iv) If the inequality in (iii) is equality for all $t \in T_s$ then, for any $\alpha \in \mathbf{R}^k$ satisfying

$$\langle B^*(t)\Phi^*(t)h'(x(t_1))^*\alpha - \lambda(t), \zeta(t)\rangle \geq 0 \quad (\zeta \in \mathcal{V}_s,\ t \in T_s),$$

there exists $w \in \mathcal{V}([t_0, s])$ such that $\langle z(s), \Phi^*(s)h'(x(t_1))^*\alpha - q(s)\rangle < 0$, where $\Phi\colon T \to \mathbf{R}^{n\times n}$ satisfies $\dot{\Phi}(t) = -\Phi(t)A(t)$ $(t \in T)$, $\Phi(t_1) = I_n$, and z is the solution of $\dot{z}(t) = A(t)z(t) + B(t)w(t)$, $z(t_0) = 0$.

The main result in [32, Theorem 5.1], relating this set to the second order condition, is the following.

Theorem 6.2. *Let $(x, u, p, \mu) \in \mathcal{E}$, $\gamma \in \mathcal{K}(x(t_1), p(t_1))$, and suppose that (x, u) is $\mathcal{T}_1$-normal. Then $(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{W}_1(u)) \Rightarrow \mathcal{G}_2(x, u, p, \mu, \gamma) \cap (t_0, t_1) = \emptyset$.*

As one readily verifies, this new set contains that of Loewen and Zheng with respect to $\mathcal{W}_1(u)$. It is illustrative to give a simple proof of this fact.

Theorem 6.3. $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k \Rightarrow \mathcal{G}_1(\mathcal{W}_1(u); x, u, p, \mu, \gamma) \subset \mathcal{G}_2(x, u, p, \mu, \gamma)$.

Proof. Set $\mathcal{V} := \mathcal{W}_1(u)$ and let $s \in \mathcal{G}_1(\mathcal{V}; x, u, p, \mu, \gamma)$. Let $(y, v) \in Y_s(\mathcal{V}; x, u)$ and $q \in X_s$ satisfy 5.1. If 5.1(a) holds then $s \in \mathcal{G}_2(x, u, p, \mu, \gamma)$. If 5.1(a) does not hold then there exists $(z, w) \in Y(\mathcal{V}; x, u)$ satisfying 5.1(b). Let $M(t) := B^*(t)\Phi^*(t)h'(x(t_1))^*$ $(t \in T)$ and let $\alpha \in \mathbf{R}^k$ satisfy $\langle M(t)\alpha - \lambda(t), \zeta(t)\rangle \geq 0$ $(\zeta \in \mathcal{V}_s,\ t \in T_s)$. Using the facts that $\langle z(t_1), h'(x(t_1))^*\alpha\rangle = 0$, $\langle M(t)\alpha, w(t)\rangle \geq 0$ $(t \in T_s)$, and

$$z(t) = \int_{t_0}^{t} \Phi(t)^{-1}\Phi(\tau)B(\tau)w(\tau)d\tau \quad (t \in T),$$

we have

$$\langle z(s), \Phi^*(s)h'(x(t_1))^*\alpha\rangle = \int_{t_0}^{s} \langle M(t)\alpha, w(t)\rangle dt = -\int_{s}^{t_1} \langle M(t)\alpha, w(t)\rangle dt \leq 0$$

and so $\langle z(s), \Phi^*(s)h'(x(t_1))^*\alpha - q(s)\rangle < 0$. □

The two undesirable features of $\mathcal{G}_1$ explained above remain present with respect to $\mathcal{G}_2$. Its nonemptiness, in the normal case, implies the existence of a negative second variation, but the converse remains open. Also, verifying membership of this set may be much more difficult than checking if $x \notin \mathcal{H}(\mathcal{W}_1(u))$. As before, these two features can be solved by introducing a new set containing $\mathcal{G}_2$.

Definition 6.4. Given $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$ set $\mathcal{V} := \mathcal{W}_1(u)$ and let $\mathcal{S}_2(x, u, p, \mu, \gamma)$ be the set of points $s \in [t_0, t_1)$ for which there exists $(y, v) \in Y_s(\mathcal{V}; x, u)$ such that

(i) $\mathcal{F}_s((y, v), (y, v)) \leq 0$.
(ii) If there exists $q \in X_s$ such that $(y, v, q) \in \tilde{\mathcal{E}}_1(s)$, then there exists $w \in \mathcal{V}([t_0, s])$ with $\langle z(s), q(s)\rangle > 0$ where z is the solution of $\dot{z}(t) = A(t)z(t) + B(t)w(t)$, $z(t_0) = 0$.

Theorem 6.5. *Let $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbf{R}^k$. Then the following holds:*

(a) $\mathcal{S}_1(x, u, p, \mu, \gamma) \cup \mathcal{G}_2(x, u, p, \mu, \gamma) \subset \mathcal{S}_2(x, u, p, \mu, \gamma)$.
(b) *If (x, u) is $\mathcal{T}_1$-normal then $(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{W}_1(u)) \Leftrightarrow \mathcal{S}_2(x, u, p, \mu, \gamma) = \emptyset$.*

Proof. (a): Let $s \in \mathcal{S}_1(x, u, p, \mu, \gamma)$ and let (y, v) and (z, w) be as in 5.3. Suppose there exists $q \in X_s$ satisfying 6.4(ii) (if it does not exist then $s \in \mathcal{S}_2(x, u, p, \mu, \gamma)$). Therefore

$$\begin{aligned}
0 &> \mathcal{F}_s((z, w), (y, v)) \\
&\geq \langle z(t_1), \Lambda_\gamma y(t_1)\rangle + \int_s^{t_1} \{\langle z(t), \dot{q}(t) + A^*(t)q(t)\rangle + \langle w(t), B^*(t)q(t)\rangle\} dt \\
&= \langle z(t_1), \Lambda_\gamma y(t_1) + q(t_1)\rangle - \langle z(s), q(s)\rangle = -\langle z(s), q(s)\rangle
\end{aligned}$$

and this proves the first contention.

Now, let $s \in \mathcal{G}_2(x, u, p, \mu, \gamma)$, set $\mathcal{V} := \mathcal{W}_1(u)$, and let $(y, v) \in Y_s(\mathcal{V}; x, u)$ and $q \in X_s$ be as in 6.1. As in the proof of 5.4(a), clearly $\mathcal{F}_s((y, v), (y, v)) \leq 0$. Now, suppose there exists $q_1 \in X_s$ such that $(y, v, q_1) \in \tilde{\mathcal{E}}_1(s)$. Note first that

$$\begin{aligned}
&\mathcal{F}_s((y, v), (y, v)) \\
&= \langle y(t_1), \Lambda_\gamma y(t_1)\rangle + \int_s^{t_1} \{\langle y(t), \dot{q}_1(t) + A^*(t)q_1(t)\rangle + \langle v(t), B^*(t)q_1(t)\rangle\} dt = 0
\end{aligned}$$

and also

$$\begin{aligned}
&\mathcal{F}_s((y,v),(y,v)) \\
&= \langle y(t_1),\Lambda_\gamma y(t_1)\rangle + \int_s^{t_1}\{\langle y(t),\dot q(t)+A^*(t)q(t)\rangle + \langle v(t),B^*(t)q(t)-\lambda(t)\rangle\}dt \\
&= -\int_s^{t_1}\langle v(t),\lambda(t)\rangle dt \le 0.
\end{aligned}$$

Therefore the inequality in 6.1(iii) is equality for all $t\in T_s$. Now, let $l, l_1\in\mathbf{R}^k$ be such that

$$q^*(t_1) = -y^*(t_1)\Lambda_\gamma - l^*h'(x(t_1)) \quad\text{and}\quad q_1^*(t_1) = -y^*(t_1)\Lambda_\gamma - l_1^*h'(x(t_1)),$$

and define $r(t) := q(t)-q_1(t)$ $(t\in T_s)$ and $\alpha := l_1 - l$. Since $\dot r(t)+A^*(t)r(t)=0$ we have

$$r(t) = \Phi^*(t)r(t_1) = \Phi^*(t)h'(x(t_1))^*\alpha \quad (t\in T_s).$$

Hence, for all $\zeta\in\mathcal{V}_s$ and $t\in T_s$,

$$\langle B^*(t)\Phi^*(t)h'(x(t_1))^*\alpha,\zeta(t)\rangle = \langle\lambda(t),\zeta(t)\rangle - \langle B^*(t)q_1(t)-\rho(t),\zeta(t)\rangle \ge \langle\lambda(t),\zeta(t)\rangle.$$

By 6.1(iv), there exists $w\in\mathcal{V}([t_0,s])$ such that, if z is the solution of $\dot z(t) = A(t)z(t) + B(t)w(t)$ $(t\in[t_0,s])$, $z(t_0)=0$, then

$$\langle z(s),\Phi^*(s)h'(x(t_1))^*\alpha - q(s)\rangle = \langle z(s), r(s)-q(s)\rangle = -\langle z(s),q_1(s)\rangle < 0.$$

This shows that $s\in\mathcal{S}_2(x,u,p,\mu,\gamma)$.

(*b*) "$\Rightarrow$": Suppose there exists $s\in\mathcal{S}_2(x,u,p,\mu,\gamma)$. Let $(y,v)\in Y_s(\mathcal{W}_1(u);x,u)$ be as in 6.4 and define $(\zeta(t),\eta(t)) := (0,0)$ if $t\in[t_0,s]$, $(\zeta(t),\eta(t)) := (y(t),v(t))$ if $t\in[s,t_1]$. Note that, by 6.4(i),

$$J((x,u);(\zeta,\eta)) = \mathcal{F}_s((y,v),(y,v)) \le 0.$$

Strict inequality contradicts the assumption $(x,u,p,\mu,\gamma)\in\mathcal{H}(\mathcal{W}_1(u))$ and, therefore, (ζ,η) solves $(\tilde{\mathrm{P}}_1)$. By Lemma 3.2, there exists $q_1\in X$ such that $(\zeta,\eta,q_1)\in\tilde{\mathcal{E}}_1$. In particular, this implies that

$$\dot q_1(t) + A^*(t)q_1(t) = -H_{xx}(t)\zeta(t) - H_{xu}(t)\eta(t) \quad (t\in T),$$

$$\langle B^*(t)q_1(t) + H_{ux}(t)\zeta(t) + H_{uu}(t)\eta(t), z(t)\rangle \le 0 \quad (z\in\mathcal{W}_1(u),\ t\in T),$$

and $-[q(t_1)+\Lambda_\gamma y(t_1)]\in\mathcal{N}(x(t_1))$. Let q be the restriction of q_1 to $[s,t_1]$. By 6.4(ii), there exists $w\in\mathcal{W}_1(u)([t_0,s])$ with $\langle z(s),q(s)\rangle > 0$ where z is the solution of $\dot z(t) = A(t)z(t)+B(t)w(t)$ $(t\in[t_0,s])$, $z(t_0)=0$. Note that, in particular, $s>t_0$. Now, since $\dot q_1(t)+A^*(t)q_1(t) = 0$ $(t\in[t_0,s])$, we have $q_1(s) = \Phi^*(s)\Phi^{*-1}(t)q_1(t)$ $(t\in[t_0,s])$, and so

$$0 < \langle z(s),q(s)\rangle = \int_{t_0}^s\langle\Phi^{-1}(s)\Phi(t)B(t)w(t),q_1(s)\rangle dt = \int_{t_0}^s\langle w(t),B^*(t)q_1(t)\rangle dt.$$

But $\langle w(t),B^*(t)q_1(t)\rangle\le 0$ for all $t\in[t_0,s]$ and we reach a contradiction.

(*b*) "$\Leftarrow$": Suppose $(x,u,p,\mu,\gamma)\notin\mathcal{H}(\mathcal{V})$. Let $(y,v)\in Y(\mathcal{V};x,u)$ with $J((x,u);(y,v))<0$. Clearly 6.4(ii) does not apply since, otherwise, $J((x,u);(y,v)) = \mathcal{F}_{t_0}((y,v),(y,v)) = 0$. Thus $t_0\in\mathcal{S}_2(x,u,p,\mu,\gamma)$. □

Remark 6.6. A simple consequence of this theorem is that, if (x,u) is $\mathcal{T}_1$-normal and $(x,u,p,\mu,\gamma) \in \mathcal{H}(\mathcal{W}_1(u))$, then $\mathcal{G}_2(x,u,p,\mu,\gamma) = \emptyset$. Thus, a necessary condition for optimality is, under normality assumptions, the emptiness of $\mathcal{G}_2(x,u,p,\mu,\gamma)$ in the half-open interval $[t_0,t_1)$. This remark was already made with respect to $\mathcal{G}_1$. We mention it because in [32], as well as in [19], it is emphasized that this necessary condition holds for points in the *interior* of the time interval. This assertion may be rather misleading since, as it is well-known, Jacobi's necessary condition (in the classical context) only applies for points in the *interior* of the time interval. On the other hand, Jacobi's strengthened condition states that there are no conjugate points in the half-open interval, and this condition leads to sufficiency. With the assertions of [32] and [19] one is thus tempted to think that a similar result might hold for the sets $\mathcal{G}_1$ or $\mathcal{G}_2$.

7. Examples

In [19] one finds four examples which, according to the authors, illustrate how their generalized conjugate point "gives much more information and is sharper than that of [33]." On the other hand, [32] provides two examples for which the results of [19] cannot be applied because they lie outside its scope. One involves a nonconvex control set and, in the other, both endpoints vary. The conclusion in [32] is then that the new results are "much sharper." Let us make a few remarks on these conclusions. We begin by briefly studying two of the examples (7.1 and 7.3) provided in [19].

Example 7.1. Let $a > \pi/2$ and consider the free-endpoint problem of minimizing

$$I(x,u) = \int_0^a \{u^2(t) - x^2(t)\}dt$$

subject to $x(0) = 0$ and $\dot{x}(t) = u(t)$, $u(t) \geq 0$ $(t \in [0,a])$.

In this case $H(t,x,u,p,\mu) = (p+\mu)u - u^2 + x^2$ and $M(x,u)$ consists of all $(p,\mu) \in X \times \mathcal{U}$ satisfying

$$\dot{p}(t) = -2x(t), \quad p(t) + \mu(t) = 2u(t), \quad p(a) = 0$$

with $\mu(t) \geq 0$ and $\mu(t)u(t) = 0$ $(t \in [0,a])$. Thus $(x_0,u_0,p,\mu) := (0,0,0,0) \in \mathcal{E}$. Now, $\mathcal{T}_0(u_0(t)) = \{0\}$ since $I(u_0(t)) = \{1\}$ and so $Y(\mathcal{W}_0(u_0); x_0,u_0) = \{(0,0)\}$. Thus Theorem 2.2 gives no information with respect to (x_0,u_0). On the other hand, since $\Gamma_\mu(t) = \emptyset$, we have $\mathcal{T}_1(u_0(t)) = \mathbf{R}^+$ and so

$$Y(\mathcal{W}_1(u_0); x_0,u_0) = \{(y,v) \in X \times \mathcal{U} \mid v(t) \geq 0,\ \dot{y}(t) = v(t)\ (t \in [0,a]) \text{ and } y(0) = 0\}.$$

Let $(y(t),v(t)) := (\sin t, \cos t)$ $(t \in [0,\pi/2])$, $(y(t),v(t)) := (1,0)$ $(t \in [\pi/2,a])$. Then (y,v) is a $\mathcal{W}_1(u_0)$-admissible variation for which

$$J((x_0,u_0);(y,v)) = 2\int_0^a \{v^2(t) - y^2(t)\}dt = -2(a - \pi/2) < 0$$

and therefore, by Theorem 2.4, (x_0,u_0) is not optimal.

Note that this conclusion has been reached without any reference to conjugate points. It follows simply because (y,v) belongs to $Y(\mathcal{W}_1(u_0);(x_0,u_0))$ where $\mathcal{W}_1(u_0) = \{v \in \mathcal{U} \mid$

$v(t) \geq 0\}$, and the second variation with respect to (x_0, u_0) along (y, v) is negative. On the other hand, $Y(\mathcal{W}_0(u_0); x_0, u_0) = \{(0,0)\}$ since $\mathcal{W}_0(u_0) = \{0\}$, and so the second order condition of [33] gives no information at all. As mentioned before, we could actually adapt the approach to conjugacy given in [19] to that of [33] by replacing $\mathcal{W}_0(u_0)$ with $\mathcal{W}_1(u_0)$ in the definition of $\mathcal{G}_0$ (and making the necessary changes) but, in view of the conclusion reached, it looks rather unnecessary to prove nonemptiness of $\mathcal{G}_0$ or $\mathcal{G}_1$ (in terms of $\mathcal{W}_1$). In other words, this example shows that the two second order conditions may differ, but it does not illustrate why one approach to conjugacy is sharper than the other.

Let us also mention that the proof given in [19] to show that $\mathcal{G}_1(\mathcal{W}_1(u_0)) \cap (0, a) \neq \emptyset$ is much more cumbersome than the simple proof given above to show nonoptimality of (x_0, u_0).

Example 7.2. Consider the fixed-endpoint calculus of variations problem of minimizing

$$I(x,u) = \int_0^a \{u_2^2(t) - x_1^2(t)\}dt$$

subject to $x(0) = x(a) = 0$, $\dot{x}(t) = u(t)$ and $u(t) \in \mathbf{R}^2$ $(t \in [0, a])$.

In this case, $H(t, x, u, p, \mu) = p_1u_1 + p_2u_2 - u_2^2 + x_1^2$. Thus, if we set $(y_1(t), v_1(t)) := (t, 1)$ $(t \in [0, a/2])$, $(y_1(t), v_1(t)) := (a - t, -1)$ $(t \in [a/2, a])$, and $y_2 \equiv v_2 \equiv 0$ then, for any $(x, u) \in Z$,

$$J((x,u);(y,v)) = 2\int_0^a \{v_2^2(t) - y_1^2(t)\}dt = -a^3/6 < 0.$$

Since for any $u \in \mathcal{U}$, $\mathcal{T}_0(u(t)) = \mathcal{T}_1(u(t)) = \mathbf{R}^2$, implying that $\mathcal{W}_0(u) = \mathcal{W}_1(u) = \mathcal{U}$, we conclude that the problem has no solution.

Note that, as in the previous example, the conclusion follows without making use of any notion of conjugacy. As before, it is trivial to show that there exists an admissible variation with a negative second variation. To show, on the other hand, that $\mathcal{G}_1(\mathcal{U}) \neq \emptyset$ (which is certainly the case), requires a much more elaborate study. That this example shows that the approach to conjugacy introduced in [19] is indeed sharper than that of [33] follows since any $(x, u) \in Z$ is singular and, therefore, we cannot apply Theorem 4.3. The example, however, yields little enlightenment to the theory of conjugate points since there is no need to invoke it at all.

The study of the other two examples (7.2 and 7.4) given in [19] is similar to that of Example 7.1. In the first case, $\mathcal{T}_0(u_0(t)) = \{0\} \times \mathbf{R}$ and $\mathcal{T}_1(u_0(t)) = \mathbf{R}^+ \times \mathbf{R}$. In the other, $\mathcal{T}_0(u_0(t)) = \{(0,0)\}$ and $\mathcal{T}_1(u_0(t)) = \mathbf{R}^- \times \mathbf{R}^+$. As in Example 7.1, the process under consideration is nonsingular, so that the theory of [33] in terms of $\mathcal{W}_1$ can be applied but, once more, one can trivially verify (without any notion of conjugacy) the existence of an admissible variation yielding a negative second variation.

Let us turn now to the conclusion given by [32]. First of all, note from Theorem 6.2 that the theory of [32] yields a necessary condition for optimality assuming the process under consideration is $\mathcal{T}_1$-*normal* (compare with Theorems 5.2 and 5.4 where normality is not required). In particular, for the linear fixed-endpoint problem (where the dynamics are of the form $\dot{x}(t) = A(t)x(t) + B(t)u(t)$ and second order conditions are easily obtained without normality assumptions), the sets $\mathcal{G}_1$ and $\mathcal{S}_1$ can be applied in the abnormal case but not $\mathcal{G}_2$. In other words, there are problems which lie beyond the scope of [32] but not of [19] or the set $\mathcal{S}_1$ introduced in this paper. A full discussion of this fact can be found in [25]. On the other hand, it is not clear if, even under normality assumptions, $\mathcal{G}_2$ is sharper than $\mathcal{G}_1$ when both sets are applied to the same problem. This question remains open not only for the problem we are considering but also that of [32] since the definition of $\mathcal{G}_1$ can certainly be extended to that case.

Now, the first example given in [32] deals with a nonconvex control set and so, according to the author, the results of [19] do not apply. However, the example corresponds to a free-endpoint problem and in [19] one reads: "In the free-endpoint case, by the special structure of the problem, we can drop the convexity hypothesis on U and choose an admissible direction set $\mathcal{W}_1(u)$." This contradicts the claim of Zeidan. In any case, as explained before, the second order condition derived in [32] allows us to replace $\mathcal{V}$ with $\mathcal{W}_1(u)$ in Theorem 2.3 for the general case (and not only the three cases mentioned in [19]). About the example it is shown that, for certain $(x_0, u_0, p, \mu) \in \mathcal{E}$ and $\gamma \in \mathcal{K}(x_0(t_1), p(t_1))$, (x_0, u_0) is $\mathcal{T}_1$-normal and any $s \in (0, 1)$ satisfies the conditions given in Definition 6.1 with a strict inequality in 6.1(iii). This shows that any $s \in (0, 1)$ belongs to $\mathcal{G}_2$ (and also to $\mathcal{G}_1(\mathcal{W}_1(u_0))$) and also, straightforwardly, that the second variation with respect to (x_0, u_0) is negative along certain $\mathcal{W}_1(u_0)$-admissible variation. Once more, we come to the question of how useful are these notions of conjugate points since, at least in all these examples, one can trivially contradict the second order condition "$(x, u, p, \mu, \gamma) \in \mathcal{H}(\mathcal{V})$," while it becomes much more complicated to prove nonemptiness of any of the sets introduced in those references.

One can find other examples (see [3]) for which also the failure of the second order condition is trivially verified but it becomes not only more complicated but even uncertain how to check the existence of points in $\mathcal{G}_1$ and $\mathcal{G}_2$. In those examples, an admissible variation (y, v) yielding a negative second variation may not be used in the definition of those sets. This, of course, cannot occur with the sets $\mathcal{S}_i$ $(i = 0, 1, 2)$ since the existence of such (y, v) automatically implies that the point t_0 belongs to them.

Let us now provide two simple calculus of variations problems which illustrate both the usefulness of the sets introduced in this paper and the difficulties that arise in trying to apply those of [33], [19] and [32]. For these examples, one can find in a straightforward way an admissible variation (y, v) for which $J((x, u); (y, v))$ vanishes and (y, v) does not satisfy Jacobi's equation, implying that the point t_0 belongs to $\mathcal{S}_0$ and $\mathcal{S}_2$. It can also be verified that it belongs to $\mathcal{S}_1$. We have chosen this property for simplicity of exposition, but an application of the other properties enjoyed by these sets (in particular when inequality control constraints are present, or in the event that (y, v) satisfies Jacobi's equation), which simplify those of $\mathcal{G}_0$ and $\mathcal{G}_1$, can also be illustrated (see, for example, [3]).

Example 7.3. Let $a \neq 0$, $b \in \mathbf{R}$, set $t_0 := -b/a$, $t_1 := \pi - b/a$, $T = [t_0, t_1]$, and consider the problem of minimizing

$$I(x, u) = \frac{1}{2}x^2(t_1) + \frac{1}{2}\int_{t_0}^{t_1} (at + b)\{u^2(t) - x^2(t)\}dt$$

subject to $\dot{x}(t) = u(t)$ $(t \in T)$, $x(0) = 0$, $x^2(t_1) + x(t_1) = 0$.

Suppose $(x, u, p, \mu) \in \mathcal{E}$ with $(x, u) \in Z_e$ and $\gamma \in \mathcal{K}(x(t_1), p(t_1))$. Clearly $\mathcal{W}_0(u) =$

$\mathcal{W}_1(u) = \mathcal{U}$,

$$J((x,u);(y,v)) = \int_{t_0}^{t_1} (at+b)\{v^2(t) - y^2(t)\}dt \quad ((y,v) \in Z),$$

and the set of $\mathcal{U}$-admissible variations is given by

$$Y(\mathcal{U};x,u) = \{(y,v) \in X \times \mathcal{U} \mid \dot{y}(t) = v(t) \ (t \in T), \ y(t_0) = y(t_1) = 0\}.$$

At first sight it is not clear if there exists $(y,v) \in Y(\mathcal{U};x,u)$ yielding a negative second variation with respect to (x,u). However, we can show its existence by a simple application of Theorems 4.16 or 6.5. To prove it, let $y(t) := \sin(t+b/a)$, $v(t) := \cos(t+b/a)$ $(t \in T)$. Then, as one readily verifies,

$$\mathcal{F}_{t_0}((y,v),(y,v)) = J((x,u);(y,v)) = \int_{t_0}^{t_1} (at+b)\{\cos 2(t+b/a)\}dt = 0.$$

Thus conditions 4.15(i) or 6.4(i) hold with $s = t_0$. Now, the set $\tilde{\mathcal{E}}_0$ (which in this case coincides with $\tilde{\mathcal{E}}_1$) is given by those $(y,v,q) \in X \times \mathcal{U} \times X$ satisfying

$$\dot{y}(t) = v(t), \quad \dot{q}(t) = -(at+b)y(t), \quad q(t) = (at+b)v(t) \quad (t \in T).$$

These relations are equivalent, if $y \in C^2$, to the second order differential equation

$$(at+b)\ddot{y}(t) + a\dot{y}(t) + (at+b)y(t) = 0 \quad (t \in T) \tag{3}$$

which is certainly not satisfied by our choice of (y,v). We conclude that $s = t_0$ belongs to $\mathcal{S}_0(x,u,p,\mu,\gamma)$ and $\mathcal{S}_2(x,u,p,\mu,\gamma)$ and so, by Theorems 4.16 or 6.5, $(x,u,p,\mu,\gamma) \notin \mathcal{H}(\mathcal{U})$, proving the claim. It implies, in particular, that the problem has no solution.

Note that, in this example, both $\mathcal{S}_0$ and $\mathcal{S}_2$ have the advantage over $\mathcal{S}_1$ that we do not have to check the existence of $(z,w) \in Y(\mathcal{U};x,u)$ satisfying $\mathcal{F}_{t_0}((z,w),(y,v)) < 0$. However, it is a simple fact to prove that also this condition holds. Indeed, for any $(z,w) \in Y(\mathcal{U};x,u)$ we have

$$\begin{aligned}\mathcal{F}_{t_0}((z,w),(y,v)) &= \int_{t_0}^{t_1} (at+b)\{\dot{z}(t)\cos(t+b/a) - z(t)\sin(t+b/a)\}dt \\ &= -a\int_{t_0}^{t_1} z(t)\cos(t+b/a)dt\end{aligned}$$

and thus the required inequality clearly occurs if, for example, $z(t) = a\sin(t+b/a)\cos(t+b/a)$, $w(t) = \dot{z}(t)$ $(t \in T)$. Hence $t_0 \in \mathcal{S}_1(\mathcal{U};x,u,p,\mu,\gamma)$. Observe, on the other hand, that once the two conditions defining membership of $\mathcal{S}_1$ are satisfied, the proof of Theorem 5.4(b) gives us a methodology for finding an admissible variation which yields a negative second variation.

So far we have shown in a trivial way, by making use of the sets introduced in this paper, that the problem has no solution. Let us point out that, for a wide range of problems (like this example), one can establish simple criteria for finding $(y,v) \in Y_s(\mathcal{W}_1(u);x,u)$ for which $\mathcal{F}_s((y,v),(y,v))$ vanishes, but the corresponding second condition in the definition of the sets $\mathcal{S}_0$ or $\mathcal{S}_2$ is immaterial (see [3]) in which case, therefore, the point s belongs to them.

Let us now try to apply the necessary conditions obtained in terms of the different sets of "generalized conjugate points" previously defined. To begin with, the main result of [33] (in terms of $\mathcal{G}_0$) cannot be applied because any $(x,u) \in Z$ is singular since $H_{uu}(t) = -(at+b)$ vanishes at $t = t_0$. For the sets of [19] (in terms of $\mathcal{G}_1$) and [32] (in terms of $\mathcal{G}_2$), if a point $s \in [t_0, t_1)$ belongs to any of them then, necessarily, there exists a nonzero $(y,v,q) \in Z_s \times X_s$ with $y(s) = y(t_1) = 0$ such that

$$\dot{y}(t) = v(t), \quad \dot{q}(t) = -(at+b)y(t), \quad v(t)[q(t) - (at+b)v(t)] \geq 0 \quad (t \in [s, t_1]).$$

At this stage we pose the question of how such functions can be found. One could of course try by testing simple functions like the one defined above, or functions with a piecewise (nonzero) constant derivative. They yield, however, a contradiction in the definitions of the two sets. For example, if $y(t) = \sin(t + b/a)$, $v(t) = \cos(t + b/a)$ $(t \in T)$ as before, and $s = t_0$, it is required the existence of a constant $c \in \mathbf{R}$ such that

$$\dot{y}(t)\left[c - \int_{t_0}^{t} (a\tau+b)y(\tau)d\tau - (at+b)\dot{y}(t)\right] = \cos(t+b/a)[c - \sin(t+b/a)] \geq 0 \quad \text{for all } t \in [t_0, t_1].$$

Clearly such a constant does not exist and so the $\mathcal{U}$-admissible variation (y,v), which was used to prove nonemptiness of S_i $(i = 0, 1, 2)$, must be ruled out.

As mentioned before, the approach to conjugacy in terms of $\mathcal{G}_1$ and $\mathcal{G}_2$ seems to be more general than approaches (such as those of $\mathcal{G}_0$, $\mathcal{P}$ or $\mathcal{Q}$ of Section 4) where one is interested in finding a nonzero solution of Jacobi's system, that is, a nonzero $(y,v,q) \in \tilde{\mathcal{E}}_0$. In particular, for the example we are dealing with, for certain nonzero $(y,q) \in Z_s$ satisfying $y(s) = y(t_1) = 0$ and $\dot{q}(t) = -(at+b)y(t)$ $(t \in [s, t_1])$, the former requires the inequality

$$\dot{y}(t)[q(t) - (at+b)\dot{y}(t)] \geq 0 \quad (t \in [s, t_1])$$

while, for the latter, the equality $q(t) = (at+b)\dot{y}(t)$ $(t \in [s, t_1])$ must hold. However, there seems to be no other general criterion for proving nonemptiness of $\mathcal{G}_1$ or $\mathcal{G}_2$ than finding precisely a nonzero solution of Jacobi's system. For this example, in particular, it corresponds to finding a nonnull solution y of (3) with $y(s) = y(t_1) = 0$ but, even for such a simple problem, the general solution to (3) is rather complicated. To give some idea, referring to [22, Equation 2.1.2.103], it is given by

$$y(t) = e^{it}\mathcal{J}(1/2, 1; -2i(at+b)/a)$$

where $\mathcal{J}(\alpha, \beta; t)$ is an arbitrary solution of the degenerate hypergeometric equation

$$t\ddot{y}(t) + (\beta - t)\dot{y}(t) - \alpha y(t) = 0$$

(see [22, Equation 2.1.2.65] for its solution with $\alpha = 1/2$ and $\beta = 1$).

Example 7.4. Let $T = [0, 1]$, $n \geq 3$ an integer, and consider the problem of minimizing

$$I(x,u) = \frac{1}{2}\int_0^1 (\sin t)\{\alpha(t)u^2(t) - \beta(t)x^2(t)\}dt$$

subject to $\dot{x}(t) = u(t)$ $(t \in T)$ and $x(0) = x(1) = 0$, where $\beta(t) = t^{n-2}$ $(t \in T)$ and

$$\alpha(t) = \begin{cases} t^n & \text{if } t \in [0, 1/2) \\ t^{n-2}(1-t)^2 & \text{if } t \in [1/2, 1]. \end{cases}$$

Note that, in this example, $L_{uu}(t, x, u) = \alpha(t) \sin t$ is continuous on the whole interval $[0, 1]$, but its derivative is discontinuous at $t = 1/2$. Also, any process is singular since $H_{uu}(t) = -\alpha(t) \sin t$ vanishes at 0 and 1, so that the theory of conjugacy in terms of $\mathcal{G}_0$ cannot be applied. For the sets $\mathcal{G}_1$ and $\mathcal{G}_2$ we have that, if $s \in [0, 1)$ belongs to any of them then, necessarily, there exists a nonzero $(y, v, q) \in Z_s \times X_s$ with $y(s) = y(1) = 0$ such that

$$\dot{y}(t) = v(t), \quad \dot{q}(t) = -\beta(t)y(t) \sin t, \quad v(t)[q(t) - \alpha(t)v(t) \sin t] \geq 0 \quad (t \in [s, 1]).$$

We arrive again to the question of how such functions can be found, and try by solving Jacobi's equation which, in this case, corresponds on the interval $[0, 1/2)$ to

$$[t^n \sin t]\ddot{y}(t) + [t^n \cos t + nt^{n-1} \sin t]\dot{y}(t) + [t^{n-2} \sin t]y(t) = 0.$$

We omit the equation on $[1/2, 1]$ and proceed no further with the task of trying to prove the (uncertain) nonemptiness of these sets. Instead, simply observe that, if we set $(y(t), v(t)) := (t, 1)$ $(t \in [0, 1/2])$, $(y(t), v(t)) := (1 - t, -1)$ $(t \in [1/2, 1])$ then, clearly,

$$J((x, u); (y, v)) = \int_0^1 (\sin t)\{\alpha(t)v^2(t) - \beta(t)y^2(t)\}dt = 0.$$

Since (y, v) does not satisfy Jacobi's equation, we conclude that $t_0 = 0$ belongs to $\mathcal{S}_0$ and $\mathcal{S}_2$. As in the previous example, one readily verifies that the use of this function yields a contradiction in the definition of the sets $\mathcal{G}_1$ and $\mathcal{G}_2$.

References

[1] R. Berlanga, J. F. Rosenblueth: Jacobi's condition for singular extremals: An extended notion of conjugate points, Appl. Math. Lett. 15 (2002) 453–458.

[2] R. Berlanga, J. F. Rosenblueth: A Sturm-Liouville approach applicable to different notions of conjugacy, Appl. Math. Lett. 17 (2004) 467–472.

[3] R. Berlanga, J. F. Rosenblueth: Extended conjugate points in the calculus of variations, IMA J. Math. Control Inf. 21 (2004) 159–173.

[4] P. Bernhard: La théorie de la seconde variation et le problème linéaire quadratique, in: Advances in Hamiltonian Systems (Rome, 1981), J. P. Aubin et al. (ed.), Birkhäuser, Boston (1983) 109–142.

[5] J. V. Breakwell, Y. C. Ho: On the conjugate point condition for the control problem, Int. J. Eng. Sci. 2 (1965) 565–579.

[6] N. Caroff, H. Frankowska: Conjugate points and shocks in nonlinear optimal control, Trans. Amer. Math. Soc. 348 (1996) 3133–3153.

[7] F. H. Clarke: Methods of Dynamic and Nonsmooth Optimization, CBMS-NSF Regional Conference Series in Applied Mathematics 57, SIAM, Philadelphia (1989).

[8] F. H. Clarke: Optimization and Nonsmooth Analysis, Classics in Applied Mathematics 5, SIAM, Philadelphia (1990).

[9] A. V. Dmitruk: The Euler-Jacobi equation in variational calculus, Mat. Zametki 20 (1976) 847–858 (in Russian); Math. Notes 20 (1976) 1032–1038 (in English).

[10] A. V. Dmitruk: Jacobi-type conditions for the problem of Bolza with inequalities, Mat. Zametki 35 (1984) 813–827 (in Russian); Math. Notes 35 (1984) 427–435 (in English).

[11] E. G. Gilbert, D. S. Bernstein: Second-order necessary conditions in optimal control: Accessory-problem results without normality conditions, J. Optimization Theory Appl. 41 (1983) 75–106.

[12] M. R. Hestenes: Applications of the theory of quadratic forms in Hilbert space to the calculus of variations, Pacific J. Math. 1 (1951) 525–581.

[13] M. R. Hestenes: Quadratic variational theory and linear elliptic partial differential equations, Trans. Amer. Math. Soc. 101 (1961) 306–350.

[14] M. R. Hestenes: Calculus of Variations and Optimal Control Theory, John Wiley & Sons, New York (1966).

[15] M. R. Hestenes: Quadratic variational theory, in: Control Theory and the Calculus of Variations, A. V. Balakrishnan (ed.), Academic Press, New York (1969) 1–37.

[16] M. R. Hestenes: Quadratic control problems, J. Optimization Theory Appl. 17 (1975) 1–42.

[17] M. R. Hestenes: On quadratic control problems, in: Calculus of Variations and Control Theory, D. L. Russell (ed.), Academic Press, New York (1976) 289–304.

[18] M. R. Hestenes: Singular quadratic variational problems, J. Optimization Theory Appl. 41 (1983) 123–137.

[19] P. D. Loewen, H. Zheng: Generalized conjugate points for optimal control problems, Nonlinear Anal., Theory Methods Appl. 22 (1994) 771–791.

[20] D. G. Luenberger: Linear and Nonlinear Programming, 2nd Ed., Addison-Wesley, Reading (1984).

[21] L. W. Neustadt: Optimization. A Theory of Necessary Conditions, Princeton University Press, Princeton (1976).

[22] A. D. Polyanin, V. F. Zaitsev: Handbook of Exact Solutions for Ordinary Differential Equations, CRC Press, Boca Raton (1995).

[23] M. Popescu: Singular normal extremals and conjugate points for Bolza functionals, J. Optimization Theory Appl. 115 (2002) 267–282.

[24] J. F. Rosenblueth: Conjugate intervals for singular trajectories in the calculus of variations, in: Proceedings of the XXXIV National Congress of the Mexican Mathematical Society, G. Contreras et al. (ed.), Aportaciones Matemáticas, Comunicaciones 30, Sociedad Matemática Mexicana, México (2002) 81–98.

[25] J. F. Rosenblueth: Conjugate intervals for the linear fixed-endpoint control problem, J. Optimization Theory Appl. 116 (2003) 393–420.

[26] J. F. Rosenblueth: A conjugate journey in optimal control, Int. Math. Forum 2 (2007) 633–674.

[27] J. F. Rosenblueth: Three improved notions of conjugate points in optimal control, IMA J. Math. Control Inf. 22 (2005) 489–515.

[28] G. Stefani, P. Zezza: Optimality conditions for a constrained optimal control problem, SIAM J. Control Optimization 34 (1996) 635–659.

[29] G. Stefani, P. Zezza: Constrained regular LQ-control problems, SIAM J. Control Optimization 35 (1997) 876–900.

[30] J. Warga: A second-order Lagrangian condition for restricted control problems, J. Optimization Theory Appl. 24 (1978) 475–483.

[31] V. Zeidan: The Riccati equation for optimal control problems with mixed state-control constraints: Necessity and sufficiency, SIAM J. Control Optimization 32 (1994) 1297–1321.

[32] V. Zeidan: Admissible directions and generalized coupled points for optimal control problems, Nonlinear Anal., Theory Methods Appl. 26 (1996) 479–507.

[33] V. Zeidan, P. Zezza: The conjugate point condition for smooth control sets, J. Math. Anal. Appl. 132 (1988) 572–589.

[34] V. Zeidan, P. Zezza: Necessary conditions for optimal control problems: Conjugate points, SIAM J. Control Optimization 26 (1988) 592–608.

[35] V. Zeidan, P. Zezza: Conjugate points and optimal control: Counterexamples, IEEE Trans. Autom. Control 34 (1989) 254–255.

[36] V. Zeidan, P. Zezza: Coupled points in optimal control theory, IEEE Trans. Autom. Control 36 (1991) 1276–1281.