Three improved notions of conjugate points in optimal control

JAVIER F. ROSENBLUETH†

Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, Apartado Postal 20-726, México DF 01000, México

[Received on 15 April 2004; accepted on 25 October 2004]

Three widely quoted approaches to conjugacy in optimal control, due to Zeidan and Zezza, Loewen and Zheng and Zeidan, are studied. 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. No strong normality or non-singularity assumptions are imposed. Moreover, the new sets introduced characterize completely a second-order condition, a property which has remained uncertain for the previous sets.

Keywords: calculus of variations; Jacobi’s necessary condition; generalized conjugate points; non-singular extremals; optimal control; normality.

Oh reader, if that thou canst read and know!
’Tis not enough to spell or even to read
To constitute a reader; there must go
Virtues of which both you and I have need.
Firstly, begin with the beginning (though
That clause is hard); and secondly, proceed;
Thirdly, commence not with the end, or sinning
In this sort, end at least with the beginning.

Byron, Don Juan, XIII, 73

1. Introduction

The reader of this paper is asked to enter into a rather controversial theory. It corresponds to the question of how to generalize the notion of conjugate points in the calculus of variations to an optimal control setting. In order to clearly understand the statement of this question, let us begin by briefly explaining its origins as well as its main role played in the classical theory.

For the basic fixed-end-point problem in the calculus of variations, one is given an interval \( T = [t_0, t_1] \) in \( \mathbb{R} \), two points \( \xi_0, \xi_1 \) in \( \mathbb{R}^n \) and a function \( L \) mapping \( T \times \mathbb{R}^n \times \mathbb{R}^n \) to \( \mathbb{R} \), and one is interested in minimizing a functional of the form

\[
I(x) = \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t)) \, dt,
\]

over the space \( X \) of all piecewise \( C^1 \) functions \( x \) mapping \( T \) to \( \mathbb{R}^n \), and such that \( x(t_0) = \xi_0, x(t_1) = \xi_1 \).

†Email: jfrl@servidor.unam.mx

© The author 2005. Published by Oxford University Press on behalf of the Institute of Mathematics and its Applications. All rights reserved.
Similarly, if \( L \) and the definition itself of conjugate points, may vary considerably.

For this problem, first- and second-order conditions can be easily established under certain smoothness assumptions on the integrand. In particular, if \( L \) is of class \( C^1 \), any solution of the problem belongs to \( \mathcal{E} \), the set of all \( x \in X \) satisfying \( I'(x; y) = 0 \) for all \( y \in Y \), where

\[
I'(x; y) = \int_{t_0}^{t_1} \left( L_x(\bar{x}(t))y(t) + L_{\bar{x}}(\bar{x}(t))\dot{y}(t) \right) \, dt
\]

is the first variation of \( I \) along \( x \), \( (\bar{x}(t)) \) is short for \((t, x(t), \dot{x}(t))\), and

\[
Y = \{ y \in X \mid y(t_0) = y(t_1) = 0 \}.
\]

Similarly, if \( L \) is of class \( C^2 \), any solution of the problem belongs to \( \mathcal{H} \), the set of all \( x \in X \) satisfying \( I''(x; y) \geq 0 \) for all \( y \in Y \), where

\[
I''(x; y) = \int_{t_0}^{t_1} \left\{ \langle y, L_{xx}(\bar{x}(t))y \rangle + 2\langle y, L_{x\bar{x}}(\bar{x}(t))\dot{y} \rangle + \langle \dot{y}, L_{\bar{x}\bar{x}}(\bar{x}(t))\ddot{y} \rangle \right\} \, dt
\]

is the second variation of \( I \) along \( x \).

Now, for a wide range of problems, verifying these conditions may be extremely difficult, and one is thus interested in finding characterizations which simplify them. For the set \( \mathcal{E} \), this is successfully achieved by means of Euler’s equation. If \( x \in X \) and \( L \in C^1 \), then \( x \in \mathcal{E} \Leftrightarrow \) there exists \( c \in \mathbb{R}^n \) such that

\[
L_{\bar{x}}(\bar{x}(t)) = \int_{t_0}^{t_1} L_x(s) \, ds + c \quad (t \in T).
\]

For the set \( \mathcal{H} \), the theory of Jacobi plays a fundamental role. For simplicity of exposition, and because it is one of the best known approaches, let us follow that of Hestenes (1966).

Given \( x \in X \), define a point \( s \in (t_0, t_1] \) as conjugate to \( t_0 \) with respect to \( x \) if there exists \( y \in X \) with \( y(t_0) = y(s) = 0 \), \( y \not\equiv 0 \) on \((t_0, s)\) and, for all \( t \in [t_0, s] \),

\[
q(t) := L_{x\bar{x}}(\bar{x}(t))\dot{y}(t) + L_{\bar{x}\bar{x}}(\bar{x}(t))y(t) \Rightarrow \dot{q}(t) = L_{xx}(\bar{x}(t))y(t) + L_{x\bar{x}}(\bar{x}(t))\ddot{y}(t).
\]

Denote by \( \mathcal{C}(x) \) the set of points conjugate to \( t_0 \) with respect to \( x \), let \( \mathcal{L} \) be the set of all \( x \in X \) such that \( L_{xx}(\bar{x}(t)) \geq 0 \) \((t \in T)\) (Legendre’s condition), and call a trajectory non-singular if \(|L_{x\bar{x}}(\bar{x}(t))| \neq 0 \) \((t \in T)\). It is well known that, if \( L \in C^2 \) and \( x \in \mathcal{E} \cap C^1 \) is non-singular, then \( x \in \mathcal{H} \Leftrightarrow x \in \mathcal{L} \) and \( \mathcal{C}(x) \cap (t_0, t_1) = \emptyset \).

In Hestenes (1966), the elements of \( \mathcal{E} \cap C^1 \) are called extremals, and the condition of Jacobi corresponds to the emptiness of \( \mathcal{C}(x) \cap (t_0, t_1) \). Thus, according to the above result, this condition becomes necessary for optimality if the trajectory under consideration is a non-singular extremal. The non-singularity assumption is crucial. For example, if \( L(t, x, \dot{x}) = x\dot{x} \) for all \((t, x, \dot{x}) \in T \times \mathbb{R}^2\), then for any \( x \in X \), \( \mathcal{C}(x) = (t_0, t_1] \), but \( x \in \mathcal{H} \) since \( I''(x; y) = 0 \) for all \( y \in Y \). Let us also mention that the strengthened condition of Jacobi, stating that \( \mathcal{C}(x) = \emptyset \), yields to sufficiency. Indeed, for local minima, the classical sets of sufficient conditions include \( \mathcal{H} \), the set of all \( x \in X \) satisfying \( I''(x; y) > 0 \) for all \( y \in Y \), \( y \not\equiv 0 \). Similar to the above result we have that, if \( L \in C^2 \) and \( x \) is a non-singular extremal, then \( x \in \mathcal{H} \Leftrightarrow x \in \mathcal{L} \) and \( \mathcal{C}(x) = \emptyset \).

This summary of the classical theory will be our starting point, though the theory of conjugacy is rather controversial even for this case. In Rosenblueth (2005), the reader can find several references of different (not equivalent) ways of stating and proving ‘the’ necessary condition of Jacobi for the basic problem in the calculus of variations, where the assumptions on the sets and functions of the problem, and the definition itself of conjugate points, may vary considerably.

It might be extremely complicated to compare these, or more, approaches to conjugacy. 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 & Zezza (1988a), where an optimal control problem with a fixed-end-point and equality constraints in the control is considered. Basically, (see the details in Section 2), one is interested in minimizing a functional of the form

\[
I(x, u) = g(x(t_1)) + \int_{t_0}^{t_1} L(t, x(t), u(t)) \, dt,
\]

subject to

(a) \((x, u): T \rightarrow \mathbb{R}^n \times \mathbb{R}^m\) with \(x\) piecewise \(C^1\), \(u\) piecewise continuous.

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

(c) \(x(t_0) = x_0, h(x(t_1)) = 0\).

(d) \(u(t) \in U \, (t \in T)\), where \(U = \{u \in \mathbb{R}^m \mid \varphi(u) = 0\}\).

The approach proposed in Zeidan & Zezza (1988a) is quite successful in several respects. For this problem, which we label (P), the fundamental notions of the classical theory have been generalized in the following sense. First, it is obtained a set \(\mathcal{H}\) of processes \((x, u)\) satisfying a second-order condition, and it is shown that all solutions that are ‘normal extremals’ belong to it. Then, it is defined a set \(\mathcal{G}_0(x, u)\) of ‘generalized conjugate points’ and by making use of the so-called ‘accessory problem’ it is proved that under certain two-sided normality conditions (always satisfied in the calculus of variations context), if a ‘non-singular’ normal extremal \((x, u)\) solves the problem (and therefore it belongs to \(\mathcal{H}\)), then the set \(\mathcal{G}_0(x, u) \cap (t_0, t_1)\) is empty.

Some fundamental aspects of this theory were studied in 1994 by Loewen and Zheng (Loewen & Zheng, 1994) who questioned, in particular, the second-order conditions obtained, as well as the non-singularity and two-sided normality assumptions. Let us briefly summarize their main contributions in these respects.

The question of establishing second-order conditions is crucial since the problem posed consists precisely of finding a characterization of these conditions. In Zeidan & Zezza (1988a), it is claimed in an ‘important remark’, that the case when inequality constraints are present (i.e. the control set \(U\) is given by those \(u \in \mathbb{R}^m\) satisfying \(\varphi_\alpha(u) = 0 \, (\alpha \in I_1)\), \(\varphi_\beta(u) \leq 0 \, (\beta \in I_2)\), with \(I_1, I_2\) disjoint index sets), is within the scope of the paper where, in fact, only the equality constraint case is studied. According to Loewen & Zheng (1994), ‘this is neither clear nor convincing’, and the presence of this kind of constraints is a fundamental feature in the theory developed in Loewen & Zheng (1994). The second-order condition obtained in Loewen & Zheng (1994), in the presence of inequality constraints, is shown to be necessary for optimality if the set \(U\) is convex. Now, as mentioned before, the theory of Zeidan & Zezza (1988a) generalizes the classical one in the sense that, when reduced to the calculus of variations problem, both coincide. Loewen and Zheng, on the other hand, define a set \(\mathcal{G}_1(x, u)\) of ‘generalized conjugate points’ and, without making use of the accessory problem, prove that \((x, u) \in \mathcal{H} \Rightarrow \mathcal{G}_1(x, u) \cap (t_0, t_1) = \emptyset\), independent of non-singularity or strong normality assumptions. In the calculus of variations setting, this set contains (but in general does not coincide with) that of classical conjugate points in the
underlying open time interval under non-singularity assumptions. Thus, their approach generalizes not only the classical theory to optimal control but also to singular trajectories in calculus of variations.

This paper had an important impact in the work of Zeidan who, so far, had given possible generalizations of conjugacy from a classical point of view (see also, e.g. Zeidan, 1994; Zeidan & Zezza, 1988b, 1991). In 1996, Zeidan (1996) introduced a new set $G_2(x, u)$ of ‘generalized coupled points’, containing that of Loewen & Zheng (1994) and defined for certain classes of optimal control problems where both end-points vary (in the problem posed above, $g(x(t_1))$ is replaced by $g(x(t_0), x(t_1))$ and condition (c) by $h(x(t_0), x(t_1)) = 0$). These problems include inequality constraints, as in Loewen & Zheng (1994), but the control set does not have to be convex. For the normal case, it is shown, again through the accessory problem, that a necessary condition for optimality is the non-existence of such points in $(t_0, t_1)$.

Now, the sets of conjugate points defined in Zeidan & Zezza (1988a), Loewen & Zheng (1994) and Zeidan (1996) do solve certain fundamental questions of conjugacy, but they present several undesirable features. For the set $G_0$, the non-singularity and strong normality assumptions reduce its range of applicability. Though this aspect is no longer present in $G_1$ and $G_2$, the non-emptiness 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 non-emptiness of these sets may be much more difficult than verifying directly if that condition holds.

Our main objective in this paper is to find an ‘appropriate’ notion of conjugate points, for certain optimal control problems, solving the above undesirable features. To be able to concentrate on the different approaches, and to keep notational complexity to a minimum, we shall deal with the problem posed above with equality constraints. It is important to mention, however, that the different approaches to conjugacy presented in this paper can be generalized to more complex problems. The theory we develop certainly depends on the second-order conditions, but not the approaches themselves.

Thus, we shall study the sets $G_0$, $G_1$ and $G_2$ for problem (P), and improve them in several respects. To begin with, we modify the assumptions on the functions of the problem which are simply ‘recalled’ in Zeidan & Zezza (1988a), Loewen & Zheng (1994) and Zeidan (1996). Based on the well-known book by Neustadt (1976), this modification enlarges considerably the range of problems to which the first- and second-order conditions of these references can be applied. We then introduce three sets of points $S_0$, $S_1$ and $S_2$ with the following properties. The first contains, under non-singularity and strong normality assumptions, the set $G_0 \cap (t_0, t_1)$, while the other two contain $G_1$ and $G_2$, respectively. Moreover, the second-order condition in terms of $H$ is shown to be equivalent to the emptiness of $S_1$ in general, and that of $S_0 \cup S_2$ for the normal case. Finally, we illustrate the fact that verifying membership of $G_0$, $G_1$ or $G_2$ may be extremely difficult, in some cases perhaps even a hopeless task, while for $S_0$, $S_1$ and $S_2$ one can devise simple criteria which, in some cases, can make it an easy or perhaps even a trivial task.

2. Preliminaries

In this section, we pose the optimal control problem we shall be concerned with throughout the paper. It coincides with the one studied in Zeidan & Zezza (1988a). First- and second-order necessary conditions are stated from Zeidan & Zezza (1988a), as well as the basic assumptions on the sets and functions delimiting the problem. We end with a brief discussion on the assumptions, enlarging the range of problems to which these conditions can be applied.
2.1 The problem

Suppose we are given an interval \( T := [t_0, t_1] \) in \( \mathbb{R} \), a point \( \tilde{\zeta}_0 \in \mathbb{R}^n \), open sets \( O \subset \mathbb{R}^n \) and \( V \subset \mathbb{R}^m \) and functions

\[
(g, h): O \rightarrow \mathbb{R} \times \mathbb{R}^k \quad (k \leq n), \quad (L, f): T \times O \times V \rightarrow \mathbb{R} \times \mathbb{R}^n, \quad \varphi: V \rightarrow \mathbb{R}^q \quad (q \leq m).
\]

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 \( V \), set \( Z := X(T, O) \times \mathcal{U}(T, V) \), and define

\[
D := \{(x, u) \in Z \mid \dot{x}(t) = f(t, x(t), u(t)) \quad (t \in T)\},
\]

\[
Z_c(\mathcal{A}) := \{(x, u) \in D \mid (t, x(t), u(t)) \in \mathcal{A} \quad (t \in T)\}, \quad x(t_0) = \tilde{\zeta}_0, \quad x(t_1) \in C, \quad \mathcal{A} = \{(t, x, u) \in T \times O \times V \mid \varphi(u) = 0\} \quad \text{and} \quad C = \{x \in O \mid h(x) = 0\}.
\]

Let \( I: Z \rightarrow \mathbb{R} \) 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 deal with, which we label \( P(\mathcal{A}) \), is that of minimizing \( I \) over \( Z_c(\mathcal{A}) \).

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

\[
\begin{align*}
\text{minimize} \quad & I(x, u) = g(x(t_1)) + \int_{t_0}^{t_1} L(t, x(t), u(t)) \, dt \\
\text{subject to} \quad & (x, u) \in X(T, O) \times \mathcal{U}(T, V). \\
& \dot{x}(t) = f(t, x(t), u(t)) \quad (t \in T). \\
& x(t_0) = \tilde{\zeta}_0, \quad h(x(t_1)) = 0. \\
& \varphi(u(t)) = 0 \quad (t \in T).
\end{align*}
\]

2.2 Solutions

Elements of \( Z \) will be called processes, and a process \((x, u)\) solves \( P(\mathcal{A}) \) if \((x, u) \in Z_c(\mathcal{A})\) and \( I(x, u) \leq I(y, u) \) for all \((y, u) \in Z_c(\mathcal{A})\). Without loss of generality, the theory to follow will be applied to global solutions of the problem instead of, as in Zeidan & Zezza (1988a) and Zeidan (1996), to weak local minima. The reason is simply that the first- and second-order necessary conditions we state hold for any open sets \( O \subset \mathbb{R}^n \) and \( V \subset \mathbb{R}^m \). Thus, shrinking these sets if necessary, the same conditions remain valid for local minima.

2.3 Assumptions

Let \( F := (L, f) \).

**ASSUMPTION A1**

\[
\begin{align*}
\text{a) } & g, \ h, \ \varphi \text{ are } C^1 \text{ and } F(t, \cdot, \cdot) \text{ is } C^1(O \times V) \text{ for all } t \in T. \\
\text{b) } & \text{For all } (x, u) \in O \times V, \ F \text{ and its first-order derivatives in } (x, u) \text{ are piecewise continuous in } t. \\
\text{c) } & \text{There exists an integrable function } \alpha: T \rightarrow \mathbb{R} \text{ such that, for all } (t, x, u) \in T \times O \times V, \\
& |F(t, x, u)| + |F_x(t, x, u)| + |F_u(t, x, u)| \leq \alpha(t).
\end{align*}
\]
ASSUMPTION A2

(a) \( g, h, \varphi \) are \( C^2 \) and \( F(t, \cdot, \cdot) \) is \( C^2(O \times V) \) for all \( t \in T \).
(b) For all \((x, u) \in O \times V \), \( F \) and its second-order derivatives in \((x, u)\) are piecewise continuous in \( t \).
(c) There exists an integrable function \( \beta : T \to \mathbb{R} \) such that, for all \((t, x, u) \in T \times O \times V \),
\[
|F_{xx}(t, x, u)| + |F_{xu}(t, x, u)| + |F_{uu}(t, x, u)| \leq \beta(t).
\]

2.4 Necessary conditions

In order to state the necessary conditions for \( P(A) \) given in Zeidan & Zezza (1988a), we shall define several sets and functions which allow us to write them 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, a statement concerning a ‘regular weak normal extremal’ in Zeidan & Zezza (1988a) may simply be incorrect in Loewen & Zheng (1994), or even Zeidan (1996), where it has a different meaning. Only a few concepts will receive a certain name (such as that of normality) when we find convenient to do so and no confusion should arise.

- For any \( S \subset T \) and \( r \in \mathbb{N} \), let \( X(S, \mathbb{R}^r) \) (resp. \( \mathcal{U}(S, \mathbb{R}^r) \)) be the space of piecewise \( C^1 \) (resp. piecewise continuous) functions mapping \( S \) to \( \mathbb{R}^r \). For simplicity, set \( X := X(T, \mathbb{R}^n), \mathcal{U} := \mathcal{U}(T, \mathbb{R}^m) \).
- For \( u \in V \) and \( x \in O \), define (‘\( * \)’ denotes transpose)
\[
\mathcal{T}(u) := \{ v \in \mathbb{R}^m \mid \varphi'(u)v = 0 \},
\]
\[
\mathcal{N}(x) := \{ p \in \mathbb{R}^n \mid p = h'(x)^* \gamma \text{ for some } \gamma \in \mathbb{R}^k \}.
\]
- Given \((x, u) \in Z \) let \( A(t) := f_x(\tilde{x}(t)), B(t) := f_u(\tilde{x}(t)) \) \((t \in T)\) (we use the notation \((\tilde{x}(t))\) to represent \((t, x(t), u(t))\)), and define
\[
D(x, u) := \{ (y, v) \in X \times V \mid \dot{y}(t) = A(t)y(t) + B(t)v(t) \mid t \in T \},
\]
where
\[
V := \{ v \in \mathcal{U} \mid v(t) \in \mathcal{T}(u(t)) \mid t \in T \}
\]
The set of admissible variations along \((x, u)\) will be given by
\[
Y(x, u) := \{ (y, v) \in D(x, u) \mid y(t_0) = 0, h'(x(t_1))y(t_1) = 0 \}.
\]
- A process \((x, u)\) will be said to be normal to \( P(A) \) if there is no non-null solution \( p \) on \( T \) of the system
\[
\begin{align*}
(a) \quad & \dot{p}(t) + A^*(t)p(t) = 0 \quad (t \in T), \\
(b) \quad & \langle B^*(t)p(t), \zeta \rangle = 0 \quad (\zeta \in \mathcal{T}(u(t)), \ t \in T), \\
(c) \quad & -p(t_1) \in \mathcal{N}(x(t_1)).
\end{align*}
\]
- For all \((t, x, u, p, \mu) \) in \( T \times O \times V \times \mathbb{R}^n \times \mathbb{R}^q \) 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 \((p, \mu) \in X \times \mathcal{U}(T, \mathbb{R}^q)\) satisfying

(i) \(\dot{p}(t) + H_x^r(x(t), p(t), \mu(t)) = 0 \quad (t \in T)\),

(ii) \(H_u(x(t), p(t), \mu(t)) = 0 \quad (t \in T)\),

(iii) \(-[p(t_1) + g'(x(t_1))]^* \in \mathcal{N}(x(t_1))\).

For any \((x, \gamma) \in O \times \mathbb{R}^k\), let \(a_\gamma(x) := g(x) + \gamma^* h(x)\) and note that, if

\[\mathcal{K}(x, p) := \{\gamma \in \mathbb{R}^k \mid p^* = -a_\gamma'(x)\} \quad ((x, p) \in O \times \mathbb{R}^n),\]

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

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

\[\mathcal{E} := \{(x, u, p, \mu) \in \mathcal{X} \mid (x, u) \in D \text{ and } (p, \mu) \in M(x, u)\},\]

\[\mathcal{H} := \{(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k \mid J((x, u); (y, v)) \geq 0, \text{ for all } (y, v) \in \mathcal{Y}(x, u)\},\]

where

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

\[A_\gamma := a_\gamma''(x(t_1)) \text{ and, for all } (t, y, v) \in T \times \mathbb{R}^n \times \mathbb{R}^m,

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

where \(H(t)\) denotes \(H(x(t), p(t), \mu(t))\).

The next result gives first- and second-order necessary conditions for a normal solution to \(P(A)\). It corresponds to Theorem 5.1 of Zeidan & Zezza (1988a).

**Theorem 2.1** Suppose (A1) holds, \((x, u)\) is a normal solution to \(P(A)\), and \(h'(x(t_1))\) and \(\varphi'(u(t))\) \((t \in T)\) are of full rank. Then there exists a unique \((p, \mu, \gamma) \in X \times \mathcal{U}(T, \mathbb{R}^q) \times \mathbb{R}^k\) such that \((x, u, p, \mu) \in \mathcal{E}\) and \(\gamma \in \mathcal{K}(x(t_1), p(t_1))\). If also (A2) holds, then \((x, u, p, \mu, \gamma) \in \mathcal{H}\).

### 2.5 A change on the assumptions

Assumptions (A1) and (A2) are simply ‘recalled’ in Zeidan & Zezza (1988a). Then the Pontryagin principle (which implies the first part of Theorem 2.1) is stated in Zeidan & Zezza (1988a) under assumption (A1). No reference is mentioned at all except for Luenberger (1984) which does not deal with that principle. In Zeidan (1996), on the other hand, the Pontryagin principle is said to be taken from Clarke (1989, 1990) or Hestenes (1966), but the assumptions used in the three references differ considerably from (A1). In Loewen & Zheng (1994), the ‘standing hypotheses’ are similar, but stronger, to those of Zeidan & Zezza (1988a) and Zeidan (1996).

Apparently (from the list of references in Zeidan & Zezza, 1988a; Zeidan, 1996), the assumptions and the principle are taken from Gilbert & Bernstein (1983) where first- and second-order conditions are obtained for a certain Mayer optimal control problem. On the other hand, some results of Neustadt (1976) play a fundamental role in the techniques employed in Gilbert & Bernstein (1983). However, in Neustadt (1976), first-order necessary conditions are derived (in one of several cases) replacing assumption (A1c) by (c').
(c') For each compact subset \( O_c \) of \( O \) and each compact subset \( V_c \) of \( V \), there exists an integrable function \( \alpha : T \to \mathbb{R} \) 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 Neustadt, 1976, p. 283), for the problem we are dealing with, if \( F \) is continuous and continuously differentiable with respect to \((x, u)\), then \( \alpha \) may even be taken to be constant in (c'). This is certainly not the case with respect to (c). More importantly, when dealing with the accessory problem (see the following section) the assumption (A1c), with respect to the integrand \( \Omega \), may not hold in general, implying that the first-order conditions of Zeidan & Zezza (1988a) or Gilbert & Bernstein (1983) may not be applied.

In the remaining part of this paper, we assume that both (A1) and (A2) hold, but replacing (A1c) (and in the obvious way also (A2c)) as above. With this change in the assumptions, as one can verify from the proofs, Theorem 2.1 (based on Neustadt, 1976; Gilbert & Bernstein, 1983; Zeidan & Zezza, 1988a) still holds. Also, given \((x, u) \in Z_c(A)\), we assume from now on that \( h'(x(t)) \) and \( \varphi'(u(t)) \) \((t \in T)\) are of full rank.

3. Zeidan and Zezza

The notion of conjugate points given in Zeidan & Zezza (1988a) is strongly based on the so-called accessory problem to \( P(A) \). Given \((x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k\), this problem, which we label (AP), corresponds to that of minimizing \( J((x, u); \cdot)/2 \) over \( Y(x, u) \), i.e.

\[
\text{minimize } K(y, v) := \frac{1}{2} \int_{t_0}^{t_1} \Omega(t, y(t), v(t)) \, dt \text{ subject to }
\]

(i) \( \dot{y}(t) = A(t)y(t) + B(t)v(t) \) \((t \in T)\).

(ii) \( y(t_0) = 0, \ h'(x(t_1))y(t_1) = 0 \).

Clearly, necessary conditions for this problem may not follow from Theorem 2.1. To begin with, though no constraints of the form \( \varphi(v(t)) = 0 \) are present, as in problem \( P(A) \), a process \((y, v)\) for (AP) is admissible only if \( v \) belongs to \( \mathcal{Y} \), a subset of \( \mathcal{U} = \mathcal{U}(T, \mathbb{R}^m) \). For such problems, we could try to apply the results of Gilbert & Bernstein (1983) but, as mentioned above, the assumption (A1c), with respect to the integrand \( \Omega \), may not hold in general. Under the modified assumptions, let us now define the corresponding sets and functions for (AP) which yield the first-order necessary conditions obtained from Neustadt (1976).

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

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

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

(i) \( \dot{q}(t) + \tilde{H}_y^*(\tilde{y}(t), q(t)) = 0 \) \((t \in T)\),

(ii) \( \langle \tilde{H}_y^*(\tilde{y}(t), q(t)), \zeta \rangle = 0 \) \((\zeta \in \mathcal{T}(u(t)), t \in T)\),

(iii) \(-[q(t_1) + A \gamma y(t_1)] \in \mathcal{N}(x(t_1))\),

and define the set (depending on \((x, u, p, \mu, \gamma)\))

\[
\mathcal{E} := \{(y, v, q) \in \mathcal{X} \times \mathcal{U} \times \mathcal{X} \mid (y, v) \in D(x, u) \text{ and } q \in \tilde{M}(y, v)\}.
\]
Note that \((y, v, q) \in X \times U \times X\) belongs to \(\tilde{E}\) if and only if

1. \(\dot{y}(t) = A(t)y(t) + B(t)v(t), \ v(t) \in T(u(t)) \ (t \in T)\).
2. \(\dot{q}(t) + A^*(t)q(t) = -H_{xX}(t)y(t) - H_{xU}(t)v(t) \ (t \in T)\).
3. \((B^*(t)q(t) + H_{ux}(t)y(t) + H_{uu}(t)v(t), \zeta) = 0 \ (\zeta \in T(u(t)), \ t \in T)\).
4. \(q^*(t_1) = -y^*(t_1)A_y - l^*h'(x(t_1))\) for some \(l \in \mathbb{R}^k\).

The following result gives first-order conditions for a solution of the accessory problem. It can be easily derived (assuming that (A1) and (A2) (modified) hold, transforming the problem into a Mayer one, and using the fact that \(T(u(t))\) is a subspace) from the necessary conditions obtained in Neustadt (1976).

**Lemma 3.1** Let \((x, u, p, \mu, \gamma) \in X \times \mathbb{R}^k\) and suppose \((x, u)\) is normal to \(P(A)\). If \((y, v)\) solves \((AP)\) then there exists \(q \in X\) such that \((y, v, q) \in \tilde{E}\).

Let us now introduce the notion of conjugacy given in Zeidan & Zezza (1988a). It is, in some sense (there are others, as we shall explain later), a ‘natural’ generalization of that concept in the context of calculus of variations, where (see Section 1 and Hestenes, 1966) a point \(s \in (t_0, t_1]\) is conjugate to \(t_0\) if there exists a non-trivial secondary extremal (i.e. an extremal for \((AP)\) vanishing at \(t_0\) and \(s\). In Zeidan & Zezza (1988a), Loewen & Zheng (1994) and Zeidan (1996), contrary to what happens in most texts treating the classical theory, conjugacy is defined not with respect to the initial but the final point of the interval. In calculus of variations, this modification is trivial but, for problem \(P(A)\) in particular, several inequalities appearing in the definitions introduced in those papers make it rather cumbersome. For comparison reasons, however, we shall adopt the approach given in these references.

**Definition 3.2** For any \((x, u, p, \mu, \gamma) \in X \times \mathbb{R}^k\), let \(G_0(x, u, p, \mu, \gamma)\) be the set of points \(s \in [t_0, t_1]\) for which there exists \((y, v, q) \in \tilde{E}\) with \(y(s) = 0, h'(x(t_1))y(t_1) = 0\) and \((y, v, q) \neq (0, 0, 0)\).

In Zeidan & Zezza (1988a), if \((x, u, p, \mu) \in \tilde{E}, \gamma \in K(x(t_1), p(t_1))\), and \((x, u)\) is normal to \(P(A)\), the elements of \(G_0(x, u, p, \mu, \gamma)\) are called points conjugate to \(t_1\).

The necessary condition of Jacobi for optimality (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, 1966, it is non-singular). This condition can be generalized by defining

\[ \mathcal{L}' := \{(x, u, p, \mu) \in X | G^*(t)H_{uu}(\tilde{x}(t), p(t), \mu(t))G(t) < 0, \forall t \in T\}, \]

where \(G\) is piecewise continuous on \(T\) and \(G(t)\) is a matrix whose columns form an orthonormal basis for \(T(u(t)) \ (t \in T)\). However, as trivial examples show (see Zeidan & Zezza, 1989), even if \((x, u, p, \mu)\) belongs to \(\mathcal{L}'\) and \((x, u)\) is normal, the non-existence of such points in the open interval may not be necessary for optimality. Let us mention that the theory developed in Zeidan & Zezza (1988a) is the continuation of three counterexamples given in Zeidan & Zezza (1989) to the notions of conjugacy introduced in Bernhard (1983) and Breakwell & Ho (1965), where it is assured repeatedly that both papers are incorrect (this claim is denied by Bernhard in a reply following Zeidan & Zezza, 1989).

To understand what else might be needed, let us first characterize \(\tilde{E}\), for the non-singular case, in terms of a linear system. The following result is obtained by a direct substitution.
PROPOSITION 3.3 Let \((x, u, p, \mu, \gamma) \in \mathcal{L} \times \mathbb{R}^k\) and \((y, q) \in X \times X\) with 

\[-[q(t_1) + A\gamma y(t_1)] \in N(x(t_1))\] 

Then the following are equivalent:

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

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

\[
\dot{y}(t) = \alpha(t)y(t) + \beta(t)q(t), \quad \dot{q}(t) = \gamma(t)y(t) - A^*q(t) \quad (t \in T),
\]

which we label (J), where

\[
\alpha(t) = A(t) - B(t)\Psi(t)H_{ux}(t), \quad \beta(t) = -B(t)\Psi(t)B^*(t), \quad \gamma(t) = H_{xu}(t)\Psi(t)H_{ux}(t) - H_{xx}(t),
\]

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

Proof. \(\mathcal{N}\) will be denoted by \(\mathcal{N}(\mathcal{P})\).

PROPOSITION 3.4 Let \((x, u, p, \mu, \gamma) \in \mathcal{L} \times \mathbb{R}^k\) and \((y, q) \in X \times X\) with 

\[-[q(t_1) + A\gamma y(t_1)] \in N(x(t_1))\] 

Then the following holds:

(a) \(\dot{p}(t) + A^*p(t) = 0 \quad (t \in [a, b])\),

(b) \(\langle B^*(t)p(t), \zeta \rangle = 0 \quad (\zeta \in \mathcal{T}(u(t)), \ t \in [a, b])\),

(c) \(-p(t_1) \in \mathcal{N}(x(t_1))\) (in case \(b = t_1\)),

has no non-null solution \(p\) on \([a, b]\). The set of all \((x, u) \in \mathcal{Z}\) which are normal to \(\mathcal{P}(\mathcal{A})\) on \([a, b]\) will be denoted by \(N[a, b]\). We continue saying that ‘\((x, u)\) is normal to \(\mathcal{P}(\mathcal{A})\)’ if \((x, u)\) is normal to \(\mathcal{P}(\mathcal{A})\) on \(T\).

We now make a simple observation which helps to understand what lies behind the main result of Zeidan & Zezza (1988a).

PROPOSITION 3.5 Suppose \(s \in (t_0, t_1)\) and \((x, u, p, \mu, \gamma)\) and \((y, v, q)\) are such that

(i) \((x, u, p, \mu) \in \mathcal{L}, (x, u, p, \mu, \gamma) \in \mathcal{H}\), and \((x, u)\) is normal to \(\mathcal{P}(\mathcal{A})\).

(ii) \((y, v, q) \in \dot{\mathcal{E}}\) 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

\[
J((x, u); (z, w)) = \langle y(t_1), A\gamma y(t_1) \rangle + \int_{s}^{t_1} 2\Omega(t, y(t), v(t)) \, dt = 0,
\]

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

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

\[
\dot{r}(t) + A^*r(t) = 0, \quad \langle B^*(t)[q(t) - r(t)], \zeta \rangle = 0, \quad (\zeta \in \mathcal{T}(u(t)), \ t \in [s, t_1]).
\]
(a) Suppose \((x, u) \in N[a, s]\). By (1), \(r(t) = 0\) for all \(t \in [a, s]\). 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,

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

Since \((x, u, p, \mu) \in L'\), 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 \(t \in [s, s]\). Proceeding as above with \([s, t_1]\) replaced by \([s, s]\), we conclude that \(v(t) = 0\) on \([s, s]\).

(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 normality of \((x, u)\) on \(T\) implies that also \(q \equiv 0\).

We are now in a position to clearly understand the main theorem of Zeidan & Zezza (1988a). Theorem 6.1 of that paper corresponds to the following result, whose proof follows straightforwardly from Proposition 3.5(c).

**Theorem 3.6** Suppose \((x, u)\) solves \(P(A)\) 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 \(\mathcal{G}_0(x, u, p, \mu, \gamma) \cap (t_0, t_1) = \emptyset\).

We shall now improve this theorem in several directions.

To begin with, let us mention that, according to Remark 6.1 of Zeidan & Zezza (1988a), ‘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’. On the other hand, in the introduction of Loewen & Zheng (1994), one reads: ‘The claim of Zeidan and Zezza that the normality conditions on both ends are indispensable is because their definition of conjugate points [...] forces them to impose this condition’. This clearly follows from Proposition 3.5. In the definition given in Zeidan & Zezza (1988a), the corresponding \((y, v, q) \in \mathcal{E}'\) 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 3.5(a), it suffices to have \(y \not\equiv 0\) on \([s, t_1]\), and the non-existence of such points in the open interval follows, as in Theorem 3.6, if \((x, u) \in N[t_0, b]\) for all \(b \in (t_0, t_1]\). The same occurs, by 3.5(b), if one imposes the condition \(y \not\equiv 0\) on \([t_0, s]\) and \((x, u) \in N[a, t_1]\) for all \(a \in [t_0, t_1]\). Let us formalize these facts.

**Definition 3.7** For any \((x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k\), let \(\mathcal{C}_0(x, u, p, \mu, \gamma)\) be the set of points \(s \in [t_0, t_1]\) for which there exists \((y, v, q) \in \mathcal{E}'\) with \(y(s) = 0\), \(h'(x(t_1))y(t_1) = 0\) and \(y \not\equiv 0\) on \([t_0, s]\) (resp. \([s, t_1]\)).

Clearly, under the assumptions of Theorem 3.6, \(\mathcal{G}_0 \subseteq \mathcal{C}_0 \cup \mathcal{C}_1\) (all depending on \((x, u, p, \mu, \gamma)\)). Hence, this theorem is a corollary of the following result.

**Theorem 3.8** Suppose \((x, u)\) solves \(P(A)\), \((x, u, p, \mu) \in \mathcal{E} \cap \mathcal{L}'\) and \(\gamma \in \mathcal{K}(x(t_1), p(t_1))\). Then

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

(b) \((x, u) \in N[t_0, b]\) 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 $C_0$ and $C_1$ improve the set $G_0$ of conjugate points provided in Zeidan & Zezza (1988a) because there are problems for which the assumptions of Theorem 3.8(a) or 3.8(b) hold but not both. For such problems, the results of Zeidan & Zezza (1988a) give no information while Theorem 3.8 may be useful to detect non-optimality. Let us illustrate this fact with a simple example.

**EXAMPLE 3.9** 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

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

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

Here, $T = [-\pi, \pi]$, $n$, $m$, $k$ and $q$ equal 1, $\zeta_0 = 0$, $\mathcal{A} = T \times \mathbb{R}^2$, $h(x) = x$,

$$L(t, x, u) = \frac{1}{2}(u^2 - x^2) \quad \text{and} \quad f(t, x, u) = \lambda(t)u \quad ((t, x, u) \in \mathcal{A}).$$

Let $(x, u) \in Z$. Note that $(x, u)$ is normal to $P(\mathcal{A})$ on $[a, b]$ if the equations $\dot{p}(t) = 0$, $\lambda(t)p(t) = 0$ have no non-null 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 $(x, u, p, \mu, \gamma) \in \mathcal{E}$ and $\gamma \in \mathcal{K}(x(t_1), p(t_1))$, we have $(x, u, p, \mu, \gamma) \in \mathcal{L}'$ and the assumptions of Theorem 3.8(a) hold. By this theorem, the problem has no solution if $C_0(x, u, p, \mu, \gamma) \cap (-\pi, \pi) \neq \emptyset$. On the other hand, $(x, u) \not\in N[-\pi, b]$ if $b \in (-\pi, -1]$ and, therefore, we cannot apply Theorem 3.8(b) (nor, in consequence, Theorem 3.6).

Now, as one readily verifies, $\tilde{\mathcal{E}}$ is given by those $(y, v, q) \in Z \times X$ satisfying

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

If we try to apply Theorem 3.8(a), we need to find $s \in (-\pi, \pi)$ and $(y, v, q) \in \tilde{\mathcal{E}}$ 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

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

This example also motivates a second direction for improving the set $G_0$ of Zeidan & Zezza (1988a). As one easily deduces from the proof of Proposition 3.5, the properties required for an element $(y, v, q) \in \tilde{\mathcal{E}}$ with $y(s) = 0$ and $h'(x(t_1))y(t_1) = 0$ actually have 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, non-singularity and the non-vanishing of $y$ may be imposed only in certain subintervals of $T$.

With this in mind, let us introduce the following notation. For any $s \in [t_0, t_1]$, let $T_s := [s, t_1]$, $X_s := X(T_s, \mathbb{R}^n)$, $U_s := U(T_s, \mathbb{R}^m)$, and set $Z_s := X_s \times U_s$. Given $(x, u, p, \mu, \gamma) \in X \times \mathbb{R}^{k}$, let $\tilde{\mathcal{E}}_s$ be
the set of all \((y, v, q) \in Z_s \times X_s\) satisfying

(i) \(\dot{y}(t) = A(t)y(t) + B(t)v(t), v(t) \in T(u(t)) (t \in T_{s})\).

(ii) \(\dot{q}(t) + A^*(t)q(t) = -H_{sx}(t)y(t) - H_{xu}(t)v(t) (t \in T_{s})\).

(iii) \((B^*(t)q(t) + H_{ux}(t)y(t) + H_{uu}(t)v(t), z) = 0 (z \in T(u(t)), t \in T_{s})\).

(iv) \(q^*(t_1) = -y^*(t_1)A_{\gamma} - l^*h'(x(t_1))\) for some \(l \in \mathbb{R}^k\).

For any \(s < t_0\) we set \(\tilde{E}_s := \tilde{E}\).

**DEFINITION 3.10** Given \((x, u, p, \mu, \gamma) \in X \times \mathbb{R}^k\), let \(P(x, u, p, \mu, \gamma)\) 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 \((y, v, q) \in \tilde{E}_{s-\epsilon}\) with \(y(s) = 0, h'(x(t_1))y(t_1) = 0\) and \(y \neq 0\) on \([s - \epsilon, s] \cap T\).

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

Note that, under the assumptions of Theorem 3.8(a), \(C_0 \subset P\) and, under the assumptions of Theorem 3.8(b), \(C_1 \subset P\). Given \(s \in C_0 \cup C_1\), the result follows simply by choosing \(\epsilon = \max[s - t_0, t_1 - s]\).

**THEOREM 3.11** If \((x, u, p, \mu, \gamma) \in \mathcal{H}\) with \((x, u)\) normal to \(P(A)\), then \(P(x, u, p, \mu, \gamma) \cap (t_0, t_1) = \emptyset\).

**Proof.** We proceed basically as in Proposition 3.5. Suppose there exists \(s \in P(x, u, p, \mu, \gamma) \cap (t_0, t_1)\). Let \(\epsilon > 0\) be as in Definition 3.10 and suppose that 3.10(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 \((y, v, q) \in \tilde{E}_{s-\epsilon}\) with \(y(s) = 0, h'(x(t_1))y(t_1) = 0\) and \(y \neq 0\) on \([s - \epsilon, s] \cap T\). 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 (AP), by Lemma 3.1 there exists \(r \in X\) such that \((z, w, r) \in \tilde{E}\). In particular, this implies that

\[ [\dot{q}(t) - \dot{r}(t)] + A^*(t)[q(t) - r(t)] = 0, \]

\[ (B^*(t)y(t) - r(t), \zeta) = 0 \quad (\zeta \in T(u(t)), t \in [s, t_1]). \]

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 3.10(b) holds and \((y, v, q) \in \tilde{E}_s\) is such that

\[ y(s) = 0, \quad h'(x(t_1))y(t_1) = 0, \quad y \neq 0 \quad \text{on} \quad [s, s + \epsilon], \]

there exists \(r \in X\) such that \((z, w, r) \in \tilde{E}_{s}\), where \((z, w)\) is defined as above. Thus,

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

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 \(t \in [s - \epsilon, s + \epsilon]\), contradicting (b).

Let us return to Example 3.9 for which neither Theorem 3.6 nor Theorem 3.8(b) can be applied. In trying to apply Theorem 3.8(a), we faced the problem of finding \((y, q) \in X \times X\) satisfying certain differential equation in the entire interval \([-\pi, \pi]\). That this is no longer required follows by an application of Theorem 3.11. Indeed, let \((y(t), v(t), q(t)) := (\sin t, \cos t, \cos t)\) for all \(t \in [0, \pi]\). Clearly,
\[ (y, v, q) \in \mathcal{E}_0. \] Let \( 0 < \delta \leq \pi \). Since \((x, u) \in N[-\epsilon, 0], y \neq 0 \) on \([0, \epsilon]\) and \( H_{uu}(\hat{x}(t), p(t), 1) = -1 \) for all \( t \in [-\epsilon, \epsilon]\), the point \( s = 0 \) satisfies the conditions of 3.10(b) and so \( 0 \in \mathcal{P}(x, u, p, \mu, \gamma) \). By Theorem 3.11, 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 non-singularity assumption of Theorems 3.6 or 3.8 fails and, as the previous example illustrates, this may also occur if that condition holds. However, as the following trivial example shows, one may face problems for which the theory related to \( \mathcal{P} \) cannot be applied.

**Example 3.12** 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 3.6 or 3.8 since \( H_{uu}(t) = -\lambda(t) \) (\( t \in [-\pi, \pi] \)) and so \( \mathcal{L}' = \emptyset \) (i.e., any \((x, u, p, \mu, \gamma)\) is singular). Now, if \((y, v, q)\) belongs to \( \mathcal{E}_0 \), 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 \( \alpha \in \mathbb{R}, (y(t), q(t)) = (\alpha \sin t, \alpha \cos t) \) (\( t \in [0, \pi] \)), 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 \( t \in [s - \epsilon, s + \epsilon] \).

We end this section by introducing a new set of points which may be used in situations like the one mentioned above. For this set, its emptiness in the open time interval is necessary for normal minimizers without even one-sided normality assumptions.

**Definition 3.13** Given \((x, u, p, \mu, \gamma) \in X \times \mathbb{R}^k \), let \( \mathcal{Q}(x, u, p, \mu, \gamma) \) be the set of points \( s \in [t_0, t_1] \) for which there exists \((y, v, q) \in \mathcal{E}_0 \) with \( y(s) = 0, h'(x(t_1))y(t_1) = 0, \langle H_{uu}(s) + \nu(s), \zeta \rangle \neq 0 \) for some \( \zeta \in T(u(s)) \), and \( B \) is continuous at \( s \).

**Theorem 3.14** If \((x, u, p, \mu, \gamma) \in \mathcal{H} \) with \((x, u)\) normal to \( \mathcal{P}(\mathcal{A}) \), 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 \((y, v, q) \in \mathcal{E}_0 \) and \( \zeta \in T(u(s)) \) be as in Definition 3.13. 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 \((\mathcal{A})\), by Lemma 3.1 there exists \( r \in X \) such that \((z, w, r) \in \mathcal{E} \). In particular, this implies that

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

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

Returning to Example 3.12 note that, if \((y(t), v(t), q(t)) = (\sin t, \cos t, \cos t) \) (\( t \in [0, \pi] \)), then \((y, v, q) \in \mathcal{E}_0 \) for \( s = 0 \), \( H_{uu}(0+) + \nu(0) = -\lambda(0) \neq 0 \) and \( B = 1 \) is continuous at \( s = 0 \). Therefore, \( 0 \in \mathcal{Q}(x, u, p, \mu, \gamma) \) and, by Theorem 3.14, this problem has no solution.

4. Loewen and Zheng

The underlying ideas in the approach to conjugacy given in the previous section are simple. Assuming that \((x, u, p, \mu, \gamma) \in \mathcal{H} \) with \((x, u)\) normal to \( \mathcal{P}(\mathcal{A}) \), it requires the existence of a non-trivial “secondary extremal” \((y, v, q)\) (satisfying the conditions of \( \mathcal{E} \) in a subinterval of \( T \)) with \( y(s) = 0 \) and \( h'(x(t_1))y(t_1) = 0 \). Extending \((y, v)\) to zero outside \([s, t_1]\), one obtains a solution \((z, w)\) of the accessory
problem, and an application of the maximum principle yields the existence of a piecewise $C^1$ function $r$ such that $(z, w, r) \in \tilde{E}$. The definition of $P$ relies on the uniqueness of solutions of the linear system (J) which is satisfied (at least locally) by $(z, w, r)$. The set $Q$ depends basically on the continuity of $r$.

An entirely different approach was devised by Loewen & Zheng (1994). The optimal control problem they consider is precisely $P(A)$ but involving equality and inequality constraints in the control, i.e. $A = T \times O \times U$, where

$$U = \{u \in V \mid \Phi_\alpha(u) = 0 (\alpha \in I_1), \Phi_\beta(u) \leq 0 (\beta \in I_2)\},$$

with $I_1$, $I_2$ disjoint index sets. Second-order conditions are derived assuming convexity of $U$. Their approach is then based, essentially, on a set of ‘generalized conjugate points’ whose non-emptiness implies the existence of a negative second variation.

To introduce this set of points, we shall find convenient to use the following notation. For all $s \in [t_0, t_1)$ recall that $T_s = [s, t_1]$ and $X_s = X(T_s, \mathbb{R}^n)$. Given $(x, u) \in Z$, let as before $A(t) := f_s(\tilde{x}(t))$, $B(t) := f_s(\tilde{x}(t)) (t \in T)$ and define

$$D_s(x, u) := \{(y, v) \in X_s \times \mathcal{V}_s \mid \dot{y}(t) = A(t)y(t) + B(t)v(t) (t \in T_s)\},$$

where $\mathcal{V}_s := \mathcal{V}(T_s)$ and, for any $S \subset T$,

$$\mathcal{V}(S) := \{v \in \mathcal{U}(S, \mathbb{R}^m) \mid v(t) \in T(u(t)) (t \in S)\}.$$

Let

$$Y_s(x, u) := \{(y, v) \in D_s(x, u) \mid y(s) = 0, h'(x(t_1))y(t_1) = 0\}.$$

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

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

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

Let us now introduce, for problem $P(A)$, the set of ‘generalized conjugate points’ defined in Loewen & Zheng (1994).

**Definition 4.1** For any $(x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k$, let $\mathcal{G}_1(x, u, p, \mu, \gamma)$ be the set of points $s \in [t_0, t_1)$ for which there exist $(y, v) \in Y_s(x, u)$ 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) $q(s) \neq 0, -[q(t_1) + A_y y(t_1)] \in N(x(t_1))$, 

(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 $(z, w) \in Y(x, u)$ such that

(i) $\langle z(s), q(s) \rangle > 0$.

(ii) $\langle w(t), \lambda(t) \rangle \geq 0 (t \in T_s)$. 

Downloaded from http://imamci.oxfordjournals.org/ at Universidad Autonoma de Mexico on June 28, 2013
The main result in Loewen & Zheng (1994, Theorem 4.3) relating this set to the condition
\((x, u, p, \mu, \gamma) \in H\) is the following.

**Theorem 4.2** \((x, u, p, \mu, \gamma) \in H \Rightarrow G_1(x, u, p, \mu, \gamma) \cap (t_0, t_1) = \emptyset.\)

In view of this result, the emptiness of \(G_1\) in \((t_0, t_1)\) is necessary for optimality even if the (normal) process under consideration is singular. Several examples given in Loewen & Zheng (1994) deal precisely with singular processes, so that the classical approach of Zeidan & Zezza (1988a) fails to detect non-optimality, while this fact follows by proving non-emptiness of \(G_1\). Note that this last statement also applies to the sets \(P\) and \(Q\) of Section 3.

Let us briefly explain the main difference between \(G_1\) and any of the previous sets of points. To begin with, it is required the existence of \((y, v, q)\) with \(y(s) = 0\) and \(h'(x(t_1))y(t_1) = 0\), satisfying the same conditions defining membership of \(\tilde{E}_s\) except for
\[
(B^*(t)q(t) + H_{ux}(t)y(t) + H_{uu}(t)v(t), z) = 0 \quad (z \in T(u(t)), t \in T_s).
\]

Instead, it is required that
\[
(B^*(t)q(t) + H_{ux}(t)y(t) + H_{uu}(t)v(t), v) \geq 0 \quad (t \in T_s).
\]

Of course, if (3) holds, in which case \((y, v, q)\) is a secondary extremal, then (4) 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 \(t \in [0, s]\) and \((\zeta(t), \eta(t)) := (y(t), v(t))\) if \(t \in [s, t_1]\), then \(J((x, u); (\zeta, \eta)) \leq 0\). If 4.1(a) holds, i.e. if the inequality in (4) is strict on a set of positive measure, then \((x, u, p, \mu, \gamma) \notin H\). Otherwise, 4.1(b) holds and, setting \((y_a, v_a) := (z + a\zeta, w + a\eta)\) it follows that, for some appropriate \(a, J((x, u); (y_a, v_a)) < 0\) and so, once again, \((x, u, p, \mu, \gamma) \notin H\).

Though this approach seems to be successful for certain optimal control problems, let us point out two undesirable features mentioned in the introduction. First, its non-emptiness is established in Loewen & Zheng (1994) merely as a sufficient condition for the existence of negative second variations. It is thus not clear if there are problems with negative second variations for which this set is empty. Second, it is clear that the main objective of introducing a characterization of a second-order condition should be to obtain a simpler way of verifying it. However, one can easily find examples for which to solve the question of non-emptiness of this set may be much more difficult than verifying directly if that condition holds (see, e.g. Berlanga & Rosenblueth, 2002, 2004a,b; Rosenblueth, 2002, 2003, 2005).

These two features can be solved by defining a new set which generalizes a set applicable to the calculus of variations problem, first introduced in Berlanga & Rosenblueth (2002), and generalized for optimal control problems without constraints in Rosenblueth (2005). It is based on the following observation. We are interested in finding certain admissible variation \((y, v)\) in \(Y(x, u)\) for which \(J((x, u); (y, v)) < 0\). This is accomplished if the integrand of the second variation satisfies the required inequality, which is essentially what the set of Loewen & Zheng (1994) imposes. The set of points we now define, \(R(x, u, p, \mu, \gamma)\), imposes that condition in the integral itself. As we shall see presently, this implies, in particular, that the function \(q \in X_s\) of Definition 4.1 is not required at all. We refer the reader to Berlanga & Rosenblueth (2004b) for a wide range of problems for which verifying membership of \(R\) is trivial but that of \(G_1\) may be extremely difficult or perhaps even a hopeless task.

- Given \(s \in [t_0, t_1]\) and \((x, u, p, \mu, \gamma) \in X \times \mathbb{R}^k\), consider the bilinear form
  \[
  F_s((z, w), (y, v)) = \langle z(t_1), A_y y(t_1) \rangle + \int_s^{t_1} \{ [z(t), \sigma(t)] + [w(t), \rho(t)] \} \, dt.
  \]
DEFINITION 4.3 For any \((x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k\), let \(\mathcal{R}(x, u, p, \mu, \gamma)\) be the set of points \(s \in [t_0, t_1)\) for which there exists \((y, v) \in Y_s(x, u)\) such that
\[
\begin{align*}
\mathcal{F}_s((y, v), (y, v)) &\leq 0. \\
\text{(ii)} &\quad \text{There exists } (z, w) \in Y(x, u) \text{ such that } \mathcal{F}_s((z, w), (y, v)) \neq 0.
\end{align*}
\]

THEOREM 4.4 Let \((x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k\). Then the following holds:

\(\begin{align*}
\text{(a)} &\quad \mathcal{G}_1(x, u, p, \mu, \gamma) \subset \mathcal{R}(x, u, p, \mu, \gamma). \\
\text{(b)} &\quad (x, u, p, \mu, \gamma) \in \mathcal{H} \Leftrightarrow \mathcal{R}(x, u, p, \mu, \gamma) = \emptyset.
\end{align*}\)

\textbf{Proof.}

\(\begin{align*}
\text{(a)} &\quad \text{Let } s \in \mathcal{G}_1(x, u, p, \mu, \gamma) \text{ and let } (y, v) \in Y_s(x, u) \text{ and } q \in X_s \text{ be as in Definition 4.1. Condition 4.3(i) follows since}
\quad \mathcal{F}_s((y, v), (y, v)) = \langle y(t_1), A_\gamma y(t_1) \rangle + \int_{s}^{t_1} \{\langle y(t), \dot{q}(t) + A^* t q(t) \rangle + \langle v(t), B^* t q(t) - \lambda(t) \rangle \} dt \\
\quad \qquad + \langle v(t), B^* t q(t) - \lambda(t) \rangle dt = \langle y(t_1), A_\gamma y(t_1) + q(t_1) \rangle - \int_{s}^{t_1} \langle v(t), \lambda(t) \rangle dt \leq 0.
\end{align*}\)

If 4.1(a) holds, then the inequality in 4.3(i) is strict and so \(s \in \mathcal{R}(x, u, p, \mu, \gamma)\). If 4.1(b) holds, then there exists \((z, w) \in Y(x, u)\) such that \(\langle z(s), q(s) \rangle > 0\) and \(\langle w(t), \lambda(t) \rangle \geq 0\) \((t \in T_s)\). Therefore,
\[
\mathcal{F}_s((z, w), (y, v)) \leq \langle z(t_1), A_\gamma y(t_1) + q(t_1) \rangle - \langle z(s), q(s) \rangle - \int_{s}^{t_1} \langle w(t), \lambda(t) \rangle dt < 0.
\]

\(\begin{align*}
\text{(b)} &\quad \Rightarrow: \text{Suppose there exists } s \in \mathcal{R}(x, u, p, \mu, \gamma). \text{ Let } (y, v) \text{ and } (z, w) \text{ be as in Definition 4.3, and let } (\zeta(t), \eta(t)) := (0, 0) \text{ if } t \in [t_0, s], (\zeta(t), \eta(t)) := (y(t), v(t)) \text{ if } t \in [s, t_1]. \text{ By 4.3(i)},
\quad J((x, u); (\zeta, \eta)) = \langle \zeta(t_1), A_\gamma \zeta(t_1) \rangle + \int_{t_0}^{t_1} 2\Omega(t, \zeta(t), \eta(t)) dt \\
\quad \quad = \mathcal{F}_s((y, v), (y, v)) \leq 0.
\end{align*}\)

Set \(k := J((x, u); (z, w)), \beta := \mathcal{F}_s((z, w), (y, v)) \text{ and } \alpha := -(-\beta + k/2\beta). \text{ Then } (y_\alpha, v_\alpha) := (z + \alpha \zeta, w + \alpha \eta) \text{ belongs to } Y(x, u) \text{ and}
\[
J((x, u); (y_\alpha, v_\alpha)) = \langle y_\alpha(t_1), A_\gamma y_\alpha(t_1) \rangle + \int_{t_0}^{t_1} 2\Omega(t, y_\alpha(t), v_\alpha(t)) dt \\
\quad = k + \alpha^2 J((x, u); (\zeta, \eta)) + 2\alpha \mathcal{F}_s((z, w), (y, v)) \\
\quad \leq k + 2\alpha k = -2\beta^2 < 0.
\]

\(\begin{align*}
\text{(b)} &\quad \Leftarrow: \text{Suppose } (x, u, p, \mu, \gamma) \notin \mathcal{H}. \text{ Let } (y, v) \in Y(x, u) \text{ be such that } J((x, u); (y, v)) < 0 \text{ and let } (z, w) \equiv (y, v). \text{ Then } t_0 \in \mathcal{R}(x, u, p, \mu, \gamma). \quad \square
\end{align*}\)

Note that, in view of this result, \((x, u, p, \mu, \gamma) \in \mathcal{H} \Rightarrow \mathcal{G}_1(x, u, p, \mu, \gamma) = \emptyset\). In other words, a necessary condition for optimality is the non-existence of generalized conjugate points in the half-open
interval \([t_0, t_1]\), and not only in \((t_0, t_1)\). Note also that, in the proof of Theorem 4.4(a), the condition 
\(q(s) \neq 0\) is never used so that, in Definition 4.1, it can be removed. The reason is that, if 4.1(a) holds, 
then it is unnecessary and, if 4.1(b) holds, it is redundant.

Now, from the proof of Theorem 4.4, observe that to obtain a negative admissible variation we have 
imposed the condition \(F_s((y, v), (y, v)) \leq 0\), which is implied by the first conditions in the definition 
of Loewen & Zheng (1994). However, we can actually consider functions for which this bilinear form 
is positive and, when perturbed in an appropriate way, may yield a negative second variation. This idea 
motivates the introduction of the following set.

**Definition 4.5** For any \((x, u, p, \mu, \gamma) \in X \times \mathbb{R}^k\), let \(S_1(x, u, p, \mu, \gamma)\) be the set of points \(s \in [t_0, t_1]\) 
for which there exists \((y, v) \in Y_s(x, u)\) such that either (i) or (ii) holds:

(i) \(F_s((y, v), (y, v)) < 0\).

(ii) There exists \((z, w) \in Y(x, u)\) such that

\[
F_s((z, w), (y, v))^2 > F_s((y, v), (y, v))J((x, u); (z, w)).
\]

**Theorem 4.6** Let \((x, u, p, \mu, \gamma) \in X \times \mathbb{R}^k\). Then the following holds:

(a) \(R(x, u, p, \mu, \gamma) \subset S_1(x, u, p, \mu, \gamma)\).

(b) \((x, u, p, \mu, \gamma) \in H \iff S_1(x, u, p, \mu, \gamma) = \emptyset\).

**Proof.** (a) and (b) ‘\(\Rightarrow\)’ are clear. To prove (b) ‘\(\Rightarrow\)’, suppose there exists \(s \in S_1(x, u, p, \mu, \gamma)\). Let 
\((y, v) \in Y_s(x, u)\) be as in Definition 4.5. Since \((x, u, p, \mu, \gamma) \in H, 4.5(i)\) cannot hold. Thus, \(a := F_s((y, v), (y, v)) \geq 0\) and there exists \((z, w) \in Y(x, u)\) such that \(\beta^2 > ak\), where

\[
\beta = F_s((z, w), (y, v)) \quad \text{and} \quad k = J((x, u); (z, w)).
\]

If \(a = 0\) then \(\beta \neq 0\) and so \(s \in R(x, u, p, \mu, \gamma)\), contradicting 4.4(b). Thus, \(a > 0\). Set \((\zeta(t), \eta(t)) := (0, 0)\) if \(t \in [0, s]\), \((\zeta(t), \eta(t)) := (y(t), v(t))\) if \(t \in [s, t_1]\), and let \(\alpha := -\beta/a\). Then \((y_a, v_a) := (z + a_\zeta, \omega + a_\eta)\) belongs to \(Y(x, u)\) and

\[
J((x, u); (y_a, v_a)) = k + a\alpha^2 + 2a\beta = (ak - \beta^2)/a < 0.
\]

\[\square\]

**5. Zeidan**

As a clear response to the paper by Loewen & Zheng (1994), a new set of ‘generalized coupled 
points’ was introduced by Zeidan (1996). It is defined for certain classes of optimal control problems 
where both end-points vary and the control set is not necessarily convex, thus including the problem of 
Loewen & Zheng (1994) as a particular case. For the normal case, it is shown that a necessary 
condition for optimality is again the non-existence of such points in \((t_0, t_1)\). When reduced to problem P(A) 
(and making the necessary changes since Zeidan (1996) refers to a minimum instead of a maximum 
principle), this set corresponds to the following.

**Definition 5.1** For any \((x, u, p, \mu, \gamma) \in X \times \mathbb{R}^k\), let \(G_2(x, u, p, \mu, \gamma)\) be the set of points \(s \in [t_0, t_1]\) 
for which there exist \((y, v) \in Y_s(x, u)\) 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) \(-[q(t_1) + A_T y(t_1)] \in N(x(t_1))\).
(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 \( \alpha \in \mathbb{R}^k \) satisfying

\[
\langle B^*(t)\Phi^*(t)h'(x(t_1))^*\alpha - \lambda(t), \zeta(t) \rangle \geq 0 \quad (\zeta \in V_s, \ t \in T_s),
\]

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 \( \Phi : T \rightarrow \mathbb{R}^{n \times n} \) satisfies \( \Phi(t) = -\Phi(t)A(t) \ (t \in T) \), \( \Phi(t_1) = I_n \), and \( z \) is the solution of

\[
\dot{z}(t) = A(t)z(t) + B(t)w(t) \quad (t \in [t_0, s]), \quad z(t_0) = 0.
\]

The main result in Zeidan (1996, Theorem 5.1), relating this set to the second-order condition, is the following theorem.

**Theorem 5.2** Let \((x, u, p, \mu) \in \mathcal{E}, \gamma \in \mathcal{K}(x(t_1), p(t_1))\), and suppose that \((x, u)\) is normal to \( P(A) \). Then \((x, u, p, \mu, \gamma) \in \mathcal{H} \Rightarrow \mathcal{G}_2(x, u, p, \mu, \gamma) \cap (t_0, t_1) = \emptyset\).

As one readily verifies, this new set contains the set introduced by Loewen and Zheng. Let us give a simple proof of this fact.

**Theorem 5.3** \((x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k \Rightarrow \mathcal{G}_1(x, u, p, \mu, \gamma) \subset \mathcal{G}_2(x, u, p, \mu, \gamma)\).

**Proof.** Let \( s \in \mathcal{G}_1(x, u, p, \mu, \gamma) \) and let \((y, v) \in \mathcal{Y}_s(x, u)\) and \( q \in X_s\) be as in Definition 4.1. If 4.1(a) holds, then \( s \in \mathcal{G}_2(x, u, p, \mu, \gamma) \). If 4.1(a) does not hold, then there exists \((z, w) \in Y(x, u)\) satisfying 4.1(b(i)) and 4.1(b(ii)). Let \( \mathcal{H}(t) := B^*(t)\Phi^*(t)h'(x(t_1))^* \ (t \in T) \) and let \( \alpha \in \mathbb{R}^k \) satisfy

\[
\langle \mathcal{H}(t)\alpha - \lambda(t), \zeta(t) \rangle \geq 0 \quad (\zeta \in V_s, \ t \in T_s).
\]

Using the facts that \( \langle z(t_1), h'(x(t_1))^*\alpha \rangle = 0 \), \( \langle \mathcal{H}(t)\alpha, w(t) \rangle \geq 0 \ (t \in T_s) \), and

\[
z(t) = \int_{t_0}^{t} \dot{\Phi}(t)^{-1}\Phi(\tau)B(\tau)v(\tau) \ d\tau \quad (t \in T_s),
\]

we have

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

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

In Zeidan (1996), two examples are provided for which the results of Loewen & Zheng (1994) cannot be applied because both examples lie outside its scope. One involves a non-convex control set and, in the other, both end-points vary. The conclusion in Zeidan (1996) is then that the new results are ‘much sharper’.

Let us make a few remarks on this conclusion. First of all, note from Theorem 5.2 that the theory of Zeidan (1996) yields a necessary condition for optimality assuming that the process under consideration is normal (compare with Theorems 4.2, 4.4 and 4.6 where normality is not required). In particular, for the linear fixed-end-point 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 \( \mathcal{G}_1, \mathcal{R} \) and \( \mathcal{S}_1 \) can be applied in the abnormal case but not \( \mathcal{G}_2 \). Thus, there are problems which lie beyond the scope of Zeidan (1996) but not of Loewen & Zheng (1994) or the sets introduced in this paper. A full discussion of this fact can be found in Rosenblueth (2003).
On the other hand, it is not clear if, even under normality assumptions, \( G_2 \) is sharper than \( G_1 \) when both sets are applied to the same problem. This question remains open not only for the problem considered in Loewen & Zheng (1994) but also that of Zeidan (1996) since the definition of \( G_1 \) can certainly be extended, once second-order conditions like those of Zeidan (1996) are obtained, to the (more general) problem of Zeidan.

Now, even for the problem we are considering, the two undesirable features of \( G_1 \) explained above remain present with respect to \( G_2 \). Its non-emptiness, 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 \( x \not\in \mathcal{H} \). As before, these two features can be solved by the introduction of a new set containing \( G_2 \).

**Definition 5.4** For any \((x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k\), let \( S_2(x, u, p, \mu, \gamma) \) be the set of points \( s \in [t_0, t_1] \) for which there exists \((y, v) \in Y_s(x, u)\) such that
\[
\begin{align*}
  (i) & \quad \mathcal{F}_s((y, v), (y, v)) \leq 0, \\
  (ii) & \quad \text{if there exists } q \in X_s \text{ such that } \dot{q}(t) + A^*(t)q(t) = \sigma(t), \\
  & \quad \langle B^*(t)q(t) - \rho(t), \zeta \rangle = 0 \quad (\zeta \in \mathcal{T}(u(t)), \ t \in T_s), \\
  \end{align*}
\]
and \(-[q(t_1) + A_\gamma y(t_1)] \in \mathcal{N}(x(t_1)))\), then there exists \( \omega \in \mathcal{V}([t_0, s]) \) with \( \langle z(s), q(s) \rangle \neq 0 \), where \( z \) is the solution of
\[
\dot{z}(t) = A(t)z(t) + B(t)\omega(t) \quad (t \in [t_0, s]), \quad z(t_0) = 0.
\]

**Theorem 5.5** Let \((x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k\). Then the following holds:

(a) \( \mathcal{R}(x, u, p, \mu, \gamma) \cup G_2(x, u, p, \mu, \gamma) \subset S_2(x, u, p, \mu, \gamma) \).

(b) If \((x, u)\) is normal to \( P(A) \) then \((x, u, p, \mu, \gamma) \in \mathcal{H} \iff S_2(x, u, p, \mu, \gamma) = \emptyset \).

**Proof.**

(a) Let \( s \in \mathcal{R}(x, u, p, \mu, \gamma) \) and let \((y, v)\) and \((z, w)\) be as in Definition 4.3. Suppose there exists \( q \in X_s \) such that \( \dot{q}(t) + A^*(t)q(t) = \sigma(t), \)
\[
\langle B^*(t)q(t) - \rho(t), \zeta \rangle = 0 \quad (\zeta \in \mathcal{T}(u(t)), \ t \in T_s),
\]
and \(-[q(t_1) + A_\gamma y(t_1)] \in \mathcal{N}(x(t_1)))\) (if it does not exist, then \( s \in S_2(x, u, p, \mu, \gamma) \)). Therefore,
\[
0 \neq \mathcal{F}_s((z, w), (y, v)) = \langle z(t_1), A_\gamma y(t_1) \rangle + \int_{t_0}^{t_1} \langle \dot{z}(t), q(t) + A^*(t)q(t) \rangle \ dt
\]
\[
+ \langle w(t), B^*(t)q(t) \rangle \ dt
\]
\[
= \langle z(t_1), A_\gamma y(t_1) + q(t_1) \rangle - \langle z(s), q(s) \rangle = -(z(s), q(s)),
\]
and this proves the first contention.

Now, let \( s \in G_2(x, u, p, \mu, \gamma) \) and let \((y, v) \in Y_s(x, u)\) and \( q \in X_s \) be as in Definition 5.1. As in the proof of Theorem 4.4(a), we have \( \mathcal{F}_s((y, v), (y, v)) \leq 0 \). Now, suppose there exists \( q_1 \in X_s \) such that \( q_1(t) + A^*(t)q(t) = \sigma(t), \)
\[
\langle B^*(t)q_1(t) - \rho(t), \zeta \rangle = 0 \quad (\zeta \in \mathcal{T}(u(t)), \ t \in T_s),
\]
and \(-[q_1(t_1) + A_\gamma y(t_1)] \in \mathcal{N}(x(t_1)))\). Note first that
\[
\mathcal{F}_s((y, v), (y, v)) = \langle y(t_1), A_\gamma y(t_1) \rangle + \int_{t_0}^{t_1} \langle \dot{y}(t), q_1(t) + A^*(t)q_1(t) \rangle
\]
\[
+ \langle v(t), B^*(t)q_1(t) \rangle \ dt = 0,
\]
and also
\[\mathcal{F}_s((y, v), (y, v)) = \langle y(t_1), A_\gamma y(t_1) \rangle + \int_{t_1}^{t} \langle \dot{y}(t), \dot{A}_\gamma y(t) + A^*(t)q(t) \rangle dt + \langle v(t), B^*(t)q(t) - \lambda(t) \rangle dt = - \int_{t_1}^{t} \langle v(t), \lambda(t) \rangle dt \leq 0.\]

Therefore, the inequality in 5.1(iii) is equality for all \( t \in T_s \). Now, let \( l, l_1 \in \mathbb{R}^k \) be such that
\[q^*(t_1) = -y^*(t_1)A_\gamma - l^*h'(x(t_1)) \quad \text{and} \quad q_1^*(t_1) = -y^*(t_1)A_\gamma - l_1^*h'(x(t_1)),\]
and define \( r(t) := q(t) - q_1(t) \) \( (t \in T_s) \) and \( \alpha := l - l_1 \). 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 \( \zeta \in V_s \) and \( t \in T_s \),
\[\langle B^*(t)\Phi^*(t)h'(x(t_1))^*\alpha, \zeta(t) \rangle = \langle \lambda(t), \zeta(t) \rangle - \langle B^*(t)q_1(t) - \rho(t), \zeta(t) \rangle = \langle \lambda(t), \zeta(t) \rangle.
\]
By 5.1(iv), there exists \( w \in \mathcal{V}([t_0, s]) \) such that, if \( z \) is the solution of
\[\dot{z}(t) = A(t)z(t) + B(t)w(t) \quad (t \in [t_0, s]), \quad z(t_0) = 0,
\]
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 \( s \in S_2(x, u, p, \mu, \gamma) \).

(b) \( \Rightarrow \): Suppose there exists \( s \in S_2(x, u, p, \mu, \gamma) \). Let \((y, v) \in Y_s(x, u)\) be as in Definition 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 \( t \in [s, t_1] \). Note that, by 5.3(i),
\[J((x, u); (\zeta, \eta)) = \mathcal{F}_s((y, v), (y, v)) \leq 0.
\]
Strict inequality contradicts the assumption \((x, u, p, \mu, \gamma) \in \mathcal{H}\) and, therefore, \((\zeta, \eta)\) solves (AP). By Lemma 3.1, there exists \( q_1 \in X \) such that \((\zeta, \eta, q_1) \in \mathcal{E}\). In particular, this implies that
\[\dot{q}_1(t) + A^*(t)q_1(t) = -H_{xx}(t)\zeta(t) - H_{xu}(t)\eta(t) \quad (t \in T),
\]
\[B^*(t)q_1(t) + H_{ux}(t)\zeta(t) + H_{uu}(t)\eta(t), \quad z(t) = 0 \quad (z \in T(u(t)), \quad t \in T),\]
and \(-[q(t_1) + A_\gamma y(t_1)] \in \mathcal{N}(x(t_1))\). Let \( q \) be the restriction of \( q_1 \) to \([s, t_1]\). By 5.3(ii), there exists \( w \in \mathcal{V}([t_0, s]) \) with \( \langle z(s), q(s) \rangle \neq 0 \), where \( z \) is the solution of
\[\dot{z}(t) = A(t)z(t) + B(t)w(t) \quad (t \in [t_0, s]), \quad z(t_0) = 0.
\]
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
\[0 \neq \langle z(s), q(s) \rangle = \int_{t_0}^{s} \langle \Phi^{-1}(s)\Phi(t)B(t)w(t), q_1(s) \rangle dt = \int_{t_0}^{s} \langle w(t), B^*(t)q_1(t) \rangle dt.
\]
But \( \langle w(t), B^*(t)q_1(t) \rangle = 0 \) for all \( t \in [t_0, s] \) and we reach a contradiction.
(b) ‘\(\iff\)’: Suppose \((x, u, p, \mu, \gamma) \not\in \mathcal{H}\). Let \((y, v) \in Y(x, u)\) be such that \(J((x, u); (y, v)) < 0\).

Clearly, condition 5.3(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, \(t_0 \in S_2(x, u, p, \mu, \gamma)\). \(\square\)

A simple consequence of this theorem is that, if \((x, u)\) is normal to \(P(A)\), then \((x, u, p, \mu, \gamma) \in \mathcal{H} \iff \mathcal{G}_2(x, u, p, \mu, \gamma) = \emptyset\). 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 Zeidan (1996), as well as in Loewen & Zheng (1994), 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 Loewen & Zheng (1994) and Zeidan (1996) one is thus tempted to think that a similar result might hold for the sets \(\mathcal{G}_1\) or \(\mathcal{G}_2\).

Now, one of the basic ideas in the definition of \(S_2\) is the following. Let \((x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k\) with \((x, u)\) normal to \(P(A)\) and suppose that, for some \(s \in [t_0, t_1)\) and \((y, v) \in Y_s(x, u)\), we have \(\mathcal{F}_s((y, v), (y, v)) = 0\). Then, if there does not exist \(q \in X_s\) such that \((y, v, q) \in \mathcal{E}_s\), Lemma 3.1 implies that \((x, u, p, \mu, \gamma) \not\in \mathcal{H}\). Thus, for certain problems, it suffices to find \((y, v)\) as above to conclude non-optimality of the process under consideration. This idea can be incorporated to the sets \(\mathcal{P}\) and \(\mathcal{Q}\) of Section 3 as follows.

**DEFINITION 5.6** For any \((x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k\), let \(S_0(x, u, p, \mu, \gamma)\) be the set of points \(s \in [t_0, t_1)\) for which there exists \((y, v) \in Y_s(x, u)\) such that

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

(ii) If there exists \(q \in X_s\) such that

\[\langle \dot{B}^*(t)q(t) - \rho(t), \zeta \rangle = 0 \quad (\zeta \in T(u(t)), \ t \in T_s)\]

and

\[\langle q(t_1) + A\gamma y(t_1) \rangle \in \mathcal{N}(x(t_1))\]

then \(s > t_0\) and \(s \in \mathcal{P}(x, u, p, \mu, \gamma) \cup \mathcal{Q}(x, u, p, \mu, \gamma)\).

The idea explained above together with an application of Theorems 3.11 and 3.14 clearly yield the following result.

**THEOREM 5.7** Suppose \((x, u, p, \mu, \gamma) \in \mathcal{X} \times \mathbb{R}^k\) and \((x, u)\) is normal to \(P(A)\). Then \((x, u, p, \mu, \gamma) \in \mathcal{H} \iff S_0(x, u, p, \mu, \gamma) = \emptyset\).

Let us summarize the results we have obtained so far. The three improved notions of conjugate points introduced in this paper correspond to \(S_0\), \(S_1\) and \(S_2\). In brief, these sets contain \(G_0 \cap (t_0, t_1)\) (under non-singularity and strong normality assumptions), \(G_1\) and \(G_2\), respectively, i.e. those defined by Zeidan & Zezza (1988a), Loewen & Zheng (1994) and Zeidan (1996). Moreover, the second-order necessary condition in terms of \(\mathcal{H}\) is equivalent to the emptiness of \(S_1\) in general, and that of \(S_0 \cup S_2\) for the normal case. Finally, verifying membership of \(G_0, G_1\) or \(G_2\) may be extremely difficult, in some cases perhaps even a hopeless task, while for \(S_0, S_1\) and \(S_2\) one can devise simple criteria which, in some cases, can make it an easy or perhaps even a trivial task.

6. Two examples

In this last section, we provide two examples which illustrate some of the advantages of the sets introduced in this paper, compared with those of Zeidan & Zezza (1988a), Loewen & Zheng (1994) and
Zeidan (1996). We have chosen simple problems from calculus of variations (and not problems with all the constraints of the general problem studied throughout the paper), allowing the reader to concentrate on the essential differences of the sets involved.

**Example 6.1** Let $\psi : [0, \pi] \to \mathbb{R}$ be any positive, continuous function, and consider the problem of minimizing

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

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

Here, $T = [0, \pi], \xi_0 = 0, O = V = \mathbb{R}, h(x) = x, L(t, x, u) = \psi(t)[u^2 \sin^2 t - x^2 \cos^2 t]/2$ and $f(t, x, u) = u$.

Since $H(t, x, u, p, \mu) = pu - \psi(t)[u^2 \sin^2 t - x^2 \cos^2 t]/2$, $(x, u, p, \mu) \in E$ if and only if

$$\dot{x}(t) = u(t), \quad \dot{p}(t) + \psi(t)x(t) \cos^2 t = 0, \quad p(t) = \psi(t)u(t) \sin^2 t \quad (t \in T).$$

Thus, for any $\mu \in U$, $(x_0, u_0, p, \mu) \equiv (0, 0, 0, \mu)$ belongs to $E$. Also, any process $(x, u)$ is normal, and

$$J((x, u); (y, v)) = \int_0^\pi \psi(t)(v^2(t) \sin^2 t - y^2(t) \cos^2 t) \, dt.$$

Let us try to apply the necessary conditions obtained in terms of the different sets of ‘generalized conjugate points’ to see if the admissible process $(x_0, u_0)$ may be a solution to the problem.

To begin with, the main result of Zeidan & Zezza (1988a, Theorem 3.6, in terms of $G_0$) cannot be applied because any $(x, u) \in Z$ is singular since $H_{uu}(t) = -\psi(t) \sin^2 t$ $(t \in T)$ vanishes at the end-points. For the sets of Loewen & Zheng (1994, Theorem 4.2, in terms of $G_1$) and Zeidan (1996, Theorem 5.2, in terms of $G_2$), if a point $s \in [0, \pi)$ belongs to any of them then, necessarily, there exists $(y, q) \in X_s \times X_s$ with $y(s) = \psi(s) = 0$ such that

$$\dot{q}(t) = -\psi(t)y(t) \cos^2 t, \quad \dot{q}(t)[q(t) - \psi(t) \dot{y}(t) \sin^2 t] \geq 0 \quad (t \in [s, \pi]).$$

At this early stage, we pose the question of how such functions can be found. There seems to be no criterion at all except, possibly, to find a non-zero solution of Jacobi’s system, i.e. a pair $(y, q) \neq (0, 0)$ satisfying

$$\dot{q}(t) = -\psi(t)y(t) \cos^2 t, \quad q(t) = \psi(t) \dot{y}(t) \sin^2 t \quad (t \in [s, \pi]).$$

But recall that $\psi$ is any positive continuous function on $[0, \pi]$ and it is thus uncertain how to find the general solution of such an equation. Note that, if we add the assumption that $\psi$ is differentiable, the second-order differential equation that this approach requires to solve is given, for all $t \in [s, \pi]$, by

$$[\psi(t) \sin^2 t]\dot{y}(t) + [\psi(t) \sin^2 t + 2\psi(t) \sin t \cos t] \dot{y}(t) + [\psi(t) \cos^2 t]y(t) = 0,$$

with $y(s) = \psi(s) = 0$, $y \neq 0$.

Let us turn now to the new sets introduced in this paper. Setting $F_{s}(z, y) := F_{s}((z, w), (y, v))$ for $\dot{z} = w$ and $\dot{y} = v$, we have

$$F_{s}(z, y) = \int_s^\pi \psi(t)[\dot{z}(t)\dot{y}(t) \sin^2 t - z(t)y(t) \cos^2 t] \, dt.$$

Obviously, $F_{s}(y, y) = 0$ for $y(t) = \sin t$ $(t \in [0, \pi])$ and $s = 0$. The requirement of the second condition in the definition of $S_0$ or $S_2$ must be met if there exists $q \in X$ such that

$$\dot{q}(t) = -\psi(t)y \cos^2 t, \quad q(t) = \psi(t) \cos t \sin^2 t \quad (t \in [0, \pi]).$$
Since \( q(0) = 0 \), this corresponds to

\[
q(t) = -\int_0^t \psi(\tau) \sin \tau \cos^2 \tau \, d\tau = \psi(t) \cos t \sin^2 t \quad (t \in [0, \pi]).
\]

From the first equality, we have, in particular, that \( q(\pi/2) \) is strictly negative while from the second it vanishes. Hence, \( s = 0 \) belongs to \( S_0 \cup S_2 \). Since any process is normal to \( P(A) \), and the second variation is independent of the process \( (x, u) \), Theorem 5.5 (or 5.7) implies that this problem has no solution.

Let us mention at this point that, for a wide range of problems (like this example), one can establish simple criteria for finding \( (y, v) \in Y(x, u) \) for which \( F_S((y, v), (y, v)) \) vanishes, but the corresponding second condition in the definition of \( S_0 \) or \( S_2 \) is immaterial (see Berlanga & Rosenblueth, 2004b) in which case, therefore, the point \( s \) belongs to any of these sets.

Now, note that both \( S_0 \) and \( S_2 \) have the advantage over \( S_1 \) that we do not have to check the existence of \( (z, w) \in Y(x, u) \) satisfying \( F_S(z, y) \neq 0 \). For this example, however, it is a simple fact to prove that this condition holds. Indeed, if we set, e.g. \( z(t) = \sin 2t \) for \( t \in [0, \pi/2] \) and \( z(t) = 0 \) for \( t \in [\pi/2, \pi] \), then

\[
F_0(z, y) = \int_0^{\pi/2} \psi(t) (\dot{z}(t) \cos t \sin^2 t - z(t) \sin t \cos^2 t) \, dt
= -2 \int_0^{\pi/2} \psi(t) \cos t \sin^4 t \, dt \neq 0,
\]

and so \( 0 \in \mathcal{R} \subset S_1 \).

**Example 6.2** Consider the problem of minimizing

\[
I(x, u) = \frac{1}{2} \int_0^\pi \psi(t) [u^2(t) - 4x_1^2(t)] \, dt \quad ((x, u) = (x_1, x_2, u)),
\]

subject to \( (\dot{x}_1(t), \dot{x}_2(t)) = (u(t), x_1(t)) \) \( (t \in [0, \pi]) \) and \( x(0) = x(\pi) = 0 \), where \( a = \pi/4 \) and

\[
\psi(t) = t - ia, \quad \text{for all } t \in [ia, (i + 1)a] \quad (i = 0, 1, 2, 3).
\]

In this case, \( T = [0, \pi], n = k = 2, m = 1, O = \mathbb{R}^2, V = \mathbb{R}, \xi_0 = 0, h(x_1, x_2) = (x_1, x_2), \)

\( L(t, x, u) = \psi(t) [u^2 - 4x_1^2]/2 \) and \( f(t, x, u) = (u, x_1) \).

We have \( T(u) = \mathbb{R}, N(x_1, x_2) = \mathbb{R}^2, V = U(T, \mathbb{R}), \)

\[
A(t) = f_x(\tilde{x}(t)) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad B(t) = f_u(\tilde{x}(t)) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},
\]

and \( Y(x_1, x_2, u) \) is given by those \( (y, v) = (y_1, y_2, v) \in X \times U \) with

\[
(y_1(t), y_2(t)) = (v(t), y_1(t)) \quad (t \in [0, \pi]), \quad y(0) = y(\pi) = 0.
\]

Note that any process is normal since \( p = (p_1, p_2) \equiv 0 \) is the only solution of

\[
\begin{pmatrix}
\dot{p}_1(t) + p_2(t) \\
\dot{p}_2(t)
\end{pmatrix} = \begin{pmatrix}
0 \\
0
\end{pmatrix}, \quad p_1(t) = 0.
\]
Now, $H(t, x, u, p, \mu) = p_1 u + p_2 x_1 - \psi(t)(u^2 - 4x_1^2)/2$, so that $(x, u, p, \mu) \in \mathcal{E}$ if and only if, for all $t \in [0, \pi]$, 
\[
\begin{pmatrix}
\dot{x}_1(t) \\
\dot{x}_2(t)
\end{pmatrix} = \begin{pmatrix}
u(t) \\
\dot{y}_2(t)
\end{pmatrix}, \quad \begin{pmatrix}
\dot{p}_1(t) \\
\dot{p}_2(t)
\end{pmatrix} = \begin{pmatrix}
-p_2(t) - 4\psi(t)x_1(t) \\
0
\end{pmatrix},
\]
and $p_1(t) = \psi(t)u(t)$. Thus, $(x_0, u_0, p, \mu) = (0, 0, 0, 0)$ belongs to $\mathcal{E}$. Since $H_{uu}(t) = -\psi(t)$ vanishes at the points $0, \pi/4, \pi/2$ and $3\pi/4$, any process is singular and so we cannot apply the results of Zeidan & Zezza (1988a). For the theory of Loewen & Zheng (1994) and Zeidan (1996) observe that, if a point $s \in [0, \pi)$ belongs to $\mathcal{G}_1$ or $\mathcal{G}_2$ then, necessarily, there exist $(y_1, y_2, v) \in X \times \mathcal{U}$ and $(q_1, q_2) \in X$ such that $y(s) = y(\pi) = 0$ and, for all $t \in [s, \pi]$,
\[
\begin{pmatrix}
\dot{y}_1(t) \\
\dot{y}_2(t)
\end{pmatrix} = \begin{pmatrix}
v(t) \\
\dot{y}_2(t)
\end{pmatrix}, \quad \begin{pmatrix}
\dot{q}_1(t) + q_2(t) \\
\dot{q}_2(t)
\end{pmatrix} = \begin{pmatrix}
-4\psi(t)y_1(t) \\
0
\end{pmatrix},
\]
and $v(t)(q_1(t) - \psi(t)v(t)) \geq 0$. We again leave open the question of how such functions can be found, and turn to the sets introduced in this paper. We have
\[
\mathcal{F}_s((z, w), (y, v)) = \int_s^\pi \psi(t)\{w(t)v(t) - 4z_1(t)y_1(t)\} \, dt.
\]
Set $(y, v) = (y_1, y_2, v)$, where
\[
y_1(t) = \sin 2t, \quad y_2(t) = (1 - \cos 2t)/2 \quad \text{and} \quad v(t) = 2\cos 2t.
\]
Then $(y, v) \in \mathcal{Y}(x, u)$ with $s = 0$ for any $(x, u) \in X$ and, as one readily verifies,
\[
\mathcal{F}_0((y, v), (y, v)) = \int_0^\pi 4\psi(t)\cos 4t \, dt = \sum_{i=0}^3 \int_{ia}^{(i+1)a} 4(t - ia)\cos 4t \, dt = 0.
\]
Moreover, there does not exist $(q_1, q_2) \in X$ satisfying
\[
\dot{q}_1(t) + q_2(t) + 4\psi(t)y_1(t) = 0, \quad \dot{q}_2(t) = 0, \quad q_1(t) - \psi(t)v(t) = 0 \quad (t \in [0, \pi])
\]
for, if this were the case, we would have
\[
q_1(t) = c - \int_0^t 4\psi(\tau) \sin 2\tau \, d\tau = 2\psi(t)\cos 2t \quad (t \in [0, \pi]),
\]
for some constant $c \in \mathbb{R}$. But this implies that $c = q_1(0) = 0$ and, therefore,
\[
q_1(0) = -\int_0^a 4\tau \sin 2\tau \, d\tau = 0,
\]
which is certainly not the case. As in the previous example, this implies that $s = 0$ belongs to $\mathcal{S}_0 \cup \mathcal{S}_2$. By Theorem 5.5 (or 5.7), this problem has no solution.

Note that we left open the question of how to find functions satisfying the conditions defining membership of the sets $\mathcal{G}_1$ and $\mathcal{G}_2$. It remains open but, for this example, let us show that the functions used to prove that $\mathcal{S}_0$ and $\mathcal{S}_2$ are non-empty cannot be applied to those other sets. Indeed, if this were the case, there would exist a constant $c \in \mathbb{R}$ such that, if
\[
a(t) := c - \int_0^t 4\psi(\tau) \sin 2\tau \, d\tau - 2\psi(t)\cos 2t,
\]
then \( \alpha(t) \geq 0 \) \((t \in [0, a])\) and \( \alpha(t) \leq 0 \) \((t \in [a, 2a])\). This implies, in particular, that \( c - \sin 2t \geq 0 \) \((t \in [0, a])\) and \( c - \sin 2t \leq 0 \) \((t \in [a, 2a])\), and we reach a contradiction.

REFERENCES


