Journal of Global Optimization **13:** 109–122, 1998.

On Global Optimality Conditions for Nonlinear Optimal Control Problems

F.H. CLARKE[1], J.-B. HIRIART-URRUTY[2] and Yu.S. LEDYAEV[3]
[1]*Institut Desargues, Université Lyon I (Bâtiment 101), 69622 Villeurbanne, France (e-mail: clarke@jonas.univ-lyon1.fr);* [2]*Université Paul Sabatier, 31062 Toulouse, France (e-mail: jbhu@cict.fr);* [3]*Steklov Institute of Mathematics, Moscow 117966, Russia (e-mail: ledyaev@mi.ras.ru)*

(Received 2 December 1997; accepted 27 January 1998)

Abstract. Let a trajectory and control pair $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ maximize globally the functional $g(x(T))$ in the basic optimal control problem. Then (evidently) any pair $(\mathbf{x}, \mathbf{u})$ from the level set of the functional g corresponding to the value $g(\bar{x}(T))$ is also globally optimal and satisfies the Pontryagin maximum principle. It is shown that this necessary condition for global optimality of $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ turns out to be a sufficient one under the additional assumption of nondegeneracy of the maximum principle for every pair $(\mathbf{x}, \mathbf{u})$ from the above-mentioned level set. In particular, if the pair $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ satisfies the Pontryagin maximum principle which is nondegenerate in the sense that for the Hamiltonian H, we have along the pair $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$

$$\max_u H \not\equiv \min_u H \quad \text{on } [0, T],$$

and if there is no another pair $(\mathbf{x}, \mathbf{u})$ such that $g(x(T)) = g(\bar{x}(T))$, then $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ is a global maximizer.

Key words: Optimal control, Pontryagin maximum principle, Global optimality

1. Introduction

Let us consider a control system

$$\dot{x}(t) = f(x(t), u(t)), \quad x(0) = x_0, \tag{1.1}$$

where a measurable function $u : [0, T] \to \mathbb{U}$ with values in a compact set $\mathbb{U} \subset \mathbb{R}^l$ is called a *control* and an absolutely continuous function $x : [0, T] \to \mathbb{R}^n$ satisfying (1.1) is called a *trajectory* corresponding to the control $\mathbf{u}$. Any pair $(\mathbf{x}, \mathbf{u})$ satisfying (1.1) is called *an admissible pair*.

In this paper we provide a test to verify that a given admissible pair $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ is a global solution of the following *basic optimal control problem*.

PROBLEM $(\mathcal{P})$ *Maximize the functional*

$$g(x(T)) \tag{1.2}$$

over the set of all admissible pairs $(\mathbf{x}, \mathbf{u})$.

The best known test for global optimality stems from the Legendre-Carathéodory approach to sufficient conditions in the calculus of variations and is based on the *verification function* technique (for an account of the earlier application of this method in optimal control see Krotov, 1993).

In general, a differentiable verification function can fail to exist but it was shown (see Clarke, 1983) that there always exists a nonsmooth one. A particular choice is a certain value function, which establishes the link between the verification function technique and Bellman's dynamic programming approach (see Bellman, 1957). As with that approach, the method of verification functions becomes more problematic as the dimension n increases.

The alternative to the dynaming programming method for finding optimal solutions of optimal control problems is based upon the Pontryagin maximum principle (see Pontryagin et al., 1962) which gives necessary conditions satisfied by the optimal control $\bar{\mathbf{u}}$ and the trajectory $\bar{\mathbf{x}}$ in terms of the Hamiltonian (or Pontryagin pseudo-Hamiltonian in Clarke, 1983)

$$H(p, x, u) = \langle p, f(x, u)\rangle$$

and solution $\bar{\mathbf{p}}$ of the adjoint equation

$$\dot{\bar{p}}(t) = -H_x'(\bar{p}(t), \bar{x}(t), \bar{u}(t)), \quad \bar{p}(T) = g_x'(\bar{x}(T))$$

in the following way.

PONTRYAGIN MAXIMUM PRINCIPLE. *If a pair* $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ *is an optimal solution for the optimal control problem* ($\mathcal{P}$) *then for a.a. (almost all)* $t \in [0, T]$ *one has*

$$H(\bar{p}(t), \bar{x}(t), \bar{u}(t)) = \max_{u \in U} H(\bar{p}(t), \bar{x}(t), u).$$

The complexity of finding the trajectory and control satisfying the maximum principle does not increase essentially with the growth of the dimension n of the state vector. But we should note that, in general, the maximum principle is not a *sufficient* condition for global optimality of $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$. It is only a necessary condition for optimality and, moreover, it is a necessary condition for merely local optimality of $\bar{\mathbf{x}}$.

Nevertheless, the maximum principle becomes a sufficient condition for optimality of $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ under some additional assumptions about the control system (1.1) and the functional (1.2). For example, let us assume that the function g is concave and the system (1.1) is linear

$$\dot{x}(t) = Ax(t) + h(u), \tag{1.3}$$

where $h : \mathbb{U} \to \mathbb{R}^n$ is a continuous function. It is well known that in this case the Pontryagin maximum principle is a sufficient condition for global optimality of $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$.

In the present note we consider sufficient conditions for global optimality of admissible pairs $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ which are based on the maximum principle. But we make no structural assumptions on the problem's data such as linearity. Instead, conditions for global optimality are stated in terms of the level set $\mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ of the function g (1.2) consisting of all admissible pairs $(\mathbf{x}, \mathbf{u})$ of the control system (1.1) satisfying the equality

$$g(x(T)) = g(\bar{x}(T)). \tag{1.4}$$

It is clear that if $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ is a globally optimal admissible pair, then every pair $(\mathbf{x}, \mathbf{u})$ from $\mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ is also globally optimal. This implies that each such $(\mathbf{x}, \mathbf{u})$ satisfies the Pontryagin maximum principle:

$$H(p(t), x(t), u(t)) = \max_{u \in U} H(p(t), x(t), u), \quad \text{for a.a.} \quad t \in [0, T], \tag{1.5}$$

where $p(t)$ is the solution of the adjoint equation

$$\dot{p}(t) = -H'_x(p(t), x(t), u(t)), \quad p(T) = g'_x(x(T)). \tag{1.6}$$

Thus we have that a necessary condition for global optimality of $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ is the fulfillment of the maximum principle (1.5)–(1.6) for *any* pair $(\mathbf{x}, \mathbf{u})$ from the level set $\mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{u}})$.

The main result of this note asserts that this necessary condition for global optimality of the pair $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ becomes *sufficient* under the additional assumption that the maximum principle (1.5) is *nondegenerate* for each pair $(\mathbf{x}, \mathbf{u}) \in \mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{u}})$, namely: for any $(\mathbf{x}, \mathbf{u}) \in \mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ there exists $\tau \in [0, T]$ such that

$$\max_{u \in U} H(x(\tau), p(\tau), u) > \min_{u \in U} H(x(\tau), p(\tau), u). \tag{1.7}$$

This regularity assumption means that

$$\max_{u \in U} H(x(t), p(t), u) \not\equiv \min_{u \in U} H(x(t), p(t), u)$$

on $[0, T]$ and that the maximum principle (1.5) gives a non-trivial characterization of the control $\mathbf{u}$.

Strictly speaking, these sufficient conditions for global optimality are valid under the assumption that the attainability set

$$\mathcal{A}_T(x_0) = \{x(T) : \mathbf{x} \text{ is an arbitrary trajectory of (1.1)}\}$$

is closed . This is true , for example, when $f(x, u)$ is affine in u and $\mathbb{U}$ is convex, or for a linear system (1.3).

In the next section we use the concept of *relaxed control* (Warga, 1972) to formulate analogous necessary and sufficient conditions for global optimality in the general case of a nonlinear control system (1.1).

We should note that an approach to global optimality using level sets of the maximizing functional was suggested in Strekalovskii (1990) for the general problem of maximizing a *convex* functional ℓ on a set C. Necessary conditions from Strekalovskii (1990) for $\bar{x}$ to be a global maximizer become sufficient under the additional assumption that $\bar{x}$ is not a global minimum of ℓ on C. These conditions must be verified at any point x (even outside of C) such that $\ell(x) = \ell(\bar{x})$. This increase in the number of necessary conditions to verify is beneficial when they are used to eliminate $\bar{x}$ as a candidate for global maximizer. However, it is desirable to verify a lesser number of sufficient conditions to insure that $\bar{x}$ is a global maximizer.

Such conditions were suggested in Hiriart-Urruty and Ledyaev (1996) where the problem of maximizing a *general nonlinear* functional ℓ on a *convex* closed set C was considered. These sufficient conditions must be verified at points of a relevant level set *only in* C. Thus, even in the case of a convex functional ℓ, the number of sufficient conditions to verify is less than in Strekalovskii (1990).

Optimal control problems ($\mathcal{P}$) with *convex* functions g for general control systems(1.1) were considered in Strekalovskii (1995). Under the assumption that the attainability set $\mathcal{A}_T(x_0)$ is closed (it is assumed erroneously in Strekalovskii (1995, p.88) that lipschitzness of $f(x, u)$ is enough for this) the set of necessary conditions for global optimality is obtained (it includes relations stated in terms of points lying outside of the attainability set). It was shown, that under the additional hypothesis that an admissible pair in question $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ is not a global minimum, these necessary conditions become sufficient ones.

In this note the general optimal control problem with *general nonlinear* function g is considered. Since the set of relaxed controls is convex, the approach based on Hiriart-Urruty and Ledyaev (1996) can be adapted to this problem. The sufficient condition for global optimality is formulated in terms of the Pontryagin maximum principle and an additional mild nondegeneracy hypothesis for this maximum principle.

These results are also closely related to conditions for global optimality for the basic problem in calculus of variations such as obtained in Clarke et al. (1997).

2. Basic assumptions and main theorem

Let R^n be the n-dimensional euclidean space with inner product $\langle \cdot, \cdot \rangle$ and norm $|x| = \langle x, x \rangle^{1/2}$. The space of continuous functions $x : [0, T] \to \mathbb{R}^n$ equipped with the norm

$$\|\mathbf{x}\| = \max_{t \in [0,T]} |x(t)|$$

is denoted by C^n, the space of continuous functions $\phi : \mathbb{U} \to \mathbb{R}$ with the norm

$$\|\phi\|_C = \max_{u \in U} |\phi(u)|$$

is denoted by $C(U)$.

The following assumptions provide the existence and uniqueness of an absolutely continuous solution **x** of the equation (1.1) for any control **u**.

HYPOTHESIS A.

A1. *The function* $f(x,u)$ *and its partial derivative* $f'_x(u,v)$ *are continuous on* $\mathbb{R}^n \times U$.

A2. *There exists a constant* $a > 0$ *such that for any* $(x,u) \in \mathbb{R}^n \times \mathbb{U}$

$$\langle x, f(x,u)\rangle \leq a(1+|x|^2).$$

Such a solution is called a trajectory corresponding to the control **u**. In general, the set of all trajectories in C^n is not closed in C^n, so we cannot assert even the existence of an optimal solution in the optimal control problem $(\mathscr{P})$. To overcome this difficulty we use the concept of a relaxed control. This concept originated from L.C.Young's theory of generalized curves . In the context of classical optimal control theory it was introduced in Warga (1962).

Let $\mathbb{U} \subset R^p$ be a compact set , let frm $(\mathbb{U})$ denote the linear space of Radon measures m on $\mathbb{U}$, that is, finite regular Borel measures on $\mathbb{U}$. The weak norm $\|\cdot\|_w$ in frm $(\mathbb{U})$ is defined as follow

$$\|m\|_w = \sum_{i=1}^{\infty} \frac{1}{2^i(1+\|\phi_i\|_C)} \left| \int_{\mathbb{U}} \phi_i(u)m(du) \right|, \tag{2.1}$$

where $\{\phi_i\}_{i=1}^{\infty}$ is a dense countable subset of $C(\mathbb{U})$. The set $M = \text{rpm}(\mathbb{U})$ of Radon probability measures is convex and compact in the space (frm $(\mathbb{U}), \|\cdot\|_w)$. A measurable function $m : [0,T] \to M$ is called a *relaxed control*. It is known (see Warga, 1972; Krasovskii and Subbotin, 1974) that $m : [0,T] \to M$ is measurable if and only if the function

$$t \to \int_{\mathbb{U}} \phi(u)m(t|du)$$

is measurable for any $\phi \in C(\mathbb{U})$.

The set $\mathcal{M}$ of all relaxed controls is convex and sequentially weakly* compact (see Krasovskii and Subbotin, 1974). We recall that weak* convergence of the sequence of relaxed controls $\mathbf{m}_i$ to the relaxed control **m** means that

$$\lim_{i\to\infty} \int_0^T dt \int_{\mathbb{U}} \phi(t,u)m_i(t|du) = \int_0^T dt \int_{\mathbb{U}} \phi(t,u)m(t|du)$$

for any function $\phi(t,u)$ such that the mapping $t \in [0,T] \to \phi(t,\cdot) \in C(\mathbb{U})$ is measurable.

Under Hypothesis A, for an arbitrary relaxed control **m** there exists a unique solution **x** of the following equation

$$\dot{x}(t) = \tilde{f}(x(t), m(t)), \quad x(0) = x_0, \tag{2.2}$$

where

$$\tilde{f}(x, m) = \int_{\mathbb{U}} f(x, u)m(du).$$

The pair $(\mathbf{x}, \mathbf{m})$ defined by (2.2) is called an *admissible pair*.

The important property of the solution $\mathbf{x_m}$ of (2.2) corresponding to the relaxed control $\mathbf{m}$ is its continuous dependence upon $\mathbf{m}$ in the topology of weak* convergence on $\mathcal{M}$. Since the set of relaxed controls $\mathcal{M}$ is weakly* compact , this implies that the set of all trajectories corresponding to relaxed controls is compact in C^n. This implies that there always exists an optimal solution $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ for the following optimal control problem

PROBLEM $(\mathcal{P}_{rel})$ *Maximize the functional (1.2) over the set of all admissible pairs* $(\mathbf{x}, \mathbf{m})$ *for the system (2.2)* .

The connection between the original optimal control problem $(\mathcal{P})$ and the optimal control problem $(\mathcal{P}_{rel})$ is established by the fact that any trajectory $\mathbf{x_m}$ of the control system (2.2) can be approximated in C^n by trajectories $\mathbf{x_u}$ of (1.1), namely: for any $\mathbf{m} \in \mathcal{M}$ and any $\epsilon > 0$ there exists a control $\mathbf{u}$ such that

$$\|\mathbf{x_m} - \mathbf{x_u}\| < \epsilon.$$

Thus, if $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ is an optimal solution for the optimal control problem $(\mathcal{P})$, then the pair $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ is an optimal solution for the optimal control problem $(\mathcal{P}_{rel})$ where

$$\bar{m}(t) = \delta_{\bar{u}(t)},$$

and where $\delta_{\bar{u}}(du)$ is the Dirac measure concentrated on vector $\bar{u}$.

It was mentioned before that we could formulate conditions for global optimality without using the concept of relaxed control if we assume that the function $f(x, u)$ is affine in u and $\mathbb{U}$ is convex, or if the control system is linear (1.3). Note that the introduction of relaxed controls in (2.2) is a way to 'linearize' an original control system in the control variable since $\tilde{f}(x, m)$ in (2.2) is affine with respect to the new control variable m.

Define the level set of the functional (1.2) consisting of admissible pairs $(\mathbf{x}, \mathbf{m})$ for the control system (2.2) as follows:

$$\mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{m}}) = \{(\mathbf{x}, \mathbf{m}) : g(x(T)) = g(\bar{x}(T))\}. \tag{2.3}$$

For the pair $(\mathbf{x}, \mathbf{m})$ we define a solution $\mathbf{p}$ of the adjoint equation

$$\dot{p}(t) = -\tilde{H}'_x(p(t), x(t), m(t)), \quad p(T) = g'(x(T)), \tag{2.4}$$

where

$$\tilde{H}(p, x, m) = \langle p, \tilde{f}(x, m)\rangle.$$

Finally we make the following assumption about the function g in (1.2)

HYPOTHESIS B. *The function $g(x)$ is differentiable.*

The main result of this note is given by Theorem 2.1 below.

THEOREM 2.1. *Let Hypotheses A,B hold. If a pair $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ is globally optimal in the optimal control problem $(\mathcal{P}_{rel})$ then the following condition C1 holds:*

C1. For any $(\mathbf{x}, \mathbf{m}) \in \mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ and $\mathbf{p}$ defined by (2.4) one has

$$\tilde{H}(p(t), x(t), m(t)) = \max_{u \in U} H(p(t), x(t), u) \quad \textit{for a.a. } t \in [0, T] \tag{2.5}$$

The condition C1 is sufficient for optimality of a pair $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ if the following condition C2 of nondegeneracy of the maximum principle (2.5) holds:

C2. For any $(\mathbf{x}, \mathbf{m}) \in \mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ and $\mathbf{p}$ defined by (2.4) one has

$$\max_{u \in \mathbb{U}} H(x(t), p(t), u) \not\equiv \min_{u \in \mathbb{U}} H(x(t), p(t), u) \quad \textit{on } [0, T]. \tag{2.6}$$

Thus, summarily speaking, condition C1 is necessary for global optimality of $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ and under the additional condition C2 it becomes sufficient, as well.

Of course, the fact that conditions (2.5) and (2.6) should be verified for every admissible pair from the level set $\mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ complicates the practical application of such sufficient conditions. We consider now one special case when it is enough to check the relation (2.6) only for an admissible pair $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ itself.

Consider a linear control system (1.3) and a convex function g (1.2). In this case the optimal control problem is called *linear-convex*.

PROPOSITION 2.1. For a linear-convex optimal control problem, condition C2 holds for an admissible pair $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ whenever this pair is not a global minimizer of the functional $g(x(T))$.

Proof. Since $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ is not a global minimizer there exists an admissible pair $(\mathbf{x}_0, \mathbf{m}_0)$ such that

$$g(x_0(T)) < g(\bar{x}(T)).$$

Fix arbitrary $(\mathbf{x}, \mathbf{m})$ from $\mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ and let $p(t)$ satisfy (2.4). Then it follows from the convexity of g for the linear system (1.3) that

$$\begin{aligned}
&\int_0^T (\min_{u \in \mathbb{U}} H(p(t), x(t), u) - \max_{u \in \mathbb{U}} H(p(t), x(t), u))dt \\
&\quad \leq \int_0^T (\tilde{H}(p(t), x(t), m_0(t)) - \tilde{H}(p(t), x(t), m(t))dt \\
&\quad = \langle g'_x(x(T)), x_0(T) - x(T) \rangle \leq g(x_0(T)) - g(x(T)) \\
&\quad = g(x_0(T)) - g(\bar{x}(T)) < 0.
\end{aligned}$$

This strict inequality implies (2.6). □

It is well known that for a linear-convex optimal control problem an admissible pair $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ delivers a global minimum to functional (1.2) if and only if it satisfies the following minimum principle

$$\tilde{H}(\bar{p}(t), \bar{x}(t), \bar{m}(t)) = \min_{u \in \mathbb{U}} H(\bar{p}(t), \bar{x}(t), u) \quad \text{for a.a. } t \in [0, T]. \tag{2.7}$$

This means that if $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ does not satisfy the minimum principle, then due to Proposition 2.1 condition C2 holds for it. If, in addition, C1 holds for $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ then due to Theorem 2.1 $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ is a global maximizer.

It is clear that the converse is also valid for the linear-convex optimal control problem $(\mathcal{P})$ if it is non-trivial, in the sense that

$$\max\ g(x(T)) > \min\ g(x(T)), \tag{2.8}$$

where maximum and minimum are taken over the set of all admissible trajectories.

Indeed, let $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ be a global maximizer for such a problem ; then it is not a global minimizer and does not satisfy the minimum principle (2.7). Thus, we have proved

COROLLARY 2.1. *Let the linear-convex optimal control problem satisfy (2.8). Then the admissible pair* $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ *is global maximizer if and only if it does not satisfy the minimum principle (2.7) and condition C1 holds.*

Note that for the optimal control problem for a linear system (1.3) we can use an ordinary measurable control $\mathbf{u}$ instead of a relaxed one. Then the maximum principle (1.5) replaces (2.5) and should be verified for any admissible pair $(\mathbf{x}, \mathbf{u})$ satisfying (1.4).

3. Proof of the main theorem

If $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ is a globally optimal pair in the optimal control problem $(\mathcal{P}_{rel})$ then any $(\mathbf{x}, \mathbf{m}) \in \mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ is also globally optimal. Then condition C1 follows immediately from the classical Pontryagin maximum principle. Nevertheless we present here a proof of (2.5) based on the sliding variations method, not only for the sake of completeness of exposition, but also because this is a convenient way to introduce new notations.

For fixed $(\mathbf{x}, \mathbf{m}) \in \mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ we choose an arbitrary relaxed control $\mathbf{m}'$ and define a sliding variation $\mathbf{m}^\lambda$ of the relaxed control $\mathbf{m}$

$$m^\lambda(t) = (1 - \lambda)m(t) + \lambda m'(t) \tag{3.1}$$

where $\lambda \in [0, 1]$.

Due to convexity of $\mathcal{M}$ the function $\mathbf{m}^\lambda$ is a relaxed control and the trajectory $\mathbf{x}^\lambda$ of (2.2) corresponding to this control satisfies the following equation

$$\dot{x}^\lambda(t) = \tilde{f}(x^\lambda(t), m(t)) + \lambda(\tilde{f}(x^\lambda(t), m'(t)) - \tilde{f}(x^\lambda(t), m(t))), \quad x^\lambda(0) = x_0.$$

It is easy to see that

$$x^\lambda(t) = x(t) + \lambda\Delta(t) + r(\lambda, t),$$

where $\Delta(t)$ is a solution of the following 'equation in variations'

$$\dot{\Delta}(t) = \tilde{f}'_x(x(t), m(t))\Delta(t) + \tilde{f}(x(t), m'(t)) - \tilde{f}(x(t), m(t)), \quad \Delta(0) = 0, \tag{3.2}$$

and $r(\lambda, t)/\lambda$ converges to 0 uniformly on $[0, T]$ as $\lambda \downarrow 0$.

Since $(\mathbf{x}, \mathbf{m})$ is optimal we have that

$$\lim_{\lambda\downarrow 0} \frac{g(x^\lambda(T)) - g(x(T))}{\lambda} \leq 0.$$

In view of the representation for $x^\lambda(t)$ it implies that

$$0 \geq \langle g'(x(T)), \Delta(T)\rangle = \int_0^T [\tilde{H}(p(t), x(t), m'(t)) - \tilde{H}(p(t), x(t), m(t))]dt, \tag{3.3}$$

where $p(t)$ is defined in (2.4).

The last relation in (3.3) follows directly from properties of solutions of the adjoint equation (2.4) and the representation below for $\Delta(t)$

$$\Delta(t) = \int_0^t \Phi(t, s)[\tilde{f}(x(s), m'(s)) - \tilde{f}(x(s), m(s))]ds,$$

where $\Phi(t, s)$ is the fundamental matrix solution of linear equation

$$\dot{z}(t) = \tilde{f}'_x(x(t), m(t))z(t).$$

It is clear that for any Radon probability measure $m \in M$

$$\tilde{H}(p, x, m) \leq \max_{u\in\mathbb{U}} H(p, x, u).$$

It follows from Filippov Theorem on measurable selectors (see Warga, 1972; Clarke, 1983) that there exists measurable control $\mathbf{u}'$ such that

$$H(p(t), x(t), u'(t)) = \max_{u\in\mathbb{U}} H(p(t), x(t), u) \quad \text{for a.a. } t \in [0, T].$$

By choosing the relaxed control $m'(t) = \delta_{u'(t)}$ we derive from (3.3)

$$0 \geq \int_0^T [\max_{u\in\mathbb{U}} H(p(t), x(t), u) - \tilde{H}(p(t), x(t), m(t))]dt.$$

Because of the previous inequality we have that the integrand in this relation is non-negative, which implies that it equals 0 almost everywhere on $[0, T]$. Thus, $(\mathbf{x}, \mathbf{m})$ satisfies the maximum principle (2.5) and condition C1 is proved.

Now we prove that under condition C2 on the nondegeneracy of maximum principle the condition C1 is sufficient for global optimality of $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$.

Let us assume to the contrary that there exists an admissible pair $(\hat{\mathbf{x}}, \hat{\mathbf{m}})$ such that

$$g(\hat{x}(T)) > g(\bar{x}(T)). \tag{3.4}$$

Consider the following optimal control problem:
Minimize

$$J(\mathbf{m}') = \int_0^T \|m'(t) - \hat{m}(t)\|_w dt \tag{3.5}$$

over the set of all admissible pairs $(\mathbf{x}', \mathbf{m}')$ *of (2.2) satisfying*

$$g(x'(T)) \leq g(\bar{x}(T)). \tag{3.6}$$

Since the functional (3.5) is weakly* lower semicontinuous on $\mathcal{M}$ and $g(x'(T))$ continuously depends upon $\mathbf{m}'$, there exists an optimal solution of this problem (3.5)–(3.6). We denote the optimal solution for this problem by $(\mathbf{x}, \mathbf{m})$.

LEMMA 3.1. *Under condition (3.4) the pair* $(\mathbf{x}, \mathbf{m})$ *belongs to the level set* $\mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ *defined in (2.3).*

Let us assume that

$$g(x(T)) < g(\bar{x}(T)),$$

and choose the following variation $\mathbf{m}^\lambda$ of $\mathbf{m}$ for $\lambda \in [0, 1]$

$$m^\lambda(t) = (1 - \lambda)m(t) + \lambda\hat{m}(t).$$

Since the trajectory $\mathbf{x}^\lambda$ corresponding to the relaxed control $\mathbf{m}^\lambda$ continuously depends upon λ it follows from the previous inequality that for $\lambda \in (0, 1]$ small enough,

$$g(x^\lambda(T)) < g(\bar{x}(T)).$$

This implies that $(\mathbf{x}^\lambda, \mathbf{m}^\lambda)$ satisfies the constraint (3.6), but because of the obvious relation

$$J(\mathbf{m}^\lambda) = (1 - \lambda)J(\mathbf{m}) < J(\mathbf{m})$$

we conclude that $\mathbf{m}$ is not an optimal solution for the optimal control problem (3.5)–(3.6). This contradiction implies that the assertion of the Lemma is valid.

In order to derive the necessary optimality conditions characterizing the optimal solution $(\mathbf{x}, \mathbf{m})$ of the control problem (3.5)–(3.6), we consider the following control problem

Minimize

$$I(\mathbf{m}') = \max\{J(\mathbf{m}') - J(\mathbf{m}), g(x'(T)) - g(\bar{x}(T))\} \tag{3.7}$$

over the set of all admissible pairs $(\mathbf{x}', \mathbf{m}')$ *of (2.2).*

It is easy to see that the optimal solution $(\mathbf{x}, \mathbf{m})$ for the optimal control problem (3.5)–(3.6) is also an optimal solution for this control problem. This implies that

$$I'(\mathbf{m}; \mathbf{m}') := \lim_{\lambda \downarrow 0} \frac{I(\mathbf{m}^\lambda) - I(\mathbf{m})}{\lambda} \geq 0, \tag{3.8}$$

where the variation $\mathbf{m}^\lambda$ of $\mathbf{m}$ is defined by (3.1) for an arbitrary relaxed control $\mathbf{m}'$.

Now we consider a set Γ_0 of functions $\gamma(t, u)$ defined as follows

$$\gamma(t, u) = \sum_{i=1}^{\infty} \frac{1}{2^i(1 + \|\phi_i\|_C)} \rho_i(t)\phi_i(u),$$

where the measurable functions $\{\rho_i\}_{i\geq 1}$ satisfy the following relations for a.a. $t \in [0, T]$

$$-1 \leq \rho_i(t) \leq 1.$$

Thus, every element γ of Γ_0 is determined by the the measurable functions $\{\rho_i\}_{i\geq 1}$ which are called components. Let γ^n, $n = 1, 2, \ldots$, be a sequence of elements of Γ_0 with components $\{\rho_i^n\}_{i\geq 1}$.

By definition, the sequence γ^n converges to $\gamma \in \Gamma_0$ when for every $i \geq 1$ the sequence of measurable functions ρ_i^n weakly converges to ρ_i as $n \to \infty$. Note that for this notion of convergence the set Γ_0 is sequentially compact.

Let a set Γ consist of all $\gamma \in \Gamma_0$ such that

$$\begin{aligned}\rho_i(t) &\left(\int_{\mathbb{U}} \phi_i(u) m(t|du) - \int_{\mathbb{U}} \phi_i(u)\hat{m}(t|du) \right) \\ &= \left| \int_{\mathbb{U}} \phi_i(u) m(t|du) - \int_{\mathbb{U}} \phi_i(u)\hat{m}(t|du) \right|.\end{aligned}$$

For any $\gamma \in \Gamma$ one has

$$\begin{aligned}&\int_{\mathbb{U}} \gamma(t, u) m(t|du) - \int_{\mathbb{U}} \gamma(t, u)\hat{m}(t|du) \\ &\quad = \|m(t) - \hat{m}(t)\|_w \quad \text{for a.a. } t \in [0, T].\end{aligned} \tag{3.9}$$

Note that the set Γ is convex and sequentially compact, as well.

By using sequential compactness of Γ_0, Γ and a standard technique for finding directional derivatives of a *max*-like functional (3.7) (see Clarke, 1983) we can calculate the limit in (3.8) in terms of the set Γ

$$I'(\mathbf{m}; \mathbf{m}') = \max_{\alpha \in [0,1]} \max_{\gamma \in \Gamma} G(\alpha, \gamma, \mathbf{m}') \tag{3.10}$$

where

$$\begin{aligned} G(\alpha, \gamma, \mathbf{m}') :=& (\alpha \int_0^T (\int_U \gamma(t, u) m'(t|du) - \int_U \gamma(t, u) m(t|du)) dt \\ & + (1 - \alpha)\langle g'(x(T)), \Delta(T)\rangle \end{aligned}$$

and $\Delta(t)$ is a solution of the equation (3.2).

It follows from (3.8) and (3.10) that

$$\min_{\mathbf{m}' \in \mathcal{M}} \max_{\alpha \in [0,1]} \max_{\gamma \in \Gamma} G(\alpha, \gamma, \mathbf{m}') \geq 0.$$

Note that G is a linear functional of each of its variables when the two other variables are fixed. This means that we can use the non-symmetric minimax theorem from Borwein and Zhuang (1986):

$$\min_{\mathbf{m}' \in \mathcal{M}} \max_{\alpha \in [0,1]} \max_{\gamma \in \Gamma} G = \max_{\alpha \in [0,1]} \min_{\mathbf{m}' \in \mathcal{M}} \max_{\gamma \in \Gamma} G = \max_{\alpha \in [0,1]} \max_{\gamma \in \Gamma} \min_{\mathbf{m}' \in \mathcal{M}} G$$

to obtain from the previous inequality that there exists $\alpha \in [0, 1]$ and $\gamma \in \Gamma$ such that the following relation holds

$$\min_{\mathbf{m}' \in \mathcal{M}} G(\alpha, \gamma, \mathbf{m}') \geq 0.$$

This implies that for any $\mathbf{m}' \in \mathcal{M}$

$$\begin{aligned} 0 \leq& \alpha \int_0^T (\int_{\mathbb{U}} \gamma(t, u) m'(t|du) - \int_{\mathbb{U}} \gamma(t, u) m(t|du)) dt \\ & + (1 - \alpha) \int_0^T (\tilde{H}(p(t), x(t), m'(t)) - \tilde{H}(p(t), x(t), m(t))) dt \end{aligned} \tag{3.11}$$

where $p(t)$ is defined in (2.4). (Note that we used the representation for $\langle g'(x(T)), \Delta(T)\rangle$ contained in (3.3) to write G in the form (3.11)).

Recall that due to Lemma 3.1, $(\mathbf{x}, \mathbf{m})$ belongs to $\mathcal{L}_g(\bar{\mathbf{x}}, \bar{\mathbf{m}})$. This means that $(\mathbf{x}, \mathbf{m})$ satisfies the maximum principle (2.5) and the nondegeneracy condition (2.6).

We claim that

$$\alpha > 0. \tag{3.12}$$

Put $m'(t) = \delta_{u'(t)}$ where the control $\mathbf{u}'$ satisfies the relation

$$H(p(t), x(t), u'(t)) = \min_{u \in \mathbb{U}} H(p(t), x(t), u) \quad \text{for a.a. } t \in [0, T].$$

If $\alpha = 0$ then it follows from (3.11) for $\mathbf{m}'$ just defined and from the maximum principle (2.5) that

$$0 \leq \int_0^T [\min_{u \in \mathbb{U}} H(p(t), x(t), u) - \max_{u \in \mathbb{U}} H(p(t), x(t), u)]dt.$$

But this contradicts (2.6) for $(\mathbf{x}, \mathbf{m})$. This contradiction implies that (3.12) is true.

We use (3.12) to derive from (3.11) and the maximum principle (2.5) that one has for any $\mathbf{m}' \in \mathcal{M}$

$$\int_0^T \left(\int_{\mathbb{U}} \gamma(t,u)m(t|du) - \int_{\mathbb{U}} \gamma(t,u)m'(t|du) \right) dt$$
$$\leq \frac{1-\alpha}{\alpha} \int_0^T [\tilde{H}(p(t), x(t), m'(t)) - \max_{u \in \mathbb{U}} H(p(t), x(t), u)]dt \leq 0.$$

Then we replace $\mathbf{m}'$ by $\hat{\mathbf{m}}$ in this relation and use the property (3.9) of $\gamma(t,u)$ to obtain

$$\int_0^T \|\hat{m}(t) - m(t)\|_w dt \leq 0.$$

This means that $\hat{\mathbf{m}} = \mathbf{m}$ which contradicts (3.4) and the definition of $\hat{\mathbf{m}}$. Thus, the pair $(\bar{\mathbf{x}}, \bar{\mathbf{m}})$ is globally optimal.

Acknowledgments

Supported in part by Russian Fund for Fundamental Research Grant 96-01-00219, by the NSERC of Canada and FCAR du Québec.

References

Bellman, R. (1957), *Dynamic Programming*, Princeton University Press, Princeton.

Borwein, J.M. and Zhuang, D. (1986), On Fan's minimax theorem, *Mathematical Programming* 34: 232–234.

Clarke, F.H. (1983), *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York. Third edition in: *Classics in Applied Mathematics*, volume 5. SIAM, Philadelphia, 1990.

Clarke, F.H., Hiriart-Urruty, J.-B. and Ledyaev, Yu.S. (1997), On global optimality in the calculus of variations, *Nonlinear Analysis* 28: 1187–1192.

Hiriart-Urruty, J.-B. and Ledyaev, Yu.S. (1996), A note on the characterization of the global maxima of a (tangentially) convex function over a convex set, *J. Convex Analysis* 3: 55–61.

Krasovskii N.N. and Subbotin, A.I. (1974), *Positional Differential Games*, Nauka, Moscow [in Russian]. French translation: (1979) *Jeux differentiels*, Editions Mir, Moscow. Revised English translation: (1988) *Game-Theoretical Control Problems*, Springer Verlag, Berlin / New York.

Krotov, V.F. (1993), Global methods in optimal control theory, pp. 74-121, in A.B. Kurzhanskii (ed.), *Advances in Nonlinear Dynamics and Control: A Report from Russia, Progress in Systems and Control Theory, 17*, Birkhäuser, Boston.

Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V. and Mishchenko E.F. (1962), *The Mathematical Theory of Optimal Processes*, Interscience, New York.

Strekalovskii, A.S. (1990), On problems of global extremum in nonconvex extremal problems, *Izvestia Vyshikh Uchebnykh Zavedenii* 74–80. English translation: (1990), *Soviet Mathematics. (Iz.VUZ)*, 34: 82–90.

Strekalovskii, A.S. (1995), On global maximum of a convex terminal functional in optimal control problems, *J. Global Optimization.* 7: 75–91.

Warga, J. (1962), Relaxed variational problems, *J. Mathematical Analysis and Applications* 4: 111–128.

Warga, J. (1972), *Optimal Control of Differential and Functional Equations*, Academic Press, New York.