Avgerinos, E.P. and Papageorgiou, N.S.
Osaka J. Math.
27 (1990), 745–767

OPTIMAL CONTROL AND RELAXATION FOR A CLASS OF NONLINEAR DISTRIBUTED PARAMETER SYSTEMS

EVGENIOS P. AVGERINOS and NIKOLAOS S. PAPAGEORGIOU*

(Received May 8, 1989)
(Revised May 1, 1990)

1. Introduction

The study of optimal control of distributed parameter systems started in the early sixties and since then there have been numerous important developments on the subject. Most of the works concentrate on linear systems. Selectively we mention the important works of Egorov [24], Friedman [26], [27] and the classical book of Lions [34]. Nonlinear systems were considered by Cesari [14], [15], [16], Lions [33], Ahmed [2] and Hou [30] among others. A comprehensive presentation of the nonlinear theory, together with the more recent trends in the study of distributed parameter systems can be found in the book of Ahmed-Teo [6]. Very recently, in a series of interesting papers, Ahmed [3], [4], [5], studied the existence of optimal controls for large classes of semilinear systems, as well as their relaxation properties.

In this paper we continue on the road paved by those three recent works of Ahmed. We study the Lagrange optimal control problem for a large class of nonlinear systems governed by Volterra integrodifferential evolution equations. Our hypotheses are general enough to incorporate both parabolic and hyperbolic distributed parameter control systems. First we establish the existence of optimal controls by considering systems in which the control appears linearly in the dynamics. Since the dynamic equation of our system is not "instantaneous", but also has memory incorporated in the integral part of the integrodifferential equation, we encounter serious technical difficulties if we try to handle systems with the control entering nonlinearly. The otherwise powerful and elegant "Cesari-Rockafellar reduction technique" does not seem to work here. Then we consider completely nonlinear systems. Now in order to guarantee the existence of optimal controls, we have to pass to a larger system with measure valued controls, known as the "relaxed system". For this augmented system, we prove that optimal controls exist and in addition, under

* Research supported by NSF Grant DMS-8802688

mild hypotheses the value of the relaxed problem equals that of the original one. Again the memory feature of our system does not allow us to have an alternative "deparametrized" version of the relaxed problem. Finally we prove a density (relaxation) result relating the trajectories of the original and relaxed systems. This result illustrates that the relaxed problem is the "closure" of the original one. Two more results of this nature are also proved. Finally some examples are worked out in detail.

2. Preliminaries

Let (Ω, Σ) be a measurable space and X a separable Banach space. Throughout this paper we will be using the following notations:

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

and

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

A multifunction $F\colon \Omega \to P_f(X)$ is said to be measurable, if for every $z \in X$, $\omega \to d(z, F(\omega)) = \inf\{\|z-x\|\colon x \in F(\omega)\}$ is measurable. Also a multifunction $G\colon \Omega \to 2^X \setminus \{\emptyset\}$ is said to be "graph measurable", if $\operatorname{Gr} G = \{(\omega, x) \in \Omega \times X\colon x \in G(\omega)\} \in \Sigma \times B(X)$, with $B(X)$ being the Borel σ-field of X. For closed valued multifunctions, measurability implies graph measurability. The converse is true if there exists a complete σ-finite measure $\mu(\cdot)$ on (Ω, Σ). For details we refer to Wagner [42], Hiai-Umegaki [28] and Levin [32].

Now let (Ω, Σ, μ) be a finite measure space and $F\colon \Omega \to 2^X \setminus \{\emptyset\}$ a multifunction. By $S_F^p (1 \leq p \leq \infty)$ we will denote the set of all $L^p(X)$-selectors of $F(\cdot)$; i.e. $S_F^p = \{f \in L^p(X)\colon f(\omega) \in F(\omega)\ \mu\text{-a.e.}\}$. This set may be empty. It is easy to see that it is nonempty if and only if $F(\cdot)$ is graph measurable and $\omega \to \inf\{\|x\|\colon x \in F(\omega)\} \in L_+^p$. So if $F(\cdot)$ is graph measurable and $\omega \to \sup\{\|x\|\colon x \in F(\omega)\}$ belongs in L_+^1 (in which case the multifunction is said to be "integrably bounded"), then $S_F^1 \neq \phi$. Using S_F^1 we can define a set valued integral for $F(\cdot)$, by setting $\int_\Omega F(\omega)\, d\mu(\omega) = \{\int_\Omega f(\omega)\, d\mu(\omega)\colon f \in S_F^1\}$, where the vector integrals of the right hand side are defined in the sense of Bochner.

Recall (see for example Barbu [10]), that the duality map of a Banach space X, is the mapping $\hat{F}\colon X \to 2^{X^*}$ defined by $\hat{F}(x) = \{x^* \in X^*\colon (x^*, x) = \|x\|^2 = \|x^*\|^2\}$. It is clear that for every $x \in X$, $\hat{F}(x)$ is closed, convex, bounded and furthermore the Hahn-Banach theorem, tells us that $\hat{F}(x) \neq \emptyset$. Note that when X^* is strictly convex, then the duality map is single valued. Using $\hat{F}(\cdot)$ we can define a semi-inner product $(\cdot, \cdot)_-$ on X, by setting $(x, y)_- = \inf\{(x^*, y)\colon x^* \in \hat{F}(x)\}$. An operator $A\colon D(A) \subseteq X \to 2^X$ is said to be "dissipative" if and only if $(x-x', y-y')_- \leq 0$ for all $(x, y), (x', y') \in \operatorname{Gr} A$. The operator $A(\cdot)$ is said to be "m-dissipative" if and only if it is dissipative and $R(I-\lambda A) = X$ for each $\lambda \geq 0$. It is well known that an m-dissipative operator $A(\cdot)$ generates a semigroup of nonlinear, nonexapansive maps $\{S(t): \overline{D(A)} \to \overline{D(A)}, t \geq 0\}$, via the Crandall-Liggett exponential formula: $S(t)x = \lim_{x \to \infty} (I - \frac{t}{n} A)^{-n} x$ for each $x \in \overline{D(A)}$. Finally recall that by J_λ, $\lambda > 0$ we denote the resolvent of A i.e. $J_\lambda = (I - \lambda A)^{-1}$. We know that this is a nonexapnsive map defined on all X and $\lim_{\lambda \to 0} J_\lambda x = x$ for all $x \in \overline{D(A)}$.

Let Z be a separable, complete metric space (a Polish space) and let $B(Z)$ be its Borel σ-field. By $M_+^1(Z)$ we will denote the space of probability measures on Z. A transition probability is a function $\lambda: \Omega \times B(Z) \to [0, 1]$ s.t. for all $A \in B(Z)$, $\lambda(\cdot, A)$ is Σ-measurable and for every $\omega \in \Omega$, $\lambda(\omega, \cdot) \in M_+^1(Z)$. If Z is compact, this definition is equivalent to saying that $\omega \to \lambda(\omega, \cdot)$ from Ω into $M_+^1(Z)$ with the narrow (weak) topology, is measurable. We will denote the set of all transition probabilities from (Ω, Σ, μ) into $(Z, B(Z))$, by $R(\Omega, Z)$. Following Balder [9] (see also Warga [43]), we can define a topology on $R(\Omega, Z)$ as follows: Let $f: \Omega \times Z \to \boldsymbol{R}$ be a Caratheodory, L^1-function (i.e. $\omega \to f(\omega, x)$ is measurable, $x \to f(\omega, x)$ is continuous and $|f(\omega, x)| \leq a(\omega)$ μ-a.e. with $a(\cdot) \in L_+^1$) and let $I_f: R(\Omega, Z) \to \boldsymbol{R}$ be the integral functional defined by $I_f(\lambda) = \int_\Omega \int_Z f(\omega, z) \lambda(\omega)(dz)\, d\mu(\omega)$. The weakest topology on $R(\Omega, Z)$ that makes the above functionals (for any Caratheodory L^1-integrand) continuous, is called the weak topology on $R(\Omega, Z)$. Observe that when Ω is a singleton, then $R(\Omega, Z) = M_+^1(Z)$ and the weak topology just defined, is nothing else but the well known narrow topology ("topologie etroite" in the Bourbaki terminology [12], see also Choquet [18]) on $M_+^1(Z)$. Suppose Z is a compact metric space. Then the Caratheodory, L^1-integrands can be identified with the Lebesgue-Bochner space $L^1(C(Z))$. To see this, associate to each Caragheodory L^1-integrand $f(\cdot, \cdot)$, the map $\omega \to f(\omega, \cdot) \in C(Z)$. The measurability of this map follows from the lemma in [36]. From the Dinculeanu-Foias theorem, we know that $L^1(C(Z))^* = L^\infty(M(Z))$ where $M(Z)$ is equipped with the narrow topology (Warga [43] calls this result "Dunford-Petts theorem; see [43], p. 268, see also Dunford-Schwartz [22], p. 503). Then the weak topology on $R(\Omega, Z)$ defined above conicides with the relative w*$(L^\infty(M(Z)), L^1(C(Z)))$-topology.

3. Existence of optimal controls

Let $T = [0, b]$, X a separable with uniformly convex dual, hence reflexive (X models the state space) and Y a separable reflexive Banach space (it models the control space).

The nonlinear optimal control problem under consideration is the following:

$$\left\{\begin{array}{l} J(x,u)=\int_0^b L(t,x(t),u(t))\,dt\to\inf=m \\ \text{s.t. } \dot{x}(t)\in Ax(t)+\int_0^t K(t-s)f(s,x(s))u(s)\,ds \quad \text{a.e.} \\ \quad x(0)=x_0,\ u(t)\in U(t,x(t)) \text{ a.e., } u(\cdot) \text{ measurable} \end{array}\right\} (*)$$

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

H(A): $A: D(A)\subseteq X\to 2^X$ is an m-dissipative operator s.t. $(I-A)^{-1}$ is comprct.

H(f): $f: T\times X\to \mathcal{L}(Y,X)$

(1) $t\to f(t,x)$ is measurable,

(2) $x\to f(t,x)^*$ is continuous from X into $\mathcal{L}(X^*,Y^*)$ endowed with the strong operator topology,

(3) $\|f(t,x)\|_{\mathcal{L}(X,Y)}\leq a(t)+b(t)\|x\|$ a.e. with $a(\cdot)\in L^2_+$, $b(\cdot)\in L^\infty_+$.

H(K): $K: T\to\mathcal{L}(X)$ is a C^1-function, where $\mathcal{L}(X)$ is endowed with the operator norm topology.

H(U): $U: T\times X\to P_{fc}(Y)$ is a multifunction s.t.

(1) $(t,x)\to U(t,x)$ is graph measurable,

(2) Gr $U(t,\cdot)=\{(x,u)\in X\times Y: u\in U(t,x)\}$ is closed in $X\times Y_w$,

(3) $U(t,x)\subseteq W$ a.e. with $W\in P_{wkc}(Y)$.

H(L): $L: T\times X\times Y\to\overline{\boldsymbol{R}}=\boldsymbol{R}\cup\{+\infty\}$ is an integrand s.t.

(1) $(t,x,u)\to L(t,x,u)$ is measurable,

(2) $(x,u)\to L(t,x,u)$ is l.s.c. from $X\times Y_w$ into $\overline{\boldsymbol{R}}$ and convex in u,

(3) $\psi_1(t)+\psi_2(t)\|x\|+\psi_3(t)\|u\|\leq L(t,x,u)$ a.e. with $\psi_1(\cdot)$, $\psi_2(\cdot)$, $\psi_3(\cdot)\in L^1$.

In order for our problem to have content, we will need the following feasibility hypothesis:

H_α: There exists admissible "state-control" pair (x,u) s.t. $J(x,u)<\infty$.

Now we are ready for our existence result concerning optimal sontrol problem (*). By $P(x_0)$ we will denote the set of admissible "state-control" pairs of (*).

Theorem 3.1. *If hypotheses* H(A), H(f), H(K), H(U), H(L) *and* H_α *hold, then there exists* $(x,u)\in P(x_0)$ *s.t.* $J(x,u)=m$.

Proof. First we are going to determine an a priori bound for the trajectories of our system. So let $x(\cdot)\in C(T,X)$ be such a trajectory. From Benilan [11], we know that

$$\|x(t)-S(t)x_0\|\leq\int_0^t\|g(s)\|ds$$

where $\{S(t)(\cdot)\}_{t\in T}$ is the semigroup of nonlinear contractions generated by

$A(\cdot)$ and $g(s)=\int_0^s K(s-r)\,f(r, x(r))\,u(r)\,dr$. So we have:

$$\begin{aligned}||x(t)|| &\leq ||S(t)\,x_0||+\int_0^t\int_0^s ||K(s-r)||\,||f(r, x(r))\,u(r)||\,dr\,ds\\ &\leq ||S(t)\,x_0||+\int_0^t\int_0^s N(a(r)+b(r)||x(r)||)\,|W|\,dr\,ds\end{aligned}$$

where $|W|=\sup\{||u||: u\in W\}$ and $||K(t)||_{\mathcal{L}(X)}\leq N$ for all $t\in T$. Such an $N>0$ exists, since by hypothesis H(K), $K(\cdot)$ is C^1. Note that since $t\to S(t)\,x_0$ is continuous, there exists $M_1>0$ s.t. $||S(t)\,x_0||\leq M_1$. Therefore we can write that

$$||x(t)||\leq M_1+N|W|b\,||a||_1+N|W|\int_0^t\int_0^s b(r)\,||x(r)||\,dr\,ds\,.$$

Invoking Theorem 1 of Pachpatte [35], we know that there exists $M>0$ s.t. for all $x(\cdot)\in P_1(x_0)=\{$set of trajectories of (*)$\}$ and all $t\in T$, $||x(t)||\leq M$.

Next let $\{(x_n, u_n)\}_{n\geq 1}\subseteq P(x_0)$ be a minimizing sequence for our problem. We will show that $\{x_n(\cdot)\}_{n\geq 1}$ is relatively compact in $C(T, X)$. To this end let $t\in T\setminus\{b\}$ and $h>0$ s.t. $t+h\in T$. We know (see Barbu [10] and Benilan [11]), that:

$$||x_n(t+h)-x_n(t)||\leq ||x_n(h)-x_0||+\int_0^t ||g_n(s+h)-g_n(s)||\,ds,\ n\geq 1\,,$$

where $g_n(s)=\int_0^s K(s-r)\,f(r, x_n(r))\,u_n(r)\,dr$. So for all $n\geq 1$ we have:

$$\begin{aligned}&||x_n(t+h)-x_n(t)||\\ &\leq ||x_n(h)-S(h)\,x_0||+||S(h)\,x_0-x_0||+\int_0^t\Big|\Big|\int_0^{s+h} K(s+h-r)\,f(r, x_n(r))\,u_n(r)\,dr\\ &\quad -\int_0^s K(s-r)\,f(r, x_n(r))\,u_n(r)\,dr\Big|\Big|\,ds\\ &\leq ||x_n(h)-S(h)\,x_0||+||S(h)\,x_0-x_0||+\int_0^t\int_s^{s+h} ||K(s+h-r)||\,\psi(r)\,dr\,ds\\ &\quad +\int_0^t\int_0^s ||K(s+h-r)-K(s-r)||\,\psi(r)\,dr\,ds\,,\end{aligned}$$

where $\psi(t)=(a(t)+b(t)\,M)|W|$. Clearly $\psi(\cdot)\in L^1_+$. Then from the above inequalities, we have:

$$||x_n(t+h)-x_n(t)||\to 0 \text{ as } h\to 0^+, \text{ uniformly in } n\geq 1\,.$$

Similarly we can get that $||x_n(t)-x_n(t-h)||\to 0$ as $h\to 0^+$, uniformly in $n\geq 1$. Thus we deduce that $\{x_n(\cdot)\}_{n\geq 1}$ is equicontinuous.

Next let $\lambda>0$, $t\in T$. From Brézis [13], we know that

$$||x_n(t)-J_\lambda\,x_n(t)||\leq\frac{2}{s}\left(1+\frac{\lambda}{s}\right)\int_0^s ||S(r)\,x_n(t)-x_n(t)||\,dr,\ n\geq 1\,.$$

But note that

$$\|S(r)\,x_n(t)-x_n(t)\|\le\|S(r)\,x_n(t)-S(r)\,x_n(t-r)\|+\|S(r)\,x_n(t-r)-x_n(t)\|$$
$$\le\|x_n(t)-x_n(t-r)\|+\int_{t-r}^{t}\Big\|\int_0^s K(s-\tau)\,f(\tau,\,x_n(\tau))\,u_n(\tau)\,d\tau\Big\|\,ds\,.$$

Given that $\{x_n(\cdot)\}_{n\ge1}$ is an equicontinuous family, there exists $\mu(\cdot)$ increasing function s.t. $\mu(r)\to0$ as $r\to0$ and $\|x_n(t)-x_n(t-r)\|\le\mu(r)$ for all $n\ge1$. Hence

$$\|S(r)\,x_n(t)-x_n(t)\|\le\mu(r)+\int_{t-r}^{t}\int_0^s N\,\psi(r)\,d\tau\,ds=\hat{\mu}(r),\ n\ge1\,,$$

with $\hat{\mu}(\cdot)$ increasing and $\hat{\mu}(r)\to0$ as $r\to0$. So finally for $\lambda=\beta s$, $\beta\in\boldsymbol{N}_+$, we have

$$\|x_n(t)-J_\lambda\,x_n(t)\|\le\frac{2}{s}\,(1+\beta)\int_0^s\hat{\mu}(r)\,dr\leqq\frac{2}{s}\,(1+\beta)\,\hat{\mu}(s)\,s=2(1+\beta)\,\hat{\mu}\Big(\frac{\lambda}{\beta}\Big)$$
$$\to0\quad\text{as}\quad\lambda\to0$$

and the convergence is uniform in $n\ge1$.

But by hypothesis J_1 is compact and so by the resolvent identity J_λ for $\lambda>0$ is compact too. So the identity map on $\overline{\{x_n(t)\}}_{n\ge1}$ is compact $\Rightarrow\overline{\{x_n(t)\}}_{n\ge1}$ is compact for all $t\in T$. Invoking the Arzela-Ascoli theorem, we have that $\{x_n(\cdot)\}_{n\ge1}$ is relatively compact in $C(T,X)$. So by passing to a subsequence if necessary, we may assume that $x_n\to x$ in $C(T,X)$.

We have seen in an earlier estimation that for all $n\ge1$ and $h\in T$ s.t. $t+h\in T$ we have:

$$\|x_n(t+h)-x_n(t)\|\le\|x_n(h)-S(h)\,x_0\|+\|S(h)\,x_0-x_0\|+\int_0^t\|g_n(s+h)-g_n(s)\|\,ds$$

where as before $g_n(s)=\int_0^s K(s-r)\,f(r,\,x_n(r))\,u_n(r)\,dr$. Note that

$$\|x_n(h)-S(h)\,x_0\|\le\int_0^h\|g_n(s)\|\,ds=\int_0^h\Big\|\int_0^s K(s-r)\,f(r,\,x_n(r))\,u_n(r)\,dr\Big\|\,ds$$
$$\le\int_0^h N\int_0^s\psi(r)\,dr\,ds\le\int_0^h N\|\psi\|_1\,ds=N\|\psi\|_1\,h\,.$$

Also we know (see Barbu [10]), that:

$$\|S(h)\,x_0-x_0\|\le h\,\|Ax_0\|\,.$$

Finally set $\hat{f}_n(r)=f(r,\,x_n(r))\,u_n(r)$. We have:

$$\int_0^t\|g_n(s+h)-g_n(s)\|\,ds=\int_0^t\Big\|\int_0^{s+h}K(s+h-r)\,\hat{f}_n(r)\,dr-\int_0^s K(s-r)\,\hat{f}_n(r)\,dr\Big\|\,ds$$
$$\le\int_0^t\int_s^{s+h}N\,\psi(r)\,dr\,ds+\int_0^t\int_0^s\|K(s+h+r)-K(s-r)\|\,\psi(r)\,dr\,ds\,,\quad n\ge1\,,$$
$$\Rightarrow\Big\|\frac{x_n(t+h)-x_n(t)}{h}\Big\|\le N\int_0^t\frac{1}{h}\int_s^{r+h}\psi(r)\,dr\,ds$$

$$+\int_0^t \int_0^s \left\| \frac{K(s+h-r)-K(s-r)}{h} \right\| \psi(r)\, dr\, ds + |Ax_0| \quad n \geq 1\,,$$

$$\Rightarrow \|\dot{x}_n(t)\| \leq N\|\psi\|_1 + \int_0^t \int_0^s K'(s-r)\, \psi(r)\, dr\, ds + |Ax_0|$$

$$\leq N\|\psi\|_1 + N'\, \|\psi\|_1\, b + |Ax_0| = N_1, \quad \text{a.e.}\,.$$

where $\|K'(t)\| \leq N'$, $t \in T$ (recall that by hypothesis H(K), $K(\cdot) \in C^1(T, \mathcal{L}(X))$). Therefore for every $n \geq 1$ and almost every $t \in T$, we have

$$\|\dot{x}_n(t)\| \leq N_1\,.$$

Thus by passing to a subsequence if necessary, we may assume that $\dot{x}_n \xrightarrow{w} \dot{x}$ in $L^2(X)$.

Also since $S^2_{U(\cdot, x_n(\cdot))} \subseteq S^2_W$ and the latter is sequentially weakly compact in $L^2(Y)$, we may assume that $u_n \xrightarrow{w} u$ in $L^2(Y)$. From Theorem 3.1 of [37], we know that $u(t) \in \overline{\text{conv}}\ \text{w-}\overline{\lim}\ \{u_n(t)\}_{n\geq 1} \subseteq \overline{\text{conv}}\ \text{w-}\overline{\lim}\ U(t, x_n(t))$ a.e. and because of hypothesis H(U), we have $\text{w-}\overline{\lim}\ U(t, x_n(t)) \subseteq U(t, x(t))$ a.e.. So $u(t) \in U(t, x(t))$ a.e. $\Rightarrow u(\cdot) \in S^2_{U(\cdot, x(\cdot))}$.

Next let $\eta_n(t) = \int_0^t K(t-s) f(s, x_n(s))\, u_n(s)\, ds$ and $\eta(t) = \int_0^t K(t-s) f(s, x(s))\, u(s)\, ds$. From hypothesis H(f), we have $\eta_n(t) \xrightarrow{w} \eta(t)$ for all $t \in T \Rightarrow \eta_n \xrightarrow{w} \eta$ in $L^2(X)$. Observe that

$$(x_n, \dot{x}_n - \eta_n) \in \text{Gr}\, \hat{A}$$

where $\hat{A}$ is the realization (lifting) of A on $L^2(X)$. We know (see Barbu [10]), that $\hat{A}$ is m-dissipative too. Since $L^2(X)^* = L^2(X^*)$ and $L^2(X^*)$ is uniformly convex (see Day [20]), from Proposition 3.5, p. 75 of Barbu [10] we know that $\text{Gr}\, \hat{A}$ is demiclosed (i.e. closed in $L^2(X) \times L^2(X)_w$). Since $(x_n, \dot{x}_n - \eta_n) \xrightarrow{s \times w} (x, \dot{x} - \eta)$ in $L^2(X) \times L^2(X)$, we have $(x, \dot{x} - \eta) \in \text{Gr}\, \hat{A} \Rightarrow \dot{x}(t) \in Ax(t) + \eta(t)$ a.e. $\Rightarrow$ (x, u) is an admissible "state-control" pair.

Finally we will show that $J(x, u) \leq \underline{\lim}\, J(x_n, u_n)$. Recall that W with the weak topology (denoted henceforth by W_w) is compact metrizable (see Dunford-Schwartz [22]). So from Lemma 2 of Balder [8], we know that we can find L_m: $T \times X \times W_w \to \boldsymbol{R}$ Caratheodory integrands s.t. $L_m \uparrow L$ and as $m \to \infty$ and $\psi_1(t) + \psi_2(t)\, \|x\| + \psi_3(t)\, \|u\| \leq L_m(t, x, u) \leq m$ a.e.. Let $\{\delta_{u_n(\cdot)}(\cdot)\}_{n \geq 1} \subseteq L^\infty(T, M(W_w))$ be the Dirac transition probabilities associated with the functions $u_n(\cdot)$, $n \geq 1$. Then from Alaoglu's theorem and Theorem 1., p. 426 of Dunford-Schwartz [22], we may assume by passing to a subsequence if necessary, that $\delta_{u_n} \xrightarrow{w^*} \lambda$ in $L^\infty(T, M(W_w))$. We also claim that $L^n_m(t)(\cdot) = L_m(t, x_n(t), \cdot) \to L_m(t)(\cdot) = L_m(t, x(t), \cdot)$ in $C(W_w)$ as $n \to \infty$ for every $m \geq 1$. To this end note that

$$\sup_{u\in W}|L_m(t, x_n(t), u)-L_m(t, x(t), u)| = |L_m(t, x_n(t), v_n)-L_m(t, x(t), v_n)|$$

since $L_m(t, x_n(t), \cdot)-L_m(t, x(t), \cdot)$ is continuous on W_w. Again we may assume that $u_n \xrightarrow{w} u\in W$. Then because of the continuity of $L_m(t, \cdot, \cdot)$ on $X\times W_w$ we get:

$$|L_m(t, x_n(t), v_n)-L_m(t, x(t), v_n)| \to 0 \quad \text{as } n\to\infty\,,$$
$$\Rightarrow L_m^n(t)(\cdot)\to L_m(t)(\cdot) \text{ in } C(W_w) \quad \text{as } n\to\infty\,.$$

Hence by the dominated convergence theorem, we get that

$$L_m^n(\cdot)\to L_m(\cdot) \text{ in } L^1(T, C(W_w)) \quad \text{as } n\to\infty\,.$$

Denote by $\langle\cdot,\cdot\rangle$ the duality brackets for the pair $(L^1(T, C(W_w)), L^\infty(T, M(W_w)))$. We have:

$$\langle L_m^n, \delta_{u_n}\rangle \to \langle L_m, \lambda\rangle \quad \text{as } n\to\infty$$

and note that $\langle L_m^n, \delta_{u_n}\rangle=\int_0^b\int_W L_m(t, x_n(t), u)\,\delta_{u_n(t)}(du)\,dt=\int_0^b L_m(t, x_n(t), u_n(t))\,dt$, while $\langle L_m, \lambda\rangle=\int_0^b\int_W L_m(t, x(t), u)\,\lambda(t)\,(du)\,dt$.

Also by the monotone convergence theorem we have

$$\int_0^b\int_W L_m(t, x(t), u)\,\lambda(t)\,(du)\,dt \uparrow \int_0^b\int_W L(t, x, (t), u)\,\lambda(t)\,(du)\,dt \quad \text{as } m\to\infty\,.$$

By a diagonalization process, we have

$$\int_0^b L_{m(n)}(t, x_n(t), u_n(t))\,dt \to \int_0^b\int_W L(t, x(t), u)\,\lambda(t)\,(du)\,dt \quad \text{as } n\to\infty\,.$$

Recall that $u_n \xrightarrow{w} u$ in $L^2(Y)$. So for all $A\in B(T)$ we have

$$\int_A u_n(t)\,dt \xrightarrow{w} \int_A u(t)\,dt \quad \text{as } n\to\infty$$

and for every $n\geq 1$ $u_n(t)=\int_W u\,\delta_{u_n(t)}(du)$. Since $\delta_{u_n(\cdot)}(\cdot)\xrightarrow{w^*}\lambda$ in $L^\infty(T, M(W_w))$, we have:

$$\int_A\int_W u\,\delta_{u_n(t)}(du)\,dt \xrightarrow{w} \int_A\int_W u\lambda(t)\,(du)\,dt \quad \text{as } n\to\infty$$
$$\Rightarrow \int_A u(t)\,dt = \int_A\int_W u\lambda(t)\,(du)\,dt \quad \text{for all} \quad A\in B(T)\,,$$
$$\Rightarrow u(t) = \int_W u\lambda(t)\,(du) \quad \text{a.e.}\,.$$

Since $L(t, x, \cdot)$ is convex, from Jensen's inequality, we get

$$\int_0^b L(t, x(t), u(t))\, dt = \int_0^b L(t, x(t), \int_W u\lambda(t)\,(du))\, dt$$
$$\leq \int_0^b \int_W L(t, x(t)\, u)\, \lambda(t)\,(du)\, dt$$
$$\Rightarrow \int_0^b L(t, x(t)\, u(t))\, dt \leq \underline{\lim} \int_0^b L(t, x_n(t), u_n(t))\, dt$$
$$\Rightarrow J(x, u) \leq \underline{\lim}\, J(x_n, u_n) = m\,.$$

Since (x, u) is admissible, we conclude that $J(x, u)=m$.

REMARKS. (1) Given $u \in S_W^2$, the evolution inclusion describing the dynamics of (*), has a strong solution $x(\cdot) \in C(T, X)$. This follows from a standard fixed point argument as in Hirano [29] (Theorem 2.1 and Corollary 2.1).

(2) It remains an open problem whether instead of the linearity with respect to $u(\cdot)$ of the dynamic equation, we can have a more general convexity hypothesis on an appropirate orientor field (analogous to "property Q″ of Cesari [7]) and then apply the "Cesari-Rockafellar reduction technique" to establish the existence of optimal admissible pairs (see [40] where this approach is applied to a more restricted class of problems with no "memory").

4. Relaxed problem

In this section we drop the nice features that problem (*) had. Namely we no longer assume that the control function appears linearly in the dynamics and that the cost integrand is convex in u. Then in order to guarantee the existence of an optimal admissible pair, we need to pass to a larger, "convexified" system, known as "relaxed system". This new augmented system has measure valued controls. More specifically the relaxed system has the following form:

$$\left\{\begin{array}{l} J_r(x, \lambda) = \int_0^b \int_W L(t, x(t), z)\, \lambda(t)\,(dz)\, dt \to \inf = m_r \\ \text{s.t. } \dot{x}(t) \in Ax(t) + \int_0^t K(t-s) \int_W f(s, x(s), z)\, \lambda(s)\,(dz)\, ds \quad \text{a.e.} \\ \quad x(0) = x_0,\ \lambda(t) \in \Sigma(t) = \{\mu \in M_+^1(W_w) \colon \mu(U(t)) = 1\} \\ \qquad \lambda(\cdot) \quad \text{is measurable} \end{array}\right\} (*)_r$$

Note that in this problem $f(t, x, \cdot)$ is nonlinear and the original control constraint set is state independent (open loop). So hypotheses H(f), H(U), now have the following forms.

H(f)′: $f\colon T \times X \times Y \to X$ is a map s.t.
(1) $t \to f(t, x, u)$ is measurable,
(2) $(x, u) \to f(t, x, u)$ is sequentially continuous from $X \times Y_w$ into X_w (where X_w and Y_w denote the spaces X and Y with their respective weak topologies),

(3) $\|f(t, x, u)\| \leq a(t)+b(t)\,(\|x\|+\|u\|)$ a.e. $a(\cdot)\in L^2_+$, $b(\cdot)\in L^\infty_+$.

H(U)′: $U\colon T\to P_{fc}(Y)$ is a measurable multifunction s.t. $U(t)\subseteq W$ a.e. with $W\in P_{wkc}(Y)$.

First we will show that the relaxed optimal control problem always has a solution. By $P_r(x_0)$ we will denote the set of admissible "state-control" pairs for $(*)_r$.

Theorem 4.1. *If hypotheses* H(A), H(f)′, H(K), H(U)′, H(L) *and* H_α *hold, then there exists* $(x, \lambda)\in P_r(x_0)$ s.t. $J_r(x, \lambda)=m_r$.

Proof. Let $\{(x_n, \lambda_n)\}_{n\geq 1}\subseteq P_r(x_0)$ be a minimizing sequence for $(*)_r$. As in the proof of Theorem 3.1 we can show that $\{x_n(\cdot)\}_{n\geq 1}$ is relatively compact in $C(T, X)$ and $\{\dot{x}_n(\cdot)\}_{n\geq 1}$ is uniformly bounded in $L^2(X)$, thus relatively sequentially weakly compact in $L^2(X)$. Also $\{\lambda_n(\cdot)\}_{n\geq 1}$ is relatively sequentially w*-compact in $L^\infty(T, M(W_w))$ (see Dunford-Schwartz [22], Theorem 2, p. 434). So we may assume that $x_n\to x$ in $C(T, X)$, $\dot{x}_n \xrightarrow{w} \dot{x}$ in $L^2(X)$ and $\lambda_n \xrightarrow{w^*} \lambda$ in $L^\infty(T, M(W_w))$. We claim that $(x, \lambda)\in P_r(x_0)$. By hypothesis for every $n\geq 1$, we have:

$$\dot{x}_n(t)-g_n(t)\in Ax_n(t)\,, \quad x_n(0) = x_0$$

where $g_n(t)=\int_0^t K(t-s)\int_W f(s, x_n(s), z)\,\lambda_n(s)(dz)\,ds$. Let $x^*\in X^*$. We have $(x^*, g_n(t))=\int_0^t (x^*, K(t-s)\int_W f(s, x_n(s), z)\,\lambda_n(s)(dz))\,ds=\int_0^t\int_W (x^*, K(t-s) f(s, x_n(s), z))\,\lambda_n(s)(dz)\,ds=\langle h_n(t, x^*)(\cdot), \lambda_n(\cdot)\rangle_{[0,t]}$, with $h_n(t, x^*)\in L^1(T, C(W_w))$ defined by $h_n(t, x^*)(s)(\cdot)=(x^*, K(t-s)f(s, x_n(s), \cdot))$ and $\lambda_n(\cdot)\in S_\Sigma=$ {set of measurable selectors of $\Sigma(\cdot)$}. So $\lambda\in S_\Sigma$. Also as in the proof of Theorem 3.1 we can show that $h_n(t, x^*)(\cdot)\to h(t, x^*)(\cdot)$ in $L^1(T, C(W_w))$, with $h(t, x^*)(s)(\cdot)=(x^*, K(t-s)f(s, x(s), \cdot))$. Hence we have:

$$\begin{aligned}&\langle h_n(t, x^*), \lambda_n\rangle_{[0,t]} \to \langle h(t, x^*), \lambda\rangle_{[0,t]}\,,\\ &\Rightarrow g_n(t) \xrightarrow{w} g(t) = \int_0^t K(t-s)\int_W f(s, x(s), z)\,\lambda(s)(dz)\,ds \quad \text{for all} \quad t\in T\,,\\ &\Rightarrow g_n \xrightarrow{w} g \text{ in } L^2(X)\,.\end{aligned}$$

Thus we get

$$(x_n, \dot{x}_n-g_n) \xrightarrow{s\times w} (x, \dot{x}-g) \quad \text{in } L^2(X)\times L^2(X)$$

and for all $n\geq 1$ $(x_n, \dot{x}_n-g_n)\in \mathrm{Gr}\,\hat{A}$, which is demiclosed since $\hat{A}$ being the realization (lifting) of A on $L^2(X)$, is m-dissipative and $L^2(X)^*=L^2(X^*)$ is uniformly convex (see Barbu [10]). Hence $(x, \dot{x}-g)\in \mathrm{Gr}\,A \Rightarrow \dot{x}(t)\in Ax(t)+g(t)$

a.e. $\Rightarrow \dot{x}(t) \in A(t) + \int_0^t K(t-s) \int_W f(s, x(s), z)\, \lambda(s)\, (dz)\, ds \Rightarrow (x, \lambda) \in P_r(x_0)$ as claimed.

Finally by approximating $L(t, x, u)$ with Caratheodory integrands on $T \times X \times W_w$, we can show as in the proof of Theorem 3.1, that

$$\begin{aligned} J_r(x, \lambda) &= \int_0^b \int_W L(t, x(t), z)\, \lambda(t)\, (dz)\, dt \leq \varliminf \int_0^b \int_W L(t, x_n(t), z)\, \lambda_n(t)\, (dz)\, dt \\ &= \varliminf J_r(x_n, \lambda_n) = m_r \\ &\Rightarrow m_r = J_r(x, \lambda) \\ &\Rightarrow (x, \lambda) \in P_r(x_0) \text{ is the desired solution of } (*)_r. \end{aligned}$$

The augmented relaxed system will be useful if its value, equals that of the original problem; i.e. $m = m_r$. Namely we want to show that the relaxed system captures the asymptotic behavior of the minimizing sequences of the original problem. Such a result is usually called "relaxation theorem". Under some additional hypotheses on the cost integrand $L(\cdot, \cdot, \cdot)$, we can have a "relaxation theorem". So we will need the following stronger version of hypothesis H(L).

H(L)′: $L: T \times X \times Y \to \boldsymbol{R}$ is an integrand s.t.
(1) $t \to L(t, x, u)$ is measurable,
(2) $(x, u) \to L(t, x, u)$ is continuous on $X \times Y_w$,
(3) $|L(t, x, u)| \leq \psi_1(t) + \psi_2(t)\, (\|x\|^2 + \|u\|^2)$ a.e. with $\psi_1(\cdot), \psi_2(\cdot) \in L^1_+$.

If the control constraint set is state independent (open loop), then from Hirano's existence result [29] (see also remark (1) after the proof of Theorem 3.1), we deduce that hypothesis H_α is satisfied by all $(x, u) \in P(x_0) \neq \emptyset$.

Theorem 4.2. *If hypotheses* H(A), H(f)″, H(K), H(U)′ *and* H(L)′ *hold, then problem* $(*)_r$ *has a solution and* $m = m_r$.

Proof. The existence of a solution for problem $(*)_r$ follows from Theorem 4.1. Let (x, λ) be an optimal "state-control" pair for $(*)_r$. Invoking Corollary 3 of Balder [9], we can find $u_n(\cdot) \in S^1_U$ s.t. $\delta_{u_n} \to \lambda$ weakly in $R(T, W_w)$ (see section 2). Since W_w is compact, metrizable, the weak topology on $R(T, W_w)$ coincides with the relative $w^*(L^\infty(T, M(W_w)), L^1(T, C(W_w)))$-topology. So $\delta_{u_n} \xrightarrow{w^*} \lambda$. Let $x_n(\cdot) \in C(T, X)$ be the trajectory generated by $u_n(\cdot)$, $n \geq 1$. Exactly as in the proof of Theorem 3.1, via the Arzela-Ascoli theorem, we can show that $\{x_n(\cdot)\}_{n \geq 1}$ is relatively compact in $C(T, X)$ (see Proposition 5.1). So by passing to a subsequence if necessary, we may assume that $x_n \to x$ in $C(T, X)$. As before, set $\hat{L}_n(t)(\cdot) = L(t, x_n(t), \cdot)$ and $\hat{L}(t)(\cdot) = L(t, x(t), \cdot)$. We have $\hat{L}_n \to \hat{L}$ in $L^1(T, C(W_w))$. Hence

$$\langle \hat{L}_n, \delta_{u_n} \rangle = \int_0^b L(t, x_n(t), u_n(t))\, dt \to \langle \hat{L}, \lambda \rangle = \int_0^b \int_W L(t, x(t), z)\, \lambda(t)\,(dz)\, dt$$
$$\Rightarrow J_r(x_n, \delta_{u_n}) = J(x_n, u_n) \to J_r(x, \lambda) = m_r \quad \text{as} \quad n \to \infty$$
$$\Rightarrow m \leq m_r\,.$$

On the other hand we always have $m_r \leq m$. Hence we conclude that $m = m_r$.

Again the memory feature of our system causes serious technical difficulties in our attempt to introduce an alternative, control free formulation of the relaxed system. This was done by Ekeland [24] for memoryless systems driven by semilinear elliptic equations and by Papageorgiou [38] for memoryless, completely nonlinear systems driven by evolution inclusions.

5. Relaxation theorems

In this section we will show that the equivalent relaxed problems $(*)_r$ and $(*)'_r$ are the "closure" of the original problem.

This is first illustrated by a density result, which shows that the original trajectories are dense in the relaxed ones for the $C(T, X)$-topology.

For this we will need a stronger hypothesis on $f(\cdot, \cdot, \cdot)$.

H(f)'': $f\colon T \times X \times Y \to X$ is a map s.t.
(1) $t \to f(t, x, u)$ is measurable,
(2) $(x, u) \to f(t, x, u)$ is continuous from $X \times Y_w$ into X_w,
(3) $\|f(t, x, u) - f(t, x', u)\| \leq \eta(t)\, \|x - x'\|$ a.e. for all $u \in U(t)$ and with $\eta(\cdot) \in L^1_+$,
(4) $\|f(t, x, u)\| \leq a(t) + b(t)\,(\|x\| + \|u\|)$ a.e. with $a(\cdot) \in L^2_+$, $b(\cdot) \in L^\infty_+$.

Recall that by $P_1(x_0)$ we denote the set of trajectories of the original system and by $P_{r1}(x_0)$ the set of relaxed trajectories.

Let $h \in L^2(X)$. By $|h|_w$ we will denote the weak norm on $L^2(X)$, which is defined by $|h|_w = \sup \{\|\int_t^{t'} h(s)\, ds\| : t, t' \in T\}$. Suppose that $\sup_{n \geq 1} \|h_n\|_2 < \infty$ and $h_n \xrightarrow{|\cdot|_w} h$. We claim $h_n \xrightarrow{w} h$ in $L^2(X)$. To see this let $s\colon T \to X$ be a step function i.e. $s(t) = \sum_{k=1}^m \chi_{[t_{k-1}, t_k]}\, r_k$, $s(b) = r_m$, with $t_0 = 0$, $t_m = b$. Then

$$|(h_n - h, s)| \leq \sum_{k=1}^n \|\int_{t_{k-1}}^{t_k} (h_n(s) - h(s))\, ds\|\, \|r_k\| \leq |h_n - h|_w \sum_{k=1}^m \|r_k\| \to 0 \text{ as } n \to \infty\,.$$

To conclude the proof of the claim, recall that step functions are dense in $L^2(X)$.

Theorem 5.1. *If hypotheses* H(A), H(f)'', H(K) *and* H(U)' *hold, then* $P_{r1}(x_0) = \overline{P_1(x_0)}$, *the closure in* $C(T, X)$.

Proof. Let $x(\cdot)\in P_{r1}(x_0)$. Then by definition we have

$$\left\{\begin{array}{l} \dot{x}(t)\in Ax(t)+\int_0^t\int_W K(t-s)f(s, x(s), z)\,\lambda(s)\,(dz)\,ds \\ x(0)=x_0,\ \lambda(\cdot)\in S_\Sigma \end{array}\right\}$$

Using Theorem 12.11, p. 221 of Choquet [18] we get, that if $F(t, x(t))= f(t, x(t), U(t))$, then

$$\overline{\text{conv}}\, F(t, x(t)) = \{\int_W f(t, x(t), z)\,\lambda(dz)\colon \lambda\in\Sigma(t)\}$$

and so

$$S^1_{\overline{\text{conv}}\, F(\cdot, x(\cdot))} = \{t\to\int_W f(t, x(t), z)\,\lambda(t)\,(dz)\colon \lambda(\cdot)\in S_\Sigma\}\ .$$

Invoking Theorem 2 of Chuong [19], we can find $h_n\in S^1_{F(\cdot, x(\cdot))}$ s.t. $h_n \xrightarrow{|\cdot|_w} h$ where $h(t)=\int_W f(t, x(t), z)\,\lambda(t)\,(dz)$. Because of hypothesis H(f)″ (4), $h_n\in L^2(X)$ and $\sup_{n\geq 1}\|h_n\|_2<\infty$. So we have $h_n \xrightarrow{w} h$ in $L^2(X)$. Let $R_n(t)=\{u\in U(t)\colon h_n(t)= f(t, x(t), u)\}$, $n\geq 1$. From the definition of $F(t, x)$, we see that $R_n(t)\neq\emptyset$ for all $t\in T$ and all $n\geq 1$. Also from hypotheses H(f)″ and H(U)′, we have Gr $R_n= \{(t, u)\in T\times W\colon h_n(t)=f(t, x(t), u)\}\cap \text{Gr}\, U\in B(T)\times B(W)$. Apply Aumann's selection theorem, to find $u_n\colon T\to Y$ measurable s.t. $u_n(t)\in R_n(t)$ a.e. Then $h_n(t)=f(t, x(t), u_n(t))$ a.e. Let $x_n(\cdot)$ be the unique strong solution of

$$\left\{\begin{array}{l} \dot{x}_n(t)\in Ax_n(t)+\int_0^t K(t-s)f(s, x_n(s), u_n(s))\,ds \\ x_n(0)=x_0 \end{array}\right\}$$

(see Hirano [29]). Then $\{x_n(\cdot)\}_{n\geq 1}\subseteq P_{r1}(x_0)$. Arguing as in the proof of Theorem 3.1, we can show that $\{x_n(\cdot)\}_{n\geq 1}$ is relatively compact in $C(T, X)$. So by passing to a subsequence if necessary, we may assume that $x_n\to y$ in $C(T, X)$.

For every $n\geq 1$ we have

$$\dot{x}_n(t)-\int_0^t K(t-s)f(s, x_n(s), u_n(s))\,ds\in Ax_n(t)\quad \text{a.e.}$$

and
$$\dot{x}(t)-\int_0^t K(t-s)\int_W f(s, x(s), z)\,\lambda(s)\,(dz)\,ds\in Ax(t)\quad \text{a.e..}$$

Since by hypothesis H(A), $A(\cdot)$ is m-dissipative, we have:

$$(\dot{x}_n(t)-\int_0^t K(t-s)f(s, x_n(s), u_n(s))\,ds-\dot{x}(t)+\int_0^t K(t-s)\int_W f(s, x(s), z)\,\lambda(s)\,(dz)\,ds\ ,$$
$$\hat{F}(x_n(t)-x(t)))\leq 0\quad \text{a.e.}$$

where $\hat{F}\colon X\to X^*$ is the duality map, which is single valued since by hypothesis X^* is uniformly convex (see section 2).

So we have:

$$\begin{aligned}&(\dot{x}_n(t)-\dot{x}(t), \hat{F}(x_n(t)-x(t)))\\ &\leq (\int_0^t K(t-s) f(s, x_n(s), u_n(s))\, ds-\int_0^t K(t-s) \int_W f(s, x(s), z)\, \lambda(s)\, (dz)\, ds,\\ &\quad \hat{F}(x_n(t)-x(t))) \quad \text{a.e.}.\end{aligned}$$

From Kato [31] (see also Barbu [10], Lemma 1.2, p. 100), we know that:

$$(\dot{x}_n(t)-\dot{x}(t), \hat{F}(x_n(t)-x(t))) = ||x_n(t)-x(t)|| \frac{d}{dt} ||x_n(t)-x(t)|| \quad \text{a.e.}.$$

Also if we set $q_n(s)=f(s, x_n(s), u_n(s))$, we have:

$$\begin{aligned}&(\int_0^t K(t-s)\, q_n(s)\, ds-\int_0^t K(t-s)\, h_n(s)\, ds+\int_0^t K(t-s)\, h_n(s)\, ds\\ &\quad -\int_0^t K(t-s)\, h(s)\, ds, \hat{F}(x_n(t)-x(t)))\\ &= (\int_0^t K(t-s)\, [q_n(s)-h_n(s)]\, ds, \hat{F}(x_n(t)-x(t)))\\ &\quad +(\int_0^t K(t-s)\, [h_n(s)-h(s)]\, ds, \hat{F}(x_n(t)-x(t))).\end{aligned}$$

Hence we can write that

$$\begin{aligned}&||x_n(t)-x(t)|| \frac{d}{dt} ||x_n(t)-x(t)||\\ &\leq (\int_0^t K(t-s)\, [q_n(s)-h_n(s)]\, ds+\int_0^t K(t-s)\, [h_n(s)-h(s)]\, ds,\\ &\quad \hat{F}(x_n(t)-x(t))) \quad \text{a.e.}.\end{aligned}$$

Integrating the above inequality, we get:

$$\begin{aligned}&||x_n(t)-x(t)||^2\\ &\leq 2\int_0^t (\int_0^s K(s-r)\, (q_n(r)-h_n(r))\, dr, \hat{F}(x_n(s))-x(s)))\, ds\\ &\quad +2\int_0^t (\int_0^s K(s-r)\, (h_n(r)--h(r))\, dr, \hat{F}(x_n(s)-x(s)))\, ds.\end{aligned}$$

Since $h_n \xrightarrow{w} h$ in $L^2(X)$ and $\hat{F}(x_n(\cdot)-x(\cdot))$ converges in $C(T, X^*)$, we see that

$$2\int_0^t (\int_0^s K(s-r)\, (h_n(r)-h(r))\, dr, \hat{F}(x_n(s)-x(s)))\, ds \to 0 \quad \text{as} \quad n \to \infty.$$

Also by hypothesis H(f)″ (3), we have:

$$\begin{aligned}&2\int_0^t (\int_0^s K(s-r)\, (q_n(r)-h_n(r))\, dr, \hat{F}(x_n(s)-\hat{x}(s)))\, ds\\ &\leq 2\int_0^t ||K||_\infty \int_0^s \eta(r)\, ||x_n(r)-x(r)||\, dr\, ||x_n(s)-x(s)||\, ds.\end{aligned}$$

Consider the following norm on $C(T, X)$

$$||w||_0 = \sup_{t \in T} [\exp(-\lambda \int_0^t \eta(s)\, ds)\, ||w(t)||],\ w(\cdot) \in C(T, X),\ \lambda \in \boldsymbol{R}_+ .$$

Clearly this is equivalent to the usual supremum norm $||\cdot||_\infty$. Using norm $||\cdot||_0$, we have:

$$\begin{aligned}
&2\int_0^t ||K||_\infty \int_0^s \eta(r) \exp(-\lambda \int_0^r \eta(v)\, dv) \exp(\lambda \int_0^r \eta(v)\, dv)\, ||x_n(r)-x(r)||\, dr \\
&\quad ||x_n(s)-x(s)||\, ds \\
&\leq 2\int_0^t ||K||_\infty\, ||x_n - x||_0 (\int_0^s \eta(r) \exp(\lambda \int_0^r \eta(v)\, dv)\, dr)\, ||x_n(s)-x(s)||\, ds \\
&\leq 2\int_0^t ||K||_\infty\, ||x_n - x||_0 (\int_0^s \frac{1}{\lambda}\, d(\exp(\lambda \int_0^r \eta(v)\, dv)))\, ||x_n(s)-x(s)||\, ds \\
&\leq 2\int_0^t ||K||_\infty\, ||x_n - x||_0 \frac{1}{\lambda} (\exp \lambda \int_0^s \eta(r)\, dr)\, ||x_n(s)-x(s)||\, ds .
\end{aligned}$$

Without any loss of generality, we may assume that $1 < \eta(t)$ for all $t \in T$. Hence we have that:

$$\begin{aligned}
&\frac{2}{\lambda}\int_0^t ||K||_\infty\, ||x_n - x||_0 \exp(\lambda \int_0^s \eta(r)\, dr)\, ||x_n(s)-x(s)||\, ds \\
&\leq \frac{2}{\lambda}\int_0^t ||K||_\infty\, ||x_n - x||_0\, \eta(s) \exp(2\lambda \int_0^s \eta(r)\, dr) \exp(-\lambda \int_0^s \eta(r)\, dr) \\
&\quad ||x_n(s)-x(s)||\, ds \\
&\leq \frac{2}{\lambda}\int_0^t ||K||_\infty\, ||x_n - x||_0^2\, \eta(s) \exp(2\lambda \int_0^s \eta(r)\, dr)\, ds \\
&\leq \frac{||x_n - x_0||_0^2\, ||K||_\infty}{\lambda} \frac{2}{2\lambda} \exp(2\lambda \int_0^t \eta(s)\, ds) \\
&= \frac{||x_n - x_0||_0^2\, ||K||_\infty}{\lambda^2} \exp(2\lambda \int_0^t \eta(s)\, ds) .
\end{aligned}$$

Hence finally we have:

$$\begin{aligned}
||x_n(t)-x(t)||^2 &\leq \frac{||x_n - x_0||_0^2\, ||K||_\infty}{\lambda^2} \exp(2\lambda \int_0^t \eta(s)\, ds) + 2\int_0^t (\int_0^s K(s-r)\, (h_n(r) \\
&\quad - h(r))\, dr,\ \hat{F}(x_n(s)-x(s))) .
\end{aligned}$$

Passing to the limit as $n \to \infty$, we get

$$\begin{aligned}
&||y(t)-x(t)||^2 \leq \frac{||y-x||_0^2\, ||K||_\infty}{\lambda^2} \exp(2\lambda \int_0^t \eta(s)\, ds) \\
&\Rightarrow (\exp(-\lambda \int_0^t \eta(s)\, ds)\, ||y(t)-x(t)||)^2 \leq \frac{||y-x||_0^2\, ||K||_\infty}{\lambda^2} \\
&\Rightarrow \exp(-\lambda \int_0^t \eta(s)\, ds)\, ||y(t)-x(t)|| \leq \frac{||y-x||_0\, ||K||_\infty^{1/2}}{\lambda^2} \quad \text{for all} \quad t \in T
\end{aligned}$$

$$\Rightarrow \|y-x\|_0 \leq \frac{\|K\|_\infty^{1/2}}{\lambda} \|y-x\|_0 .$$

Choose $\lambda > \|K\|_\infty^{1/2}$. Then from the above inequality we deduce that $y=x \Rightarrow x \in \overline{P_1(x_0)} \Rightarrow P_{r1}(x_0) \subseteq \overline{P_1(x_0)}$. But $P_1(x_0) \subseteq P_{r1}(x_0)$ and as in the proof of Theorem 3.1 we can show that $P_{r1}(x_0)$ is relatively compact in $C(T, X)$, while as in the proof of Theorem 4.1, we can show that $P_{r1}(x_0)$ is closed in $C(T, X)$. So $\overline{P_1(x_0)} \subseteq P_{r1}(x_0) \Rightarrow \overline{P_1(x_0)} = P_{r1}(x_0)$ as claimed by the theorem.

REMARKS. (1) Simple continuity of $x \to f(t, x, u)$ is not enough in order to prove the above density result. In the theory of differential inclusions there is a nice two dimensional example that illustrates this. For details we refer to the book of Aubin-Cellina [7].

(2) To our knowledge, the first such density result for distributed parameter systems was proved by Ahmed [3] for a class of semilinear systems with no memory. However as it was indicated in [39], there are some problems with his proof.

Next we will derive two theorems that illustrate that the relaxed problem describes the asymptotic behavior of the minimizing sequences of the original one. For this we will need the following result, which is a continuous dependence result of the elements in $P_{r1}(x_0)$ on the relaxed controls that generate them. So consider the following evolution with $\lambda \in L^\infty(T, M(W_w))$:

$$\left\{ \begin{array}{ll} \dot{x}(t) \in Ax(t) + \int_0^t K(t-s) \int_W f(s, x(s), z)\, \lambda(s)\, (dz)\, ds & \text{a.e.} \\ x(0) = x_0 & \end{array} \right\} (**)_\lambda$$

From Hirano [29] we know that this has a strong solution, which is unique if hypothesis H(f)″ (3) is valid. Denote this solution by $x(\lambda)(\cdot) \in C(T, X)$.

Proposition 5.1. *If hypotheses* H(A), H(f)″ *and* H(K) *hold, then* $\lambda \to x(\lambda)$ *is sequentially continuous from* $L^\infty(T, M(W_w))_{w^*}$ *into* $C(T, X)$.

Proof. Let $\lambda_n \overset{w^*}{\to} \lambda$ in $L^\infty(T, M(W_w))$. Set $x_n(\cdot) = x(\lambda_n)(\cdot)$. Then as before we can show that $\{x_n(\cdot)\}_{n\geq 1}$ is relatively compact in $C(T, X)$ and so we may assume that $x_n \to \hat{x}$ in $C(T, X)$. Also as in the Proof of Theorem 3.1, we may assume $\dot{x}_n \overset{w}{\to} \dot{\hat{x}}$ in $L^2(X)$. Working as in Theorem 4.1 with the "lifting" $\hat{A}$, we finally get $\hat{x} = x(\lambda) \Rightarrow \lambda \to x(\lambda)$ is sequentially continuous as claimed.

Now we are ready for the first limit result.

Theorem 5.2. *If hypotheses* H(A), H(f)″, H(K), H(U)′, H(L)′ *hold and* $(x, \lambda) \in P_r(x_0)$, *then there exists a sequence* $\{(x_n, u_n)\}_{n\geq 1} \subseteq P(x_0)$ *s.t.* $J(x_n, u_n) \to J_r(x, \lambda)$

Proof. Using Corollary 3 of Balder [9], we can find $u_n(\cdot)\in S_U^1$ s.t. $\delta_{u_n}\xrightarrow{w^*}\lambda$ in $L^\infty(T, M(W_w))$. Let $x_n(\cdot)=x(\delta_{u_n})(\cdot)$. Then $(x_n, u_n)\in P(x_0)$ and from Proposition 5.1 we have $x_n\to x$ in $C(T, X)$. Then because of hypothesis H(L)$'$, as in the proof of Theorem 4.2, we have $J(x_n, u_n)=J_r(x_n, \delta_{u_n})\to J_r(x, \lambda)$ as $n\to\infty$.

From this result we deduce the second limit theorem.

Theorem 5.3. *If hypotheses* H(A), H(f)$'$, H(K), H(U)$'$, H(L)$'$ *hold and* $(x, \lambda)\in P_r(x_0)$ *solves* $(*)_r$, *then* $m=m_r$ *and there exists a minimizing sequence* $\{(x_n, u_n)\}_{n\geq 1}\subseteq P(x_0)$ *of the original problem s.t.* $x_n\to x$ *in* $C(T, X)$ *and* $\delta_{u_n}\xrightarrow{w^*}\lambda$ *in* $L^\infty(T, M(W_w))$.

Proof. Is an immediate consequence of Theorems 4.2 and 5.2.

6. Examples

In this section we present in detail two examples. The first is a hyperbolic control system and the second is a parabolic one.

EXAMPLE #1. "Hyperbolic optimal control problem".

Let $T=[0, b]$ be the time interval and $V=[0, a]$ the space interval.

Consider the following optimal control problem, with dynamics defined on $T\times V$ by hyperbolic integro-partial differential equation:

$$J_1(x, u) = \int_0^b\int_0^a L\left(t, z, x(t, z), \frac{\partial x(t, z)}{\partial t}, u(t, z)\right) dz\, dt \to \inf = m_1$$

$$\text{s.t. } \frac{\partial^2 x(t, z)}{\partial t^2} = \frac{\partial^2 x(t, z)}{\partial z^2} + \int_0^t K(t-s) f(s, z, x(s, z))\, u(s, z)\, ds \quad \text{on} \quad T\times V \quad (***)_1$$

$$x(t, 0) = x(t, a) = 0$$

$$x(0, z) = x_0(z) \quad \text{and} \quad \left.\frac{\partial x(t, z)}{\partial t}\right|_{t=0} = x_1(z) \quad \text{on} \quad \{0\}\times V$$

$$|u(t, z)| \leq M \text{ a.e. on } T\times V.$$

We will need the following hypotheses on the data of $(***)_1$.

H(K)$_1$: K: $T\to \boldsymbol{R}$ is a C^1-function.

H(f)$_1$: f: $T\times V\times \boldsymbol{R}\to\boldsymbol{R}$ is a function s.t.

(1) $(t, z)\to f(t, z, x)$ is measurable,

(2) $x\to f(t, z, x)$ is continuous,

(3) $|f(t, z, x)|\leq a(t, z)+b(t)|x|$ with $a(\cdot, \cdot)\in L^2(T\times V)$, $b(\cdot)\in L^\infty(T)$.

H(L)$_1$: L: $T\times V\times\boldsymbol{R}\times\boldsymbol{R}\times\boldsymbol{R}\to\bar{\boldsymbol{R}}=\boldsymbol{R}\cup\{+\infty\}$ is an integrand s.t.

(1) $(t, z, x, y, u)\to L(t, z, x, y, u)$ is measurable,

(2) $(x, y, u)\to L(t, z, x, y, u)$ is l.s.c. and convex in u,

(3) $\psi_1(t, z)\leq L(t, z, x, y, u)$ a.e. with $\psi_1(\cdot, \cdot)\in L^1(T\times V)$.

H_0: $x_0 \in H_0^1(V)$ and $x_1 \in L^2(V)$.

Also we will need the following "feasibility" hypothesis:

H_{a_1}: There exists admissible "state-control" pair for $(***)_1$ s.t. $J(x, u) < \infty$.

Let $X = H_0^1(V) \times L^2(V)$. This is a Hilbert space with inner product defined by

$$\langle (y_1, y_2), (\bar{y}_1, \bar{y}_2) \rangle = \int_0^a y_1'(z)\, \bar{y}_1'(z)\, dz + \int_0^a y_2(z)\, \bar{y}_2(z)\, dz .$$

Also X is separable, since both $H_0^1(V)$ and $L^2(V)$ are.

Define the operator A: $D(A) \subseteq X \to X$ by

$$A = \begin{bmatrix} 0 & 1 \\ d^2/dz^2 & 0 \end{bmatrix}$$

with $D(A) = \{(y_1, y_2) \in X : y_1 \in H_0^1(V) \cap H^2(V), y_2 \in H_0^1(V)\}$.

From Barbu [10] we know that $A(\cdot)$ defined above is m-dissipative. We claim that $(I-A)^{-1}$ is compact. To this end let $B \subseteq X$ be bounded. We need to show that $\overline{(I-A)^{-1}(B)}$ is compact in X. But note that $(y_1, y_2) \in (I-A)^{-1}(B) = R(B)$ if and only if

$$\begin{aligned} y_1(z) - y_2(z) &= \hat{y}_1(z) \\ y_2(z) - \frac{d^2}{dz^2} y_1(z) &= \hat{y}_2(z) \end{aligned}$$

where $\hat{y}_1, \hat{y}_2 \in B$. Adding those two equalities, we get that $y_1(z) - d^2 y_1(z)/dz^2 = \hat{y}_1(z) + \hat{y}_2(z)$. From this and Poincare's inequality, we get that $\{y_1 : y_1 \in \text{proj}_1 R(B)\}$ is bounded in $H_0^1(V)$. Thus from Sobolev's embedding theorem (see for example Adams [1]), we conclude that $\overline{R(B)}$ is compact in $X \Rightarrow (I-A)^{-1}$ is compact and so we have satisfied hypothesis H(A).

Set $Y = L^2(V)$ and define $\hat{f}$: $T \times X \times Y \to X$ by

$$\hat{f}(t, y)\, u(\cdot) = (0, f(t, \cdot, y_1(\cdot))\, u(\cdot)) .$$

It is routine to check that $\hat{f}(\cdot, \cdot, \cdot)$ defined as above, satisfies H(f).

Here the control constraint set is $U = \{u \in Y : |u(z)| \leq M\} \subseteq L^2(V)$ and let $\hat{x}_0 = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} \in X$. Also let $\hat{K}$: $T \to \mathcal{L}(X)$ be defined by $\hat{K}(t) = (0, K(t)\, I)$. Because of hypothesis $H(K)_1$, $\hat{K}(\cdot)$ is a C^1-map from T into $\mathcal{L}(X)$.

Finally let $\hat{L}$: $T \times X \times Y \to \mathbf{R}$ be defined by

$$\hat{L}(t, y, u) = \int_0^a L(t, z, y_1(z), y_2(z), u(z))\, dz$$

From Balder [8] (see also Ekeland-Temam [25]), we know that $\hat{L}(\cdot, \cdot, \cdot)$ satisfies hypothesis H(L).

Rewrite $(***)_1$ as the following abstract control problem:

$$\left.\begin{cases} \hat{J}_1(x,u)=\int_0^b \hat{L}(t,x(t),u(t))\,dt \to \inf = m_1 \\ \text{s.t. } \dot{x}(t)\in Ax(t)+\int_0^t \hat{K}(t-s)\hat{f}(s,x(s))\,u(s)\,ds \text{ a.e.} \\ \quad x(0)=x_0 \\ \quad u(t)\in U \text{ a.e. } u(\cdot) \text{ is measurable} \end{cases}\right\} (***)_1'$$

Note that this is a special case of problem (*). So we can apply Theorem 3.1 to get a solution of $(***)_1'$ and so of $(***)_1$ too.

Theorem 6.1. *If hypotheses* $H(K)_1$, $H(f)_1$, $H(L)_1$, H_0 *and* H_{a_1} *hold, then there exists an admissible control* $u(\cdot)\in L^\infty(T,L^2(V))$ *and a corresponding trajectory* $x(\cdot)\in AC(T,H_0^1(V))$ *with* $x(t)\in H^2(V)$ a.e. *s.t.* $J(x,u)=m_1$,

EXAMPLE #2. "Parabolic optimal control problem".

Let $T=[0,b]$ and let V be a bounded domain in $\boldsymbol{R}^n$ with smooth boundary $\partial V=\Gamma$. On $T\times V$ we consider the following parabolic optimal control problem:

$$\left.\begin{cases} J_2(x,u)=\int_0^b\int_V L(t,z,x(t,z),u(t,z))\,dz\,dt\to\inf=m_2 \\ \text{s.t. } \dfrac{\partial x(t,z)}{\partial t}+\sum_{|\alpha|\le m}(-1)^{|\alpha|}D^\alpha A_\alpha(z,\eta(x(z)))=\int_0^t K(t-s)f(s,z,x(s,z))\,u(s,z)\,ds \\ \quad D^\beta x(t,z)=0 \text{ on } T\times\Gamma \quad \text{for} \quad |\beta|\le m-1 \\ \quad x(0,z)=x_0(z) \text{ on } \{0\}\times V, \\ \quad |u(t,z)|\le M,\ u(\cdot,\cdot) \quad \text{is measurable} \end{cases}\right\}$$

$$(***)_2$$

Here $\alpha=(\alpha_1,\cdots,\alpha_n)$ is an n-tuple of nonneagtive integers (multi-index), $|\alpha|=a_1+\cdots+\alpha_n$ (the length of the multi-index), $D^\alpha=D_1^{\alpha_1}\cdots D_n^{\alpha_n}$ with $D_i=\partial/\partial z_i$ and $\eta(x)=\{D^\alpha x(\cdot): |\alpha|\le m\}$.

We will need the following hypotheses on the data of $(***)_1$:

$H(A)_1$: A_α: $V\times\boldsymbol{R}^{n_m}\to\boldsymbol{R}$ are functions s.t.

(1) $z\to A_\alpha(z,\eta)$ is measurable,

(2) $\eta\to A_\alpha(z,\eta)$ is continuous,

(3) $|A_\alpha(z,\eta)|\le c_1\|\eta\|+\phi_1(z)$, $c_1>0$, $\phi_1(\cdot)\in L^2(V)$,

(4) $\sum_{|\alpha|\le m}(A_\alpha(z,\eta)-A_\alpha(z,\eta'))(\eta_\alpha-\eta_\alpha')\ge c_2\sum_{|\alpha|\le m}|\eta_\alpha-\eta_\alpha'|^2$ a.e. with $c_2>0$

Note that $n_m=\dfrac{(n+m)!}{n!\,m!}$.

$H(L)_2$: L: $T\times V\times\boldsymbol{R}\to\boldsymbol{R}$ is a normal integrand, convex in u s.t. $\phi(t,z)\le L(t,z,x)$ a.e. with $\phi(\cdot,\cdot)\in L^1(T\times V)$.

Also hypotheses $H(K)_1$, $H(f)_1$ and H_{a_1} (for $J_2(\cdot, \cdot)$) from Example #2 are valid.

Consider the Dirichlet form $a\colon H_0^m(V)\times H_0^m(V)\to \boldsymbol{R}$ associated with the nonlinear elliptic differential operator of our problem. We have

$$a(x, y) = \sum_{|\alpha|\leq m}\int_V A_\alpha(z, \eta(x(z)))\, D^\alpha y(z)\, dz$$

for all $(x, y)\in H_0^1(V)\times H_0^1(V)$.

Because of hypothesis H(A) (3) and Krasnoselski's theorem, for every multi-index α with $|\alpha|\leq m$, we have $A_\alpha(\cdot, \eta(x(\cdot))\in L^2(V)$. So using the Cauchy-Schwartz and Minkowski inequalities, we get:

$$\begin{aligned}
&|\int_V A_\alpha(z, \eta(x(z)))\, D^\alpha y(z)\, dz|\\
&\leq (\int_V |A_\alpha(z, \eta(x(z)))|^2\, dz)^{1/2} (\int_V |D^\alpha y(z)|^2\, dz)^{1/2}\\
&\leq [c_1 \sum_{|\gamma|\leq m} (\int_V |D^\gamma x(z)|^2\, dz)^{1/2} + (\int_V \phi(z)^2\, dz)^{1/2}] (\int_V |D^\alpha y(z)|^2\, dz)^{1/2}\\
&\leq [c_1 ||x|||_{H_0^m(V)} + ||\phi||_2]\, ||y||_{H_0^m(V)}\,.
\end{aligned}$$

Since α was an arbitrary multi-index of length $\leq m$, we get

$$|a(x, y)| \leq (\hat{c}_1 ||x||_{H_0^m(V)} + \hat{c}\, ||\phi||_2)\, ||y||_{H_0^m(V)},\ \hat{c}_1, \hat{c}>0\,.$$

From this inequality we deduce that for every $x(\cdot)\in H_0^m(V)$, $a(x, \cdot)$ is a continuous, linear map from $H_0^m(V)$ into $H^{-m}(V)=H_0^m(V)^*$. Hence there exists a generally nonlinear operator $A\colon H_0^m(V)\to H^{-m}(V)$ s.t.

$$a(x, y) = (A(x), y)\,.$$

Also using hypothesis H(A), (4), it is easy to see that

$$(A(x_1)-A(x_2), x_1-x_2) \geq \hat{c}_2\, ||x_1-x_2||_{H_0^m(V)}\,, \quad \hat{c}_2>0$$

i.e. $A(\cdot)$ is strongly monotone.

Let $X=L^2(V)$ and let $\hat{A}\colon D(\hat{A})\subseteq X\to X$ be defined by

$$\hat{A}x = Ax$$

for all $x\in D(\hat{A})=\{y\in X\colon Ay\in X\}$. From Barbu [10], we know that $\hat{A}$ is maximal monotone (hence m-accretive). Also from the strong monotonicity property, we see that $\hat{A}$ is coercive and since $H_0^m(V)\hookrightarrow L^2(V)$ compactly, it is easy to see that $(I+A)^{-1}$ is compact.

Let $Y=L^2(V)$ (the control space) and define:

(i) $\hat{f}\colon T\times X\times Y\to X$ by $\hat{f}(t, x)\, u(\cdot) = f(t, \cdot, x(\cdot))\, u(\cdot)$,

(ii) $\hat{L}\colon T\times X\times Y\to \bar{\boldsymbol{R}}$ by $\hat{L}(t, x, u) = \int_V L(t, z, x(z), u(z))\, dz$,

(iii) $\hat{K}$: $T \to \mathcal{L}(Y)$ by $\hat{K}(t) = K(t)I$,
(iv) $U = \{u \in Y: |u(z)| \leq M \text{ a.e.}\}$,
(v) $\hat{x}_0 = x_0(\cdot) \in L^2(V)$.
Rewrite problem $(***)_2$ in the following abstract form:

$$\left.\begin{cases} \hat{J}_2(x, u) = \int_0^b \hat{L}(t, x(t), u(t))\, dt \to \inf = m_2 \\ \text{s.t. } \dot{x}(t) \in -\hat{A}x(t) + \int_0^t \hat{K}(t-s)\hat{f}(s, x(s))\, u(s)\, ds \text{ a.e.} \\ \quad x(0) = \hat{x}_0,\ u(t) \in U \text{ a.e. } u(\cdot) \text{ is measurable} \end{cases}\right\} (***)_2'$$

Again this is a special case of problem (*). Apply Theorem 3.1 to get the following existence result:

Theorem 6.2. *If hypotheses* $\mathrm{H(A)_1}$, $\mathrm{H(K)_1}$, $\mathrm{H(f)_1}$, $\mathrm{H(L)_2}$ and H_{a_1} *hold, then there exists an admissible control* $u(\cdot, \cdot) \in L^\infty(T, L^2(V))$ *and a corresponding trajectory* $x(\cdot, \cdot) \in AC(T, L^2(V))$ *s.t.* $J_2(x, u) = m_2$.

ACKNOWLEDGEMENT: The authors are deeply indepted to the anonymous referee for his careful reading of the manuscript and for his many corrections, remarks and constructive criticisms.

References

[1] R. Adams: Sobolev Spaces, Academic Press, New York, 1975.
[2] N.U. Ahmed: *Optimal control of a class of strongly nonlinear parabolic systems,* J. Math. Anal. Appl. **16** (1977), 188–207.
[3] N.U. Ahmed: *Properties of relaxed trajectories for a class of nonlinear evolution equations on a Banach space,* SIAM J. Control. Optim. **21** (1983), 953–967.
[4] N.U. Ahmed: *Existence of optimal controls for a class of systems governed by differential inclusions on a Banach space,* J. Optim. Theory Appl. **50** (1986), 213–237.
[5] N.U. Ahmed: *Existence of optimal relaxed controls for differential inclusions on Banach space in Nonlinear Analysis and Applications,* ed. V. Lakshmikantham, Proc. 7th Intern. Conf. on Nonl. Anal. and Appl., Lecture Notes in Pure and Applied Math, Vol 109, Marcel Dekker Inc, 1987, 39–49.
[6] N.U. Ahmed and K.L. Teo: Optimal Control of Distributed Parameter Systems, North Holland, New York, 1981.
[7] J.-P. Aubin and A. Cellina: Differential Inclusion, Springer, Berlin, 1984.
[8] E. Balder: *Lower semicontinuity of integral functionals with nonconvex integrands by relaxation-compactification,* SIAM J. Control Optim. **19** (1981), 533–541.
[9] E. Balder: *A general denseness result for relaxed control theory,* Bull. Austr. Math. Soc. **30** (1984), 463–475.
[10] V. Barbu: Nonlinear Semigroups and Differential Equations in Banach Space, Noordhoff International Publishing, Leyden, The Netherlands, 1976.
[11] P. Bénilan: *Solutions integrales d'équations d'évolution dans un espace de Banach,* C.R. Acad. Sci. Paris, t. **274** (1972), 47–50.

[12] N. Bourbaki: Eléments de Mathématique: Integration, Chapitre III Second edition, Hermann, Paris, 1965.

[13] H. Brézis: *New results concerning monotone operators and nonlinear semigroups,* Proceedings of R.I.M.S., Analysis of Nonlinear Problems, Kyoto University, 1975, 2–25.

[14] L. Cesari: *Existence theorems for multidimensional Lagrange problems,* J. Optim. Theory and Appl. **1** (1967), 87–112.

[15] L. Cesari: *Existence theorems for abstract multidimensional control problems,* J. Optim Theory and Appl. **6** (1970), 210–236.

[16] L. Cesari: *Existence theorems for problems of optimization with distributed and boundary controls,* Proc. Intern. Congress of Mathematicians, Nice (1970), Vol. 3 (1971), 157–161.

[17] L. Cesari: Optimization Theory and Applications, Springer, New York, 1983.

[18] G. Choquet: Lectures on Analysis, Vol. II, Benjamin, London, 1969.

[19] P.-V. Chuong: *Some results on density of extreme selections for measurable multifunctions,* Math. Nachrichten **125** (1986), 310–326.

[20] M. Day: *Some more uniformly convex spaces,* Bull. Amer. Math. Soc. **47** (1941), 504–517.

[21] J. Diestel and J.J. Uhl: *Vector Measures,* Math. Surveys, Vol. 15, Providence, R.I., 1977.

[22] N. Dunford and J. Schwartz: Linear Operators I, Wiley, New York, 1958.

[23] A. Egorov: *Optimal control in a Banach space,* Math. Systems Theory **1** (1962), 347–352.

[24] I. Ekeland: *Sur le contrôle optimale de systèmes gouvernés par des equations elliptiques,* J. Funct. Anal. **9** (1972), 1–62.

[25] I. Ekeland and R. Temam: Convex Analysis and Variational Problems, North Holland, Amsterdam, 1974.

[26] A. Friedman: *Optimal control for parabolic euqations,* J. Math. Anal. Appl. **18** (1967), 479–491.

[27] A. Friedman: *Optimal control in Banach spaces,* J. Math. Anal. Appl. **19** (1967), 35–55.

[28] F. Hiai and H. Umegaki: *Integrals, conditional expectations and martingales of multivalued functions,* J. Multiv. Anal. **7** (1977), 149–182.

[29] N. Hirano: *Local existence theorems for nonlinear differential equations,* SIAM J. Math. Anal. **14** (1983), 117–125.

[30] S.-H. Hou: *Existence theorems for optimal control problems in Banach spaces,* Nonl. Analysis—T.M.A. **7** (1983), 239–257.

[31] T. Kato: *Nonlinear semigroups and evolution equations,* J. Math. Soc. Japan **19** (1967), 508–520.

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

[33] J.-L. Lions: *Optimisation pour certaines classes d'équations d'évolution nonlinéaires,* Annali di Mat. Pura ed Appl. LXII (1966), 275–294.

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

[35] B. Pachpatte: *A note on Grownall-Bellmann inequality,* J. Math. Anal. Appl. **44** (1973), 758–762.

[36] N.S. Papageorgiou: *Random fixed point theorems for measurable multifunctions in Banach spaces,* Proc. Amer. Math. Soc. **97** (1987), 507–514.

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

[38] N.S. Papageorgiou: *Optimal control of nonlinear evolution inclusions,* J. Optimization Theory Appl. **67** (2) (1990)—in press.

[39] N.S. Papageorgiou: *Properties of the relaxed trajectories of evolution equations and optimal control,* **27** (1989), 267–288.

[40] N.S. Papageorgiou: *Existence of optimal controls for nonlinear distributed parameter systems,* Funkcialaj Ekvacioj **32** (1989), 429–437.

[41] R.T. Rockafellar: *Conjugate Duality and Optimization,* Conference Board in Mathematical Sciences Series, no 16, SIAM Publications, Philadelphia, 1974.

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

[43] J. Warga: Optimal Control of Differential and Functional Equations, Academic Press, New York, 1972.

Evgenios P. Avgerinos
University of Thessaloniki
School of Technology
Thessaloniki 54006, GREECE

Nikolaos S. Papageorgiou
University of California
1015 Department of Mathematics
Davis, California 95616, U.S.A.