Control and Cybernetics

vol. **23** (1994) No. 1/2

Stability and convergence in nonlinear control

by

Andrzej Nowakowski

Institute of Mathematics
Łódź University
Łódź
Poland

Nonsmooth optimal control problem is considered. To study it a smooth approximation is proposed. Stability and convergence to the original problem using field theory and dynamic programming is investigated.

1. Introduction

We consider the optimal control problem

$$\text{Minimize } g(x(1)) \tag{1.1}$$

subject to

$$\dot{x}(t) \in F(t, x(t)) \qquad \text{a.e.}, \tag{1.2}$$

$$x(0) = x_0 \tag{1.3}$$

expressed in terms of:

- a nonempty subset $\Omega \subset [0,1] \times X$, X – a Hilbert space with the norm $\|\cdot\|$
- function $F(\cdot,\cdot)$ with domain Ω which takes as values subsets of X,
- a point $x_0 \in X$ and
- a function $g(\cdot) : \{x : (1,x) \in \Omega\} \to R$.

An absolutely continuous function $x : I \to X$ where I is a subinterval of $[0,1]$ with right end 1 which satisfies (1.2) and has its graph in Ω is an admissible trajectory.

The basic hypotheses (H1) we assume are the following:

(i) g is lower semicontinuous in X;

(ii) $F(\cdot,\cdot)$ takes as values nonempty weakly compact subset of X and is continuous in the sense that

$$dist(F(t',x'),F(t,x)) \to 0 \text{ if } (t',x') \to (t,x) \text{ in } \Omega;$$

(iii) there exists a constant k such that for any t, x, x'

$$dist(F(t,x'),F(t,x)) \leq k\|x' -' x\|$$

whenever the left–hand side is defined ($dist(A,B)$ — the Hausdorff distance)

(iv) there exists a constant r such that

$$|F(t,x)| = \sup\{\|v\| : v \in F(t,x)\} \leq r \text{ for all } (t,x) \in \Omega.$$

We define the Hamiltonian function $H : \Omega \times X \to R$

$$H(t,x,p) = \max\{\langle p,v\rangle;\ v \in F(t,x)\} \tag{1.4}$$

where $\langle\cdot,\cdot\rangle$ is a scalar product in X; we identify X with its dual. Under our hypotheses $(x,p) \to H(t,x,p)$ is locally Lipschitzian, for each t, on its domain of definition.

In order to motivate what follows we state here conditions which in many cases are considered as first order necessary optimality conditions (see e.g. Clarke, Vinter 1983).

We say that an admissible trajectory $x(t)$ satisfies the first order necessary conditions if there exist an absolutely continuous function $y(t)$ with values in X and a number $-1 \leq y^0 \leq 0$ such that

$$(-\dot{y}(t), \dot{x}(t)) \in \partial H(t, x(t), y(t)) \qquad \text{a.e. in } [0,1], \tag{1.5}$$

$$\begin{aligned} &(-y(1), y^0) \quad \text{is normal to epi } g \text{ at the point } (x(1), g(x(1))), \\ &y^0 + |y(1)| \quad \text{is nonzero,} \end{aligned}$$

where ∂H refers to the generalized gradient of $(x,y) \to H(t,x,y)$ for each fixed t (see Clarke 1983), epi g means the epigraph of g.

From (1.4) we see that, in general, $H(t,\cdot,\cdot)$ cannot be more smooth than it is indicated. To study our problem by more refined method we need H to be more smooth. This is why we propose to approximate H by a better function e.g. by such a function for which we are able to solve corresponding equation of type (1.5) directly or numerically. What we then need is to prove stability of our approximation and its convergence. Thus we take a smooth function near H, we write for it equations (1.5) and next we show that under additional geometric assumptions on a family of solutions of that equations we find trajectories which approximate a solution of problem (1.1)–(1.3).

Therefore let $H(t,x,y,h)$ be a family of real functions defined in $\Omega \times X \times W$, where W is a set of parameters h, such that $H(t,\cdot,\cdot,h)$ is at least of C^2 in the set of its definition for all $t \in [0,1]$, $h \in W$, $t \to H(t,x,y,h)$ is measurable for $x, y \in X$, $h \in W$,

$$\sup_{x,y} |H(t,x,y) - H(t,x,y,h)| \le \varepsilon(h), \qquad t \in [0,1], \ \ h \in W \tag{1.6}$$

and let $g(x,h) : X \times W \to R$ be such that $g(\cdot,h)$ is of C^1,

$$|g(x) - g(x,h)| \le \varepsilon(h), \qquad x \in X, \ \ h \in W \tag{1.7}$$

where $\varepsilon(h)$ are positive real numbers, such that if h tends to zero in any sense, then $\varepsilon(h) \to 0$.

Of course, we could first approximate our problem (1.1)–(1.3) and then require that the corresponding Hamiltonian (1.4) be a smooth function, however, from the practical point of view, it seems to us that our approach is more convenient.

Further, we shall study the solutions of the equations

$$-\dot{y}(t) = H_x(t, x(t), y(t), h) \qquad \text{a.e.,} \tag{1.8}$$

$$\dot{x}(t) = H_y(t, x(t), y(t), h) \qquad \text{a.e.,} \tag{1.9}$$

$$y(1) = y^0 g_x(x(1), h) \tag{1.10}$$

$$H(t, x(t), y(t), h) = y(t) H_y(t, x(t), y(t), h). \tag{1.11}$$

On the basis of Nowakowski (1988), we describe a dual approach to the field of extremal and Hilbert integral for equations (1.8)–(1.11) and we show that some members of the field approximate a minimum of our problem (1.1)–(1.3). In the last section we describe dynamic programming approach to the same problem.

2. General notions

An admissible trajectory $x(t)$ defined in the appropriate subinterval of $[0,1]$ with right end at 1 will be termed a line of flight (briefly l.f.), if there exist along $x(t)$ a conjugate function $y(t)$, absolutely continuous in t with values in X, and a number $-1 \le y^0 \le 0$ such that $\|y(t)\| + |y^0|$ is nonvanishing and the trio $x(t), y(t), y^0$ satisfies (1.8)–(1.11).

We introduce a new coordinate $x^0 = g(x(1))$ where $x(1)$ is a value of an admissible trajectory at 1. For a given l.f. $x(t)$, $x^0 = g(x(1), h)$ and we put $z(t) = (-x^0, x(t))$ and $p(t) = (y^0, y(t))$ for the corresponding conjugate function $y(t)$ the number y^0. Then the pair $z(t), p(t)$ will be called canonical pair (compare Nowakowski 1988, p. 736), (briefly c.p.).

Further, denote by $P \subset R^2 \times X$ a set covered by graphs of $p(t)$ such that $z(t), p(t)$ is a c.p., which in the sequel may be reduced to a smaller one; let $T \subset R \times X$ denote a set covered by graphs of corresponding l.f. $x(t)$.

If $(t_0, p_0) \in P$, then we write $V(t_0, p_0)$ for the value of

$$\langle -x_0(t_0), y_0(t_0) \rangle + y_0^0 g(x_0(1), h) = \langle -z_0(t_0), p_0(t_0) \rangle_z \tag{2.1}$$

where $z_0(t) = (-x_0^0, x_0(t))$, $p_0(t) = (y_0^0, y_0(t))$ is a c.p. such that $p_0(t_0) = p_0$. Of course, the map $(t,p) \to V(t,p)$ in P might be a multifunction and this is why we assume the following hypothesis:

(H2) the set P is such that the map $(t,p) \to V(t,p)$ is single–valued in P.

For each $(t_0, x_0) \in T$, $P(t_0, x_0)$ denote the set of values of all those $p(t)$ at t_0 for which $z(t) = (-x^0, x(t))$, $p(t)$ is some c.p. and $x(t_0) = x_0$. It is natural to expect that $P(t,x)$, $(t,x) \in T$, may not be single–valued. This is why by $p(t,x) \in P(t,x)$, $(t,x) \in T$, will be denoted single–valued selections of $P(t,x)$. Let us fix $h \in W$ and set $\varepsilon = \varepsilon(h)$,

$$f(t,x,y) = H_y(t,x,y,h), \qquad (t,x) \in \Omega, \quad y \in X, \tag{2.2}$$

$$L(t,x,y) = yf(t,x,y) - H(t,x,y,h), \qquad (t,x) \in \Omega, \quad y \in X. \tag{2.3}$$

To study any family of arcs of c.p. depending on a parameter σ, let us define on an open set $G \subset Y$, Y is another Hilbert space, a pair of continuous functions $t^-(\sigma)$, $t^+(\sigma)$, $0 \le t^-(\sigma) < t^+(\sigma) \le 1$, $\sigma \in G$. We assume that $t^+(\sigma)$ is C^1 in G. We further suppose that G is a projection of a certain set $\tilde{G} \subset \tilde{M}$, $\tilde{M}$ is a metric space, whose elements will be denoted by (σ, ρ). $\tilde{G}$ does not have to be necessarily open; instead of that, we assume that the operation of the projection is standard. The operation of projecting $\tilde{G}$ onto G is standard if the following condition is satisfied (see Young 1969, p. 266):

> given any point $(\sigma_0, \rho_0) \in \tilde{G}$, and any small enough curve $\gamma \subset G$ which issues from σ_0, there exists on γ a continuous function $\rho(\sigma)$ such that $\rho(\sigma_0) = \rho_0$ and that all points of the form $(\sigma, \rho(\sigma))$ for $\sigma \in \gamma$ lie in $\tilde{G}$.

Let $S^- = \{(t,\sigma) : t = t^-(\sigma) \ge 0,\ \sigma \in G\}$, $S = \{(t,\sigma) : t^-(\sigma) < t < t^+(\sigma),\ \sigma \in G\}$, $S^+ = \{(t,\sigma) : t = t^+(\sigma) \le 1,\ \sigma \in G\}$, $[S] = S^- \cup S \cup S^+$. Similarly, we denote by S^{*-}, S^*, S^{*+} the sets of (t,σ,ρ) for which t satisfies the same conditions as in S^-, S, S^+, respectively, and $(\sigma,\rho) \in \tilde{G}$; $[S^*] = S^{*-} \cup S^* \cup S^{*+}$.

3. Canonical spray

First of all, we shall construct a family Σ of arcs of c.p. depending on parameters (σ, ρ) described by functions

$$z(t,\sigma),\ p(t,\sigma,\rho), \qquad (t,\sigma) \in S, \quad (t,\sigma,\rho) \in S^*$$

for which the study of this family is the nearest to the classical considerations. The definition of the functions $z(t,\sigma)$, $p(t,\sigma,\rho)$ will be supposed extended to the sets $[S]$, $[S^*]$, respectively. The sets of pairs (t,x) where $x = x(t,\sigma)$ with (t,σ) belonging to $S^-, S, S^+, [S]$ will be denoted by $E^-, E, E^+, [E]$, respectively; $E^{*-}, E^*, E^{*+}, [E^*]$ will denote the sets of values of pairs $(t, p(t,\sigma,\rho))$

with (t,σ,ρ) in $S^{*-}, S^*, S^{*+}, [S^*]$, whereas those of pairs $(t, z(t,\sigma))$ with (t,σ) in $S^-, S, S^+, [S]$ will be $D^-, D, D^+, [D]$.

Finally, we write (when $(t,\sigma,\rho) \in [S^*]$)

$$\tilde{L}(t,\sigma,\rho), \tilde{f}(t,\sigma,\rho), \tilde{V}^+(\sigma)$$

for the expressions

$$L(t, x(t,\sigma), y(t,\sigma,\rho)), f(t, x(t,\sigma), y(t,\sigma,\rho)), V(t^+(\sigma), p(t^+(\sigma),\sigma,\rho(\sigma))).$$

We assume the following hypotheses on the family Σ:

(H3) The function $z(t,\sigma)$ is C^1 in $[S]$. For given (σ_0,ρ_0) in $\tilde{G}$ and any small neighbourhood $G_0 \subset G$ of σ_0, there exists in G_0 a function $\rho(\sigma)$ such that $\rho(\sigma_0) = \rho_0$, all points $(\sigma, \rho(\sigma))$ for $\sigma \in G_0$ lie in $\tilde{G}$, and $\bar{p}(t,\sigma) = p(t,\sigma,\rho(\sigma))$ is C^1 in $F(\sigma_0,\rho_0) = \{(t,\sigma) : t^-(\sigma) \le t \le t^+(\sigma),\ \sigma \in G_0\}$.

(H4) For each $(\sigma_0,\rho_0) \in \tilde{G}$ the functions $\bar{L}(t,\sigma) = \tilde{L}(t,\sigma,\rho(\sigma))$, $\bar{f}(t,\sigma) = \tilde{f}(t,\sigma,\rho(\sigma))$ are continuous in $F(\sigma_0,\rho_0)$ and they have continuous derivatives $\bar{L}_\sigma, \bar{f}_\sigma$ there.

(H5) The maps $S^- \to D^-$, $S \to D$ defined by $(t,\sigma) \to (t, z(t,\sigma))$ have the following property: given any arc $C_z \subset D^-$ (or $C_z \subset D$) with the description $t_1 \le \tau \le t_2$, $(-x^0, x(\tau))$ where $x(\tau)$ is an arc of the admissible trajectory $x(t)$, $x^0 = g(x(1))$, issuing from $(t_1, z(t_1,\sigma_1))$, there exists a rectifiable curve $\Gamma \subset S^-$ (or $\Gamma \subset S$) issuing from (t_1,σ_1) such that every small arc of C_z issuing from $(t_1, z(t_1,\sigma_1))$ is the image under the map $(t,\sigma) \to (t, z(t,\sigma))$ of a small arc of Γ issuing from (t_1,σ_1).

(H6) For each fixed $(\sigma,\rho) \in \tilde{G}$ and for $x = x(t,\sigma)$ we have: for each $t' \in (t^-(\sigma), t^+(\sigma))$ and each vector $(\alpha,\beta) \in R \times Y$, $\beta \in Y$, $\alpha^2 + \|\beta\|^2 = 1$, there exists a function $\alpha(t)$ of bounded variation, defined in $[t', t^+(\sigma)]$ with values $\alpha(t') = \alpha$, $\alpha(t) \in R$ for $t \in (t', t^+(\sigma))$, $\alpha(t^+(\sigma)) = t^+_\sigma(\sigma)\beta$, such that

$$|\bar{L}_\sigma(t,\sigma)\beta| \le -\varepsilon\left(\frac{\partial}{\partial t}\right)((1 + \|x_t(t,\sigma)\|^2)(\alpha(t))^2 + \|x_\sigma(t,\sigma)\beta\|^2)^{1/2}$$

for almost all t in $[t', t^+(\sigma)]$. (We assume that the derivative on the right–hand side of the last inequality exists.)

The hypothesis (H6) is used to approximate our original problem (1.1)–(1.3). Notice that, in view of (1.6), (1.9), (1.11) and (2.2) and Theorem 2.2 from Ekeland (1974), hypothesis (H6) is not essentially strong. It is formulated in that form for our convenience in calculations.

If hypotheses (H3)–(H6) together with those on $t^-(\sigma), t^+(\sigma)$, G, $\tilde{G}$ are satisfied, the family Σ is called canonical spray.

For $(t,x) \in [E]$ let $P_\Sigma(t,x)$ denote the sets of values of $p(t,\sigma)$ at those $(t,\sigma) \in S$ for which $x(t,\sigma) = x$.

REMARK 3.1 *Let C_z denote any small arc contained in D^- or D, with the description $t_1 \le \tau \le t_2$, $(-x^0, x(\tau))$ where $x(t)$, $t \in [0,1]$, is an admissible trajectory with $x(0) = x_0$, $x^0 = g(x(1))$, issuing from $(t_1, z(t_1, \sigma_1))$. We also represent C_z in terms of its arc length s as $t = t(s)$, $z = z(s) = (-x^0, x(s))$, $s \in [0, s_{C_z}]$. Let further Γ denote a rectifiable curve in S^- or S such that small arc of C_z issuing from $(t_1, z(t_1, \sigma_1))$ are, in accordance with (H5), the images under the map $(t, \sigma) \to (t, z(t, \sigma))$ of small arcs of Γ issuing from (t_1, σ_1). We represent Γ in terms of its arc length λ by functions $\bar{t}(\lambda), \bar{\sigma}(\lambda)$, so that the point (t_1, σ_1) corresponds to $\lambda = 0$. We can then define a continuous increasing function $s(\lambda)$ having its inverse $\lambda(s)$, which satisfies the relation*

$$t(s(\lambda)) = \bar{t}(\lambda), \qquad z(s(\lambda)) = z(\bar{t}(\lambda), \bar{\sigma}(\lambda)).$$

In turn, let C_p be the image under the map $(t, \sigma) \to (t, \bar{p}(t, \sigma))$ of Γ issuing from

$$(t_1, \bar{p}(t_1, \sigma_1)) = (t_1, p_1) = (t_1, y_1^0, y_1)$$

where $\bar{p}(t, \sigma) = p(t, \sigma, \rho(\sigma))$ with $\rho(\sigma)$ suitably chosen in accordance with (H3). We easily see that to small arcs of Γ issuing from (t_1, σ_1) there correspond small arcs of C_p issuing from (t_1, p_1). Thus we can express the final points of the small arcs of C_p as a function of s $(t(s), p(s))$. Denote by (t_2, p_2) the terminal point of C_p which corresponds to that of C_z $(t_2, -x^0, x(t_2))$.

To simplify further considerations we assume, *for this section only* , the following hypothesis.

(H_{S3}) We are given any point $(\sigma_0, \rho_0) \in S^{*+}$ and any sufficiently small curve $\gamma \subset G$ which issues from σ_0 with the description $\sigma(\lambda)$, $\lambda \in [0, v]$; $\sigma(\lambda)$ is a Lipschitz function, $\sigma(0) = \sigma_0$, 0 is the point of approximate continuity of $\sigma(\lambda)$. Then

$$|\langle \bar{p}\ (t^+(\sigma_0), \sigma_0), z_\sigma(t^+(\sigma_0), \sigma_0)\sigma_\lambda(0)\rangle_z| \le \varepsilon((1 + \|x_t(t^+(\sigma_0), \sigma_0)\|^2) \cdot$$
$$\cdot (t_\sigma^+(\sigma_0)\sigma_\lambda(0))^2 + \|x_\sigma(t^+(\sigma_0), \sigma_0)\sigma_\lambda(0)\|^2)^{1/2}.$$

LEMMA 3.1 *Let C_z, C_p be one of the arcs described in Remark 3.1. Then, along C_p, $V(t, p)$ is bounded. There exists along C_z Borel measurable function $p_\Sigma(t, x) \in$*
$P_\Sigma(t, x)$. Moreover, the functions $p_\Sigma(t, x)$, $f(t, x, y_\Sigma(t, x))$ are bounded along C_z.

PROOF. By the definition of C_p it is the image under the map $(t, \sigma) \to (t, \bar{p}(t, \sigma))$ of some $\Gamma \subset [S]$. Therefore we can treat $V(t, p)$ as a function $V(t, \bar{p}(t, \sigma))$ along Γ which by (H3) is continuous. The graph of Γ is a compact set in $[0, 1] \times Y$. Hence $V(t, p)$ is bounded along C_p. Applying the measurable selection theorem from Castaing, Valadier (1977) to the multifunction $(t, x) \to \{(t, \sigma) \in \Gamma : x(t, \sigma) = x\}$

defined on $x(\tau)$, $t_1 \leq \tau \leq t_2$, and putting this selection into $\bar{p}(t,\sigma)$ we obtain the function $p_\Sigma(t,x)$ as it is required in the assertion of the lemma. The proof of the last assertion is analogous to the first one.

LEMMA 3.2 *Let Γ denote any small rectifiable curve in $[S]$ with (t_0,σ_0) as the initial point and (t_1,σ_1) as the terminal one. Then there exist a set $G_0 \subset G$ and $\rho(\sigma)$ in G_0 (see (H3)) such that*

$$\int_\Gamma \tfrac{d}{dt}\langle z(t,\sigma),\bar{p}(t,\sigma)\rangle_z dt + \tfrac{d}{d\sigma}\langle z(t,\sigma),p(t,\sigma)\rangle_z d\sigma = \\ = V(t_0,\bar{p}(t_0,\sigma_0)) - V(t_1,\bar{p}(t_1,\sigma_1))$$

where $\langle z,p\rangle_z = -x^0y^0 + \langle x,y\rangle$ and $\bar{p}(t,\sigma) = p(t,\sigma,\rho(\sigma))$.

The proof follows directly from the definition of the function $V(t,p)$.

LEMMA 3.3 *On each arc of the canonical pair of Σ we have: for each $t' \in (t^-(\sigma),t^+(\sigma))$ and each vector $(\alpha,\beta) \in R \times Y$, $\alpha^2 + \|\beta\|^2 = 1$, there exists a function $\alpha(t)$ of bounded variation, defined in $[t',t^+(\sigma)]$, with values $\alpha(t') = \alpha$, $\alpha(t) \in R$ for $t \in (t',t^+(\sigma))$, $\alpha(t^+(\sigma) = t^+_\sigma(\sigma)\beta$, such that*

$$\left| \left(\tfrac{\partial}{\partial t}\right) \langle p(t,\sigma,\rho), z_\sigma(t,\sigma)\beta\rangle_z \right| \leq \\ \leq -\varepsilon\left(\frac{\partial}{\partial t}\right)((1+\|x_t(t,\sigma)\|^2)(\alpha(t))^2 + \|x_\sigma(t,\sigma)\beta\|^2)^{1/2} \tag{3.1}$$

for almost all t in $[t',t^+(\sigma)]$.

PROOF. Let (t',σ',ρ') be any point of S^* and $z'(t) = (-x^0,x'(t))$, $p'(t) = (y^{0'},y'(t))$ the corresponding values of the functions $z(t,\sigma')$, $p(t,\sigma',\rho')$, $t \in [t',t^+(\sigma))$. By performing indifferent orders to the operations of integration in t and differentiation in σ on relation (1.9) and taking notation (2.2), we get the following relation

$$\frac{\partial}{\partial t}x_\sigma(t,\sigma) = \bar{f}_\sigma, \tag{3.2}$$

calculated at the point (t,σ',ρ'), $t \in [t',t^+(\sigma))$. From (1.8) we obtain at (t,σ',ρ') for almost all t in $[t',t^+(\sigma))$ and $\beta \in Y$

$$\langle\frac{\partial}{\partial t}y'(t), x_\sigma(t,\sigma)\beta\rangle = -\langle y'(t), H_x(t,x'(t),y'(t))x_\sigma(t,\sigma)\beta\rangle, \tag{3.3}$$

and by the definition of y^0, we have at this point

$$-\frac{\partial}{\partial t}y^{0'}x^0_\sigma(\sigma)\beta = 0, \qquad -\frac{\partial}{\partial t}x^0_\sigma(\sigma)\beta y^{0'} = 0 \tag{3.4}$$

We add both sides of the last three equalities with (3.2) multiplied by $y'(t)$. As a result we obtain at the same (t,σ',ρ')

$$\frac{\partial}{\partial t}\langle p(t,\sigma,\rho), z_\sigma(t,\sigma)\beta\rangle_z = \bar{L}_\sigma(t,\sigma)\beta.$$

Taking into account hypothesis (H6), we obtain (3.1).

COROLLARY 3.1 *Let any point $(\sigma_0,\rho_0)\in S^*$ and any $\alpha\in[-1,1]$ be given. Let $\gamma\subset G$ be any sufficiently small curve which issues from σ_0, with the description $\sigma(\lambda)$, $\lambda\in[0,v]$, and such that $\sigma(\lambda)$ is a Lipschitz function $\sigma(0)=\sigma_0$, 0 is the point of approximate continuity of $\sigma_\lambda(\lambda)$ and $\alpha^2+\|\sigma_\lambda(0)\|^2=1$. Then*

$$\begin{aligned}|\langle p(t,\sigma_0,\rho_0), z_\sigma(t,\sigma_0)\sigma_\lambda(0)\rangle_z| &\le \varepsilon((1+\|x_t(t,\sigma_0)\|^2)\alpha^2+\\ &+\|x_\sigma(t,\sigma_0)\sigma_\lambda(0)\|^2)^{1/2}\end{aligned} \tag{3.5}$$

for all t in $(t^-(\sigma_0),t^+(\sigma_0))$.

PROOF. Let $t'\in(t^-(\sigma_0),t^+(\sigma_0))$ be arbitrarily fixed and let $\beta=\sigma_\lambda(0)$. Integrating (3.1) and using (H_{S3}) we find (3.5).

THEOREM 3.1 *Let C_z and C_p be as described in Remark 3.1. Then the following relation*

$$\begin{aligned}|V(t_1,p_1)\ &-V(t_2,p_2)-\langle x(t_2),y_2\rangle+\langle x(t_1),y_1\rangle+x^0y_2^0-x^0y_1^0\\ &-\textstyle\int_{C_z}\langle y_\Sigma(t,x), f(t,x,y_\Sigma(t,x))\rangle dt-\langle p_\Sigma(t,x),dz\rangle_z|\\ &\le\varepsilon\textstyle\int_{t_1}^{t_2}(1+\|\dot{x}(t)\|^2)^{1/2}dt\end{aligned} \tag{3.6}$$

holds for some $p_\Sigma(t,x)$, $(t,x)\in[E]$.

PROOF. Let $e(s)=\left(\frac{dt}{ds},\frac{dz}{ds}\right)$ stand for the direction of the tangent to C_z defined for a.e. s in $[0,s_{C_z}]$. Let s_0 be any point in $[0,s_{C_z}]$ such that $e(s)$ and $p(s)$ are approximately continuous at it. We set $t_0=t(s_0)$, $x_0=x(s_0)$, $e_0=e(s_0)$, $\dot{t}_0=dt(s_0)/ds$, $\dot{z}_0=dz(s_0)/ds$. Let $p_0=(y_0^0,y_0)$ be any admissible vector from the set $P(t_0,x_0)$ such that $p(t_0,\sigma_0)=p_0$ for any (t_0,σ_0) belonging to the graph of Γ. We also put $f_0=f(t_0,x_0,y_0)$ and let λ_0 be such that $\sigma_0=\bar{\sigma}(\lambda_0)$.

Denote by γ a sufficiently small arc Γ issuing from (t_0,σ_0) defined in the interval $I=[\lambda_0,\lambda_2]$ of values of λ, i.e. the functions $\bar{t}(\lambda)$, $\bar{\sigma}(\lambda)$ are restricted now to the interval I. Denote by ΔV the difference in $V(t,p)$ at the ends of a small arc C_p issuing from (t_0,p_0) and being the image of γ, and by Δs the corresponding difference in s. By Corollary 3.1 and taking into account Remark 3.1 we obtain

$$\begin{aligned}\left|\int_\gamma\langle p,z_\sigma\rangle_z d\sigma\right| &\le\varepsilon\int_I((1+\|x_t(\bar{t}(\lambda),\bar{\sigma}(\lambda))\|^2)(\bar{t}_\lambda(\lambda))^2\\ &+\|x_\sigma(\bar{t}(\lambda),\bar{\sigma}(\lambda))\bar{\sigma}_\lambda(\lambda)\|^2)^{1/2}d\lambda.\end{aligned}$$

Using Lemma 3.2 we further infer

$$\begin{aligned}\Big|\Delta V &+ \langle x(\bar{t}(\lambda_2)), y(\bar{t}(\lambda_2), \bar{\sigma}(\lambda_2), \rho(\bar{\sigma}(\lambda_2)))\rangle - \langle x(t_1), y_1\rangle - x^0 y^0(\bar{\sigma}(\lambda_2), \rho(\bar{\sigma}(\lambda_2))) \\ &+ x^0 y_1^0 + \int_I \Big[\langle y(\bar{t}(\lambda), \bar{\sigma}(\lambda), \rho(\bar{\sigma}(\lambda))), \bar{f}(\bar{t}(\lambda), \bar{\sigma}(\lambda))\rangle \frac{dt}{ds} \\ &- \Big\langle \bar{p}(\bar{t}(\lambda), \bar{\sigma}(\lambda)), \frac{d}{ds} z\Big\rangle_z\Big] \, ds(\lambda)\Big| \leq \varepsilon \Delta s.\end{aligned} \tag{3.7}$$

Since $p(t, \sigma, \rho(\sigma)) = \bar{p}(t, \sigma)$, ($\rho(\sigma)$ being suitably chosen), $\bar{f}(t, \sigma)$ are continuous on γ we deduce that they are bounded on I. This, along with the last inequality, imply the uniform boundedness of the ratio $\Delta V/\Delta s$ for all sufficiently small Δs. Thus $s \to V(t(s), p(s))$ is locally Lipschitz. If we show that

$$\lim_{\Delta s \to 0} \frac{1}{\Delta s} \int_I \Big[\langle y, \bar{f}\rangle \frac{dt}{ds} - \Big\langle \bar{p}, \frac{dz}{ds}\Big\rangle_z\Big] ds(\lambda) = [\langle y_0, f_0\rangle \dot{t}_0 - \langle p_0, \dot{z}_0\rangle_z], \tag{3.8}$$

then (3.6) will follow from (3.7). But to prove (3.8) it is enough to repeat the argumentation from the proof of Lemma 25.3 in Young 1969, vol. II, p. 274.

In order to be able to take into consideration more than one spray of c.p. we need one more hypothesis:

(H7) The map $S^{*-} \to E^{*-}$ defined by $(t, \sigma, \rho) \to (t, p(t, \sigma, \rho))$ is descriptive in the following sense: given any sufficiently small rectifiable curve $C \subset E^{*-}$ issuing from $(t_0, p(t_0, \sigma_0, \rho_0))$, there exists a sufficiently small rectifiable curve $\Gamma \subset S^-$ issuing from (t_0, σ_0) such that every small arc of C issuing from $(t_0, p(t_0, \sigma_0, \rho_0))$ is the image under the map $(t, \sigma) \to (t, p(t, \sigma, \rho(\sigma)))$ of a small arc of Γ issuing from (t_0, σ_0) where $\rho(\sigma)$ is as in (H3).

For $(t, p) \in [E^*]$, let $Z_\Sigma(t, p)$ stand for the set of values of $z(t, \sigma)$ at those $(t, \sigma, \rho) \in [S^*]$ for which $p(t, \sigma, \rho) = p$. Similarly as Lemma 3.1 we obtain the following lemma.

LEMMA 3.4 *Let C be a rectifiable curve lying, together with its terminal points, in E^{*-} Then, along C, $V(t, p)$ is bounded and there exists along it a Borel measurable function $z_\Sigma(t, p) \in Z_\Sigma(t, p)$. Moreover, the functions $z_\Sigma(t, p)$, $f(t, x_\Sigma(t, p), y)$ are bounded along it.*

We put $\bar{V}(s) = V(t(s), p(s))$ along any rectifiable curve C in E^{*-}, with the arc length description $t = t(s)$, $p = p(s)$, $0 \leq s \leq s_C$.

THEOREM 3.2 *The function $\bar{V}(s)$ is absolutely continuous along C and, for almost all s in $[0, s_C]$,*

$$\Big|\frac{d}{ds}\bar{V}(s) + \langle y(s), f(t(s), x_\Sigma(t(s), p(s)), y(s))\rangle \frac{dt}{ds} + \Big\langle z_\Sigma(t, p), \frac{dp}{ds}\Big\rangle_z\Big| \leq \varepsilon \tag{3.9}$$

for each single-valued selection $z_\Sigma(t, p)$ of $Z_\Sigma(t, p)$.

PROOF. The proof is similar to the proof of Theorem 3.1 (see also the proof of Theorem 1' in Nowakowski 1988). Thus we only sketch it. For convenience, we assume that 0 is a point of approximate continuity of the derivative $\left(\frac{dt}{ds}, \frac{dp}{ds}\right)$ of the function $(t(s), p(s))$. Denote by Γ a rectifiable curve in S such that small arcs of C issuing from $(t(0), p(0))$ are, in accordance with (H7), the images under the map $(t, \sigma) \to (t, \bar{p}(t, \sigma))$ of small arcs γ of Γ issuing from $(t(0), \sigma_0)$, where σ_0 is such that $\bar{p}(t(0), \sigma_0) = p(0)$. Let now $t = \bar{t}(v)$, $\sigma = \bar{\sigma}(v)$, $v \in [0, v_\gamma]$, be the arc length parametric description of γ, such that the point $(t(0), \sigma_0)$ should correspond to the value of 0. Define a continuous increasing function $s = s(v)$, $v \in [0, v_\gamma]$, such that $s(0) = 0$ which satisfies in $[0, s_\gamma]$ the relations

$$t(s(v)) = \bar{t}(v), \qquad p(s(v)) = \bar{p}(\bar{t}(v), \bar{\sigma}(v)). \tag{3.10}$$

Denote by Δs and ΔV the corresponding difference in s and in $\bar{V}(s)$ at the ends of a small arc of C issuing from $(t(0), p(0))$, being the image of γ. By Corollary 3.1

$$|\langle \bar{p}(\bar{t}(v), \bar{\sigma}(v)), z_\sigma(\bar{t}(v), \bar{\sigma}(v))\, \bar{\sigma}_v(v)\rangle_z| \le \varepsilon((1 + \|x_t(\bar{t}(v), \bar{\sigma}(v))\|^2)(\bar{t}_v(v))^2 + \\ + \|x_\sigma(\bar{t}(v), \bar{\sigma}(v))\bar{\sigma}_v(v)\|^2)^{1/2}, \qquad v \in [0, v_\gamma].$$

Hence, and from Lemma 3.2, we conclude, taking account of (3.10), that

$$\Big|\Delta V + \int_\gamma \frac{d}{dt}\langle z(t,\sigma), \bar{p}(t,\sigma)\rangle_z dt + \frac{d}{d\sigma}\langle z(t,\sigma), \bar{p}(t,\sigma)\rangle_z d\sigma \\ - \int_\gamma \langle \bar{p}(t,\sigma), z_\sigma(t,\sigma) d\sigma\rangle\Big| \le \varepsilon \Delta s$$

and further

$$\Big|\Delta V + \int_0^{v_\gamma} \Big(\langle (\bar{t}(v), \bar{\sigma}(v), \rho(\bar{\sigma}(v))), \bar{f}(\bar{t}(v), \bar{\sigma}(v))\rangle \frac{dt}{ds} \\ + \Big\langle z(\bar{t}(v), \bar{\sigma}(v)), \frac{d}{ds} p\Big\rangle_z \Big) ds(v)\Big| \le \varepsilon \Delta s.$$

Proceeding quite analogously as in the corresponding part of Theorem 3.1, we find the assertion of the theorem.

4. A chain of c.p.

In the preceding section we described and discussed a fixed spray of c.p. Σ. However, the family of l.f. defined in Section 2 may consist of a greater number of sprays of c.p. satisfying conditions (H3)–(H7), whose graphs of trajectories are contained in T.

We recall (see Young 1969, vol. II, §27) that a finite or countable sequence of sprays of c.p.

$$\Sigma_1, \Sigma_2, \dots, \Sigma_N, \dots$$

will be termed a chain of c.p. if, for $i = 1, 2, \ldots, N, \ldots$, they fit together in the inverse order so that the set E_i^{*-} corresponding to Σ_i contains E_{i+1}^{*+} corresponding to Σ_{i+1}.

Now, we are in position to prove that in a given chain of c.p. hypothesis H_{S3} is satisfied in each S_i^{*+}, $i = 1, 2, \ldots$.

LEMMA 4.1 *In each* S_1^{*+} $\langle p, z_{\sigma_1}\rangle_z \equiv 0$.

PROOF. Be the definition of c.p. and (1.10) $y(1, \sigma_1, \rho_1) = y^0(\sigma_1, \rho_1) g_x(x(1, \sigma_1), h)$. If we multiply the last equality by $x_{\sigma_1}(1, \sigma_1)$, then we obtain the assertion of the lemma.

LEMMA 4.2 *We are given any point* $(\sigma_2^0, \rho_2^0) \in S_2^{*+}$ *and any sufficiently small curve* $\gamma \subset G_2$ *which issues from* σ_2^0 *with description* $\sigma_2(v)$, $v \in [0, v_\gamma]$; $\sigma_2(v)$ *is a Lipschitz function,* $\sigma_2(0) = \sigma_2^0$, *0 is the point of approximate continuity of* $\sigma_2(v)$. *Then*

$$|\langle p(t^+(\sigma_2^0), \sigma_2^0, \rho_2^0), z_{\sigma_2}(t^+(\sigma_2^0), \sigma_2^0)\sigma_{2v}(0)\rangle_z| \leq$$
$$\leq \varepsilon((1 + \|x_t(t^+(\sigma_2^0), \sigma_2^0)\|^2 (t_{\sigma_2}^+(\sigma_2^0)\sigma_{2v}(0))^2 + \tag{4.1}$$
$$+\|x_\sigma(t^+(\sigma_2^0), \sigma_2^0)\sigma_{2v}(0)\|^2)^{1/2}).$$

PROOF. By (H3) there exists $G_{20} \supset \gamma$ and $\rho_2(\sigma_2)$ in G_{20} such that $\rho_2(\sigma_2^0) = \rho_2^0$ and $\bar{p}(t, \sigma_2) = p(t, \sigma_2, \rho_2(\sigma_2))$. Let C be the image of γ in E_2^{*+} under the map $(t, \sigma_2) \to (t, \bar{p}(t, \sigma_2))$ with ends (t_0, p_0), (t_1, p_1). Since $E_2^{*+} \subset E_1^{*-}$ therefore C is a rectifiable curve lying, together with its terminal points, in E_1^{*-}. This is why we can apply to it Theorem 3.2. Integrating (3.9) along C we get for any single-valued selection $z_{\Sigma_1}(t, p)$ of $Z_{\Sigma_1}(t, p)$, $(t, p) \in E^{*-}$

$$\left| V(t_1, p_1) - V(t_0, p_0) + \int_C \langle y, f(t, x_{\Sigma_1}(t, p), y)\rangle dt + \langle z_{\Sigma_1}(t, p), dp\rangle_z \right| \leq \varepsilon l, \tag{4.2}$$

where l is the length of C. Taking into account that C is the image of γ and $x_t(t, \sigma_2) = \bar{f}(t, \sigma_2)$ we further find from (4.2)

$$\left| \int_\gamma \langle \bar{p}(t, \sigma_2), z_{\sigma_2}(t, \sigma_2) d\sigma_2\rangle_z \right| \leq$$
$$\leq \varepsilon \int_0^{v_\gamma} ((1 + \|x_t(t^+(\sigma_2(v)), (\sigma_2(v))\|^2)(t_{\sigma_2}^+(\sigma_2(v))^2 + \tag{4.3}$$
$$+\|x_{\sigma_2}(t^+(\sigma_2(v)), (\sigma_2(v))\sigma_{2v}(v)\|^2)^{1/2} dv.$$

Dividing both sides of (4.3) by v_γ and contracting γ to the initial point $\sigma_2(0)$ we obtain (4.1).

Using the induction from Lemmas 4.1, 4.2 and Corollary 3.1 we infer the following proposition.

PROPOSITION 4.1 *For each* S_i^{*+}, $i = 1, 2, \ldots, N \ldots$, *of a given chain of c.p. the assertions of hypothesis* H_{S3} *are satisfied.*

5. A concourse of c.p.

The concept of a concourse of c.p. originates from L.C. Young (1969), vol. II, §28 where there are many details on it. Here we only give a sketch of this theory to formulate further results.

Denote by T_n, $n = 1, 2, \ldots$, a finite or countable system of disjoint subsets of T whose union is T and such that each T_n should be a subset of some E_i or E^- of a chain or a subset of a few such sets of different chains. Let N be the family of all arcs of admissible trajectories $x(t)$, $t \in [0,1]$ such that $x(0) = x_0$ and their graphs are contained in T. An arc from N will be called a fragment if its interior portion lies in some T_n. The class of such fragments is denoted by N_0. We need a form of an addition of fragments — the fusion of curves C_1 and C_2 from N_0: if the final point of C_1 is the initial point of C_2, we term fusion of C_1, C_2 a curve C made up of two adjucent arcs, consisting of C_1 and C_2, in that order. If the class N can be derived from N_0 by a finite fusion, then the set T will be termed the unimpaired union of the sets T_n.

A concourse of c.p. is a finite or countable infinite system of chains of c.p. such that T is the unimpaired union of the sets of the type E_i^-, E_i of these chains.

Suppose that a concourse of c.p. exists. Let C_x denote any arc of an admissible trajectory $x(t)$, $t \in [0,1]$, such that $x(0) = x_0$ and the graph of $x(t)$ is contained in T. We assume C_x defined in $[t_1, t_2] \subset [0,1]$ and for $x(t)$ we set $x^0 = g(x(1))$. Define C_z as an arc with the description $t_1 \le t \le t_2$, $(-x^0, x(t))$. By hypothesis, there is a decomposition of T into disjoint subsets T_n, each of which is a subset of some sets of the type E_j^-, E_j of the chains of c.p. of our concourse. We define the families N and N_0 as above. Of course, our C_x belongs to N. Denote further by $\bar{C}_x$ a subarc of C_x defined in $[\bar{t}_1, \bar{t}_2]$ which belongs to N_0; $\bar{C}_z$ is a subarc of C_z corresponding to $\bar{C}_x$. Let Σ be any spray of c.p. of one of our chains such that $\bar{C}_x$ meets either the set E^- or the set E of Σ, i.e. $\bar{C}_x$ lies in some T_n wholly contained in E^- or in E. In accordance with (H5) and Remark 3.1, there is a rectifiable curve $\bar{C}_p$ corresponding to the arc $\bar{C}_z$, contained in the set E^{*-} or E^* of Σ, with ends $(\bar{t}_1, \bar{p}_1)$, $(\bar{t}_2, \bar{p}_2)$. Hence, by Theorem 3.1, we have equality (3.6) for $\bar{C}_z$, $\bar{C}_p$. The arc C_x is a finite fusion of members of N_0, thus there is a rectifiable curve C_p corresponding to C_z with ends $(t_1, p_1) = (t_1, y_1^0, y_1)$, $(t_2, p_2) = (t_2, y_2^0, y_2)$ and for which (3.6) still holds. In this manner we have proved the following theorem.

THEOREM 5.1 *With the above hypothesis and notations, the relation*

$$\Big| V(t_1, p_1) - V(t_2, p_2) - \langle x(t_2), y_2\rangle + \langle x(t_1), y_1\rangle + x^0 y_2^0 - x^0 y_1^0 - \\ - \int_{t_1}^{t_2} \left(\langle y(t, x(t)), f(t, x(t), y(t, x))\rangle - \langle y(t, x(t)), \dot{x}(t)\rangle\right) dt \Big| \le \\ \le \varepsilon \int_{t_1}^{t_2} (1 + \|\dot{x}(t)\|^2)^{1/2} dt \tag{5.1}$$

holds for some single-valued selection $p(t,x)$ of $P(t,x)$, $(t,x) \in T$.

Denote by $G(0,x_0)$ the set of values of $g(x(1),h)$ for all l.f. of a concourse of c.p.

THEOREM 5.2 *Suppose that a concourse of c.p. exists and that there exists c.p. $z_\varepsilon(t), p_\varepsilon(t)$, $t \in [0,1]$, $x_\varepsilon(0) = x_0$ being a member of our concourse of c.p. and $x_\varepsilon^0 = g(x_\varepsilon(1),h) = \min G(0,x_0)$. Let $K(x_0)$ be the set of those $x(1)$ for which the graph of admissible trajectories $x(t)$, $x(0) = x_0$ are contained in T. Then*

$$\begin{aligned} -y_\varepsilon^0(g(x_\varepsilon(1)) - g(x(1)) &\le \varepsilon \left(3 + \int_0^1 (1 + \|\dot{x}(t)\|^2)^{1/2} dt\right) \\ &\le \varepsilon \left(3 + \sqrt{1+r^2}\right) \end{aligned} \tag{5.2}$$

for all $x(1)$ in $K(x_0)$.

PROOF. We apply Theorem 5.1. Let $x(t)$, $t \in [0,1]$, $x(0) = x_0$, $x(1) \in K(x_0)$ be given. Put $x^0 = g(x(1))$, C_z is the arc $(-x^0, x(t))$, $t \in [0,1]$, C_p corresponding to it rectifiable curve in P with ends $p_\varepsilon(0)$, $p_2 = (y_2^0, y_2)$. From (5.1) we obtain for our case

$$\begin{aligned} &-x_\varepsilon^0 y_\varepsilon^0 + x_\varepsilon(0) y_\varepsilon(0) + g(x(1),h) y_2^0 - x(1) y_2 + x(1) y_2 - x(0) y_\varepsilon(0) - x^0 y_2^0 + \\ &+ x^0 y_\varepsilon^0 + \int_0^1 (\langle y(t,x(t)), f(t,x(t),y(t,x(t)))\rangle - \langle y(t,x(t)), \dot{x}(t)\rangle) dt \\ &\le \varepsilon \int_0^1 (1 + \|\dot{x}(t)\|^2)^{1/2} dt. \end{aligned} \tag{5.3}$$

By (1.4), (1.6) and (1.11) the integral in (5.3) is greater or equal to $-\varepsilon$, by (1.7) $|g(x(1),h) - x^0| \le \varepsilon$. Thus from (5.3) we get

$$-y_\varepsilon^0(g(x_\varepsilon(1),h) - g(x(1))) \le \varepsilon(2 + \int_0^1 (1 + \|\dot{x}(t)\|^2)^{1/2} dt)$$

and since $|g(x_\varepsilon(1),h) - g(x_\varepsilon(1))| \le \varepsilon$ we obtain (5.2).

6. The dynamic programming approach

Let all assumptions written down in Section 1 be fulfilled. In addition suppose Ω to be an open set. Let $x(t)$, $t \in [0,1]$, $x(0) = x_0$, be an admissible trajectory. A result customarily associated with the name of Carathéodory, but which has appeared in a variety of guises virtually from the inception of the calculus of variations (see Young 1969, Chapt. 1) provides a sufficient condition that $x(t)$ be optimal, expressed in terms of a solution to the Hamilton–Jacobi equation

$$G_t(t,x) + H(t,x,-G_x(t,x)) = 0 \tag{6.1}$$

with boundary condition

$$G(1,x) = g(x). \tag{6.2}$$

The sufficient condition (applied under suitable conditions on $F(\cdot,\cdot)$) is that there exists a continuously differentiable function $G(\cdot,\cdot)$ which satisfies (6.1) for all $(t,x) \in \Omega$ and (6.2) for all x belonging to the projection of Ω onto X, and is such that

$$G(0,x_0) = g(x(1)).$$

The question arises: how widely applicable is the Carathéodory condition? If we consider the condition essentially as stated above then the answer is disappointing. It is easy to construct examples of problems in modern control theory whose solutions cannot be characterized in this way. The main reason is that the function $H(t,\cdot,\cdot)$ is merely Lipschitz continuous and so solutions to (6.1) are often at most Lipschitz continuous. This is why in the last case the nonsmooth analysis appears to be very fruitful (see e.g. Clarke, Vinter 1983). However, one question remains far to be solved: how to find a solution to the generalization of equation (6.1) :

$$\min_{(\alpha,\beta)\in\partial G(t,x)} \{\alpha + \min_{e\in F(t,x)} \{\langle e,\beta\rangle\}\} = 0.$$

(The last equation is studied in Clarke, Vinter 1983.)

We propose a different approach. First, we propose to study the equation

$$G_t(t,x) - H(t,x,-G_x(t,x),h) = 0, \qquad (t,x) \in \Omega \tag{6.3}$$

with boundary condition

$$G(1,x) = g(x,h). \tag{6.4}$$

$H(t,x,y,h)$ is a smooth function of (x,y) described in Section 1 and we can choose it in such a way that we are able to solve (6.3), (6.4) directly or at least to assert that $G(\cdot,\cdot)$ is of C^1 or that there exists a numerical solution of (6.3), (6.4) (see e.g. Fleming 1969). Next, using (1.6), we easily check that this G satisfies the following inequality (with original $H(t,x,y)$!)

$$-\varepsilon(h) \le G_t(t,x) - H(t,x,-G_x(t,x)) \le \varepsilon(h), \qquad (t,x) \in \Omega. \tag{6.5}$$

Define now a new function

$$G_\varepsilon(t,x) = G(t,x) + \varepsilon(h)(1-t). \tag{6.6}$$

Then it satisfies

$$-2\varepsilon(h) \le G_{\varepsilon t}(t,x) - H(t,x,-G_{\varepsilon x}(t,x)) \le 0, \qquad (t,x) \in \Omega. \tag{6.7}$$

It turns out (see Proposition 6.2) that $G_\varepsilon(t,x)$ is an ε-value function in Ω i.e. it satisfies

$$S(t,x) \le G_\varepsilon(t,x) \le S(t,x) + 3\varepsilon(h), \qquad (t,x) \in \Omega, \tag{6.8}$$

$$g(x) - \varepsilon(h) \le G_\varepsilon(1,x) \le g(x) + \varepsilon(h), \qquad (1,x) \in \Omega, \tag{6.9}$$

where $S(t,x) = \inf\{g(x(1)) : x(\tau),\ \tau \in [t,1],$ admissible trajectory, $x(t) = x\}$ is the value function. If we find an admissible trajectory $x_\varepsilon(t)$, $t \in [0,1]$, $x_\varepsilon(0) = x_0$ satisfying

$$G_\varepsilon(0,x_0) \ge g(x_\varepsilon(1)), \tag{6.10}$$

then we call it ε–optimal trajectory associated with $G_\varepsilon(t,x)$.

Therefore the above simple procedure allows us to find an approximate solution to our problem (1.1)–(1.3).

We begin with the reformulation, in terms of ε–functions, of known propositions from dynamic programming (see e.g. Fleming, Rishel 1975).

Proposition 6.1 *Let $K(t,x)$ be any real-valued function defined in Ω such that $K(1,x) = g(x,h)$. Let $(t_0,x_0) \in \Omega$ be a given initial condition, and suppose that for each admissible trajectory $x(t)$, $t \in [t_0,1]$, $x(t_0) = x^0$, $K(t,x)$ is finite in $[t_0,1]$ and*

$$K(t_1,x_1) \le K(t_2,x_2) + 2\varepsilon(h) \tag{6.11}$$

for each $t_0 \le t_1 \le t_2 \le 1$. If an admissible trajectory $x_\varepsilon^0(t)$, $t \in [0,1]$, $x_\varepsilon^0(0) = x_0$ is such that

$$K(0,x_0) \ge g(x_\varepsilon^0(1)) \tag{6.12}$$

and

$$K(t,x_\varepsilon^0(t)) \le g(x_\varepsilon^0(1)) + 3\varepsilon(h), \qquad 0 < t \le 1, \tag{6.13}$$

then $x_\varepsilon^0(t)$ is an optimal trajectory for $G_\varepsilon(t,x) = K(t,x)$.

Proof. Let $x(t)$, $t \in [0,1]$, $x(0) = x_0$ be any admissible trajectory. Then

$$K(0,x_0) \le g(x(1)) + 3\varepsilon(h).$$

Thus $K(0,x_0) \le S(0,x_0) + 3\varepsilon(h)$. For $x_\varepsilon^0(t)$, $K(0,x_0) \ge g(x_\varepsilon^0(1))$, so $x_\varepsilon^0(t)$ is an ε–optimal trajectory for $G_\varepsilon(t,x) = K(t,x)$.

Proposition 6.2 *Let $K(t,x)$, $(t,x) \in \Omega$ be a C^1 solution to the following inequality*

$$-2\varepsilon(h) \le K_t(t,x) - H(t,x,-K_x(t,x)) \le 0, \qquad (t,x) \in \Omega, \tag{6.14}$$

which satisfies the boundary condition

$$K(1,x) = g(x,h), \qquad (1,x) \in \Omega.$$

If $x_\varepsilon(t)$, $t \in [0,1]$, $x(0) = x_0$ *is an admissible pair such that*

$$-2\varepsilon(h) \le K_t(t, x_\varepsilon(t)) - H(t, x_\varepsilon(t), -K_x(t, x_\varepsilon(t))) \le 0, \quad t \in [0,1], \tag{6.15}$$

then $x_\varepsilon(t)$ *is an optimal trajectory for the* ε*–value function* $G_\varepsilon(t,x) = K(t,x)$.

PROOF. By (6.14) and (1.4) for an admissible trajectory $x(t)$

$$\left(\frac{d}{dt}\right) K(t,x(t)) = K_t(t,x(t)) + \langle K_x(t,x(t)), \dot{x}(t)\rangle \ge -2\varepsilon(h). \tag{6.16}$$

Integrating (6.13) in $[t_1, t_2]$ we obtain (6.11) and along $x_\varepsilon(t)$ in $[0,1]$ we get (6.12) and (6.13). Thus $x_\varepsilon(t)$ is an ε–optimal trajectory for $G_\varepsilon(t,x) = K(t,x)$.

7. Stability and convergence

In Sections 2–6 we described two procedures of calculating an approximate solution to problem (1.1)–(1.3). The question which appears in natural way is how the solutions behave when $\varepsilon(h) \to 0$ as h tends to zero. This is just the problem of stability of our approximations.

Let W be any topological space containing zero and such that convergence to zero of elements of W makes sense. We assume $H(t,x,y,0) = H(t,x,y)$, ($H(t,x,y,h)$ and $H(t,x,y)$ are those from Section 1) and $\varepsilon(h) \to 0$ as $h \to 0$. Then condition (1.6) means that $H(t,x,y,h) \to H(t,x,y)$ as $h \to 0$ uniformly with respect to (t,x,y). For each $H(t,x,y,h)$ satisfying (1.6) by each of two procedures we can calculate an ε–optimal trajectory $x_{\varepsilon(h)}(t)$.

We say that our approximation is stable if for each $\varepsilon > 0$ there exists $M > 0$ such that for each $H(t,x,y,h_1), H(t,x,y,h_2)$ satisfying (1.6) with $\varepsilon(h) = \varepsilon$ and corresponding to them $x_{\varepsilon(h_1)}(t), x_{\varepsilon(h_2)}(t)$

$$|g(x_{\varepsilon(h_1)}(1)) - g(x_{\varepsilon(h_2)}(1))| \le M\varepsilon. \tag{7.1}$$

It turns out that what we have proved in Theorem 5.2 and Proposition 6.2 (see also (6.7)) is just the stability of approximation (1.6). In the first case we assume for our convenience that $y_\varepsilon^0 = -1$ in (5.2).

PROPOSITION 7.1 *The procedures described in Sections 2–5 and Section 6 are stable.*

PROOF. Let $x_{\varepsilon(h_1)}(t)$ and $x_{\varepsilon(h_2)}(t)$ be two ε–optimal trajectories as stated in Theorem 5.2 or Proposition 6.2 that correspond to $H(t,x,y,h_1)$ and $H(t,x,y,h_2)$ and which satisfy (1.6) with $\varepsilon(h) = \varepsilon$. By (5.2) and (6.8)

$$g(x_{\varepsilon(h_1)}(1)) \le \inf_{K(x_0)} g(x(1)) + \frac{1}{2}M\varepsilon,$$

$$g(x_{\varepsilon(h_2)}(1)) \leq \inf_{K(x_0)} g(x(1)) + \frac{1}{2}M\varepsilon,$$

for some $M > 0$. These imply (7.1).

The next problem is a convergence of these approximations to a solution of (1.1)–(1.3).

THEOREM 7.1 *Assume g to be weakly lower semicontinuous in X and in (ii), Section 1, $x' \to x$ weakly or $X = R^n$. Let $\{x_{\varepsilon(h_i)}\}_{i=1}^{\infty}$ be a sequence of ε-optimal trajectories corresponding to the approximation $\{H(t,x,y,h_i)\}_{i=1}^{\infty}$ with $\varepsilon(h_i) \to 0$, $h_i \to 0$ as $i \to \infty$ satisfying (1.6). Then there exists a subsequence of $\{x_{\varepsilon(h_i)}\}_{i=1}^{\infty}$ which we denote again by $\{x_{\varepsilon(h_i)}\}_{i=1}^{\infty}$ converging weakly, in the space of absolutely continuous function $A^2(X)$ with $\dot{x} \in L^2(0,T;X)$, to a solution of (1.1)–(1.3).*

PROOF. Since all $x_{\varepsilon(h_i)}(t)$ satisfy (1.2) and we assumed basic hypotheses therefore $\{\dot{x}_{\varepsilon(h_i)}(\cdot)\}_{i=1}^{\infty}$ is bounded in $L^{\infty}(0,T;X)$ and in $L^2(0,T;X)$. Hence $\{x_{\varepsilon(h_i)}(\cdot)\}_{i=1}^{\infty}$ is bounded in $A^2(X)$ and there exists a subsequence of it weakly convergent in $A^2(X)$ to an $\bar{x} \in A^2(X)$. By (5.2) or (6.8)

$$g(x_{\varepsilon(h_i)}(1)) \leq \inf_{K(x_0)} g(x(1)) + M\varepsilon(h_i), \qquad i = 1, 2 \ldots,$$

for some $M > 0$ independent from h_i. By the assumption on g or X

$$g(\bar{x}(1)) \leq \inf_{K(x_0)} g(x(1)).$$

From the basic assumption we infer that $\bar{x}(t)$ satisfies (1.2) and (1.3), i.e. $\bar{x}(t)$ is a solution of (1.1)–(1.3).

References

CASTAING C. AND VALADIER M. (1977) Convex Analysis and Measurable Multifunctions, Lecture Notes in Math., 580, Springer.

CLARKE F.H. (1983) Optimization and Nonsmooth Analysis, Wiley.

CLARKE F.H. AND VINTER R.B. (1983) Local optimality conditions and Lipschitzian solutions to the Hamilton–Jacobi equation, *SIAM J. Control and Optimization*, **21**, 856–870.

EKELAND I. (1974) On the variational principle, *Journal of Mathematical Analysis and Applications* **47**, 324–353.

FLEMING W.H. (1969) The Cauchy problem for a nonlinear first–order partial diferential equation, *J. Differential Equation* **5**, 515–530.

FLEMING W.H. AND RISHEL R.W. (1975) Deterministic and stochastic optimal control, Springer, N.Y.

NOWAKOWSKI A. (1988) Field theories in the modern calculus of variations, *Trans. Amer. Math. Soc.* **309**, 725–752.

Young L.C. (1969) Lectures on the Calculus of Variations and Optimal Control Theory, Saunders, Philadelphia.