Admissible directions in optimal control under uniqueness assumptions

JAVIER F. ROSENBLUETH
IIMAS, Universidad Nacional Autónma de México
Apartado Postal 20-126, CDMX 01000
MEXICO

Abstract: It is well-known that, for a mathematical programming problem involving equality and inequality constraints, the uniqueness of a Lagrange multiplier associated with a local solution implies, under certain smoothness assumptions, second order necessary optimality conditions. Those conditions hold on a set of critical directions defined by those points satisfying the constraints and for which the minimizing function and the standard Lagrangian coincide. No similar links between uniqueness of multipliers and second order conditions seem to have been established for optimal control problems. In this paper, we provide some results in this direction. In particular, we study and completely solve a natural conjecture which provides, under uniqueness assumptions, nonnegative second variations on a classical cone of admissible directions.

Key–Words: Optimal control, uniqueness of Lagrange multipliers, second order conditions, normality

Received: January 10, 2023. Revised: September 13, 2023. Accepted: October 16, 2023. Published: November 15, 2023.

1 Introduction

It is well-known that, for a mathematical programming problem involving equality and inequality constraints, a second order necessary condition for a local solution x_0 holds, under certain smoothness assumptions, on the tangent cone $T_{S_1}(x_0)$ at x_0 of the set S_1 of points satisfying the constraints and for which the minimizing function and the standard Lagrangian coincide. Thus, if x_0 is a *regular* point of S_1 (in the sense that the tangent cone and the set of tangential constraints coincide), the second order condition holds on the set $R_{S_1}(x_0)$ of tangential constraints (or critical directions) at x_0 with respect to S_1.

Regularity can be achieved in different ways. In particular it holds if there is only one Lagrange multiplier associated with the local minimizer. A simple line of reasoning, to support this statement, can be given as follows. First of all, the uniqueness of the multiplier is equivalent to the Mangasarian-Fromovitz constraint qualification with respect to S_1 at x_0. Second, this constraint qualification is equivalent to the condition of normality of x_0 relative to the set S_1. Finally (and this is a crucial result in the theory of mathematical programming), normality implies regularity. Thus, if the set of Lagrange multipliers associated with x_0 is a singleton, then x_0 is a regular point of S_1 and a second order condition holds on $R_{S_1}(x_0)$.

For clarity of exposition, and for comparison reasons, we shall find convenient to "unfold" in the next few lines some of these concepts, main characters, statements and implications.

Consider the nonlinear programming problem, which we label (N), of minimizing f on the set S, where $f, g_i\colon \mathbf{R}^n \to \mathbf{R}$ $(i \in A \cup B)$ are given functions, $A = \{1, \ldots, p\}$, $B = \{p+1, \ldots, m\}$, and

$$S := \{x \in \mathbf{R}^n \mid g_\alpha(x) \le 0 \ (\alpha \in A),$$

$$g_\beta(x) = 0 \ (\beta \in B)\}.$$

We assume, as in [1], [2], [3], that the functions delimiting the problem are continuously differentiable and, when second derivatives occur, they are twice continuously differentiable (for weaker assumptions see, for example, [4], [5], [6], [7]).

Let us begin with the Karush-Kuhn-Tucker (KKT) conditions or first order Lagrange multiplier rule. Denote by $\Lambda(f, x_0)$ the set of all $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbf{R}^m$ satisfying

i. $\lambda_\alpha \ge 0$ and $\lambda_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A)$.

ii. If $F(x) := f(x) + \langle \lambda, g(x) \rangle$ then $F'(x_0) = 0$.

Here, the function F is the standard Lagrangian, g is the function mapping $\mathbf{R}^n$ to $\mathbf{R}^m$ whose components are $g_1, \ldots, g_m$, $\langle \lambda, g(x) \rangle = \sum_1^m \lambda_i g_i(x)$ is the standard inner product, and $\lambda_1, \ldots, \lambda_m$ are the Kuhn-Tucker or Lagrange multipliers.

The KKT conditions become first order necessary conditions for a local solution if a certain *constraint qualification* is imposed. In other words, if x_0 is a local solution to the problem, the nonemptiness of $\Lambda(f, x_0)$ can be assured if some extra assumption on x_0 and the constraints is added to the hypothesis.

Denote by

$$I(x) := \{\alpha \in A \mid g_\alpha(x) = 0\} \quad (x \in S)$$

the set of *active* (or *effective* or *binding*) *indices* at x. From the theory of convex cones (see, for example, [6], [7]) or using the Farkas-Minkowski theorem of the alternative (see [5]), it can be shown that, if

$$R_S(x_0) := \{h \in \mathbf{R}^n \mid g'_\alpha(x_0; h) \leq 0 \ (\alpha \in I(x_0)), \quad g'_\beta(x_0; h) = 0 \ (\beta \in B)\}$$

denotes the set of vectors satisfying the *tangential constraints* at x_0 (see [6], [7]), also called the *linearized tangent cone* or the *cone of locally constrained directions* (see [5]), then

$$\Lambda(f, x_0) \neq \emptyset \Leftrightarrow f'(x_0; h) \geq 0 \text{ for all } h \in R_S(x_0)$$

or, equivalently, $\Lambda(f, x_0) \neq \emptyset \Leftrightarrow -f'(x_0) \in R_S^*(x_0)$ where $B^* := \{z \in \mathbf{R}^n \mid \langle y, z\rangle \leq 0 \text{ for all } y \in B\}$ is the (closed convex) *dual* or *polar cone* of $B \subset \mathbf{R}^n$.

One can find in the literature different answers to the question of how, if x_0 is a local minimum, the relation $-f'(x_0) \in R_S^*(x_0)$ can be assured. In other words, there are different assumptions on the constraints which ensure that the condition $-f'(x_0) \in R_S^*(x_0)$ is a necessary optimality condition for our problem. Not all constraint qualifications coincide, indeed, and a rather intricate web of implications and equivalences has been established (see, for example, [5]).

We shall first give to that question a simple answer (that is, a constraint qualification) by making use of the tangent cone. Moreover, as we shall see, this approach will lead us also to the derivation of second order conditions.

The definition we choose of *tangent cone* is the one given by Hestenes in [6]. As shown in [5], it is equivalent to the one introduced by Bouligand (also known as *contingent cone*). A sequence $\{x_q\} \subset \mathbf{R}^n$ is said to *converge to* x_0 *in the direction* h if h is a unit vector, $x_q \neq x_0$, and

$$\lim_{q\to\infty} |x_q - x_0| = 0, \qquad \lim_{q\to\infty} \frac{x_q - x_0}{|x_q - x_0|} = h.$$

The *tangent cone of* S *at* x_0, denoted by $T_S(x_0)$, is the (closed) cone determined by the unit vectors h for which there exists a sequence $\{x_q\}$ in S converging to x_0 in the direction h. Equivalently (see [7]), $T_S(x_0)$ is the set of all $h \in \mathbf{R}^n$ for which there exist sequences $\{x_q\}$ in S and $\{t_q\}$ of positive numbers, such that

$$\lim_{q\to\infty} t_q = 0, \qquad \lim_{q\to\infty} \frac{x_q - x_0}{t_q} = h.$$

Note that, if $\{x_q\}$ converges to x_0 in the direction h and f has a differential at x_0, then

$$\lim_{q\to\infty} \frac{f(x_q) - f(x_0)}{|x_q - x_0|} = f'(x_0; h).$$

Also, if f has a second differential at x_0, then

$$\lim_{q\to\infty} \frac{f(x_q) - f(x_0) - f'(x_0; x_q - x_0)}{|x_q - x_0|^2} = \tfrac{1}{2} f''(x_0; h).$$

Suppose that x_0 solves (N) locally, $h \in T_S(x_0)$ is a unit vector, and $\{x_q\} \subset S$ is a sequence converging to x_0 in the direction h. Hence, $f(x_q) \geq f(x_0)$ for large values of q and, therefore,

$$0 \leq \lim_{q\to\infty} \frac{f(x_q) - f(x_0)}{|x_q - x_0|} = f'(x_0; h).$$

If also $f'(x_0) = 0$, then

$$0 \leq \lim_{q\to\infty} \frac{f(x_q) - f(x_0)}{|x_q - x_0|^2} = \tfrac{1}{2} f''(x_0; h).$$

This proves the following basic result on first and second order necessary conditions (see, for example, [5, p 303], [6, p 27-28], [7, p 214, 220]).

Theorem 1. *Suppose* x_0 *solves* (N) *locally. Then the following holds.*

a. $f'(x_0; h) \geq 0$ *for all* $h \in T_S(x_0)$.

b. *If* $f'(x_0) = 0$*, then* $f''(x_0; h) \geq 0$ *for all* $h \in T_S(x_0)$.

The dual cone $T_S^*(x_0)$ of the tangent cone of S at x_0 is called the *normal cone* of S at x_0. By the first part of Theorem 1, if x_0 is a local minimum point of a C^1 function f on a set S (actually, merely differentiability at x_0 is required) then the negative gradient $-f'(x_0)$ is an outer normal of S at x_0, that is, $-f'(x_0) \in T_S^*(x_0)$.

Note that $T_S(x_0) \subset R_S(x_0)$ since, if $h \in T_S(x_0)$ is a unit vector, and $\{x_q\} \subset S$ a sequence converging to x_0 in the direction h, then

$$\lim_{q\to\infty} \frac{g_\gamma(x_q) - g_\gamma(x_0)}{|x_q - x_0|} = g'_\gamma(x_0; h) \quad (\gamma \in A \cup B)$$

and, therefore, $g'_\alpha(x_0; h) \le 0$ ($\alpha \in I(x_0)$) and $g'_\beta(x_0; h) = 0$ ($\beta \in B$). Of course, the converse may not hold. For example, if $S = \{x : x^3 \le 0\}$, then $T_S(0) = \{h : h \le 0\}$ but $R_S(0) = \mathbf{R}$.

Hence, we have $R^*_S(x_0) \subset T^*_S(x_0)$. If these two dual cones coincide and x_0 solves (N) locally, then

$$-f'(x_0) \in T^*_S(x_0) = R^*_S(x_0)$$

and so $\Lambda(f, x_0) \neq \emptyset$. The condition

$$R^*_S(x_0) = T^*_S(x_0)$$

is referred to in [7] as *quasi-regularity* and in [5] as the *Guignard-Gould-Tolle* constraint qualification. According to [5, p 264], it "enjoys the best situation with respect to the problem, being the most general constraint qualification for the feasible set of problem (N)."

The notion of "regularity" as defined in [6, p 35], [7, p 221], known also as *Abadie's* constraint qualification (see, for example, [5]), is a crucial one. A point $x_0 \in S$ will be said to be *regular* with respect to S if $T_S(x_0)$ and $R_S(x_0)$ coincide. Clearly, regularity implies quasi-regularity, but one may find quasi-regular points of S which are not regular. These cases, however, are exceptional. To give a simple example (see [7]), this occurs with the origin with respect to the set

$$S = \{(x, y) : x \le 0,\ x + y \le 0,\ x^2 = y^2\}.$$

The same arguments given above yield the following result, the first order Lagrange multiplier rule for a regular solution. It is a direct consequence of Theorem 1(a) and the definition of regularity.

Theorem 2. *If x_0 solves (N) locally and is a regular point of S, then $\Lambda(f, x_0) \neq \emptyset$.*

The second order Lagrange multiplier rule is a straightforward consequence of Theorem 1(b). In what follows, $F(x) = f(x) + \langle\lambda, g(x)\rangle$ denotes (as before) the Lagrangian with respect to $\lambda \in \Lambda(f, x_0)$.

Theorem 3. *Suppose that $x_0 \in S$ and $\lambda \in \Lambda(f, x_0)$. If x_0 solves (N) locally, then $F''(x_0; h) \ge 0$ for all $h \in T_{S_1}(x_0)$ where $S_1 := \{x \in S \mid F(x) = f(x)\}$. In particular, if x_0 is a regular point of S_1, then $F''(x_0; h) \ge 0$ for all $h \in R_{S_1}(x_0)$.*

Note that, if $\lambda \in \Lambda(f, x_0)$ and $\Gamma := \{\alpha \in A \mid \lambda_\alpha > 0\}$, then

$$\begin{aligned} &S_1\ [= S_1(\lambda)] \\ &= \{x \in \mathbf{R}^n \mid g_\alpha(x) \le 0\ (\alpha \in A,\ \lambda_\alpha = 0), \\ &\qquad g_\beta(x) = 0\ (\beta \in \Gamma \cup B)\} \\ &= \{x \in S \mid g_\alpha(x) = 0\ (\alpha \in \Gamma)\}. \end{aligned}$$

Therefore, by definition of tangential constraints, we have

$$\begin{aligned} &R_{S_1}(x_0) \\ &= \{h \in \mathbf{R}^n \mid g'_\alpha(x_0; h) \le 0\ (\alpha \in I(x_0),\ \lambda_\alpha = 0), \\ &\qquad g'_\beta(x_0; h) = 0\ (\beta \in \Gamma \cup B)\} \\ &= \{h \in R_S(x_0) \mid g'_\alpha(x_0; h) = 0\ (\alpha \in \Gamma)\} \\ &= \{h \in R_S(x_0) \mid f'(x_0; h) = 0\}. \end{aligned}$$

In general, it may be difficult to test for quasi-regularity or even regularity, and some criteria implying these conditions is required. As pointed out in [6], "it is customary in the calculus of variations to call a condition on the gradients $g'_1(x_0), \ldots, g'_m(x_0)$ a *normality condition* if it implies regularity at x_0."

In [7], a point $x_0 \in S$ is said to be *normal* relative to S if $\lambda = 0$ is the only solution of

i. $\lambda_\alpha \ge 0$ and $\lambda_\alpha g_\alpha(x_0) = 0$ ($\alpha \in A$).
ii. $\sum_1^m \lambda_i g'_i(x_0) = 0$.

Normality implies regularity. A proof of this fact can be found in [7] where another condition, called properness, is used to establish this implication. A point $x_0 \in S$ is said to be *proper* relative to S if the set $\{g'_\beta(x_0) \mid \beta \in B\}$ is linearly independent and, if $p > 0$, there exists h such that

$$g'_\alpha(x_0; h) < 0\ (\alpha \in I(x_0)),\ g'_\beta(x_0; h) = 0\ (\beta \in B).$$

As shown in [7, p 241], properness and normality are equivalent. This is also proved in [5, p 256] by using theorems of alternative (see Motzkin in [5, Theorem 2.4.19]), and in [4, p 43] for inequality constraints, in terms of positive linear independence, by an application of the Hahn-Banach separation theorem.

Since normality is equivalent to properness and they imply regularity, Theorem 2 tells us that both are constraint qualifications. They are also referred to in the literature as the *Cottle-Dragomirescu* and the *Mangasarian-Fromovitz* constraint qualifications respectively. In view of Theorem 2, we have the following classical result.

Theorem 4. *If x_0 solves (N) locally and is a normal point of S, then $\Lambda(f, x_0) \neq \emptyset$.*

It is important to point out that the conclusion that normality is a constraint qualification can be reached without making use of Theorem 2 and the notion of regularity. The well-known Fritz John necessary optimality condition (see, for example, [1], [5], [7], [8] and, for a nonsmooth multiplier rule, see [4]) allows the cost multiplier to vanish. It states that, if x_0 solves (N) locally, then there exist $\lambda_0 \geq 0$ and $\lambda \in \mathbf{R}^m$, not both zero, such that

i. $\lambda_\alpha \geq 0$ and $\lambda_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A)$.
ii. $F_0(x) := \lambda_0 f(x) + \langle \lambda, g(x) \rangle \Rightarrow F_0'(x_0) = 0$.

In view of this result, if x_0 solves (N) locally and is a normal point of S, then $\lambda_0 > 0$ (since, by normality relative to S, if $\lambda_0 = 0$ then $\lambda = 0$, contradicting the nontriviality condition) and the multipliers can be chosen so that $\lambda_0 = 1$, thus implying nonemptiness of $\Lambda(f, x_0)$.

We are now in a position to state a fundamental result on second order necessary conditions under normality assumptions. It is an immediate consequence of Theorem 3 and the well-known link between normality and regularity. For the application of this result note that, given $\lambda \in \Lambda(f, x_0)$, x_0 is a normal point of $S_1(\lambda)$ if $\mu = 0$ is the only solution of

i. $\mu_\alpha \geq 0$ and $\mu_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A,\ \lambda_\alpha = 0)$.
ii. $\sum_1^m \mu_i g_i'(x_0) = 0$.

Theorem 5. *Suppose that* $x_0 \in S$ *and* $\lambda \in \Lambda(f, x_0)$. *If* x_0 *solves* (N) *locally and is a normal point of* $S_1(\lambda)$, *then* $F''(x_0; h) \geq 0$ *for all* $h \in R_{S_1}(x_0)$.

Let us now introduce one of the main characters which plays a leading role in the theory: the *linear independence* constraint qualification (LI). For a given point $x_0 \in S$, it asks the set $\{g_\gamma'(x_0) \mid \gamma \in I(x_0) \cup B\}$ to be linearly independent.

This is, by the way, the definition of *normality* given in [6], in contrast with the definition in [7] which we introduced before. As one readily verifies, it corresponds to normality of x_0 with respect to (neither S nor S_1 but) the set of equality constraints for active indices:

$$S_0\,[= S_0(x_0)\,]$$

$$:= \{x \in \mathbf{R}^n \mid g_\gamma(x) = 0\ (\gamma \in I(x_0) \cup B)\}.$$

This follows since, by definition, $x \in S_0(x_0)$ is normal relative to $S_0(x_0)$ if $\lambda = 0$ is the only solution of

i. $\lambda_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A)$.
ii. $\sum_1^m \lambda_i g_i'(x) = 0$.

This is equivalent to LI, the condition that the linear equations $g_\gamma'(x_0; h) = 0$ $(\gamma \in I(x_0) \cup B)$ in h be linearly independent.

Note that, given $x_0 \in S$ and $\lambda \in \mathbf{R}^m$ with $\lambda_\alpha \geq 0$ $(\alpha \in A)$, we have $R_{S_0}(x_0) \subset R_{S_1}(x_0) \subset R_S(x_0)$ where, by definition,

$$R_{S_0}(x_0) = \{h \in \mathbf{R}^n \mid g_\gamma'(x_0; h) = 0$$

$$(\gamma \in I(x_0) \cup B)\}.$$

Also, if x_0 is a normal point of S_0, then it is a normal point of S_1, and hence a normal point of S.

In most textbooks (see a thorough explanation and references in [1]) the main result on second order necessary conditions differs from Theorem 5 in two fundamental aspects: it is derived under the assumption of the linear independence constraint qualification (that is, normality relative to S_0 instead of S_1) and the second order condition holds on the set of critical directions $R_{S_0}(x_0)$ (instead of $R_{S_1}(x_0)$). Thus, in this result, the assumptions are stronger and the conclusions weaker than those of Theorem 5 since, usually, normality relative to S_0 is stronger than normality relative to S_1 and $R_{S_1}(x_0)$ contains properly the set $R_{S_0}(x_0)$. As pointed out in [1], “The source of this weaker result can be attributed to the traditional way of treating the active inequality constraints as equality constraints.” Explicitly, this rather “well-worn result” (see [1]) is the following.

Theorem 6. *Suppose* $x_0 \in S$ *and* $\lambda \in \Lambda(f, x_0)$. *If* x_0 *solves* (N) *locally and is a normal point of* $S_0(x_0)$, *then* $F''(x_0; h) \geq 0$ *for all* $h \in R_{S_0}(x_0)$.

The set of critical directions in Theorem 5 is then, in general, bigger than that of Theorem 6. However, as the following examples show, even under the strong assumption of Theorem 6 (normality relative to $S_0(x_0)$ or the LI constraint qualification), the set of critical directions cannot be “too big”: we may very well have negative second variations on $R_S(x_0)$.

Example 1. Consider the problem of minimizing $f(x_1, x_2) = x_1$ subject to

$$g_1(x_1, x_2) = -x_1^2 - x_2 \leq 0,$$

$$g_2(x_1, x_2) = x_1 + x_2 = 0.$$

Clearly $x_0 = (1, -1)$ is a local solution to the problem and LI is satisfied since the gradients $g_1'(x_0) = (-2, -1)$ and $g_2'(x_0) = (1, 1)$ are linearly independent. We have

$$F(x_1, x_2) = x_1 + \lambda_1(-x_1^2 - x_2) + \lambda_2(x_1 + x_2)$$

and so $F'(x_1, x_2) = (1 - 2\lambda_1 x_1 + \lambda_2, -\lambda_1 + \lambda_2)$. Thus $F'(x_0) = 0$ implies $\lambda_1 = \lambda_2 = 1$. Also,

$$F''(x_1, x_2) = \begin{pmatrix} -2\lambda_1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$$

and so $F''(x_0; h, k) = -2h^2$. Now, we have

$$g_1'(x_1, x_2) = (-2x_1, -1), \quad g_1'(x_1, x_2) = (1, 1)$$

and so

$$R_S(x_0) = \{(h, k) \mid -2h - k \le 0,\ h + k = 0\}.$$

Therefore $(1, -1) \in R_S(x_0)$ but $F''(x_0; 1, -1) = -2 < 0$. ▮

Example 2. Consider the problem of minimizing $f(x) = -x^3$ subject to $g(x) = x - 1 \le 0$.

Clearly $x_0 = 1$ is a (global) solution and LI holds since $g'(x_0) = 1 \ne 0$. We have

$$F(x) = f(x) + \lambda g(x) = -x^3 + \lambda x - \lambda$$

and so $F'(x) = -3x^2 + \lambda$ and $F'(x_0) = 0$ implies that $\lambda = 3$. Also $F''(x_0; h) = -6h^2$. Since

$$R_S(x_0) = \{h \mid g'(x_0; h) = h \le 0\}$$

we have $-1 \in R_S(x_0)$ and $F''(x_0; -1) = -6 < 0$. ▮

We can now explain, in a clear and succinct way, the role played by the uniqueness of Lagrange multipliers in the intricate web of constraint qualifications.

Let us begin with a simple argument which implies uniqueness. Suppose $x_0 \in S$ and $\lambda \in \Lambda(f, x_0)$. If also $\bar\lambda \in \Lambda(f, x_0)$, set $\mu := \bar\lambda - \lambda$ and observe that

$$\mu_\alpha g_\alpha(x_0) = \bar\lambda_\alpha g_\alpha(x_0) - \lambda_\alpha g_\alpha(x_0) = 0 \ (\alpha \in A),$$

$$\sum_1^m \mu_i g_i'(x_0) = \sum_1^m \bar\lambda_i g_i'(x_0) - \sum_1^m \lambda_i g_i'(x_0) = 0.$$

Thus, if x_0 is normal relative to $S_0(x_0)$, then $\mu = 0$ and $\Lambda(f, x_0)$ would be a singleton. In other words, the following result holds.

Theorem 7. *Suppose $x_0 \in S$ and $\lambda \in \Lambda(f, x_0)$. If x_0 is normal relative to $S_0(x_0)$, then $\Lambda(f, x_0) = \{\lambda\}$.*

In view of Theorem 4, if x_0 affords a local minimum to f on S and x_0 is normal relative to S_0 (and so normal relative to S) then there exists $\lambda \in \Lambda(f, x_0)$. We have just proved that, in this event, $\lambda \in \Lambda(f, x_0)$ is unique. As one readily verifies, if normality is assumed relative to S, the existence of $\lambda \in \Lambda(f, x_0)$ can be assured, but it may not be unique. Here we arrive at a crucial result in the theory.

In [2] it is shown that uniqueness of the Lagrange multiplier associated with the point x_0 can be achieved by an assumption weaker than that of normality relative to $S_0(x_0)$, namely, normality relative to $S_1(\lambda)$. Moreover, this assumption is not only sufficient for the uniqueness of the multiplier but also necessary. This is the content of the following result.

Theorem 8. *Suppose that $x_0 \in S$ and $\lambda \in \Lambda(f, x_0)$. Then x_0 is normal relative to $S_1(\lambda)$ if and only if $\Lambda(f, x_0) = \{\lambda\}$.*

In [2], the statement of this result is expressed in terms of a condition called the *strict Mangasarian-Fromovitz* constraint qualification which is no other than properness relative to $S_1(\lambda)$. The proof given in that paper, however, relies precisely on the equivalence between normality and properness.

Combining this result with Theorem 5, we finally reach the main result on second order necessary conditions under uniqueness assumptions.

Theorem 9. *Suppose that $x_0 \in S$ and $\lambda \in \Lambda(f, x_0)$. If x_0 solves (N) locally and $\Lambda(f, x_0) = \{\lambda\}$, then $F''(x_0; h) \ge 0$ for all $h \in R_{S_1}(x_0)$.*

It is worth mentioning that an entirely different approach, found in [3], is based on the idea that, since constraint qualifications are independent of the objective function f, if a constraint qualification implies a certain property for the Lagrange multipliers, this property will hold *for all* objective functions for which x_0 affords a local minimum. If we define $\mathcal{F}(x_0)$ as the set of all $f \in C^1(\mathbf{R}^n, \mathbf{R})$ such that x_0 affords a local minimum to f on S, the result on uniqueness of Lagrange multipliers given in [3] states that, if $x_0 \in S$, then $\Lambda(f, x_0)$ is a singleton for all $f \in \mathcal{F}(x_0) \Leftrightarrow x_0$ is normal relative to $S_0(x_0)$.

Uniqueness of multipliers in different areas of constrained optimization has been studied from different points of view in order to deal with, to mention a few, problems subject to cone constraints [9], sensitivity analysis of optimization problems [10], composite optimization [11], sets of functions with a common local minimum [3], or derivation of second order conditions [2], [5], [6], [7], [12], [13], [14].

In this paper we shall be concerned with uniqueness of multipliers for certain classes of optimal control problems. Our main objective will be to state the main aspects of the theory in that context and give some answers to questions related to possible links between uniqueness of multipliers, admissible directions, and second order optimality conditions.

The stage for such problems, involving equality and inequality constraints in the control functions, is rather different than that for the finite dimensional case. In particular, it has been proved that the corresponding Mangasarian-Fromovitz constraint qualification (also called properness) and the condition of normality are equivalent [15], as in the mathematical programming problem, but these conditions do not necessary imply regularity [16]. Moreover, the Mangasarian-Fromovitz constraint qualification with respect to S_1 implies uniqueness of the Lagrange multiplier, but the converse may not hold [17], [18].

It is of interest to know if, for optimal control problems, the uniqueness of Lagrange multipliers implies or not second order conditions on a specific set of critical directions. Let us move forward to that question.

2 Statement of the problem

In the remaining of this paper we shall deal with an optimal control problem with fixed endpoints, posed over piecewise C^1 trajectories and piecewise continuous controls, and involving inequalities and equalities in the control functions.

We believe this problem, though (relatively) simple compared with other formulations, captures the essence of the question connecting uniqueness of multipliers with second order conditions. The problem is not simple, by any means. Actually, the difficulties encountered for this kind of optimal control problems are of a much subtler nature than those for (N). As explained in [4, p 335], the type of constraints we shall now deal with (even for the classical Lagrange problem in the calculus of variations with equality constraints) "make the constrained optimal control problem much more complex than (N), or even the isoperimetric problem. In part, this is because we now have infinitely many constraints, one for each t."

To state the problem, suppose we are given an interval $T := [t_0, t_1]$ in $\mathbf{R}$, two points ξ_0, ξ_1 in $\mathbf{R}^n$, and functions L and f mapping $T \times \mathbf{R}^n \times \mathbf{R}^m$ to $\mathbf{R}$ and $\mathbf{R}^n$ respectively, and $\varphi = (\varphi_1, \ldots, \varphi_q)$ mapping $\mathbf{R}^m$ to $\mathbf{R}^q$. Denote by X the space of piecewise C^1 functions mapping T to $\mathbf{R}^n$, and by $\mathcal{U}_k$ the space of piecewise continuous functions mapping T to $\mathbf{R}^k$ ($k \in \mathbf{N}$). Let $Z := X \times \mathcal{U}_m$, and set

$$D := \{(x, u) \in Z \mid \dot{x}(t) = f(t, x(t), u(t)) \ (t \in T),$$
$$x(t_0) = \xi_0, \ x(t_1) = \xi_1\},$$
$$S := \{(x, u) \in D \mid \varphi_\alpha(u(t)) \leq 0,$$
$$\varphi_\beta(u(t)) = 0 \ (\alpha \in R, \ \beta \in Q, \ t \in T)\}$$

where $R = \{1, \ldots, r\}$ and $Q = \{r+1, \ldots, q\}$. For all (x, u) in Z, let

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

The problem we shall deal with, which we label (P), is that of minimizing I over S.

A *control function* u is an element of $\mathcal{U}_m$ and a *state trajectory* x corresponding to a control function u is an element of X which solves the differential equation $\dot{x}(t) = f(t, x(t), u(t))$ ($t \in T$). A pair (x, u) comprising a state trajectory x and an associated control function u is referred to as a *process*. If u satisfies the control constraints $u(t) \in U$ ($t \in T$), where

$$U := \{u \in \mathbf{R}^m \mid \varphi_\alpha(u) \leq 0 \ (\alpha \in R),$$
$$\varphi_\beta(u) = 0 \ (\beta \in Q)\},$$

and x satisfies the endpoint constraints $x(t_0) = \xi_0$ and $x(t_1) = \xi_1$, then the process (x, u) is said to be *admissible*, and it *solves* (P) if it achieves the minimum value of I over all admissible processes.

With respect to the functions delimiting the problem, we assume that, if $F := (L, f)$, then $F(t, \cdot, \cdot)$ is C^2 for all $t \in T$ and φ is C^2; $F(\cdot, x, u)$ and its derivatives in (x, u) are piecewise continuous and there exists an integrable function $\alpha: T \to \mathbf{R}$ such that, at any point $(t, x, u) \in T \times \mathbf{R}^n \times \mathbf{R}^m$,

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

Moreover, the $q \times (m + r)$-dimensional matrix

$$\begin{pmatrix} \dfrac{\partial \varphi_i}{\partial u^k} & \delta_{i\alpha}\varphi_\alpha \end{pmatrix}$$

($i = 1, \ldots, q$; $\alpha = 1, \ldots, r$; $k = 1, \ldots, m$), has rank q on U (here $\delta_{\alpha\alpha} = 1$, $\delta_{\alpha\beta} = 0$ ($\alpha \neq \beta$)). This condition is equivalent to the condition that, at each point u in U, the matrix

$$\begin{pmatrix} \dfrac{\partial \varphi_i}{\partial u^k} \end{pmatrix} \qquad (i = i_1, \ldots, i_p; \ k = 1, \ldots, m)$$

has rank p, where $i_1, \ldots, i_p$ are the indices $i \in \{1, \ldots, q\}$ such that $\varphi_i(u) = 0$ (see [12] for details). Finally, given $(x_0, u_0) \in S$, we shall denote by $[t]$ the point $(t, x_0(t), u_0(t))$ and set $A(t) := f_x[t]$ and $B(t) := f_u[t]$. In what follows, the notation '*' will be used to denote transpose.

3 Normality and first order conditions

First order necessary conditions are well established and one version (consequence of the Maximum Principle) can be stated as follows (see, for example, [6]). This result can be seen as the analogous of the Fritz John necessary optimality condition in mathematical programming mentioned in the introduction.

For all $(t, x, u, p, \mu, \lambda)$ in $T \times \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^n \times \mathbf{R}^q \times \mathbf{R}$ define the Hamiltonian function as

$$H(t, x, u, p, \mu, \lambda) :=$$

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

Theorem 10. *Suppose* (x_0, u_0) *solves* (P). *Then there exist* $\lambda_0 \geq 0$, $p \in X$, *and* $\mu \in \mathcal{U}_q$, *not vanishing simultaneously on* T, *such that*

a. $\mu_\alpha(t) \geq 0$ *and* $\mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ t \in T)$*;*

b. $\dot p(t) = -H_x^*(t, x_0(t), u_0(t), p(t), \mu(t), \lambda_0)$ *on every interval of continuity of* u_0*;*

c. $H_u(t, x_0(t), u_0(t), p(t), \mu(t), \lambda_0) = 0$ $(t \in T)$.

In this theorem, the case $\lambda_0 = 1$, as in the theory of mathematical programming, is particularly relevant. Denote by $\mathcal{E}$ the set of all $(x_0, u_0, p, \mu) \in Z \times X \times \mathcal{U}_q$ satisfying

a. $\mu_\alpha(t) \geq 0$ and $\mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ t \in T)$;

b. $\dot p(t) = -A^*(t)p(t) + L_x^*[t]$ $(t \in T)$;

c. $B^*(t)p(t) = L_u^*[t] + \varphi'^*(u_0(t))\mu(t)$ $(t \in T)$.

Given $(x_0, u_0) \in S$, let $\Lambda(x_0, u_0)$ be the set of all $(p, \mu) \in X \times \mathcal{U}_q$ such that $(x_0, u_0, p, \mu) \in \mathcal{E}$. The elements of $\mathcal{E}$ will be called *extremals* and of $\Lambda(x_0, u_0)$ *Lagrange multipliers* (no confusion should arise with the notation and terminology used in Section 1).

As for the finite dimensional case, given a solution (x_0, u_0) to the problem, nonemptiness of $\Lambda(x_0, u_0)$ requires the assumption of a certain constraint qualification applied to the constraints and the admissible process. By analogy with the mathematical programming problem, we define normality relative to S by imposing the null solution as the unique solution to the system given in Theorem 10 when the cost multiplier vanishes. This is a natural extension of the definition of normality given for problem (N).

We shall say that $(x_0, u_0) \in S$ is *normal relative to* S if, given $(p, \mu) \in X \times \mathcal{U}_q$ satisfying

i. $\mu_\alpha(t) \geq 0$ and $\mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ t \in T)$;

ii. $\dot p(t) = -A^*(t)p(t)$
$[= -H_x^*(t, x_0(t), u_0(t), p(t), \mu(t), 0)]$ $(t \in T)$;

iii. $0 = B^*(t)p(t) - \varphi'^*(u_0(t))\mu(t)$
$[= H_u^*(t, x_0(t), u_0(t), p(t), \mu(t), 0)]$ $(t \in T)$,

then $p \equiv 0$. Note that, in this event, also $\mu \equiv 0$.

From Theorem 10 we conclude that, if (x_0, u_0) solves (P) and is a normal process relative to S, then $\Lambda(x_0, u_0)$ is nonempty, that is, there exists $(p, \mu) \in X \times \mathcal{U}_q$ such that (x_0, u_0, p, μ) is an extremal.

The sets S_0 and S_1, which played a fundamental role before, have also their counterpart in optimal control. Denote the set of *active indices at* u by

$$I_a(u) := \{\alpha \in R \mid \varphi_\alpha(u) = 0\} \quad (u \in \mathbf{R}^m).$$

Given $u_0 \in \mathcal{U}_m$, let

$$S_0\,[= S_0(u_0)\,] := \{(x, u) \in D \mid \varphi_i(u(t)) = 0$$

$$(i \in I_a(u_0(t)) \cup Q,\ t \in T)\}.$$

For $\mu \in \mathcal{U}_q$ with $\mu_\alpha(t) \geq 0$ $(\alpha \in R,\ t \in T)$, define

$$S_1\,[= S_1(\mu)\,] := \{(x, u) \in D \mid$$

$$\varphi_\alpha(u(t)) \leq 0\ (\alpha \in R,\ \mu_\alpha(t) = 0,\ t \in T),$$

$$\varphi_\beta(u(t)) = 0\ (\beta \in R \text{ with } \mu_\beta(t) > 0,$$

$$\text{or } \beta \in Q,\ t \in T)\}.$$

Note that

$$S_1(\mu) = \{(x, u) \in S \mid$$

$$\varphi_\alpha(u(t)) = 0\ (\alpha \in R,\ \mu_\alpha(t) > 0,\ t \in T)\}.$$

Since normality is defined relative to any set of processes given by inequality and equality constraints, it can be applied to the sets $S_0(u_0)$ and $S_1(\mu)$. By definition, we obtain the following conditions.

An admissible process (x_0, u_0) is normal relative to $S_0(u_0)$ if, given $(p, \mu) \in X \times \mathcal{U}_q$ satisfying

i. $\mu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ t \in T)$;

ii. $\dot p(t) = -A^*(t)p(t)$ $(t \in T)$;

iii. $B^*(t)p(t) = \varphi'^*(u_0(t))\mu(t)$ $(t \in T)$,

then $p \equiv 0$. As before, this implies that $\mu \equiv 0$.

Similarly, an admissible process (x_0, u_0) is normal relative to $S_1(\mu)$ if, given $(q, \nu) \in X \times \mathcal{U}_q$ satisfying

i. $\nu_\alpha(t) \geq 0$ and $\nu_\alpha(t)\varphi_\alpha(u_0(t)) = 0$ $(\alpha \in R,\ \mu_\alpha(t) = 0,\ t \in T)$;

ii. $\dot{q}(t) = -A^*(t)q(t)$ $(t \in T)$;

iii. $B^*(t)q(t) = \varphi'^*(u_0(t))\nu(t)$ $(t \in T)$,

then $q \equiv 0$. In this event, again, we also have $\nu \equiv 0$.

As one readily verifies, if $(x_0, u_0) \in S$ is normal relative to $S_0(u_0)$, then $\Lambda(x_0, u_0)$ is a singleton. This assumption, however, can be weakened, and the same conclusion will follow if, given $(p, \mu) \in \Lambda(x_0, u_0)$, the admissible process is normal relative to $S_1(\mu)$. The converse, however, may not hold. There are problems for which the pair (p, μ) is unique in $\Lambda(x_0, u_0)$ but (x_0, u_0) fails to be normal relative to $S_1(\mu)$. A full account of these results can be found in [18]. For a characterization of the uniqueness of multipliers in terms of normality of yet another set of constraints we refer to [17].

4 Uniqueness and second order conditions

We shall find convenient to express the normality conditions of the previous section in terms of certain convex cones. Given $\mu \in \mathbf{R}^q$ define the following subsets of indices of R:

$$\Gamma_0(\mu) := \{\alpha \in R \mid \mu_\alpha = 0\},$$

$$\Gamma_p(\mu) := \{\alpha \in R \mid \mu_\alpha > 0\}.$$

For all $u \in \mathbf{R}^m$ and $\mu \in \mathbf{R}^q$, consider the following sets

$$\tau_0(u) := \{h \in \mathbf{R}^m \mid \varphi_i'(u)h = 0\ (i \in I_a(u) \cup Q)\},$$

$$\tau_1(u, \mu) := \{h \in \mathbf{R}^m \mid \varphi_i'(u)h \leq 0\ (i \in I_a(u) \cap \Gamma_0(\mu)),\ \varphi_j'(u)h = 0\ (j \in \Gamma_p(\mu) \cup Q)\},$$

$$\tau(u) := \{h \in \mathbf{R}^m \mid \varphi_i'(u)h \leq 0\ (i \in I_a(u)),\ \varphi_j'(u)h = 0\ (j \in Q)\}.$$

As shown in [13], [14], (x_0, u_0) is normal with respect to $S_0(u_0)$ if $z \equiv 0$ is the only solution of the system

$$\dot{z}(t) = -A^*(t)z(t), \quad z^*(t)B(t)h = 0 \quad \text{for all } h \in \tau_0(u_0(t)) \quad (t \in T).$$

Similarly, (x_0, u_0, μ) is normal with respect to $S_1(\mu)$ if $z \equiv 0$ is the only solution of the system

$$\dot{z}(t) = -A^*(t)z(t), \quad z^*(t)B(t)h \leq 0 \quad \text{for all } h \in \tau_1(u_0(t), \mu(t)) \quad (t \in T).$$

Finally, (x_0, u_0) is normal with respect to S if $z \equiv 0$ is the only solution of the system

$$\dot{z}(t) = -A^*(t)z(t), \quad z^*(t)B(t)h \leq 0 \quad \text{for all } h \in \tau(u_0(t)) \quad (t \in T).$$

For second order necessary conditions, we consider the quadratic integral defined, for all $(x, u, p, \mu) \in Z \times X \times \mathcal{U}_q$, by

$$J((x, u, p, \mu); (y, v)) := \int_{t_0}^{t_1} 2\Omega(t, y(t), v(t))dt\ ((y, v) \in Z)$$

where, for all $(t, y, v) \in T \times \mathbf{R}^n \times \mathbf{R}^m$,

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

and $H(t)$ denotes $H(t, x(t), u(t), p(t), \mu(t), 1)$.

The next result is well-known in the literature (see, for example, [12]). It guarantees uniqueness of the Lagrange multiplier and provides second order necessary conditions.

Theorem 11. *Suppose (x_0, u_0) solves (P) and $(p, \mu) \in \Lambda(x_0, u_0)$. If (x_0, u_0) is normal relative to $S_0(u_0)$ then (p, μ) is unique and $J((x_0, u_0, p, \mu); (y, v)) \geq 0$ for all $(y, v) \in Z$ satisfying*

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

ii. $y(t_0) = y(t_1) = 0$;

iii. $v(t) \in \tau_0(u_0(t))$ $(t \in T)$.

This result is, clearly, the counterpart of Theorem 6 which deals with normality with respect to S_0 and guarantees nonnegativity on the set R_{S_0}.

It is worth mentioning that some examples found in the literature (see [13], [14]) show that we may have a solution (x_0, u_0) to the problem which is normal relative to S, a pair (p, μ) such that $(x_0, u_0, p, \mu) \in \mathcal{E}$, and $(y, v) \in Z$ satisfying (i), (ii) and (iii) of this theorem, but $J((x_0, u_0, p, \mu); (y, v)) < 0$. In other words, the conclusion of Theorem 11 may not hold if we assume that the solution to the problem is normal with respect to S.

Those references provide also examples where a solution which is normal relative to $S_1(\mu)$ may yield a negative second variation on $\tau(u_0(t))$. In other words, there exists an admissible process (x_0, u_0) which solves the problem and is normal with respect to $S_1(\mu)$ with $(p, \mu) \in \Lambda(x_0, u_0)$ but, for some $(y, v) \in Z$ satisfying (i), (ii) of Theorem 11 with $v(t) \in \tau(u_0(t))$, we have $J((x_0, u_0, p, \mu); (y, v)) < 0$. It is natural to ask if the same conclusion can be reached assuming only uniqueness of the multiplier.

More generally, it is of interest to know if the second order necessary condition given in Theorem 11 holds in a larger set and/or under weaker assumptions. In particular, we would like to know if the theorem remains valid assuming uniqueness of the multiplier instead of the strong normality assumption on (x_0, u_0). Explicitly, the question is if the following result holds.

Conjecture 1. *Suppose* (x_0, u_0) *solves* (P) *and* $(p, \mu) \in \Lambda(x_0, u_0)$. *If* (p, μ) *is unique, that is, if* $\Lambda(x_0, u_0) = \{(p, \mu)\}$, *then* $J((x_0, u_0, p, \mu); (y, v)) \geq 0$ *for all* $(y, v) \in Z$ *satisfying*

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

ii. $y(t_0) = y(t_1) = 0$;

iii. $v(t) \in \tau(u_0(t))$ $(t \in T)$.

5 The example

In this final section we provide a new and illustrative result in the direction signaled by the conjecture. It corresponds to an example where a solution (x_0, u_0) to the problem has a unique multiplier $(p, \mu) \in \Lambda(x_0, u_0)$ but $J((x_0, u_0, p, \mu); (y, v)) < 0$ for some $(y, v) \in Z$ satisfying

$$\dot{y}(t) = A(t)y(t) + B(t)v(t) \ \ (t \in T),$$

$$y(t_0) = y(t_1) = 0,$$

$$v(t) \in \tau(u_0(t)) \ \ (t \in T).$$

In other words, the conclusion of the conjecture may not be true.

Example 3. Consider the problem of minimizing $I(x, u) = \int_{-1}^{1} b(t)u^3(t)dt$ subject to

$$\dot{x}(t) = u(t) \ \ (t \in [-1, 1]), \quad x(-1) = x(1) = 0,$$

$$u^2(t) \leq 1 \ \ (t \in [-1, 1])$$

where

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

In this case $T = [-1, 1]$, $n = m = 1$, $\xi_0 = \xi_1 = 0$ and, for all $(t, x, u) \in T \times \mathbf{R} \times \mathbf{R}$,

$$L(t, x, u) = b(t)u^3, \ f(t, x, u) = u, \ \varphi(u) = u^2 - 1.$$

Let

$$x_0(t) := \begin{cases} t + 1 & \text{if } t \in [-1, 0] \\ 1 - t & \text{if } t \in [0, 1] \end{cases}$$

$$u_0(t) := \begin{cases} 1 & \text{if } t \in [-1, 0] \\ -1 & \text{if } t \in (0, 1]. \end{cases}$$

Clearly (x_0, u_0) is a solution to the problem.

Now, if $(x_0, u_0, p, \mu) \in \mathcal{E}$, then $\mu(t) \geq 0$, $\dot{p}(t) = 0$,

$$p(t) = 3b(t)u_0^2(t) + 2u_0(t)\mu(t) \quad (t \in [-1, 1]).$$

Thus p is a constant satisfying

$$p = \begin{cases} 2\mu(t) & \text{if } t \in [-1, 0] \\ 3t^2 - 2\mu(t) & \text{if } t \in (0, 1]. \end{cases}$$

Since $\mu(t) \geq 0$ for all $t \in T$, from the first relation we have $p \geq 0$ and, from the second, $p \leq 3t^2$ for all $t \in (0, 1]$ and so $p \leq 0$. Thus $p \equiv 0$ and therefore

$$\mu(t) = \begin{cases} 0 & \text{if } t \in [-1, 0] \\ (3/2)t^2 & \text{if } t \in [0, 1]. \end{cases}$$

This implies that (p, μ) is the only pair such that $(x_0, u_0, p, \mu) \in \mathcal{E}$.

We have $H(t, x, u, p, \mu) = pu - b(t)u^3 - (u^2 - 1)\mu$ and so

$$H_u(t, x, u, p, \mu) = p - 3b(t)u^2 - 2\mu u,$$

$$H_{uu}(t, x, u, p, \mu) = -6b(t)u - 2\mu.$$

Therefore

$$H_{uu}(t, x_0(t), u_0(t), p(t), \mu(t))$$

$$= \begin{cases} 0 & \text{if } t \in [-1, 0] \\ 3t^2 & \text{if } t \in [0, 1]. \end{cases}$$

This implies that, for all $(x, u, p, \mu) \in Z \times X \times \mathcal{U}_1$ and $(y, v) \in Z$,

$$J((x, u, p, \mu); (y, v)) = -\int_0^1 3t^2 v^2(t) dt.$$

Since $\varphi'(u) = 2u$, we have

$$\tau_0(u_0(t)) = \{h \mid 2u_0(t)h = 0\} = \{0\},$$

$\tau_1(u_0(t), \mu(t))$ is given by

$$\{h \mid 2u_0(t)h \leq 0\} = \{h \mid h \leq 0\} \quad \text{if } t \in [-1, 0],$$
$$\{h \mid 2u_0(t)h = 0\} = \{0\} \quad \text{if } t \in (0, 1],$$

and $\tau(u_0(t))$ is given by

$$\{h \mid 2u_0(t)h \leq 0\} = \{h \mid h \leq 0\} \quad \text{if } t \in [-1, 0],$$
$$\{h \mid 2u_0(t)h \leq 0\} = \{h \mid h \geq 0\} \quad \text{if } t \in (0, 1].$$

Let

$$y(t) := \begin{cases} -t - 1 & \text{if } t \in [-1, 0] \\ t - 1 & \text{if } t \in [0, 1] \end{cases}$$
$$v(t) := \begin{cases} -1 & \text{if } t \in [-1, 0] \\ 1 & \text{if } t \in (0, 1]. \end{cases}$$

Then (y, v) solves $\dot{y}(t) = v(t)$ $(t \in T)$, $y(-1) = y(1) = 0$, $v(t) \in \tau(u_0(t))$, and

$$J((x_0, u_0, p, \mu); (y, v)) = -\int_0^1 3t^2 v^2(t) dt < 0. \blacksquare$$

Let us end this paper with an open question, a new conjecture. As mentioned before, it is of interest to know if Theorem 11 holds in a larger set and/or under weaker assumptions. In particular, we would like to know if Conjecture 1 is valid if we replace condition (iii) by

$v(t) \in \tau_1(u_0(t), \mu(t))$ $(t \in T)$.

Explicitly, the question is if the following result holds.

Conjecture 2. *Suppose* (x_0, u_0) *solves* (P) *and* $(p, \mu) \in \Lambda(x_0, u_0)$. *If* (p, μ) *is unique, that is, if* $\Lambda(x_0, u_0) = \{(p, \mu)\}$, *then* $J((x_0, u_0, p, \mu); (y, v)) \geq 0$ *for all* $(y, v) \in Z$ *satisfying*

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

ii. $y(t_0) = y(t_1) = 0$;

iii. $v(t) \in \tau_1(u_0(t), \mu(t))$ $(t \in T)$.

References:

[1] Ben-Tal A (1980) *Second-order and related extremality conditions in nonlinear programming*, Journal of Optimization Theory & Applications, **31**: 143-165

[2] Kyparisis J (1985) *On uniqueness of Kuhn-Tucker multipliers in nonlinear programming*, Mathematical Programming, **32**: 242-246

[3] Wachsmuth G (2013) *On LICQ and the uniqueness of Lagrange multipliers*, Operations Research Letters, **41**: 78-80

[4] Clarke F (2013) *Functional Analysis, Calculus of Variations and Optimal Control*, Springer-Verlag, London

[5] Giorgi G, Guerraggio A, Thierfelder J (2004) *Mathematics of Optimization: Smooth and Nonsmooth Case*, Elsevier, Amsterdam

[6] Hestenes MR (1966) *Calculus of Variations and Optimal Control Theory*, John Wiley, New York

[7] Hestenes MR (1975) *Optimization Theory, The Finite Dimensional Case*, John Wiley, New York

[8] McShane EJ (1973) *The Lagrange Multiplier Rule*, The American Mathematical Monthly, **80**: 922-925

[9] Shapiro A (1997) *On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints*, SIAM Journal of Optimization, **7**: 508-518

[10] Kyparisis J (1990) *Sensitivity analysis for nonlinear programs and variational inequalities with non-unique multipliers*, Mathematics of Operations Research, **15**: 286-298

[11] Deng S (1996) *On uniqueness of Lagrange multipliers in composite optimization*, Journal of Mathematical Analysis & Applications, **201**: 689-696

[12] de Pinho MR, Rosenblueth JF (2007) *Mixed constraints in optimal control: an implicit function theorem approach*, IMA Journal of Mathematical Control and Information, **24:** 197-218

[13] Rosenblueth JF, Sánchez Licea G (2012) *Negative second variations for problems with inequality control constraints*, 3rd European Conference for the Applied Mathematics and Informatics, Montreux, Suiza, WSEAS Press, Mathematical Methods for Information Science and Economics, 150-155

[14] Rosenblueth JF, Sánchez Licea G (2013) *Cones of critical directions in optimal control*, International Journal of Applied Mathematics & Informatics, **7**: 55-67

[15] Cortez KL, Rosenblueth JF (2021) *Normality, controllability and properness in optimal control*, Applied Mathematics & Optimization, **84**: 159-173

[16] Cortez KL, Rosenblueth JF (2019) *The broken link between normality and regularity in the calculus of variations*, Systems & Control Letters, **124**: 27-32

[17] Becerril JA, Cortez KL, Rosenblueth JF (2019) *Uniqueness of multipliers in optimal control: the missing piece*, IMA Journal of Mathematical Control & Information, **36**: 1395-1411

[18] Cortez KL, Rosenblueth JF (2018) *Normality and uniqueness of Lagrange multipliers*, Discrete & Continuous Dynamical Systems Series A, **38**: 3271-3290

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)
The author carried out all the research of this article.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself
No funding was received for conducting this study.

Conflicts of Interest
The author has no conflicts of interest to declare that are relevant to the content of this article.