MPE — Volume 1, pp. 179-191
Reprints available directly from the publisher
Photocopying permitted by license only

NECESSARY CONDITIONS OF OPTIMALITY FOR INFINITE DIMENSIONAL UNCERTAIN SYSTEMS

N. U. AHMED

Department of Electrical Engineering and Department of Mathematics, University of Ottawa, Ottawa, Ontario, K1N 6N5 Canada

X. XIANG

Department of Mathematics, Guizhou University, Guiyang, Guizhou, P. R. China

(Received 10 November 1994)

In this paper we consider optimal control problem for infinite dimensional uncertain systems. Necessary conditions of optimality are presented under the assumption that the principal operator is the infinitesimal generator of a strongly continuous semigroup of linear operators in a reflexive Banach space. Further, a computational algorithm suitable for computing the optimal policies is also given.

AMS No.: 93C25
Keywords: Uncertain systems, optimal controls, semigroups, operators, necessary conditions

1. INTRODUCTION

Many physical systems, such as thermodynamics, electrodynamics, and population biology are governed by differential equations, integro differential equations, or more generally functional differential inclusions on Banach spaces. The abstract mathematical model for such systems can be described as follows:

$$\begin{cases} \dot{x}(t) \in Ax(t) + F(t, x(t)) \\ x(0) = x^0 \end{cases} \tag{1}$$

where A is typically a linear unbounded operator in a suitable Banach space and F a multi-valued map. An associated control system may be described as

$$\begin{cases} \dot{x} \in Ax(t) + G(t, x(t), u(t)) \\ x(0) = x^0 \end{cases} \tag{2}$$

where G is a multi-valued map and u is a suitable function representing the control actions. Many engineering systems with incomplete mathematical model can be formulated as differential inclusions rather than differential equations. This is because of either uncertainty in the measurement of system parameters or their random fluctuation. Evolution inequalities rather than equations also can be formulated as differential inclusions.

In general we consider the mild solutions of (2). For each admissible control u, let $X(u)$ denote the set of solutions of (2) corresponding to u. A natural problem is to find $u_0 \in U_{ad}$ (admissible controls) so that

$$J_0(u_0) = \inf_{u \in U_{ad}} J_0(u) \tag{3}$$

where

$$J_0(u) = \sup_{x \in X(u)} \int_I l(t, x(t), u(t))dt.$$

Clearly, this is a minimax problem.

Recently, optimal control of systems governed by evolution inclusions on infinite dimensional spaces have been studied [3, 6, 7]. we studied the question of existence of optimal controls for infinite-dimensional uncertain systems (see [4]). To the knowledge of the authors, very little seems to be known about the necessary conditions of optimality for infinite dimensional uncertain systems.

In this paper we will present some necessary conditions of optimality for a special class of system (2) where the multi-valued map G is given by

$$G(t, x, u) \equiv \{\int_\Sigma g(t, x, u, \sigma)\mu(d\sigma) \qquad \mu \in M(\Sigma)\}$$

with $M(\Sigma)$ being the space of probability measures. Here A is assumed to be the infinitesimal generator of a strongly continuous semigroup. This is a broad class of systems representing physical processes with parametric uncertainty, incompleteness of the mathematical model, and systems with randomly fluctuating parameters, and so on. The problem for the system designer is to find a control policy that minimizes the maximum risk or maximizes the minimum revenue.

2. NOTATIONS AND ASSUMPTIONS

Let X be a reflexive Banach space considered as the state space; and Y another reflexive Banach space where the control take their values from. For any interval $I = [0,T]$, $T < \infty$, $L_p(I,Y)$, $1 \leq p < \infty$ will denote the Banach space of strongly measurable $Y-$ valued functions having pth power summable norms. For any two Banach spaces X and Y, $L(X, Y)$ will denote the space of bounded linear operators from X to Y.

Let Σ be a compact Polish space and $M(\Sigma)$ the space of probability measures on Σ. A sequence $\mu_n \in M(\Sigma)$ is said to converge weakly to $\mu \in M(\Sigma)$ if

$$\int_\Sigma g(\sigma)\mu_n(d\sigma) \to \int_\Sigma g(\sigma)\mu(d\sigma) \qquad \text{in } X$$

for every $g \in C(\Sigma, X)$, the space of continuous functions from Σ to X.

For the necessary conditions we shall introduce the following assumptions:

(A1) A is the infinitesimal generator of a strongly continuous semigrup $T(t), t \geq 0$, in X.

(A2) U: $I \to CC(Y)$ = class of nonempty, closed, convex subset of Y is a measurable multifunction satisfying $U(t) \subseteq \mathcal{U}$ for almost all $t \in I$, where $\mathcal{U}$ is a fixed weakly compact convex subset of Y. For the admissible controls, we choose the set $U_{ad} \equiv \{u \in L_p(Y)$: $u(t) \in U(t)$ $a,e\}$ $(1 \leq p \leq \infty)$.

(A3) $\tilde{g}$: $I \times X \times Y \times \Sigma \to X$, for $\mu \in M(\Sigma)$, $g(t, x, u, \mu) \equiv \int_\Sigma \tilde{g}(t, x, u, \sigma)\mu(d\sigma)$,

1) $t \to g(.,x, u, \mu)$ is a measurable function,

2) For any given $u \in \mathcal{U}_{ad}$, $\mu \in M(\Sigma)$, there exists a $K \in L^+_1(I)$ s.t.

$$\| g(t, x, u, \mu) \|_X \leq K(t)(1 + \| x \|_X),$$

3) g is Frechet differentiable with respect to x and u with the Frechet derivatives g_x, g_u bounded measurable in t on I and continuous and bounded on bounded subsets of $X \times Y$.

(A4) l: $I \times X \times Y \to R\cup\{+\infty\}$ is continuous, and Frechet differentiable in both x and u on X and Y, respectively, so that $l_x \in L_1(I, X^*)$ and $l_u \in L_q(I, Y^*)$ in the neighborhood of the optimal trajectories.

Under the assumptions above we consider the following system

$$\begin{cases} \dot{x}(t) \in Ax(t) + G(t, x(t), u(t)) \\ x(0) = x^0 \end{cases} \tag{4}$$

where $G(t, x, u) \equiv \{g(t, x, u, \mu), \quad \mu \in M(\Sigma)\}$.

DEFINITION 2.1 A function x is said to be a mild solution of the problem (4) corresponding to $u \in U_{ad}$, if

1) $x \in C(I, X)$,

2) there exists a $\mu \in M(\Sigma)$ so that x satisfies the integral equation:

$$x(t) = T(t)x^0 + \int_0^t T(t - \tau)g(\tau, x(\tau), u(\tau), \mu)d\tau.$$

Define $X(u) \equiv \{x \mid x$ is a mild solution of (4) corresponding to $u\}$.

Our problem (P) is to find $u_0 \in U_{ad}$ s.t.

$$J_0(u_0) = \inf_{u \in U_{ad}} J_0(u)$$

where $J_0(u) = \sup_{x \in X(u)} \int_I l(t,x,u)dt$.

3. PREPARATORY RESULTS

Before discussing the necessary conditions, we need some preliminary results.

THEOREM 3.1 *Suppose the Assumptions (A1)–(A3) hold. Then, for every $x^0 \in X$, $u \in U_{ad}$ and $\mu \in M(\Sigma)$ the following Cauchy problem*

$$\begin{cases} \dot{x} = Ax(t) + g(t, x(t), u(t), \mu) \\ x(0) = x^0 \end{cases} \tag{5}$$

has a unique mild solution.

Proof. For proof see [8, 2].

Occasionally, for fixed $x_0 \in X$, we use the notation $x(u, \mu)$ to denote the solution of (5) corresponding to $u \in U_{ad}$ and $\mu \in M(\Sigma)$.

In the following two Theorems we give some results on the continuous dependence of solutions on controls and parameters.

THEOREM 3.2. *Suppose the Assumptions (A1)–(A3) hold and $x^0 \in X$ and $\mu_0 \in M(\Sigma)$ be given. Then if $u_n \to u_0$ in $L_p(Y)$ $(1 \le p \le \infty)$, the solution $x_n = x(u_n, \mu_0) \to x_0 = x(u_0, \mu_0)$ in $C(I, X)$.*

Proof. Define

$$R_n(t) \equiv \int_0^1 g_x(t, x_0(t) + s(x_n(t) - x_0(t)), u_n(t), \mu_0)ds$$

$$S_n(t) \equiv \int_0^1 g_u(t, x_0(t), u_0(t) + s(u_n(t) - u_0(t)), \mu_0)ds$$

Since

$$\begin{aligned} x_n(t) - x_0(t) &= \int_0^t T(t-\tau)[g(\tau, x_n(\tau), u_n(\tau), \mu_0) - g(\tau, x_0(\tau), u_0(\tau), \mu_0)]d\tau \\ &= \int_0^t T(t-\tau)R_n(\tau)(x_n(\tau) - x_0(\tau))d\tau \\ &+ \int_0^t T(t-\tau)S_n(\tau)(u_n(\tau) - u_0(\tau))d\tau \end{aligned}$$

the assertion follows from Assumptions (A1), (A3), and Gronwall inequality.

THEOREM 3.3. *Suppose the Assumptions (A1)–(A3) hold and $x^0 \in X$ and $\mu_0 \in M(\Sigma)$. For any $\mu \in M(\Sigma)$ and $\varepsilon \ge 0$, let $\mu_\varepsilon = \mu_0 + \varepsilon(\mu - \mu_0)$. Then as $\varepsilon \to 0$, $x_\varepsilon = x(u_0, \mu_\varepsilon) \to x_0 = x(u_0, \mu_0)$ in $C(I, X)$.*

Proof. Define

$$L_\varepsilon(\tau) \equiv \int_0^1 g_x(\tau, x_0(\tau) + s(x_\varepsilon(\tau) - x_0(\tau)), u_0(\tau), \mu_\varepsilon)ds.$$

Then we have

$$x_\varepsilon(t) - x_0(t) = \int_0^t T(t-\tau)[g(\tau, x_\varepsilon(\tau), u_0(\tau), \mu_\varepsilon) - g(\tau, x_0(\tau), u_0(\tau), \mu_0)]d\tau$$

$$= \int_0^t T(t-\tau)[g(\tau, x_\varepsilon(\tau), u_0(\tau), \mu_\varepsilon) - g(\tau, x_0(\tau), u_0(\tau), \mu_\varepsilon)]d\tau$$

$$+ \int_0^t T(t-\tau)[g(\tau, x_0(\tau), u_0(\tau), \mu_\varepsilon) - g(\tau, x_0(\tau), u_0(\tau), \mu_0)]d\tau$$

$$= \int_0^t T(t-\tau)L_\varepsilon(\tau)(x_\varepsilon(\tau) - x_0(\tau))d\tau$$

$$+ \varepsilon \int_0^t T(t-\tau)[g(\tau, x_0(\tau), u_0(\tau), \mu) - g(\tau, x_0(\tau), u_0(\tau), \mu_0)]d\tau.$$

By virtue of assumptions (A1) and (A3), there exist constants $M > 0$, $L > 0$, so that

$$\| x_\varepsilon(t) - x_0(t) \| X \leq \varepsilon L + M \int_0^t \| x_\varepsilon(\tau) - x_0(\tau) \|_x d\tau.$$

Hence, by Gronwall Lemma, we have

$$x_\varepsilon \to x_0 \quad in \quad C(I,X)$$

as $\varepsilon \to 0$.

Remark 3.4. In fact, if μ_n converges weakly to $\mu \in M(\Sigma)$, one can easily verify that $x(u_0, \mu_n) \to x(u_0, \mu_0)$ in $C(I, X)$.

Let X^* denote the topological dual of X and X^*_w the space X^* endowed with the w^* topology. Let $C(I, X^*_w)$ denote the topological space of w^*-continuous X^*-valued functions defined on the interval I. Let $\langle\ ,\ \rangle_{z^*,z}$ denote the duality pairing between Z^* and Z, where Z is any Banach space. Let A^* denote the adjoint of the operator A.

As usual in the study of optimal control problems, we need an associated adjoint problem. In the following theorem we present an existence result for the associated adjoint Cauchy problem.

Theorem 3.5. *Let $B \in C(I, L(X))$, $f \in L_1(I, X^*)$. Then the adjoint problem*:

$$\begin{cases} \dot{\psi} + (A^* + B(t)^*)\psi = -f \\ \psi(T) = 0 \end{cases} \tag{6}$$

has a unique solution $\psi \in C(I, X^*_w)$ in the weak sense.

Proof. By virtue of perturbation theory of semigroups, there exists a strongly continuous transition operator $S(t, \tau)$, $0 \leq \tau < t < \infty$, generated by $(A + B(.))$ (see Theorem 2.4.4 of [1]). Since X is a reflexive Banach space, the adjoint transition operator $S^*(t, \tau)$, $0 \leq \tau < t < \infty$, is also strongly continuous and its generator is given by $A^* + B^*(t)$, $t \geq 0$. Define

$$\psi(t) = \int_t^T S^*(\tau,t)f(\tau)d\tau.$$

Clearly this is the mild solution of equation (6).

For any $h \in D(A)$,

$$\langle \psi(t), h \rangle_{X^*,X} = \int_t^T \langle S^*(\tau,t)f(\tau), h \rangle_{X^*,X} d\tau,$$

and hence

$$\frac{d}{dt}\langle \psi(t), h \rangle_{X^*,X} = -\langle f(t), h \rangle_{X^*,X} - \int_t^T \langle S^*(\tau,t)f(\tau), (A + B(t))h \rangle_{X^*,X} d\tau$$

Thus, for all $h \in D(A)$,

$$\frac{d}{dt}\langle \psi(t), h \rangle_{X^*,X} + \langle \psi(t), (A + B(t))h \rangle_{X^*,X} = -\langle f(t), h \rangle_{X^*,X},$$

for almost $t \in [0, T]$. Clearly $\psi(T) = 0$. Thus, ψ, as defined above, solves the problem (6) also in the weak sense.

4. NECESSARY CONDITIONS OF OPTIMALITY

In this section we present our main results on the necessary conditions of optimality for the problem (P) as stated in section 2. In what follows we shall assume that an optimal control exists (see [4]).

DEFINITION 4.1. For problem (P), suppose $(u_0, \mu_0) \in U_{ad} \times M(\Sigma)$ satisfies

$$\inf_{u \in U_{ad}} \sup_{\mu \in M(\Sigma)} J(u, \mu) = \sup_{\mu \in M(\Sigma)} \inf_{u \in U_{ad}} J(u, \mu) = J(u_0, \mu_0)$$

where $J(u, \mu) = \int_I l(t, x(u, \mu), u)dt$ with $x(u, \mu)$ being the solution of equation (5). We call (u_0, μ_0) optimal (saddle) solution of problem (P).

Since (u_0, μ_0) is an optimal solution, the following system of inequalities must hold:

$$J(u_0, \mu) \le J(u_0, \mu_0) \le J(u, \mu_0) \ \forall\, u \in U_{ad},\ \mu \in M(\Sigma). \tag{7}$$

THEOREM 4.1 *(Necessary Conditions) Suppose the assumptions (A1)–(A4) hold. Then, in order that (u_0, μ_0) be the optimal solution of problem (P), it is necessary that there exists a $\Psi \in C(I, X^*_w)$ such that the following equations and inequalities hold:*

(1) $x_0(t) = T(t)x^0 + \int_0^t T(t - \tau)g(\tau, x_0(\tau), u_0(\tau), \mu_0)d\tau;$

(2) $\dot{\Psi}(t) + (A^* + B^*(t))\Psi(t) = -l_x^0(t),\ \Psi(T) = 0,$

where $B(t) = g_x(t, x_0(t), u_0(t), \mu_0),\ l_x^0(t) = l_x(t, x_0(t), u_0(t));$

(3) $\int_I \langle C^*(t)\Psi(t) + l_u^0(t), u(t) - u_0(t) \rangle_{Y^*,Y} dt \ge 0$ *for all* $u \in U_{ad}$,

where $C(t) = g_u(t, x_0(t), u_0(t), \mu_0)$, $l_u^0(t) = l_u(t, x_0(t), u_0(t))$;

(4) $\int_I \langle \Psi(t), g(t,x_0(t), u_0(t), \mu)\rangle_{X^*,X} dt \leq \int_I \langle \Psi(t), g(t,x_0(t), u_0(t), u_0(t), \mu_0)\rangle_{X^*,X} dt$

for all $\mu \in M(\Sigma)$.

Proof. Let (u_0, μ_0) be an optimal (saddle) solution of problem (P) and $x_0 \equiv x(u_0, \mu_0)$. By convexity of U_{ad}, for any $u \in U_{ad}$, $u^\varepsilon \equiv u_0 + \varepsilon(u - u_0) \in U_{ad}$ for $0 \leq \varepsilon \leq 1$. According to Theorem 3.1, the state equation (5) has a unique mild solution $x^\varepsilon = x(u^\varepsilon, \mu_0)$ corresponding to the control u^ε and parameter μ_0.

Using the second part of the inequality (7), we have

$$\int_I l(t,x^\varepsilon(t), u^\varepsilon(t))dt - \int_I l(t, x_0(t), u_0(t))dt \geq 0. \tag{8}$$

Define $y^\varepsilon = (x^\varepsilon - x_0)/\varepsilon$. Note that y^ε satisfies the integral equation

$$z(t) = \int_0^t T(t-\tau)(\int_0^1 g_x(\tau, x_0(\tau), + s(x^\varepsilon(\tau) - x_0(\tau)), u_0(\tau), \mu_0)ds)z(\tau)d\tau$$
$$+ \int_0^t T(t-\tau)(\int_0^t g_u(\tau, x^\varepsilon(\tau), u_0(\tau) + s(u^\varepsilon(\tau) - u_0(\tau)), \mu_0)ds)(u(\tau) - u_0(\tau))d\tau.$$

By virtue of assumptions (A3) and Theorem 3.2, one can justify taking ε to zero in the above equation to obtain

$$y(t) = \int_0^t T(t-\tau)B(\tau)d\tau + \int_0^t T(t-\tau)C(\tau)(u(\tau) - u_0(\tau))d\tau. \tag{9}$$

The integral equation (9) has a unique solution $y \in C(I, X)$ (see Theorem 2.4.3 of [1]). Hence y is a mild solution of the equation

$$\begin{cases} \dot{y}(t) = Ay + B(t)y + C(t)(u(t) - u_0(t)) \\ y(0) = 0. \end{cases} \tag{10}$$

Note that y is the Gateaux differential of x in the direction $u - u_0$.

By use of hypothesis (A4), and some elementary computations, one obtains from (8) the following inequality

$$\int_I \langle l_X^0, y\rangle_{X^*,X} dt + \langle l_u^0, u - u_0\rangle_{Y^*,Y} dt \geq 0. \tag{11}$$

By virtue of (A4), $l_x^0 \in L_1(I, X^*)$, and hence, by Theorem 3.5, the adjoint equation

$$\begin{cases} \dot{\psi} + (A^* + B^*)\psi = -\, l_x^0 \\ \psi(T) = 0 \end{cases} \tag{12}$$

has a unique weak solution $\psi \in C(I, X_w^*)$.

Since the solution $y(t)$ need not belong to $D(A)$, we use the Yosida approximation of the identity, $I_n = nR(n, A)$ where $R(\lambda, A)$ is the resolvent of A corresponding to $\lambda \in \rho(A)$. It is well known that $I_n \to I$ (Identity operator) in the strong operator topology in $L(X)$, and for any $x \in X$, $I_n x \in D(A)$ for $n \in \rho(A)$. Consider the following equation

$$\begin{cases} \dot{y}_n = Ay_n + (I_n B)y_n + I_n C(u - u_0) \\ y_n(0) = 0. \end{cases} \tag{13}$$

Equation (13) has a unique strong solution y_n with $y_n(t) \in D(A)$ for almost all $t \in I$ provided $n \in \rho(A)$. Note that a strong solution is obviously a mild solution, and, hence

$$y_n(t) = \int_0^t T(t-\tau)I_n B(\tau)y_n(\tau)d\tau + \int_0^t T(t-\tau)I_n C(\tau)(u(\tau) - u_o(\tau)\, d\tau.$$

Using Gronwall inequality it is easy to verify that $y_n \to y$ in the usual topology of $C(I, X)$. By considering ψ as the limit of strong solutions of equation (12) corresponding to Holder continuous approximations of l_x^0 we have

$$\begin{aligned} \int_I \langle l_x^0, y\rangle_{X^*,X} dt &= \lim_{n\to\infty} \int_I \langle l_x^0, y_n\rangle_{X^*,X} dt \\ &= \lim_{n\to\infty} \int_I \langle -\dot{\psi} - (A^* + B^*)\psi, y_n\rangle_{X^*,X} dt \\ &= \lim_{n\to\infty} \int_I \langle \psi, \dot{y}_n - Ay_n - (I_n B)y_n - By_n + I_n By_n\rangle_{X^*,X}\, dt \\ &= \lim_{n\to\infty} \int_I \langle \psi, I_n C(u - u_0) - By_n + I_n By_n\rangle_{X^*,X}\, dt \\ &= \int_I \langle \psi, C(u - u_0)\rangle_{X^*,X}\, dt. \end{aligned} \tag{14}$$

Here we have used the strong convergence of I_n to I and uniform convergence of y_n to y and the following estimate

$$\| I_n By_n - By_n \|_X \le \| I_n By_n - I_n By \|_X + \| I_n By - By \|_X + \| By - By_n \|_X.$$

Combining (11) and (14), we have

$$\int_I \langle \psi, C(u - u_0)\rangle_{X^*,X}\, dt + \int_I \langle l_u^0, u - u_0\rangle_{Y^*,Y}\, dt \ge 0.$$

This proves inequality (3) of the Theorem.

Since (u_0, μ_0) is an optimal solution of the problem (P), it follows from the first part of the inequality (7) that

$$\int_I l(t, x(u_0, \mu), u_0)\, dt - \int_I l(t, x(u_0, \mu_0), u_0)\, dt \le 0 \text{ for all } \mu \in M(\Sigma). \tag{15}$$

For any $\mu \in M(\bar{z})$, let $\mu_\varepsilon \equiv \mu_0 + \varepsilon(\mu - \mu_0) \in M(\Sigma)$, $0 \le \varepsilon \le 1$, and let $x_\varepsilon \equiv x(u_0, \mu_\varepsilon)$ be the unique mild solution of the state equation corresponding to the control u_0 and parameter

μ_ε. Thus it follows from (15) that

$$\int_I l(t, x_\varepsilon, u_0)\,dt - \int_I l(t, x_0, u_0)\,dt \le 0, \qquad \text{for } 0 \le \varepsilon \le 1. \tag{16}$$

Define $\omega^\varepsilon \equiv (x_\varepsilon - x_0)/\varepsilon$. Note that ω^ε satisfies the integral equation

$$\omega^\varepsilon(t) = \int_0^t T(t-\tau)\Big[\int_0^1 g_x(\tau, x_0(\tau) + s(x_\varepsilon(\tau) - x_0(\tau)), u_0(\tau), \mu^\varepsilon)ds\Big]\omega^\varepsilon(\tau)\,d\tau$$
$$+ \int_0^t T(t-\tau)[g(\tau, x_0(\tau), u_0(\tau), \mu) - g(\tau, x_0(\tau), \mu_0)]d\tau.$$

By virtue of assumptions (A3) and Theorem 3.3 one can justify taking ε to zero in the above equation to obtain

$$\omega(t) = \int_0^t T(t-\tau)B(\tau)\omega(\tau)\,d\tau$$
$$+ \int_0^t T(t-\tau)[g(\tau, x_0(\tau), u_0(\tau), \mu) - g(\tau, x_0(\tau), u_0(\tau), \mu_0)]\,d\tau \tag{17}$$

This integral equation has a unique solution $\omega \in C(I, X)$.

Following similar arguments as in the case of control proving inequality (3), we can verify that

$$\int_I \langle \psi, g(t, x_0, u_0, \mu)\rangle_{X^*,X} \le \int_I \langle \psi, g(t, x_0, u_0, \mu_0)\rangle_{X^*,X} dt$$

for all $\mu \in M(\Sigma)$ where ψ satisfies the adjoint equation (2) in the weak sense. This completes the proof of Theorem 4.1.

In the above result we assumed that l is Frechet differentiable in the control variable. In case $l(t, x, u)$ is merely continuous in u and Frechet differentiable in x and $U \subset Y$ is a closed bounded convex set, we can prove Pontryagin type necessary conditions of optimality using well known Ekland's variational principle. Define

$$M \equiv \{u : I \to Y, \text{ strongly measurable: } u(t) \in U \qquad a.e\}$$

with the topology induced by the metric

$$\rho(u, v) \equiv \lambda\{t \in I : u(t) \ne v(t)\},$$

where λ denotes the Lebesgue measure. Since U is a closed subset of a Banach space, the set M, with the metric ρ as defined above, is a complete metric space.

We need the continuous dependence of solutions on control.

LEMMA 4.2 *Suppose the Assumptions (A1) and (A3) hold, with (A3) modified by replacing the Frechet differentiability of g in u by mere continuity and boundedness. Let $U_{ad} = M$. Then the semilinear system (5) has a unique mild solution for every $u \in M$ and $\mu \in$*

$M(\Sigma)$. The mapping $u \to x(u) = x(u, \mu)$ (for fixed μ) is continuous from M to $C(I, X)$ in the respective metric topologies and further there exists a constant β such that

$$\|x(u) - x(v)\|_{C(I,X)} \leq \beta\, \rho(u, v)$$

for all $u, v \in M$.

Proof. The proof of existence of of mild solutions is standard, see [2,8]. Let $x(t; u)$ and $x(t; v)$ denote the solutions corresponding to u and v respectively. let $\sigma \equiv \{t \in I: u(t) \neq v(t)\}$. We have

$$x(t; u) - x(t; v) = \int_0^t T(t - \tau)[g(\tau, x(\tau; u), u(\tau), \mu) - g(\tau, x(\tau; v), v(\tau), \mu]\, d\tau$$

$$= \int_0^t T(t - \tau) \int_0^1 g_x(\tau, x(\tau; v) + s(x(\tau; u) - x(\tau; v)), u(\tau), \mu)\, ds((x(t; u) - x(t; v))\, d\tau$$

$$+ \int_\sigma T(t - \tau)[g(\tau, x(\tau, v), u(\tau), \mu) - g(\tau, x(\tau; v), v(\tau), \mu)]\, d\tau$$

It follows from our assumptions that there exist constants a, b, such that

$$\| x(t; u) - x(t; v)\|_X \leq a \int_o^t \| x(\tau; u) - x(\tau; v)\|_X\, d\tau + b\, \rho\, (u, v).$$

Thus the assertion follows from Gronwall inequality.

THEOREM 4.3 *Suppose the assumptions of Lemma 4.2 hold and further $u \to l(t, x, u)$ is merely continuous and $x \to l(t, x, u)$ is continuously Frechet differentiable with $l_x \in L_1(I, X^*)$. Then the optimality conditions (1), (2), (4) of Theorem 4.1 hold and (3) is replaced by (3)′*

$$(3)' \qquad l(t, x_0(t), u_0(t)) + \langle \psi(t), g(t, x_0(t), u_0(t), \mu_0)\rangle_{Y^*,Y}$$
$$\leq l(t, x_0(t), v) + \langle \psi(t), g(t, x_0(t), v, \mu_0)\rangle_{Y^*,Y}$$

for all $v \in U$ and almost all $t \in I$.

Proof. Since (u_0, μ_0) is optimal, again by the second part of the inequality (7), we have

$$\int_I l(t, x(t; u), u)\, dt - \int_I l(t, x_0, u_0)\, dt \geq 0 \;\; \forall \;\; u \in M.$$

For any measurable set $\sigma \subset I$ and $v \in U$, define

$$u^\sigma(t) = \begin{cases} u_0(t) & t \in I \backslash \sigma \\ v & t \in \sigma. \end{cases}$$

Let x^σ be the solution of the system (5) corresponding to u^σ and μ_0.

Then

$$\int_I l(t, x^\sigma, u^\sigma)\, dt - \int_I l(t, x_0, u_0)\, dt = \int_\sigma l(t, x^\sigma, v)\, dt - \int_\sigma l(t, x_0, u_0)\, dt$$
$$+ \int_{I\setminus\sigma} [l(t, x^\sigma, u_0) - l(t, x_0, u_0)]\, dt \geq 0 \tag{18}$$

By virtue of Frechet differentiability of l, we have

$$\int_{I\setminus\sigma} [l(t, x^\sigma, u_0) - l(t, x_0, u_0)]\, dt = \int_{I\setminus\sigma} \langle l_x (t, x_0, u_0), x^\sigma - x_0\rangle\, dt + o(\lambda(\sigma))$$
$$= \int_I \langle l_x (t, x_0, u_0), x^\sigma - x_0\rangle\, dt + o(\lambda(\sigma)),$$

where o(.) stands for small order of approximation.

Hence expression (18) reduces to

$$\int_\sigma l(t, x_0, u_0)\, dt \leq \int_\sigma l(t, x^\sigma, u_0) + \int_I \langle l_x(t, x_0, u_0), x^\sigma - x_0\rangle\, dt + o(\lambda(\sigma)). \tag{19}$$

Using the adjoint equation (2) of Theorem 4.1 and following similar arguments as in that Theorem, one can verify that

$$\int_I \langle l_x (t, x_0, u_0), x^\sigma - x_0\rangle\, dt$$
$$= \int_I \langle \psi, g(t, x^\sigma, u^\sigma, \mu_0) - g(t, x_0, u_0, \mu_0) - B(x^\sigma - x_0)\rangle dt$$
$$= \int_I \langle \psi, g(t, x_0, u^\sigma, \mu_0) - g(t, x_0, u_0, \mu_0)\rangle\, dt + o(\lambda(\sigma))$$
$$= \int_\sigma \langle \psi, g(t, x_0, u, \mu_0)\rangle\, dt - \int_\sigma \langle \psi, g(t, x_0, u_0, \mu_0)\rangle\, dt + o(\lambda(\sigma)).$$

Thus the expression (19) reduces to

$$\int_\sigma l(t, x_0, u_0)\, dt + \int_\sigma \langle \psi, g(t, x_0, u_0, \mu_0)\rangle\, dt$$
$$\leq \int_\sigma l(t, x_0, v)\, dt + \int_\sigma \langle \psi, g(t, x_0, v, \mu_0)\rangle\, dt + 0\, (\lambda(\sigma)) \tag{20}$$

Let t be any Lebesgue density point of u_0 and σ any measurable set containing t shrinking to the one point set $\{t\}$ as $\lambda(\sigma) \rightarrow 0$. Dividing (20) by $\lambda(\sigma)$ and letting it converge to zero, we obtain the inequality (3)$'$. This completes the proof.

Remark 4.4. In case $-A$ is the generator of an analytic semigroup, the assumptions on g can be considerably relaxed admitting unbounded nonlinear operators which are relatively bounded with respect to fractional powers of A. In that situation the preceding results can be improved as in [3, Theorem 14, p. 401].

Remark 4.5. All the results presented here do also hold for time varying operators $A(t)$, $t \geq 0$ generating bounded evolution operators $U(t, s)$, $0 \leq s \leq t < \infty$, in X.

5. COMPUTATIONAL ALGORITHM

Based on the necessary conditions of optimality given by Theorem 4.1, we can construct an algorithm for computing the optimal solution of the problem (P). For this purpose, we require the duality maps.

Since Y is a reflexive Banach space, $L_q(I, Y^*) \equiv L_q(Y^*)$ is the dual of $L_p(Y)$ $(1 < p, q < \infty)$. The map ν_1: $L_q(Y^*) \to L_p(Y)$ denotes the duality map, that is, for $\xi \in L_q(Y^*)$

$$\nu_1(\xi) \equiv \{\eta \in L_p(Y) : \langle \xi, \eta \rangle_{L_q(Y^*), L_p(Y)} = \| \xi \|^2_{L_q(Y^*)} = \| \eta \|^2_{L_p(Y)} \}.$$

If Y is strictly convex then $L_p(Y)$ is also strictly convex and the duality map ν_1 is uniquely defined. Since Σ is a compact Polish space, the dual of $C(\Sigma)$ is given by $M_{rca}(\Sigma)$ which is the space of regular countably additive bounded measures on Σ. Thus we may define the duality map $\nu_2 : C(\Sigma) \to M_{rca}(\Sigma)$ such that for any $\zeta \in C(\Sigma)$

$$\nu_2(\zeta) \equiv \{\mu \in M_{rca}(\Sigma) : \langle \zeta, \mu \rangle_{C(\Sigma), M_{rca}(\Sigma)} = \| \zeta \|^2_{C(\Sigma)} \}.$$

For Frechet differentiable l, we can define $D_u J(.) \equiv (C^* \psi_0 + \ell_u^0)(.)$. The inequality (3) of Theorem 4.1 is then equivalent to following inequality

$$J_u'(u_0, \mu_0; u - u_0) \equiv \langle D_u J, u - u_0 \rangle = \int_I \langle C^* \psi_0 + \ell_{u}^0, u - u_0 \rangle_{Y^*, Y} \, dt \geq 0$$

for all $u \in U_{ad}$.

Similarly, the Frechet differential of J with respect to μ is given by

$$D_\mu J(\sigma) \equiv \int_I \langle \psi_0(t), \tilde{g}(t, x_0(t), u_0(t), \sigma) \rangle_{X^*, X} \, dt.$$

Then the inequality (4) of Theorem 4.1 is equivalent to following inequality

$$J_\mu'(u_0, \mu_0; \mu - \mu_0) = \langle D_\mu J(u_0, \mu_0), \mu - \mu_0 \rangle_{C(\Sigma), M_{rca}(\Sigma)} \leq 0.$$

Now the algorithm may be stated as follows:

Algorithm 5.1:

Step 1. Suppose that at the nth stage, the control and the measure (parameter), is given by $\{u^n, \mu^n\}$, $u^n \in L_p(Y)$, $\mu^n \in M(\Sigma) \subset M_{rca}(\Sigma)$.

Step 2. Use $\{u^n, \mu^n\}$ to determine $\{x^n, \psi^n\}$ where x^n is the solution of equation (5) corresponding to $\{u^n, \mu^n\}$ and ψ^n is the solution of the adjoint equation (2) of Theorem 4.1.

Step 3. Compute $D_u J(u^n, \mu^n)$ and $D_\mu J(u^n, \mu^n)$.

Step 4. Define

$$u^{n+1} \equiv u^n - \varepsilon\tilde{u}_n, \qquad \tilde{u}_n \,\varepsilon\, v_1(D_u J(u^n, \mu^n)), \qquad \notin > 0,$$
$$u^{n+1} \equiv \mu^n + \gamma\tilde{\mu}_n, \qquad \tilde{u}_n \,\varepsilon\, v_2(D_\mu J(u^n, \mu^n)), \qquad \gamma > 0,$$

choosing ε, γ sufficiently small so that

$$\begin{aligned} J(u^{n+1}, \mu^n) &\equiv J(u^n, \mu^n) - \varepsilon\langle D_u J(u^n, \mu^n), \tilde{u}^n\rangle + o(\varepsilon) \\ &= J(u^n, \mu^n) - \varepsilon \, \| D_u J(u^n, \mu^n) \|^2_{L_q(Y^*)} + o(\varepsilon) \leq J(u^n, \mu^n) \\ J(u^{n+1}, \mu^n) &\equiv J(u^n, \mu^{n+1}) - \gamma\langle D_u J(u^n \mu^n), \tilde{\mu}^n\rangle + o(\gamma) \\ &= J(u^n, \mu^{n+1}) - \gamma \, \| D_\mu J(u^n, \mu^n) \|^2_{C(\Sigma)} + o(\gamma) \leq J(u^n, \mu^{n+1}). \end{aligned}$$

Step 5. Solve the state equation corresponding to the pair $\{u^{n+1}, \mu^{n+1}\}$ and compute $J(u^{n+1}, \mu^{n+1})$ using the following expression:

$$J(u^{n+1}, \mu^{n+1}) = \int_I l(t, x^{n+1}(t), u^{n+1}(t))\, dt.$$

If $| J(u^{n+1}, \mu^{n+1}) - J(u^n, \mu^n) | < \delta$ for some preassigned small positivie number δ, stop; otherwise go back to step 2 with new control u^{n+1} and the new measure μ^{n+1}

Remark 5.2. If $U_{ad} = \{u \in L_p(Y) : u(t) \in U(t)\ a.e\}$, $U(t) \subseteq \mathcal{U}$(see Assumption (A2)), then in Step 4

$$u^{n+1} = P_u(u^n - \varepsilon\tilde{u}_n), \ \tilde{u}_n \in v_1(D_u J(u^n, \mu^n)),$$

where P_u is the projection operator from Y to $\mathcal{U}$.

References

1. N. U. Ahmed and K. L. Teo, *Optimal control of distributed parameter systems*, North Holland, Amsterdam, 1981.
2. N. U. Ahmed, Semigroup theory with applications to systems and control, *Pitman Research Notes in Math Series* ***246***(1991).
3. N. U. Ahmed, *Optimal control of infinite-dimensional systems governed by integro differential equations*: *Differential equations, dynamical systems and control science* (A Festschrift in Honor of Lawrence Markus) Marcel Dekker, New York, ***152***(1994) 383–402.
4. N. U. Ahmed and X. Xiang, Optimal control of infinite dimensional uncertain systems, *J. Optimization Theory and Applications*, ***80***(1994) 261–272.
5. N. U. Ahmed, *Optimization and identification of systems governed by evolution equations on Banach space*, *Pitman Research Notes in Math. Series* ***184***(1988).
6. J. Warga, *Optimal control of differential and functional equations*, Academic Press, New York/London (1972).
7. E. P. Avgerinos and N. S. Papageorgiou, Optimal control and relaxation for a class of nonlinear distributed parameter systems, *Osaka J. Math.*, ***27***(1990) 745–767.
8. A. Pazy, *Semigroup of linear operators and applications to partial differential equations*, Springer-Verlag, Heidelberg/Berlin/New York (1983).