JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 50, No. 2, AUGUST 1986

Existence of Optimal Controls for a Class of Systems Governed by Differential Inclusions on a Banach Space[1]

N. U. AHMED[2]

Communicated by L. Cesari

Abstract. Using Cesari's approach, we prove the existence of optimal controls for a class of systems governed by differential inclusions on a Banach space having the Radon–Nikodym property. Theorem 3.1 gives the existence result for optimal relaxed controls under fairly general assumptions on the system and the admissible controls. This result depends on a fundamental result (Theorem 2.1) that proves the existence of mild solutions of differential inclusions on a Banach space, which has also independent interest. Further, the preparatory results, such as Lemma 3.1 and Lemma 3.2, are also useful in the study of time-optimal and terminal control problems.

For illustration of the results, we present two examples, one on distributed controls for a class of systems governed by nonlinear parabolic equations and the other on boundary controls with discontinuous boundary operator.

Key Words. Differential inclusions, Banach spaces having Radon-Nikodym property, mild solutions, Polish spaces, relaxed controls, measure-valued controls, Cesari property, distributed controls, boundary controls.

1. Introduction

Let Γ be a compact subset of a Polish space B, and let $M = M(\Gamma)$ denote the space of bounded positive Radon measures, in particular probability measures, on the Borel σ-field of B, with support Γ. Suppose that M is given the usual w^*-topology, which is metrizable, making M itself into a Polish space. Let X be a Banach space; let I be the interval $[0, T]$, $T < \infty$;

[1] This work is supported in part by the National Science and Engineering Council of Canada under Grant No. 7109.

[2] Professor, Department of Electrical Engineering, University of Ottawa, Ottawa, Ontario, Canada.

0022-3239/86/0800-0213$05.00/0

and let f be a multivalued mapping defined on $I \times X \times M$, with values $f(t, x, \nu) \in 2^X$, the space of all nonempty subsets of X. More technical assumptions will be given in the sequel.

Let $\mathcal{M}$ denote the space of w^*-measurable M-valued functions defined on I, that is,

$$\mathcal{M} \equiv \Big\{\{\mu_t, t \in I\}: \mu_t \in M, \text{ a.e., and } t \to \mu_t(h) \equiv \int_\Gamma h(\sigma)\mu_t(d\sigma)$$

$$\text{is measurable for each } h \in C(\Gamma), \text{ the space of continuous functions on } \Gamma\Big\}.$$

Let $L: I \times X \times M \to R$ be given by

$$L(t, x, \nu) \equiv \int_\Gamma L(t, x, \sigma)\nu(d\sigma).$$

Let $A(t)$, $t \geq 0$, be a family of linear operators with domain $D(A(t)) \subset X$ and range $R(A(t)) \subset X$. Consider the system governed by the inclusion relation on X,

$$dx/dt + A(t)x(t) \in f(t, x(t), \mu_t), \tag{1a}$$

$$x(0) = x_0, \tag{1b}$$

with $\mu \in \mathcal{M}$ and $x_0 \in X$.

Define, for each $\mu \in \mathcal{M}$, the functional

$$J(\mu) \equiv \int_0^T L(t, x(t), \mu_t)\, dt, \tag{2}$$

with x being the solution of the evolution equation (1) corresponding to μ. We consider the question of the existence of optimal relaxed controls (control measures) $\mu^0 \in \mathcal{M}$, in the sense that

$$J(\mu^0) \equiv \int_0^T L(t, x^0(t), \mu_t^0)\, dt \leq J(\mu) \equiv \int_0^T L(t, x(t), \mu_t)\, dt, \tag{3}$$

for all pairs $\{\mu, x\}$ satisfying the evolution equation (inclusion) (1), where x^0 is also a solution of the system (1), corresponding to μ^0.

We wish to prove the existence of optimal controls under quite general conditions on the nonlinear operator f and the cost integrand L. This result is given in Theorem 3.1 after we have presented in Section 2 some basic results on the question of existence of solutions of differential inclusions on Banach space. This result, given in Theorem 2.1, has independent interest.

We have used upper semicontinuity properties, the Cesari property, measurable selection theorems, the Radon–Nikodym property, and the Kakutani–Fan fixed point theorem for the existence results. For earlier results in the area of existence of optimal controls for nonlinear evolution equations in Banach space, the reader is referred to Refs. 1–11. For illustration of our results, we have presented two examples, one on distributed controls of nonlinear parabolic equations and the other on boundary controls with a multivalued boundary operator.

A question of significant theoretical and practical interest is whether or not the optimal relaxed controls can be approximated by ordinary controls, that is, measurable functions defined on I with values in Γ. In a recent paper (Ref. 1), it was shown that, for a very general class of nonlinear evolution equations on a Banach space, this approximation property holds. In the present context, this remains an open question.

Basic Notations. For any Banach space X, let $\|x\|_X$ denote the norm of the element x of X and let X^* denote its (topological) dual. Let $\mathscr{L}(X)$ represent the space of bounded linear operators in X, with $\|B\|_{\mathscr{L}(X)}$ denoting the norm of the element B of $\mathscr{L}(X)$. For a topological vector space Z, we use 2^Z to denote the set of all nonempty subsets of Z and $\mathrm{CC}(Z)$ to denote the class of closed convex subsets of Z. If $F \in 2^Z$, its closure is denoted by cl F and its closed convex hull is denoted by clco F. If Z is also a Banach space, $\mathrm{WC}(Z)$ will represent the class of weakly compact subsets of Z and $\mathrm{WCC}(Z)$ the class of weakly compact and convex subsets of Z. For any bounded interval $I \subset R$, and X a Banach space, $C(I, X)$ denotes the vector space of continuous functions on I with values in X. Furnished with the norm topology given by

$$\|f\|_{C(I,X)} \equiv \sup\{\|f(t)\|_X,\ t \in I\},$$

$C(I, X)$ is a Banach space. Similarly, for any p, $1 \le p < \infty$, $L_p(I, X)$ denotes the space of equivalence classes of strongly measurable X-valued Bochner integrable functions $\{f\}$, so that

$$\|f\|_{L_p(I,X)} \equiv \left(\int_I \|f(t)\|_X^p\right)^{1/p} < \infty.$$

Furnished with this norm topology, $L_p(I, X)$ is also a Banach space. This is also true for $p = \infty$ with the norm

$$\|f\|_{L_\infty(I,X)} \equiv \text{ess. sup}\{\|f(t)\|_X,\ t \in I\}.$$

$L_p^{\text{loc}}([0, \infty); X)$ is a locally convex complete topological vector space whose elements when restricted to finite intervals $J \subset [0, \infty)$ belong to $L_p(J, X)$. If X is a Banach space with its dual X^* satisfying the Radon–Nikodym property (Ref. 12) and if the Lebesgue measure of I is finite, then the dual of $L_p(I, X)$ is given by $L_q(I, X^*)$, provided

$$(1/p)+(1/q)=1 \quad \text{and} \quad 1 \leq p < \infty.$$

In particular, $Y^* = L_\infty(I, X^*)$, whenever $Y = L_1(I, X)$. If X is a reflexive Banach space and $1 < p < \infty$, then $L_p(I, X)$ is also reflexive. For any open bounded connected subset Ω of R^n, any nonnegative integer m and $1 \leq p < \infty$, $W^{m,p}(\Omega) \subset L_p(\Omega)$ denotes the standard Sobolev spaces, with $W_0^{m,p}$ denoting the closure of $C_0^\infty(\Omega)$ in the topology of $W^{m,p}$. The dual of $W^{m,p}$ is denoted by $(W^{m,p})^*$ and that of $W_0^{m,p}$ is given by $W^{-m,q}$ for $(1/p)+(1/q)=1$, where the elements of $W^{-m,q}$ are genuine distributions on Ω. In fact, the Sovolev spaces $W^{s,p}$ are defined for all $s \in R$. For $p=2$, we use H^s for $W^{s,2}$. For a nonempty set E of a normed space Z, we define

$$\|E\|^0 \equiv \sup\{\|\xi\|_Z : \xi \in E\} \subset [0, \infty) \cup \{\infty\}.$$

Let F denote a measurable set-valued function (or multifunction), defined on I and taking values $F(t) \in 2^Z$. By the integral of the function F, we mean

$$\int_I F(t)\, dt \equiv \left\{ \int_I f(t)\, dt, f \text{ a measurable selection of } F \right\}.$$

The multifunction F is said to be pth power integrably bounded, $1 \leq p \leq \infty$, if there exists a nonnegative function $g \in L_p(I) \equiv L_p(I, R)$, such that

$$\|F(t)\|^0 \equiv \sup\{\|y\|, y \in F(t)\} \leq g(t), \qquad \text{a.e.}$$

A multifunction H mapping Z_1 to 2^{Z_2}, where Z_1 and Z_2 are any two topological vector spaces with topologies τ_1 and τ_2, respectively, is said to be uppersemicontinuous with respect to inclusion (USC) at $\xi_0 \in Z_1$ if, for every τ_2-neighborhood $N_2(H(\xi_0))$ of $H(\xi_0)$, there exists a τ_1-neighborhood $N_1(\xi_0)$ of ξ_0, such that

$$H(\xi) \subset N_2(H(\xi_0)), \qquad \text{for all } \xi \in N_1(\xi_0).$$

If this is true for all $\xi \in K \subset Z_1$, then H is said to be USC on K. For further details on multifunctions, the reader is referred to Refs. 2 and 9 and the references therein.

We have used the notation $\xi^n \xrightarrow{w(s)} \xi^0$ to indicate that ξ^n converges to ξ^0 weakly (strongly) as $n \to \infty$. For any Banach space E with dual E^*, the duality pairing is denoted by $(f, g)_{E^*,E}$ or $(g, f)_{E,E^*}$ for $f \in E^*$ and $g \in E$. For any set F, co F denotes the convex hull of F and clco F stands for closure of co F.

2. Existence of Solutions of the Differential Inclusion

In this section, we consider the question of the existence of solutions of the differential inclusion (1). For convenience of reference, we quote some standard results on the existence of mild solutions of the evolution equation

$$dx/dt = -A(t)x + g(t), \qquad t \in I, \tag{4a}$$

$$x(0) = x_0, \tag{4b}$$

where $x_0 \in X$ and $g \in L_1(I, X)$ (Ref. 13, p. 108 and Ref. 2, p. 100). First, we consider the Cauchy problem

$$dx/dt + A(t)x = 0, \qquad x(0) = x_0, \qquad t \geq 0,$$

in the Banach space X.

Lemma 2.1. Suppose that the operator A satisfies the following assumptions:

(ai) the domain $\mathrm{D}(A) = \mathrm{D}(A(t))$ is independent of t and is dense in X; and, for each t, $A(t)$ is a closed operator;

(aii) there exists a constant C independent of t such that, for each $t \in I$, the resolvent $\mathrm{E}(\lambda, A(t)) \equiv (\lambda I + A(t))^{-1}$ of $A(t)$ exists for all λ with real $\lambda \leq 0$ and

$$\|\mathrm{R}(\lambda, A(t))\|_{\mathscr{L}(X)} \geq C/(1+|\lambda|), \qquad \mathrm{Re}\,\lambda \leq 0;$$

(aiii) there exists an $\alpha \in (0, 1)$ such that, for all t, τ, $s \in J$,

$$\|(A(t) - A(\tau))A^{-1}(s)\|_{\mathscr{L}(X)} \geq c|t-\tau|^{\alpha}.$$

Then, there exists an evolution operator $U(t, \tau)| \in \mathscr{L}(X)$, $0 \leq \tau \leq t \leq T$, satisfying the following conditions:

(ci) $$(\partial/\partial t)U(t, \tau) + A(t)U(t, \tau) = 0, \qquad \tau < t \leq T, \tag{5a}$$

$$U(\tau, \tau) = I \text{ (identity in } \mathscr{L}(X)), \qquad \tau \geq 0; \tag{5b}$$

(cii) $U(t, \tau)$ is strongly continuous in t, τ, for $0 \leq \tau \leq t \leq T$;

(ciii) $(\partial/\partial t)U(t, \tau)$ exists in the strong topology and belongs to $\mathscr{L}(X)$ for $0 \leq \tau < t \leq T$ and is also strongly continuous in t for $t > \tau$;

(civ) there exists a $C = C(T) > 0$ such that, for any $\rho \in [0, 1]$,

$$\|A^{\rho}(t)U(t, \tau)\|_{\mathscr{L}(X)} \leq C/(t-\tau)^{\rho}, \qquad \text{for } 0 \leq \tau < t \leq T;$$

(cv) for each $x_0 \in X$, the equation

$$dx/dt + A(t)x = 0, \qquad x(0) = x_0,$$

has a unique solution $x \in \mathrm{C}(I, X)$, with $\dot{x} \in \mathrm{C}((0, T], X)$, given by

$$x(t) = U(t, 0)x_0, \qquad t \in I.$$

Remark 2.1. We must emphasize that the evolution operator U exists under milder hypothesis (Ref. 14, Theorem 3.1, p. 116).

Under an additional hypothesis, one can prove the existence of a mild solution of the nonhomogeneous Cauchy problem (4). This is given by the following lemma.

Lemma 2.2. Suppose that the operator A satisfies the hypotheses of Lemma 2.1 and $g \in L_1(I, X)$ and $x_0 \in X$. Then, the evolution equation (4) has a unique mild solution $x \in C(I, X)$, given by

$$x(t) = U(t, 0)x_0 + \int_0^t U(t, \theta)g(\theta)\, d\theta. \tag{6}$$

Remark 2.2. In fact, the mild solution (6) is defined for all $t \geq 0$ and for each $g \in L_p^{\text{loc}}([0, \infty); X)$ for all $1 \leq p \leq \infty$ (Ref. 2, p. 100).

For each $t \in I$ and $x_0 \in X$, define the operator S_t as

$$S_t g \equiv U(t, 0)x_0 + \int_0^t U(t, \tau)g(\tau)\, d\tau. \tag{7}$$

Clearly, S_t is an affine map and takes $L_1([0, t], X)$ into X for each $t \in I$, and $t \to S_t g$ is a continuous X-valued function on any compact interval of $[0, \infty)$.

We need one more result before we can prove the existence of solutions of the differential inclusion (1). Let Y denote $L_1(I, X)$; and let $CC(X)$ and $CC(Y)$ denote the class of closed convex subsets of X and Y, respectively. Define, for each given $\mu \in \mathcal{M}$, the multifunction $f_\mu(t, x) \equiv f(t, x, \mu_t)$, and write Eq. (1) as

$$\dot{x} \in -A(t)x + f_\mu(t, x), \qquad t \geq 0.$$

Let $\rho \in (0, 1)$, and suppose that $D_\rho \equiv D(A^\rho(t))$ is independent of $t \in I$.

Lemma 2.3. Let X be a Banach space, with its dual X^* satisfying the Radon-Nikodym property (RNP), and suppose that the multifunction f satisfies the following properties:

(ai) $f(t, x, \nu) \in CC(X)$ for all $(t, x, \nu) \in I \times D_\rho \times M$;

(aii) $t \to f(t, x, \nu)$ is a measurable multifunction on I for each $x \in D_\rho$ and $\nu \in M$;

(aiii) $x \to f(t, x, \nu)$ is upper semicontinuous (USC) with respect to inclusion on D_ρ, uniformly in $I \times M$;

(aiv) $\nu \to f(t, x, \nu)$ is upper semicontinuous (in the sense of inclusion) with respect to the w^*-topology on M for each fixed $t \in I$ and $x \in D_\rho$;

(av) there exists a constant $k \in [0, \infty)$, such that, for each $x \in D_\rho$,

$$\|f(t, x, \nu)\|^0 \equiv \sup\{\|y\|_X : y \in f(t, x, \nu)\} \leq k(1 + \|A^\rho(t)x\|).$$

Then, for each $\mu \in \mathcal{M}$ and $g \in L_1(I, X)$, so that $S_t g \in D_\rho$, we have:

(ci) the mapping $t \to F_\mu(t) \equiv f_\mu(t, S_t g)$ is a measurable and integrably bounded set-valued function on I with values in $\mathrm{CC}(X)$;

(cii) $H_\mu(g) \equiv \{h \in Y: h(t) \in f_\mu(t, S_t g) \text{a.e.}\} \in \mathrm{CC}(Y)$;

(ciii) $g \to H_\mu(g)$ is USC (upper semicontinuous with respect to inclusion) on Y to 2^Y.

Proof. Since, for each $g \in Y \equiv L_1(I, X)$, $t \to S_t g$ is continuous, and μ is a w^*-measurable M-valued function, the measurability of the set-valued map $t \to F_\mu(t)$ follows from the assumptions (aii), (aiii), (aiv). We show that it is integrably bounded. Indeed, it follows from (civ) of Lemma 2.1 that

$$\|A^\rho(t) S_t g\|_X \leq (C\|x_0\|/t^\rho) + C \int_0^t (1/(t-\tau)^\rho)\|g(\tau)\| \, d\tau; \tag{8}$$

and hence, one can easily verify that

$$\int_0^T \|A^\rho(t) S_t g\|_X \, dt < \infty, \qquad \text{for any } g \in Y.$$

Therefore, the measurable set-valued function F_μ is integrably bounded since, by assumption (av),

$$\|F_\mu(t)\|^0 = \|f_\mu(t, S_t g)\|^0 \leq K(1 + \|A^\rho(t) S_t g\|), \qquad \text{a.e.} \tag{9}$$

In fact, using (8) in (9), one can show that there exists a constant c_1 independent of g such that

$$\int_0^T \|F_\mu(t)\|^0 \, dt \leq c_1(1 + \|g\|_Y). \tag{10}$$

Closure and convexity of $F_\mu(t)$, $t \in I$, is a direct consequence of our assumption (ai). This proves (ci). We now prove (cii). Convexity of $H_\mu(g)$ is immediate. For its closure, let $\{h_n\} \in H_\mu(g)$ and suppose that $h_n \xrightarrow{w} h_0$ in $Y \equiv L_1(I, X)$. Note that the weak topology on $L_1(I, X)$ is generated by elements of $L_\infty(I, X^*)$, since, by the Radon–Nikodym property of X^*, $Y^* = L_\infty(I, X^*)$. Since $h_n \in H_\mu(g)$, $h_n(t) \in F_\mu(t)$, a.e.; consequently, by the integrability of the set-valued map F_μ, we have

$$[1/l(J)] \int_J h_n(t) \, dt \in [1/l(J)] \int_J F_\mu(t) \, dt, \qquad \text{for all } n, \tag{11}$$

and for any Lebesgue measurable set $J \subset I$ with Lebesgue measure $l(J)$.

By the integral of the set-valued map F_μ, we mean, as usual, the integrals of measurable selections of F_μ. Since $F_\mu(t) \in \mathrm{CC}(X)$, for all $t \in I$, it is clear that $[1/l(J)] \int_J F_\mu(t)\, dt$ is convex. Therefore, by Mazur's theorem, its strong closure coincides with its weak closure; hence,

$$w-\lim_n [1/l(J)] \int_J h_n(t)\, dt$$

$$= [1/l(J)] \int_J h_0(t)\, dt \in [1/l(J)]\, \mathrm{cl} \int_J F_\mu(t)\, dt. \tag{12}$$

This is true for any measurable set $J \subset I$; consequently, $h_0(t) \in \mathrm{clco}\, F_\mu(t)$, a.e. Since $F_\mu(t)$ itself is closed convex, $h_0(t) \in F_\mu(t)$, a.e.; hence, $h_0 \in H_\mu(g)$, proving its closure. Thus, $H_\mu(g) \in \mathrm{CC}(Y)$, for each $g \in Y$. (ciii) follows from the facts that $x \to f(t, x, \nu)$ is USC on D_ρ, uniformly in $t \in I$ and $\nu \in M$, and that S_t is an affine continuous map from $L_1(0, t; X)$ to X with $S_t g \in D_\rho$. □

We now present our main result of this section. We need the following definition.

Definition 2.1. For each $x_0 \in X$ and $\mu \in \mathcal{M}$, an element $x_\mu \in C(I, X)$ is said to be a mild solution of the differential inclusion (1) if there exists a $g^\mu \in Y \equiv L_1(I, X)$ such that $g^\mu(t) \in f_\mu(t, x_\mu(t))$, a.e., and x_μ is a mild solution of the evolution equation

$$\dot{x} = -A(t)x + g^\mu, \qquad x(0) = x_0.$$

Theorem 2.1. Suppose that the hypotheses of Lemma 2.1 and Lemma 2.3 hold, and let $x_0 \in X$ be given. Suppose that, for each $\mu \in \mathcal{M}$, there exists a nonempty set $K_\mu \in \mathrm{WCC}(Y)$, the class of closed convex and weakly sequentially compact subsets of Y, such that $K_\mu \cap H_\mu(g) \neq \varnothing$, for all $g \in K_\mu$, where

$$H_\mu(g) \equiv \{h \in Y: h(t) \in f_\mu(t, S_t g), \text{ a.e.}\}.$$

Then, the system (1) has a mild solution for each $\mu \in \mathcal{M}$.

Proof. Let $x_0 \in X$ and $\mu \in \mathcal{M}$ be given. Then, by virtue of (cii) and (ciii) of Lemma 2.3, for each $g \in Y$, $H_\mu(g) \in \mathrm{CC}(Y)$ and $g \to H_\mu(g)$ is USC from Y to 2^Y. Further, under the assumption of the theorem, there exists a $K_\mu \in \mathrm{WCC}(Y)$ such that $H_\mu(g) \cap K_\mu \neq \varnothing$, for all $g \in K_\mu$. Hence, it follows from a version of the Kakutani-Fan fixed-point theorem (Ref. 15, Corollary 2 of Theorem 6.3, p. 75) that H_μ has at least one fixed point in K_μ. Let

$g^\mu \in K_\mu$ be any such fixed point, i.e., $g^\mu \in H_\mu(g^\mu)$. Then, clearly, $g_\mu(t) \in f_\mu(t, S_t g^\mu)$, a.e. Define x^μ with values $x^\mu(t) = S_t g^\mu$, $t \in I$. Clearly, $x^\mu \in C(I, X)$, with $x^\mu(0) = x_0$; and, by Lemma 2.2, it is the unique mild soiution of the evolution equation

$$\dot{x} = -A(t)x + g^\mu(t), \qquad t \geq 0.$$

Since $x^\mu(0) = x_0$ and $g^\mu(t) \in f_\mu(t, x^\mu(t))$, a.e., it follows from Definition 2.1 that x^μ is a mild solution of the differential inclusion

$$\dot{x} \in -A(t)x + f_\mu(t, x), \qquad x(0) = x_0, \qquad t \in I.$$

This completes the proof. □

It is evident that differential inclusions need not have unique solutions. Hence, the following result has meaning.

Corollary 2.1. Let $x_0 \in X$ and $\mu \in \mathcal{M}$ be fixed, and suppose that the assumptions of Theorem 2.1 hold. Then, the set of all solutions of the differential inclusion (1) corresponding to the given x_0 and μ, denoted C_{μ,x_0}, is a closed subset of $C(I, X)$.

Proof. Let

$$G_{\mu,x_0} \equiv \{g \in K_\mu : g \in H_\mu(g)\}$$

denote the set of all fixed points of H_μ. This is a closed set. Indeed, let $\{g^n\}$ be any sequence in G_{μ,x_0} converging weakly (in Y) to g^0. Clearly, $g^0 \in K_\mu$; and, by USC (upper semicontinuity) of H_μ, we have $H_\mu(g^n) \subset H_\mu^\epsilon(g^0)$ [ϵ-neighborhood of $H_\mu(g^0)$ in Y], for all n sufficiently large. Hence, $g^0 \in H_\mu^\epsilon(g^0)$, for any $\epsilon > 0$. Since $H_\mu(g^0)$ is a closed convex set, we have $g^0 \in H_\mu(g^0)$, and hence $g^0 \in G_{\mu,x_0}$. By Theorem 2.1, the set of mild solutions of the system (1) corresponding to μ and x_0 is given by

$$C_{\mu,x_0} \equiv \{x \in C(I, X) : x(t) = S_t g,\ t \in I,\ g \in G_{\mu,x_0}\}. \tag{13}$$

Being the image of a closed set under affine continuous mapping, C_{μ,x_0} is a closed subset of $C(I, X)$. □

The following remark gives an elementary sufficient condition for the existence of a weakly compact convex set K_μ of Y satisfying the hypothesis of Theorem 2.1.

Remark 2.3. Suppose that X is a reflexive Banach space, $\rho \in [0, 1)$, and $1 < p < (1/\rho)$; and take $L_p(I, X)$ for Y. Clearly, Y is a reflexive Banach space. For this choice of ρ and p, it is readily verified, using inequality (8), that $t \to A^\rho(t) S_t g$ belongs to $L_p(I, X)$ for each $g \in L_p(I, X)$. Hence, it follows from inequality (9) that $H_\mu(g) \subset L_p(I, X) = Y$ and that there exists a constant C_2 such that

$$\|H_\mu(g)\|^0 = \sup\{\|h\|_Y: h \in H_\mu(g)\} \leq C_2(1 + \|g\|_Y).$$

Under this situation, a sufficient condition for the existence of a weakly compact convex set $K_\mu \subset Y$ such that $K_\mu \cap H_\mu(g) \neq \emptyset$ is that there exists a finite number $r_0 = r_0(\mu) > 0$ and a number $\alpha = \alpha(\mu) \in [0, 1]$ such that

$$\|H_\mu(g)\|^0 \leq \alpha \|g\|_Y, \qquad \text{for all } g \in Y, \text{ with } \|g\|_Y > r_0.$$

Letting B_r denote the closed ball of radius r in Y, it is clear that

$$\sup\{\|H_\mu(g)\|^0, g \in B_{r_0}\} \triangleq r_1 \leq C_2(1 + r_0).$$

Hence, for $\tilde{r} \equiv \max\{r_0, r_1\}$, it follows that $H_\mu(g) \subset B_{\tilde{r}}$, for all $g \in B_{\tilde{r}}$. Since $B_{\tilde{r}}$ is weakly compact, the existence of K_μ follows.

Remark 2.4. In general, differential inclusions do not have unique solutions. However, if the multivalued operator $-f_\mu$ is hypermaximal accretive, and if both X and X^* are uniformly convex, then uniqueness may hold (Ref. 15, Theorem 9.23, p. 152).

3. Existence of Optimal Controls

In this section, we present an existence theorem for optimal relaxed controls. Define, for each $t \in I_0 \equiv (0, T]$ and $x \in D_\rho$, the set-valued map Q with values

$$Q(t, x) \equiv \{(\lambda, y) \in R \times X: \lambda \geq L(t, x, \nu), y \in f(t, x, \nu), \text{ for some } \nu \in M(\Gamma)\}. \tag{14}$$

For any subset $N \subset D_\rho$, define

$$Q(t, N) = \bigcup \{Q(t, \xi), \xi \in N\};$$

and, for any $x^* \in D_\rho$, let $N_\epsilon(x^*)$ denote the intersection of D_ρ with ϵ-neighborhood of x^*. Q is said to satisfy the weak Cesari property at x^* if

$$\bigcap_{\epsilon>0} \text{clco } Q(t, N_\epsilon(x^*)) \subset Q(t, x^*), \qquad \text{for } t \in I_0. \tag{15}$$

Theorem 3.1. Consider the optimal control problem (1)-(3), and suppose that the following assumptions hold in addition to the hypotheses of Theorem 2.1:

(ai) both X and its dual X^* satisfy the Radon-Nikodym property (RNP);

(aii) for each ball $B_r \subset X$, the set-valued map F_r with values

$$F_r(t) \equiv \{y \in X: y \in f(t, x, \nu), \text{ for some } x \in E_r \equiv B_r \cap D_p \text{ and } \nu \in M(\Gamma)\} \equiv f(t, E_r, M)$$

is measurable; and, for each Lebesgue measurable set $J \subset I$, the set

$$\left\{\int_J h(t)\, dt: h \in Y \text{ and } h(t) \in F_r(t), \text{a.e.}\right\}$$

is a relatively weakly sequentially compact subset of X;

(aiii) for each ball $B_r \subset X$, there exists an $h \in L_1(I, R)$, possibly dependent on B_r, such that $L(t, x, \nu) \geq h(t)$, a.e. for all $x \in B_r$ and $\nu \in M$, and such that L is measurable in t, for each $\{x, \nu\} \in X \times M$, and, for almost all $t \in I$, it is continuous in x on B_r and continuous in ν on M (with respect to its w^*-topology);

(aiv) the set-valued map Q satisfies the weak Cesari property on $I_0 \times (D_p \cap B_r)$ for every finite $r > 0$.

Then, there exists an optimal control for the problem (1)-(3).

For the proof of this result, we shall use the following intermediate results.

Lemma 3.1. Let $\mathscr{X}$ denote the set of all solutions of the system (1) corresponding to the admissible (relaxed) controls $\mathscr{M}$. Then, $\mathscr{X}$ is a bounded subset of $C(I, X)$.

Proof. By Theorem 2.1, for each $\mu \in \mathscr{M}$, there exists at least one $g^\mu \in Y$ such that $g^\mu(t) \in f(t, S_t g^\mu, \mu_t)$, a.e., where $t \to S_t g^\mu$ is the mild solution of the evolution equation

$$\dot{x} = -A(t)x + g^\mu, \qquad x(0) = x_0.$$

Clearly, by assumption (av) of Lemma 2.3,

$$\|g^\mu(t)\|_X \leq k(1 + \|A^\rho(t) S_t g^\mu\|_X), \qquad \text{a.e., for all } \mu \in \mathscr{M}. \tag{16}$$

It follows from (16) and (8) that there exists a constant C_1 independent of μ such that the function ϕ_μ, with values $\phi_\mu(t) \equiv \|A^\rho(t) S_t g^\mu\|_X$, satisfies the inequality

$$\phi_\mu(t) \leq C_1/t^\rho + C_1 \int_0^t 1/(t-\theta)^\rho \phi_\mu(\theta)\, d\theta, \tag{17}$$

for $t \in I_0 = (0, T]$. From this, it follows that there exists a constant C_2, independent of μ, such that

$$\phi_\mu(t) \leq C_2/t^\rho, \qquad \text{for all } t \in (0, T] \text{ and } \mu \in \mathcal{M}. \tag{18}$$

Using this estimate in (16) and integrating over I, one can verify that there exists a constant $C_3 < \infty$, independent of μ, such that

$$\sup\left\{\int_0^T \|g^\mu(t)\|_X \, dt, \mu \in \mathcal{M}\right\} \leq C_3. \tag{19}$$

Clearly, by virtue of (19), the set

$$G \equiv \{g^\mu \in Y: g^\mu(t) \in f(t, S_t g^\mu, \mu_t), \text{ a.e. for some } \mu \in \mathcal{M}\} \tag{20}$$

is a bounded subset of $Y \equiv L_1(I, X)$. Since S_t is an affine continuous map from $L_1(0, t; X)$ to X, and since $t \to S_t g$ is a continuous X-valued function on I for each $g \in Y$, and G is bounded, the result follows. □

Remark 3.1. It is clear from (18) and the definition of ϕ_μ that

$$S_t g^\mu \in D_\rho = D(A^\rho(t)), \qquad \text{for all } t \in (0, T] \text{ and for all } \mu \in \mathcal{M}.$$

Therefore, the attainable set $S_t(G) \subset D_\rho$, $0 < t \leq T$; and also it follows from the above lemma that there exists a ball

$$B_{r_0} = \{y \in X: \|y\| < r_0\}$$

of finite radius r_0 such that $S_t(G) \subset B_{r_0}$ for all $t \in I = [0, T]$.

Remark 3.2. Note that, as a consequence of the above results, the ball B_r in assumptions (aii)–(aiv) can be taken as B_{r_0}.

Lemma 3.2. Under the assumptions of Theorem 3.1, the set G given by (20) is a relatively weakly sequentially compact subset of $Y = L_1(I, X)$.

Proof. Since both X and X^* satisfy the RNP, by a theorem due to Dunford (Ref. 12, Theorem 1, p. 101), it suffices to show that (i) G is a bounded subset of $L_1(I, X)$, (ii) G is uniformly integrable, and (iii) for each measurable set $J \subset I$, the set $\{\int_J g(t)\, dt, g \in G\}$ is a relatively weakly compact subset of X. Boundedness of G follows from (19) as in Lemma 3.1. Uniform integrability follows from (16), (18), and the definition of ϕ_μ. Indeed, for any measurable set $J \subset I$,

$$\int_J \|g^\mu(t)\|_X \, dt \leq Kl(J) + KC_2 \int_J (1/t^\rho) \, dt, \tag{21}$$

for all $\mu \in \mathcal{M}$, with K, C_2 independent of μ. Consequently,

$$\lim_{l(J) \to 0} \int_J \|g(t)\|_X \, dt = 0,$$

uniformly with respect to $g \in G$. For (iii), we take any Lebesgue measurable set $\sigma \subset I$ and define

$$G(\sigma) = \left\{ \int_\sigma h(t)\, dt,\, h \in G \right\}.$$

Since G is a bounded subset of Y and $S_t(G) \subset B_{r_0} \subset X$ for all $t \in I$ (see Remark 3.1), we take $E_{r_0} = B_{r_0} \cap D_\rho$ [see (aii) of Theorem 3.1] and define the set-valued map F_{r_0} with values

$$F_{r_0}(t) = \bigcup \{ f(t, x, \nu) : x \in E_{r_0},\, \nu \in M \},$$

which, by (aii), is measurable. Define

$$H(\sigma) = \left\{ \int_\sigma h(t)\, dt : h \in Y \text{ and } h(t) \in F_{r_0}(t), \text{ a.e.} \right\},$$

and note that, by assumption (aii) of Theorem 3.1, this is a relatively weakly (sequentially) compact subset of X. Since $G(\sigma) \subset H(\sigma)$ for any measurable set $\sigma \subset I$, it follows that $G(\sigma)$ is also a relatively weakly compact subset of X. This concludes the proof. □

Remark 3.3. In case X is a reflexive Banach space, condition (iii) in the above lemma can be omitted; in that case, assumption (aii) of Theorem 3.1 is also not required. In other words, for reflexive Banach spaces, the sufficient condition for weak sequential compactness of a subset $K \subset L_1(I, X)$ is precisely the same as for finite-dimensional spaces X.

With this preparation, we can now prove Theorem 3.1. For $\mu \in \mathcal{M}$ and $g \in Y$, define

$$\eta(\mu, g) \equiv \int_0^T L(t, S_t g, \mu_t)\, dt, \tag{22}$$

where S_t is the operator as defined in Section 2 [see Eq. (7)]. Clearly, η is an extended real-valued function, defined on $\mathcal{M} \times Y$. Define

$$\mathcal{A} \equiv \{(\mu, g) \in \mathcal{M} \times Y : 0 \in H_\mu(g) - g\} \subset \mathcal{M} \times G, \tag{23}$$

where $H_\mu(g)$ is as given in Lemma 2.3 (cii). Note that the cost functional $J(\mu)$ is the restriction of η to $\mathcal{A}$. Therefore, for the existence of an optimal control in $\mathcal{M}$, it suffices to prove the existence of a pair $(\mu^0, g^0) \in \mathcal{A}$ such that

$$\eta(\mu^0, g^0) \le \eta(\mu, g), \qquad \text{for all } (\mu, g) \in \mathcal{A}.$$

Proof of Theorem 3.1. Let $(\mu^n, g^n) \in \mathcal{A}$ be a minimizing sequence for the functional η restricted to $\mathcal{A}$. That is,

$$\lim_n \eta(\mu^n, g^n) = j_0 \equiv \inf\{\eta(\mu, g), (\mu, g) \in \mathcal{A}\}. \tag{24}$$

If $j_0 = \infty$, there is nothing to prove; so, we assume $j_0 < \infty$. Since $\{g^n\} \in G$ and since, by Lemma 3.2, G is a relatively weakly sequentially compact subset of Y, there exists a subsequence of the sequence $\{g^n\}$, relabeled as $\{g^n\}$, and an element $g^0 \in Y$ such that $g^n \overset{w}{\to} g^0$. Then, by Mazur's theorem, there exists a finite convex combination of $\{g^n\}$ that converges strongly to g^0. In particular, for each integer j, there exists an integer n_j, a set of integers $i = 1, 2, \ldots, m(j)$, and a set of nonnegative numbers $\{\alpha_{ji}, i = 1, 2, \ldots, m(j)\}$, with

$$\sum_{i=1}^{m(j)} \alpha_{ji} = 1, \qquad \text{for all } j,$$

such that

$$\psi_j(t) \equiv \sum_{i=1}^{m(j)} \alpha_{ji} g^{n_j+i} \overset{s}{\to} g^0, \qquad \text{in } Y. \tag{25}$$

Corresponding to the above sequence, define

$$l_{n_j+i}(t) \equiv L(t, S_t g^{n_j+i}, \mu_t^{n_j+i}), \qquad t \in I, \tag{26}$$

$$\lambda_j(t) \equiv \sum_{i=1}^{m(j)} \alpha_{ji} l_{n_j+i}(t), \qquad t \in I, \tag{27}$$

where $t \to S_t g^{n_j+i}$ is a mild solution of the differential inclusion (1) corresponding to the control μ^{n_j+i}. Define

$$\lambda_0(t) \equiv \varliminf_j \lambda_j(t). \tag{28}$$

It follows from Lemma 3.1 and the following remark (Remark 3.1) that

$$\{S_t g^{n_j+i}, i = 1, 2, \ldots, m(j), j = 1, 2, \ldots\} \subset B_{r_0} \subset X,$$

for sufficiently large r_0. Combining this with the fact that $\mu_t^n \in M(\Gamma)$, a.e., it follows from our assumption (aiii) (see Theorem 3.1) that there exists an $h \in L_1(I, R)$ such that $\lambda_j(t) \geq h(t)$, a.e.; consequently, $\varliminf \lambda_j(t)$ is defined a.e. Therefore, by Fatou's lemma,

$$\int_I \lambda_0(t)\, dt = \int_I \varliminf_j \lambda_j(t)\, dt \leq \varliminf_j \int_I \lambda_j(t)\, dt. \tag{29}$$

Clearly, by virtue of (24),

$$\lim_{j \to \infty} \eta(\mu^{n_j+i}, g^{n_j+i}) = j_0,$$

and hence also

$$\lim_{j \to \infty} \sum_{i=1}^{m(j)} \alpha_{ji} \eta(\mu^{n_j+i}, g^{n_j+i}) = j_0. \tag{30}$$

Thus, it follows from (27), (29), (30) that

$$\int_I \lambda_0(t)\, dt \leq j_0; \tag{31}$$

and, since $\lambda_0(t) \geq h(t)$, a.e., and $h \in L_1(I, R)$, it is clear that $\lambda_0 \in L_1(I, R)$. Now, we show that

$$(\lambda_0(t), g^0(t)) \in Q(t, S_t g^0), \qquad \text{a.e.}$$

Define

$$I_1 \equiv \{t \in I: |\lambda_0(t)| < \infty\} \cap \{t \in I: \lim_j \|\psi_j(t) - g^0(t)\|_X = 0\},$$

$$N_k \equiv \{t \in I: \mu_t^k \notin M(\Gamma)\}, \qquad N_0 = \bigcup_k N_k,$$

and set

$$I_2 = I \backslash N_0, \qquad I_3 = I_1 \cap I_2.$$

Due to (25) and the fact that $\lambda_0 \in L_1$, it is clear that the Lebesgue measure $l(I \backslash I_1) = 0$ and, by our definition of admissible controls, $l(I \backslash I_2) = 0$ also. Thus, I_3 has full Lebesgue measure. Therefore, for $t \in I_3 \backslash \{0, T\}$, there exists a subsequence, possibly dependent on t, of the sequence $\{\lambda_j\}$, again denoted by $\{\lambda_j\}$, such that

$$\lambda_j(t) \to \lambda_0(t), \qquad \text{for } t \in I_3 \backslash \{0, T\}.$$

Choosing the corresponding subsequence for the sequence $\{\psi_j\}$, we have

$$\psi_j(t) \overset{s}{\to} g^0(t), \qquad \text{in } X, \text{ for all } t \in I_3 \backslash \{0, T\}.$$

Thus, for every $t \in I_3 \backslash \{0, T\}$ and every $\epsilon > 0$, there exists an integer $\tilde{j} \equiv \tilde{j}(t, \epsilon)$ such that, for $j > \tilde{j}$, $S_t g^{n_j+i} \in N_\epsilon(S_t g^0)$, where again we use $N_\epsilon(S_t g^0)$ to denote the intersection of the ϵ-neighborhood of $S_t g^0$ with $D_\rho \equiv D(A^\rho(t))$. Since $\xi \to L(t, \xi, \nu)$ is continuous and $\xi \to f(t, \xi, \nu)$ is an USC (upper semicontinuous) multifunction on D_ρ for $(t, \nu) \in I \times M$, and since

$$\{S_t g^n, S_t g^0\} \subset B_{r_0} \cap D_\rho,$$

we have

$$Q(t, S_t g^{n_j+i}) \subset Q(t, N_\epsilon(S_t g^0)), \qquad \text{for } j > \tilde{j} \text{ and } t \in I_3 \backslash \{0, T\}.$$

It follows from the definition of Q that

$$(l_{n_j+i}(t), g^{n_j+i}(t)) \in Q(t, S_t g^{n_j+i}), \qquad \text{a.e. in } I.$$

Consequently, due to (25) and (27),

$$(\lambda_j(t), \psi_j(t)) \in \text{co } Q(t, S_t g^{n_j+i});$$

and hence, for $j > \tilde{j}$,

$$(\lambda_j(t), \psi_j(t)) \in \text{co } Q(t, N_\epsilon(S_t g^0)).$$

In view of the above facts,

$$\lim_j (\lambda_j(t), \psi_j(t)) = (\lambda_0(t), g^0(t)) \in \text{clco } Q(t, N_\epsilon(S_t g^0)),$$

for every $\epsilon > 0$, and hence

$$(\lambda_0(t), g^0(t)) \in \bigcap_{\epsilon>0} \text{clco } Q(t, N_\epsilon(S_t g^0)).$$

Therefore, by the weak Cesari property [see (aiv), Theorem 3.1],

$$(\lambda_0(t), g^0(t)) \in Q(t, S_t g^0), \qquad \text{for all } t \in I_3 \backslash \{0, T\}, \tag{32}$$

and hence almost everywhere in $I = [0, T]$. This implies that, for every $t \in I_3$, there exists a $\tilde{\mu}_t \in M$ such that

$$\lambda_0(t) \geq L(t, S_t g^0, \tilde{\mu}_t) \quad \text{and} \quad g^0(t) \in f(t, S_t g^0, \tilde{\mu}_t).$$

Since I_3 has full Lebesgue measure, we have

$$\lambda_0(t) \geq L(t, S_t g^0, \tilde{\mu}_t), \qquad \text{a.e. in } I,$$

$$g^0(t) \in f(t, S_t g^0, \tilde{\mu}_t), \qquad \text{a.e. in } I.$$

This defines a function $t \to \tilde{\mu}_t$ with values $\tilde{\mu}_t \in M \equiv M(\Gamma)$, a.e. The question that remains to be settled is whether or not a measurable (w^*-measurable) substitute for $\tilde{\mu}$ can be found. At this point, essentially we follow the arguments given in Refs. 1 and 2. Define, for $t \in I_3 \backslash \{0, T\}$, the set-valued map V with values

$$V(t) = \{\nu \in M(\Gamma) \colon \lambda_0(t) \geq L(t, S_t g^0, \nu)$$

and

$$g^0(t) \in f(t, S_t g^0, \nu)\} \subset M(\Gamma). \tag{33}$$

If we can show that $t \to V(t)$ is a measurable multifunction with closed values, then it will follow from the Ryll-Nardzweski selection theorem (Ref. 2, Theorem 1.4.5, p. 40) that there exists a measurable substitute μ^0 (i.e., $\mu^0 \in \mathcal{M}$) for $\tilde{\mu}$. But, if a set-valued function has closed values, then, by a theorem due to Himmelberg-Jacobs-Van Vleck (Ref. 2, Theorem 1.4.3, p. 39), it is measurable if and only if, for each $\epsilon > 0$, there exists a closed set $I_\epsilon \subset \tilde{I} \equiv I_3 \backslash \{0, T\}$ such that the graph of V restricted to I_ϵ is a closed subset of $I_\epsilon \times M$. The fact that, for each $t \in I_0$, $V(t)$ has closed values follows easily from the facts that

(i) $\nu \to L(t, S_t g^0, \nu)$ is continuous on M;
(ii) $\nu \to f(t, S_t g^0, \nu)$ is USC on M;
(iii) $f(t, S_t g^0, \nu) \in \mathrm{CC}(X)$.

For measurability of V, we consider its graph

$$\gamma(V) \equiv \{(t, \nu) \in I_0 \times M: \nu \in V(t)\}, \tag{34}$$

and show that, for each $\epsilon > 0$, there exists a closed set $I_\epsilon \subset \tilde{I}$ such that the restriction of V to I_ϵ, denoted V_ϵ, has a closed graph. Since $\lambda_0 \in L_1$, $g^0 \in L_1(I, X)$, and since L satisfies assumption (aiii) of Theorem 3.1 and f satisfies assumptions (aii)-(aiv) of Lemma 2.3, it is clear that, for every $\epsilon > 0$, there exists a closed set $I_\epsilon \subset \tilde{I}$ such that:

(i)′ $t \to \lambda_0(t)$ is continuous on I_ϵ;
(ii)′ $t \to g^0(t)$ is continuous on I_ϵ;
(iii)′ $(t, \nu) \to L(t, S_t g^0, \nu)$ is continuous on $I_\epsilon \times M$;
(iv)′ $(t, \nu) \to f(t, S_t g^0, \nu)$ is USC on $I_\epsilon \times M$.

Let V_ϵ denote the restriction of V to I_ϵ, and let $\gamma(V_\epsilon)$ denote its graph. We show that $\gamma(V_\epsilon)$ is a closed subset of $I_\epsilon \times M$. Let $(t_n, \nu_n) \in \gamma(V_\epsilon)$ and $t_n \to t_0$ and $\nu_n \overset{w^*}{\to} \nu_0$ [or, equivalently, in the metric topology of $M(\Gamma)$]. Since I_ϵ is closed, $t_0 \in I_\epsilon$; and, due to (i)′ and (iii)′,

$$\lambda_0(t_0) = \lim_n \lambda_0(t_n) \geq \underline{\lim}_n L(t_n, S_{t_n} g^0, \nu_n) = L(t_0, S_{t_0} g^0, \nu_0). \tag{35}$$

By virtue of (iv)′, there exists $n_0 = n_0(\epsilon)$ such that, for $n > n_0$,

$$f(t_n, S_{t_n} g^0, \nu_n) \subset f^\epsilon(t_0, S_{t_0} g^0, \nu_0),$$

where f^ϵ denotes the ϵ-nbh of the set $f(t_0, S_{t_0} g^0, \nu_0)$ in X. Since

$$g^0(t_n) \in f(t_n, S_{t_n} g^0, \nu_n), \qquad \text{for all } n,$$

and, by (ii),

$$g^0(t_n) \to g^0(t_0), \qquad \text{in } X, \text{ on } I_\epsilon,$$

it follows from the above that

$$g^0(t_0) \in f^\epsilon(t_0, S_{t_0} g^0, \nu_0).$$

Since $\epsilon > 0$ is arbitrary and $f(t_0, S_{t_0} g^0, \nu_0) \in \mathrm{CC}(X)$,

$$g^0(t_0) \in f(t_0, S_{t_0} g^0, \nu_0). \tag{36}$$

Hence, it follows from (33)-(36) that $(t_0, \nu_0) \in \gamma(V_\epsilon)$ and, consequently, V_ϵ has a closed graph. Thus, by the results mentioned above, V is a measurable set-valued map with closed values in $M(\Gamma)$. Hence, by the Ryll-Nardzweski theorem mentioned above, there exists a measurable selection of V, denoted μ^0, i.e., $\mu^0 \in \mathcal{M}$. Consequently,

$$\lambda_0(t) \geq L(t, S_t g^0, \mu_t^0), \qquad \text{a.e.},$$

and

$$g^0(t) \in f(t, S_t g^0, \mu_t^0), \qquad \text{a.e.};$$

and hence, $(\mu^0, g^0) \in \mathcal{A}$. Consequently, $j_0 = \eta(\mu^0, g^0)$, and hence (μ^0, g^0) is an optimal pair and μ^0 is an optimal (relaxed) control. This concludes the proof. □

The following corollary follows from the Dunford–Pettis theorem that justifies the equivalence

$$(L_1(I, C(\Gamma)))^* \cong L_\infty(I, \mathring{M}(\Gamma)),$$

where $\mathring{M}(\Gamma)$ is the space of finite Radon measures on Γ, with $M(\Gamma) \subset \mathring{M}(\Gamma)$ and $\mathcal{M} \subset L_\infty(I, \mathring{M}(\Gamma))$ furnished with the w^*-topology.

Corollary 3.1. Let

$$\inf\{\eta(\mu, g), (\mu, g) \in \mathcal{A}\} = j_0,$$

and define

$$\mathcal{A}_0 \equiv \{(\mu, g) \in \mathcal{A}: \eta(\mu, g) = j_0\}.$$

Then, under the hypothesis of Theorem 3.1, the set $\mathcal{A}_0$ is closed.

Remark 3.4. Our results on existence theorems also extend for the space of control measures M with variable supports Γ. For example, in case $\Gamma: I \to 2^B$ with values $\Gamma(t) \in \mathrm{CC}(\Gamma^0)$, where Γ^0 is a fixed compact subset of the Polish space B and $t \to \Gamma(t)$ is both upper and lower semicontinuous (i.e., simply continuous), then both Theorem 2.1 and Theorem 3.1 remain valid. However, for feedback controls, the problem becomes more difficult. The major difficulty lies in the proof of existence of (mild) solutions for the inclusions of evolution

$$\dot{x}(t) + A(t)x(t) \in f(t, x(t), \mu_t), \qquad t \in I,$$

$$\mu_t \in M(\Gamma(t, x(t))), \qquad t \in I,$$

where $\Gamma: I \times (D_\rho \cap B_r) \to 2^B$ (for r sufficiently large) and M is the space of bounded positive Radon measures on Γ. For the existence of mild solutions, this problem can be reformulated as a family of fixed-point problems in $Y \times \mathcal{M}$,

$$g(t) \in f(t, S_t g, \mu_t), \qquad t \in I,$$

$$\mu_t \in M(\Gamma(t, S_t g)), \qquad t \in I,$$

where $Y = L_1(I, X)$ and $\mathcal{M}$ is the space of w^*-measurable functions on I with values in the space of bounded positive Radon measures with support Γ. Given that the fixed-point problems have solutions, and assuming that $\Gamma(t, x) \in CC(\Gamma^0)$ for some fixed compact Γ^0 in B, for all $(t, x) \in I \times (D_\rho \cap B_r)$ (for r sufficiently large), and given that $(t, x) \to \Gamma(t, x)$ is both upper and lower semicontinuous with respect to inclusions, one can prove Theorem 3.1. Probably, the continuity conditions and the compactness of Γ^0 may be relaxed. In general, it would be desirable to relax the regularity assumptions (aiii) and (aiv) of Lemma 2.3.

4. Examples

In this section we present two examples, one involving distributed controls and the other involving boundary controls, to which our results apply.

Example 4.1. *Distributed Control.* Consider the parabolic control problem

$$\partial\phi/\partial t + \sum_{|\alpha| \leq 2m} a_\alpha(t, \xi) D^\alpha \phi = f(t, \xi; \phi, D_1\phi, \ldots, D_{2m-1}\phi; u),$$

$$\text{in } (0, T) \times \Omega \equiv Q, \tag{37a}$$

$$\tau\phi \equiv \{B_j\phi \equiv \sum_{|\alpha| \leq m_j} b^j_\alpha(\xi) D^\alpha \phi, 0 \leq j \leq m-1\} = 0,$$

$$\text{in } (0, T) \times \partial\Omega, \tag{37b}$$

$$\phi(0, \xi) = \phi_0(\xi), \qquad \text{in } \Omega, \tag{37c}$$

where Ω is an open bounded subset of R^n with C^1 boundary $\partial\Omega$,

$$D^\alpha \equiv D^{\alpha_1} D^{\alpha_2} \cdots D^{\alpha_n}, \qquad D_j \equiv \{D^\alpha, |\alpha| = j\},$$

α_i, j nonnegative integers,

$$|\alpha| \equiv \sum_{i=1}^n \alpha_i, \qquad D^{\alpha_i} \equiv \partial^{\alpha_i}/\partial\xi_i^{\alpha_i},$$

and m_j is the order of B_j. Define

$$A_0(t)\phi \equiv \sum_{|\alpha| \leq 2m} a_\alpha(t, \cdot) D^\alpha \phi,$$

and assume that the system $\{A_0(t), \tau, \Omega\}$ constitutes a regular elliptic boundary-value problem, with principal coefficients $a_\alpha(|\alpha| = 2m) \in C(\bar{Q})$ and lower-order coefficients $a_\alpha(|\alpha| \leq 2m-1) \in L_\infty(Q)$, and further they are all Hölder continuous in t uniformly with respect to $x \in \bar{\Omega}$. Let $1 < p < \infty$, and define the operator $A(t)$, $t \in I = [0, T]$, by

$$D(A(t)) = \{\phi \in L_p(\Omega) \colon A_0(t)\phi \in L_p(\Omega) \text{ and } \tau\phi = 0\},$$

and assume that the coefficients of B_j belong to $C^{m-m_j}(\partial\Omega)$. Then, under the above assumptions, for each $t \in I$, $A(t)$ is a closed densely defined linear operator in $X \equiv L_p(\Omega)$, and $-A(t)$ is the generator of an analytic semigroup in X (Ref. 16, p. 140). For the nonlinear operator f, we may choose B to be $L_\infty(\Omega)$ furnished with the w^*-topology, and consider Γ to be a closed bounded convex subset of B with the relative w^*-topology, thereby making it into a compact Polish space, since $L_1(\Omega)$ is separable. Define

$$\tilde{\eta}_j \equiv \{\eta_\alpha, |\alpha| = j\} \in R^{N_j},$$

for some suitable N_j, and

$$\eta \equiv \{\tilde{\eta}_j, j = 0, 1, \ldots, 2m-1\} \subset R^N, \qquad N = \Sigma N_j.$$

We consider the function $f \colon I \times \Omega \times R^N \times \Gamma \to R$, and suppose that it is measurable on $I \times \Omega$ for each $(\eta, \sigma) \in R^N \times \Gamma$ and continuous on $R^N \times \Gamma$ for almost all $(t, \xi) \in I \times \Omega$; also, we suppose that there exists a constant $C_0 \geq 0$ and a nonnegative function $w \in L_p(\Omega)$ such that

$$|f(t, \xi, \eta, \sigma)| \leq w(\xi) + C_0 \sum_{0 \leq j \leq 2m-1} |\tilde{\eta}_j|^{r_j}, \qquad \text{a.e. in } Q, \tag{38}$$

uniformly with respect to $\sigma \in \Gamma$.

The admissible values for the exponents r_j are determined by the use of Sobolev's embedding theorem which (partly) states that $W^{2m,p} \hookrightarrow W^{j,p_j}$, provided

$$(1/p_j) = (j/n) + (1/p) - (2m/n),$$

giving

$$r_j = (p_j/p).$$

In order that $r_j \geq 0$, we require that $n \geq 2mp$. Under these assumptions, it is easily verified that there exists a constant $C_1 > 0$ such that, for all $\sigma \in \Gamma$ and $t \in I$,

$$\|f(t, \cdot\,; \phi, D_1\phi, \ldots, D_{2m-1}\phi; \sigma)\|_{L_p(\Omega)}$$
$$\leq C_1 \Big\{1 + \sum_{0 \leq j \leq 2m-1} \|D_j\phi\|_{L_{p_j}(\Omega)}\Big\}, \tag{39}$$

for all $\phi \in W^{2m,p}(\Omega)$. Using this inequality, with slight abuse of notation, one can conclude that there exists a constant k, $0 < k < \infty$, such that, for all $t \in I$ and $\sigma \in \Gamma$,

$$\|f(t, x, \sigma)\|_X \leq k(1 + \|A^\rho(t)x\|_X), \tag{40}$$

for any $\rho \in [(2m-1)/2m, 1)$ and $x \in D(A^{\rho}(t))$ (constant). It follows from the above discussion that we can reformulate the system (37) as a relaxed evolution equation

$$dx/dt + A(t)x = f(t, x, \nu), \tag{41a}$$

$$x(0) = x_0 \equiv \phi_0, \qquad \nu \in M(\Gamma), \tag{41b}$$

in the Banach space X, with

$$f(t, x, \nu) \equiv \int_{\Gamma} f(t, x, \sigma)\nu(d\sigma).$$

Obviously, our general results apply to the optimal control problem (41) and (3). In case the function $(t, \xi, \eta, \sigma) \to f(t, \xi, \eta, \sigma)$ is multivalued in any one of the variables $\{\eta, \sigma\}$, we obtain a differential inclusion. In that case, the continuity condition is replaced by the upper semicontinuity with respect to set inclusions retaining, however, the growth condition (38) or equivalently (40).

In a recent paper, Seidman and Zhou (Ref. 17) have studied the question of the existence and uniqueness of optimal controls for a class of quasilinear parabolic equations of the form given in this example.

Example 4.2. *Boundary Controls.* Here, we give an example illustrating that our abstract results also apply to boundary control problems. For simplicity, we shall consider only second-order elliptic operators, though our results apply equally to higher-order problems.

Consider a control problem of the form

$$\partial\phi/\partial t + A_0\phi = 0, \qquad \text{in } Q = (0, T) \times \Omega, \tag{42a}$$

$$\partial\phi/\partial\nu + \beta_i(\phi) \ni h_i(t, u_i), \qquad \text{in } \Sigma_i = (0, T) \times \partial\Omega_i,\ i = 1, 2, \tag{42b}$$

$$\phi(0, \xi) = \phi_0(\xi), \qquad \xi \in \Omega, \tag{42c}$$

with

$$\partial\Omega_1 \cup \partial\Omega_2 = \partial\Omega, \qquad \partial\Omega_1 \cap \partial\Omega_2 = \varnothing.$$

Here, Ω is an open bounded subset of R^n, A_0 is a second-order elliptic operator, and β_i is a multivalued function from R to closed convex subsets of R with monotone graphs. The controls $u_i \in L_2(I, U_i)$, while U_i are certain separable Hilbert spaces and h_i are continuous bounded operators from $L_2(I, U_i)$ to $L_2(I, H_i)$, where H_i are certain Sobolev spaces over $\partial\Omega_i$, $i = 1, 2$. More specific assumptions will be introduced shortly. We can write this boundary-value problem as an abstract differential inclusion in a Banach space, in particular, the Hilbert space $H \equiv L_2(\Omega)$ or $(H^1)^*$. For definiteness, let

$$A_0\phi \equiv -\sum_{i,j} (a_{ij}(\xi)\phi_{\xi_i})_{\xi_j} + a_0(\xi)\phi,$$

where

$$a_{ij} \in C^1(\bar{\Omega}), \qquad a_0 \in L_\infty(\Omega), \qquad a_{ij} = a_{ji},$$

and, for some $\alpha > 0$,

$$\sum_{i,j} a_{ij}(\xi)\eta_i\eta_j \geq \alpha|\eta|^2,$$

for almost all $\xi \in \Omega$ and all $\eta \in R^n$, and $a_0 \geq 0$, a.e., on Ω. Define the bilinear form

$$a(\phi, \psi) \equiv \sum_{i,j} \int_\Omega a_{ij}(\xi)\phi_{\xi_i}\psi_{\xi_j}\, d\xi + \int_\Omega a_0(\xi)\phi\psi\, d\xi, \tag{43}$$

for ϕ, $\psi \in H^1$, and the operator A such that

$$(A\phi, \psi) = a(\phi, \psi),$$

with

$$\mathrm{D}(A) = \{\phi \in H: A\phi \in H\} \quad \text{or} \quad \mathrm{D}(A) = \{\phi \in (H^1)^*: A\phi \in (H^1)^*\}$$

and

$$A\phi = A_0\phi, \qquad \text{for } \phi \in H_0^1 \cap H^2 \text{ or } H_0^1.$$

Under these assumptions, A is the generator of an analytic semigroup both in H and $(H^1)^*$; in the first case, $\mathrm{D}(A) = H^2$; in the second case, $\mathrm{D}(A) = H^1$. We define the multivalued operator f as follows. Let $B \equiv U_1 \times U_2$ and $\Gamma \equiv \Gamma_1 \times \Gamma_2$ a closed bounded convex subset of B, and take $M(\Gamma) \equiv M(\Gamma_1) \times M(\Gamma_2)$. Let $t \to h_i(t, \sigma)$ be (strongly) measurable on I for each $\sigma \in \Gamma_i$, and $\sigma \to h_i(t, \sigma)$ continuous and bounded on Γ_i for almost all $t \in I$ with range in H_i. Define, for each $\nu_i \in M(\Gamma_i)$,

$$\nu_i(h_i)(\cdot) \equiv \int_{\Gamma_i} h_i(\cdot, \sigma)\nu_i(d\sigma),$$

amd suppose that $\nu_i(h_i) \in L_2(I, H_i)$.

For $X = H$, we have $\mathrm{D}(A) = H^2$, and it follows from integration by parts of (42) that a natural choice of H_i is $H^{1/2}(\partial\Omega_i)$, $i = 1, 2$, and that the multifunction β_i determines a multivalued operator (again denoted by β_i)

$$\beta_i : H^1(\Omega) \to 2^{H_i}.$$

The multivalued operator $f\colon I\times H^1\times M(\Gamma)\to 2^H$ is then determined by the relation

$$(f(t,\phi,\nu),\psi)_H\equiv\left\{\sum_{i=1}^{2}(w_i-\nu_i(h_i)(t),\psi)_{H_i,H_i^*}\colon w_i\in\beta_i(\phi)\right\},\tag{44}$$

which holds for all

$$\psi\in H(\psi|_{\partial\Omega_i}\in H_i^*=H^{-1/2}(\partial\Omega_i)),$$

for each given $(t,\phi,\nu)\in I\times H^1\times M$. Taking $\mathcal{M}$ as in the introduction, we can rewrite Eq. (42) as a differential inclusion in H,

$$dy/dt+Ay\in f(t,y,\mu_t),\qquad y(0)=y_0\equiv\phi_0\in H,$$

with $\mu\in\mathcal{M}$. Assuming that there exists a constant K, independent of t, such that, for all $\nu_i\in M(\Gamma_i)$, $i=1,2$,

$$\begin{aligned}&\|\beta_i(\phi)-\nu_i(h_i)\|_{H_i}^0\\&\equiv\sup\{\|w_i-\nu_i(h_i)\|_{H_i},\ w_i\in\beta_i(\phi)\}\le K(1+\|\phi\|_{H^1}),\end{aligned}\tag{45}$$

it follows from (44) that there exists a constant $\tilde{K}$, dependent on K and the bounds of a_{ij} and a_0, such that, for all $\nu\in M(\Gamma)$,

$$\|f(t,\phi,\nu)\|_H^0\le\tilde{K}(1+\|A^\rho\phi\|_H),\qquad\text{with }\rho=1/2.\tag{46}$$

Similarly, for $X\equiv(H^1)^*$, we have $\mathrm{D}(A)=H^1$; and it follows from similar arguments as above that a natural choice for H_i, in this case, is $H^{-1/2}(\partial\Omega_i)$. If β is a multivalued operator mapping H to 2^{H_i}, then

$$f\colon I\times H\times M(\Gamma)\to 2^{(H^1)^*}.$$

Assuming a similar estimate as in (45), with H replacing H^1, one has, for all $\nu\in M(\Gamma)$,

$$\|f(t,\phi,\nu)\|_{(H^1)^*}^0\le\tilde{K}(1+\|A^\rho\|_{(H^1)^*}),\qquad\text{for }\rho=1/2.\tag{47}$$

In this case, the corresponding differential inclusion is defined in $(H^1)^*$ with $y_0\equiv\phi_0\in(H^1)^*$. For practical applications, one would prefer to take $U_i\equiv L_2(K_i)$ for the control space, where K_i is any nonempty Lebesgue measurable subset of $\partial\Omega_i$ and $h_i\colon L_2(K_i)\to L_2(\partial\Omega_i)$ is a Urysohn operator, given by

$$h_i(\sigma)(\zeta)=\int_{K_i}R_i(\zeta,\xi;\sigma(\xi))\,d\xi,\qquad\zeta\in\partial\Omega_i.$$

Assuming that the Lebesgue $(n-1)$ measure of $\partial\Omega_i$ is finite and that there exists an $m_i\in L_2(\partial\Omega_i\times K_i)$ and a constant $C_i\ge 0$ such that

$$|R_i(\zeta,\xi;v)|\le m_i(\zeta,\xi)+C_i|v|,$$

for all $(\zeta, \xi) \in \partial\Omega_i \times K_i$ and $v \in R$, one can show that h_i is a continuous bounded operator from $L_2(K_i)$ to $L_2(\partial\Omega_i)$. The operator β_i may be considered to be given by the composition of the trace map $\tau_i : \phi \to \phi|_{\partial\Omega_i}$ and a multifunction $\tilde{\beta}_i$ from R to closed convex subsets of R, so that

$$|\tilde{\beta}_i(r)|^0 \equiv \sup\{|\theta|, \theta \in \tilde{\beta}(r)\} \leq C(1+|r|), \qquad \text{for some } 0 \leq C < \infty.$$

Clearly, for each $\psi \in L_2(\partial\Omega_i)$,

$$\tilde{\beta}_i(\psi) \in 2^{L_2(\partial\Omega_i)};$$

hence, for $\phi \in H^{1/2}(\Omega)$,

$$\beta_i(\phi) \equiv \tilde{\beta}_i(\tau_i\phi) \in 2^{L_2(\partial\Omega_i)}.$$

Then, for any closed bounded convex subset $\Gamma_i \subset U_i$, there exists a constant $\tilde{C}_i$ such that

$$\| \beta_i(\phi) - h_i(\sigma) \|^0_{L_2(\partial\Omega_i)} \leq \tilde{C}_i(1 + \|\phi\|_{H^{1/2}(\Omega)}), \qquad \text{for all } \sigma \in \Gamma_i;$$

consequently, for the corresponding multivalued operator f, there exists a constant $\tilde{k}$ such that

$$\|f(t, \phi, \nu)\|^0_{(H^1)^*} \leq \tilde{k}(1 + \|A^\rho\phi\|_{(H^1)^*}), \qquad \text{with } \rho = 3/4.$$

In any case, our results apply to the boundary control problem (42) and (3).

In a recent paper, Barbu (Ref. 18) has developed the necessary conditions of optimality for convex control problems involving systems of the form (42).

Remark 4.1. For recent results on the necessary conditions of optimality and the existence of time-optimal control for distributed and boundary control problems, the reader is referred to Refs. 18–21. In a recent paper (Ref. 21), the author has studied the questions of complete controllability, including the necessary conditions of optimality for a very general class of linear evolution equations on a Banach space including applications to distributed boundary control problems.

References

1. AHMED, N. U., *Properties of Relaxed Trajectories for a Class of Nonlinear Evolution Equations on a Banach Space*, SIAM Journal on Control and Optimization, Vol. 21, pp. 953–968, 1983.
2. AHMED, N. U., and TEO, K. L., *Optimal Control of Distributed Parameter Systems*, North-Holland, New York, New York, 1981.
3. AHMED, N. U., *Optimal Control of a Class of Strongly Nonlinear Parabolic Systems*, Journal of Mathematical Analysis and Applications, Vol. 61, pp. 188–207, 1977.

4. AHMED, N. U., *Some Comments on Optimal Control of a Class of Strongly Nonlinear Parabolic Systems*, Journal of Mathematical Analysis and Applications, Vol. 68, pp. 595-596, 1979.
5. AHMED, N. U., and TEO, K. L., *Optimal Control of Systems Governed by a Class of Nonlinear Evolution Equations in a Reflexive Banach Space*, Journal of Optimization Theory and Applications, Vol. 25, pp. 57-81, 1978.
6. AHMED, N. U., *Relaxed Control in Infinite Dimensional Systems*, Proceedings of IEEE-CDC, pp. 1073-1077, 1982.
7. CESARI, L., *Geometric and Analytic Views in Existence Theorems for Optimal Control in Banach Spaces, II, Distributed and Boundary Controls*, Journal of Optimization Theory and Applications, Vol. 15, pp. 467-497, 1975.
8. CESARI, L., *Geometric and Analytic Views in Existence Theorems for Optimal Controls in Banach Spaces, III, Weak Solution*, Journal of Optimization Theory and Applications, Vol. 19, pp. 185-214, 1976.
9. CESARI, L., *Optimization Theory and Applications, Problems with Ordinary Differential Equations*, Springer-Verlag, New York, New York, 1983.
10. BERKOVITZ, L. D., *Existence and Lower Closure Theorems for Abstract Control Problems*, SIAM Journal on Control and Optimization, Vol. 12, pp. 27-42, 1974.
11. WARGA, J., *Optimal Control of Differential and Functional Equations*, Academic Press, New York, New York, 1972.
12. DIESTEL, J., and UHL, JR., J. J., *Vector Measures*, Mathematical Surveys, No. 15, American Mathematical Society, Providence, Rhode Island, 1977.
13. FRIEDMAN, A., *Partial Differential Equations*, Robert E. Krieger Publishing Company, Huntington, New York, 1976.
14. KATO, T., and TANABE, H., *On the Abstract Evolution Equation*, Osaka Journal of Mathematics, Vol. 14, pp. 107-133, 1962.
15. BROWDER, F. E., *Nonlinear Operators and Nonlinear Equations of Evolution in Banach Spaces*, Proceedings of Symposium on Pure Mathematics, Vol. 8, Part 2, American Mathematical Society, Providence, Rhode Island, 1976.
16. TANABE, H., *Equations of Evolution*, Pitman, London, England, 1979.
17. SEIDMAN, T. I., and ZHOU, H. X., *Existence and Uniqueness of Optimal Controls for a Quasilinear Parabolic Equation*, SIAM Journal on Control and Optimization, Vol. 20, pp. 747-762, 1982.
18. BARBU, V., *Boundary Control Problems with Nonlinear State Equation*, SIAM Journal on Control and Optimization, Vol. 20, pp. 125-143, 1982.
19. BARBU, V., *Boundary Control Problems with Convex Cost Criterion*, SIAM Journal on Control and Optimization, Vol. 18, pp. 227-243, 1980.
20. AHMED, N. U., *A Note on the Maximum Principle for Time Optimal Control for a Class of Distributed-Boundary Control Problems*, Journal of Optimization Theory and Applications, Vol. 45, pp. 147-157, 1985.
21. AHMED, N. U., *Finite-Time Null Controllability for a Class of Linear Evolution Equations on a Banach Space with Control Constraints*, Journal of Optimization Theory and Applications, Vol. 47, pp. 129-158, 1985.