Proc. Indian Acad. Sci. (Math. Sci.), Vol. 100, No. 1, April 1990, pp. 65–77.

On infinite-dimensional control systems with state and control constraints

NIKOLAOS S PAPAGEORGIOU
1015 Department of Mathematics, University of California, Davis, California 95616, USA

MS received 12 October 1988; revised 15 December 1989

Abstract. In this paper we examine infinite-dimensional control systems governed by semilinear evolution equations and having both state and control constraints. We introduce the relaxed system and show that the original trajectories are dense in an appropriate function space in the relaxed ones. We also determine the dependence of the solution set on the initial conditions. Then using those results we establish necessary and sufficient conditions for optimality for some optimization problems. Finally we prove some controllability results.

Keywords. Relaxation; relaxed system; lower semicontinuous and upper semicontinuous multifunction; continuous selections; convex subdifferential; controllability.

1. Introduction

In this paper we consider control systems governed by semilinear evolution equations, with both state and control constraints. The study of infinite-dimensional control systems was initiated by the works of Egorov [8] and Lions [13] and since then it has attracted the interest of many mathematicians and has become one of the most important and productive areas of applied mathematics. The majority of the works treats systems with no state constraints and with time independent control constraints. Recently Barbu [2] and Mackenroth [15], studied convex optimal control problems where a linear equation and a state constraint are given.

In this work, we study systems governed by a semilinear evolution equation and having general state and control constraints of the form $x(t)\in K(t)$ and $u(t)\in U(t)$, where $K(\cdot)$, $U(\cdot)$ are the time varying sets (they can be functional constraints or inclusion constraints or combinations thereof). We define the relaxed (convexified) system, and we address the "relaxation" problem, namely we examine what is the relation between the sets of trajectories of the original and relaxed systems. Then we examine the dependence of the set of solutions of the original system on the initial condition. Combining these results, we are then able to obtain some necessary and sufficient conditions for the solutions of some nonlinear, convex optimization problems and determine some continuity properties of the corresponding value (Bellman) function. Finally we derive some necessary and sufficient conditions for controllability.

2. Preliminaries

Let X be a separable Banach space. Throughout this paper we will be using the following notations:

$$P_{f(c)}(X) = \{A \subseteq X\text{: nonempty, closed, (convex)}\}$$

and

$$P_{(w)k(c)} = \{A \subseteq X\text{: nonempty, } (w-) \text{ compact, (convex)}\}.$$

If (Ω, Σ) is a measurable space and $F\colon \Omega \to 2^X \setminus \{\emptyset\}$, we say that $F(\cdot)$ is graph measurable, if $\operatorname{Gr} F = \{(\omega, x) \in \Omega \times X\colon x \in F(\omega)\} \in \Sigma \times B(X)$, $B(X)$ being the Borel σ-field of X.

If Y, Z are Hausdorff topological spaces, a multifunction $G\colon Y \to 2^Z \setminus \{\emptyset\}$ is said to be upper semicontinuous (u.s.c.) (resp. lower semicontinuous (l.s.c.)) if for every $C \subseteq Z$ closed $G^-(C) = \{y \in Y\colon G(y) \cap C \neq \emptyset\}$ (resp. $G^+(C) = \{y \in Y\colon G(y) \subseteq C\}$) is closed in Y. Under additional hypotheses on Y, Z we can have other equivalent definitions of upper and lower semicontinuity. For details we refer to Delahaye–Denel [4]. When a multifunction is both u.s.c. and l.s.c., then it is said to be continuous.

On $P_f(X)$, we can define a generalized metric, known as the "Hausdorff metric" by setting:

$$h(A, B) = \max\{\sup(d(a, B), a \in A), \sup(d(b, A), b \in B)\}.$$

Here $A, B \in P_f(X)$ and $d(a, B) = \inf(\|a - b\|\colon b \in B)$, the distance of a from B (similarly for $d(b, A)$). It is well known that $(P_f(X), h)$ is a complete metric space. A multifunction $F\colon X \to P_f(X)$ is said to be Hausdorff continuous (h-continuous), if it is continuous from X into the metric space $(P_f(X), h)$. In general, continuity and h-continuity are not comparable notions. However they coincide if the multifunction is compact valued (i.e. has values in $P_k(X)$). This is a consequence of the fact that on $P_k(X)$ the Vietoris and Hausdorff topologies coincide (see Klein–Thompson [10], p. 41).

Now let $f\colon X \to \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ be a proper (i.e. not identically $+\infty$), convex function. Let $x \in \operatorname{dom} f = \{z \in X\colon f(z) < +\infty\}$. The subdifferential of $f(\cdot)$ at x is defined to be the set

$$\partial f(x) = \{x^* \in X^*\colon (x^*, x' - x) \leqslant f(x') - f(x), x' \in X\}$$

(see Laurent [12]).

Finally if $A \subseteq X$, then the support function of A is the function $\sigma(\cdot, A)\colon X^* \to \bar{\mathbb{R}}$ defined by

$$\sigma(x^*, A) = \sup\{(x^*, x)\colon x \in A\},$$

where $(\cdot, \cdot)$ denotes the duality brackets for the pair (X, X^*).

3. Relaxation results

The mathematical setting is the following. Let $T = [0, b]$ and H a separable Hilbert space. Also X is a dense subspace of H carrying the structure of a separable, reflexive Banach space and which embeds continuously in H. Identifying H with its dual (pivot space), we have $X \hookrightarrow H \hookrightarrow X^*$ with all embeddings continuous and dense. We will also assume that they are compact. To have an example in mind consider $X = H_0^1(0, 1)$, $H = L^2(0, 1)$ and $X^* = H^{-1}(0, 1)$. Such a triple of spaces (X, H, X^*) is often called a "Gelfand triple" or "spaces in normal position". By $\|\cdot\|$ (resp. $|\cdot|$, $\|\cdot\|_*$), we will denote the norm of X(resp. of X, X^*). Also by $(\cdot,\cdot)$ we will denote the inner product in H and by $\langle\cdot,\cdot\rangle$ the duality brackets for the pair (X, X^*). The two are compatible in the sense that if $x \in X \subseteq H$ and $h \in H \subseteq X^*$, then we have $(x, h) = \langle x, h\rangle$. By H_w we will denote the space H endowed with the weak topology. Also let Z be a separable Banach space, modelling the control space. The distributed parameter control system under consideration is the following:

$$\left.\begin{cases} \dot{x}(t) + A(t)x(t) = f(t, x(t), u(t)) \text{ a.e. on } T \\ x(0) = x_0,\ x(t) \in K(t),\ t \in T \\ u(t) \in U(t) \text{ a.e., } u(\cdot) \text{ is measurable} \end{cases}\right\} (\ast).$$

To this system we associate a larger system, with convexified dynamics, known as the "relaxed system".

$$\left.\begin{cases} \dot{x}(t) + A(t)x(t) = \int_Z f(t, x(t), z)\lambda(t)(dz) \\ x(0) = x_0, x(t) \in K(t) t \in T \\ \lambda(t) \in \Sigma(t) \text{ a.e., } \lambda(\cdot) \text{ is measurable} \end{cases}\right\} (\ast)_r.$$

Here $\Sigma(t) = \{\lambda \in M_+^1(Z): \lambda(U(t)) = 1\}$ (by $M_+^1(Z)$ we denote the space of probability measures on Z). So $\lambda(\cdot)$ is a transition probability. Note that in the dynamical equation governing $(\ast)_r$, the control function $\lambda(\cdot)$ enters linearly.

We will make the following hypotheses concerning the data of our problems.

<u>$H(A)$</u>: $A: T \times X \to X^*$ is an operator *s.t.*

(1) $A(t)(\cdot)$ is linear, continuous,
(2) $\|A(t')x - A(t)x\|_* \leqslant k|t' - t|\,\|x\|, k > 0$ for all $t', t \in T, x \in X$,
(3) $\langle A(t)x, x\rangle \geqslant c\|x\|^2$, $c > 0$ (i.e. $A(t)(\cdot)$ is strongly monotone).

<u>$H(f)$</u>: $f: T \times H \times Z \to H$ is a function *s.t.*

(1) $t \to f(t, x, z)$ is measurable,
(2) $x \to -f(t, x, z)$ is continuous, monotone,
(3) $(x, z) \to f(t, x, z)$ is continuous from $H \times Z$ into H_w,
(4) $|f(t, x, z)| \leqslant a(t) + b(t)(|x| + \|z\|)$ a.e., with $a(\cdot), b(\cdot) \in L_+^2$.

Remark. The dissipativity hypothesis $H(f)$ (2) on $f(t, \cdot, z)$ can be replaced by

(2)' $|f(t, x', z) - f(t, x, z)| \leqslant k(t)|x' - x|$ a.e., with $k(\cdot) \in L_+^1$.

<u>$H(U)$</u>: $U: T \to P_{fc}(Z)$ is a graph measurable multifunction s.t. $U(t) \subseteq W$ a.e., where $W \in P_{wkc}(Z)$.

Because of hypothesis $H(A)$, from proposition 5.5.1 of Tanabe [22], we know that $A(\cdot)(\cdot)$ generates a strongly continuous evolution operator $S: \Delta = \{(t, s) \in T \times T: 0 \leqslant s \leqslant t \leqslant b\} \to \mathscr{L}(H)$, in terms of which a solution of the Cauchy problem describing the dynamics of (*), is given by the following variation of constants formula:

$$x(t) = S(t,0)x_0 + \int_0^t S(t,s)f(s,x(s),u(s))\,ds, \quad t\in T.$$

It is well known (see for example Barbu [3], p. 167 or Lions [14]), that in this case $x(\cdot)\in W(T) = \{x\in L^2(X): \dot{x}\in L^2(X^*)\} \subseteq C(T,H)$.

We will make the following hypothesis on $S(t,s)$: $\underline{H_c}$: $S(t,s)$ is compact for $t > s$, $(t,s)\in\Delta$.

To have a typical example in mind consider the Laplacian.

Let $Q(x_0)$ be the set of trajectories of (*) and $Q_r(x_0)$ the set of trajectories of $(*)_r$. Note that given any initial control function $u(\cdot)$, we can view it as a relaxed control by considering $\delta_{u(t)}(\cdot)$, the Dirac transition probability. So $Q(x_0)\subseteq Q_r(x_0)$. In [20], with no state constraints present (i.e. $K(t) = H$ for all $t\in T$), we proved that under hypotheses $H(A)$, $H(f)$, $H(U)$, H_c. we have $\overline{Q(x_0)} = Q_r(x_0)$, the closure taken in $C(T,H)$. We would like to have the same result, in the presence of the state constraints $K(\cdot)$. For this we will need the following hypothesis on $K(\cdot)$.

$\underline{H(K)}$: $K: T\to P_{fc}(H)$ is l.s.c. and $\operatorname{int} K(t) \neq \varnothing$ for all $t\in T$.

A well known selection theorem of Michael [16], tells us that under the above hypothesis, there exists $k: T\to H$ continuous *s.t.* $k(t)\in K(t)$ for all $t\in T$ (i.e. $K(\cdot)$ admits a continuous selector). Denote by $CS(K)$ the set of continuous selectors of $K(\cdot)$.

To prove our density (relaxation) result, we will need the following three auxiliary results, which are actually of independent interest.

Lemma α. If hypothesis $H(K)$ holds and $x(\cdot)\in CS(\operatorname{int} K)$ then $t\to p(t) = d(x(t), bdK(t))$ is l.s.c. where $bdK(t)$ denotes the boundary of $K(t)$.

Remark. Because of Michael's theorem [16] (theorem 3.2″), we know that $CS(\operatorname{int} K) \neq \varnothing$. Also note that even though $K(\cdot)$ is lower semicontinuous, $bd\,K(\cdot)$ need not be such. For example let $K(t) = [0,1]$ if $t\neq 0$ and $K(0) = [0,1/2]$. Clearly $K(\cdot)$ is *l.s.c.* but $bd\,K(\cdot)$ is neither *l.s.c.* nor *u.s.c.*

Proof. We need to show that for all $\lambda > 0$, the level set

$$L(\lambda) = \{t\in T: p(t)\leqslant \lambda\}$$

is closed in T. So let $t_n\to t$, $t_n\in L(\lambda)$, $n\geqslant 1$. Let $\varepsilon > 0$, $\varepsilon < \lambda$ be given. Due to the lower semicontinuity of $K(\cdot)$ (hypothesis $H(K)$), we can find $n_0\geqslant 1$ s,t, for $n\geqslant n_0$ we have:

$$K(t)\subseteq K(t_n) + \mathring{B}_\varepsilon,$$

where $\mathring{B}_\varepsilon$ is the open ε-ball in H. Also there exists $n_1\geqslant 1$ *s.t.* for $n\geqslant n_1$ we have $|x(t_n) - x(t)| < \varepsilon$. Then for $z\in bdK(t)$ and for $n\geqslant n_2 = \max(n_0, n_1)$, we have: $p(t)\leqslant |x(t) - z| \leqslant |x(t) - x(t_n)| + |x(t_n) - z|$. Therefore $p(t) < \varepsilon + d(x(t_n), bdK(t))$. Since $x(t)\in K(t)\subseteq K(t_n) + \varepsilon\mathring{B}_1$ for $n\geqslant n_2$, we get:

$$d(x(t_n), bdK(t)) \leqslant d(x(t_n), bd(K(t_n) + \mathring{B}_\varepsilon))).$$

From this inequality we see that if there exists $\{t_{n_k}\}_{k\geq 1} \subseteq \{t_n\}_{n\geq 1}$ *s.t.*

$$d(x(t_{n_k}), bd(K(t_{n_k}) + \mathring{B}_\varepsilon)) \leqslant \lambda,$$

then in the limit as $k \to \infty$, we will get

$$p(t) < \varepsilon + \lambda.$$

Let $\varepsilon \downarrow 0$. We have $p(t) \leqslant \lambda \Rightarrow t \in L(\lambda)$. So assume that for all but finitely many $n \geqslant n_2$ we have:

$$d(x(t_n), bd(K(t_n) + \varepsilon \mathring{B}_1)) = \hat{\lambda}_n > \lambda.$$

This inequality means that $x(t_n) + \mathring{B}_{\hat{\lambda}_n - \varepsilon} + \mathring{B}_\varepsilon \subseteq K(t_n) + \mathring{B}_\varepsilon$ for $n \geqslant n_2$. So for $n \geqslant n_2$ we have

$$x(t_n) + \mathring{B}_{\hat{\lambda}_n - \varepsilon} \subseteq K(t_n) \Rightarrow d(x(t_n), bdK(t_n)) \geqslant \hat{\lambda}_n - \varepsilon$$
$$\hat{\lambda}_n \leqslant \varepsilon + d(x(t_n), bdK(t_n)) = \varepsilon + p(t_n) \Rightarrow p(t) < 2\varepsilon + p(t_n) < 2\varepsilon + \lambda.$$

Let $\varepsilon \downarrow 0$. We get $p(t) \leqslant \lambda \Rightarrow t \in L(\lambda)$ is closed $\Rightarrow p(\cdot)$ is l.s.c.. Q.E.D.

Remark. It is clear from the above proof that H can be any Banach space (note that since T is a compact metric space we do not need separability on H; see Michael [16], theorem 3.2″).

Lemma β. If hypothesis H(K) holds, then $CS(\text{int}\,K) = \text{int}\,CS(K) \neq \varnothing$.

Proof. It is clear from the definitions that $\text{int}\,CS(K) \subseteq CS(\text{int}\,K)$. We will now show that the opposite inclusion also holds. Let $x(\cdot) \in CS(\text{int}\,K)$ and set $p(t) = d(x(t), bdK(t))$. From lemma α, we know that $p(\cdot)$ is *l.s.c.*. Since T is compact, there exists $t \in T$ *s.t.*

$$\inf_{t\in T} p(t) = p(\hat{t}) = d(x(\hat{t}), bdK(\hat{t})) > 0$$

since $x(\hat{t}) \in \text{int}\,K(\hat{t})$. Let $\delta < p(\hat{t})$. Then clearly $x(\cdot) + \mathring{B}_\delta \subseteq CS(K) \Rightarrow x(\cdot) \in \text{int}\,CS(K) \Rightarrow CS(\text{int}\,K) = \text{int}\,CS(K)$ and the nonemptiness follows from proposition 2.3 and theorem 3.1‴ (c) of Michael [16]. Q.E.D.

Remark. Again H can be any Banach space.

For the next lemma assume that Y is any Banach space.

Lemma γ. If $A, B \subseteq Y$ *are such that* $\bar{A}$ *is convex,* B *is convex with int* $B \neq \varnothing$,

and $\bar{A} \cap \text{int}\,B \neq \varnothing$,
then $\overline{A \cap B} = \bar{A} \cap \bar{B}$.

Proof. From lemma 5 of Moreau [17], we know that $\sigma_{\bar{A}\cap\text{int}\,B}(\cdot) = \sigma_{\bar{A}\cap\bar{B}}(\cdot) = (\sigma_A \Box \sigma_B)(\cdot)$ (the infimal convolution in the sense of convex analysis of $\sigma_A(\cdot)$ and $\sigma_B(\cdot)$, see Laurent [12]). Since $\bar{A} \cap \text{int}\,B$ and $\bar{A} \cap \bar{B}$ are both convex, from the equality of their support functions, we get that $\overline{\bar{A}\cap \operatorname{int} B}=\bar{A}\cap\bar{B}$. Let $x\in\bar{A}\cap\operatorname{int} B$. Then we can find $x_n\in A$ s.t. $x_n\to x$ and $\delta>0$ s.t. $x+\mathring{B}_\delta\subseteq\operatorname{int} B$. So for $n\geqslant 1$ large enough, we will have that $x_n\in x+\mathring{B}_\delta\Rightarrow x_n\in A\cap\operatorname{int} B\Rightarrow x\in\overline{A\cap\operatorname{int} B}\Rightarrow\overline{\bar{A}\cap\operatorname{int} B}=\overline{A\cap\operatorname{int} B}=\bar{A}\cap\bar{B}\Rightarrow\overline{A\cap B}=\bar{A}\cap\bar{B}$ as claimed.

Q.E.D.

Now we are ready for the theorem comparing the sets $Q(x_0)$ and $Q_r(x_0)$.

Theorem 3.1. *If hypotheses $H(A)$, $H(f)$, $H(U)$, H_c and $H(K)$ hold and there exists $x(\cdot)\in Q(x_0)$ s.t. $x(t)\in\operatorname{int} K(t)$ for all $t\in T$, then $Q_r(x_0)=\overline{Q(x_0)}$ the closure taken in $C(T,H)$.*

Proof. Let $P(x_0)$ and $P_r(x_0)$ be the sets of trajectories of (*) and $(*)_r$ respectively, when no state constraints are present (i.e. $K(t)=H$ for all $t\in T$). From theorem 3.2 of [20], we know that $P_r(x_0)=\overline{P(x_0)}$, the closure in $C(T,H)$. Now observe that:

$$Q(x_0)=P(x_0)\cap CS(K)$$

and

$$Q_r(x_0)=P_r(x_0)\cap CS(K).$$

So $Q_r(x_0)$ is convex, closed in $C(T,H)$. From our interiority hypothesis and lemma γ, we have that $\operatorname{int} CS(K)\neq\varnothing$. Since $\overline{P(x_0)}=P_r(x_0)$ is convex, $CS(K)$ is convex with nonempty interior and by hypothesis $P(x_0)\cap\operatorname{int} CS(K)\neq\varnothing$, we can apply lemma β and get that $\overline{Q(x_0)}=\overline{P(x_0)\cap CS(K)}=\overline{P(x_0)}\cap\overline{CS(K)}=P_r(x_0)\cap CS(K)=Q_r(x_0)$, the closure in $C(T,H)$.

Q.E.D.

An immediate, useful consequence of the above density result is the following theorem concerning the reachable sets $R(t)=\{x(t): x(\cdot)\in Q(x_0)\}$ and $R_r(t)=\{x(t): x(\cdot)\in Q_r(x_0)\}$, $t\in T$.

Theorem 3.2. *If the hypotheses of theorem 3.1 hold, then for all $t\in T$, $\overline{R(t)}=R_r(t)$, the closure taken in the strong topology of H.*

Proof. Note that for every $t\in T$, $R(t)=e_t(Q(x_0))$ and $R_r(t)=e_t(Q_r(x_0))$, where $e_t(\cdot)$ is the evaluation of $t\in T$ map. Recalling that $e_t(\cdot)$ is continuous on $C(T,H)$ (see for example Kuratowski [11]) and using theorem 3.1, we get:

$$e_t(Q_r(x_0))=e_t(\overline{Q(x_0)})\subseteq\overline{e_t(Q(x_0))}.$$

But from [20] we know that $Q_r(x_0)$ is compact in $C(T,H)$. So we have:

$$e_t(Q_r(x_0))=\overline{e_t(Q(x_0))}$$

$$\Rightarrow R_r(t)=\overline{R(t)},\text{ the closure taken in the strong topology on } H.$$

Q.E.D.

Next we will examine the dependence of the set $Q_r(x_0)$ on the initial condition x_0. Such a study, will lead us to some interesting observations concerning the value function of certain optimization problems that we will consider in the next section.

We will need the following lemma.

Lemma δ. If V, Y are metric spaces and $G: V \to P_k(Y)$ is a multifunction s.t. for all $K \subseteq V$ compact, the restriction of $G(\cdot)$ on K is u.s.c., then $G(\cdot)$ is u.s.c..

Proof. We need to show that for all $C \subseteq Y$ nonempty, closed, we have that $G^-(C) = \{x \in V: G(x) \cap C \neq \varnothing\}$ is closed too. So let $\{x_n\}_{n \geq 1} \subseteq G^-(C)$, $x_n \to x$ in V. Then by hypothesis $G(\cdot)$ restricted on $K = \{x_n, x\}_{n \geq 1}$ is u.s.c.. Thus $G(K)$ is compact (see Klein–Thompson [10], theorem 7.4.2, p. 90). Let $y_n \in G(x_n) \cap C$. Since $\{y_n\}_{n \geq 1} \subseteq G(K)$, by passing to a subsequence if necessary, we may assume that $y_n \to y$ in Y. So $y \in \overline{\lim}\, G(x_n)$ and because $G|_K$ is u.s.c., we have $\overline{\lim}\, G(x_n) \subseteq G(x) \Rightarrow y \in G(x)$ and $y \in C \Rightarrow G(x) \cap C \neq \varnothing \Rightarrow x \in G^-(C) \Rightarrow G(\cdot)$ is indeed u.s.c.

Q.E.D.

Having this lemma, we can prove our first theorem concerning the dependence of $Q_r(x_0)$ on the initial condition x_0.

In what follows we set $F(t, x) = \cup \{f(t, x, u): u \in U(t)\}$ and $F_r(t, x) = \{\int_W f(t, x, z)\, \lambda(\mathrm{d}z): \lambda \in \Sigma(t)\}$. From the proof of theorem 3.2 in [20] we know that $\overline{\text{conv}}\, F(t, x) = F_r(t, x)$. So $F_r(t, x) \in P_{wkc}(H)$ for all $(t, x) \in T \times H$.

Theorem 3.3. *If the hypotheses of theorem 3.1 hold for all $x_0 \in V \subseteq H$, then Q_r: $V \to P_k(C(T, H))$ is u.s.c.*

Proof. In view of lemma δ and the equivalences in Delahaye–Denel [4], we only need to show that if $x_0^n \xrightarrow{s} x_0$ in V, then $\overline{\lim}\, Q_r(x_0^n) \subseteq Q_r(x_0)$.

So let $y \in \overline{\lim}\, Q_r(x_0^n)$. Then by definition (see Kuratowski [11], p. 337), we can find $y_{n_k} \in Q_r(x_0^{n_k})$ s.t. $y_{n_k} \xrightarrow{s} y$ in $C(T, H)$. Then we have:

$$y_{n_k}(t) = S(t, 0)x_0^{n_k} + \int_0^t S(t, s) g_{n_k}(s)\, \mathrm{d}s, \quad t \in T$$

where

$$g_{n_k}(\cdot) \in S^1_{F_r}(\cdot, y_{n_k}(\cdot)) = \{h \in L^1(H): h(t) \in F_r(t, y_{n_k}(t)) \text{ a.e.}\}.$$

From proposition 1, p. 47, of Aubin–Cellina [1], we know that $F(t, \cdot)$ is u.s.c. from H into H_w. So $F_r(t, \cdot)$ is u.s.c. from H into H_w. Therefore we have that $\overline{\text{conv}}\, \bigcup_{k \geq 1} F(t, y_{n_k}(t)) = L(t) \in P_{wkc}(H)$. Furthermore for every $k \geq 1$ we have:

$$\begin{aligned} \operatorname{Gr} F(\cdot, y_{n_k}(\cdot)) &= \{(t, x) \in T \times H: x \in F(t, y_{n_k}(t))\} \\ &= \{(t, x) \in T \times H: x = f(t, y_{n_k}(t), u), u \in U(t)\} \\ &= \{(t, x) \in T \times H: x - f(t, y_{n_k}(t), u) = 0,\ d(u, U(t)) = 0\}. \end{aligned}$$

Let $\theta(t, x, u) = x - f(t, y_{n_k}(t), u)$ and $\eta(t, u) = d(u, U(t))$. From $H(f)$ we have that $\theta(\cdot, \cdot, \cdot)$ is measurable on $T \times H \times W_w$, where W_w is the set W with the relative weak topology. Also from $H(U)$ we have that $\eta(\cdot, \cdot)$ is measurable on $T \times W$. Furthermore from corollary 2.4 of Edgar [7], we have that $B(Y_w) = B(Y) \Rightarrow B(Y_w) \cap W = B(Y) \cap W \Rightarrow B(W_w) = B(W)$. Hence

$$\{(t, x, u) \in T \times H \times W: \theta(t, x, u) = 0, \quad \eta(t, u) = 0\} \in B(T) \times B(X) \times B(W_w).$$

But recall that W_w is a compact Polish space (see Dunford–Schwartz [6], theorem

3, p. 434). So we can apply the Arsenin–Novikov theorem (see Dellacherie [5]) and get that:

$$\mathrm{proj}_{T\times X}\{(t,x,u)\in T\times H\times W:\theta(t,x,u)=0,\ \eta(t,u)=0\}\in B(T)\times B(H)$$
$$\Rightarrow \mathrm{Gr}\,F(\cdot,y_{n_k}(\cdot))\in B(T)\times B(H)\Rightarrow F_r(\cdot,y_{n_k}(\cdot))$$

is graph measurable $\Rightarrow L(\cdot)$ is graph measurable. Also from $H(f)$, it is easy to see that $t\to|L(t)|=\sup\{|y|:y\in L(t)\}\in L^1_+$. So we can apply proposition 3.1 of [18] and deduce that by passing to a subsequence if necessary, we may assume that

$g_{n_k}\xrightarrow{w} g$ in $L^1(H)$. Then from theorem 3.1 of [19], we have:

$$g(t)\in\overline{\mathrm{conv}}\ \text{w-}\overline{\lim}\,F_r(t,y_{n_k}(t))\subseteq F_r(t,y(t))\text{ a.e.}$$
$$\Rightarrow g\in S^1_{F_r}(\cdot,y(\cdot))$$
$$\Rightarrow y(t)=S(t,0)x_0+\int_0^t S(t,s)g(s)\,\mathrm{d}s,\quad t\in T.$$

Let $L(t)=\{\lambda\in\Sigma(t): g(t)=\int_W f(t,y(t),z)\lambda(\mathrm{d}z)\}$. From the definition of $F_r(\cdot,\cdot)$ it is clear that $L(t)\neq\varnothing$ for all $t\in T$. Let $p\colon T\times M^1_+(W_w)\to H$ be defined by $p(t,\lambda)=g(t)-\int_W f(t,y(t),z)\lambda(\mathrm{d}z)$. Clearly $p(t,\lambda)$ is measurable in t, continuous in λ. Hence since $M^1_+(W_w)$ with the usual narrow (weak) topology, is compact metrizable, we conclude that $p(t,\lambda)$ is jointly measurable. Therefore since $\mathrm{Gr}\,\Sigma\in B(T)\times B(M^1_+(W_w))$ (see [20]), we conclude that:

$$\mathrm{Gr}\,L=\{(t,\lambda)\in T\times M^1_+(W_w): p(t,\lambda)=0\}\cap\mathrm{Gr}\,\Sigma\in B(T)\times B(M^1_+(W_w)).$$

Apply Aumann's selection theorem (see Wagner [23]) to get $\lambda\colon T\to M^1_+(W_w)$ measurable s.t. $\lambda(t)\in L(t)$ for all $t\in T$. Then $g(t)=\int_W f(t,y(t),z)\lambda(t)(\mathrm{d}z)\Rightarrow y\in Q_r(x_0)\Rightarrow Q_r(\cdot)$ is indeed u.s.c. from H into $P_k(C(T,H))$.

Q.E.D.

For a simpler form of the system, in which state and control functions are separated, we can say more.

So consider the following "original" and "relaxed" systems:

$$\left.\begin{cases}\dot x(t)+A(t)x(t)=f(t,u(t))\text{ a.e. on }T\\ x(0)=x_0,\ x(t)\in K(t)t\in T,\\ u(t)\in U(t)\text{ a.e., }u(\cdot)\text{ is measurable}\end{cases}\right\}(\underline{**})$$

and

$$\left.\begin{cases}\dot x(t)+A(t)x(t)=\int_W f(t,z)\lambda(t)(\mathrm{d}z)\text{ a.e.}\\ x(0)=x_0,\ x(t)\in K(t)t\in T,\\ \lambda(t)\in\Sigma(t)\text{ a.e., }\lambda(\cdot)\text{ is measurable}\end{cases}\right\}(\underline{**})_r.$$

As before $\Sigma(t)=\{\lambda\in M^1_+(W_w):\lambda(U(t))=1\}$ and $Q(x_0)$ (resp. $Q_r(x_0)$) is the solution set of (**) (resp. $(**)_r$). Hypothesis $H(f)$ now takes the following simpler form: $\underline{H(f)_1}$: $f\colon T\times Z\to H_w$ is a map *s.t.*

(1) $t\to f(t,z)$ is continuous,

(2) $z\to f(t,z)$ is continuous from Z into H_w,

(3) $|f(t,z)|\leqslant a(t)$ a.e. with $a(\cdot)\in L^2_+$.

Theorem 3.4. *If hypotheses $H(A)$, $H(f)_1$, $H(U)$, H_c, $H(K)$ hold and there exists*

$$\hat{x}(\cdot)\in Q(x_0) \text{ s.t. } x(t)\in \operatorname{int} K(t) \text{ for all } t\in T,$$

then there exists $\delta > 0$ s.t. for all $x_0'\in(x_0 + \mathring{B}_\delta) = V$, $Q(x_0') \neq \emptyset$ and $Q(\cdot)$ is l.s.c. from V into $2^{C(T,H)}\setminus\{\emptyset\}$.

Proof. Let $\hat{x}(\cdot)$ be the interior trajectory emanating from x_0, postulated by our hypothesis. Let $\hat{u}(\cdot)\in S_U^1$ be the admissible control generating $\hat{x}(\cdot)$. From lemma β we know that there exists $\delta_1 > 0$ *s.t.* $\hat{x} + \mathring{B}_{\delta_1} \subseteq CS(K)$. Let $\delta = (\delta_1/M)$ (where $\|S(t,s)\| \leqslant M$ for all $(t,s)\in\Delta$) and for $x_0'\in x_0 + \mathring{B}_{\delta_1}$ define:

$$x(t) = S(t,0)x_0' + \int_0^t S(t,s)f(s,\hat{u}(s))\,ds$$

$$\Rightarrow \|x(t) - \hat{x}(t)\| \leqslant \|S(t,0)\|\cdot\|x_0' - x_0\| < \delta$$

$$\Rightarrow x(\cdot)\in P(x_0')\cap CS(K) = Q(x_0')$$

$$\Rightarrow Q(x_0') \neq \emptyset \text{ for all } x_0'\in x_0 + \mathring{B}_\delta.$$

Next let $x(\cdot)\in Q(x_0)$. From lemma γ we know that we can find $z_\varepsilon(\cdot)\in P(x_0)\cap\operatorname{int} CS(K)$ (i.e. an interior trajectory) s.t. $\|x - z_\varepsilon\|_\infty < \varepsilon/2$ in $C(T,H)$. Also from above we know that we can find $\delta > 0$ *s.t.* for all $x_0'\in x_0 + \mathring{B}_\delta$ there exists $y(\cdot)\in Q(x_0')$ *s.t.* $\|z_\varepsilon - y\|_\infty < \varepsilon/2$ in $C(T,H)$. Therefore $\|x - y\|_\infty < \varepsilon/2 + \varepsilon/2 = \varepsilon \Rightarrow x(\cdot)\in Q(x_0') + \mathring{B}_\varepsilon$. Since $x(\cdot)\in Q(x_0)$ was arbitrary, we get $Q(x_0) \subseteq Q(x_0') + \mathring{B}_\varepsilon$ for all $x_0'\in x_0 + \mathring{B}_\delta \Rightarrow Q(\cdot)$ is l.s.c. on V.

Q.E.D.

Combining theorems 3.3 and 3.4 we get:

Theorem 3.5. *If the hypotheses of theorem 3.4 hold for all $x_0\in V\subseteq H$, then $Q_r\colon V\to P_k(C(T,H))$ is h-continuous.*

Proof. From theorems 3.3 and 3.4 we get that $Q_r(\cdot)$ is continuous on V and because it is $P_k(C(T,H))$-valued, it is h-continuous, as claimed.

Q.E.D.

4. Optimization problems

In this section we solve some optimization problems involving the systems considered in §3.

So let $\phi\colon H\times C(T,H)\to\mathbb{R}$ be a continuous function *s.t.* for every $x_0\in H$, $\phi(x_0,\cdot)$ is convex. We consider the following optimization problem:

$$(\underline{O_1})\colon v(x_0) = \inf\{\phi(x_0,x)\colon x(\cdot)\in Q(x_0)\}.$$

Theorem 4.1. *If the hypotheses of theorem 3.1 hold, then $\hat{x}(\cdot)$ is a solution of $(\underline{O_1})$ if and only if there exists $x^*\in\partial_2\phi(x_0,x)\subseteq M(T,H)$ s.t. $((x^*,x)) = \inf\{((x^*,x))\colon x\in Q(x_0)\}$ with $((\cdot,\cdot))$ denoting the duality brackets for the pair $(C(T,H)$, $M(T,H) = (C(T,H))^*)$. Furthermore if the hypotheses of theorem 3.3 hold. then $v\colon V\to\mathbb{R}$ is u.s.c..*

Proof. Let $\delta_{Q(x_0)} : C(T,H) \to \bar{\mathbb{R}} = \mathbb{R}U\{+\infty\}$ be defined by $\delta_{Q(x_0)}(x) = 0$ if $x \in Q(x_0)$, and $+\infty$ otherwise. Exploiting the continuity of $\phi(\cdot,\cdot)$ we have:

$$\inf_{x \in Q(x_0)} \phi(x_0, x) = \inf_{x \in \overline{Q(x_0)}} \phi(x_0, x) = \inf_{x \in C(T,H)} [\phi(x_0, x) + \delta_{\overline{Q(x_0)}}(x)].$$

From theorem 3.1 we know that $\overline{Q(x_0)} = Q_r(x_0)$ and the latter is convex. So $\delta_{\overline{Q(x_0)}}(\cdot)$ is a closed convex function. Hence our problem is a convex optimization problem. So the optimality condition from convex analysis (see Laurent [12], theorem 6.4.5, p. 352), tells us that $\hat{x} \in Q(x_0)$ is a solution if and only if $0 \in \partial[\phi(x_0, \hat{x}) + \delta_{\overline{Q(x_0)}}(\hat{x})]$. Since $\phi(x_0, \cdot)$ is continuous we can apply the Moreau–Rockafellar theorem (see Laurent [12], theorem 6.6.7, p. 369) and get that $\partial[\phi(x_0, \hat{x}) + \delta_{\overline{Q(x_0)}}(\hat{x})] = \partial\phi(x_0, \hat{x}) + \partial\delta_{\overline{Q(x_0)}}(\hat{x})$. Therefore there exists $x^* \in \partial_2\phi(x_0, \hat{x})$ s.t. $-x^* \in \partial\delta_{\overline{Q(x_0)}}(\hat{x}) \Rightarrow ((-x^*, \hat{x})) = \sigma(-x^*, \overline{Q(x_0)}) \Rightarrow ((x^*, \hat{x})) = \inf\{((x^*, x)): x \in Q(x_0)\}$.

The last part of the theorem, concerning the upper semicontinuity of the value functional $v(\cdot)$, follows from theorem 3.3 of this paper and theorem 5, p. 52, of Aubin–Cellina [1].

Q.E.D.

Remark. If $\phi(x_0, x) = \phi(x) = \int_0^b f(t, x(t))\,dt$, then the conclusion of the theorem takes the following form: "$\hat{x}(\cdot) \in Q(x_0)$ solves (O_1) if and only if there exists $x^*(\cdot) \in M(T,H)$ *s.t.* $x^*(t) \in \partial f(t, \hat{x}(t))$ a.e. and $((x^*, \hat{x})) = \min\{((x^*, x)): x \in Q(x_0)\}$. This is a consequence of a result of Rockafellar [21] (theorem 22), which says that $x^* \in \partial\phi(\hat{x}) \Leftrightarrow x^*(t) \in \partial f(t, \hat{x}(t))$ a.e.

If instead we consider system (**), then we can get a stronger result concerning the value function $v(\cdot)$.

Theorem 4.2. *If the hypotheses of theorem* 3.5 *hold, then* $v: V \to \mathbb{R}$ *is continuous.*

Proof. This follows from theorem 3.5, combined with theorem 6, p. 53 of Aubin–Cellina [1].

Q.E.D.

Now we pass to a terminal cost control problem. So consider the following optimization problem:

$$(\underline{O_2}) \inf\{\psi(x(b)): x(\cdot) \in Q(x_0)\}.$$

Here $\psi: H \to \mathbb{R}$ is continuous, convex. We have the following "minimum principle", characterizing the solutions of $(\underline{O_2})$.

Theorem 4.3. *If the hypotheses of theorem* 3.1 *hold, then* $\hat{x}(\cdot) \in Q(x_0)$ *solves* $(\underline{O_2})$ *if and only if there exists* $x^* \in \partial\psi(\hat{x}(b))$ *s.t.* $(x^*, \hat{x}(b)) = \inf\{(x^*, y): y \in R(b)\}$.

Proof. Let $\phi: C(T,H) \to \mathbb{R}$ be defined by $\phi(x) = (\psi \circ e_b)(x) = \psi(x(b))$, $e_b(\cdot)$ being the evaluation at b map. From proposition 5.7, p. 27, of Ekeland–Temam [9], we know that $\partial\phi(x) = e_b^*\partial\psi(x(b))$. Rewrite $(\underline{O_2})$ as $\inf\{\phi(x): x \in Q(x_0)\}$ and apply theorem 4.1. So there exists $z^* \in \partial\phi(\hat{x}) = e_b^*\partial\psi(\hat{x}(b))$ *s.t.* $((z^*, \hat{x})) = \min\{((z^*, x)): x \in Q(x_0)\}$. But $z^* =$

$e_b^* x^*$, $x^* \in \partial\psi(\hat{x}(b))$. Hence we have:

$$\begin{aligned}((e_b^* x^*, \hat{x})) = (x^*, \hat{x}(b)) &= \min\{((e_b^* x^*, x)) : x \in Q(x_0)\} \\ &= \min\{(x^*, x(b)) : x \in Q(x_0)\} \\ &= \min\{(x^*, y) : y \in R(b)\}.\end{aligned}$$

Q.E.D.

5. Controllability

In this section we examine some controllability questions for the systems considered in this paper.

Let V_0, $V_b \subseteq H$ be nonempty. We say that (*) is "$V_0 V_b$-approximately controllable," if given $\varepsilon > 0$ there exists $x_0^\varepsilon \in V_0$ and admissible control $u^\varepsilon(\cdot) \in S_U^1$ *s.t.* if $x^\varepsilon(\cdot) \in C(T, H)$ is the corresponding trajectory emanating from x_0^ε, we have $d(x^\varepsilon(b), V_b) < \varepsilon$.

Theorem 5.1. *If the hypotheses of theorem* 3.1 *hold for every* $x_0 \in V_0$, *and system* (*) *is* $V_0 V_b$*-approximately controllable, then*

$$\sup_{x_0 \in V_0} \min_{|x^*| \leq 1} \sigma(x^*, R(x_0, b) - V_1) \geq 0, \text{ where } R(x_0 b) = \{x(b) : x(\cdot) \in Q(x_0)\}.$$

Proof. From the definition of $V_0 V_b$-approximate controllability, we know that there exist $\{x_0^n\}_{n \geq 1} \subseteq V_0$ *s.t.*

$$R(x_0^n, b) \cap (V_1 + \mathring{B}_{1/n}) \neq \varnothing.$$

So for every $x^* \in H$, $|x^*| \leq 1$, we have

$$\sigma(x^*, R(x_0^n, b)) \geq \inf_{v \in V_1} (x^*, v) - 1/n \Rightarrow \sigma(x^*, R(x_0^n, b) - V_1) \geq -1/n.$$

Let $n \to \infty$. We get $\sup_{x_0 \in V_0} \min_{|x^*| \leq 1} \sigma(x^*, R(x_0, b) - V_1) \geq 0$.

Q.E.D.

When the system is linear the above inequality condition is also sufficient for $V_0 V_b$-approximate controllability of the system.

So consider the following system.

$$\left.\begin{cases} \dot{x}(t) + A(t)x(t) = B(t)u(t) \text{ a.e. on } T \\ x(0) = x_0, \\ u \in U \subseteq L^2(Z) \end{cases}\right\} (\underline{\underline{***}})$$

Theorem 5.2. *If hypotheses* $H(A)$, $H(U)$, H_c *hold,* $B \in L^\infty(T, \mathscr{L}(Z, H))$, $V_0 = \{x_0\}$ *and* $\min_{|x^*| \leq 1} \sigma_{R(x_0, b) - V_1}(x^*) \geq 0$ *then* (***) *is* $V_0 V_b$*-approximately controllable.*

Proof. Suppose (***) is not $V_0 V_b$-approximately controllable. Then this means that $\overline{R(x_0, b)} \cap V_1 = \varnothing$. Note that because of the linearity of the system $\overline{R(x_0, b)}$ is convex, closed, while by hypothesis the same is true for V_b. Also $\overline{R(x_0, b)}$ is w-compact (see theorem 3.2). Thus we can apply the strong separation theorem and get $x^* \in H \backslash \{0\}$, $|x^*| \leqslant 1$ and $\delta > 0$ *s.t.*

$$\sigma(x^*, R(x_0, b)) < -\delta + \min_{v \in V_1} (x^*, v) \Rightarrow \sigma(x^*, R(x_0, b) - V_1) < -\delta$$

a contradiction to our hypothesis. So (***) is $V_0 V_b$-approximately controllable.

Q.E.D.

When $V_b = H$, then we say that (***) is "approximately controllable" from $V_0 = \{x_0\}$ in time b, if $\overline{R(x_0, b)} = H$.

Let $B \in L^\infty(T, \mathscr{L}(Z, X^*))$ and define $\hat{B} \in \mathscr{L}(L^2(Z), L^2(X^*))$ by setting $(\hat{B}u)(\cdot) = B(\cdot)u(\cdot)$. Also we need the following weaker form of $H(A)$. $\underline{H(A)_1}$: $A: T \times X \to X^*$ is an operator *s.t.* $A(t)(\cdot)$ is linear, continuous, strongly monotone and $\| A(\cdot) \|_{\mathscr{L}(X, X^*)} \in L^1_+$.

Theorem 5.3. *If $H(A_1)$ holds and $\hat{B}(U)$ is dense in $L^2(X^*)$, then (***) is approximately controllable from x_0 in time b.*

Proof. Without any loss of generality assume that $x_0 = 0$. Then consider the adjoint equation $\dot{v}(t) - A^*(t)v(t) = 0$ a.e. on T, $v(b) = h \in H$. From Lions [14] we know that this has a unique solution $v(\cdot) \in W(T) \subseteq L^2(X) \cap C(T, H)$. Therefore for every trajectory $x(\cdot)$ of (***) we have:

$$\int_0^b \langle \dot{x}(t), v(t) \rangle \, \mathrm{d}t + \int_0^b \langle A(t)x(t), v(t) \rangle \, \mathrm{d}t = \int_0^b \langle B(t)u(t), v(t) \rangle \, \mathrm{d}t.$$

Since $v(\cdot)$, $x(\cdot) \in W(T)$, we can integrate by parts and get that $(x(b), h) = \int_0^b \langle B(t)u(t), v(t) \rangle \, \mathrm{d}t = (\hat{B}u, v)_{L^2(X^*), L^2(X)} \Rightarrow h \in R(b, x_0)^\perp$ if and only if $(\hat{B}u, v)_{L^2(X^*), L^2(X)} = 0$ for all $u \in U$. But $\hat{B}(U)$ is dense in $L^2(X^*)$. So $v = 0$ on $T \Rightarrow h = v(b) = 0 \Rightarrow R(b, x_0)^\perp = \{0\} \Rightarrow \overline{R(b, x_0)} = H \Rightarrow$ (***) is indeed approximately controllable.

Q.E.D.

Acknowledgement

The author is grateful to the referee for his constructive criticisms and remarks. This research work is supported by NSF Grant DMS-8802688.

References

[1] Aubin J P and Cellina A, *Differential inclusions* (Berlin: Springer) (1984)

[2] Barbu V, Constrained control problems with convex cost in Hilbert space, *J. Math. Anal. Appl.* **55** (1976) 502–528

[3] Barbu V, *Nonlinear semigroups and differential equations in Banach spaces* (The Netherlands: Noordhoff International Publishing, Leyden) (1976)

[4] Delahaye J P and Denel J, The continuities of the point-to-set maps, definitions and equivalences, *Math. Program Study* **10** (1979) 8–12

[5] Dellacherie C, Ensembles analytiques: Théorèmes de separation et applications, *Semin. de probabilités IX, Univ. de Strasbourg, Lecture Notes in Math.* Vol. 465, (Berlin: Springer) (1975) 336–372

[6] Dunford N and Schwartz J, *Linear operators I* (New York: Wiley) (1958)
[7] Edgar G, Measurability in Banach spaces II, *Indiana Univ. Math. J.* **38** (1979) 559–579
[8] Egorov Yu V, Some problems in the theory of optimal control, *Sov. Math.* 3 (1962) 1080–1084
[9] Ekeland I and Temam R, *Convex analysis and variational problems* (New York: Elsevier) (1964)
[10] Klein E and Thompson A, *Theory of correspondences* (New York: Wiley) (1984)
[11] Kuratowski K, *Topology I* (New York: Academic Press) (1966)
[12] Laurent J P, *Approximation et optimisation* (Paris: Hermann) (1972)
[13] Lions J L, Optimisation pour certaines classes d' equations d' evolution non lineaires, *Ann. Mat. Pura Appl.* **72** (1966) 275–293
[14] Lions J L, *Optimal control of systems governed by partial differential equations* (Berlin: Springer) (1971)
[15] Mackenroth U, Time-optimal parabolic boundary control problems with state constraints, *Numer. Funct. Anal. Optimiz.* **3** (1981) 285–300
[16] Michael E, Continuous selections I, *Ann. Math.* **63** (1956), 361–381
[17] Moreau J, Intersection of moving convex sets in a normed space, *Math. Scand.* **36** (1975) 159–173
[18] Papageorgiou N S, On the theory of Banach space valued multifunctions. Part I: Integration and conditional expectation, *J. Multivar. Anal* **17** (1985) 185–206
[19] Papageorgiou N S, Convergence theorems for Banach space valued integrable multifunctions, *Int. J. Math. Math. Sci.* **10** (1987) 433–442
[20] Papageorgiou N S, Properties of the relaxed trajectories of evolution equations and optimal control, *SIAM J. Control Optim* **27** (1989) 267–288
[21] Rockafellar R T, Conjugate duality and optimization, *Conference Board of Math. Sciences Series* No. 16 (Philadelphia: SIAM Publications) (1974)
[22] Tanabe H, *Equations of evolution* (London: Pitman) (1979)
[23] Wagner D, Survey of measurable selection theorems, *SIAM J. Control Optim.* **15** (1977) 859–903