Journal of Applied Mathematics & Bioinformatics, vol.1, no.1, 2011, 1-20
ISSN: 1792-6602 (print), 1792-6939 (online)

Sufficiency in optimal control without the strengthened condition of Legendre

Gerardo Sánchez Licea[1]

Abstract

In this paper we derive a sufficiency theorem of an unconstrained fixed-endpoint problem of Lagrange which provides sufficient conditions for processes which do not satisfy the standard assumption of nonsingularity, that is, the new sufficiency theorem does not impose the strengthened condition of Legendre. The proof of the sufficiency result is direct in nature since the former uses explicitly the positivity of the second variation, in contrast with possible generalizations of conjugate points, solutions of certain matrix Riccati equations, invariant integrals, or the Hamiltonian-Jacobi theory.

Mathematics Subject Classification : 49K15
Keywords: Optimal control, sufficient conditions for optimality, strong minima, singular extremals, strengthened Legendre-Clebsch condition.

1 Introduction

In the classical calculus of variations it is well-known that the nonnegativity of the second variation along an arc x_0 over the set of admissible variations becomes a second order necessary condition for optimality. The theory of Jacobi

[1] Facultad de Ciencias UNAM, Departamento de Matemáticas, México D.F. 04510, México, e-mail: gesl@ciencias.unam.mx

Article Info: *Revised* : March 25, 2011. *Published online* : May 31, 2011

concerns with a characterization of the nonnegativity of the second variation over the set of admissible variations with the nonexistence of conjugate points on the underlying time interval under consideration. In concrete, a *smooth nonsingular* trajectory x_0 satisfies the fact that its second variation is nonnegative over the set of admissible variations if and only if x_0 satisfies the condition of Legendre and there are no conjugate points on x_0 in the underlying open interval. Moreover, the second variation is positive over the set of nonnull admissible variations if and only if x_0 satisfies the strengthened condition of Legendre and there are no conjugate points on x_0 in the half-open interval. One of the unfortunate features of Jacobi's theory arises when the arc under consideration is not nonsingular, that is, when it is *singular*, in this case, the theory of Jacobi is not applicable. On the other hand, all the classical sufficiency theorems in the theory of calculus of variations assume the strengthened condition of Legendre, and therefore the extremals under consideration satisfying the classical sufficiency conditions must be nonsingular.

Because of this fact, Ewing mentions in [7], that in the theory of calculus of variations, there is a gap between the necessary and sufficient conditions for optimality. In fact, in [7], Ewing devotes an entire section to problems for which Legendre strengthened condition fails. There he shows that we can partially close this gap by an elementary device discussed in [6] of adding a *penalty term*. This procedure is illustrated by means of three examples where the trajectory being examined is singular, but one can obtain a solution directly from the properties of the particular examples. However, this technique may not hold in general. Ewing states that 'although the use of the penalty term sheds light on the theory, it provides no panacea for attacking particular examples. Indeed there are no panaceas!'

In more recent years, the study of second order conditions for optimality in the theory of calculus of variations and optimal control has provided an extensive literature (see [1-5, 17-22, 24, 26-35] and references therein). In [17] sufficient conditions in the calculus of variations are obtained by using an appropriate form of local convexity. For optimal control problems, sufficiency is derived in [25, 26] by applying the positivity of the second variation as well as a generalized theory of Jacobi in terms of conjugate points, the insertion of the original optimal control problem in a Banach space exhibits in [35] an alternate sufficiency method, the construction of a bounded solution to a matrix-valued Riccati equation, a verification function satisfying the Hamilton-Jacobi equation, and a quadratic function that satisfies a Hamilton-Jacobi inequality become fundamental devices in sufficiency results given in [2, 5, 16, 18-21, 24].

A different approach which provides sufficiency in the classical isoperimetric calculus of variations fixed-endpoint problem is given by Hestenes in [15]. This method treats explicitly with the positivity of the second variation on the set of nonnull admissible variations and it is implicitly based on the concept of a directionally convergent sequence of trajectories which is in turn a generalization of the concept of directional convergence for vectors in the finite dimension case. The development of this technique as it appears in [15], as well as its application to more general problems, can be traced back to different papers of the author and McShane (see [8-14, 23]). A generalization of this method which covers optimal control problems can be found in [26, 31, 32].

In this paper we study an unconstrained fixed-endpoint optimal control problem of Lagrange. The main contribution is based on two fundamental facts. First, we show how by applying a similar technique of that given in [26, 31, 32], one can be able to obtain an itself-contained proof to a sufficiency result which is in contrast with possible generalizations of conjugate points, solutions of certain matrix Riccati equations, invariant integrals, theory of extremals, or the Hamiltonian-Jacobi theory. Second, the new sufficiency theorem does not include the standard assumption of nonsingularity, that is, the strengthened Legendre-Clebsch condition is not imposed. In particular, we refer the reader to [24] where the importance of this condition is fully explained.

The paper is organized as follows. In Section 2 we pose the problem we shall deal with, introduce some notation and basic definitions, and state the main result. In Section 3 we illustrate the usefulness of the new sufficiency theorem by means of a simple example of a singular process which is a proper strong minimum of the problem in hand. Section 4 is devoted to the proof of the main sufficiency theorem together with the statement of an auxiliary result on which the proof is strongly based. The auxiliary result, which implicitly includes a possible generalization of the notion of a directionally convergent sequence of trajectories, is established in Section 5.

2 The problem and the main result

The fixed-endpoint optimal control problem we shall study in this paper can be stated as follows. Suppose we are given an interval $T := [t_0, t_1]$ in $\mathbf{R}$, two points ξ_0 and 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.

Let $X := AC(T; \mathbf{R}^n)$ denote the space of absolutely continuous functions mapping T to $\mathbf{R}^n$, let $U := L^1(T; \mathbf{R}^m)$, set $Z := X \times U$, and denote by Z_e the set of all $(x, u) \in Z$ satisfying

a. $L(t, x(t), u(t))$ is integrable on T.

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

c. $x(t_0) = \xi_0$, $x(t_1) = \xi_1$.

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

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

For this problem, an *admissible process* is an element of Z_e, that is, a couple (x, u) comprising functions $x \in X$ and $u \in U$ which satisfy the constraints of problem (P). An admissible process (x, u) is called a *strong minimum* of problem (P) if there exists $\epsilon > 0$ such that $I(x, u) \leq I(y, v)$ for all $(y, v) \in Z_e$, $(y, v) \neq (x, u)$, with $\|y - x\|_\infty < \epsilon$. If the inequality can be replaced by a strict inequality then (x, u) is said to be a proper strong minimum of (P).

We shall assume throughout the paper that the functions L and f are continuous and of class C^2 with respect to x and u on $T \times \mathbf{R}^n \times \mathbf{R}^m$.

For the theory to follow we shall find convenient to introduce the following definitions.

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

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

- A triple (x, u, p) will be called an *extremal* if (x, u) is a process, $p \in X$,

$$\dot{p}(t) = -H_x^*(t, x(t), u(t), p(t)) \text{ (a.e. in } T) \quad \text{and} \quad H_u(t, x(t), u(t), p(t)) = 0 \ (t \in T)$$

where '*' denotes transpose. Throughout the paper all derivatives such as H_x

and H_u will be gradient row vectors and all vector-valued functions such as $p \in X$ will be considered as column vectors.

• For a given $p \in X$ define, for all $(t, x, u) \in T \times \mathbf{R}^n \times \mathbf{R}^m$,

$$F(t,x,u) := L(t,x,u) - \langle p(t), f(t,x,u)\rangle - \langle \dot{p}(t), x\rangle \quad [= -H(t,x,u,p(t)) - \langle \dot{p}(t), x\rangle].$$

With respect to F, define the functional $J : Z_e \to \mathbf{R}$ as

$$J(x,u) := \langle p(t_1), \xi_1\rangle - \langle p(t_0), \xi_0\rangle + \int_{t_0}^{t_1} F(t, x(t), u(t))dt.$$

Consider the *first variation* of J along $(x,u) \in X \times L^\infty(T; \mathbf{R}^m)$ over $(y, v) \in Z$ given by

$$J'((x,u);(y,v)) := \int_{t_0}^{t_1} \{F_x(t,x(t),u(t))y(t) + F_u(t,x(t),u(t))v(t)\}dt,$$

and the *second variation of* J along $(x,u) \in X \times L^\infty(T; \mathbf{R}^m)$ over $(y,v) \in X \times L^2(T; \mathbf{R}^m)$ given by

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

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

$$2\Omega(t,y,v) := \langle y, F_{xx}(t,x(t),u(t))y\rangle + 2\langle y, F_{xu}(t,x(t),u(t))v\rangle + \langle v, F_{uu}(t,x(t),u(t))v\rangle.$$

Also, with respect to F, denote by $\mathcal{E}$ the *Weierstrass excess function* which corresponds to

$$\mathcal{E}(t,x,u,v) := F(t,x,v) - F(t,x,u) - F_u(t,x,u)(v-u)$$

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

• A process (x,u) is *nonsingular* if the determinant $|-H_{uu}(t,x(t),u(t),p(t))| = |F_{uu}(t,x(t),u(t))|$ is different from zero for all $t \in T$. A process (x,u) satisfies the strengthened condition of Legendre if the matrix $-H_{uu}(t,x(t),u(t),p(t)) = F_{uu}(t,x(t),u(t)) > 0$ for all $t \in T$.

• For all $(x,u) \in X \times L^\infty(T; \mathbf{R}^m)$, denote by $Y(x,u)$ the class of all $(y,v) \in X \times L^2(T; \mathbf{R}^m)$ satisfying

$$\dot{y}(t) = A(t)y(t) + B(t)v(t) \text{ a.e. in } T, y(t_0) = y(t_1) = 0$$

where $A(t) := f_x(t, x(t), u(t))$, $B(t) := f_u(t, x(t), u(t))$ $(t \in T)$. Elements of $Y(x, u)$ will be called *admissible variations* along (x, u).

- For all $x \in X$ and all $u \in U$ let

$$D_1(x) := \int_{t_0}^{t_1} \varphi(\dot{x}(t))dt \qquad \text{and} \qquad D_2(u) := \int_{t_0}^{t_1} \varphi(u(t))dt$$

where

$$\varphi(c) := (1 + |c|^2)^{1/2} - 1.$$

- Define $D\colon Z \to \mathbf{R}$ by

$$D(x, u) := \max\{D_1(x), D_2(u)\},$$

and denote by $\|\cdot\| = \|\cdot\|_\infty$ the supremum norm in X.

Let us now state the main theorem of the paper. It consists of a sufficiency result for a proper strong minimum of problem (P) assuming, with respect to a given extremal, the Legendre-Clebsch condition, the positivity of the second variation along nonnull admissible variations, and two conditions related to the Weierstrass excess function.

2.1 Theorem: *Let (x_0, u_0, p) be an extremal with $u_0 \in L^\infty(T; \mathbf{R}^m)$ and suppose that the there exist $h, \epsilon > 0$ such that*

i. $F_{uu}(t, x_0(t), u_0(t)) \geq 0$ (a.e. *in* T).

ii. $J''((x_0, u_0); (y, v)) > 0$ *for all nonnull admissible variations* (y, v) *along* (x, u).

iii. *For all* $(x, u) \in Z_e$ *satisfying* $\|x - x_0\| < \epsilon$, $\mathcal{E}(t, x(t), u_0(t), u(t)) \geq 0$ (a.e. *in* T) *and*

$$\int_{t_0}^{t_1} \mathcal{E}(t, x(t), u_0(t), u(t))dt \geq hD(x - x_0, u - u_0).$$

Then there exist $\rho, \delta > 0$ such that, for all admissible processes (x, u) satisfying $\|x - x_0\| < \rho$,

$$I(x, u) \geq I(x_0, u_0) + \delta D(x - x_0, u - u_0).$$

In particular, (x_0, u_0) is a proper strong minimum of (P).

3 Example

In this section we provide an example of a fixed-endpoint problem for which an application of Theorem 2.1 shows that the singular extremal under consideration is in fact a proper strong minimum.

3.1 Example: Consider the problem of minimizing

$$I(x,u) = \int_0^4 \{2u_1^2(t) + |u_2(t)|^3 - x_1(t) - x_2(t)\}dt$$

subject to

a. $2u_1^2(t) + |u_2(t)|^3 - x_1(t) - x_2(t)$ is integrable on $[0,4]$.

b. $(\dot{x}_1(t), \dot{x}_2(t)) = (\sin^2 u_2(t), u_1(t))$ a.e. in $[0,4]$.

c. $(x_1(0), x_2(0)) = (0,0)$ and $(x_1(4), x_2(4)) = (0,-2)$.

For this case $n = m = 2$, $T = [0,4]$, $\xi_0 = (0,0)^*$, $\xi_1 = (0,-2)^*$,

$$L(t,x,u) = 2u_1^2 + |u_2|^3 - x_1 - x_2 \quad \text{and} \quad f(t,x,u) = (\sin^2 u_2, u_1)^*.$$

Let $x_0(t) = (0, -t^2/8)^*$, $u_0(t) = (-t/4, 0)^*$ $(t \in T)$. Clearly, $(x_0, u_0) \in Z_e$. We have

$$H(t,x,u,p) = p_1 \sin^2 u_2 + p_2 u_1 - 2u_1^2 - |u_2|^3 + x_1 + x_2,$$

$$H_x(t,x,u,p) = (1,1), \quad H_u(t,x,u,p) = (p_2 - 4u_1, 2p_1 \sin u_2 \cos u_2 - 3|u_2|u_2).$$

Therefore, (x_0, u_0, p) with $p(t) = (-t,-t)^*$ $(t \in T)$ is an extremal. In addition, $F(t,x,u) = 2u_1^2 + |u_2|^3 + t\sin^2 u_2 + tu_1$, and so

$$F_{uu}(t,x,u) = \begin{pmatrix} 4 & 0 \\ 0 & 6|u_2| + 2t\cos^2 u_2 - 2t\sin^2 u_2 \end{pmatrix}.$$

Observe that

$$\langle F_{uu}(t, x_0(t), u_0(t))h, h\rangle = 4h_1^2 + 2th_2^2 \quad (t \in T)$$

so that (x_0, u_0) does not satisfy the strengthened condition of Legendre but 2.1(i) holds. Now, observe that

$$A(t) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad B(t) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (t \in T)$$

and hence $(y,v)\in Y(x_0,u_0)$ implies that $(\dot y_1(t),\dot y_2(t))=(0,v_1(t))$ a.e. in T. It follows that

$$J''((x_0,u_0);(y,v))=\int_0^4\{4v_1^2(t)+2tv_2^2(t)\}dt>0$$

for all $(y,v)\in Y(x_0,u_0)$, $(y,v)\neq(0,0)$. Hence 2.1(ii) holds.

Clearly, $(x,u)\in Z_e$ implies that $x_1\equiv 0$ and $u_2(t)=k(t)\pi$ $(t\in T)$ for some integrable function $k(\cdot)$ which maps T to the set of integers. Observing that $\varphi(c)\le |c|^2/2$ for all $c\in\mathbf{R}^2$, $(x,u)\in Z_e$ implies that

$$\begin{aligned}
\mathcal{E}(t,x(t),u_0(t),u(t)) &= \mathcal{E}(t,0,x_2(t),-t/4,0,u_1(t),k(t)\pi)\\
&= 2u_1^2(t)+|k(t)|^3\pi^3+tu_1(t)+t^2/8\\
&= 2(u_1(t)+t/4)^2+|k(t)|^3\pi^3\\
&\ge 2^{-1}[(u_1(t)+t/4)^2+k^2(t)\pi^2]\\
&\ge \max\{\varphi((0,u_1(t)+t/4)^*),\varphi((u_1(t)+t/4,k(t)\pi)^*)\}\\
&= \max\{\varphi(\dot x(t)-\dot x_0(t)),\varphi(u(t)-u_0(t))\}\quad \text{(a.e. in } T).
\end{aligned}$$

Thus with any $\epsilon>0$ and $h=1$, 2.1(iii) holds. By Theorem 2.1, (x_0,u_0) is a proper strong minimum of problem (P).

4 Proof of Theorem 2.1

In this section we shall prove Theorem 2.1. We first state an auxiliary result (proved in Section 5) on which the proof of Theorem 2.1 is strongly based. Implicit on the statement of the result it is inserted a possible generalization of the notion of a directionally convergent sequence of trajectories, firstly introduced in a calculus of variations context by Hestenes in [15], page 155.

4.1 Lemma: *Let* $\{z_q:=(x_q,u_q)\}$ *be a sequence in* Z, $z_0:=(x_0,u_0)\in Z$, *and suppose that*

$$\lim_{q\to\infty}D(z_q-z_0)=0\quad\text{and}\quad d_q:=[2D(z_q-z_0)]^{1/2}>0\quad (q\in\mathbf{N}).$$

For all $q\in\mathbf{N}$ *and almost all* $t\in T$ *define*

$$w_q(t):=\max\left\{\left[1+\frac{1}{2}\varphi(\dot x_q(t)-\dot x_0(t))\right]^{1/2},\left[1+\frac{1}{2}\varphi(u_q(t)-u_0(t))\right]^{1/2}\right\}.$$

For all $q \in \mathbf{N}$ and $t \in T$ define

$$y_q(t) := \frac{x_q(t) - x_0(t)}{d_q}, \qquad v_q(t) := \frac{u_q(t) - u_0(t)}{d_q}.$$

Then the following hold:

a. *For some $v_0 \in L^2(T; \mathbf{R}^m)$ and some subsequence of $\{z_q\}$, again denoted by $\{z_q\}$, $\{v_q\}$ converges weakly to v_0 in $L^1(T; \mathbf{R}^m)$. Moreover, $\{(\dot{x}_q, u_q)\}$ converges almost uniformly to $(\dot{x}_0, u_0)$ on T and hence $w_q(t) \to 1$ almost uniformly on T.*

b. *There exist a function $\sigma_0 \in L^2(T; \mathbf{R}^n)$ and some subsequence of $\{z_q\}$, again denoted by $\{z_q\}$, such that $\{\dot{y}_q\}$ converges weakly in $L^1(T; \mathbf{R}^n)$ to σ_0. Moreover, if we define*

$$y_0(t) := \int_{t_0}^{t} \sigma_0(s) ds \quad (t \in T),$$

then $y_q(t) \to y_0(t)$ uniformly on T.

c. *Suppose $S \subset T$ is measurable and $w_q(t) \to 1$ uniformly on S. Let $R_q(\cdot), R_0(\cdot)$ be $m \times m$ real matrix-valued functions with $R_q(\cdot)$ measurable on S, $R_0(\cdot) \in L^\infty(S; \mathbf{R}^{m\times m})$, $R_q(t) \to R_0(t)$ uniformly on S, and $R_0(t) \geq 0$ $(t \in S)$. Then for some subsequence of $\{z_q\}$, again denoted by $\{z_q\}$,*

$$\liminf_{q\to\infty} \int_S \langle R_q(t) v_q(t), v_q(t)\rangle dt \geq \int_S \langle R_0(t) v_0(t), v_0(t)\rangle dt.$$

Proof of Theorem 2.1:

Let $z_0 := (x_0, u_0)$. Assume that, for all $\rho, \delta > 0$, there exists $z = (x, u) \in Z_e$ with $\|x - x_0\| < \rho$ such that

$$J(x, u) < J(x_0, u_0) + \delta D(z - z_0). \tag{1}$$

We are going to show that this contradicts (ii) of Theorem 2.1 and the statement will follow, since $I(x, u) = J(x, u)$ on Z_e.

Note that, for all $z = (x, u) \in Z_e$,

$$J(z) = J(z_0) + J'(z_0; z - z_0) + K(z) + \tilde{\mathcal{E}}(z) \tag{2}$$

where

$$\tilde{\mathcal{E}}(x, u) := \int_{t_0}^{t_1} \mathcal{E}(t, x(t), u_0(t), u(t)) dt,$$

$$K(x,u) := \int_{t_0}^{t_1} \{M(t,x(t)) + \langle u(t) - u_0(t), N(t,x(t))\rangle\} dt,$$

$$M(t,y) := F(t,y,u_0(t)) - F(t,x_0(t),u_0(t)) - F_x(t,x_0(t),u_0(t))(y - x_0(t)),$$

$$N(t,y) := F_u^*(t,y,u_0(t)) - F_u^*(t,x_0(t),u_0(t)).$$

By Taylor's theorem,

$$M(t,y) = \frac{1}{2}\langle y - x_0(t), P(t,y)(y - x_0(t))\rangle, \qquad N(t,y) = Q(t,y)(y - x_0(t)),$$

where

$$P(t,y) := 2\int_0^1 (1-\lambda) F_{xx}(t, x_0(t) + \lambda(y - x_0(t)), u_0(t)) d\lambda,$$

$$Q(t,y) := \int_0^1 F_{ux}(t, x_0(t) + \lambda(y - x_0(t)), u_0(t)) d\lambda.$$

Let us begin by proving the existence of $\alpha_0, \delta_0 > 0$ such that, for all $z = (x,u) \in Z_e$ with $\|x - x_0\| < \delta_0$,

$$|K(x,u)| \leq \alpha_0 \|x - x_0\| [1 + D(z - z_0)]. \qquad (3)$$

By using the inequality of Schwarz and the continuity of the functions P and Q we may choose $\alpha, \delta_0 > 0$ such that for all $z \in Z_e$ with $\|x - x_0\| < \delta_0$,

$$|M(t,x(t)) + \langle u(t) - u_0(t), N(t,x(t))\rangle| \leq \alpha |x(t) - x_0(t)| [1 + |u(t) - u_0(t)|^2]^{1/2} \quad (t \in T).$$

Set $\alpha_0 := \max\{\alpha, \alpha(t_1 - t_0)\}$. Then, for all $z \in Z_e$ with $\|x - x_0\| < \delta_0$,

$$|K(z)| \leq \alpha \|x - x_0\| \int_{t_0}^{t_1} [1 + \varphi(u(t) - u_0(t))] dt \leq \alpha_0 \|x - x_0\| [1 + D(z - z_0)]$$

and hence (3) holds with α_0 and δ_0 given above.

Now, by (1), for all $q \in \mathbf{N}$ there exists $z_q := (x_q, u_q) \in Z_e$ such that

$$\|x_q - x_0\| < \delta_0, \quad \|x_q - x_0\| < \frac{1}{q}, \quad J(z_q) - J(z_0) < \frac{1}{q} D(z_q - z_0). \qquad (4)$$

Since $z_q \in Z_e$, observe that the last inequality implies that $u_q(t) \neq u_0(t)$ on a set of positive measure and so $D(z_q - z_0) > 0$ $(q \in \mathbf{N})$. Since (x_0, u_0, p) is an extremal, it follows that $J'(z_0; w) = 0$ for all $w \in Z$. With this in mind, by (2), the second condition of 2.1(iii), and (3),

$$J(z_q) - J(z_0) = K(z_q) + \tilde{\mathcal{E}}(z_q) \geq -\alpha_0 \|x_q - x_0\| + D(z_q - z_0)(h - \alpha_0 \|x_q - x_0\|).$$

By (4) we obtain

$$D(z_q - z_0)\left(h - \frac{1}{q} - \frac{\alpha_0}{q}\right) < \frac{\alpha_0}{q}$$

and consequently $D(z_q - z_0) \to 0$, $q \to \infty$. Define d_q, w_q, y_q and v_q, as in Lemma 4.1, that is, $d_q := [2D(z_q - z_0)]^{1/2}$,

$$w_q(t) := \max\left\{\left[1 + \frac{1}{2}\varphi(\dot{x}_q(t) - \dot{x}_0(t))\right]^{1/2}, \left[1 + \frac{1}{2}\varphi(u_q(t) - u_0(t))\right]^{1/2}\right\},$$

$$y_q(t) := \frac{x_q(t) - x_0(t)}{d_q} \quad \text{and} \quad v_q(t) := \frac{u_q(t) - u_0(t)}{d_q}.$$

By Lemma 4.1a there exist $v_0 \in L^2(T; \mathbf{R}^m)$ and some subsequence of $\{z_q\}$, again denoted by $\{z_q\}$, such that $\{v_q\}$ converges weakly in $L^1(T; \mathbf{R}^m)$ to v_0. By Lemma 4.1b for some $\sigma_0 \in L^2(T; \mathbf{R}^n)$ and some subsequence of $\{z_q\}$, again denoted by $\{z_q\}$, if $y_0(t) := \int_{t_0}^t \sigma_0(s)ds$ $(t \in T)$, then

$$\lim_{q\to\infty} y_q(t) = y_0(t) \quad \text{uniformly on } T.$$

The theorem will be proved if we show that $J''(z_0; (y_0, v_0)) \leq 0$, $(y_0, v_0) \in Y(z_0)$ and $(y_0, v_0) \neq (0, 0)$.

The fact that $y_0(t_0) = y_0(t_1) = 0$ follows by Lemma 4.1b. By definition of the functional K, for all $q \in \mathbf{N}$,

$$\frac{K(z_q)}{d_q^2} = \int_{t_0}^{t_1}\left\{\frac{M(t, x_q(t))}{d_q^2} + \left\langle v_q(t), \frac{N(t, x_q(t))}{d_q}\right\rangle\right\}dt.$$

In view of Lemma 4.1b,

$$\lim_{q\to\infty} \frac{M(t, x_q(t))}{d_q^2} = \frac{1}{2}\langle y_0(t), F_{xx}(t, x_0(t), u_0(t))y_0(t)\rangle,$$

$$\lim_{q\to\infty} \frac{N(t, x_q(t))}{d_q} = F_{ux}(t, x_0(t), u_0(t))y_0(t)$$

both uniformly on T and, since $\{v_q\}$ converges weakly to v_0 in $L^1(T; \mathbf{R}^m)$,

$$\frac{1}{2}J''(z_0; (y_0, v_0)) = \lim_{q\to\infty} \frac{K(z_q)}{d_q^2} + \frac{1}{2}\int_{t_0}^{t_1} \langle v_0(t), F_{uu}(t, x_0(t), u_0(t))v_0(t)\rangle dt. \quad (5)$$

Let us now show that for some subsequence of $\{z_q\}$, again denoted by $\{z_q\}$,

$$\liminf_{q\to\infty} \frac{\tilde{\mathcal{E}}(z_q)}{d_q^2} \geq \frac{1}{2}\int_{t_0}^{t_1} \langle v_0(t), F_{uu}(t, x_0(t), u_0(t))v_0(t)\rangle dt. \quad (6)$$

By Lemma 4.1a, we may choose $S \subset T$ measurable such that $(\dot{x}_q(t), u_q(t)) \to (\dot{x}_0(t), u_0(t))$ uniformly on S. By Taylor's theorem, for all $t \in S$ and $q \in \mathbf{N}$, we have

$$\frac{1}{d_q^2}\mathcal{E}(t, x_q(t), u_0(t), u_q(t)) = \frac{1}{2}\langle v_q(t), R_q(t)v_q(t)\rangle$$

where

$$R_q(t) := 2\int_0^1 (1-\lambda)F_{uu}(t, x_q(t), u_0(t) + \lambda[u_q(t) - u_0(t)])d\lambda.$$

Clearly,

$$\lim_{q\to\infty} R_q(t) = R_0(t) := F_{uu}(t, x_0(t), u_0(t)) \qquad \text{uniformly on } S.$$

By 2.1(i), $R_0(t) \geq 0$ $(t \in S)$. Moreover, by the first condition of 2.1(iii) and Lemma 4.1c, for some subsequence of $\{z_q\}$, again denoted by $\{z_q\}$,

$$\liminf_{q\to\infty} \frac{\tilde{\mathcal{E}}(z_q)}{d_q^2} \geq \frac{1}{2}\int_S \langle v_0(t), F_{uu}(t, x_0(t), u_0(t))v_0(t)\rangle dt.$$

Since S can be chosen to differ from T by a set of an arbitrary small measure, and the function

$$t \mapsto \langle v_0(t), F_{uu}(t, x_0(t), u_0(t))v_0(t)\rangle$$

belongs to $L^1(T; \mathbf{R})$, this inequality holds when $S = T$, and this establishes (6). Thus, by (4), (5) and (6),

$$\frac{1}{2}J''(z_0; (y_0, v_0)) \leq \lim_{q\to\infty} \frac{K(z_q)}{d_q^2} + \liminf_{q\to\infty} \frac{\tilde{\mathcal{E}}(z_q)}{d_q^2} = \liminf_{q\to\infty} \frac{J(z_q) - J(z_0)}{d_q^2} \leq 0.$$

In addition, if $(y_0, v_0) = (0, 0)$, then

$$\lim_{q\to\infty} \frac{K(z_q)}{d_q^2} = 0$$

and so, by the second condition of 2.1(iii),

$$\frac{1}{2}h \leq \liminf_{q\to\infty} \frac{\tilde{\mathcal{E}}(z_q)}{d_q^2} \leq 0$$

contradicting the positivity of h.

Finally, to show that $(y_0, v_0) \in Y(z_0)$, observe that, by Taylor's theorem, for all $q \in \mathbf{N}$,

$$\dot{y}_q(t) = A_q(t)y_q(t) + B_q(t)v_q(t) \text{ (a.e. in } T)$$

where

$$A_q(t) = \int_0^1 f_x(t, x_0(t) + \lambda[x_q(t) - x_0(t)], u_0(t) + \lambda[u_q(t) - u_0(t)])d\lambda,$$

$$B_q(t) = \int_0^1 f_u(t, x_0(t) + \lambda[x_q(t) - x_0(t)], u_0(t) + \lambda[u_q(t) - u_0(t)])d\lambda.$$

We know by Lemma 4.1a that there exists $S \subset T$ measurable such that

$$A_q(t) \to A_0(t) := f_x(t, x_0(t), u_0(t)), \qquad B_q(t) \to B_0(t) := f_u(t, x_0(t), u_0(t))$$

both uniformly on S. Since $y_q(t) \to y_0(t)$ uniformly on S and $\{v_q\}$ converges weakly to v_0 in $L^1(S; \mathbf{R}^m)$, it follows that $\{\dot{y}_q\}$ converges weakly in $L^1(S; \mathbf{R}^n)$ to $A_0 y_0 + B_0 v_0$. By Lemma 4.1b, $\{\dot{y}_q\}$ converges weakly in $L^1(S; \mathbf{R}^n)$ to $\sigma_0 = \dot{y}_0$. Hence,

$$\dot{y}_0(t) = A_0(t) y_0(t) + B_0(t) v_0(t) \quad (t \in S).$$

Since S can be chosen to differ from T by a set of an arbitrary small measure, there cannot exist a subset of T of positive measure on which the functions y_0 and v_0 do not satisfy the differential equation $\dot{y}_0(t) = A_0(t) y_0(t) + B_0(t) v_0(t)$. Consequently,

$$\dot{y}_0(t) = A_0(t) y_0(t) + B_0(t) v_0(t) \text{ (a.e. in } T)$$

and this completes the proof. ▮

5 Proof of Lemma 4.1

(a): Observe that $\varphi(c)(2 + \varphi(c)) = |c|^2$ $(c \in \mathbf{R}^m)$. Then for all $q \in \mathbf{N}$,

$$\int_{t_0}^{t_1} \frac{|v_q(t)|^2}{w_q^2(t)} dt \leq 1. \tag{7}$$

Thus there exist $v_0 \in L^2(T; \mathbf{R}^m)$ and some subsequence of $\{z_q\}$ (we do not relabel) such that $\{v_q / w_q\}$ converges weakly to v_0 in $L^2(T; \mathbf{R}^m)$. Let $h \in L^\infty(T; \mathbf{R}^m)$. Note that, for all $q \in \mathbf{N}$,

$$\int_{t_0}^{t_1} \langle h(t), v_q(t) \rangle dt = \int_{t_0}^{t_1} \left\langle h(t), \frac{v_q(t)}{w_q(t)} \right\rangle dt + \int_{t_0}^{t_1} \left\langle h(t)[w_q(t) - 1], \frac{v_q(t)}{w_q(t)} \right\rangle dt.$$

By the inequality of Schwarz and (7),

$$\left|\int_{t_0}^{t_1}\left\langle h(t)[w_q(t)-1], \frac{v_q(t)}{w_q(t)}\right\rangle dt\right|^2 \leq \int_{t_0}^{t_1} |h(t)|^2[w_q(t)-1]^2 dt.$$

For all $q \in \mathbf{N}$, set

$$w_{1q}(t) := \left[1+\frac{1}{2}\varphi(\dot{x}_q(t)-\dot{x}_0(t))\right]^{1/2} \text{ (a.e. in } T\text{)},$$

$$w_{2q}(t) := \left[1+\frac{1}{2}\varphi(u_q(t)-u_0(t))\right]^{1/2} \ (t \in T).$$

Since for $i = 1, 2$, $w_{iq}^2(t) \geq w_{iq}(t) \geq 1$ for almost all $t \in T$, we have

$$\begin{aligned} 0 &\leq \int_{t_0}^{t_1}[w_{iq}(t)-1]dt \leq \int_{t_0}^{t_1}[w_{iq}^2(t)-1]dt \\ &\leq \max\left\{\int_{t_0}^{t_1}\varphi(\dot{x}_q(t)-\dot{x}_0(t))dt, \int_{t_0}^{t_1}\varphi(u_q(t)-u_0(t))dt\right\} = D(z_q-z_0). \end{aligned}$$

Thus it is readily seen that

$$\lim_{q\to\infty}\int_{t_0}^{t_1}[w_q(t)-1]dt = \lim_{q\to\infty}\int_{t_0}^{t_1}[w_q^2(t)-1]dt = 0.$$

Observe also that

$$\int_{t_0}^{t_1}[w_q(t)-1]^2 dt = \int_{t_0}^{t_1}[w_q^2(t)-1]dt - 2\int_{t_0}^{t_1}[w_q(t)-1]dt.$$

Consequently,

$$\lim_{q\to\infty}\int_{t_0}^{t_1}|h(t)|^2[w_q(t)-1]^2 dt = 0.$$

Since $L^\infty(T;\mathbf{R}^m) \subset L^2(T;\mathbf{R}^m)$,

$$\lim_{q\to\infty}\int_{t_0}^{t_1}\langle h(t), v_q(t)\rangle dt = \lim_{q\to\infty}\int_{t_0}^{t_1}\left\langle h(t), \frac{v_q(t)}{w_q(t)}\right\rangle dt = \int_{t_0}^{t_1}\langle h(t), v_0(t)\rangle dt,$$

that is, $\{v_q\}$ converges weakly in $L^1(T;\mathbf{R}^m)$ to v_0.

In order to prove that $\{(\dot{x}_q, u_q)\}$ converges almost uniformly to $(\dot{x}_0, u_0)$ on T, for all $x \in X$, set

$$\|x\|_1 := \int_{t_0}^{t_1}|\dot{x}(t)|dt, \qquad w(t) := \left[1+\frac{1}{2}\varphi(\dot{x}(t))\right]^{1/2} \quad \text{(a.e in } T\text{)}.$$

Observe first that

$$\int_{t_0}^{t_1} 2w^2(t)dt = 2(t_1 - t_0) + \int_{t_0}^{t_1} \varphi(\dot{x}(t))dt = 2(t_1 - t_0) + D_1(x)$$

and

$$\int_{t_0}^{t_1} \frac{|\dot{x}(t)|^2}{2w^2(t)}dt = \int_{t_0}^{t_1} \frac{|\dot{x}(t)|^2}{2 + \varphi(\dot{x}(t))}dt = \int_{t_0}^{t_1} \varphi(\dot{x}(t))dt = D_1(x).$$

By the inequality of Schwarz,

$$\|x\|_1^2 \leq \int_{t_0}^{t_1} \frac{|\dot{x}(t)|^2}{2w^2(t)}dt \int_{t_0}^{t_1} 2w^2(t)dt.$$

From these relations we have

$$\|x\|_1^2 \leq D_1(x)[2t_1 - 2t_0 + D_1(x)] \leq D(z)[2t_1 - 2t_0 + D(z)] \quad (z = (x, u)).$$

Consequently, $\|x_q - x_0\|_1 \to 0$, $q \to \infty$, and so some subsequence of $\{\dot{x}_q\}$ converges pointwisely a.e. to $\dot{x}_0$. By Egoroff's theorem, it converges to $\dot{x}_0$ almost uniformly on T.

Similarly it is readily seen that

$$\|u\|_1^2 \leq D_2(u)[2t_1 - 2t_0 + D_2(u)] \leq D(z)[2t_1 - 2t_0 + D(z)] \quad (z = (x, u))$$

implying that some subsequence of $\{u_q\}$ converges almost uniformly to u_0 on T. Thus there is some subsequence of $\{z_q\}$ (we do not relabel) such that

$$\lim_{q\to\infty} (\dot{x}_q(t), u_q(t)) = (\dot{x}_0(t), u_0(t)) \quad \text{almost uniformly on } T.$$

(b): As in (a), for all $q \in \mathbf{N}$,

$$\int_{t_0}^{t_1} \frac{|\dot{y}_q(t)|^2}{w_q^2(t)}dt \leq 1. \tag{8}$$

Hence there exists a function $\sigma_0 \in L^2(T; \mathbf{R}^n)$ such that some subsequence of $\{\dot{y}_q/w_q\}$ converges weakly in $L^2(T; \mathbf{R}^n)$ to σ_0. We conclude, by an argument similar to that used in the proof of (a), that there exists some subsequence of $\{z_q\}$ (we do not relabel) such that $\{\dot{y}_q\}$ converges weakly in $L^1(T; \mathbf{R}^n)$ to σ_0.

It remains to show that $y_q(t) \to y_0(t)$ uniformly on T. We have

$$y_q(t) = \int_{t_0}^{t} \dot{y}_q(s)ds \quad (t \in T,\ q \in \mathbf{N}),$$

and hence

$$\lim_{q\to\infty} y_q(t) = y_0(t) := \int_{t_0}^{t} \sigma_0(s)ds \quad \text{pointwisely on } T.$$

In order to prove that this convergence is uniform observe that, by (8), given a measurable set $S \subset T$,

$$\left|\int_S \dot{y}_q(t)dt\right|^2 \le \int_S \frac{|\dot{y}_q(t)|^2}{w_q^2(t)}dt \int_S w_q^2(t)dt \le \int_S w_q^2(t)dt \quad (q \in \mathbf{N}).$$

Moreover

$$\int_S w_q^2(t)dt = m(S) + \int_S [w_q^2(t) - 1]dt \quad (q \in \mathbf{N}).$$

Given a constant $\epsilon > 0$, choose $q_\epsilon \in \mathbf{N}$ such that

$$\int_{t_0}^{t_1} [w_q^2(t) - 1]dt < \frac{\epsilon^2}{2} \quad (q \ge q_\epsilon).$$

Choose $0 < \delta < \epsilon^2/2$ such that

$$m(S) < \delta \Rightarrow \left|\int_S \dot{y}_q(t)dt\right| < \epsilon \quad (q < q_\epsilon).$$

Note that if $q \ge q_\epsilon$, then

$$m(S) < \delta \Rightarrow \left|\int_S \dot{y}_q(t)dt\right|^2 \le m(S) + \int_{t_0}^{t_1} [w_q^2(t) - 1]dt < \frac{\epsilon^2}{2} + \frac{\epsilon^2}{2} = \epsilon^2,$$

and so

$$m(S) < \delta \Rightarrow \left|\int_S \dot{y}_q(t)dt\right| < \epsilon \quad (q \in \mathbf{N}).$$

Thus the sequence of functions $\{y_q\}$ is equicontinuous on T. Consequently, $y_q(t) \to y_0(t)$ uniformly on T.

(c): By hypothesis we may assume that, for all $t \in S$ and $q \in \mathbf{N}$,

$$|R_q(t) - R_0(t)|w_q^2(t) \le 1.$$

Hence

$$M_q := \sup_{t\in S} |R_q(t) - R_0(t)|w_q^2(t) < \infty \qquad (q \in \mathbf{N}).$$

Using the inequality of Schwarz it is easily seen that, for all $t \in S$ and $q \in \mathbf{N}$,

$$|\langle R_q(t)v_q(t), v_q(t)\rangle - \langle R_0(t)v_q(t), v_q(t)\rangle| \le M_q \frac{|v_q(t)|^2}{w_q^2(t)}.$$

Since $R_q(t) \to R_0(t)$, and $w_q(t) \to 1$, both uniformly on S, we have $M_q \to 0$. Therefore, by (7),

$$\liminf_{q\to\infty} \int_S \langle R_q(t)v_q(t), v_q(t)\rangle dt = \liminf_{q\to\infty} \int_S \langle R_0(t)v_q(t), v_q(t)\rangle dt.$$

But for all $t \in S$,

$$\begin{aligned}\langle R_0(t)v_q(t), v_q(t)\rangle = \langle R_0(t)v_0(t), v_0(t)\rangle &+ 2\langle v_q(t) - v_0(t), R_0(t)v_0(t)\rangle \\ &+ \langle R_0(t)(v_q(t) - v_0(t)), v_q(t) - v_0(t)\rangle.\end{aligned}$$

Since $w_q(t) \to 1$ uniformly on S, it is readily seen (see the proof of (a)) that there is some subsequence of $\{z_q\}$ (again denoted by $\{z_q\}$) such that $\{v_q\}$ converges weakly to v_0 in $L^2(S; \mathbf{R}^m)$. Since $R_0 v_0 \in L^2(S; \mathbf{R}^m)$, we have

$$\lim_{q\to\infty} \int_S \langle R_0(t)v_0(t), v_q(t) - v_0(t)\rangle dt = 0.$$

Hence

$$\begin{aligned}\liminf_{q\to\infty} \int_S \langle R_q(t)v_q(t), v_q(t)\rangle dt &= \int_S \langle R_0(t)v_0(t), v_0(t)\rangle dt \\ &+ \liminf_{q\to\infty} \int_S \langle R_0(t)(v_q(t) - v_0(t)), v_q(t) - v_0(t)\rangle dt.\end{aligned}$$

Since the last term is nonnegative the result follows. ▮

References

[1] A. A. Agrachev, G. Stefani, and P. L. Zezza, Strong optimality for a bang-bang trajectory, *SIAM J. Control Optim.*, **41**, (2002), 991-1014.

[2] V. M. Alekseev, V.M. Tikhomirov, and S. V. Fomin, *Optimal Control*, Translated from the Russian by V. M. Volosov, Contemporary Soviet Mathematics, Consultants Bureau, New York, 1987.

[3] R. Berlanga, J. F. Rosenblueth, Jacobi's Condition for Singular Extremals: An Extended Notion of Conjugate Points, *Appl. Math. Letters*, **15**, (2002), 453-458.

[4] R. Berlanga, J. F. Rosenblueth, Extended conjugate points in the calculus of variations, *IMA J. Math. Control Inform.*, **21**, (2004), 159-173.

[5] F. H. Clarke, *Methods of Dynamic and Nonsmooth Optimization*, CBMS-NSF Regional Conference Series in Applied Mathematics **57**, Society for Industrial and Applied Mathematics, Philadelphia, 1989.

[6] G. M. Ewing, Sufficient conditions for a non-regular problem in the calculus of variations,*Bull. Am. Math. Soc.*, **43**, (1937), 371-376.

[7] G. M. Ewing, *Calculus of Variations with Applications*, Dover, New York, 1985.

[8] M. R. Hestenes, Sufficient conditions for the problem of Bolza in the calculus of variations, *Trans. Amer. Math. Soc.*, **36**, (1934), 793-818.

[9] M. R. Hestenes, On sufficient conditions in the problems of Lagrange and Bolza, *Ann. of Math.*, **37**, (1936), 543-551.

[10] M. R. Hestenes, A direct sufficiency proof for the problem of Bolza in the calculus of variations, *Trans. Amer. Math. Soc.*, **42**, (1937), 141-154.

[11] M. R. Hestenes, Sufficient conditions for the isoperimetric problem of Bolza in the calculus of variations, *Trans. Amer. Math. Soc.*, **60**, (1946), 93-118.

[12] M. R. Hestenes, The Weierstrass E-function in the calculus of variations, *Trans. Amer. Math. Soc.*, **60**, (1947), 51-71.

[13] M. R. Hestenes, An indirect sufficiency proof for the problem of Bolza in nonparametric form, *Trans. Amer. Math. Soc.*, **62**, (1947), 509-535.

[14] M. R. Hestenes, Sufficient conditions for multiple integral problems in the calculus of variations, *Amer. J. Math.*, **70**, (1948), 239-276.

[15] M. R. Hestenes, *Calculus of Variations and Optimal Control Theory*, John Wiley, New York, 1966.

[16] A. D. Ioffe and V. M. Tihomirov, *Theory of Extremal Problems*, Translated from the Russian by K. Makowski, Studies in Mathematics and its Applications, **6**, North-Holland Publishing, Amsterdam, 1979.

[17] P. D. Loewen, Second-order sufficiency criteria and local convexity for equivalent problems in the calculus of variations, *J. Math. Anal. Appl.*, **146**, (1990), 512-522.

[18] K. Malanowski, Sufficient optimality conditions for optimal control subject to state constraints, *SIAM J. Control Optim.*, **35**, (1997), 205-227.

[19] K. Malanowski, H. Maurer and S. Pickenhain, Second order sufficient conditions for state- constrained optimal control problems, *J. Optim. Theory Appl.*, **123**, (2004), 595-617.

[20] H. Maurer and S. Pickenhain, Second order sufficient conditions for control problems with mixed control-state constraints, *J. Optimal Theory Appl.*, **86**, (1995), 649-667.

[21] H. Maurer and H. J. Oberle, Second order sufficient conditions for optimal control problems with free final time: the Riccati approach, *SIAM J. Control Optim.*, **41**, (2002), 380–403.

[22] D. Q. Mayne, Sufficient conditions for a control to be a strong minimum, *J. Optimal Theory Appl.*, **21**, (1977), 339-351.

[23] E. J. McShane, Sufficient conditions for a weak relative minimum in the problem of Bolza, *Trans. Amer. Math. Soc.*, **52**, (1942), 344-379.

[24] A.A. Milyutin and N. P. Osmolovskiĭ, *Calculus of Variations and Optimal Control*, Translations of Mathematical Monographs, **180**, American Mathematical Society, Providence, Rhode Island, 1998.

[25] M. Popescu, Singular normal extremals and conjugate points for Bolza functionals, *J. Optimal Theory and Appl.*, **23**, (2007), 243-256.

[26] J. F. Rosenblueth, Variational conditions and conjugate points for the fixed-endpoint control problem, *IMA J. Math. Control Inform.*, **16**, (1999), 147-163.

[27] J. F. Rosenblueth, A New Notion of Conjugacy for Isoperimetric Problems, *Appl. Math. Optim.*, **50**, (2004), 209-228.

[28] J. F. Rosenblueth, Conjugacy for normal or abnormal linear Bolza problems, *J. Math. Anal. Appl.*, **318**, (2006), 444-458.

[29] J. F. Rosenblueth, A direct approach to second order conditions for mixed equality constraints, *J. Math. Anal. Appl.*, **333**, (2007), 770-779.

[30] Maria Do Rosário de Pinho, J. F. Rosenblueth, Mixed constraints in optimal control: an implicit function theorem approach, *IMA J. Math. Control Inform.*, **24**, (2007), 197-218.

[31] J. F. Rosenblueth, G. Sánchez Licea, A direct sufficiency proof for a weak minimum in optimal control, *Appl. Math. Sci.*, **4**(6), (2010), 253-269.

[32] J. F. Rosenblueth, G. Sánchez Licea, Sufficient variational conditions for isoperimetric control problems, *Int. Math. Forum.*, **6**, (2011), 303-324.

[33] G. Stefani, On sufficient optimality conditions for a singular arc, *Proceedings of the 42nd IEEE Conference on Decision and Control*, (2003), 2746-2749

[34] G. Stefani, Minimum-time optimality of a singular arc: second order sufficient conditions, *Proceedings of the 43rd IEEE Conference on Decision and Control*, (2004), 450-454.

[35] G. Stefani and P. L. Zezza, Optimality conditions for a constrained optimal control problem, *SIAM J. Control Optim.*, **34**, (1996), 635-659.