Applied Mathematical Sciences, Vol. 4, 2010, no. 6, 253 - 269

A Direct Sufficiency Proof for a Weak Minimum in Optimal Control

Javier F Rosenblueth

IIMAS, Universidad Nacional Autónoma de México
Apartado Postal 20-726, México DF 01000
jfrl@servidor.unam.mx

Gerardo Sánchez Licea

Facultad de Ciencias, Universidad Nacional Autónoma de México
Departamento de Matemáticas, México DF 04510
gslicea@apolo.acatlan.unam.mx

Abstract

In this paper we derive a sufficiency theorem for a weak minimum of a fixed-endpoint optimal control problem. The proof is direct in nature as it deals explicitly with the positivity of the second variation, in contrast with possible generalizations of Jacobi theory, solutions to matrix-valued Riccati equations, or Hamilton-Jacobi theory. The approach we follow is essentially a generalization of the one introduced by Hestenes, in terms of directional convergent sequences of trajectories, originally posed for the problem of Bolza in the calculus of variations.

Mathematics Subject Classification: 49K15

Keywords: Optimal control, sufficient conditions for optimality, weak minima

1 Introduction

In 1966, Hestenes included in his classical text on calculus of variations and optimal control theory (see [9]) a sufficiency proof for strong and weak minima in the calculus of variations that does not make use of invariant integrals, Mayer fields, conjugate points, or the Hamilton-Jacobi theory.

The proof deals explicitly with the positivity of the second variation along admissible variations and it is applicable to cases in which Mayer fields may not exist (for example, in the case of an isoperimetric problem, the theory of Mayer fields does not apply without transforming it to a problem of Lagrange). The method is thus illustrated in the book by applying it to the fixed-endpoint isoperimetric problem. The development of this technique as it appears in [9], as well as its application to more general problems, can be traced back to different papers of the author and McShane (see [2-8, 15]).

In more recent years, one can find an extensive literature on second order sufficient conditions for certain optimal control problems (see [1, 10-14, 16, 18] and references therein). Some of the approaches include a generalized theory of Jacobi and conjugate points, the construction of a bounded solution to a matrix-valued Riccati equation, a quadratic function that satisfies a Hamilton-Jacobi inequality, or the insertion of the original optimal control problem as an abstract optimization problem in a Banach space. The sufficiency proof given by Hestenes in [9] has, however, received little attention in the context of optimal control and the main objective of this paper is to show how it can be successfully generalized to that theory.

The original proof is strongly based on the concept, introduced by Hestenes, of a directional convergent sequence of trajectories (absolutely continuous functions) which is in turn a generalization of the concept of directional convergence for vectors in the finite dimension case. This notion relies on the specific calculus of variations problem considered in [9] but, as we show in this paper, it can be modified to cover optimal control problems. Based on that notion, we provide a direct sufficiency proof for a weak minimum for the fixed-endpoint optimal control problem.

Let us point out that the result itself is not new and it was previously established in [17] by a similar technique. However, some statements of the proof given in [17] seem to be rather incomplete and, by filling some of the gaps, we provide in this paper a new and clearer proof.

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. Section 3 is devoted to the proof of the sufficiency theorem together with the statement of an auxiliary result on which the proof is based. The auxiliary result, which implicitly includes the generalization of the notion of a directional convergent sequence of trajectories, is established in Section 4.

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 $\mathbf{R}^n$, a set $\mathcal{A}$ in $T \times \mathbf{R}^n \times \mathbf{R}^m$, 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 all 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 $Z_e(\mathcal{A})$ the set of all $(x, u) \in Z$ satisfying

a. $L(\cdot, x(\cdot), u(\cdot)) \in L^1(T; \mathbf{R})$;
b. $\dot{x}(t) = f(t, x(t), u(t))$ a.e. in T;
c. $x(t_0) = \xi_0$, $x(t_1) = \xi_1$;
d. $(t, x(t), u(t)) \in \mathcal{A}$ $(t \in T)$.

The problem we shall deal with, which we label (P), is that of minimizing I over $Z_e(\mathcal{A})$, where

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

For this problem, a *process* is an element of $Z_e(\mathcal{A})$, that is, a couple (x, u) comprising functions $x \in X$ and $u \in U$ which satisfy the constraints of problem (P). A process (x, u) is called a (global) *solution* of (P) if it belongs to

$$S(\mathcal{A}) := \{(x, u) \in Z_e(\mathcal{A}) \mid I(x, u) \leq I(y, v) \text{ for all } (y, v) \in Z_e(\mathcal{A})\},$$

and a *weak minimum* of (P) if there exists $\epsilon > 0$ such that (x, u) belongs to $S(T_1((x, u); \epsilon) \cap \mathcal{A})$, where

$$T_1((x, u); \epsilon) := \{(t, y, v) \in T \times \mathbf{R}^n \times \mathbf{R}^m : |x(t) - y| < \epsilon,\ |u(t) - v| < \epsilon\}.$$

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, $p \in X$,

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

where '$*$' denotes transpose.

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

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

With respect to F, define the functional J as

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

Denote by $C := C(T; \mathbf{R}^m)$ the space of all continuous functions mapping T to $\mathbf{R}^m$ and consider the *first variation* of J with respect to $(x, u) \in X \times C$ along $(y, v) \in Z$ 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 with respect to $(x, u) \in X \times C$ along $(y, v) \in X \times L^2(T; \mathbf{R}^m)$ 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 $(t, x, u, v) \in T \times \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^m$.

- For all $(x, u) \in X \times C$, let the space of *admissible variations* be given by

$$Y(x, u) := \{(y, v) \in X \times L^2(T; \mathbf{R}^m) \mid \dot{y}(t) = A(t)y(t) + B(t)v(t) \text{ a.e. in } T, \; y(t_0) = y(t_1) = 0\}$$

where, for all $t \in T$, $A(t) := f_x(t, x(t), u(t))$ and $B(t) := f_u(t, x(t), u(t))$.

- For all $u \in U$ let

$$D(u) := \int_{t_0}^{t_1} \varphi(u(t))dt \quad \text{where} \quad \varphi(c) := (1 + |c|^2)^{1/2} - 1.$$

Denote by $\|\cdot\|_0$ the supremum norm and define $\eta \colon Z \to \mathbf{R} \cup \{+\infty\}$ by

$$\eta(x, u) := \|x\|_0 + \|u\|_0.$$

Let us now state the main theorem of the paper. It corresponds to a sufficiency result for a strict weak minimum of problem (P) assuming, with respect to a given extremal, the strengthened Legendre-Clebsch condition and the positivity of the second variation along nonnull admissible variations.

Theorem 2.1 *Let (x_0, u_0, p) be an extremal with u_0 continuous and suppose that*

i. $F_{uu}(t, x_0(t), u_0(t)) > 0$ $(t \in T)$.

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

Then there exist $\rho, \delta > 0$ such that

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

for all processes (x, u) satisfying $\eta(x - x_0, u - u_0) < \rho$. In particular, (x_0, u_0) is a strict weak minimum of (P).

3 Proof of Theorem 2.1

In this section we shall prove Theorem 2.1. We first state an auxiliary result on which the proof is strongly based. Implicit in the statement of the result we have included a generalization of the notion of a directional convergent sequence of trajectories, first introduced in a calculus of variations context by Hestenes in [9].

Lemma 3.1 *Let $\{z_q = (x_q, u_q)\}$ be a sequence in Z, $(x_0, u_0) \in Z$, and suppose that*

$$\lim_{q\to\infty} D(u_q - u_0) = 0 \quad \text{and} \quad d_q := [2D(u_q - u_0)]^{1/2} > 0 \quad (q \in \mathbf{N}).$$

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

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

$$v_q(t) := \frac{u_q(t) - u_0(t)}{d_q}.$$

Then the following hold:

a. *There exists $v_0 \in L^2(T; \mathbf{R}^m)$ such that $\{v_q\}$ converges weakly to v_0 in $L^1(T; \mathbf{R}^m)$.*

b. *Let $A_q \in L^\infty(T; \mathbf{R}^{n\times n})$ and $B_q \in L^\infty(T; \mathbf{R}^{n\times m})$ be matrix functions for which there exist constants $m_0, m_1 > 0$ such that, for all $q \in \mathbf{N}$, $\|A_q\|_\infty \leq m_0$ and $\|B_q\|_\infty \leq m_1$, and suppose that y_q satisfies the system*

$$\dot{y}(t) = A_q(t)y(t) + B_q(t)v_q(t) \ (a.e. \ in \ T), \quad y(t_0) = 0.$$

Then there exist $\sigma_0 \in L^2(T; \mathbf{R}^n)$ and a 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. Moreover, if

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

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

c. *If $f_q, f\colon T \to \mathbf{R}^m$ are measurable functions with $f \in L^\infty(T;\mathbf{R}^m)$ and $f_q(t) \to f(t)$ uniformly on T then*

$$\lim_{q\to\infty} \int_{t_0}^{t_1} \langle f_q(t), v_q(t)\rangle dt = \int_{t_0}^{t_1} \langle f(t), v_0(t)\rangle dt.$$

d. *Suppose that $w_q(t) \to 1$ uniformly on T. Let R_q, R_0 be quadratic forms with $m\times m$ associated matrices $R_q(\cdot)$ measurable on T, $R_0(\cdot) \in L^\infty(T;\mathbf{R}^{m\times m})$, $R_q(t) \to R_0(t)$ uniformly on T, and $R_0(t) \geq 0$ a.e. in T. Then*

$$\liminf_{q\to\infty} \int_{t_0}^{t_1} R_q(t; v_q(t))dt \geq \int_{t_0}^{t_1} R_0(t; v_0(t))dt.$$

Proof of Theorem 2.1:

Assume that, for all $\rho, \delta > 0$, there exists $(x,u) \in Z_e(\mathcal{A})$ with $\eta(x - x_0, u - u_0) < \rho$ such that

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

We are going to show that this contradicts (ii) of Theorem 2.1 and the first statement will follow. The second conclusion is a consequence of the first since $J(x,u) = I(x,u)$ for all $(x,u) \in Z_e(\mathcal{A})$.

Let $z_0 := (x_0,u_0)$. Note that, for all $z = (x,u) \in Z_e(\mathcal{A})$,

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

where

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

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

and the functions M and N are given by

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

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

By Taylor's theorem we have

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

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 $h, \alpha_0, \delta > 0$ such that, for all $z = (x,u) \in Z_e(\mathcal{A})$ with $\eta(z - z_0) < \delta$,

$$\mathcal{E}^*(x,u) \geq hD(u - u_0), \tag{3}$$

$$|K(x,u)| \leq \alpha_0\eta(z - z_0)[1 + D(u - u_0)]. \tag{4}$$

Since $F_{uu}(t, x_0(t), u_0(t)) > 0$ $(t \in T)$ and u_0 is continuous, there exist $h, \epsilon > 0$ such that

$$\langle c, F_{uu}(t,x,u)c\rangle \geq h|c|^2 \quad (c \in \mathbf{R}^m,\ (t,x,u) \in T_1(z_0;\epsilon)).$$

By Taylor's theorem, for all (t,x,u,v) with (t,x,u) and (t,x,v) in $T_1(z_0;\epsilon)$, we have

$$\mathcal{E}(t,x,u,v) = \int_0^1 (1-\lambda)\langle v - u, F_{uu}(t,x,u + \lambda(v-u))(v-u)\rangle d\lambda$$

Therefore

$$\mathcal{E}(t,x,u,v) \geq \frac{h}{2}|v-u|^2 \geq h\varphi(v-u)$$

and so

$$\mathcal{E}^*(z) = \int_{t_0}^{t_1} \mathcal{E}(t, x(t), u_0(t), u(t))dt \geq h\int_{t_0}^{t_1} \varphi(u(t) - u_0(t))dt = hD(u - u_0)$$

for all $z \in Z_e(\mathcal{A})$ satisfying $\eta(z - z_0) < \epsilon$. Choose $\alpha, \mu > 0$ such that, for all $z \in Z_e(\mathcal{A})$ with $\|x - x_0\|_0 < \mu$ and $t \in T$,

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

Set $\alpha_0 := \max\{\alpha, \alpha(t_1 - t_0)\}$. Then

$$|K(z)| \leq \alpha\|x - x_0\|_0 \int_{t_0}^{t_1} [1 + \varphi(u(t) - u_0(t))]dt \leq \alpha_0\eta(z - z_0)[1 + D(u - u_0)]$$

for all $z \in Z_e(\mathcal{A})$ with $\eta(z - z_0) < \mu$. Hence (3) and (4) hold with h, α_0 given above and $\delta = \min\{\epsilon, \mu\}$.

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

$$\eta(z_q - z_0) < \delta, \quad \eta(z_q - z_0) < \frac{1}{q}, \quad J(z_q) - J(z_0) < \frac{1}{q}D(u_q - u_0). \tag{5}$$

Observe that the last inequality implies that $u_q(t) \neq u_0(t)$ on a set of positive measure and so $D(u_q - u_0) > 0$ $(q \in \mathbf{N})$. Since $J'(z_0; w) = 0$ for all $w \in Z$, it follows by (2), (3) and (4) that

$$J(z_q) - J(z_0) = K(z_q) + \mathcal{E}^*(z_q) \geq -\alpha_0\eta(z_q - z_0) + D(u_q - u_0)(h - \alpha_0\eta(z_q - z_0)).$$

By (5) we obtain

$$D(u_q - u_0)\Big(h - \frac{1}{q} - \frac{\alpha_0}{q}\Big) < \frac{\alpha_0}{q}$$

and consequently $D(u_q - u_0) \to 0$, $q \to \infty$. Define d_q, w_q, y_q and v_q as in Lemma 3.1.

By Lemma 3.1(a) there exists $v_0 \in L^2(T; \mathbf{R}^m)$ such that $\{v_q\}$ converges weakly in $L^1(T; \mathbf{R}^m)$ to v_0. By Taylor's theorem, for all $q \in \mathbf{N}$ we have

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

where

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

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

By continuity of f_x and f_u there exist $m_0, m_1 > 0$ such that $\|A_q\|_\infty \le m_0$ and $\|B_q\|_\infty \le m_1$ ($q \in \mathbf{N}$). By Lemma 3.1(b) there exist $\sigma_0 \in L^2(T; \mathbf{R}^n)$ and a subsequence of $\{z_q\}$ (we do not relabel) such that, if

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

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

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

The fact that $y_0(t_0) = y_0(t_1) = 0$ follows by Lemma 3.1(b). Now, by definition of the functional K,

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

In view of Lemma 3.1(b),

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

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

both uniformly on T. This fact, together with Lemma 3.1(c), implies that

$$\frac{1}{2} J''(z_0; (y_0, v_0)) = \lim_{q\to\infty} \frac{K(z_q)}{d_q^2} + \frac{1}{2} \int_{t_0}^{t_1} \langle v_0(t), F_{uu}(t, x_0(t), u_0(t)) v_0(t) \rangle dt. \quad (6)$$

Now, by Taylor's theorem,

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

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,

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

Since $w_q(t) \to 1$ uniformly on T and $R_0(t) \geq 0$ ($t \in T$), it follows by Lemma 3.1(d) that

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

This fact, together with (5) and (6), implies that

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

In addition, if $(y_0, v_0) = 0$, then $\lim_{q\to\infty} K(z_q)/d_q^2 = 0$ and so, by (3),

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

contradicting the positivity of h.

Finally, to show that $(y_0, v_0) \in Y(z_0)$, note that $y_q(t) \to y_0(t)$,

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

all uniformly on T, and $\{v_q\}$ converges weakly to v_0 in $L^1(T; \mathbf{R}^m)$. Therefore $\{\dot{y}_q\}$ converges weakly in $L^1(T; \mathbf{R}^n)$ to $A_0 y_0 + B_0 v_0$. By Lemma 3.1(b), $\{\dot{y}_q\}$ converges weakly in $L^1(T; \mathbf{R}^n)$ to $\sigma_0 = \dot{y}_0$. Hence,

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

and this completes the proof. ∎

4 Proof of Lemma 3.1

In order to prove Lemma 3.1 we shall first establish three auxiliary results.

Set $L^r_{p\times q} := L^r(T;\mathbf{R}^{p\times q})$. For all $q\in\mathbf{N}$, let $A_q\in L^1_{n\times n}$ and define $\Gamma_q\colon L^\infty_{n\times n}\to L^\infty_{n\times n}$ as

$$\Gamma_q\Phi(t) := \int_{t_0}^t A_q(s)\Phi(s)ds \quad (\Phi\in L^\infty_{n\times n},\ t\in T).$$

It is readily seen that Γ_q is a bounded linear operator.

Lemma 4.1 *Let $A\in L^1_{n\times n}$. For all $k\in\mathbf{N}$ define $\Phi^k\colon T\to AC(T;\mathbf{R}^{n\times n})$ by*

$$\Phi^k(t) := I + \int_{t_0}^t A(s)\Phi^{k-1}(s)ds \quad \textit{and} \quad \Phi^0(t) := I\ (t\in T)$$

where I is the $n\times n$ identity matrix. Then there exists a unique $\Phi\in AC(T;\mathbf{R}^{n\times n})$ satisfying

$$\dot\Phi(t) = A(t)\Phi(t)\ (a.e.\ in\ T), \quad \Phi(t_0) = I,$$

and such that $\Phi^k(t)\to\Phi(t)$ uniformly on T.

Proof: Let us first assume that $\int_{t_0}^{t_1}|A(t)|dt<1$. Observe that, for all $k\in\mathbf{N}$,

$$\|\Phi^k-\Phi^{k-1}\|_\infty \le \left(\int_{t_0}^{t_1}|A(t)|dt\right)^k.$$

Set $M_0 := \int_{t_0}^{t_1}|A(t)|dt$ and choose m,p and N in $\mathbf{N}$ such that $m>p\ge N$. We have

$$\|\Phi^m-\Phi^p\|_\infty \le \sum_{k=p}^{m-1}\|\Phi^{k+1}-\Phi^k\|_\infty \le \sum_{k=N}^{\infty} M_0^{k+1} = \frac{M_0^{N+1}}{1-M_0}.$$

Since $M_0<1$, $M_0^{N+1}\to 0$ as $N\to\infty$. Therefore, $\{\Phi^k\}$ is a Cauchy sequence in $L^\infty_{n\times n}$. Thus, there exists $\Phi\in L^\infty_{n\times n}$ such that $\Phi^k(t)\to\Phi(t)$ uniformly on T. Now, since

$$\lim_{k\to\infty} A(s)\Phi^k(s) = A(s)\Phi(s) \quad \text{uniformly on } [t_0,t],$$

we have

$$\Phi(t) = I + \int_{t_0}^t A(s)\Phi(s)ds \quad (t\in T).$$

Consequently, $\Phi\in AC(T;\mathbf{R}^{n\times n})$ satisfies

$$\dot\Phi(t) = A(t)\Phi(t)\ (\text{a.e. in } T), \quad \Phi(t_0) = I$$

and so Φ is the unique solution of the system given above. Since $A\in L^1_{n\times n}$, the case $\int_{t_0}^{t_1}|A(t)|dt\ge 1$ can be reduced to the previous one by considering the equation on possibly shorter intervals. ∎

Lemma 4.2 *For all $q \in \mathbf{N} \cup \{0\}$, let $\Phi_q \in AC(T; \mathbf{R}^{n\times n})$ be the solution of the initial value problem*

$$\dot{\Phi}(t) = A_q(t)\Phi(t) \ (a.e.\ in\ T), \quad \Phi(t_0) = I$$

where $A_q \in L^1_{n\times n}$. If $\int_{t_0}^{t_1} |A_q(t)|dt \leq c_0$ $(q \in \mathbf{N})$ for some $c_0 > 0$ then $\{\Phi_q\}$ is bounded in $L^\infty_{n\times n}$.

Proof: Since

$$|\Gamma_q \Phi(t)| \leq \left(\int_{t_0}^{t_1} |A_q(t)|dt\right) \cdot \|\Phi\|_\infty \quad (t \in T,\ q \in \mathbf{N}),$$

it follows that

$$\sup_{q\in\mathbf{N}} \|\Gamma_q \Phi\|_\infty < \infty \quad \text{for all } \Phi \in L^\infty_{n\times n}.$$

By the Banach-Steinhaus Theorem there exists $c_1 > 0$ such that

$$\|\Gamma_q \Phi\|_\infty \leq c_1 \|\Phi\|_\infty \quad (\Phi \in L^\infty_{n\times n},\ q \in \mathbf{N}). \tag{7}$$

Let $\Phi^0: T \to L^\infty_{n\times n}$ be given by $\Phi^0(t) = I$ and, for all $k \in \mathbf{N}$, let $\Phi^k: T \to L^\infty_{n\times n}$ be such that

$$\Phi^k(t) = I + \int_{t_0}^{t} A_0(s)\Phi^{k-1}(s)ds.$$

In view of (7) we have

$$\|\Gamma_q \Phi^k\|_\infty \leq c_1 \|\Phi^k\|_\infty \quad (k \in \mathbf{N} \cup \{0\},\ q \in \mathbf{N}).$$

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

$$\Phi^{k,q}(t) := I + \int_{t_0}^{t} A_q(s)\Phi^{k-1}(s)ds = I + \Gamma_q \Phi^{k-1}(t).$$

Again by (7),

$$\|\Phi^{k,q}\|_\infty \leq \|I\|_\infty + c_1 \|\Phi^{k-1}\|_\infty \quad (k, q \in \mathbf{N}). \tag{8}$$

By Lemma 4.1, $\Phi^k(t) \to \Phi_0(t)$ uniformly on T. Thus, there exists $c_2 > 0$ such that

$$\|\Phi^k\|_\infty \leq c_2 \quad (k \in \mathbf{N}). \tag{9}$$

By (8) and (9), there exists $c_3 > 0$ such that

$$\|\Phi^{k,q}\|_\infty \leq c_3 \quad (k, q \in \mathbf{N}). \tag{10}$$

Denote by m the Lebesgue measure and suppose that $\{\Phi_q\}$ is not bounded in $L^\infty_{n\times n}$. Then, for all $n \in \mathbf{N}$, there exist $E_n \subset T$ with $0 < m(E_n)$ and $q_n \in \mathbf{N}$ such that $|\Phi_{q_n}(t)| > n$ $(t \in E_n)$. Since, for each fixed $q \in \mathbf{N}$,

$$\lim_{k\to\infty} \Phi^{k,q}(t) = \Phi_q(t) \quad \text{uniformly on } T,$$

there exists $K_q \in \mathbf{N}$ such that

$$|\Phi^{k,q}(t) - \Phi_q(t)| < 1 \quad (k \geq K_q,\ q \in \mathbf{N},\ t \in T).$$

We have

$$|\Phi^{k,q_n}(t) - \Phi_{q_n}(t)| < 1 \quad (k \geq K_{q_n},\ n \in \mathbf{N},\ t \in T).$$

By (10),

$$n < |\Phi_{q_n}(t)| < 1 + |\Phi^{K_{q_n},q_n}(t)| \leq 1 + \|\Phi^{K_{q_n},q_n}\|_\infty \leq 1 + c_3 \quad (n \in \mathbf{N},\ t \in E_n)$$

which is a contradiction. Therefore $\{\Phi_q\}$ is bounded in $L^\infty_{n\times n}$. ▮

The following result can be proved in a similar way.

Lemma 4.3 *For all $q \in \mathbf{N} \cup \{0\}$, let $\Phi_q^{-1} \in AC(T;\mathbf{R}^{n\times n})$ be the solution of the initial value problem*

$$\dot{\Phi}^{-1}(t) = -\Phi^{-1}(t)A_q(t)\ (a.e.\ in\ T), \quad \Phi^{-1}(t_0) = I$$

where $A_q \in L^1_{n\times n}$. If $\int_{t_0}^{t_1} |A_q(t)|dt \leq c_0$ $(q \in \mathbf{N})$ for some $c_0 > 0$ then $\{\Phi_q^{-1}\}$ is bounded in $L^\infty_{n\times n}$.

Note that, by Lemmas 4.2 and 4.3, there exists $c_4 > 0$ such that

$$\max\{\|\Phi_q\|_\infty, \|\Phi_q^{-1}\|_\infty\} \leq c_4 \quad (q \in \mathbf{N}). \tag{11}$$

We are now in a position to prove the auxiliary result of Section 3.

Proof of Lemma 3.1:
(a): Observe that

$$\int_{t_0}^{t_1} \frac{|v_q(t)|^2}{w_q(t)^2}dt = 1 \quad (q \in \mathbf{N}). \tag{12}$$

Thus there exist $v_0 \in L^2(T;\mathbf{R}^m)$ and a 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 $h \in L^\infty(T;\mathbf{R}^m)$ and 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 (12),

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

Since $w_q(t)^2 \geq w_q(t) \geq 1$ for all $t \in T$, we have

$$0 \leq \int_{t_0}^{t_1} [w_q(t) - 1] dt \leq \int_{t_0}^{t_1} [w_q(t)^2 - 1] dt \leq \int_{t_0}^{t_1} \varphi(u_q(t) - u_0(t)) dt = D(u_q - u_0).$$

Observe also that

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

Consequently,

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

and so, since $h \in L^\infty(T; \mathbf{R}^m)$,

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

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.

(b): Let us first show that some subsequence of $\{\dot{y}_q / w_q\}$ converges weakly in $L^2(T; \mathbf{R}^n)$ to some function σ_0. Denote by $(L^2(T; \mathbf{R}^n))'$ the dual space of $L^2(T; \mathbf{R}^n)$ and let $f \in (L^2(T; \mathbf{R}^n))'$. By the Riesz Representation Theorem we have the existence of a unique $u_f \in L^2(T; \mathbf{R}^n)$ such that

$$\begin{aligned} f\left(\frac{\dot{y}_q}{w_q}\right) &= \int_{t_0}^{t_1} \left\langle u_f(t), \frac{\dot{y}_q(t)}{w_q(t)} \right\rangle dt \\ &= \int_{t_0}^{t_1} \left\langle A_q^*(t) u_f(t), \frac{y_q(t)}{w_q(t)} \right\rangle dt + \int_{t_0}^{t_1} \left\langle B_q^*(t) u_f(t), \frac{v_q(t)}{w_q(t)} \right\rangle dt. \end{aligned}$$

By (11), the fact that $\{v_q\}$ converges weakly in $L^1(T; \mathbf{R}^m)$, and Hölder's inequality, there exist $c_4, c_5 > 0$ such that, for all $q \in \mathbf{N}$ and all $t \in T$,

$$\begin{aligned} |y_q(t)| &\leq |\Phi_q(t)| \int_{t_0}^{t_1} |\Phi_q^{-1}(t) B_q(t) v_q(t)| dt \\ &\leq \|\Phi_q\|_\infty \cdot \|\Phi_q^{-1}\|_\infty \int_{t_0}^{t_1} |B_q(t) v_q(t)| dt \\ &\leq c_4^2 \cdot \|B_q\|_\infty \int_{t_0}^{t_1} |v_q(t)| dt \leq c_4^2 \cdot m_1 \cdot c_5. \end{aligned}$$

Thus there exists $c_6 > 0$ such that $\|y_q\|_\infty \le c_6$ ($q \in \mathbf{N}$). By Hölder's inequality,

$$\begin{aligned}\left|\int_{t_0}^{t_1}\left\langle A_q^*(t)u_f(t), \frac{y_q(t)}{w_q(t)}\right\rangle dt\right| &\le \int_{t_0}^{t_1}|A_q^*(t)u_f(t)|\cdot|y_q(t)|dt \\ &\le (t_1-t_0)^{1/2}\cdot\|A_q^*\|_\infty\cdot\|u_f\|_2\cdot\|y_q\|_\infty \\ &\le (t_1-t_0)^{1/2}\cdot m_0\cdot\|u_f\|_2\cdot c_6 \quad (q\in\mathbf{N}).\end{aligned}$$

Once again, by Hölder's inequality and by (12),

$$\begin{aligned}\left|\int_{t_0}^{t_1}\left\langle B_q^*(t)u_f(t), \frac{v_q(t)}{w_q(t)}\right\rangle dt\right| &\le \|B_q^*u_f\|_2\cdot\left\|\frac{v_q}{w_q}\right\|_2 \\ &\le \|B_q^*\|_\infty\cdot\|u_f\|_2 \le m_1\cdot\|u_f\|_2 \quad (q\in\mathbf{N}).\end{aligned}$$

Therefore $\{f(\dot y_q/w_q)\}_{q\in\mathbf{N}}$ is bounded in $\mathbf{R}$ for all $f \in (L^2(T;\mathbf{R}^n))'$ and hence $\{\dot y_q/w_q\}$ is bounded in $L^2(T;\mathbf{R}^n)$. This implies the existence of $c_7 > 0$ such that

$$\int_{t_0}^{t_1}\frac{|\dot y_q(t)|^2}{w_q(t)^2}dt \le c_7 \quad (q\in\mathbf{N}). \tag{13}$$

We conclude that 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.

By an argument similar to that used in the proof of (a), it follows that there is a 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 (13), given a measurable set $S \subset T$,

$$\left|\int_S \dot y_q(t)dt\right|^2 \le \int_S\frac{|\dot y_q(t)|^2}{w_q(t)^2}dt\int_S w_q(t)^2dt \le c_7\int_S w_q(t)^2dt \quad (q\in\mathbf{N}).$$

Moreover,

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

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

$$\int_{t_0}^{t_1}[w_q(t)^2-1]dt < \frac{\epsilon^2}{2c_7} \quad (q \ge q_\epsilon).$$

Choose $0 < \delta < \epsilon^2/2c_7$ such that

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

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

$$m(S) < \delta \Rightarrow \left|\int_S \dot{y}_q(t)dt\right|^2 \leq c_7\Big(m(S) + \int_{t_0}^{t_1}[w_q(t)^2 - 1]dt\Big) < \frac{\epsilon^2}{2} + \frac{\epsilon^2}{2} = \epsilon^2$$

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 integrals $\{\int_S \dot{y}_q(t)dt\}$ and hence also the sequence of functions $\{y_q(t)\}$ are equi-absolutely continuous on T. Consequently, $y_q(t) \to y_0(t)$ uniformly on T.

(c): By (a), there exists $c_5 > 0$ such that

$$\int_{t_0}^{t_1} |v_q(t)|dt \leq c_5 \quad (q \in \mathbf{N}).$$

Since $f_q(t) \to f(t)$ uniformly on T,

$$\lim_{q\to\infty} \int_{t_0}^{t_1} \langle f_q(t) - f(t), v_q(t)\rangle dt = 0.$$

Since $f \in L^\infty(T; \mathbf{R}^m)$, by (a),

$$\lim_{q\to\infty} \int_{t_0}^{t_1} \langle f_q(t), v_q(t)\rangle dt = \int_{t_0}^{t_1} \langle f(t), v_0(t)\rangle dt.$$

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

$$\sum_{i,j=1}^m [R_q^{ij}(t) - R_0^{ij}(t)]^2 w_q(t)^4 \leq 1.$$

Hence

$$M_q := \sup_{t\in T}\Big[\sum_{i,j=1}^m [R_q^{ij}(t) - R_0^{ij}(t)]^2 w_q(t)^4\Big]^{1/2} < \infty \quad (q \in \mathbf{N}).$$

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

$$|R_q(t; v_q(t)) - R_0(t; v_q(t))| \leq M_q \frac{|v_q(t)|^2}{w_q(t)^2}.$$

Since $R_q(t) \to R_0(t)$, and $w_q(t) \to 1$, both uniformly on T, we have $M_q \to 0$. Therefore, by (12),

$$\liminf_{q\to\infty} \int_{t_0}^{t_1} R_q(t; v_q(t))dt = \liminf_{q\to\infty} \int_{t_0}^{t_1} R_0(t; v_q(t))dt.$$

But for all $t \in T$,

$$R_0(t; v_q(t)) = R_0(t; v_0(t)) + 2\langle v_q(t) - v_0(t), R_0(t)v_0(t)\rangle + R_0(t; v_q(t) - v_0(t)).$$

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

$$\lim_{q\to\infty} \int_{t_0}^{t_1} \langle R_0(t)v_0(t), v_q(t) - v_0(t)\rangle dt = 0.$$

Hence

$$\liminf_{q\to\infty} \int_{t_0}^{t_1} R_q(t; v_q(t))dt \geq \int_{t_0}^{t_1} R_0(t; v_0(t))dt + \liminf_{q\to\infty} \int_{t_0}^{t_1} R_0(t; v_q(t) - v_0(t))dt.$$

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 Journal on Control and Optimization,* **41** (2002), 991-1014.

[2] M.R. Hestenes, Sufficient conditions for the problem of Bolza in the calculus of variations, *Transactions of the American Mathematical Society,* **36** (1934), 793-818.

[3] M.R. Hestenes, On sufficient conditions in the problems of Lagrange and Bolza, *The Annals of Mathematics,* **37** (1936), 543-551.

[4] M.R. Hestenes, A direct sufficiency proof for the problem of Bolza in the calculus of variations, *Transactions of the American Mathematical Society,* **42** (1937), 141-154.

[5] M.R. Hestenes, Sufficient conditions for the isoperimetric problem of Bolza in the calculus of variations, *Transactions of the American Mathematical Society,* **60** (1946), 93-118.

[6] M.R. Hestenes, The Weierstrass E-function in the calculus of variations, *Transactions of the American Mathematical Society,* **60** (1947), 51-71.

[7] M.R. Hestenes, An indirect sufficiency proof for the problem of Bolza in nonparametric form, *Transactions of the American Mathematical Society,* **62** (1947), 509-535.

[8] M.R. Hestenes, Sufficient conditions for multiple integral problems in the calculus of variations, *American Journal of Mathematics,* **70** (1948), 239-276.

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

[10] K. Malanowski, Sufficient optimality conditions for optimal control subject to state constraints, *SIAM Journal on Control and Optimization,* **35** (1997), 205-227.

[11] K. Malanowski, H. Maurer and S. Pickenhain, Second order sufficient conditions for state-constrained optimal control problems, *Journal of Optimization Theory and Applications,* **123** (2004), 595-617.

[12] H. Maurer and S. Pickenhain, Second order sufficient conditions for control problems with mixed control-state constraints, *Journal of Optimization Theory and Applications,* **86** (1995), 649-667.

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

[14] D.Q. Mayne, Sufficient conditions for a control to be a strong minimum, *Journal of Optimization Theory and Applications,* **21** (1977), 339-351.

[15] E.J. McShane, Sufficient conditions for a weak relative minimum in the problem of Bolza, *Transactions of the American Mathematical Society,* **52** (1942), 344-379.

[16] 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.

[17] J.F. Rosenblueth, Variational conditions and conjugate points for the fixed-endpoint control problem, *IMA Journal of Mathematical Control & Information,* **16** (1999), 147-163.

[18] G. Stefani and P.L. Zezza, Optimality conditions for a constrained optimal control problem, *SIAM Journal on Control & Optimization,* **34** (1996), 635-659.

Received: May, 2009