Journal of Convex Analysis
Volume 27 (2020), No. 4, 1315–1344

A Strict Mangasarian-Fromovitz Type Condition for Isoperimetric Control Problems

Jorge A. Becerril
IIMAS, Universidad Nacional Autónoma de México, México DF 01000
jorgebecerril88@gmail.com

Javier F. Rosenblueth
IIMAS, Universidad Nacional Autónoma de México, México DF 01000
jfrl@unam.mx

Received: June 4, 2019
Accepted: February 12, 2020

The aim of this paper is to improve a well-known result on second order necessary conditions for optimal control problems involving inequality and equality isoperimetric constraints. The improvement, which widens its range of applicability, consists in weakening the classical assumption under which a certain quadratic form is nonnegative along a cone of critical directions. The new, weaker assumption is shown to be equivalent to a strict Mangasarian-Fromovitz type condition which, in turn, implies regularity relative to the contingent cone of a subset of the original set of constraints. An example is given which lies beyond the scope of previous approaches and illustrates the usefulness of our main result.

Keywords: Isoperimetric control problems, second order necessary conditions, normality, regularity.

2010 Mathematics Subject Classification: 49K15.

1. Introduction

It is well-known that a slight tightening of the Mangasarian-Fromovitz constraint qualification in nonlinear programming is a necessary and sufficient condition for the uniqueness of Lagrange multipliers. Moreover, it implies the satisfaction of second order necessary optimality conditions (see, for example, [23, 25, 26, 38]). As explained in detail in [15], the implication on necessity was first proved by Hestenes [25] and, a few years later, by Kyparisis [26].

Thus, implicit in the title of the paper, are the words "uniqueness of multipliers" and "second order necessary conditions." We shall deal in this paper with the latter, in an optimal control context.

We shall be concerned with the derivation of second order necessary conditions for a fixed endpoint optimal control problem involving inequality and equality isoperimetric constraints. A detailed study of the problem we shall deal with can be found in the classical book on calculus of variations and optimal control by Hestenes [24], where second order conditions are derived under a *strong* normality

ISSN 0944-6532 / \$ 2.50

assumption. In this paper we shall show how that assumption can be considerably weakened if it is replaced with, precisely, a strict Mangasarian-Fromovitz type condition.

First order conditions for that problem are well established. In particular, a maximum principle is stated in [24, Theorem 2.1, p 254] and proved in that text by means of a theory of derived sets and derived cones. Since then, many authors have provided different first order necessary conditions, including nonsmooth data in the functions delimiting the problem, and we refer to [11] for an excellent and comprehensive presentation of the subject.

Second order necessary conditions are also derived in [24]. The result proved in [24, Theorem 9.1, p. 285] shows that, if an extremal corresponding to an optimal strongly normal process is given, then a quadratic form will be nonnegative in a class of differentially admissible variations which takes into account the sign of the multipliers associated with the extremal. The arguments used in [24] to prove this result are based on embedding theorems of differential equations as well as, thanks to the strong normality assumption, an application of the implicit function theorem. This yields a one-parameter family of differentially admissible processes which, in turn, implies a certain type of regularity of the original process. The existence of the family of processes, and the property of regularity applicable to the curvilinear tangent space, yield the desired result.

In this paper, the tangent space for which the notion of regularity is applied corresponds not to the curvilinear one (as in the classical approach) but to a space defined in terms of a generalized notion of directionally convergent sequences. The normed space obtained in this manner coincides with the contingent cone as defined in [5, 23]. Regularity is then defined in terms of this cone and, by showing that normality implies this type of regularity, we reach the desired conclusion. A similar procedure can be found in a recent paper [8] where the isoperimetric problem of Lagrange in the calculus of variations is studied.

In the literature, one can find other approaches to establish second order necessary conditions for the isoperimetric control problem. Some deal also with implicit function theorems (see, for example, [16, 22, 28, 39]), obtaining a similar condition under an equivalent assumption of normality. Others treat the maximum of a quadratic form for different types of minima [27, 29, 30], or new notions of conjugacy and critical directions [13, 14, 28, 34–37].

Special attention deserves the research developed in [1–4] where one of the main features is that the necessary conditions obtained make sense, and are derived, without a priori assumptions of normality. A full account of these results and others related to the inverse function theorem can be found in [1]. It is important to mention that the second order conditions derived in those references are expressed in terms of the maximum of a quadratic form, over certain multipliers, for all the elements of the critical cone of the original set of constraints. In our main result, on the other hand, we are able to preserve the set of critical directions given in [24] and, at the same time, weaken the usual assumption of strong normality found in the literature.

Our approach will be direct in nature. We keep the integral constraints explicitly, as a fundamental part of the problem, allowing us to easily compare our result with the well-known theorem proved in [24]. We pose the problem and the corresponding theorem on second order conditions exactly as in [24], and then show by comparison that the main result (without any reduction or transformation) can be notably improved. Also, we have chosen this approach because there are situations where it may be more convenient to deal with isoperimetric problems than problems with side constraints. For example, as mentioned in [24, p 159], the theory of Mayer fields does not apply to an isoperimetric problem without transforming it to a problem of Lagrange, and different methods applicable to cases in which Mayer fields may not exist can certainly provide optimality conditions in a desirable way.

The paper is organized as follows. In Section 2 we pose the problem we shall deal with and state the well-known result on second order necessary conditions found in [24]. In Section 3 we "unfold" this result by explaining in detail some of its main features and show how it can be derived by applying the implicit function theorem and embedding theorems of differential equations. In Section 4 we provide a useful characterization of the notion of strong normality, a notion which was used as a crucial assumption in the well-known result on second order conditions. The characterization obtained, similar to the one given in [7] in a calculus of variations context, shows that the strong assumption actually corresponds to a much simpler notion of normality applied to a set of constraints involving only active indices. In Section 5 we weaken the strong assumption and establish our main result by showing that regularity, defined in terms of the contingent cone, is a consequence of normality. This implication is established by introducing a notion of properness (similar to the constraint qualification of Mangasarian-Fromovitz in nonlinear programming) which, as we show, is equivalent to that of normality. In Section 6 we provide an example illustrating the usefulness of our main result and for which the classical result found in the literature cannot be applied. In the last section, a brief comparison with other works is given, showing some fundamental differences with our approach.

It is worth mentioning that, for that comparison, we show explicitly how to introduce new variables in order to enlarge the system of differential equations and thus expressing the isoperimetric constraints as endpoint equality and inequality constraints. For the works mentioned in Section 7, as we explain in detail, a method applied to Banach spaces does not work for our setting. Moreover, in contrast with works for which the cost multiplier may vanish, the main discussion in our paper is precisely on conditions which imply a positive cost multiplier, a situation analogous to that in a mathematical programming context. Finally, we may find in some of those references results for which, given an admissible variation, there exists a multiplier for which the second variation is nonnegative while, in our case, the second variation is nonnegative for all admissible variations (elements of the critical cone) related to the extremal under consideration.

2. Statement of the problem and the main result

Suppose we are given an interval $T := [t_0, t_1]$ in $\mathbf{R}$, two points ξ_0, ξ_1 in $\mathbf{R}^n$, functions L, L_γ mapping $T \times \mathbf{R}^n \times \mathbf{R}^m$ to $\mathbf{R}$ and f mapping $T \times \mathbf{R}^n \times \mathbf{R}^m$ to $\mathbf{R}^n$, and scalars b_γ in $\mathbf{R}$ ($\gamma = 1, \ldots, q$).

Denote by X the space of piecewise C^1 functions mapping T to $\mathbf{R}^n$, by $\mathcal{U}$ the space of piecewise continuous functions mapping T to $\mathbf{R}^m$, set $Z := X \times \mathcal{U}$, and consider the sets

$$D := \{(x, u) \in Z \mid \dot{x}(t) = f(t, x(t), u(t))\ (t \in T),\ x(t_0) = \xi_0,\ x(t_1) = \xi_1\},$$

$$S := \{(x, u) \in D \mid I_\alpha(x, u) \leq 0\ (\alpha \in R),\ I_\beta(x, u) = 0\ (\beta \in Q)\}$$

where $R = \{1, \ldots, r\}$, $Q = \{r + 1, \ldots, q\}$, and

$$I_\gamma(x, u) = b_\gamma + \int_{t_0}^{t_1} L_\gamma(t, x(t), u(t))dt \quad ((x, u) \in Z,\ \gamma = 1, \ldots, q).$$

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

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

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

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

$$\begin{cases} x\colon T \to \mathbf{R}^n \text{ is piecewise } C^1,\ u\colon T \to \mathbf{R}^m \ \text{ is piecewise continuous;} \\ \dot{x}(t) = f(t, x(t), u(t))\ (t \in T),\ x(t_0) = \xi_0 \ \text{ and } \ x(t_1) = \xi_1; \\ I_\alpha(x, u) \leq 0\ (\alpha \in R) \ \text{ and } \ I_\beta(x, u) = 0\ (\beta \in Q). \end{cases}$$

Our main reference will be the classical text on calculus of variations and optimal control by Hestenes [24]. Implicit in the statement of the problem is a relatively open subset O of $T \times \mathbf{R}^n$ for which the domain of L, L_γ and f is $O \times \mathbf{R}^m$.

Elements of Z will be called *processes*, and a process (x, u) is *admissible* if it belongs to S and $(t, x(t)) \in O$ $(t \in T)$. A process (x, u) *solves* (P) (locally) if it is admissible and (upon shrinking O if necessary) we have $I(x, u) \leq I(y, v)$ for all admissible processes (y, v). Given $(x, u) \in Z$ we shall use the notation $(\tilde{x}(t))$ to represent $(t, x(t), u(t))$, and '$*$' denotes transpose. With respect to the smoothness of the functions delimiting the problem, we assume that L, L_γ and f are C^2.

The following definition of s-normality, which corresponds to the constraint qualification (as in nonlinear programming) on which classical second order necessary conditions for our problem are based, can be found in [24, p 273].

Definition 2.1. An admissible process (x_0, u_0) is said to be *s-normal* if there exist $q + n$ processes $(y_\sigma, v_\sigma) \in Z$ satisfying

$$\dot{y}(t) = f_x(\tilde{x}_0(t))y(t) + f_u(\tilde{x}_0(t))v(t)\ (t \in T), \quad y(t_0) = 0$$

and such that the determinant

$$\Delta := \begin{vmatrix} I'_\gamma(x_0, u_0; y_\sigma, v_\sigma) \\ y^i_\sigma(t_1) \end{vmatrix} \qquad \begin{matrix} (\gamma = 1, \dots, q;\ i = 1, \dots, n) \\ (\sigma = 1, \dots, q+n) \end{matrix}$$

is different from zero, where

$$I'_\gamma(x_0, u_0; y, v) = \int_{t_0}^{t_1} \{L_{\gamma x}(\tilde{x}_0(t))y(t) + L_{\gamma u}(\tilde{x}_0(t))v(t)\}dt \quad ((y, v) \in Z)$$

denotes the first variation of I_γ along (x_0, u_0).

For all $(t, x, u, p, \lambda, \lambda_0)$ in $T \times \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^n \times \mathbf{R}^q \times \mathbf{R}$ let

$$H(t, x, u, p, \lambda, \lambda_0) := \langle p, f(t, x, u)\rangle - \lambda_0 L(t, x, u) - \sum_1^q \lambda_\gamma L_\gamma(t, x, u)$$

and, for any $(x, u, p, \lambda) \in Z \times X \times \mathbf{R}^q$, define

$$J''(x, u; y, v) := \int_{t_0}^{t_1} 2\Omega(t, y(t), v(t))dt \quad ((y, v) \in Z)$$

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

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

and $H[t]$ denotes $H(t, x(t), u(t), p(t), \lambda, 1)$.

The main theorem on second order necessary conditions for a minimum derived in [24] corresponds to the following result (see [24, Theorem 9.1, p. 285]).

Theorem 2.2. *Let $(x_0, u_0) \in S$ and suppose there exists $(p, \lambda) \in X \times \mathbf{R}^q$ such that*

(a) $\lambda_\alpha \geq 0$ *and* $\lambda_\alpha I_\alpha(x_0, u_0) = 0$ $(\alpha \in R)$;

(b) $\dot{p}(t) = -f_x^*(\tilde{x}_0(t))p(t) + L_x^*(\tilde{x}_0(t)) + \sum_1^q \lambda_\gamma L_{\gamma x}^*(\tilde{x}_0(t))$ $(t \in T)$;

(c) $f_u^*(\tilde{x}_0(t))p(t) = L_u^*(\tilde{x}_0(t)) + \sum_1^q \lambda_\gamma L_{\gamma u}^*(\tilde{x}_0(t))$ $(t \in T)$.

If (x_0, u_0) solves (P) *and is s-normal, then $J''(x_0, u_0; y, v) \geq 0$ for all $(y, v) \in Z$ satisfying*

(i) $\dot{y}(t) = f_x(\tilde{x}_0(t))y(t) + f_u(\tilde{x}_0(t))v(t)$ $(t \in T)$ *and* $y(t_0) = y(t_1) = 0$;

(ii) $I'_\alpha(x_0, u_0; y, v) \leq 0$ $(\alpha \in R$ *with* $I_\alpha(x_0, u_0) = 0$ *and* $\lambda_\alpha = 0)$;

(iii) $I'_\beta(x_0, u_0; y, v) = 0$ $(\beta \in R$ *with* $\lambda_\beta > 0$, *or* $\beta \in Q)$.

Our objective is to show that, in this well-known theorem, one can reach exactly the same conclusion on the nonnegativity of the quadratic form, but imposing an assumption on the solution (x_0, u_0) to the problem which may, in general, be much weaker than that of s-normality.

3. Unfolding the main result

Let us "unfold" Theorem 2.2 in several respects. In this explanation we shall actually provide a proof based on the notion of curvilinear tangent cone and a simple application of the implicit function theorem.

Let us begin with the assumption on the admissible process (x_0, u_0) concerning the existence of a couple (p, λ) satisfying 2.2(a), (b) and (c). First order necessary conditions are well established and one version can be stated as follows (see [24, Theorem 2.1, p. 254]).

Theorem 3.1. *If* (x_0, u_0) *solves* (P), *there exist* $\lambda_0 \geq 0$, $\lambda \in \mathbf{R}^q$ *and* $p \in X$, *not vanishing simultaneously on* T, *such that*

(a) $\lambda_\alpha \geq 0$ *and* $\lambda_\alpha I_\alpha(x_0, u_0) = 0$ $(\alpha \in R)$;

(b) $\dot{p}(t) = -H_x^*(\tilde{x}_0(t), p(t), \lambda, \lambda_0)$ *on every interval of continuity of* u_0.

(c) $H_u(\tilde{x}_0(t), p(t), \lambda, \lambda_0) = 0$ $(t \in T)$.

As shown in [24, Theorem 7.2, p 275], if (x_0, u_0) is s-normal in Theorem 3.1, then λ_0 is positive. In this event, the multipliers λ_0, λ, p can be chosen so that $\lambda_0 = 1$ and, when so chosen, they are unique. This property defines a set $\mathcal{E}$ whose elements will be called *extremals* and which plays a crucial role in the theory that follows.

Definition 3.2. Denote by $\mathcal{E}$ the set of all $(x, u, p, \lambda) \in Z \times X \times \mathbf{R}^q$ such that

(a) $\lambda_\alpha \geq 0$ and $\lambda_\alpha I_\alpha(x, u) = 0$ $(\alpha \in R)$;

(b) $\dot{p}(t) = -f_x^*(\tilde{x}(t))p(t) + L_x^*(\tilde{x}(t)) + \sum_1^q \lambda_\gamma L_{\gamma x}^*(\tilde{x}(t))$ $(t \in T)$;

(c) $f_u^*(\tilde{x}(t))p(t) = L_u^*(\tilde{x}(t)) + \sum_1^q \lambda_\gamma L_{\gamma u}^*(\tilde{x}(t))$ $(t \in T)$.

Thus, if (x_0, u_0) is s-normal in Theorem 3.1, there exists a unique pair (p, λ) such that $(x_0, u_0, p, \lambda) \in \mathcal{E}$. The letter '$s$' in this type of normality stands for *strong* and processes with this property will also be called *strongly normal.* This notion can certainly be weakened, as we shall show below.

Note also that, in Theorem 2.2, our first assumption on the admissible process is simply that there exists (p, λ) such that (x_0, u_0, p, λ) is an extremal.

Let us now introduce the notion of curvilinear tangent space. It corresponds to a generalization we propose of the standard notion, for the finite dimensional case, as presented in [24, p 203].

Definition 3.3. Denote by $\mathcal{C}_S(x_0, u_0)$ the *curvilinear tangent space* of S at (x_0, u_0) given by those $(y, v) \in Z$ for which there exist $\delta > 0$ and $(x(\cdot, \epsilon), u(\cdot, \epsilon)) \in S$ $(0 \leq \epsilon < \delta)$ satisfying

(i) $x(t, \epsilon)$ is continuous and has continuous first and second derivatives with respect to ϵ; $\dot{x}(t, \epsilon)$ and their first and second derivatives with respect to ϵ are piecewise continuous with respect to t.

(ii) $(x(t, 0), u(t, 0)) = (x_0(t), u_0(t))$ $(t \in T)$;

(iii) $(x_\epsilon(t,0), u_\epsilon(t,0)) = (y(t), v(t))$ $(t \in T)$;

(iv) $(x(\cdot,\epsilon), u(\cdot,\epsilon)) \to (x_0, u_0)$ and $(x_\epsilon(\cdot,\epsilon), u_\epsilon(\cdot,\epsilon)) \to (y, v)$ uniformly in T as $\epsilon \to 0$.

The following result gives us straightforwardly second order necessary conditions for optimality in terms of the curvilinear tangent space of a subset of the original set of isoperimetric constraints. It should be observed that the function $J''(x_0, u_0; y, v)$ appearing both in the following result and in Theorem 2.2 is actually the second variation of the integral J defined below.

Theorem 3.4. *Let $(x_0, u_0) \in S$ and suppose there exists $(p, \lambda) \in X \times \mathbf{R}^q$ such that $(x_0, u_0, p, \lambda) \in \mathcal{E}$. Let*

$$S_1 := \{(x, u) \in S \mid I_\alpha(x, u) = 0 \quad (\alpha \in R,\ \lambda_\alpha > 0)\}.$$

If (x_0, u_0) solves (P) *then $J''(x_0, u_0; y, v) \geq 0$ for all $(y, v) \in \mathcal{C}_{S_1}(x_0, u_0)$.*

Proof: Let $c := \langle p(t_1), \xi_1\rangle - \langle p(t_0), \xi_0\rangle + \sum_1^q \lambda_\gamma b_\gamma$ and define

$$J(x, u) := c + \int_{t_0}^{t_1} F(t, x(t), u(t))dt \quad ((x, u) \in Z)$$

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

$$F(t, x, u) := L(t, x, u) - \langle p(t), f(t, x, u)\rangle + \sum_1^q \lambda_\gamma L_\gamma(t, x, u) - \langle \dot{p}(t), x\rangle.$$

Observe that $$F(t, x, u) = -H(t, x, u, p(t), \lambda, 1) - \langle \dot{p}(t), x\rangle$$

and, if $(x, u) \in S$, then

$$\begin{aligned} J(x, u) &= I(x, u) + \sum_1^q \lambda_\gamma b_\gamma + \sum_1^q \lambda_\gamma \int_{t_0}^{t_1} L_\gamma(t, x(t), u(t)dt \\ &= I(x, u) + \sum_1^q \lambda_\gamma I_\gamma(x, u). \end{aligned}$$

Note that

$$S_1 = \{(x, u) \in S \mid \lambda_\alpha I_\alpha(x, u) = 0\ (\alpha \in R)\} = \{(x, u) \in S \mid J(x, u) = I(x, u)\}.$$

Since $(x_0, u_0) \in S_1$, (x_0, u_0) minimizes J on S_1.

Now let $(y, v) \in \mathcal{C}_{S_1}(x_0, u_0)$ and let $\delta > 0$ and $(x(\cdot,\epsilon), u(\cdot,\epsilon)) \in S_1$ $(0 \leq \epsilon < \delta)$ satisfy the relations given in Definition 3.3. Then $g(\epsilon) := J(x(\cdot,\epsilon), u(\cdot,\epsilon))$, where $0 \leq \epsilon < \delta$, satisfies

$$g(\epsilon) = I(x(\cdot,\epsilon), u(\cdot,\epsilon)) \geq I(x_0, u_0) = J(x_0, u_0) = g(0) \quad (0 \leq \epsilon < \delta).$$

Note that, for all $t \in T$, $F_x(\tilde{x}_0(t)) = -H_x(\tilde{x}_0(t), p(t), \lambda(t), 1) - \dot{p}^*(t) = 0$ and $F_u(\tilde{x}_0(t)) = -H_u(\tilde{x}_0(t), p(t), \lambda(t), 1) = 0$, and therefore

$$g'(0) = J'(x_0, u_0; y, v) = \int_{t_0}^{t_1} \{F_x(\tilde{x}_0(t))y(t) + F_u(\tilde{x}_0(t))v(t)\}dt = 0$$

implying that $g''(0) \geq 0$. The result then follows since $g''(0) = J''(x_0, u_0; y, v)$. $\square$

The conditions obtained above are expressed in terms of the curvilinear tangent space of S_1 at (x_0, u_0), whose membership is usually difficult to determine. Tangential constraints help to ease this situation and play a fundamental role in the theory of second order conditions.

Definition 3.5. Given $(x_0, u_0) \in S$ let

$$\mathcal{L}(x_0, u_0) := \{(y, v) \in Z \mid \dot{y}(t) = A(t)y(t) + B(t)v(t)\ (t \in T),\ y(t_0) = y(t_1) = 0\}$$

where $A(t) = f_x(\tilde{x}_0(t))$, $B(t) = f_u(\tilde{x}_0(t))$. Denote the set of *active indices at* (x_0, u_0) by

$$I^a(x_0, u_0) := \{\gamma \in R \mid I_\gamma(x_0, u_0) = 0\},$$

and define the set of processes satisfying the *tangential constraints at* (x_0, u_0) *with respect to* S by

$$\mathcal{R}_S(x_0, u_0) := \left\{ (y, v) \in \mathcal{L}(x_0, u_0) \;\middle|\; \begin{array}{ll} I'_\alpha(x_0, u_0; y, v) \leq 0 & (\alpha \in I^a(x_0, u_0)), \\ I'_\beta(x_0, u_0; y, v) = 0 & (\beta \in Q) \end{array} \right\}.$$

Note 3.6. For all $(x_0, u_0) \in S$, $\mathcal{C}_S(x_0, u_0) \subset \mathcal{R}_S(x_0, u_0)$.

Proof: Let $(y, v) \in \mathcal{C}_S(x_0, u_0)$ and let $\delta > 0$ and $(x(\cdot, \epsilon), u(\cdot, \epsilon)) \in S$ be as in Definition 3.3. Then, for all $0 \leq \epsilon < \delta$,

$$\dot{x}(t, \epsilon) = f(t, x(t, \epsilon), u(t, \epsilon))\ (t \in T), \quad x(t_0, \epsilon) = \xi_0,\ x(t_1, \epsilon) = \xi_1$$

and so, by taking partial derivatives with respect to ϵ and evaluating at $\epsilon = 0$, (y, v) belongs to $\mathcal{L}(x_0, u_0)$. Also, for all $\epsilon \in [0, \delta)$,

$$I_\alpha(x(\cdot, \epsilon), u(\cdot, \epsilon)) \leq 0\ (\alpha \in R), \quad I_\beta(x(\cdot, \epsilon), u(\cdot, \epsilon)) = 0\ (\beta \in Q).$$

Fix $\gamma \in R \cup Q$ and set $g(\epsilon) := I_\gamma(x(\cdot, \epsilon), u(\cdot, \epsilon))$ so that $g'(0) = I'_\gamma(x_0, u_0; y, v)$. If $\gamma \in I^a(x_0, u_0)$ then $g'(0) \leq 0$. If $\gamma \in Q$ then $g'(0) = 0$. □

Definition 3.7. We say that the process (x_0, u_0) is *c-regular relative to* S if $\mathcal{C}_S(x_0, u_0) = \mathcal{R}_S(x_0, u_0)$.

The property of c-regularity is desirable for various reasons. In particular note that if, in Theorem 3.4, c-regularity relative to S_1 holds with respect to the solution (x_0, u_0) to the problem, the theorem will provide a second order condition on $\mathcal{R}_{S_1}(x_0, u_0)$, a set for which verifying membership is of a much easier nature than that of $\mathcal{C}_{S_1}(x_0, u_0)$. As we show next, c-regularity is implied precisely by s-normality.

Theorem 3.8. *If* $(x_0, u_0) \in S$ *is s-normal then it is c-regular relative to* S.

Proof: Suppose (x_0, u_0) is s-normal and let $(y, v) \in \mathcal{R}_S(x_0, u_0)$. We want to prove that $(y, v) \in \mathcal{C}_S(x_0, u_0)$. Assume, without loss of generality, that $I^a(x_0, u_0) = R$, that is, all indices are active at (x_0, u_0).

Let $N := q + n$ and let $(y_\sigma, v_\sigma) \in Z$ $(\sigma = 1, \ldots, N)$ satisfy

$$\dot{y}_\sigma(t) = A(t)y_\sigma(t) + B(t)v_\sigma(t)\ (t \in T), \quad y_\sigma(t_0) = 0$$

and such that the determinant

$$\Delta = \begin{vmatrix} I'_\gamma(x_0, u_0; y_\sigma, v_\sigma) \\ y^i_\sigma(t_1) \end{vmatrix} \qquad \begin{matrix} (\gamma = 1, \ldots, q;\ i = 1, \ldots, n) \\ (\sigma = 1, \ldots, q+n) \end{matrix}$$

is different from zero where, as before, $A(t) = f_x(\tilde{x}_0(t))$ and $B(t) = f_u(\tilde{x}_0(t))$.

Define $\quad \bar{u}(t, \epsilon, \alpha) := u_0(t) + \epsilon v(t) + \sum_{i=1}^{N} \alpha_i v_i(t) \quad (t \in T,\ \epsilon, \alpha_i \in \mathbf{R})$

where $\alpha = (\alpha_1, \ldots, \alpha_N)$. By the embedding theorem of differential equations (see, for example, [24, Theorem 14.2, p 49]), the equation

$$\dot{x}(t) = f(t, x(t), \bar{u}(t, \epsilon, \alpha)) \quad (t \in T), \quad x(t_0) = \xi_0$$

has a unique solution $\bar{x}(t, \epsilon, \alpha)$ $(t \in T,\ |\epsilon| < \eta,\ |\alpha_i| < \eta)$ such that $\bar{x}(t, 0, 0) = x_0(t)$. The function $\bar{x}(t, \epsilon, \alpha)$ is continuous and has continuous first and second derivatives with respect to the variables $\epsilon, \alpha_1, \ldots, \alpha_N$. The functions $\dot{\bar{x}}(t, \epsilon, \alpha)$ and their first and second derivatives with respect to $\epsilon, \alpha_1, \ldots, \alpha_N$ are piecewise continuous with respect to t. By differentiation with respect to ϵ and α_i at $(\epsilon, \alpha) = (0, 0)$ it is found that

$$\dot{\bar{x}}_\epsilon(t, 0, 0) = A(t)\bar{x}_\epsilon(t, 0, 0) + B(t)v(t), \quad \bar{x}_\epsilon(t_0, 0, 0) = 0,$$
$$\dot{\bar{x}}_{\alpha_i}(t, 0, 0) = A(t)\bar{x}_{\alpha_i}(t, 0, 0) + B(t)v_i(t), \quad \bar{x}_{\alpha_i}(t_0, 0, 0) = 0$$

and therefore $\quad \bar{x}_\epsilon(t, 0, 0) = y(t), \quad \bar{x}_{\alpha_i}(t, 0, 0) = y_i(t) \quad (t \in T)$.

Let $G := (-\eta, \eta)$ and define $g\colon G \times G^N \to \mathbf{R}^N$, $g = (g^1, \ldots, g^N)$, by

$$g^\gamma(\epsilon, \alpha) := I_\gamma(\bar{x}(\cdot, \epsilon, \alpha), \bar{u}(\cdot, \epsilon, \alpha)) - \epsilon I'_\gamma(x_0, u_0; y, v) \quad (\gamma = 1, \ldots, q),$$
$$g^{q+i}(\epsilon, \alpha) := \bar{x}^i(t_1, \epsilon, \alpha) - \xi^i_1 \quad (i = 1, \ldots, n).$$

Note that $g(0, 0) = 0$ and $|g_\alpha(0, 0)| = \Delta \neq 0$. By the implicit function theorem, there exist $0 < \delta < \eta$ and $\beta\colon (-\delta, \delta) \to \mathbf{R}^N$ of class C^2 such that $\beta(0) = 0$ and $g(\epsilon, \beta(\epsilon)) = 0$ $(|\epsilon| < \delta)$. We obtain, taking the derivative with respect to ϵ at $\epsilon = 0$,

$$0 = g_\epsilon(0, 0) + g_\alpha(0, 0)\beta'(0) = g_\alpha(0, 0)\beta'(0)$$

implying that $\beta'(0) = 0$. Let us show that the one-parameter family

$$x(t, \epsilon) := \bar{x}(t, \epsilon, \beta(\epsilon)), \quad u(t, \epsilon) := \bar{u}(t, \epsilon, \beta(\epsilon)) \quad (t \in T,\ |\epsilon| < \delta)$$

has the properties, with respect to (y, v), given in the definition of $\mathcal{C}_S(x_0, u_0)$. Clearly 3.3(i) and (ii) hold, as well as 3.3(iii) since

$$x_\epsilon(t, 0) = \bar{x}_\alpha(t, 0, 0)\beta'(0) + y(t) = y(t), \quad u_\epsilon(t, 0) = \bar{u}_\alpha(t, 0, 0)\beta'(0) + v(t) = v(t).$$

Moreover, the family belongs to S for $0 \leq \epsilon < \delta$ since

$$I_\gamma(x(\cdot, \epsilon), u(\cdot, \epsilon)) = \epsilon I'_\gamma(x_0, u_0; y, v) \quad (\gamma = 1, \ldots, q)$$
$$x^i(t_1, \epsilon) - \xi^i_1 = \bar{x}^i(t_1, \epsilon, \beta(t_1, \epsilon)) - \xi^i_1 = 0 \quad (i = 1, \ldots, n).$$

It remains to establish the convergence properties given in 3.3(iv), that is, we must show $(x(\cdot,\epsilon),u(\cdot,\epsilon))\to(x_0,u_0)$ and $(x_\epsilon(\cdot,\epsilon),u_\epsilon(\cdot,\epsilon))\to(y,v)$ uniformly in T as $\epsilon\to 0$. To do so, take first an arbitrary $\eta>0$ and a nonnegative constant M such that $|v(t)|\le M$ and $|v_\sigma(t)|\le M$ for each σ. Then we have

$$|u(t,\epsilon)-u_0(t)|=\left|\epsilon v(t)+\sum_\sigma \beta_\sigma(\epsilon)v_\sigma(t)\right|\le M\left[\epsilon+\left|\sum_\sigma \beta_\sigma(\epsilon)\right|\right]$$

for all $t\in T$ and, since $\beta_\sigma(\epsilon)\to 0$ as $\epsilon\to 0$, we can find $0<\epsilon_0\le\min\{\eta/2M,\delta\}$ satisfying $\left|\sum_\sigma \beta_\sigma(\epsilon)\right|<\frac{\eta}{2M}$, which implies

$$|u(t,\epsilon)-u_0(t)|<M\left[\frac{\eta}{2M}+\frac{\eta}{2M}\right]=\eta\quad\text{for all } (t,\epsilon)\in T\times[0,\epsilon_0].$$

Thus $u(\cdot,\epsilon)\to u_0$ uniformly in T as $\epsilon\to 0$. Since

$$|u_\epsilon(t,\epsilon)-v(t)|=\left|\sum_\sigma \beta'_\sigma(\epsilon)v_\sigma(t)\right|,$$

and $\beta'_\sigma(\epsilon)\to 0$ as $\epsilon\to 0$, the uniform convergence in T as $\epsilon\to 0$ of $u_\epsilon(\cdot,\epsilon)$ to v can be proved in a similar way.

Now, if we define the function F as $F(t,x,\epsilon):=f(t,x,u(t,\epsilon))$, then F is continuous in sets of the form $S_i:=[t_i,t_{i+1}]\times\mathrm{cl}\,\mathrm{U}\times[0,\delta]$, where $\{t_i\}$ is a finite partition of T such that the functions u_0, v and v_σ are all continuous in each $T_i:=[t_i,t_{i+1}]$, and U is a neighborhood of $x_0(T)$. Clearly, F is also Lipschitz continuous uniformly in each S_i in the variable x, because $f\in C^2$ also satisfies this property. Therefore, we can invoke [12, Theorem 7.4, p 29] to guarantee the uniqueness of the solutions of

$$\dot x(t)=F(t,x(t),0),\ x(t_i)=x_0(t_i)\quad(t\in S_i)$$

for each S_i. This implies that x_0 is the only solution of the equation

$$\dot x(t)=F(t,x(t),0),\ x(t_0)=\xi_0\quad(t\in T).$$

It is also clear that F satisfies the following Carathéodory type hypotheses:

(C1) $F(\cdot,x,\epsilon)$ is measurable in t for every $(x,\epsilon)\in\mathrm{cl}\,\mathrm{U}\times[0,\delta]$;

(C2) $F(t,\cdot,\epsilon)$ is continuous in x for every $(t,\epsilon)\in T\times[0,\delta]$;

(C3) there exists a Lebesgue integrable function m over T such that, for fixed $t\in T$, $|f(t,x,u)|\le m(t)$ for all $(x,u)\in\mathrm{cl}\,\mathrm{U}\times[0,\delta]$.

The uniform convergence $x(t,\epsilon)\to x_0$ in T as $\epsilon\to 0$ follows from [12, Theorem 4.2, p. 59]. To prove that the last convergence is also uniform in T as $\epsilon\to 0$, define the function K by

$$K(t,z,\epsilon):=f_x(t,x(t,\epsilon),u(t,\epsilon))z+f_u(t,x(t,\epsilon),u(t,\epsilon))u_\epsilon(t,\epsilon).$$

Note that $x_\epsilon(\cdot,\epsilon)$ is a solution of the equation $\dot z(t) = K(t,z(t),\epsilon)$, $z(t_0) = 0$, and y is a solution of the equation $\dot z(t) = K(t,z(t),0)$, $z(t_0) = 0$. Finally, since the remarks on F that made possible the use of [12, Theorem 4.2, p. 59] can also be made with respect to K, the desired convergence follows and the proof is complete. □

Recall that the set S_1, defined in the statement of Theorem 3.4, depends on a certain vector λ in $\mathbf{R}^q$ with $\lambda_\alpha \geq 0$ ($\alpha \in R$) given beforehand. It corresponds to

$$S_1\ [= S_1(\lambda)\,] = \{(x,u) \in S \mid I_\alpha(x,u) = 0\ (\alpha \in \Gamma)\}$$

where $\Gamma := \{\alpha \in R \mid \lambda_\alpha > 0\}$. Clearly we have

$$S_1(\lambda) = \{(x,u) \in D \mid I_\alpha(x,u) \leq 0\ (\alpha \in R,\ \lambda_\alpha = 0),\ I_\beta(x,u) = 0\ (\beta \in \Gamma \cup Q)\}.$$

Thus, the set of processes satisfying the tangential constraints at (x_0,u_0) with respect to $S_1(\lambda)$ is given by

$$\mathcal{R}_{S_1}(x_0,u_0) = \left\{ (y,v)\in\mathcal{L}(x_0,u_0) \;\middle|\; \begin{array}{l} I'_\alpha(x_0,u_0;y,v) \leq 0\ (\alpha\in I^a(x_0,u_0),\ \lambda_\alpha = 0), \\ I'_\beta(x_0,u_0;y,v) = 0\ (\beta \in \Gamma \cup Q) \end{array} \right\}.$$

On the other hand, observe that s-normality is defined in terms of the first variation of I_γ along (x_0,u_0) for all $\gamma = 1,\ldots,q$, that is, it does not take into account if the indices correspond to inequality or equality constraints. Thus, this notion will not vary if it is expressed in terms of S or S_1. By Theorem 3.8 we conclude that, if (x_0,u_0) is s-normal, then it is c-regular relative not only to S but also to S_1.

A combination of these remarks with Theorem 3.4 shows that the following result holds. It corresponds precisely to Theorem 2.2 rewritten in a succinct way.

Theorem 3.9. *Let $(x_0,u_0) \in S$ and suppose that there exists $(p,\lambda) \in X \times \mathbf{R}^q$ such that $(x_0,u_0,p,\lambda) \in \mathcal{E}$. If (x_0,u_0) solves* (P) *and is strongly normal, then we have $J''(x_0,u_0;y,v) \geq 0$ for all $(y,v) \in \mathcal{R}_{S_1}(x_0,u_0)$.*

4. A characterization of normality

Theorem 3.1 provides first order necessary conditions which yield in a natural way a notion of *normality relative to S* which implies the positivity of λ_0, the cost multiplier.

Definition 4.1. We shall say that $(x_0,u_0) \in S$ is *normal relative to S* if, given $(p,\lambda) \in X \times \mathbf{R}^q$ satisfying

(i) $\lambda_\alpha \geq 0$ and $\lambda_\alpha I_\alpha(x_0,u_0) = 0$ ($\alpha \in R$);

(ii) $\dot p(t) = -f_x^*(\tilde x_0(t))p(t) + \sum_1^q \lambda_\gamma L_{\gamma x}^*(\tilde x_0(t))$ $[= -H_x^*(\tilde x_0(t),p(t),\lambda,0)\,]$ $(t\in T)$;

(iii) $0 = f_u^*(\tilde x_0(t))p(t) - \sum_1^q \lambda_\gamma L_{\gamma u}^*(\tilde x_0(t))$ $[= H_u^*(\tilde x_0(t),p(t),\lambda,0)\,]$ $(t \in T)$,

then $(p,\lambda) \equiv (0,0)$.

Clearly, if (x_0, u_0) solves (P) and is normal relative to S, then $\lambda_0 > 0$ in Theorem 3.1. Dividing by this multiplier it follows that $(x_0, u_0, p, \lambda) \in \mathcal{E}$ for some $(p, \lambda) \in X \times \mathbf{R}^q$. Note that, due to the sign of the multipliers in 4.1(i), we cannot guarantee uniqueness of (p, λ). If we remove in 4.1(i) the condition that $\lambda_\alpha \geq 0$ $(\alpha \in R)$ then, as we shall see below, (p, λ) will be the unique pair satisfying $(x_0, u_0, p, \lambda) \in \mathcal{E}$.

By introducing a subset of S and applying the definition of normality relative to that subset, we obtain the desired condition. Consider the subset of S defined only by equalities for active indices, that is, given $(x_0, u_0) \in S$, let

$$S_0\ [= S_0(x_0, u_0)\,] = \{(x, u) \in D \mid I_\gamma(x, u) = 0\ (\gamma \in I^a(x_0, u_0) \cup Q)\}.$$

Note 4.2. Applying Definition 4.1 to this subset of S we deduce that $(x_0, u_0) \in S$ is normal relative to S_0 if, given $(p, \lambda) \in X \times \mathbf{R}^q$ satisfying

(i) $\lambda_\alpha I_\alpha(x_0, u_0) = 0$ $(\alpha \in R)$;

(ii) $\dot{p}(t) = -f_x^*(\tilde{x}_0(t))p(t) + \sum_1^q \lambda_\gamma L_{\gamma x}^*(\tilde{x}_0(t))$ $(t \in T)$;

(iii) $0 = f_u^*(\tilde{x}_0(t))p(t) - \sum_1^q \lambda_\gamma L_{\gamma u}^*(\tilde{x}_0(t))$ $(t \in T)$,

then $(p, \lambda) \equiv (0, 0)$.

As mentioned before, unique couples (p, λ) producing extremals (x_0, u_0, p, λ) appear under the assumption of normality relative to S_0.

Proposition 4.3. *If (x_0, u_0) solves* (P) *and is normal relative to S_0 then there exists a unique pair $(p, \lambda) \in X \times \mathbf{R}^q$ such that $(x_0, u_0, p, \lambda) \in \mathcal{E}$.*

Proof: By Theorem 3.1, $\exists \lambda_0 \geq 0$, $\lambda \in \mathbf{R}^q$ and $p \in X$, not vanishing simultaneously on T, such that

(a) $\lambda_\alpha \geq 0$ and $\lambda_\alpha I_\alpha(x_0, u_0) = 0$ $(\alpha \in R)$;

(b) $\dot{p}(t) = -f_x^*(\tilde{x}_0(t))p(t) + \lambda_0 L_x^*(\tilde{x}_0(t)) + \sum_1^q \lambda_\gamma L_{\gamma x}^*(\tilde{x}_0(t))$ $(t \in T)$.

(c) $f_u^*(\tilde{x}_0(t))p(t) = \lambda_0 L_u^*(\tilde{x}_0(t)) + \sum_1^q \lambda_\gamma L_{\gamma u}^*(\tilde{x}_0(t))$ $(t \in T)$.

By normality relative to S_0 (and hence relative to S), $\lambda_0 > 0$ and we can assume, without loss of generality, that $\lambda_0 = 1$, so that $(x_0, u_0, p, \lambda) \in \mathcal{E}$. Suppose now that $(q, \mu) \in X \times \mathbf{R}^q$ is such that $(x_0, u_0, q, \mu) \in \mathcal{E}$. Then

(i) $[\lambda_\alpha - \mu_\alpha] I_\alpha(x_0, u_0) = 0$ $(\alpha \in R)$;

(ii) $[\dot{p}(t) - \dot{q}(t)] = -f_x^*(\tilde{x}_0(t))[p(t) - q(t)] + \sum_1^q [\lambda_\gamma - \mu_\gamma] L_{\gamma x}^*(\tilde{x}_0(t))$ $(t \in T)$;

(iii) $f_u^*(\tilde{x}_0(t))[p(t) - q(t)] = \sum_1^q [\lambda_\gamma - \mu_\gamma] L_{\gamma u}^*(\tilde{x}_0(t))$ $(t \in T)$,

implying that $p \equiv q$ and $\lambda \equiv \mu$. □

Let us point out that the notion of normality relative to S_0 plays a fundamental role in the theory of optimality conditions not only due to this uniqueness property (see also [6, 9, 23, 38] for an analogous situation in mathematical programming) but also because, as we shall prove in this section, it turns out to be equivalent precisely to the notion of s-normality introduced before.

To prove this result we shall first establish a different characterization of strong normality. Let $(x_0, u_0) \in S$ and set $A(t) := f_x(\tilde{x}_0(t))$, $B(t) := f_u(\tilde{x}_0(t))$. Choose solutions of the differential systems

$$\dot{q}_\beta(t) + A^*(t)q_\beta(t) = L^*_{\beta x}(\tilde{x}_0(t)), \quad q_\beta(t_1) = 0 \quad (\beta = 1, \ldots, q),$$

$$\dot{W}(t) + W(t)A(t) = 0, \quad W(t_1) = I$$

that is,
$$\dot{q}_{\beta j}(t) + \sum_i q_{\beta i}(t)A^i_j(t) = L_{\beta x^j}(\tilde{x}_0(t)), \; q_{\beta j}(t_1) = 0 \; (j = 1, \ldots, n),$$

$$\dot{W}_{ij}(t) + \sum_h W_{ih}A^h_j(t) = 0, \quad W_{ij}(t_1) = \delta_{ij} \quad (i, j = 1, \ldots, n).$$

Define
$$F_\beta(t, x, u) = L_\beta(t, x, u) - \langle q_\beta(t), f(t, x, u)\rangle - \langle \dot{q}_\beta(t), x\rangle \quad (\beta = 1, \ldots, q)$$

$$F_{q+i}(t, x, u) = W_i(t)f(t, x, u) + \dot{W}_i(t)x = \sum_j W_{ij}f^j + \sum_j \dot{W}_{ij}x^j \; (i = 1, \ldots, n)$$

where $W_i = (W_{i1}, \ldots, W_{in})$ denotes the i-th row of W, and let

$$J_\beta(x, u) := b_\beta - \langle q_\beta(t_0), \xi_0\rangle + \int_{t_0}^{t_1} F_\beta(t, x(t), u(t))dt \quad (\beta = 1, \ldots, q),$$

$$J_{q+1}(x, u) := W_i(t_0)\xi_0 - \xi^i_1 + \int_{t_0}^{t_1} F_{q+i}(t, x(t), u(t))dt \quad (i = 1, \ldots, n).$$

Note that F_ρ has been constructed so that

$$\frac{\partial F_\rho}{\partial x} = 0 \quad \text{along } (x_0, u_0).$$

Also, if $(x, u) \in Z$ satisfies $\dot{x}(t) = f(t, x(t), u(t))$ $(t \in T)$, we have

$$J_\beta(x, u) = I_\beta(x, u) + \langle q_\beta(t_0), x(t_0) - \xi_0\rangle \quad (\beta = 1, \ldots, q),$$

$$J_{q+i}(x, u) = x^i(t_1) - \xi^i_1 - W_i(t_0)(x(t_0) - \xi_0) \quad (i = 1, \ldots, n)$$

and so, if also $x(t_0) = \xi_0$, then

$$J_\beta(x, u) = I_\beta(x, u) \; (\beta = 1, \ldots, q) \;\text{ and }\; J_{q+i}(x, u) = x^i(t_1) - \xi^i_1 \; (i = 1, \ldots, n).$$

Without loss of generality, $I_\alpha(x_0, u_0) = 0$ $(\alpha \in R)$. Thus $J_\rho(x_0, u_0) = 0$ where $\rho = 1, \ldots, q + n)$. Let $v \in \mathcal{U}_m$ and define

$$u(t, \epsilon) := u_0(t) + \epsilon v(t) \quad (t \in T).$$

By the embedding theorem of differential equations, the equations

$$\dot{x}(t) = f(t, x(t), u(t, \epsilon)) \; (t \in T), \quad x(t_0) = \xi_0 \tag{1}$$

have a one-parameter family of solutions $x(\cdot, \epsilon)$ for $|\epsilon| \le \epsilon_0$ with $x(t, 0) = x_0(t)$. The functions $x^i(t, \epsilon)$ are continuous and have continuous derivatives with respect to ϵ. The process

$$y(t) = x_\epsilon(t, 0), \quad v(t) = u_\epsilon(t, 0) \quad (t \in T)$$

is called the *variation of the family* $(x(\cdot,\epsilon),u(\cdot,\epsilon))$ *along* (x_0,u_0). It satisfies

$$\dot y(t) = A(t)y(t) + B(t)v(t), \quad y(t_0) = 0 \tag{2}$$

where $$A^i_j(t) = \frac{\partial f^i}{\partial x^j}, \quad B^i_k(t) = \frac{\partial f^i}{\partial u^k} \quad \text{at } (t, x_0(t), u_0(t)).$$

Equation (2) is the *equation of variation of* (1) *along* (x_0,u_0).

The first variations of I_β, J_ρ along (x_0,u_0) are given by

$$\begin{aligned} I'_\beta(x_0,u_0;y,v) &= \frac{d}{d\epsilon} I_\beta(x(\cdot,\epsilon),u(\cdot,\epsilon))\big|_{\epsilon=0} \\ &= \int_{t_0}^{t_1} \{L_{\beta x}(\tilde x_0(t))y(t) + L_{\beta u}(\tilde x_0(t))v(t)\}dt \quad (\beta = 1,\dots,q) \end{aligned}$$

$$\begin{aligned} J'_\rho(x_0,u_0;y,v) &= \frac{d}{d\epsilon} J_\rho(x(\cdot,\epsilon),u(\cdot,\epsilon))\big|_{\epsilon=0} \\ &= \int_{t_0}^{t_1} \sum_k F_{\rho k}(t) v^k(t)dt = \int_{t_0}^{t_1} \frac{\partial F_\rho}{\partial u}(\tilde x_0(t))v(t)dt \ (\rho = 1,\dots,q+n) \end{aligned}$$

where $$F_{\rho k}(t) = \frac{\partial F_\rho}{\partial u^k}(t, x_0(t), u_0(t)) \quad (t \in T).$$

Note that

$$F_{\beta k}(t) = \frac{\partial L_\beta}{\partial u^k} - \sum_j q_{\beta j}(t) B^j_k(t) = \frac{\partial L_\beta}{\partial u^k} - q^*_\beta(t)\frac{\partial f}{\partial u^k} \quad (\beta = 1,\dots,q),$$

$$F_{\delta k}(t) = \sum_j W_{ij}(t)B^j_k(t) = W_i(t)\frac{\partial f}{\partial u_k} \quad (\delta = q+i,\ i = 1,\dots,n).$$

Therefore $$\frac{\partial F_\beta}{\partial u} = L_{\beta u} - q^*_\beta(t)B(t) \quad (\beta = 1,\dots,q),$$

$$\frac{\partial F_\delta}{\partial u} = W_i(t)B(t) \quad (\delta = q+i,\ i = 1,\dots,n).$$

By (1) we know that, for $|\epsilon| \le \epsilon_0$,

$$J_\beta(x(\cdot,\epsilon),u(\cdot,\epsilon)) = I_\beta(x(\cdot,\epsilon),u(\cdot,\epsilon)) \quad (\beta = 1,\dots,q),$$

and $$J_{q+i}(x(\cdot,\epsilon),u(\cdot,\epsilon)) = x^i(t_1,\epsilon) - \xi^i_1 \quad (i = 1,\dots,n).$$

It follows that

$$J'_\beta(x_0,u_0;y,v) = I'_\beta(x_0,u_0;y,v) \quad (\beta = 1,\dots,q),$$

and $$J'_{q+i}(x_0,u_0;y,v) = y^i(t_1) \quad (i = 1,\dots,n)$$

subject to conditions (2).

The definition of strongly normal corresponds to the existence of $q+n$ solutions (y_σ, v_σ) of

$$\dot y(t) = A(t)y(t) + B(t)v(t), \quad y(t_0) = 0$$

such that the determinant

$$\Delta = \begin{vmatrix} I'_\gamma(x_0, u_0; y_\sigma, v_\sigma) \\ y^i_\sigma(t_1) \end{vmatrix} \qquad \begin{matrix} (\gamma = 1, \dots, q;\ i = 1, \dots, n) \\ (\sigma = 1, \dots, q+n) \end{matrix}$$

is different from zero. Note that

$$\Delta = \left| J'_\rho((x_0, u_0); (y_\sigma, v_\sigma)) \right| = \left| \int_{t_0}^{t_1} \frac{\partial F_\rho}{\partial u}(\tilde x_0(t)) v_\sigma(t) dt \right|.$$

In this determinant, v_σ can be chosen arbitrarily, while the corresponding variations y_σ are determined by (2). In particular, we may choose

$$v_\sigma(t) := \frac{\partial F^*_\sigma}{\partial u}(\tilde x_0(t)) \quad (\sigma = 1, \dots, q+n)$$

and so (x_0, u_0) is s-normal if and only if the $q+n$ functions

$$v_\gamma(t) = L^*_{\gamma u}(\tilde x_0(t)) - B^*(t) q_\gamma(t) \quad (\gamma = 1, \dots, q)$$

$$v_{q+i}(t) = B^*(t) W^*_i(t) \quad (i = 1, \dots, n)$$

are linearly independent on T (see also [24, p 266] where some of these arguments are used to prove the maximum principle given in [24]).

We are now in position to prove the equivalence between s-normality and normality relative to S_0.

Theorem 4.4. *(x_0, u_0) is strongly normal $\Leftrightarrow$ (x_0, u_0) is normal relative to S_0.*

Proof: Assume, without loss of generality, that all indices are active, that is, $I^a(x_0, u_0) = R$.

'$\Rightarrow$': Suppose $(p, \lambda) \in X \times \mathbf{R}^q$ satisfies

(i) $\dot p(t) = -f^*_x(\tilde x_0(t)) p(t) + \sum_1^q \lambda_\gamma L^*_{\gamma x}(\tilde x_0(t))$ $(t \in T)$;

(ii) $0 = f^*_u(\tilde x_0(t)) p(t) - \sum_1^q \lambda_\gamma L^*_{\gamma u}(\tilde x_0(t))$ $(t \in T)$.

We want to prove that $(p, \lambda) \equiv (0, 0)$. Let $\lambda_{q+i} := -p_i(t_1)$ $(i = 1, \dots, n)$ and define the function

$$z(t) := \sum_{\gamma=1}^q \lambda_\gamma q_\gamma(t) - \sum_{i=1}^n \lambda_{q+i} W^*_i(t)$$

where q_γ and W_i are as before. Note that

$$\begin{aligned} \dot z(t) &= \sum_1^q \lambda_\gamma \dot q_\gamma(t) - \sum_1^n \lambda_{q+i} \dot W^*_i(t) \\ &= \sum_1^q \lambda_\gamma [-A^*(t) q_\gamma(t) + L^*_{\gamma x}(\tilde x_0(t))] + \sum_1^n \lambda_{q+i} A^*(t) W^*_i(t) \end{aligned}$$

$$= -A^*(t)z(t) + \sum_1^q \lambda_\gamma L^*_{\gamma x}(\tilde{x}_0(t))$$

implying that, if $w := p - z$, then $\dot{w}(t) = -A^*(t)w(t)$ $(t \in T)$. Since $z_i(t_1) = -\lambda_{q+i} = p_i(t_1)$, we have $w(t_1) = p(t_1) - z(t_1) = 0$ and, therefore, $w \equiv 0$. This implies that

$$p(t) = \sum_{\gamma=1}^q \lambda_\gamma q_\gamma(t) - \sum_{i=1}^n \lambda_{q+i} W_i^*(t).$$

From this equality and (ii) above, we have for all $t \in T$

$$\sum_1^q \lambda_\gamma L^*_{\gamma u}(\tilde{x}_0(t)) = B^*(t)p(t) = \sum_1^q \lambda_\gamma B^*(t)q_\gamma(t) - \sum_i^n \lambda_{q+i} B^*(t)W_i^*(t),$$

i.e. $$\sum_1^{q+n} \lambda_\sigma v_\sigma(t) = \sum_1^q \lambda_\gamma \big(L^*_{\gamma u}(\tilde{x}_0(t)) - B^*(t)q_\gamma(t)\big) + \sum_1^n \lambda_{q+i} B^*(t)W_i^*(t) = 0.$$

Since (x_0, u_0) is strongly normal, $\lambda_\sigma = 0$ for all $\sigma = 1, \ldots, q+n$ and so we get $(p, \lambda) \equiv (0, 0)$.

'$\Leftarrow$': Suppose now that (x_0, u_0) is normal relative to S_0 and, for some λ_σ in $\mathbf{R}$ $(\sigma = 1, \ldots, q+n)$, $\sum_1^{q+n} \lambda_\sigma v_\sigma(t) = 0$ $(t \in T)$. Define

$$p(t) := \sum_{\gamma=1}^q \lambda_\gamma q_\gamma(t) - \sum_{i=1}^n \lambda_{q+i} W_i^*(t) \quad (t \in T)$$

and note that, as before, $\dot{p}(t) = -A^*(t)p(t) + \sum_1^q \lambda_\gamma L^*_{\gamma x}(\tilde{x}_0(t))$. Moreover,

$$B^*(t)p(t) = \sum_1^q \lambda_\gamma B^*(t)q_\gamma(t) - \sum_1^n \lambda_{q+i} B^*(t)W_i^*(t) = \sum_1^q \lambda_\gamma L^*_{\gamma u}(\tilde{x}_0(t)).$$

Hence $p \equiv 0$ and $\lambda_\gamma = 0$ $(\gamma = 1, \ldots, q)$. But this also implies that we have $\lambda_{q+i} = -p_i(t_1) = 0$ $(i = 1, \ldots, n)$. □

5. Normality implies regularity

The notion of normality relative to S can, of course, be applied also to the set $S_1 = S_1(\lambda)$ for a given $\lambda \in \mathbf{R}^q$. Recall that, if $\Gamma = \{\alpha \in R \mid \lambda_\alpha > 0\}$, then

$$S_1(\lambda) = \{(x, u) \in D \mid I_\alpha(x, u) \le 0 \ (\alpha \in R,\ \lambda_\alpha = 0),\ I_\beta(x, u) = 0 \ (\beta \in \Gamma \cup Q)\}.$$

A process (x_0, u_0) is normal relative to $S_1(\lambda)$ if, given $(p, \mu) \in X \times \mathbf{R}^q$ satisfying

(i) $\mu_\alpha \ge 0$ and $\mu_\alpha I_\alpha(x_0, u_0) = 0$ $(\alpha \in R,\ \lambda_\alpha = 0)$;

(ii) $\dot{p}(t) = -f_x^*(\tilde{x}_0(t))p(t) + \sum_1^q \lambda_\gamma L^*_{\gamma x}(\tilde{x}_0(t))$ $(t \in T)$;

(iii) $0 = f_u^*(\tilde{x}_0(t))p(t) - \sum_1^q \lambda_\gamma L^*_{\gamma u}(\tilde{x}_0(t))$ $(t \in T)$,

then $(p, \mu) \equiv (0, 0)$.

As mentioned in [8], a fundamental question posed in the literature (see [13–16, 28, 34–37] for other problems in optimal control) is if, in Theorem 3.9 (or Theorem 2.2), the assumption of strong normality can be replaced with that of normality relative to S_1 without altering the set $\mathcal{R}_{S_1}(x_0, u_0)$ of critical directions where the second order conditions hold. In other words, the question is if the following result is valid or not.

Conjecture 5.1. *Let $(x_0, u_0) \in S$ and suppose there exists $(p, \lambda) \in X \times \mathbf{R}^q$ such that $(x_0, u_0, p, \lambda) \in \mathcal{E}$. If (x_0, u_0) solves* (P) *and is normal relative to $S_1(\lambda)$, then $J''(x_0, u_0; y, v) \geq 0$ for all $(y, v) \in \mathcal{R}_{S_1}(x_0, u_0)$.*

Before attempting to solve this conjecture, let us first introduce a notion of tangency in terms not of curvilinear but of sequential tangents. It corresponds to a generalization of the notion given by Hestenes [24, 25] for finite dimensional spaces (see also [5, 23] for equivalent definitions).

Definition 5.2. For any $z = (x, u) \in Z$, consider the norm

$$\|z\| := \|(x, u)\| = \sup_{t \in T}\big(|x(t)| + |\dot{x}(t)|\big) + \sup_{t \in T} |u(t)|.$$

We shall say that a sequence $\{z_q = (x_q, u_q)\} \subset Z$ *converges to* $z_0 = (x_0, u_0)$ *in the direction* $w = (y, v)$ if w is a unit process ($\|w\| = 1$), $z_q \neq z_0$, and

$$\lim_{q \to \infty} \|z_q - z_0\| = 0, \qquad \lim_{q \to \infty} \frac{z_q - z_0}{\|z_q - z_0\|} = w.$$

Given $z_0 \in S \subset Z$, the *tangent* or *contingent cone of S at z_0*, denoted by $\mathcal{T}_S(z_0)$, is the cone determined by the unit vectors w for which there exists a sequence $\{z_q\}$ in S converging to z_0 in the direction w. Equivalently, $\mathcal{T}_S(z_0)$ is the set of all $w \in Z$ for which there exist a sequence $\{z_q\}$ in S and a sequence $\{\epsilon_q\}$ of positive numbers such that

$$\lim_{q \to \infty} \epsilon_q = 0, \qquad \lim_{q \to \infty} \frac{z_q - z_0}{\epsilon_q} = w. \tag{3}$$

For our original isoperimetric control problem (P), where

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

the first and second variations of I are given by

$$I'(x, u; y, v) := \int_{t_0}^{t_1} \{L_x(\tilde{x}(t))y(t) + L_u(\tilde{x}(t))v(t)\}dt$$

$$I''(x, u; y, v) := \int_{t_0}^{t_1} \{\langle y(t), L_{xx}(\tilde{x}(t))y(t)\rangle + 2\langle y(t), L_{xu}(\tilde{x}(t))v(t)\rangle + \langle v(t), L_{uu}(\tilde{x}(t))v(t)\rangle\}dt.$$

A proof similar to that of [24, p 74] shows that

$$I(z) = I(z_0) + I'(z_0; z - z_0) + r_1(z_0; z - z_0)$$

$$I(z) = I(z_0) + I'(z_0; z - z_0) + \tfrac{1}{2}I''(z_0; z - z_0) + r_2(z_0; z - z_0)$$

where $$\lim_{z \to z_0} \frac{r_1(z_0; z - z_0)}{\|z - z_0\|} = 0, \qquad \lim_{z \to z_0} \frac{r_2(z_0; z - z_0)}{\|z - z_0\|^2} = 0.$$

Therefore, if $\{z_q\}$ converges to z_0 in the direction w, then

$$\lim_{q \to \infty} \frac{I(z_q) - I(z_0)}{\|z_q - z_0\|} = I'(z_0; w), \quad \text{and}$$

$$\lim_{q\to\infty} \frac{I(z_q) - I(z_0) - I'(z_0; z_q - z_0)}{\|z_q - z_0\|^2} = \tfrac{1}{2} I''(z_0; w).$$

From these relations we obtain the following second order conditions for optimality. Compare this result with Theorem 3.4.

Theorem 5.3. *Let* $(x_0, u_0) \in S$ *and suppose there exists* $(p, \lambda) \in X \times \mathbf{R}^q$ *such that* $(x_0, u_0, p, \lambda) \in \mathcal{E}$. *If* (x_0, u_0) *solves* (P) *then* $J''(x_0, u_0; y, v) \geq 0$ *for all* $(y, v) \in \mathcal{T}_{S_1}(x_0, u_0)$.

Proof: As in the proof of Theorem 3.4, let $c := \langle p(t_1), \xi_1\rangle - \langle p(t_0), \xi_0\rangle + \sum_1^q \lambda_\gamma b_\gamma$ and

$$J(x, u) := c + \int_{t_0}^{t_1} F(t, x(t), u(t))dt \quad ((x, u) \in Z)$$

where $F(t, x, u) = -H(t, x, u, p(t), \lambda, 1) - \langle \dot p(t), x\rangle$. We have

$$S_1 = \{(x, u) \in S \mid J(x, u) = I(x, u)\}$$

and so (x_0, u_0) minimizes J on S_1. If $(y, v) \in \mathcal{T}_{S_1}(x_0, u_0)$, since $J'(x_0, u_0; y, v) = 0$, we have $J''(x_0, u_0; y, v) \geq 0$. □

Now, as one readily verifies, $\mathcal{T}_S(x_0, u_0) \subset \mathcal{R}_S(x_0, u_0)$ where

$$\mathcal{R}_S(x_0, u_0) = \left\{ (y, v) \in \mathcal{L}(x_0, u_0) \;\middle|\; \begin{array}{l} I'_\alpha(x_0, u_0; y, v) \leq 0 \;\; (\alpha \in I^a(x_0, u_0)), \\ I'_\beta(x_0, u_0; y, v) = 0 \;\; (\beta \in Q) \end{array} \right\}$$

denotes, as before, the set of processes satisfying the tangential constraints at (x_0, u_0) with respect to S. We shall say that a process $(x_0, u_0) \in S$ is *regular relative to* S if $\mathcal{T}_S(x_0, u_0) = \mathcal{R}_S(x_0, u_0)$. Thus, if we assume in Theorem 5.3 that the optimal process (x_0, u_0) is regular relative to S_1, then $J''(x_0, u_0; y, v) \geq 0$ for all $(y, v) \in \mathcal{R}_{S_1}(x_0, u_0)$.

This notion of regularity plays a fundamental role in the theory to follow. To begin with, it implies nonemptiness of $\mathcal{E}$. To prove it, let us first state an auxiliary result on linear functionals (see [24], p 13).

Lemma 5.4. *Let* $L, L_1, \ldots, L_m$ *be linear forms on* X*, a real vector space, and let*

$$\mathcal{R} = \{x \in X \mid L_\alpha(x) \leq 0 \; (\alpha \in A),\; L_\beta(x) = 0 \;\; (\beta \in B)\}$$

where $A = \{1, \ldots, p\}$ *and* $B = \{p+1, \ldots, m\}$. *Suppose that* $L(x) \geq 0$ *for all* $x \in \mathcal{R}$. *Then there exist* $\lambda_\alpha \geq 0$ $(\alpha \in A)$ *and* $\lambda_\beta \in \mathbf{R}$ $(\beta \in B)$ *such that* $L(x) + \sum_1^m \lambda_i L_i(x) = 0$ $(x \in X)$.

Theorem 5.5. *Suppose* (x_0, u_0) *solves* (P). *If* (x_0, u_0) *is regular relative to* S*, then* $\mathcal{E}$ *is nonempty.*

Proof: Take $(y, v) \in \mathcal{T}_S(x_0, u_0) = \mathcal{R}_S(x_0, u_0)$ and $\{(x_q, u_q)\} \in S$ and $\{\epsilon_q\}$ satisfying (3). Since (x_0, u_0) is a local minimum of (P),

$$0 \leq \lim_{q\to\infty} \frac{I(x_q, u_q) - I(x_0, u_0)}{\epsilon_q} = I'(x_0, u_0; y, v).$$

Observe that

$$\mathcal{R}_S(x_0,u_0) = \left\{ (y,v) \in \mathcal{L}_0(x_0,u_0) \;\middle|\; \begin{array}{ll} F_\alpha(y,v) \le 0 & (\alpha \in I^a(x_0,u_0)), \\ F_\beta(y,v) = 0 & (\beta = r+1,\dots,q+n) \end{array} \right\},$$

where $\mathcal{L}_0(x_0,u_0)$ is the real vector space defined by

$$\mathcal{L}_0(x_0,u_0) = \{(y,v) \in Z \mid \dot y(t) = A(t)y(t) + B(t)v(t),\ y(t_0) = 0\},$$

and the functionals $F_\gamma\colon \mathcal{L}_0(x_0,u_0) \to \mathbf{R}$ are defined by

$$F_\gamma(y,v) = I'_\gamma(x_0,u_0;y,v) \quad \text{and} \quad F_{q+i}(y,v) = y^i(t_1)$$

for $\gamma \in I^a(x_0,u_0) \cup Q$ and $i = 1,\dots,n$. By Lemma 5.4, there exist multipliers $\lambda_\alpha \ge 0$ $(\alpha \in R)$ and $\lambda_\beta \in \mathbf{R}$ $(\beta = r+1,\dots,q+n)$ such that we have $\lambda_\alpha = 0$ if $\alpha \in R \sim I^a(x_0,u_0)$ and

$$I'(x_0,u_0;y,v) + \sum_{\gamma=1}^{q+n} \lambda_\gamma F_\gamma(y,v) = 0 \quad \text{for all } (y,v) \in \mathcal{L}_0(x_0,u_0).$$

Define p as the only solution of the differential equation

$$\dot p(t) = -A^*(t)p(t) + L_x^*(\tilde x_0(t)) + \sum_{\gamma=1}^{q} \lambda_\gamma L_{\gamma x}^*(\tilde x_0(t))$$

and endpoint condition $p^i(t_1) = -\lambda_{q+i}$ $(i = 1,\dots,n)$. Setting $\lambda = (\lambda_1,\dots,\lambda_q)^*$, we see at once that (p,λ) satisfies 3.2(a) and 3.2(b). Only 3.2(c) is left to verify. On one hand, from equation (1), we get

$$\begin{aligned}\sum_{i=1}^{q} p^i(t_1)y^i(t_1) = \int_{t_0}^{t_1} \Bigg\{ \Bigg(L_x(\tilde x_0(t)) + \sum_{\gamma=1}^{q} \lambda_\gamma L_{\gamma x}(\tilde x_0(t)) \Bigg) y(t) \\ + \Bigg(L_u(\tilde x_0(t)) + \sum_{\gamma=1}^{q} \lambda_\gamma L_{\gamma u}(\tilde x_0(t)) \Bigg) v(t) \Bigg\} dt,\end{aligned}$$

while on the other hand, from equation (2),

$$\begin{aligned}\sum_{i=1}^{q} p^i(t_1)y^i(t_1) &= \int_{t_0}^{t_1} \{\dot p^*(t)y(t) + p^*(t)\dot y(t)\}\, dt \\ &= \int_{t_0}^{t_1} \Bigg\{ \Bigg(-p^*(t)A(t) + L_x(\tilde x_0(t)) + \sum_{\gamma=1}^{q} \lambda_\gamma L_{\gamma x}(\tilde x_0(t)) \Bigg) y(t) \\ &\qquad + p^*(t)[A(t)y(t) + B(t)v(t)] \Bigg\} dt \\ &= \int_{t_0}^{t_1} \Bigg\{ \Bigg(L_x(\tilde x_0(t)) + \sum_{\gamma=1}^{q} \lambda_\gamma L_{\gamma x}(\tilde x_0(t)) \Bigg) y(t) + p^*(t)B(t)v(t) \Bigg\} dt.\end{aligned}$$

These two equations imply that for all $(y,v) \in \mathcal{L}_0(x_0,u_0)$

$$\int_{t_0}^{t_1} \Bigg(L_u(\tilde x_0(t)) + \sum_{\gamma=1}^{q} \lambda_\gamma L_{\gamma u}(\tilde x_0(t)) - p^*(t)B(t) \Bigg) v(t)dt = 0,$$

but for every $v \in \mathcal{U}$, there exists a function $y \in X$ such that $(y, v) \in \mathcal{L}_0(x_0, u_0)$. Hence, the above equation is true for all $v \in \mathcal{U}$, implying the pointwise equality

$$B^*(t)p(t) = L_u^*(\tilde{x}_0(t)) + \sum_{\gamma=1}^{q} \lambda_\gamma L_{\gamma u}^*(\tilde{x}_0(t)) \ (t \in [t_0, t_1]),$$

which is precisely 3.2(c). □

Note at this point that, if we prove that normality relative to S implies regularity relative to S, then the conjecture posed above will be established. Let us set out to prove this statement.

Recall that $(x_0, u_0) \in S$ is *normal relative to* S if, given $(p, \lambda) \in X \times \mathbf{R}^q$ satisfying

(i) $\lambda_\alpha \geq 0$ and $\lambda_\alpha I_\alpha(x_0, u_0) = 0$ $(\alpha \in R)$;

(ii) $\dot{p}(t) = -f_x^*(\tilde{x}_0(t))p(t) + \sum_1^q \lambda_\gamma L_{\gamma x}^*(\tilde{x}_0(t))$ $(t \in T)$;

(iii) $0 = f_u^*(\tilde{x}_0(t))p(t) - \sum_1^q \lambda_\gamma L_{\gamma u}^*(\tilde{x}_0(t))$ $(t \in T)$,

then $(p, \lambda) \equiv (0, 0)$.

Let us introduce a concept which, as we shall see below, characterizes normality relative to S. Note that this notion, when applied to the constraints defining the set S, corresponds to a Mangasarian-Fromovitz type condition. On the other hand, if it is applied to the constraints defining $S_1(\lambda)$, it becomes a strict Mangasarian-Fromovitz type condition.

Definition 5.6. We call $(x_0, u_0) \in S$ *proper relative to* S if

(a) (x_0, u_0) is normal relative to $S_e := \{(x, u) \in D \mid I_\beta(x, u) = 0 \ (\beta \in Q)\}$;

(b) there exists $(y, v) \in \mathcal{L}(x_0, u_0)$ such that $I'_\alpha(x_0, u_0; y, v) < 0$ $(\alpha \in I^a(x_0, u_0))$ and $I'_\beta(x_0, u_0; y, v) = 0$ $(\beta \in Q)$.

To prove that properness and normality are equivalent we shall make use, as before, of a well-known auxiliary result on linear functionals (see [25, p 201]).

Lemma 5.7. *Let $L_1, \ldots, L_m$ be linear forms on X, a real vector space, and let*

$$\mathcal{R} = \{x \in X \mid L_\alpha(x) \leq 0 \ (\alpha \in A), \ L_\beta(x) = 0 \ (\beta \in B)\}$$

where $A = \{1, \ldots, p\}$ and $B = \{p+1, \ldots, m\}$. Suppose that $L_i(x) = 0$ for all $x \in \mathcal{R}$ and $i \in A \cup B$. Then there exist $\lambda_\alpha > 0$ $(\alpha \in A)$ and $\lambda_\beta \in \mathbf{R}$ $(\beta \in B)$ such that $\sum_1^m \lambda_i L_i(x) = 0$ $(x \in X)$.

Theorem 5.8. *(x_0, u_0) is normal relative to S $\Leftrightarrow$ (x_0, u_0) is proper relative to S.*

Proof: Without loss of generality assume that all constraints are active at (x_0, u_0), that is, $R = I^a(x_0, u_0)$.

'$\Leftarrow$': Suppose (x_0, u_0) is proper relative to S. Let (p, λ) satisfy (i)–(iii) above. If $R = \emptyset$, then $S = S_e$ and the result follows. If $R \neq \emptyset$, choose a process (y, v) satisfying 5.6(b). Since

$$\dot{p}^*(t) + p^*(t)A(t) = \sum_1^q \lambda_\gamma L_{\gamma x}(\tilde{x}_0(t)) \quad \text{and} \quad p^*(t)B(t) = \sum_1^q \lambda_\gamma L_{\gamma u}(\tilde{x}_0(t)),$$

we have $$\frac{d}{dt}\langle p(t), y(t)\rangle = \sum_1^q \lambda_\gamma [L_{\gamma x}(\tilde{x}_0(t))y(t) + L_{\gamma u}(\tilde{x}_0(t))v(t)]$$

and therefore $$0 = \sum_1^q \lambda_\gamma I'_\gamma(x_0, u_0; y, v) = \sum_1^r \lambda_\gamma I'_\gamma(x_0, u_0; y, v).$$

But $I'_\gamma(x_0, u_0; y, v) < 0$ $(\gamma \in R)$ and so $\lambda_\gamma = 0$ $(\gamma \in R)$. The result then follows by 5.6(a).

'$\Rightarrow$': Suppose (x_0, u_0) is normal relative to S. Clearly, 5.6(a) holds. Without loss of generality, $R \neq \emptyset$. Define

$$C := \{\alpha \in R \mid I'_\alpha(x_0, u_0; y, v) < 0 \text{ for some } (y, v) \in \mathcal{R}_S(x_0, u_0)\}$$

and let $E := R \sim C = \{\alpha \in R \mid \alpha \notin C\}$. Note that

$$I'_\gamma(x_0, u_0; y, v) = 0 \quad \text{for all } \gamma \in E \cup Q \text{ and } (y, v) \in \mathcal{R}_S(x_0, u_0).$$

Consider now the following two sets involving isoperimetric constraints:

$$V := \{(x, u) \in D \mid I_\alpha(x, u) \leq 0 \ (\alpha \in E),\ I_\beta(x, u) = 0 \ (\beta \in Q)\},$$
$$V_0 := \{(x, u) \in D \mid I_\gamma(x, u) = 0 \ (\gamma \in E \cup Q)\}$$

as well as their corresponding sets of tangential constraints at (x_0, u_0):

$$\mathcal{R}_V(x_0, u_0) = \left\{ (y, v) \in \mathcal{L}(x_0, u_0) \;\middle|\; \begin{array}{l} I'_\alpha(x_0, u_0; y, v) \leq 0 \ (\alpha \in E), \\ I'_\beta(x_0, u_0; y, v) = 0 \ (\beta \in Q) \end{array} \right\}$$

$$\mathcal{R}_{V_0}(x_0, u_0) = \{(y, v) \in \mathcal{L}(x_0, u_0) \mid I'_\gamma(x_0, u_0; y, v) = 0 \ (\gamma \in E \cup Q)\}.$$

Let us prove that $\mathcal{R}_V(x_0, u_0) = \mathcal{R}_{V_0}(x_0, u_0)$. To do so, consider the following subset of S:

$$\tilde{S} = \{(x, u) \in D \mid I_\alpha(x, u) \leq 0 \ (\alpha \in C),\ I_\gamma(x, u) = 0 \ (\gamma \in E \cup Q)\}.$$

and the tangential constraints at (x_0, u_0) associated with $\tilde{S}$:

$$\mathcal{R}_{\tilde{S}}(x_0, u_0) = \left\{ (y, v) \in \mathcal{L}(x_0, u_0) \;\middle|\; \begin{array}{l} I'_\alpha(x_0, u_0; y, v) \leq 0 \ (\alpha \in C), \\ I'_\gamma(x_0, u_0; y, v) = 0 \ (\gamma \in E \cup Q) \end{array} \right\}.$$

As one readily verifies, $\mathcal{R}_S(x_0, u_0) = \mathcal{R}_{\tilde{S}}(x_0, u_0)$. We can assume, without loss of generality, that $C \neq \emptyset$ for, otherwise,

$$\mathcal{R}_{\tilde{S}}(x_0, u_0) = \mathcal{R}_{V_0}(x_0, u_0) \quad \text{and} \quad \mathcal{R}_S(x_0, u_0) = \mathcal{R}_V(x_0, u_0).$$

For $\alpha \in C$, choose a process $(y_\alpha, v_\alpha) \in \mathcal{R}_S(x_0, u_0)$ for which $I'_\alpha(x_0, u_0; y_\alpha, v_\alpha) < 0$, and define $(\hat{y}, \hat{v}) := \sum_{\alpha \in C}(y_\alpha, v_\alpha)$. Observe that

$$I'_\alpha(x_0, u_0; \hat{y}, \hat{v}) < 0 \ (\alpha \in C) \quad \text{and} \quad I'_\gamma(x_0, u_0; \hat{y}, \hat{v}) = 0 \ (\gamma \in E \cup Q).$$

Let $(y, v) \not\equiv (0, 0)$ in $\mathcal{R}_V(x_0, u_0)$, and let $\epsilon > 0$ be such that

$$I'_\alpha(x_0, u_0; y + \epsilon\hat{y}, v + \epsilon\hat{v}) = I'_\alpha(x_0, u_0; y, v) + \epsilon I'_\alpha(x_0, u_0; \hat{y}, \hat{v}) \leq 0 \quad (\alpha \in C).$$

Clearly, we have

$$I'_\alpha(x_0,u_0;y+\epsilon\hat{y},v+\epsilon\hat{v}) = I'_\alpha(x_0,u_0;y,v) \le 0 \quad (\alpha\in E),$$

and
$$I'_\beta(x_0,u_0;y+\epsilon\hat{y},v+\epsilon\hat{v}) = 0 \quad (\beta\in Q)$$

and therefore $(y+\epsilon\hat{y},v+\epsilon\hat{v}) \in \mathcal{R}_S(x_0,u_0) = \mathcal{R}_{\tilde{S}}(x_0,u_0)$. Hence

$$I'_\gamma(x_0,u_0;y+\epsilon\hat{y},v+\epsilon\hat{v}) = I'_\gamma(x_0,u_0;y,v) = 0 \quad (\gamma\in E\cup Q)$$

and so $(y,v)\in\mathcal{R}_{V_0}(x_0,u_0)$.

Now, an application of Lemma 5.7 yields the existence of $\lambda_\alpha > 0$ ($\alpha\in E$) and $\lambda_\beta\in\mathbf{R}$ ($\beta\in Q$) such that, for all $(y,v)\in\mathcal{L}(x_0,u_0)$,

$$\sum_{\alpha\in E}\lambda_\alpha I'_\alpha(x_0,u_0;y,v) + \sum_{\beta\in Q}\lambda_\beta I'_\beta(x_0,u_0;y,v) = 0.$$

Note that, in view of our normality assumption, we have $E=\emptyset$. As before, we can then find $(y,v)\in\mathcal{L}(x_0,u_0)$ satisfying 5.6(b) by selecting, for each $\alpha\in R$, a process $(y_\alpha,v_\alpha)\in\mathcal{R}_S(x_0,u_0)$ such that $I'_\alpha(x_0,u_0;y_\alpha,v_\alpha)<0$, and setting $(y,v) := \sum_{\alpha\in R}(y_\alpha,v_\alpha)$. □

An important property of the tangent cone, which we shall use in the next result, is that it is closed. To prove it, suppose $(y,v)\not\equiv(0,0)$, $\{(y_q,v_q)\}\subset\mathcal{T}_S(x_0,u_0)$ and that (y_q,v_q) converges to (y,v). Since $(y_q,v_q)/\|(y_q,v_q)\|\to(y,v)/\|(y,v)\|$, we may assume without loss of generality that $\|(y,v)\| = \|(y_q,v_q)\| = 1$. Let $(x_q,u_q)\in S$ satisfy

$$0<\epsilon_q := \|(x_q,u_q)-(x_0,u_0)\| < \frac{1}{q} \quad\text{and}\quad \left\|\frac{(x_q,u_q)-(x_0,u_0)}{\epsilon_q} - (y_q,v_q)\right\| < \frac{1}{q}.$$

Note that

$$\begin{aligned}\left\|\frac{(x_q,u_q)-(x_0,u_0)}{\epsilon_q} - (y,v)\right\| &\le \left\|\frac{(x_q,u_q)-(x_0,u_0)}{\epsilon_q} - (y_q,v_q)\right\| + \|(y_q,v_q)-(y,v)\| \\ &< \frac{1}{q} + \|(y_q,v_q)-(y,v)\|.\end{aligned}$$

Thus (3) holds and so $(y,v)\in\mathcal{T}_S(x_0,u_0)$. This proves the claim.

Theorem 5.9. *If (x_0,u_0) is proper relative to S, then (x_0,u_0) is regular relative to S.*

Proof: Assume, without loss of generality, that $R = I^a(x_0,u_0)$. By 5.6(b), $R=C$, that is, $E=\emptyset$. By Theorem 3.8 and the characterization of normality given in the previous section, 3.5(a) implies that (x_0,u_0) is c-regular relative to S_e. Since, as one readily verifies, $\mathcal{C}_S(x_0,u_0)\subset\mathcal{T}_S(x_0,u_0)$, this implies that (x_0,u_0) is regular relative to $S_e = V_0 = \{(x,u)\in D \mid I_\beta(x,u)=0\ (\beta\in Q)\}$, that is,

$$\mathcal{T}_{V_0}(x_0,u_0) = \mathcal{R}_{V_0}(x_0,u_0) = \{(y,v)\in\mathcal{L}(x_0,u_0) \mid I'_\beta(x_0,u_0;y,v)=0\ (\beta\in Q)\}.$$

If $R=\emptyset$, then $V_0=S$ and the result follows. If $R\ne\emptyset$, consider the set

$$K_S(x_0,u_0) := \left\{(y,v)\in\mathcal{L}(x_0,u_0) \;\middle|\; \begin{array}{l} I'_\alpha(x_0,u_0;y,v)<0\ (\alpha\in R), \\ I'_\beta(x_0,u_0;y,v)=0\ (\beta\in Q)\end{array}\right\}.$$

As we show next, $K_S(x_0,u_0)\subset\mathcal{T}_S(x_0,u_0)$. Indeed, if $(y,v)\in K_S(x_0,u_0)$, then, since $(y,v)\in\mathcal{R}_{V_0}(x_0,u_0)=\mathcal{T}_{V_0}(x_0,u_0)$, there exist sequences $\{(x_q,u_q)\}$ in V_0 and $\{\epsilon_q>0\}$ such that (3) holds. Therefore

$$\lim_{q\to\infty}\frac{I_\gamma(x_q,u_q)}{\epsilon_q}=\lim_{q\to\infty}\frac{I_\gamma(x_q,u_q)-I_\gamma(x_0,u_0)}{\epsilon_q}=I'_\gamma(x_0,u_0;y,v)\quad(\gamma=1,\dots,q).$$

Since $I'_\alpha(x_0,u_0;y,v)<0$ $(\alpha\in R)$, there exists N such that, for all $q\geq N$, $I_\alpha(x_q,u_q)<0$ $(\alpha\in R)$. Since $\{x_q,u_q\}\subset V_0$, we have $(x_q,u_q)\in S$ for all $q\geq N$. Hence, $(y,v)\in\mathcal{T}_S(x_0,u_0)$. To end the proof, let $(\hat y,\hat v)\in\mathcal{R}_S(x_0,u_0)$ and let (y,v) satisfy 5.6(b) so that $(y,v)\in K_S(x_0,u_0)$. Then

$$(\hat y+\epsilon y,\hat v+\epsilon v)\in K_S(x_0,u_0)\subset\mathcal{T}_S(x_0,u_0)\quad(\epsilon>0)$$

and, since $\mathcal{T}_S(x_0,u_0)$ is closed, $(\hat y,\hat v)\in\mathcal{T}_S(x_0,u_0)$. □

We have proved that normality implies regularity. In particular this implies, as explained before, that Conjecture 5.1 is a theorem. We have reached our main result.

Theorem 5.10. *Let $(x_0,u_0)\in S$ and suppose there exists $(p,\lambda)\in X\times\mathbf{R}^q$ such that $(x_0,u_0,p,\lambda)\in\mathcal{E}$. If (x_0,u_0) solves* (P) *and is normal relative to $S_1(\lambda)$, then $J''(x_0,u_0;y,v)\geq 0$ for all $(y,v)\in\mathcal{R}_{S_1}(x_0,u_0)$.*

Proof: By Theorem 5.3, $J''(x_0,u_0;y,v)\geq 0$ for all $(y,v)\in\mathcal{T}_{S_1}(x_0,u_0)$. By Theorem 5.8, (x_0,u_0) is proper relative to S_1. By Theorem 5.9, (x_0,u_0) is regular relative to S_1 and the result follows. □

6. Example

In this section we provide a simple example for which the classical theory cannot be applied since, for the extremal (x_0,u_0,p,λ) under consideration, (x_0,u_0) is not strongly normal. In Theorem 2.2 (or Theorem 3.9) this assumption is basic. The extremal, however, is such that (x_0,u_0) is normal relative to $S_1(\lambda)$ and so, by applying Theorem 5.10, we can conclude that (x_0,u_0) is not a solution to the problem.

Example 6.1. Consider the isoperimetric control problem (P) of minimizing

$$I(x,u)=\frac{1}{2}\int_0^1\{x_1^2(t)+x_2^2(t)-u_1^2(t)-u_2^2(t)\}dt$$

subject to

$$\begin{cases}\dot x_i(t)=u_i(t)(x_i(t)+1)\quad(t\in T),\quad x_i(0)=x_i(1)=0\quad(i=1,2),\\ \displaystyle\int_0^1\{x_1(t)+x_2(t)\}dt\leq 0\quad\text{and}\quad\int_0^1\{x_1(t)+x_2(t)+u_1^3(t)+u_2^3(t)\}dt\leq 0.\end{cases}$$

In this case, $T=[0,1]$, $n=m=r=2$, $\xi_0=\xi_1=0$, $b_1=b_2=0$,

$$f_i(t,x,u)=u_i(x_i+1)\quad(i=1,2),\quad L(t,x,u)=(x_1^2+x_2^2-u_1^2-u_2^2)/2,$$

$$L_1(t,x,u)=x_1+x_2,\quad L_2(t,x,u)=x_1+x_2+u_1^3+u_2^3.$$

We want to see if the process $(x_0, u_0) \equiv (0, 0)$, which is clearly admissible, is a solution to the problem.

In the analysis of this example we shall prove the following claims:

(A) (x_0, u_0) is normal relative to S but not normal relative to S_0 (or strongly normal).

(B) (x_0, u_0, p, λ) with $(p, \lambda) \equiv (0, 0)$ is an extremal.

(C) With $\lambda = 0$, (x_0, u_0) is normal relative to $S_1(\lambda)$.

(D) $\exists (y, v) \in \mathcal{R}_{S_1}(x_0, u_0)$ such that $J''(x_0, u_0; y, v) < 0$.

Let us begin with the normality assumptions. By Definition 4.1, (x_0, u_0) is normal relative to the set $S = \{(x, u) \in D \mid I_\alpha(x, u) \le 0 \ (\alpha = 1, 2)\}$ if $(p, \lambda) \equiv (0, 0)$ is the only solution of the system

(i) $\lambda_\alpha \ge 0$ and $\lambda_\alpha I_\alpha(x_0, u_0) = 0$ $(\alpha = 1, 2)$;

(ii) $\dot{p}(t) = -f_x^*(\tilde{x}_0(t))p(t) + \sum_1^2 \lambda_\gamma L_{\gamma x}^*(\tilde{x}_0(t))$ $(t \in T)$;

(iii) $0 = f_u^*(\tilde{x}_0(t))p(t) - \sum_1^2 \lambda_\gamma L_{\gamma u}^*(\tilde{x}_0(t))$ $(t \in T)$.

Since $I_\alpha(x_0, u_0) = 0$ $(\alpha = 1, 2)$, condition (i) simply says that $\lambda_1, \lambda_2 \ge 0$. For condition (ii), note that

$$f_x(t, x, u) = \begin{pmatrix} u_1 & 0 \\ 0 & u_2 \end{pmatrix}, \quad L_{1x}(t, x, u) = L_{2x}(t, x, u) = (1, 1)$$

and therefore $\begin{pmatrix} \dot{p}_1(t) \\ \dot{p}_2(t) \end{pmatrix} = \begin{pmatrix} -p_1(t)u_{01}(t) + \lambda_1 + \lambda_2 \\ -p_2(t)u_{02}(t) + \lambda_1 + \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 + \lambda_2 \\ \lambda_1 + \lambda_2 \end{pmatrix}.$

Similarly,

$$f_u(t, x, u) = \begin{pmatrix} x_1 + 1 & 0 \\ 0 & x_2 + 1 \end{pmatrix}, \quad L_{1u}(t, x, u) = (0, 0), \quad L_{2u}(t, x, u) = (3u_1^2, 3u_2^2)$$

and so (iii) corresponds to

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} p_1(t)(x_{01}(t) + 1) - 3\lambda_2 u_{01}^2(t) \\ p_2(t)(x_{02}(t) + 1) - 3\lambda_2 u_{02}^2(t) \end{pmatrix} = \begin{pmatrix} p_1(t) \\ p_2(t) \end{pmatrix}.$$

Since, by (iii), $p \equiv 0$, we have by (ii) that $\lambda_1 + \lambda_2 = 0$ and, by (i), $\lambda_1 = \lambda_2 = 0$. We conclude that (x_0, u_0) is normal relative to S. However, it is not normal relative to the set

$$S_0 = \{(x, u) \in D \mid I_\alpha(x, u) = 0 \ (\alpha = 1, 2)\}$$

since clearly, in Note 4.2, we can let $(\lambda_1, \lambda_2) = (1, -1)$ and $(p, \lambda) \equiv (0, \lambda) \not\equiv 0$ satisfies conditions 4.2(i)–(iii). This proves (A).

Observe now that the set $\mathcal{E}$ of extremals is given by those (x, u, p, λ) satisfying

(a) $\lambda_\alpha \ge 0$ and $\lambda_\alpha I_\alpha(x, u) = 0$ $(\alpha = 1, 2)$;

(b) $\dot{p}(t) = -f_x^*(\tilde{x}(t))p(t) + L_x^*(\tilde{x}(t)) + \sum_1^2 \lambda_\gamma L_{\gamma x}^*(\tilde{x}(t))$ $(t \in T)$;

(c) $f_u^*(\tilde{x}(t))p(t) = L_u^*(\tilde{x}(t)) + \sum_1^2 \lambda_\gamma L_{\gamma u}^*(\tilde{x}(t))$ $(t \in T)$.

Since $L_x(t,x,u) = (x_1, x_2)$ and $L_u(t,x,u) = (-u_1, -u_2)$ we clearly have, in view of the above relations, that (x_0, u_0, p, λ) with $(p, \lambda) \equiv (0,0)$ is an extremal. This proves (B). With respect to this extremal, note that $S_1(\lambda) = S$, and so (x_0, u_0) is also normal relative to S_1, Thus (C) holds.

Let us now evaluate the second variation of J along (x_0, u_0). Recall that

$$J''(x,u;y,v) = \int_{t_0}^{t_1} 2\Omega(t, y(t), v(t))dt((y,v) \in Z)$$

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

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

and $H[t]$ denotes $H(t, x(t), u(t), p(t), \lambda, 1)$. For this example, we have

$$\begin{aligned} H(t,x,u,p,\lambda,1) = p_1u_1(x_1+1) + p_2u_2(x_2+1) - (x_1^2 + x_2^2 - u_1^2 - u_2^2)/2 \\ - \lambda_1(x_1+x_2) - \lambda_2(x_1 + x_2 + u_1^3 + u_2^3) \end{aligned}$$

and therefore

$$\begin{aligned} H_x(t,x,u,p,\lambda,1) &= (p_1u_1 - x_1 - \lambda_1 - \lambda_2, p_2u_2 - x_2 - \lambda_1 - \lambda_2), \\ H_u(t,x,u,p,\lambda,1) &= (p_1(x_1+1) + u_1 - 3\lambda_2u_1^2, p_2(x_2+1) + u_2 - 3\lambda_2u_2^2). \end{aligned}$$

The second partial derivatives of H evaluated at $(t,x,u,p,\lambda,1)$ are given by

$$H_{xx} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},\ H_{uu} = \begin{pmatrix} 1-6\lambda_2u_1 & 0 \\ 0 & 1-6\lambda_2u_2 \end{pmatrix},\ H_{xu} = H_{ux} = \begin{pmatrix} p_1 & 0 \\ 0 & p_2 \end{pmatrix},$$

so that, with respect to the extremal (x_0, u_0, p, λ) with $(p, \lambda) \equiv (0,0)$, we have

$$J''(x_0, u_0; y, v) = \int_0^1 \{y_1^2(t) + y_2^2(t) - v_1^2(t) - v_2^2(t)\}dt$$

Let us now turn to the set of tangential constraints at (x_0, u_0) with respect to S_1. It corresponds to

$$\mathcal{R}_{S_1}(x_0,u_0) = \{(y,v) \in \mathcal{L}(x_0,u_0) \mid I'_\alpha(x_0,u_0;y,v) \leq 0\ (\alpha = 1,2)\}$$

where $(y,v) \in Z$ belongs to $\mathcal{L}(x_0, u_0)$ if and only if

$$\begin{pmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \end{pmatrix} = f_x(\tilde{x}_0(t)) \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} + f_u(\tilde{x}_0(t)) \begin{pmatrix} v_1(t) \\ v_2(t) \end{pmatrix} \quad (t \in T)$$

and $y_i(0) = y_i(1) = 0$ $(i = 1,2)$. In our case, with $(x_0, u_0) \equiv (0,0)$, the linear equation corresponds simply to $\dot{y}_i(t) = v_i(t)$ $(i = 1,2)$. On the other hand, we have

$$I'_1(x_0,u_0;y,v) = \int_0^1 \{y_1(t) + y_2(t)\}dt, \quad \text{and}$$

$$I'_2(x_0,u_0;y,v) = \int_0^1 \{y_1(t) + y_2(t) + 3u_{01}^2v_1(t) + 3u_{02}^2v_2(t)\}dt = \int_0^1 \{y_1(t) + y_2(t)\}dt.$$

Consider the process (y,v) defined for $i=1,2$ by

$$y_i(t) := \begin{cases} -t & \text{if } t\in[0,1/2] \\ t-1 & \text{if } t\in[1/2,1] \end{cases} \qquad v_i(t) := \begin{cases} -1 & \text{if } t\in[0,1/2) \\ 1 & \text{if } t\in[1/2,1]. \end{cases}$$

Note that

$$I_1'(x_0,u_0;y,v) = I_2'(x_0,u_0;y,v) = \int_0^{1/2} -2t dt + \int_{1/2}^1 2(t-1)dt = -\frac{1}{2} < 0$$

and so $(y,v)\in\mathcal{R}_{S_1}(x_0,u_0)$. Let us now evaluate $J''(x_0,u_0;y,v)$. We have

$$J''(x_0,u_0;y,v) = 2\int_0^{1/2}(t^2-1)dt + 2\int_{1/2}^1\{(t-1)^2-1\}dt = -\frac{11}{6} < 0$$

and so (D) holds.

From these four statements, now proved, we reach the following conclusion. Since (x_0,u_0) is not strongly normal, we cannot apply Theorem 2.2, the classical theorem on second order necessary conditions. However, by an application of Theorem 5.10, we deduce that (x_0,u_0) is not a solution to the problem.

7. Comparison with other works

In dealing with integral constraints, a usual technique found in the literature is based on the introduction of new variables in order to obtain an enlarged system of differential equations, thus converting the original Lagrange isoperimetric problem into a Mayer problem with endpoint equality and inequality constraints. Explicitly, define $\hat{f} = (\hat{f}^0,\hat{f}^1,\ldots,\hat{f}^{n+q})$ for $i=1,\ldots,n$ and $\gamma=1,\ldots,q$ as

$$\hat{f}^0(t,\hat{x},u) = L(t,x,u), \quad \hat{f}^i(t,\hat{x},u) = f^i(t,x,u), \quad \text{and} \quad \hat{f}^{n+\gamma}(t,\hat{x},u) = L_\gamma(t,x,u),$$

where $\hat{x} = (x^0,x^1,\ldots,x^{n+q})\in\mathbf{R}^{n+q+1}$. Then the isoperimetric Lagrange control problem posed in Section 2 is equivalent to the Mayer optimal control problem of minimizing $g_0(\hat{x}(t_1))$ subject to

$$\begin{cases} \dot{\hat{x}}(t) = \hat{f}(t,\hat{x}(t),u(t)) \ (t\in T), \\ \hat{x}(t_0) = \hat{\xi}_0, \\ g_{n+\alpha}(\hat{x}(t_1)) \le 0 \ (\alpha\in R) \text{ and } g_i(\hat{x}(t_1)) = 0 \ (i=1,\ldots,n;\ i=n+\beta\ (\beta\in Q)) \end{cases}$$

where $\hat{\xi}_0 = (0,\xi_0^1,\ldots,\xi_0^n,b_1,\ldots,b_q)$ and the function g mapping $\mathbf{R}^{n+q+1}$ to $\mathbf{R}^{n+q+1}$ is given by

$$g_0(\hat{x}) = x^0, \quad g_i(\hat{x}) = x^i - \xi_1^i \ \ (i=1,\ldots,n), \quad g_{n+\gamma}(\hat{x}) = x^{n+\gamma} \ \ (\gamma=1,\ldots,q).$$

In the literature, particularly in recent papers of Osmolovskii, Frankowska, Dmitruk and Bonnans, second order necessary conditions have been obtained for problems involving not only endpoint state constraints (as in the problem posed above) but also pointwise control or even mixed constraints. We refer the reader to [10, 17–21, 31–33] for the impressive body of recent, new results on the subject. There are, however, at least three main differences between those results and our main theorem.

• To begin with, the control functions in our problem are piecewise continuous and so do not form a Banach space. In the references mentioned above, results from abstract optimization on Banach spaces are applied to the optimal control problem, a method which does not work for our setting.

• Second, in some of those references the control set may be an arbitrary compact subset of $\mathbf{R}^m$; in others, one may find conditions of linear independence of the gradients of active control constraints, and so on. In all cases, however, the (endpoint or terminal) Lagrange function involves a nonnegative cost multiplier. Thus, no extra conditions are imposed on the functions delimiting the endpoint constraints. On the other hand, in our result, the main discussion is precisely on conditions which imply a positive cost multiplier. As we show, the condition imposed in [24] is too strong, and we are able to weaken it without altering the critical cone. In other words, we are providing a second order constraint qualification weaker than the one given in [24], and this is the main result of the paper.

A similar situation occurs in mathematical programming. The well-known Karush-Kuhn-Tucker conditions (or first order Lagrange multiplier rule) affect the constraints but not the cost function, while the Fritz John necessary condition includes a nonnegative cost multiplier. The former holds under additional assumptions (constraint qualifications) while the latter holds without them. In our paper, we discuss different (second order) constraint qualifications in order to have a positive cost multiplier, while in the above references no extra assumptions are added and the cost multiplier may be zero.

• Finally, in our result, given a Lagrange multiplier corresponding to a solution to the problem, we show that the second variation is nonnegative for all admissible variations (elements of the critical cone) related to the extremal, contrary to what is shown in some of those references, namely, that given an admissible variation there exists a multiplier for which the second variation is nonnegative.

To provide a concrete example illustrating some of these ideas, let us compare the results obtained in this paper with the main result provided in [33]. Define (P^∞) as the problem stated in Section 2 when minimization takes place in the space $W := W^{1,1}([t_0, t_1]; \mathbf{R}^n) \times L^\infty([t_0, t_1]; \mathbf{R}^m)$ rather than Z. Suppose that $(x_0, u_0) \in Z$ is a local solution of both (P) and (P^∞). In this particular setting, the results from [33] and this paper are applicable. As one readily verifies, through the transformation given above, the set $\Lambda(x_0, u_0)$ defined in Section 2 of [33] becomes the set of all $(p, \lambda, \lambda_0) \in X \times \mathbf{R}^q \times \mathbf{R}$ such that

(i) $\lambda_0 \geq 0$, $\lambda_\alpha \geq 0$ and $\lambda_\alpha I_\alpha(x_0, u_0) = 0$ $(\alpha \in R)$;

(ii) $\lambda_0 + \sum_{\alpha=1}^{r} \lambda_\alpha + \sum_{\beta=r+1}^{q} |\lambda_\beta| = 1$;

(iii) $-\dot{p}(t) = H_x^*(t, x_0(t), u_0(t), p, \lambda, \lambda_0)$;

(iv) $H_u(t, x_0(t), u_0(t), p, \lambda, \lambda_0) = 0$.

The set of critical directions $\mathcal{C}$, as defined in [33], corresponds to the set of all $(y, v) \in W$ such that

(i) $\dot{y}(t) = A(t)y(t) + B(t)v(t)$ a.e. in $[t_0, t_1]$;

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

(iii) $I'_\alpha(x_0, u_0; y, v) \le 0$ $(\alpha \in \{0\} \cup I^a(x_0, u_0))$ and $I'_\beta(x_0, u_0; y, v) = 0$ $(\beta \in Q)$.

According to [33, Theorem 1], the set $\Lambda(x_0, u_0)$ is nonempty and, for all $(y, v) \in \mathcal{C}$, there exists $(p, \lambda, \lambda_0) \in \Lambda(x_0, u_0)$ such that

$$\int_{t_0}^{t_1} 2\Omega(t, y(t), v(t); p, \lambda, \lambda_0) \ge 0.$$

The dependence of the integrand on (p, λ, λ_0) means that the second derivatives of the Hamiltonian are being evaluated along $(t, x_0(t), u_0(t), p(t), \lambda, \lambda_0)$ (see Sec. 2).

Our efforts go in the opposite direction. Instead of fixing a critical direction and asserting the nonnegativity of the above functional for some multiplier in Λ (where λ_0 can be zero), we fix a multiplier in $\mathcal{E}$ (where $\lambda_0 = 1$) and construct a cone of critical directions $\mathcal{T}_{S_1}$ where the above functional is nonnegative in every such direction (Theorem 5.3). Of course, the set $\mathcal{E}$ can be empty while Λ is not. Regularity relative to S does remedy this situation (see Theorem 5.5). The stronger hypothesis of regularity relative to S_1 simplifies considerably verifying whether an element belongs to $\mathcal{T}_{S_1}$ or not, while normality relative to S_1 implies regularity relative to S_1, a task much easier to verify in practice than the latter.

Acknowledgement. The second author is grateful to *Dirección General de Asuntos del Personal Académico*, from *Universidad Nacional Autónoma de México*, for the support given as part of the *PASPA* program during a sabbatical stay at the Department of Mathematical Sciences, University of Bath, United Kingdom.

References

[1] A. V. Arutyunov: *Smooth abnormal problems in extremum theory and analysis*, Russian Math. Surveys 67 (2012) 3–62.

[2] A. V. Arutyunov, D. Y. Karamzin: *Necessary conditions for a weak minimum in an optimal control problem with mixed constraints*, Differential Equations 41 (2005) 1532–1543.

[3] A. V. Arutyunov, F. L. Pereira: *Second-order necessary optimality conditions for problems without a priori normality assumptions*, Math. Oper. Research 31 (2006) 1–12.

[4] A. V. Arutyunov, Y. S. Vereshchagina: *On necessary second-order conditions in optimal control problems (in Russian)*, Diff. Uravneniya 38 (2002) 1443–1450; English translation in Diff. Equations 38 (2002) 1531–1540.

[5] J. P. Aubin, H. Frankowska: *Set-Valued Analysis*, Birkhäuser, Boston (1990).

[6] M. S. Bazaraa, H. D. Sherali, C. M. Shetty: *Nonlinear Programming – Theory and Algorithms*, Wiley, New York (1993).

[7] J. A. Becerril, J. F. Rosenblueth: *Necessity for isoperimetric inequality constraints*, Discrete Cont. Dyn. Systems, Ser. A 37 (2017) 1129–1158.

[8] J. A. Becerril, J. F. Rosenblueth: *The importance of being normal, regular and proper in the calculus of variations*, J. Optimization Theory Appl. 172 (2017) 759–773.

[9] A. Ben-Tal: *Second-order and related extremality conditions in nonlinear programming*, J. Optim. Theory Appl. 31 (1980) 143–165.

[10] J. F. Bonnans, N. P. Osmolovskii: *Quadratic growth conditions in optimal control problems*, Contemp. Math. 514 (2010) 85–98.

[11] F. Clarke: *Functional Analysis, Calculus of Variations and Optimal Control*, Springer, London (2013).

[12] E. A. Coddington, N. Levinson: *Theory of Ordinary Differential Equations*, McGraw-Hill, New York (1987).

[13] K. L. Cortez, J. F. Rosenblueth: *A second order constraint qualification for certain classes of optimal control problems*, WSEAS Trans. Systems Control 11 (2016) 419–424.

[14] K. L. Cortez, J. F. Rosenblueth: *Extended critical directions for time-control constrained problems*, Int. J. Circuits Systems Signal Processing 11 (2017) 1–11.

[15] K. L. Cortez, J. F. Rosenblueth: *Normality and uniqueness of Lagrange multipliers*, Discrete Cont. Dyn. Systems, Ser. A 38 (2018) 3169–3188.

[16] M. R. de Pinho, J. F. Rosenblueth: *Mixed constraints in optimal control: an implicit function theorem approach*, IMA J. Math. Control Information 24 (2007) 197–218.

[17] A. V. Dmitruk, N. P. Osmolovskii: *Necessary conditions for a weak minimum in optimal control problems with integral equations subject to state and mixed constraints*, SIAM J. Control Optimization 52 (2014) 3437–3462.

[18] A. V. Dmitruk, N. P. Osmolovskii: *Necessary conditions for a weak minimum in optimal control problems with integral equations on a variable time interval*, Discrete Cont. Dyn. Systems 35 (2015) 4323–4343.

[19] A. V. Dmitruk, N. P. Osmolovskii: *Necessary conditions for a weak minimum in a general optimal control problem with integral equations on a variable time interval*, Math. Control Rel. Fields 7 (2017) 507–535.

[20] H. Frankowska, N. P. Osmolovskii: *Second-order necessary conditions for a strong local minimum in a control problem with general control constraints*, Appl. Math. Optimization 80 (2019) 135–-164.

[21] H. Frankowska, N. P. Osmolovskii: *Strong local minimizers in optimal control problems with state constraints: second-order necessary conditions*, SIAM J. Control Optimization 56 (2018) 2353–2376.

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

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

[24] M. R. Hestenes: *Calculus of Variations and Optimal Control Theory*, Wiley, New York (1966).

[25] M. R. Hestenes: *Optimization Theory. The Finite Dimensional Case*, Wiley, New York (1975).

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

[27] E. Levitin, A. Milyutin, N. P. Osomolovskii: *Conditions of high order for a local minimum for problems with constraints*, Russian Math. Surveys 33 (1978) 97–168.

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

[29] A. A. Milyutin, N. P. Osmolovskii: *Calculus of Variations and Optimal Control*, Translations of Mathematical Monographs 180, American Mathematical Society, Providence (1998).

[30] N. P. Osmolovskii: *Second order conditions for a weak local minimum in an optimal control problem (necessity, sufficiency)*, Soviet Math Dokl. 16 (1975) 1480–1484.

[31] N. P. Osmolovskii: *Second order optimality conditions in optimal control problems with mixed inequality type constraints on a variable time interval*, Contemp. Mathematics 619 (2014) 141–155.

[32] N. P. Osmolovskii: *On second-order necessary conditions for broken extremals*, J. Optimization Theory Appl. 164 (2015) 379–406.

[33] N. P. Osmolovskii: *Necessary second-order conditions for a weak local minimum in a problem with endpoint and control constraints*, J. Math. Analysis Appl. 457 (2018) 1613–1633.

[34] J. F. Rosenblueth: *A new notion of conjugacy for isoperimetric problems*, Appl. Math. Optimization 50 (2004) 209–228.

[35] J. F. Rosenblueth: *Convex cones and conjugacy for inequality control constraints*, J. Convex Analysis 14 (2007) 361–393.

[36] J. F. Rosenblueth: *Modified critical directions for inequality control constraints*, WSEAS Trans. Systems Control 10 (2015) 215–227.

[37] J. F. Rosenblueth, G. Sánchez Licea: *Cones of critical directions in optimal control*, Int. J. Appl. Math. Informatics 7 (2013) 55–67.

[38] G. Wachsmuth: *On LICQ and the uniqueness of Lagrange multipliers*, Oper. Res. Letters 41 (2013) 78–80.

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