Czechoslovak Mathematical Journal

Antonella Fiacca; Nikolaos S. Papageorgiou; Francesca Papalini

On the existence of optimal controls for nonlinear infinite dimensional systems

Czechoslovak Mathematical Journal, Vol. 48 (1998), No. 2, 291–312

Persistent URL: `http://dml.cz/dmlcz/127417`

Czechoslovak Mathematical Journal, 48 (123) (1998), 291–312

ON THE EXISTENCE OF OPTIMAL CONTROLS FOR NONLINEAR INFINITE DIMENSIONAL SYSTEMS

Antonella Fiacca, Perugia, Nikolaos S. Papageorgiou, Athens, Francesca Papalini, Perugia

(Received September 6, 1995)

Abstract. We consider nonlinear systems with a priori feedback. We establish the existence of admissible pairs and then we show that the Lagrange optimal control problem admits an optimal pair. As application we work out in detail two examples of optimal control problems for nonlinear parabolic partial differential equations.

Keywords: evolution triple, optimal control, monotone operator, hemicontinuous operator, parabolic system, property (Q)

MSC 2000: 34G20

1. Introduction

In this paper we examine nonlinear systems with a priori feedback and establish the existence of optimal pairs for a Lagrange optimal control problem. This is achieved by assuming a convexity condition on an appropriate orientor field. It is well known from the finite dimensional theory (see the books of Berkovitz [4] and Cesari [6]) that such a convexity condition is indispensable in general. It was first introduced by Cesari [5] and is known as property (Q). In this note we show that under reasonably mild hypotheses on the data, this property is actually equivalent to simply assuming that the orientor field has closed and convex values. Our main results are three existence theorems. In the first two we establish the nonemptiness of the set of admissible pairs and in the third we show that the Lagrange optimal control problem admits an optimal pair. Our hypotheses are very general and natural. Finally, we illustrate our abstract theory by elaborating in detail two examples of optimal control problems monitored by nonlinear parabolic partial differential equations.

In addition to extending the finite dimensional theory presented in the books of Berkovitz [4] and Cesari [6], our research also extends the infinite dimensional work of Ahmed-Teo [1] who treated semilinear systems under more restrictive hypotheses on the data (see Section 5.4).

2. Mathematical preliminaries

Let (Ω, Σ, μ) be a measure space and $(X, \|\cdot\|)$ a separable Banach space. Throughout this paper, we will use the following notation:

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

$$P_{(w)k(c)}(X) = \{A \subset X\colon\ A \text{ nonempty, (weakly) compact (convex)}\}.$$

A multifunction $F\colon \Omega \to P_f(X)$ is said to be measurable if, for all $x \in X$, the function $\omega \mapsto d(x, F(\omega)) = \inf\{\|x - z\|\colon\ z \in F(\omega)\}$ is measurable. A multifunction $F\colon \Omega \to P_f(X)$ is said to be graph measurable if $\operatorname{Gr} F = \{(\omega, x) \in \Omega \times X\colon\ x \in F(\omega)\} \in \Sigma \times B(X)$, with $B(X)$ being to the Bore σ-field of X. For $P_f(X)$-valued multifunctions, measurability implies graph measurability, while the converse is true if there is a σ-finite measure μ on (Ω, Σ) with respect to which Σ is complete (actually, graph measurability implies measurability under the more general condition that $\Sigma = \hat{\Sigma}$ (= the universal σ-field) with no explicit reference to any measure on Σ).

We define $S_F^p (1 \leqslant p \leqslant \infty)$ to be the set of all $L^p(\Omega, X)$-selectors of $F(\cdot)$, i.e. $S_F^p = \{f \in L^p(\Omega, X)\colon\ f(\omega) \in F(\omega)\ \mu\text{-a.e}\}$. Note that for a graph measurable multifunction $F\colon \Omega \to P_f(X)$, S_F^p is nonempty if and only if the function $\omega \mapsto \inf\{\|z\|\colon\ z \in F(\omega)\}$ belongs to $L^p(\Omega, \mathbb{R}^+)$.

On $P_f(X)$ we can define a generalized metric, known in the literature as the "Hausdorff metric", by setting, for $A, B \in P_f(X)$,

$$h(A,B) = \max\Big\{h^*(A,B) = \sup\{d(a,B)\colon\ a \in A\},\ h^*(B,A) = \sup\{d(b,A)\colon\ b \in B\}\Big\}$$

(recall that $d(a,B) = \inf\{\|a - b\|\colon\ b \in B\}$; similarly for $d(b,A)$). The metric space $(P_f(X), h)$ is complete. A multifunction $F\colon X \to P_f(X)$ is said to be Hausdorff continuous (H-continuous) if it is continuous from X into $(P_f(X), h)$.

Given $\varepsilon > 0$ and $x \in X$, we put $\dot{B}_\varepsilon(x) = \{y \in X\colon\ \|x - y\| < \varepsilon\}$ and if $A \subset X$, we denote by A_ε the set $A_\varepsilon = \{y \in X\colon\ d(y, A) < \varepsilon\}$.

Let Y, Z be Hausdorff topological spaces. A multifunction $F\colon Y \to 2^Z \setminus \{\emptyset\}$ is said to be lower semicontinuous (l.s.c.) (upper semicontinuous (u.s.c.)), if for all $U \subset Z$ open $F^-(U) = \{y \in Y\colon\ F(y) \cap U \neq \emptyset\}$ (resp. $F^+(U) = \{y \in Y\colon\ F(y) \subset U\}$) is open in Y.

Now let H be a Hilbert space and let X be a dense subspace of H carrying the structure of a separable, reflexive Banach space, which embeds into H continuously. Identifying H with its dual (pivot space), we have $X \hookrightarrow H \hookrightarrow X^*$, with all embeddings being continuous and dense. Such a triple of spaces is known in the literature as "evolution triple" or "Gelfand triple" (see Zeidler [15]). We will also assume that the embedding of X into H is also compact (in fact, this implies that $H \hookrightarrow X^*$ is compact, too). To have a concrete example in mind, let m be a positive integer and $2 \leqslant p \leqslant \infty$. Let $Z \subset \mathbb{R}^N$ be a bounded domain and set $X = W_0^{m,p}(Z, \mathbb{R})$, $H = L^2(Z, \mathbb{R})$ and $X^* = W^{-m,q}(Z, \mathbb{R})$ where $\frac{1}{p} + \frac{1}{q} = 1$. Then from the Sobolev embedding theorem, we know that (X, H, X^*) is an evolution triple and all embeddings are compact. By $|\cdot|(\|\cdot\|_*)$ we will denote the norm of H (resp. of X^*). Also by $(\cdot,\cdot)$ we will denote the inner product of H and by $\langle\cdot,\cdot\rangle$ the duality brackets of the pair (X, X^*). The two are compatible in the sense that $\langle\cdot,\cdot\rangle /_{X\times H} = (\cdot,\cdot)$.

Let $1 < p, q < \infty, \frac{1}{p} + \frac{1}{q} = 1, T = [0, b]$; we define

$$W_{pq}(T) = \{x \in L^p(T, X)\colon\ \dot{x} \in L^q(T, X^*)\}.$$

The derivative involved in this definition is understood in the sense of vector valued distributions. Equipped with the norm $\|x\|_{W_{pq}} = [\|x\|^2_{L^p(T,X)} + \|\dot{x}\|^2_{L^q(T,X^*)}]^{\frac{1}{2}}$, the space $W_{pq}(T)$ becomes a separable, reflexive Banach space. It is well known that $W_{pq}(T)$ embeds continuously into $C(T, H)$, i.e. every element in $W_{pq}(T)$ has a unique representative in $C(T, H)$. Since we have assumed that $X \hookrightarrow H$ compactly, we have that $W_{pq}(T) \hookrightarrow L^p(T, H)$ compactly (see [15], p. 450).

3. Optimal control problem

Let $T = [0, b]$ and let (X, H, X^*) be an evolution triple as in Section 2, with all embeddings being compact. We consider the following Lagrange optimal control problem:

$$\begin{aligned} (1) \qquad & J(x,u) = \int_0^b L(t, x(t), u(t))\,\mathrm{d}t \to \inf J = m;\\ & \text{s.t. } \dot{x}(t) + A(t, x(t)) = f(t, x(t), u(t)) \text{ a.e.};\\ & x(0) = x_0;\\ & u(t) \in U(t, x(t)) \text{ a.e., } u(\cdot) \text{ is measurable.}\end{aligned}$$

The control space is modelled by a separable reflexive Banach space $(Y, \|\cdot\|_Y)$. We will need the following hypotheses on the data of (1):

$H(A)$: $A\colon T \times X \to X^*$ is an operator such that

(1) for all $x \in X$, $t \mapsto A(t,x)$ is measurable;
(2) for a.e. $t \in T$, $x \mapsto A(t,x)$ is hemicontinuous and monotone;
(3) $\exists\, a_1 \in L^q(T,\mathbb{R}^+)$ and $c_1 \geqslant 0$ such that $\|A(t,x)\|_* \leqslant a_1(t) + c_1\|x\|^{p-1}$ for all $x \in X$ and for a.e. $t \in T$, with $2 \leqslant p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$;
(4) $\exists\, c \geqslant 0$ such that $\langle A(t,x), x\rangle \geqslant c\|x\|^p$ for all $x \in X$ and for a.e. $t \in T$.

$H(f)$: $f\colon T \times H \times Y \to H$ is a function such that
(1) $\forall\,(x,u) \in H \times Y$, $t \mapsto f(t,x,u)$ is measurable;
(2) for a.e. $t \in T$, $(x,u) \mapsto f(t,x,u)$ is continuous;
(3) $\forall\, a, c \in L^q(T,\mathbb{R}^+)$ such that
$|f(t,x,u)| \leqslant a(t) + c(t)[|x|^{2/q} + \|u\|_Y]$ for all $(x,u) \in H \times Y$ and for a.e. $t \in T$,

$H(U)$: $U\colon T \times H \to P_f(Y)$ is a multifunction such that
(1) $(t,x) \mapsto U(t,x)$ is graph measurable;
(2) for a.e. $t \in T$, $x \mapsto U(t,x)$ is H-continuous;
(3) $\exists\, c_2 > 0$: $|U(t,x)| = \{\|u\|_Y\colon u \in U(t,x)\} \leqslant c_2(1 + |x|^{2/q})$ for all $x \in X$ and for a.e. $t \in T$.

Remark. Note that hypothesis $H(U)(2)$ implies that $x \mapsto U(t,x)$ is l.s.c. and has a closed graph.

$H(L)$: $L\colon T \times H \times Y \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is an integrand such that
(1) $(t,x,u) \mapsto L(t,x,u)$ is measurable;
(2) for a.e. $t \in T$, $(x,u) \mapsto L(t,x,u)$ is l.s.c.;
(3) $\exists\, \psi \in L^1(T,\mathbb{R})$ and $\beta \geqslant 0$:
$\psi(t) - \beta|x| \leqslant L(t,x,u)$ for all $u \in U(t,x)$, $x \in H$ and for a.e. $t \in T$.

Moreover, we introduce the following convexity hypothesis on the data:

H_c: the multifunction $Q\colon T \times H \to 2^{H\times\mathbb{R}}$, defined by

$$Q(t,x) = \{[h,\eta] \in H \times \mathbb{R}\colon \exists\, u \in U(t,x) \text{ with } L(t,x,u) \leqslant \eta \text{ and } h = f(t,x,u)\}$$

is such that $x \mapsto Q(t,x)$ has property (Q) of Cesari (see [6], p. 292).

Remark. Recall that, in our case, property (Q) of Cesari means that for all $t \in T$

$$Q(t,x) = \bigcap_{\delta>0} \overline{\mathrm{co}} \bigcup_{x' \in \dot{B}_\delta(x)} Q(t,x'),\ \forall\, x \in H.$$

A pair $[x,u] \in W_{pq}(T) \times L^1(T,Y)$ is called an "admissible pair" if the two functions satisfy all the constraints of problem (1). In that case x is called an admissible state (or trajectory) and u and admissible control. An admissible pair which minimizes the Lagrange cost functional among pairs, is called an "optimal pair".

The convexity hypothesis which we mentioned in the introduction and which is essential in establishing the existence of an optimal pair for (1) is hypothesis H_c.

The next simple lemma tells us that, under reasonable hypotheses on f, L and U, this hypothesis is actually equivalent to simply assuming that for every $(t,x) \in T \times H$, the set $Q(t,x)$ is closed and convex.

Lemma 1. *If given $x \in H$ and $\varepsilon > 0$ there exists a $\delta = \delta(x) > 0$ such that $\forall\ (x,u)$, $(x',u') \in \operatorname{Gr} U(t,\cdot)$ with $|x - x'| < \delta, \|u - u'\|_Y < \delta \Rightarrow |f(t,x,u) - f(t,x',u')| < \varepsilon$ and $L(t,x',u) > L(t,x,u) - \varepsilon$ and $U(t,\cdot)$ is H-continuous then hypothesis H_c is equivalent to assuming that for all $(t,x) \in T \times H$, $Q(t,x)$ is closed and convex.*

P r o o f. Clearly we only need to show that if $Q(t,x)$ is closed and convex for all $(t,x) \in T \times H$, then hypothesis H_c is satisfied. According to Theorem 5.5 of [9], if we show that $Q(t,\cdot)$ is h^*-u.s.c. then we are done. So we need to show that if $x_n \to x$ in H, then $h^*(Q(t,x_n), Q(t,x)) \to 0$ as $n \to \infty$. Fix $(t,x) \in T \times H$, let $\varepsilon > 0$ be given and let $\delta = \delta(x) > 0$ be the one postulated by our hypothesis on $f(t,\cdot,\cdot)$ and $L(t,\cdot,\cdot)$. Since by hypothesis $U(t,\cdot)$ is H-continuous, we can find $n_0 \geqslant 1$ such that for $n \geqslant n_0$ we have $|x_n - x| < \delta$ and $h(U(t,x_n), U(t,x)) < \delta$. Given $[\nu_n, \eta_n] \in Q(t,x_n)$, by definition we can find $u_n \in U(t,x_n)$ such that $\nu_n = f(t,x_n,u_n)$ and $L(t,x_n,u_n) \leqslant \eta_n$. Then $d_Y(u_n, U(t,x)) < \delta$ for $n \geqslant n_0$. So we can find $u'_n \in U(t,x)$ such that $\|u_n - u'_n\|_Y < \delta$. Hence from our hypothesis on $f(t,\cdot,\cdot)$ and $L(t,\cdot,\cdot)$ we have

$$|f(t,x_n,u_n) - f(t,x,u'_n)| < \varepsilon \text{ and } L(t,x,u'_n) - \varepsilon < L(t,x_n,u_n) \leqslant \eta_n, \quad \forall\ n \geqslant n_0.$$

Therefore $[\nu_n, \eta_n] \in Q(t,x) + \varepsilon B_1$, where $B_1 = \{[\nu,\eta] \in H \times \mathbb{R}\colon\ |\nu| + |\eta| < 1\}$, $\forall\ n \geqslant n_0$. Since $[\nu_n, \eta_n] \in Q(t,x_n)$ was arbitrary, we deduce that $h^*(Q(t,x_n)Q(t,x)) \leqslant \varepsilon$, $\forall\ n \geqslant n_0$. So $Q(t,\cdot)$ is h^*-u.s.c. and then $Q(t,\cdot)$ has property (Q). □

Next we will establish the nonemptiness of the set $P(x_0) \subset W_{pq}(T) \times L^1(T,Y)$ of admissible pairs.

Theorem 1. *If hypothesis $H(A)$, $H(f)$, $H(U)$ with (2) replaced by (2′): for a.e. $t \in T$, $x \mapsto U(t,x)$ is l.s.c. hold and for all $(t,x) \in T \times H$, $f(t,x,U(t,x))$ is closed, then $P(x_0) \neq \emptyset$ and $\operatorname{proj}_1 P(x_0) = P_1(x_0)$ is relatively sequentially w-compact in $W_{pq}(T)$.*

P r o o f. Let $F\colon\ T \times H \to P_f(H)$ be defined by

$$F(t,x) = f(t,x,U(t,x)).$$

First we will show that for every $x\colon T\to H$ Lebesgue measurable, the multifunction $t\mapsto F(t,x(t))$ is Lebesgue measurable, too. To this end let $u_n\colon T\to Y$, $n\geqslant 1$, be Lebesgue measurable selectors of $U(\cdot,x(\cdot))$ such that $U(t,x(t))=\overline{\{u_n(t)\}}_{n\geqslant 1}$. The existence of such a sequence of measurable selectors follows from hypothesis $H(U)(1)$ and Theorem 4.2 of Wagner [14]. Then hypothesis $H(f)(2)$ tells us that

$$F(t,x(t))=\overline{\{f(t,x(t),u_n(t))\}}_{n\geqslant 1}.$$

Because of $H(f)(1)$ and (2), $(t,x,u)\mapsto f(t,x,u)$ is jointly measurable and so, for every $n\geqslant 1$, $t\mapsto f(t,x(t),u_n(t))$ is measurable. A final appeal to Theorem 4.2 of [14] yields the Lebesgue measurability of $t\to F(t,x(t))$. Next we will show that for every $t\in T$, $x\mapsto F(t,x)$ is l.s.c. To prove this we need to show that if $x_n\to x$, then $F(t,x)\subset \underline{\lim}F(t,x_n)$. So let $\nu\in F(t,x)$, then $\exists\ u\in U(t,x)$ such that $\nu=f(t,x,u)$.

Because of hypothesis $H(U)(2')$ we get that there exist $u_n\in U(t,x_n)$ such that $u_n\to u$ in Y. Set $\nu_n=f(t,x_n,u_n)$. Then $\nu_n\to f(t,x,u)=\nu$ and $\nu_n\in F(t,x_n)$. Hence we have shown that $x\mapsto F(t,x)$ is l.s.c.

Because of hypothesis $H(f)(3)$ and $H(U)(3)$ we have

$$\begin{aligned}|F(t,x)|&\leqslant a(t)+c(t)|x|^{2/q}+c(t)c_2+c(t)c_2|x|^{2/q}\\&\leqslant \hat a(t)+\hat c(t)|x|^{2/q}\ \text{a.e. on } T,\ \text{where } \hat a(\cdot)=a(\cdot)+c_2\ c(\cdot)\in L^q(T,\mathbb{R}^+)\\&\qquad\qquad\qquad\qquad\text{and } \hat c(\cdot)=c(\cdot)+c_2\ c(\cdot)\in L^q(T,\mathbb{R}^+).\end{aligned}$$

Now consider the evolution inclusion

$$\begin{aligned}&\dot x(t)+A(t,x(t))\in F(t,x(t))\ \text{a.e.};\\&x(0)=x_0.\end{aligned}\tag{2}$$

By a solution of (2) we mean a function $x\in W_{pq}(T)$ such that $\dot x(t)+A(t,x(t))=f(t)$ a.e., $x(0)=x_0$ with $f\in L^q(T,H)$, $f(t)\in F(t,x(t))$ a.e. Recall that $W_{pq}(T)$ embeds into $C(T,H)$ continuously, so that the initial condition makes sense. First we will obtain an a priori bound for the solutions of (2). So let $x\in W_{pq}(T)$ be such a solution (denote the solution set of (2) by $P_1'(x_0)$). From

$$\dot x(t)+A(t,x(t))=f(t)\ \text{a.e.},\ x(0)=x_0,\ \text{with } f\in S^q_{F(\cdot,x(\cdot))},$$

it follows that

$$\langle\dot x(t),x(t)\rangle+\langle A(t,x(t)),x(t)\rangle=(f(t),x(t))\ \text{a.e.}$$

This implies

$$\frac12\frac{\mathrm d}{\mathrm dt}|x(t)|^2+c\|x(t)\|^p\leqslant |f(t)|\ |x(t)|\leqslant |f(t)|\beta_1\|x(t)\|\ \text{a.e.},$$

where $\beta_1 > 0$ is such that $|\cdot| \leqslant \beta_1 \|\cdot\|$ (recall that X embeds into H continuously).

Apply Cauchy's inequality with $\varepsilon > 0$ and right hand side to get

$$\frac{\mathrm{d}}{\mathrm{d}t}|x(t)|^2 + 2c\|x(t)\|^p \leqslant 2\beta_1\Big(\frac{\varepsilon^q}{q}|f(t)|^q + \frac{1}{\varepsilon^p p}\|x(t)\|^p\Big) \text{ a.e.}$$

Let $\varepsilon > 0$ be such that $\frac{2\beta_1}{\varepsilon^p p} = 2c$. Then we have

$$\frac{\mathrm{d}}{\mathrm{d}t}|x(t)|^2 \leqslant \hat{c}_0|f(t)|^q \text{ a.e.}$$

with $\hat{c}_0 = \frac{2\beta_1}{Q}\big(\frac{\beta_1}{cp}\big)^{1/(p-1)}$, hence by integrating we have

$$\begin{aligned}|x(t)|^2 &\leqslant |x_0|^2 + \hat{c}_0\int_0^t (\hat{a}(s) + \hat{c}(s)|x(s)|^{2/q})^q \,\mathrm{d}s\\ &\leqslant |x_0|^2 + 2^{q-1}\hat{c}_0\int_0^t \hat{a}(s)^q \,\mathrm{d}s + 2^{q-1}\hat{c}_0\int_0^t \hat{c}(s)^q|x(s)|^2 \,\mathrm{d}s.\end{aligned}$$

Applying Gronwall's lemma, we get $M_1 > 0$ such that for all $x(\cdot) \in P_1'(x_0)$ we have

$$\|x(\cdot)\|_{C(T,H)} \leqslant M_1. \tag{3}$$

Then we get

$$\frac{\mathrm{d}}{\mathrm{d}t}|x(t)|^2 + 2c\|x(t)\|^p \leqslant 2|f(t)|\beta_1 M_1 \text{ a.e.}$$

and hence

$$2c\int_0^b \|x(s)\|^p \,\mathrm{d}s \leqslant |x_0|^2 + 2\beta_1 M_1\int_0^b (\hat{a}(s) + \hat{c}(s)M_1^{2/q})\,\mathrm{d}s.$$

Therefore there exists $M_2 > 0$ such that for all $x(\cdot) \in P_1'(x_0)$

$$\|x(\cdot)\|_{L^p(T,X)} \leqslant M_2. \tag{4}$$

Since $\dot{x}(t) = -A(t, x(t)) + f(t)$ a.e., by using $H(A)(3)$, the fact that $|f(t)| \leqslant \hat{a}(t) + \hat{c}(t)|x|^{2/q}$ a.e. and bounds (3) and (4) above, we get an $M_3 > 0$ such that for all $x(\cdot) \in P_1'(x_0)$

$$\|\dot{x}(\cdot)\|_{L^q(T,X^*)} \leqslant M_3. \tag{5}$$

From (4) and (5) above we have that $P_1'(x_0)$ is bounded in $W_{pq}(T)$, hence relatively sequentially weakly compact.

Now because of (3), without any loss of generality we may assume that

$$|F(t,x)| \leqslant \hat{a}(t) + \hat{c}(t)M_1^{2/q} = \varphi(t) \text{ a.e.,}$$

$\varphi(\cdot) \in L^q(T,\mathbb{R}^+)$. Let $V = \{h \in L^q(T,H)\colon\ |h(t)| \leqslant \varphi(t) \text{ a.e.}\}$ and let $\xi\colon\ L^q(T,T) \to W_{pq}(T)$ be the map which to each $h \in L^q(T,H)$ assigns the unique solution of $\dot{x}(t) + A(t,x(t)) = h(t)$ a.e., $x(0) = x_0$ (see [3], Th. 4.2, p. 167). Let $K = \xi(V)$. We claim that K is weakly compact in $W_{pq}(T)$. Indeed, by a similar a priori estimation as above we get that K is bounded in $W_{pq}(T)$, hence relatively sequentially weakly compact in $W_{pq}(T)$. Let $\{x_n\}_{n\geqslant 1} \subset K$ and assume that $x_n \to x$ weakly in $W_{pq}(T)$. Be definition we have

$$\dot{x}_n(t) + A(t,x_n(t)) = f_n(t) \text{ a.e.}$$
$$x_n(0) = x_0.$$

By passing to a subsequence if necessary, we may assume that $f_n \to f$ weakly in $L^q(T,H)$. Let $\hat{A}(\cdot)$ be the Nemitsky (superposition) operator corresponding to $A(\cdot,\cdot)$. So $\hat{A}(x)(\cdot) = A(\cdot,x(\cdot))$ for every measurable $x\colon\ T \to X$. Because of hypothesis $H(A)(3)$ and bound (4) above, we see that $\{\hat{A}(x_n)\}_{n\geqslant 1}$ is bounded in the reflexive Lebesgue-Bochner space $L^q(T,X^*)$. So we may assume that $\hat{A}(x_n) \to v$ weakly in $L^q(T,X^*)$. Let $((\cdot,\cdot))_0$ be the duality brackets for the pair $(L^p(T,X), L^q(T,X^*))$ and $((\cdot,\cdot))_{L^2(T,H)}$ the inner product on the separable Hilbert space $L^2(T,H)$. From the properties of the evolution triple we have that $((\cdot,\cdot))_{0/L^p(T,H)\times L^2(T,H)} = ((\cdot,\cdot))_{L^2(T,H)}$. From the integration by parts formula for functions in W_{pq} (T) ([15], Proposition 23.23) we have

$$\int_0^b \langle \dot{x}_n(s) - \dot{x}(s), x_n(s) - x(s)\rangle \,\mathrm{d}s = \frac{1}{2}|x_n(b) - x(b)|^2$$

and hence we get

$$((\dot{x}_n, x_n - x))_0 = \frac{1}{2}|x_n(b) - x(b)|^2 + ((\dot{x}, x_n - x))_0.$$

Note that $((\dot{x}, x_n - x))_0 \to 0$. Also $\{x_n\}_{n\geqslant 1}$ is bounded in $W_{pq}(T)$ and since the latter embeds compactly in $L^p(T,H)$ and continuously in $C(T,H)$, we also have $\frac{1}{2}|x_n(b) - x(b)|^2 \to 0$ as $n \to \infty$. Therefore $((\dot{x}_n, x_n - x))_0 \to 0$ as $n \to \infty$. So we have

$$((\hat{A}(x_n), x_n - x))_0 = ((f_n, x_n - x))_{L^2(T,H)} - ((\dot{x}_n, x_n - x))_0 \to 0 \text{ as } n \to \infty.$$

But because of hypothesis $H(A)$, $\hat{A}$ is hemicontinuous, monotone, hence has property (M) (see [15], pp. 583–584). Thus $v = \hat{A}(x)$, i.e. $\hat{A}(x_n) \to \hat{A}(x)$ weakly in $L^q(T,X^*)$. Therefore for every $h \in L^p(T,X)$ we have

$$((\dot{x}_n, h))_0 + ((\hat{A}(x_n), h))_0 = ((f_n, h))_0$$

and

$$((\dot{x},h))_0 + ((\hat{A}(x),h))_0 = ((f,h))_0.$$

It follows that

$$\dot{x}(t) + A(t,x(t)) = f(t) \text{ a.e., } x(0) = x_0 \text{ and } f \in V$$

and so $x \in K$, i.e. K is w-compact in $W_{pq}(T)$ and $\xi(\cdot)$ is sequentially weakly continuous. Let $\hat{K} = \overline{\mathrm{co}}\, K \in P_{kc}(L^p(T,H))$ and define $R\colon \hat{K} \to P_f(L^q(T,H))$ by

$$R(x) = S^q_{F(\cdot,x(\cdot))}.$$

Using Fryszkowski's selection theorem (see [8]), we can find a continuous function $r\colon \hat{K} \to L^q(T,H)$ such that $r(x) \in R(x)$ for all $x \in \hat{K}$. Then $\varrho = \xi_0 r\colon \hat{K} \to \hat{K}$ is continuous. Apply Schauder's fixed point theorem to get a point $x \in \hat{K}$ such that $\varrho(x) = x$. Clearly $x(\cdot) \in P_1'(x_0)$.

Finally, let $x \in \xi(g)$, $g \in L^q(T,H)$, $g(t) \in F(t,x(t))$ a.e. and let $\Delta\colon T \to 2^Y$ be defined by

$$\Delta(t) = \{u \in U(t,x(t))\colon\ g(t) = f(t,x(t),u)\}.$$

Clearly $\Delta(t) \neq \emptyset$ a.e. and from hypotheses $H(f)$ and $H(U)$ we have that $\mathrm{Gr}\,\Delta \in B(T) \times B(Y)$.

Apply Aumann's selection theorem (see [14], Theorem 5.10) to get $u\colon T \to Y$ measurable such that $u(t) \in \Delta(t)$ a.e.. Then $g(t) = f(t,x(t),u(t))$ a.e. and so $x(\cdot) \in P_1(x_0) \neq \emptyset$. Finally, since by the above argument $P_1(x_0) = P_1'(x_0)$, we have that $P_1'(x_0)$ is relatively w-compact in $W_{pq}(T)$ (hence relatively compact in $L^p(T,H)$). □

Although the above result suffices for the optimal control problem considered in the paper, we also mention another result, which offers more. We will need the following hypotheses on the data:

$H(f)_1$: $f\colon T \times H \times Y \to H$ is a function such that
- (1) $\forall\ (x,u) \in H \times Y$, $t \mapsto f(t,x,u)$ is measurable;
- (2) for a.e. $t \in T$, $(x,u) \mapsto f(t,x,u)$ is uniformly continuous;
- (3) $\exists\ a,\ c \in L^q(T,\mathbb{R}^+)$ such that
  $|f(t,x,u)| \leqslant a(t) + c(t)[|x|^{2/q} + \|u\|_Y]$ for all $(x,u) \in H \times Y$ and for a.e. $t \in T$.

$H(U)_1$: $U\colon T \times H \to P_f(Y)$ is a multifunction such that
- (1) $(t,x) \mapsto U(t,x)$ is graph measurable;
- (2) for a.e. $t \in T$, $x \mapsto U(t,x)$ is h^*-u.s.c.;

(3) $\exists\, c_2 > 0$: $|U(t,x)| \leqslant c_2(1+|x|^{2/q})$ for all $x \in H$ and for a.e. $t \in T$.

H_c': for almost all $t \in T$ and for all $x \in H$, the set $f(t,x,U(t,x))$ is closed and convex.

Then we have the following result concerning the sets $P_1(x_0)$ and $P(x_0)$.

Theorem 2. *If hypotheses $H(A)$, $H(f)_1$, $H(U)_1$ and H_c' hold, then $P(x_0) \neq \emptyset$ and $P_1(x_0)$ is weakly compact in $W_{pq}(T)$.*

Proof. Let $F\colon T \times H \to P_{fc}(H)$ be the multifunction defined by $F(t,x) = f(t,x,U(t,x))$. As in the proof of Theorem 1, we can show that if $x\colon T \to H$ is Lebesgue measurable, then $t \mapsto F(t,x(t))$ is Lebesgue measurable, too. Next we will show that, for a.e. $t \in T$, we have

$$\bigcap_{\delta>0} \overline{\mathrm{co}}\, f(t, \dot{B}_\delta(x), U(t, \dot{B}_\delta(x))) = f(t,x,U(t,x)),\ x \in H.$$

To this end, note that because of hypothesis $H(f)_1(2)$, given $\varepsilon > 0$, we can find $\delta_1 > 0$ such that $f(t, \dot{B}_{\delta_1}(x), U(t,x)_{\delta_1}) \subset f(t,x,U(t,x))_\varepsilon$, while from hypothesis $H(U)_1(2)$, we can find $0 < \delta \leqslant \delta_1$ such that $U(t, \dot{B}_\delta(x)) \subset U(t,x)_{\delta_1}$. So using hypothesis H_c', we get

$$\bigcap_{\delta>0} \overline{\mathrm{co}}\, f(t, \dot{B}_\delta(x), U(t, \dot{B}_\delta(x))) = f(t,x,U(t,x)) = F(t,x),\ x \in H.$$

Now we can show that the graph of $F(t,\cdot)$ is sequentially closed in $H \times H_w$ (H_w denotes the space H equipped with the weak topology). Indeed, let $[x_n, \nu_n] \to [x,\nu]$ in $H \times H_w$, $\nu_n \in F(t,x_n)$, $n \geqslant 1$. Then there exist $u_n \in U(t,x_n)$ such that $\nu_n = f(t,x_n,u_n)$. Now given $\delta > 0$ we can find $n_0(\delta) = n_0 \geqslant 1$ such that $x_n \in \dot{B}_\delta(x)$ and $u_n \in U(t, \dot{B}_\delta(x))$. Hence for $n \geqslant n_0$ we have

$$\nu_n = f(t,x_n,u_n) \in f(t, \dot{B}_\delta(x), U(t, \dot{B}_\delta(x))).$$

Using the property established above together with Mazur's lemma (since $\nu_n \to \nu$ in H_w) we get that $\nu \in F(t,x)$.

So $\operatorname{Gr} F(t,\cdot)$ is sequentially closed in $H \times H_w$.

With the same priori estimation as in the proof of Theorem 1, we can show that $P_1(x_0)$ is bounded in $W_{pq}(T)$ and without any loss of generality we can assume that

$$|F(t,x)| \leqslant \varphi(t) \text{ a.e., where } \varphi(\cdot) \in L^q(T, \mathbb{R}^+).$$

Let $V = \{h \in L^q(T,H)\colon\ |h(t)| \leqslant \varphi(t) \text{ a.e.}\}$. This set, furnished with the weak topology, is compact metrizable. Let $R\colon\ V \to P_{fc}(V)$ be defined by $R(h) = S^1_{F(\cdot,\xi(h)(\cdot))}$, where $\xi\colon\ L^p(T,H) \to W_{pq}(T)$ is the solution map, as in the proof of Theorem 1.

We claim that $R(\cdot)$ is u.s.c. on V equipped with the relative weak $L^q(T,H)$-topology. Since V topologized like that is w-compact metrizable, it suffices to show that $\operatorname{Gr} R$ is closed in $V \times V$. So let $[h_n, g_n] \in \operatorname{Gr} R$, $n \geqslant 1$ and assume that $[h_g, g_n] \to [h, g]$ weakly in $L^q(T,H) \times L^q(T,H)$. Then by definition $g_n(t) \in F(t, \xi(h_n)(t))$ a.e. From the proof of Theorem 1 we know that $\xi(h_n) \to \xi(h)$ in $C(T,H)$. Then using Theorem 3.1 of [13], we get

$$g(t) \in \overline{\operatorname{co}}\, w - \overline{\lim}\, F(t, \xi(h_n)(t)) \subset F(t, \xi(h)(t)) \text{ a.e.}$$

the last inclusion being a consequence of the fact that $\operatorname{Gr} F(t,\cdot)$ is sequentially closed in $H \times H_w$. So $[h,g] \in \operatorname{Gr} \mathbb{R}$ and then $R(\cdot)$ is u.s.c. Applying the Kakutani-KyFan fixed point theorem we get that there exists $h \in V$ such that $h \in R(h)$. Then $x = \xi(h) \in P_1(x_0)$. As in the proof of Theorem 1, via Aumann's selection theorem, we can get $u\colon\ T \to Y$ measurable such that $[x,u] \in P(x_0)$, i.e. $P(x_0) \neq \emptyset$. Finally, let $x_n \in P_1(x_0)$ and assume that $x_n \to x$ weakly in $W_{pq}(T)$. Then $x_n = \xi(g_n)$ with $g_n \in F(t, x_n(t))$. By passing to a subsequence if necessary, we may assume that $g_n \to g$ weakly in $L^q(T,H)$. Note that, since $x_n \to x$ weakly in $W_{pq}(T), x_n \to x$ in $L^p(T,H)$ and by passing to a subsequence if necessary we can achieve $x_n(t) \to x(t)$ a.e. in H. Then as before via Theorem 3.1 of [13] we get $g(t) \in F(t, x(t))$ a.e. Also $x = \xi(g)$ and hence $x \in P_1(x_0)$; so $P_1(x_0)$ is w-closed in $W_{pq}(T)$, and thus, since $P_1(x_0)$ is bounded in $W_{pq}(T)$, we have that $P_1(x_0)$ is w-compact in $W_{pq}(T)$. $\square$

Remark. Since $W_{pq}(T)$ embeds compactly into $L^p(T,H)$, we have that $P_1(x_0)$ is compact in $L^p(T,H)$ and also in $L^2(T,H)$ (recall that $p \geqslant 2$).

Now we can present our result on the existence of optimal pairs for problem (1).

Theorem 3. *If hypotheses $H(A)$, $H(f)$, $H(U)$, H_c, $H(L)$ hold and the set $f(t,x, U(t,x))$ is closed for every $(t,x) \in T \times H$, then problem (1) admits an optimal pair.*

P r o o f. Let $\Gamma\colon\ T \times X \times X^* \to 2^Y$ be defined by

$$\Gamma(t,x,\nu) = \{u \in U(t,x)\colon\ \nu + A(t,x) = f(t,x,u)\}.$$

Then $\operatorname{Gr}\Gamma = \{(t,x,\nu,u) \in T \times X \times X^* \times Y\colon\ (t,x,u) \in \operatorname{Gr} U\colon\ \nu + A(t,x) = f(t,x,u)\}$, $\operatorname{Gr}\Gamma \in B(T) \times B(X) \times B(X^*) \times B(Y)$ (cf. hypotheses $H(A)$, $H(f)$ and $H(U)$). Define $p\colon\ T \times X \times X^* \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ by

$$p(t,x,\nu) = \inf\{L(t,x,u)\colon\ u \in \Gamma(t,x,\nu)\},$$

where as always we use the convention that the infimum over the empty set is $+\infty$.

Now we will establish some properties of p.

(i) $(t,x,\nu)\mapsto p(t,x,\nu)$ is superpositionally measurable, i.e. every $x\colon T\to X$ and $\nu\colon T\to X^*$ Lebesgue measurable functions, $t\mapsto p(t,x(t),\nu(t))$ is Lebesgue measurable, too.

So let $x\colon T\to X$ and $\nu\colon T\to X^*$ be two Lebesgue measurable functions. We need to show that for every $\theta\in\mathbb{R}$

$$\Delta(\theta)=\{t\in T\colon\ p(t,x(t),\nu(t))\leqslant\theta\}\in L(T),$$

with $L(T)$ being the Lebesgue σ-field of T. Observe that

$$\Delta(\theta)=\bigcap_{n\geqslant 1}\operatorname{proj}_T E_n(\theta),\ \text{where}$$

$$E_n(\theta)=\Big\{(t,u)\in T\times Y\colon\ L(t,x(t),u)\leqslant\theta+\frac{1}{n},\ (t,x(t),\nu(t),u)\in\operatorname{Gr}\Gamma\Big\}.$$

Because of hypothesis $H(L)(1)$, $(t,u)\mapsto L(t,x(t),u)$ is measurable while we have already established above that Γ is graph measurable. Hence $E_n(\theta)\in L(T)\times B(Y)$ and so by the "projection theorem" (see [10] or [14]) we have $\operatorname{proj}_T E_n(\theta)\in L(T)$ and thus $\bigcap_{n\geqslant 1}\operatorname{proj}_T E_n(\theta)=\Delta(\theta)\in L(T)$.

ii) $(x,\nu)\mapsto p(t,x,\nu)$ is sequentially l.s.c. on $X_w\times X^*$ (X_w denotes the space X equipped with the weak topology).

We need to show that for every $\theta\in\mathbb{R}$ the level set

$$K(\theta)=\{[x,\nu]\in X\times X^*\colon\ p(t,x,\nu)\leqslant\theta\}$$

is sequentially closed on $X_w\times X^*$. So let $\{[x_n,\nu_n]\}_{n\geqslant 1}\subset K(\theta)$, $[x_n,\nu_n]\to[x,\nu]$ in $X_w\times X^*$, hence $x_n\to x$ in H. By definition there exist, by passing to a subsequence if necessary, $u_n\in\Gamma(t,x_n,\nu_n)$, $n\geqslant 1$, such that

$$L(t,x_n,u_n)<p(t,x_n,\nu_n)+\frac{1}{n}\leqslant\theta+\frac{1}{n}.$$

Then $u_n\in U(t,x_n)$ and $\nu_n+A(t,x_n)=f(t,x_n,u_n)$. Because of hypotheses $H(U)(3)$ and $H(f)(3)$ and by passing to a subsequence if necessary, we may assume that $h_n=f(t,x_n,u_n)\to h$ weakly in H. By virtue of Mazur's lemma, we can find $\lambda_n^k\geqslant 0$, $\sum\limits_{k=0}^{m_n}\lambda_n^k=1$ such that

$$\hat{h}_n=\sum_{k=0}^{m_n}\lambda_n^k h_{n+k}\to h\ \text{in}\ H.$$

Note that, for every $n \geqslant 1$, we have

$$\hat{h}_n \in \overline{\mathrm{co}} \bigcup_{k=n}^{\infty} f(t, x_k, u_k)$$

whence we get

$$h \in \bigcap_{n \in \mathbb{N}} \overline{\mathrm{co}} \bigcup_{k=n}^{\infty} f(t, x_k, u_k).$$

Also, if $\eta_n = L(t, x_n, u_n)$, we set

$$\hat{\eta}_n = \sum_{k=0}^{m_n} \lambda_n^k \eta_{n+k}$$

and let $\hat{\eta} = \overline{\lim}\, \hat{\eta}_n$, which is finite because of hypothesis $H(L)(3)$ and the choice of u_n.

Then thanks to hypothesis H_c, we get $[h, \hat{\eta}] \in Q(t, x)$ and so there exists $u \in U(t, x)$ such that $h = f(t, x, u)$ and $L(t, x, u) \leqslant \hat{\eta}$.

Note that because of hypothesis $H(A)(3)$ $\{A(t, x_n)\}_{n \geqslant 1}$ is bounded in X^* and so by passing to a subsequence if necessary, we may assume that $A(t, x_n) \to w$ weakly in X^*. Further we have

$$\langle A(t, x_n), x_n - x \rangle = \langle f(t, x_n, u_n) - \nu_n, x_n - x \rangle = (f(t, x_n, u_n), x_n - x) - \langle \nu_n, x_n - x \rangle .$$

Recall that $f(t, x_n, u_n) \to h$ weakly in H, $x_n \to x$ weakly in X, $x_n \to x$ in H and $\nu_n \to \nu$ in X^*. So we have

$$\lim_{n \to \infty} \langle A(t, x_n), x_n - x \rangle = 0.$$

But $A(t, \cdot)$, being monotone and hemicontinuous (see hypothesis $H(A)(2)$), has property (M) (see [15], pp. 383–384). So $w = A(t, x)$. Therefore we have

$$\nu_n + A(t, x_n) \to \nu + A(t, x) \text{ weakly in } X^*$$

and then

$$\nu + A(t, x) = f(t, x, u), \text{ where } u \in U(t, x) \text{ and } L(t, x, u) \leqslant \hat{\eta}.$$

Finally, since

$$\sum_{k=0}^{m_n} \lambda_n^k \, \eta_{n+k} \leqslant \sum_{k=0}^{m_n} \lambda_n^k \Big(\theta + \frac{1}{n+k} \Big) \leqslant \theta + \frac{1}{n},$$

we get $L(t,x,u) \leqslant \theta$. Hence we obtain that $p(t,x,\nu) \leqslant \theta$ and so $(x,\nu) \mapsto p(t,x,\nu)$ is sequentially l.s.c. on $X_w \times X^*$.

(iii) $\nu \mapsto p(t,x,\nu)$ is convex.

Observe that epi $p(t,x,\cdot) = \{[\nu,\eta] \in X^* \times \mathbb{R}\colon\ p(t,x,\nu) \leqslant \eta\} = \bigcap\limits_{\varepsilon>0} \{[\nu,\eta] \in X^* \times \mathbb{R}\colon$ $\exists\ u \in U(t,x)\colon\ L(t,x,u) \leqslant \eta + \varepsilon,\ \nu + A(t,x) = f(t,x,u)\}$ and the last intersection is convex because of hypothesis H_c.

Now let $\hat{p}\colon\ T \times H \times X^* \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ be defined by

$$\hat{p}(t,x,\nu) = \begin{cases} p(t,x,\nu) & \text{if} \quad x \in X; \\ +\infty & \text{if} \quad x \in H \setminus X. \end{cases}$$

Our claim is that $(x,\nu) \mapsto \hat{p}(t,x,\nu)$ is l.s.c. on $H \times X^*$. Indeed, let $[x_n,\nu_n] \to [x,\nu]$ in $H \times X^*$ and assume that $\hat{p}(t,x_n,\nu_n) \leqslant \theta$, $\forall\ n \geqslant 1$, $\theta \in \mathbb{R}$. Then $x_n \in x$, $\forall\ n \geqslant 1$ and $\nu_n + A(t,x_n) = f(t,x_n,u_n)$ with $u_n \in U(t,x_n)$. Because of hypotheses $H(U)(3)$, $H(f)(3)$ and $H(A)(4)$, we have that $\{x_n\}_{n\geqslant 1}$ is bounded on X and so we may assume that $x_n \to x$ weakly in X. Then property (ii) of p together with the equality $\hat{p}(t,x_n,\nu_n) = p(t,x_n,\nu_n)$ implies

$$\underline{\lim}\ \hat{p}(t,x_n,\nu_n) \geqslant \hat{p}(t,x,\nu).$$

It follows that $\hat{p}(t,x,\nu) \leqslant \theta$ and so

$$(x,\nu) \mapsto \hat{p}(t,x,\nu) \text{ is l.s.c. as claimed.}$$

Remark that $\hat{p}(t,x,\cdot)$ is convex (cf. property (iii) of p) and that

$$\psi(t) - \beta|x| \leqslant p(t,x,\nu) \text{ a.e. on } T,\ \forall\ (x,v) \in X \times X^* \text{ (cf. hypothesis}$$

$H(L)(3)$). Suppose that $m < +\infty$ and let $\{[x_n,u_n]\}_{n\geqslant 1} \subset W_{pq}(T) \times L^1(T,Y)$ be a minimizing sequence for the optimal control problem (1), i.e. $J(x_n,u_n) \to m$. From the a priori bounds established in the proof of Theorem 1, we know that $\{x_n\}_{n\geqslant 1}$ is bounded in $W_{pq}(T)$, hence $\{\dot{x}_n\}_{n\geqslant 1}$ is bounded in $L^q(T,X^*)$. By passing to a subsequence if necessary, we may assume that $x_n \to x$ weakly in $W_{pq}(T)$, hence $x_n \to x$ in $L^p(T,X)$ and $\dot{x}_n \to \dot{x}$ weakly in $L^q(T,X^*)$. Applying Theorem 2.1 of [2], we get

$$-\infty < \int_0^b \hat{p}(t,x(t),\dot{x}(t)\,\mathrm{d}t \leqslant \underline{\lim} \int_0^b p(t,x_n(t),\dot{x}_n(t))\,\mathrm{d}t \leqslant \underline{\lim}\ J(x_n,u_n) = m < \infty.$$

From the above inequalities we see that by redefining, if necessary, $\hat{p}(t,x(t),\dot{x}(t))$ on a Lebesgue-null subset on T, we may assume that $\hat{p}(t,x(t),\dot{x}(t)) = p(t,x(t),\dot{x}(t))$ is finite for all $t \in T$. By a straightforward application of Aumann's selection theorem, for every $k \geqslant 1$ we can find a measurable function $u_k\colon T \to Y$ such that $u_k(t) \in \Gamma(t, x(t), \dot{x}(t))$ a.e. and

$$L(t, x(t), u_k(t)) \leqslant p(t, x(t), \dot{x}(t)) + \frac{1}{k} \text{ a.e.}$$

Let $\hat{L}_k\colon T \to \mathbb{R}$ be defined by $\hat{L}_k(t) = L(t, x(t), u_k(t))$. Note that

$$\psi(t) - \beta|x(t)| \leqslant \hat{L}_k(t) \leqslant p(t, x(t), \dot{x}(t)) + \frac{1}{k} \text{ a.e.}$$

and so $\{\hat{L}_k\}_{k\geqslant 1}$ is uniformly integrable in $L^1(T, \mathbb{R})$. Hence from the classical Dunford-Pettis compactness criterion, we may assume that $\hat{L}_k \to \hat{L}$ weakly in $L^1(T, \mathbb{R})$, $\hat{L} \in L^1(T, \mathbb{R})$.

Remark that $[\dot{x}(t) + A(t, x(t)), \hat{L}_k(t)] \in Q(t, x(t))$ a.e. Then as before, using Mazur's lemma on the sequence $\{\hat{L}_k\}_{k\geqslant 1}$ weakly convergent in $L^1(T, \mathbb{R})$, recalling that from a strong convergent sequence in $L^1(T, \mathbb{R})$ we can always extract a subsequence converging almost everywhere and using hypothesis H_c, we conclude that

$$[\dot{x}(t) + A(t, x(t)), \hat{L}(t)] \in Q(t, x(t)) \text{ a.e.}$$

Another application of Aumann's selection theorem yields a measurable function $u\colon T \to Y$ such that $u(t) \in U(t, x(t))$ a.e., $\dot{x}(t) + A(t, x(t)) = f(t, x(t), u(t))$ a.e. and $\hat{L}(t) \geqslant L(t, x(t), u(t))$ a.e.

Then since $\int_0^b p(t, x(t), \dot{x}(t))\,\mathrm{d}t + \frac{b}{k} \geqslant \int_0^b \hat{L}_k(t)\,\mathrm{d}t$, we have

$$\int_0^b \hat{L}_k(t)\,\mathrm{d}t \leqslant \int_0^b p(t, x(t), \dot{x}(t))\,\mathrm{d}t \leqslant m$$

and so we get

$$J(x, u) \leqslant m.$$

However $[x, u] \in P(x_0)$. Hence $J(x, u) = m$ and consequently $[x, u]$ is optimal. $\square$

4. Examples

We present two applications of nonlinear parabolic optimal control problems

Let $T = [0, b]$ and let Z be a bounded domain in $\mathbb{R}^N$ with smooth boundary $\Gamma = \partial Z$. Let $2 \leqslant p < \infty$. We consider the distributed parameters optimal control problem

$$
(6)\qquad \begin{cases} \displaystyle\int_0^b \int_Z L(t, z, x(t, z), u(t, z))\,\mathrm{d}z\,\mathrm{d}t \to \inf = m; \\ \text{s.t. } \displaystyle\frac{\partial x}{\partial t} - \sum_{k=1}^{N} D_k(\alpha(t, z)|D_k x|^{p-2} D_k x) = f(t, z, x(t, z), u(t, z)), \\ \qquad \text{a.e. on } T \times Z; \\ x|_{T\times\Gamma} = 0,\ x(0, z) = x_0(z),\ \|u(t, z)\| \leqslant \gamma(t, z\|x(t, \cdot)\|_2 \text{ a.e. on } T \times Z, \end{cases}
$$

where $D_k = \frac{\partial}{\partial z_k}$ and $|\cdot|$ denotes the absolute value.

We will need the following hypotheses on the data:

$H(\alpha)$: $\alpha\colon T \times Z \to \mathbb{R}$ is a measurable function such that there exist $\beta_1, \beta_2 \in \mathbb{R}$:

$$0 < \beta_1 \leqslant \alpha(t, z) \leqslant \beta_2 \text{ a.e. on } T \times Z.$$

$H(f)_2$: $f\colon T \times Z \times \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}$ is a function such that

(1) $\forall\ (x, u) \in \mathbb{R} \times \mathbb{R}^k$, $(t, z) \mapsto f(t, z, x, u)$ is measurable;

(2) $\exists\ k \in L^1(T, L^\infty(Z, \mathbb{R}^+))$ such that
$|f(t, z, x, u) - f(t, z, x', u')| \leqslant k(t, z)(|x - x'| + \|u - u'\|)$
a.e. in $T \times Z$ and $\forall\ (x, u)$, $(x', u') \in \mathbb{R} \times \mathbb{R}^k$;

(3) $\exists\ a \in L^q(T, Z^2(Z, \mathbb{R}^+))$ and $c \in L^q(T, L^\infty(Z, \mathbb{R}^\infty))$:
$|f(t, z, x, u)| \leqslant a(t, z) + c(t, z)(|x| + \|u\|)$, a.e. in $T \times Z$, $\forall\ (x, u) \in \mathbb{R} \times \mathbb{R}^k$.

$H(\gamma)$: $\gamma\colon T \times Z \times \mathbb{R}^+ \to \mathbb{R}^+$ is a function such that

(1) $\forall\ r \in \mathbb{R}^+$, $(t, z) \mapsto \gamma(t, z, r)$ is measurable;

(2) $\forall\ (t, z) \in T \times Z$, $r \mapsto \gamma(t, z, r)$ is continuous;

(3) there exists $\theta \in L^\infty(T \times Z, \mathbb{R}^+)$ such that

$$\gamma(t, z, r) \leqslant \theta(t, z)(1 + r) \text{ a.e. on } T \times Z.$$

$H(L)_1$: $L\colon T \times Z \times \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}$ is an integrand such that

(1) $\forall\ (x, u) \in \mathbb{R} \times \mathbb{R}^k$, $(t, z) \mapsto L(t, z, x, u)$ is measurable;

(2) $\forall\, M > 0\ \exists\, w_M\colon\ T \times Z \times \mathbb{R}^+ \times \mathbb{R}_0^+$ such that

$(2)_1$ $\forall\, (r, r') \in \mathbb{R}^+ \times \mathbb{R}$, $(t, z) \mapsto w_M(t, z, r, r')$ is measurable;
$(2)_2$ $\forall\, (t, z) \in T \times Z$, $(r, r') \mapsto w_M(t, z, r, r')$ is continuous at $(0, 0)$ and $w_M(t, z, 0, 0) = 0$;
$(2)_3$ $\exists\, \varphi_M\colon\ T \times Z \to \mathbb{R}$: $\varphi_M(t, \cdot) \in L^1(Z, \mathbb{R})$ and $w_M(t, z, r, r') \leqslant \varphi_M(t, z)\ \forall\, t \in T$ and a.e. on Z;
$(2)_4$ $|L(t, z, x, u) - L(t, z, x', u')| \leqslant w_M(t, z, |x - x'|, \|u - u'\|)$ a.e. in $T \times Z$ and $\forall\, (x, u), (x', u') \in \mathbb{R} \times \mathbb{R}^k$ such that $\|u\| \leqslant M$ and $\|u'\| \leqslant M$;
(3) $\exists\, \psi \in L^1(T \times Z, \mathbb{R})$ and $\beta \geqslant 0$:
$\psi(t) - \beta(|x| + \|u\|) \leqslant L(t, z, x, u)$ a.e. on $T \times Z$, $\forall\, (x, u) \in \mathbb{R} \times \mathbb{R}^k$;
(4) $\forall\, (t, z, x) \in T \times Z \times \mathbb{R}$, $u \mapsto L(t, z, x, u)$ is convex.

H_c'': $\forall\, x \in L^2(Z, \mathbb{R})$, the function $u \mapsto f(t, z, x(z), u)$ is such that $\forall\, u,\ u' \in \mathbb{R}^k$ with $\|u\|, \|u'\| \leqslant \gamma(t, z, \|x\|_2)$ and $\forall\, \lambda \in [0, 1]$ we have $f(t, z, x(z), \lambda u + (1 - \lambda) u') = \lambda f(t, z, x(z), u) + (1 - \lambda) f(t, z, x(z), u')$;

H_0: $x_0 \in L^2(Z, \mathbb{R})$.

Then we have the following existence result concerning problem (6)

Theorem 4. *If hypotheses $H(\alpha)$, $H(\gamma)$, $H(f)_2$, $H(L)_1$, H_c'' and H_0 hold, then problem* (6) *admits an optimal pair $[x, u] \in C(T, L^2(Z, \mathbb{R})) \times L^2(T \times Z, \mathbb{R}^k)$ such that $\frac{\partial x}{\partial t} \in L^q(T, W^{-1,q}(Z, \mathbb{R}))$.*

Proof. In this case $X = W_0^{1,p}(Z, \mathbb{R})$, $H = L^2(Z, \mathbb{R})$ and $X^* = W^{-1,q}(Z, \mathbb{R})$. From the Sobolev embedding theorem we know that (X, H, X^*) is an evolution triple with all embeddings being compact. Let $\theta\colon\ T \times W_0^{1,p}(Z, \mathbb{R}) \times W_0^{1,p}(Z, \mathbb{R}) \to \mathbb{R}$ be the time varying Dirichlet form defined by

$$\theta(t, x, y) = \int_Z \sum_{k=1}^N \alpha(t, z) |D_k x|^{p-2} D_k x D_k y \,\mathrm{d}z.$$

Using Hölder's inequality and hypothesis $H(\alpha)$, we get

$$|\theta(t, x, y)| \leqslant \beta_2 \sum_{k=1}^N \left(\int_Z |D_k x|^p \,\mathrm{d}z \right)^{1/q} \cdot \left(\int_Z |D_k y|^p \,\mathrm{d}z \right)^{1/p}.$$

Recall that $\sum\limits_{k=1}^N \|D_k x\|_p$ is an equivalent norm on $W_0^{1,p}(Z, \mathbb{R})$; therefore $\exists\, \hat{c}_1 > 0$: $|\theta(t, x, y)| \leqslant \hat{c}_1 \|x\|^{p-1} \|y\|$.

So we can define an operator $A\colon T \times W_0^{1,p}(Z,\mathbb{R}) \to W^{-1,q}(Z,\mathbb{R})$ by

$$\langle A(t,x), y\rangle = \theta(t,x,y), \ \forall\, y \in W_0^{1,p}(Z,\mathbb{R}).$$

Observe that we have just proved that

$$\|A(t,x)\|_* \leqslant \hat{c}_1 \|x\|^{p-1}.$$

Using the elementary inequality

$$2^{2-p}|\mu-\nu|^p \leqslant (\mu|\mu|^{p-2} - \nu|\nu|^{p-2})(\mu-\nu), \ \mu,\nu \in \mathbb{R}, \ \forall\, p \geqslant 2,$$

we get that there exists $\hat{c}_2 > 0$ such that

$$\hat{c}_2\|x-y\|^p \leqslant \theta(t,x,x-y) - \theta(t,y,x-y) = \langle A(t,x) - A(t,y), x-y\rangle \qquad \forall\, x,y \in W_0^{1,p}(Z,\mathbb{R}).$$

It follows that $x \mapsto A(t,x)$ is strongly monotone.

Note that by Fubini's theorem $t \mapsto \langle A(t,x), y\rangle$ is measurable for every $y \in W_0^{1,p}(Z,\mathbb{R})$ and hence $t \mapsto A(t,x)$ is weakly measurable. Since $W^{-1,q}(Z,\mathbb{R})$ is a separable, reflexive Banach space, using the Pettis measurability theorem we deduce that $t \mapsto A(t,x)$ is measurable. Also it is easily verified that $x \mapsto A(t,x)$ is continuous from $W_0^{1,p}(Z,\mathbb{R})$ into $W^{-1,q}(Z,\mathbb{R})$ and that $\langle A(t,x), x\rangle \geqslant \bar{c}\|x\|^p$ for some $\bar{c} > 0$. Thus we have satisfied hypothesis $H(A)$. Next, let $Y = L^2(Z,\mathbb{R}^k)$ and define $\hat{f}\colon T \times H \times Y \to H$ by

$$\hat{f}(t,x,u)(\cdot) = f(t,\cdot,x(\cdot),u(\cdot)),$$

i.e. $\hat{f}$ is the Nemitsky (superposition) operator corresponding to the function f. Using $H(f)_2(3)$ we obtain, by virtue of Young's inequality, that there exist $\hat{a}, \hat{c} \in L^q(T,\mathbb{R}^+)$ such that

$$|\hat{f}(t,x,u)| \leqslant \hat{a}(t) + \hat{c}(t)[|x| + \|u\|_Y] \leqslant \hat{a}_1(t) + c(t)[\|x\|^{2/q} + \|u\|_Y] \text{ a.e. on } T,$$

where $\hat{a}_1 = \hat{a} + \hat{c} \in L^q(T,\mathbb{R}^+)$. ☐

Further, from Krasnoselskii's theorem we know that $(x,u) \mapsto \hat{f}(t,x,u)$ is continuous and by $H(f)_2$, using Fubini's theorem we get that $t \mapsto \hat{f}(t,x,u)$ is measurable. So we have satisfied hypothesis $H(f)$.

Next let $U\colon T \times H \to P_{wkc}(Y)$ be defined by

$$U(t,x) = \{u \in Y\colon \|u(z)\| \leqslant \gamma(t,z,|x|) \text{ a.e. on } Z\}.$$

Because of hypothesis $H(\gamma)$, we readily see that $t \mapsto U(t,x)$ is measurable, while $x \mapsto U(t,x)$ is H-continuous; therefore $(t,x) \mapsto U(t,x)$ is jointly measurable (cf. Theorem 3.3 of [12]). Futhermore, hypothesis $H(\gamma)(3)$ gives that $\exists\, \hat{\theta} > 0$:

$$|U(t,x)| \leqslant \hat{\theta}(1+|x|) \leqslant \hat{\theta}_1(1+|x|^{2/q})$$

with $\hat{\theta}_1 = 2\hat{\theta}$. So we have satisfied hypothesis $H(U)$.

Next, let $\hat{L}\colon T \times H \times Y \to \overline{\mathbb{R}}$ be defined by

$$\hat{L}(t,x,u) = \int_Z L(t,z,x(z),u(z))\,\mathrm{d}z.$$

Then from Fubini's theorem, $t \mapsto \hat{L}(t,x,u)$ is measurable, while from hypothesis $H(L)_1(2)$ we get that $(x,u) \mapsto \hat{L}(t,x,u)$ is continuous. Hence $(t,x,u) \mapsto \hat{L}(t,x,u)$ is measurable. Also, from hypothesis $H(L)_1(3)$, we have that there exist $\hat{\psi} \in L^1(T,\mathbb{R})$ and $\hat{\beta} > 0$ such that $\hat{\psi}(t) - \hat{\beta}|x| \leqslant \hat{L}(t,x,u)$ a.e. on T, $\forall\, x \in H$, $\forall\, u \in U(t,x)$. Therefore also hypothesis $H(L)$ is satisfied. Finally, note that by virtue of hypotheses $H(f)_2(2), (L)_1(2)-(4)$ and H_c'' and Lemma 1, we see that the convexity hypothesis H_c is satisfied (observe that $u \mapsto L(t,x,u)$ is convex) and by H_c'' we have that $f(t,x,U(t,x))$ is closed $\forall\, (t,x) \in T \times H$.

Now we can rewrite (6) as the equivalent abstract optimal control problem

$$(6') \qquad \begin{aligned} &\hat{J}(x,u) = \int_0^b \hat{L}(t,x(t),u(t))\,\mathrm{d}t \to \inf = m;\\ &\text{s.t. } \dot{x}(t) + A(t,x(t)) = \hat{f}(t,x(t)u(t)) \text{ a.e.};\\ &x(0) = x_0 \in H = L^2(Z,\mathbb{R});\\ &u(t) \in U(t,x(t)) \text{ a.e., } u(\cdot) - \text{measurable.} \end{aligned}$$

Apply Theorem 3 to get an optimal pair $[x,u] \in C(T,L^2(Z,\mathbb{R})) \times L^2(T\times Z,\mathbb{R})$. For the optimal state $x(t,z)$ we know that $\frac{\partial x}{\partial t} \in L^q(T,W^{-1,q}(Z,\mathbb{R}))$.

Again let $T=[0,b]$ and let Z be a bounded domain in $\mathbb{R}^N$ with smooth boundary $\Gamma = \partial Z$. We consider the optimal control problem

$$(7) \qquad \begin{aligned} &\frac{1}{2}\int_0^b\int_Z |x(t,z)-\hat{x}(t,z)|^2\,\mathrm{d}z\,\mathrm{d}t + \frac{1}{2}\int_0^b\int_Z \|u(t,z)\|^2\,\mathrm{d}z\,\mathrm{d}t \to \inf = m;\\ &\frac{\partial x}{\partial t} - \sum_{k=1}^N D_k(\alpha(\|Dx\|_N^2)D_kx) = f(t,z,x(t,z)) + b(t,z)u(t,z),\\ &\text{a.e. on } T\times Z;\\ &x|_{T\times\Gamma} = 0,\ x(0,z) = x_0(z) \text{ a.e. on } Z;\\ &u(t,z) \in U(t,z,\|x(t,\cdot)\|_2) \text{ a.e. on } T\times Z. \end{aligned}$$

Here $D = \text{grad}$ and $\|Dx\|_N^2 = \sum\limits_{k=1}^{N} |D_k x|^2$.

Problems of this type arise in nonlinear elasticity.

We will need the following hypotheses on the data of (7):

$H(\alpha)$: $\alpha\colon \mathbb{R}^+ \to \mathbb{R}^+$ is continuous and there exist $h > 0$ and $d > 0$ such that $0 \leqslant \alpha(r) \leqslant h$ for all $r \geqslant 0$ and for all $\hat{t} = (t_k)_{k=1}^N$, $\hat{s} = (s_k)_{k=1}^N$ we have

$$\sum_{k=1}^{N} (t_k - s_k)(\alpha(\|t\|^2)t_k - \alpha(\|s\|^2)s_k) \geqslant d\|t-s\|^2.$$

$H(f)_3$: $f\colon T \times Z \times \mathbb{R} \to \mathbb{R}$ is a function such that

(1) $\forall\, x \in R$, $(t,z) \mapsto f(t,z,x)$ is measurable;

(2) $\exists\, k \in L^1(T, L^\infty(Z, \mathbb{R}^+))$:
$|f(t,z,x) - f(t,z,x')| \leqslant k(t,z)|x - x'|$, $\forall\, x, x' \in \mathbb{R}$, a.e. on $T \times Z$;

(3) $\exists\, a \in L^2(T \times Z, \mathbb{R}^+)$, $c \in L^2(T, L^\infty(Z, \mathbb{R}^+))$ such that
$|f(t,z,x)| \leqslant a(t,z) + c(t,z)|x|$, a.e. on $T \times Z$, $\forall\, x \in \mathbb{R}$.

$H(b)$: $b \in L^\infty(T \times Z, \mathbb{R}^k)$.

$H(U)_2$: $U\colon T \times Z \times \mathbb{R}^+ \to P_{kc}(\mathbb{R}^k)$ is a multifunction such that

(1) $\forall\, x \in \mathbb{R}$, $(t,z) \mapsto U(t,z,r)$ is measurable;

(2) $\forall\, (t,z) \in T \times Z$, $r \mapsto U(t,z,r)$ is continuous;

(3) $\exists\, \theta \in L^\infty(T \times Z, \mathbb{R}^+)$ such that
$|U(t,z,r)| \leqslant \theta(t,z)(1+r)$ a.e. on $T \times Z$, $\forall\, x \in \mathbb{R}^+$.

H_0: $x_0 \in L^2(Z, \mathbb{R})$;
$\hat{H}$: $\hat{x} \in L^2(T \times Z, \mathbb{R})$ (this is the target function).

We have the following existence result concerning (7):

Theorem 5. *If hypotheses $H(\alpha)$, $H(f)_3$, $H(b)$, $H(U)_2$, H_0 and $\hat{H}$ hold, then problem (7) admits an optimal pair $[x,u] \in C(T, L^2(Z,\mathbb{R})) \times L^2(T \times Z, \mathbb{R}^k)$ such that $\frac{\partial x}{\partial t} \in L^2(T, H^{-1}(Z, \mathbb{R}))$.*

Proof. In this case the evolution triple consists of $X = H_0^1(Z,\mathbb{R})$, $H = L^2(Z,\mathbb{R})$ and $X^* = H^{-1}(Z,\mathbb{R})$. Then let $\theta\colon H_0^1(Z,\mathbb{R}) \times H_0^1(Z,\mathbb{R}) \to \mathbb{R}$ be the Dirichlet form defined by

$$\theta(x,y) = \int_Z \alpha(\|Dx\|_N^2)(Dx, Dy)\,\mathrm{d}z,$$

where $(\cdot,\cdot)$ denotes the scalar product on $\mathbb{R}^N$.

Using hypothesis $H(\alpha)$, we get that there exists $\hat{h} > 0$ such that

$$|\theta(x,y)| \leqslant \hat{h}\|x\|\ \|y\|.$$

Consequently, we can define $A\colon H_0^1(Z,\mathbb{R}) \to H^{-1}(Z,\mathbb{R})$ by

$$\langle A(x), y\rangle = \theta(x,y),\ \forall\ y \in H_0^1(Z,\mathbb{R})$$

and from above we know that $\|A(x)\|* \leqslant \hat{h}\|x\|$. Furthermore, as in the previous example, it is easy to see that $A(\cdot)$ verifies hypothesis $H(A)$, with $p = 2$. Let $Y = L^2(Z,\mathbb{R}^k)$ and define $\hat{f}\colon T \times H \times Y \to H$ by

$$\hat{f}(t,x,u)(z) = f(t,z,x(z)) + b(t,z)u(z),\ \forall\ z \in Z.$$

Clearly, hypothesis $H(f)$ is satisfied.

Now let $\hat{U}(t,x) = \{u \in Y\colon\ u(z) \in U(t,z,|x|) \text{ a.e. on } Z\}$. Using hypothesis $H(U)_2$ and Theorem 3.1 and 4.4 of [13], we see that $\hat{U}(t,\cdot)$ is H-continuous. Furthermore, for every $\nu \in Y$ we have

$$d(\nu, \hat{U}(t,x))^2 = \int_Z d(\nu(z), U(t,z,|x|))^2\,\mathrm{d}z.$$

Therefore $(t,x,\nu) \mapsto d(\nu,\hat{U}(t,x))$ is a Carathéodory function, hence it is jointly measurable and then $(t,x) \mapsto \hat{U}(t,x)$ is measurable. Note that $\exists\ \hat{\theta} > 0$ such that $|\hat{U}(t,x)| \leqslant \hat{\theta}(1+|x|)$ (cf. hypothesis $H(U)_2(3)$.

Let $L\colon T \times H \times Y \to \mathbb{R}$ be defined by

$$L(t,x,u) = \frac{1}{2}\int_Z |x(z) - \hat{x}(t,z)|^2\,\mathrm{d}z + \frac{1}{2}\int_Z \|u(z)\|^2\,\mathrm{d}z.$$

It is easy to see that hypothesis $H(L)$ is satisfied; moreover, by using Lemma 1, also hypothesis H_c is verified.

Rewriting problem (7) in the equivalent form

$$(7') \qquad \begin{aligned} & J(x,u) = \int_0^b L(t,x(t),u(t))\,\mathrm{d}t \to \inf = m;\\ & \text{s.t. } \dot{x}(t) + A(x(t)) = \hat{f}(t,x(t),u(t)) \text{ a.e.};\\ & x(0) = x_0 \in H = L^2(Z,\mathbb{R});\\ & u(t) \in \hat{U}(t,x(t)) \text{ a.e., } u(\cdot)\text{-measurable,} \end{aligned}$$

and taking into account that $\hat{f}(t,x,\hat{U}(t,x))$ is closed, $\forall\ (t,x) \in T \times H$, we can apply Theorem 3 to get the existence of an optimal pair with the requested properties. □

References

[1] *N. U. Ahmed – K. L. Teo*: Optimal Control of Distributed Parameter Systems. North Holland, New York, 1981.

[2] *E. Balder*: Necessary and sufficient conditions for L_1-strong-weak lower semicontinuity of integral functionals. Nonlinear Analysis *11* (1987), 1399–1404.

[3] *V. Barbu*: Nonlinear Semigroups and Differential Equations in Banach Spaces. Noordhoff, Leyden, 1976.

[4] *L. D. Berkovitz*: Optimal Control Theory. Springer-Verlag, New York, 1974.

[5] *L. Cesari*: Existence theorems for weak and usual optimal solutions in Lagrange problems with unilateral constraints - I. Trans. Amer. Math. Society *124* (1966), 369–412.

[6] *L. Cesari*: Optimization-Theory and Applications. Springer-Verlag, New York, 1983.

[7] *G. A. Edgar*: Measurability in a Banach space - II. Indiana J. Math. *28* (1979), 559–579.

[8] *A. Fryszkowski*: The generalization of Cellina's Fixed Point Theorem. Studia Math. *78* (1983), 213–215.

[9] *S. H. Hou*: On Property (Q) and other semicontinuity properties of multifunctions. Pacific Journal Math. *103* (1982), 39–56.

[10] *V. Levin*: Borel sections of many valued maps. Siberian Math. J. *19* (1979), 434–438.

[11] *J. L. Lions*: Optimal Control of Systems Governed by Partial Differential Equations. Springer, Berlin, 1971.

[12] *N. S. Papageorgiou*: On measurable multifunctions with applications to random multivalued equations. Math. Japonica *32* (1987), 437–464.

[13] *N. S. Papageorgiou*: Convergence theorems for Banach space-valued integrable multifunctions. Intern. J. Math. and Math. Sci. *10* (1987), 433–442.

[14] *D. Wagner*: Survey on measurable selection theorems. SIAM J. Control Optim. *15* (1977), 859–903.

[15] *E. Zeidler*: Nonlinear Functional Analysis and Its Application II. Springer, New York, 1990.

Authors' addresses: Antonella Fiacca, Department of Mathematics of Perugia University, Via Vanvitelli 1, Perugia, 06123, Italy; Nikolaos S. Papageorgiou, National Technical University, Department of Mathematics, Zografou Campus, Athens 15780, Greece; Francesca Papalini, Department of Mathematics of Perugia University, Via Vanvitelli 1, Perugia, 06123, Italy.