JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **120**, 169–210 (1986)

Optimal Control of Linear Retarded Systems in Banach Spaces

SHIN-ICHI NAKAGIRI

Department of Applied Mathematics, Faculty of Engineering, Kobe University, 657, Kobe, Nada, Japan

Submitted by George Leitmann

This paper deals with standard optimal control problems, namely, the fixed time integral convex cost problem and the time optimal control problem for linear retarded systems in Banach spaces. For the basis of optimal control theory the fundamental solution is constructed and a variation of constant formula of (mild) solutions is established. After the controlled system description and the formulation of optimal control problems are given, the retarded adjoint system is introduced. For the integral convex cost problem two existence theorems of optimal controls and necessary conditions of optimality are given. These conditions are precisely characterized by the solution of retarded adjoint system. The "pointwise" maximum principle for time varying control domain is derived from the optimality conditions. The bang-bang principle is also established for the terminal value cost problem under some regularity condition of the adjoint system. For the time optimal control problem to a target set an existence theorem is shown. In the case where the target set has interior, the maximum principle and the bang-bang principle are established for the time optimal control. Finally, a convergence theorem of time optimal controls to a point target set is given. This paper also contains illustrative examples which give technologically important control problems.

1. INTRODUCTION

There exists a considerable literature which studies optimal control problems of control systems in infinite-dimensional spaces (see the books [1, 14, 26] for results and extensive references cited therein). Most studies have been devoted to the systems without time delay, and the papers treating the systems with delay are not many [2, 27, 28, 34, 35]. Furthermore in the above literature the continuous retardation effect is not in consideration and the concept of fundamental solution (or Green function) is not used, so that the calculations to obtain the existence and optimality conditions of optimal controls are complicated. On the other hand, optimal control theory of retarded systems in finite-dimensional space is widely

0022-247X/86 $3.00

developed in many references [5–8, 10–13, 15, 16, 19, 22, 31, 36] and at a satisfactory level. This paper is intended to fill the gap existing between finite- and infinite-dimensional retarded systems.

In this paper we study some standard optimal control problems, namely, the fixed time integral convex cost problem and the time optimal control problem for general linear retarded systems in reflexive Banach spaces. We shall present results on the existence of optimal controls, necessary conditions of optimality, maximum principle, and bang-bang principle for the optimal control problems. For the basis of our study we construct the fundamental solution and establish a variation of constant formula of (mild) solutions for the free system. We introduce the retarded adjoint system and give the representation formula of adjoint states in terms of the fundamental solution. Our treatise to solve the control problems is based on the formula and the adjoint system plays a central role in the description of optimality conditions as developed in [26]. However, unlike in [26] the method of integration by parts is not used; instead the simpler Fubini theorem is used in this paper.

This paper is devided into eight sections plus appendices. Section 2 gives the background of our optimal control theory. The notations and terminology to be used in the paper are given in Subsection 2.1; for the free system the existence, uniqueness, and a variation of constant formula for mild solutions are given in Subsection 2.2; some further results on the existence of strong and weak solutions are also given in Subsection 2.2. In Section 3, we give the controlled system description and the formulation of the optimal control problems to be investigated and introduce the retarded adjoint system. The purpose of Section 3 is to establish the representation formula for the adjoint state. Two existence theorems of optimal controls are given in Section 4; one is for bounded control set and the other is for unbounded control set. In Section 5, we present the necessary conditions of optimality which are described by the adjoint state and integral inequality. Two applications of the main theorem (Theorem 5.1) are given; one gives a feedback control law for the regulator problem and the other gives a uniqueness of the optimal control of the averaging observation control problem. Section 6 is devoted to studying the "pointwise" maximum principle. The maximum principle for time varying control domain is derived from the optimality conditions in Section 5 by the variational technique. Some examples of the maximum principle for technologically important costs are also given in Section 6. In Section 7, the bang-bang principle for terminal value problem and its applications to uniqueness and expression of the optimal control are given under some regularity condition of the adjoint system. Section 8 deals with the time optimal control problem to a target set. Under very general conditions on the target set and the controlled system an existence theorem of the time optimal control is given. In the case where the target set has nonempty interior, the maximum principle and the bang-band principle are established with some examples. A convergence theorem of time optimal controls to a point target set is also given in Section 8. Appendices collect proofs of some results which are needed in our optimal control theory.

2. Fundamental Theorems on Linear Functional Differential Equations

2.1. *Notation and Terminology*

First we give the notations and terminology used throughout this paper. Let R be the set of real numbers and let R^+ be the set of non-negative numbers. Let X and Y be real (separable) Banach spaces with norms $|\cdot|$ and $|\cdot|_Y$, respectively. The adjoint spaces of X, Y are denoted by X^*, Y^* and their norms are denoted by $|\cdot|_*$ and $|\cdot|_{Y^*}$. For a densely defined closed linear operator A on X, its adjoint operator on X^* is denoted by A^*. We write the duality pairing between X and X^* by $<, >$ and the pairing between Y and Y^* by $<, >_{Y,Y^*}$. Let $\mathscr{L}(X, Y)$ be the Banach space of bounded linear operators from X into Y. When $X = Y$, $\mathscr{L}(X, Y)$ is denoted by $\mathscr{L}(X)$. Their operator norms are denoted by $\|\cdot\|$.

Given an interval $I \subset R$, we denote by $L_p(I; X)$ and $C(I: X)$ the usual Banach space of X-valued measurable functions which are p-Bochner integrable $(1 \leqslant p < \infty)$ or essentially bounded $(p = \infty)$ on I and the Banach space of strongly continuous functions on I, respectively. The norm of $L_p(I; X)$ is denoted by $\|\cdot\|_{p,I}$. For each integer $k \geqslant 1$, $W_p^{(k)}(I; X)$ denotes the Sobolev space of X-valued measurable functions x on I such that x and its distributional derivatives up to order k belong to $L_p(I; X)$. $L_p^{\text{loc}}(R^+; X)$ (resp. $C(R^+; X)$) will denote the Frèchet space of functions which belongs to $L_p([0, T]; X)$ (resp. $C([0, T]; X)$) for any $T > 0$. Let $M_p(I; X)$ denote the product Banach space $X \times L_p(I; X)$ with norm

$$\|g\|_{M_p(I;X)} = \begin{cases} (|g^0|^p + \|g^1\|_{p,I}^p)^{1/p} & \text{if } \ 1 \leqslant p < \infty, \\ |g^0| + \|g^1\|_{\infty,I} & \text{if } \ p = \infty, \end{cases} \quad g = (g^0, g^1) \in M_p(I; X).$$

The function χ_I means the characteristic function of the interval I. For a measurable function $x: R^+ \to X$, its Laplace transform $\hat{x}$ is defined by $\hat{x}(\lambda) = \int_{R^+} e^{-\lambda t}\, x(t)\, dt$, whenever the integral exists. If x is measurable and satisfies $|x(t)| \leqslant M e^{\rho t}$, $t \in R^+$ for some $M > 0$ and $\rho \in R$, then $\hat{x}(\lambda)$ can be defined in $\operatorname{Re} \lambda > \rho$ and is analytic there.

2.2. Existence, Uniqueness, and a Variation of Constant Formula for Mild Solutions

In this subsection we present some basic results on existence, uniqueness, and a representation formula of (mild) solutions for linear functional differential equations in Banach spaces. Let $h>0$ be fixed and $I_h=[-h,0]$. For notational brevity we write the space $M_p(I_h; X)$ by M_p. Consider the following free (or non-controlled) system which is described by a linear functional differential equation on X:

$$(\mathrm{E})\quad \begin{cases} \dfrac{dx(t)}{dt}=A_0x(t)+\displaystyle\int_{-h}^{0} d\eta(s)\,x(s+t)+f(t) \qquad \text{a.e.}\quad t>0 & (2.1)\\ x(0)=g^0, \qquad x(s)=g^1(s) \qquad \text{a.e.}\quad s\in[-h,0), & (2.2)\end{cases}$$

where $g=(g^0,g^1)\in M_p, f\in L_q^{\mathrm{loc}}(R^+;X)$, p, $q\in[1,\infty]$, A_0 generates a strongly continuous semigroup $\{T(t);\ t\geqslant 0\}$ on X, and η is the Stieltjes measure given by

$$\eta(s)=-\sum_{r=1}^{m}\chi_{(-\infty,\,-h_r]}(s)\,A_r-\int_s^0 D(\xi)\,d\xi, \qquad s\in I_h. \tag{2.3}$$

Here in (2.3) it is assumed that $0\leqslant h_1<\cdots<h_m\leqslant h$ are non-negative constants, $A_r\in\mathscr{L}(X)$ $(r=1,\dots,m)$, and $D(\cdot)\in L_1(I_h;\mathscr{L}(X))$. Then the delayed term $\int_{I_h}d\eta(s)\,x(s+t)$ is written by

$$\sum_{r=1}^{m}A_rx(t-h_r)+\int_{-h}^{0}D(s)\,x(t+s)\,ds.$$

The integral kernel $D(s)$ in η is assumed to satisfy

$$H_0^p\colon D(\cdot)\in L_{p'}(I_h;\mathscr{L}(X)), \qquad 1/p+1/p'=1.$$

Instead of (E) we consider the following functional integral equation:

(IE) $x(t)$

$$=\begin{cases} T(t)\,g^0+\displaystyle\int_0^t T(t-s)f(s)\,ds+\int_0^t T(t-s)\int_{-h}^{0}d\eta(\xi)\,x(s+\xi)\,ds, & t\geqslant 0\\ g^1(t) \qquad \text{a.e. } t\in[-h,0). & (2.4)\end{cases}$$

THEOREM 2.1. *Let $g=(g^0,g^1)\in M_p, f\in L_q^{\mathrm{loc}}(R^+;X)$, $1\leqslant p,q\leqslant\infty$, and the assumption H_0^p be satisfied. Then there exists a unique solution $x(t)=x(t;f,g)$, $t\in[-h,\infty)$ of* (IE) *which satisfies*

(i) $x(\cdot) \in C(R^+; X)$; (2.5)

(ii) $|x(t; f, g)| \leqslant (M_0 \|g\|_{M_p} + M_1 \|f(\cdot)\|_{q,[0,t]}) e^{\gamma_0 t}, \qquad t \geqslant 0,$ (2.6)

where M_0, M_1, γ_0 are constants depending only on p, q, η, and A_0.

Proof. We shall show the existence and uniqueness of a solution of (IE) by the contraction mapping theorem. Let $b > 0$ be fixed and define the mapping $\mathscr{K}: L_p([-h, b]; X) \to L_p([-h, b]; X)$ by $(\mathscr{K}x)(t) =$ the right-hand side of (2.4) a.e. $t \in [-h, b]$ for each $x \in L_p([-h, b]; X)$. First we shall show that $\mathscr{K}$ is into. Relating to the term $\int_{-h}^{0} d\eta(\xi)\, x(s+\xi)$ in (2.4) we define the operator $E_b: L_p([-h, b]; X) \to L_p([0, b]; X)$ by

$$(E_b x)(s) = \int_{-h}^{0} d\eta(\xi)\, x(s+\xi) \qquad \text{a.e. } s \in [0, b]. \tag{2.7}$$

Using Hölder inequality and H_0^p, we obtain

$$\begin{aligned} \|E_b x\|_{p,[0,b]} &\leqslant \sum_{r=1}^{m} \left(\int_0^b \|A_r\|^p\, |x(t-h_r)|^p\, dt \right)^{1/p} \\ &\quad + \left(\int_0^b \left(\int_{-h}^{0} \|D(\xi)\| \cdot |x(t+\xi)|\, d\xi \right)^p dt \right)^{1/p} \\ &\leqslant \left(\sum_{r=1}^{m} \|A_r\| + \|D(\cdot)\|_{p',I_h} \cdot b^{1/p} \right) \|x(\cdot)\|_{p,[-h,b]}, \end{aligned} \tag{2.8}$$

for $1 \leqslant p < \infty$, where $1/p + 1/p' = 1$. It is easy to see that the inequality (2.8) is also true for $p = \infty$. Hence E_b is bounded and

$$\|E_b\| \leqslant \left(\sum_{r=1}^{m} \|A_r\| + b^{1/p} \cdot \|D(\cdot)\|_{p',I_h} \right). \tag{2.9}$$

Thus $\int_{-h}^{0} d\eta(\xi)\, x(\cdot+\xi) \in L_p([0, b]; X)$, and hence it is verified by assumption that $\mathscr{K}x \in C([0, b]; X) \cap L_p([-h, b]; X)$. That is, $\mathscr{K}$ is into. We next show that $\mathscr{K}$ is a contraction mapping for b small. Let x, $y \in L_p([-h, b]; X)$. Since $T(t)$ is a C_0-semigroup, there exist $M > 0$ and $\omega > 0$ such that

$$\|T(t)\| \leqslant M e^{\omega t}, \qquad t \geqslant 0. \tag{2.10}$$

Then from (2.9) and (2.10) it follows by using Hölder inequality that

$$\begin{aligned} \|\mathscr{K}x - \mathscr{K}y\|_{p,[-h,b]} &\leqslant \left(\int_0^b \left(\int_0^t \|T(t-s)\|\, |E_b(x-y)(s)|\, ds \right)^p dt \right)^{1/p} \\ &\leqslant M \left\{ \frac{1}{p'\omega} (e^{p'\omega b} - 1) \right\}^{1/p'} \cdot b^{1/p} \cdot \|E_b\| \cdot \|x-y\|_{p,[-h,b]}. \end{aligned} \tag{2.11}$$

In (2.11) if $p \neq \infty$, $b^{1/p} \to 0$ as $b \to 0$ and if $p = \infty$, $e^{\omega b} - 1 \to 0$ as $b \to 0$. Hence $\mathscr{K}$ is a contraction for sufficiently small $b > 0$. This proves the local existence and uniqueness of the solution of (IE). To prove the global existence, we derive an a priori estimate of this solution. Let $x(t)$ be a solution of (IE) on the interval $[-h, a], a \geqslant h$. Then $E_a x = \int_{-h}^{0} d\eta(\xi)\, x(\cdot + \xi) \in L_p([0, a]; X)$ can be written by

$$(E_a x)(s) = \begin{cases} \displaystyle\int_{-h}^{-s-0} d\eta(\xi)\, g^1(s+\xi) + \int_{-s}^{0} d\eta(\xi)\, x(s+\xi) & \text{a.e. } s \in [0, h] \\ \displaystyle\int_{-h}^{0} d\eta(\xi)\, x(s+\xi), & s \in [h, a]. \end{cases} \tag{2.12}$$

We put

$$\begin{aligned} (E_a^1 x)(s) &= \int_{-h}^{-s-0} d\eta(\xi)\, g^1(s+\xi) \\ &= \sum_{r=1}^{m} A_r\, \chi_{(-h_r, 0]}(-s)\, g^1(s-h_r) + \int_{-h}^{-s} D(\xi)\, g^1(s+\xi)\, d\xi \\ &\qquad \text{a.e. } s \in [0, h]. \end{aligned} \tag{2.13}$$

Applying similar calculations as in (2.7) to (2.13), we have

$$\|E_a^1 x\|_{p, [0, h]} \leqslant \|E_a\| \cdot \|g^1(\cdot)\|_{p, I_h}. \tag{2.14}$$

Since $x(t)$ is continuous for $t \geqslant 0$, we see easily that

$$|(E_a x - \chi_{[0, h]}\, E_a^1 x)(s)| \leqslant (\operatorname{Var} \eta) \sup_{0 \leqslant \xi \leqslant s} |x(\xi)|, \qquad s \in [0, a], \tag{2.15}$$

where $\operatorname{Var} \eta = \sum_{r=1}^{m} \|A_r\| + \int_{-h}^{0} \|D(s)\|\, ds$. Then, making use of Hölder inequality several times, we can obtain from (2.4), (2.14), and (2.15) that

$$\begin{aligned} |x(t)| &\leqslant \|T(t)\| \cdot |g^0| + \|T(\cdot)\|_{q', [0, t]} \cdot \|f(\cdot)\|_{q, [0, t]} \\ &\quad + \int_0^t \|T(t-s)\| \cdot |E_a x(s)|\, ds \\ &\leqslant M e^{\omega t}\, |g^0| + M_1\, e^{\omega t}\, \|f(\cdot)\|_{q, [0, t]} + M_2\, e^{\omega t}\, \|E_a\| \cdot \|g^1(\cdot)\|_{p, I_h} \\ &\quad + M (\operatorname{Var} \eta) \int_0^t e^{\omega(t-s)} \left(\sup_{0 \leqslant \xi \leqslant s} |x(\xi)| \right) ds, \qquad t \in [0, a], \end{aligned} \tag{2.16}$$

where $1/q+1/q'=1$ and M_1, M_2 are some positive constants. We now apply Gronwall's inequality to (2.16) and obtain

$$|x(t)| \leqslant (M_0 \|g\|_{M_p} + M_1 \|f(\cdot)\|_{q,[0,t]})\, e^{\gamma_0 t}, \qquad t \in [0, a], \qquad (2.17)$$

where $\gamma_0 = \omega + M(\operatorname{Var} \eta)$ and $M_0, M_1 > 0$. Since $a \geqslant h$ can be chosen arbitrarily large, the global existence of the solution with the estimate (ii) is proved.

Remark 2.1. Let the mapping $\mathscr{S}\colon M_p \times L_q^{\mathrm{loc}}(R^+; X) \to C(R^+; X)$ be defined by $\mathscr{S}(g,f)(t) = x(t;f,g)$, $t \geqslant 0$. Then Theorem 2.1 says that $\mathscr{S}$ is linear and continuous. The estimate (2.6) permits us to define the Laplace transform $\hat{x}(\lambda;f,g)$ of $x(t;f,g)$ if $f \in L_q(R^+; X)$.

The solution $x(t;f,g)$ is called a mild solution of (E). We now define the fundamental solution $G(t)$ of (E) by

$$G(t)\, g^0 = \begin{cases} x(t; 0, (g^0, 0)), & t \geqslant 0 \\ 0, & t < 0 \end{cases} \qquad \text{for} \quad g^0 \in X. \qquad (2.18)$$

It is easily checked that under the condition $\int_{-h}^0 \|D(s)\|\, ds < \infty$ the fundamental solution $G(t)$ can be constructed. The definition (2.18) implies that $G(t)$ is a unique solution of

$$G(t) = \begin{cases} T(t) + \int_0^t T(t-s) \int_{-h}^0 d\eta(\xi)\, G(\xi+s)\, ds, & t \geqslant 0 \\ 0, & t < 0 \end{cases} \qquad (2.19)$$

in $\mathscr{L}(X)$, where 0 is the null operator on X. By virtue of Theorem 2.1, $G(t)$ is strongly continuous on R^+ and satisfies

$$\|G(t)\| \leqslant M e^{(\omega + M \cdot \operatorname{Var} \eta) t}, \qquad t \geqslant 0. \qquad (2.20)$$

The main theorem in this section is the following variation of constant formula for mild solutions of (E).

THEOREM 2.2. *Let $g = (g^0, g^1) \in M_p, f \in L_q^{\mathrm{loc}}(R^+; X)$, $1 \leqslant p, q \leqslant \infty$, and H_0^p be satisfied. Then the mild solution $x(t) = x(t;f,g)$ of* (E) *is represented by*

$$x(t) = G(t)\, g^0 + \int_{-h}^0 U_t(s)\, g^1(s)\, ds + \int_0^t G(t-s) f(s)\, ds, \qquad t \geqslant 0, \qquad (2.21)$$

where

$$U_t(s) = \sum_{r=1}^{m} G(t-s-h_r)\, A_r \chi_{[-h_r, 0]}(s)$$
$$+ \int_{-h}^{s} G(t-s+\xi)\, D(\xi)\, d\xi, \qquad s \in I_h. \tag{2.22}$$

Proof. We shall prove (2.21) by assuming $f \in L_q(R^+; X) \cap L_1(R^+; X)$. Put $y(t) =$ the right-hand side of (2.21) for $t \geqslant 0$ and $y(t) = g^1(t)$ for a.e. $t \in [-h, 0)$. Since H_0^p is satisfied, $U_t(s)$ in (2.22) belongs to $L_{p'}(I_h; \mathscr{L}(X))$, $1/p + 1/p' = 1$. Hence by (2.20), $y(\cdot) \in C(R^+; X) \cap L_p(I_h; X)$. It is possible to prove, by using Hölder inequality and (2.20), that $y(t)$ satisfies the similar inequality as in (2.6), the Laplace transform $\hat{y}(\lambda)$ of $y(t)$ can be defined for $\operatorname{Re} \lambda > \gamma_0$. We transform $y(t)$ by Fubini's theorem as

$$y(t) = G(t)\, g^0 + \int_0^t G(t-s) \left(\sum_{r=1}^{m} \chi_{[0, h_r]}(s)\, A_r g^1(s-h_r) \right) ds$$
$$+ \int_0^t G(t-s)\, \chi_{[0,h]}(s) \left(\int_{-h}^{-s} G(\xi)\, g^1(\xi+s)\, d\xi \right) ds$$
$$+ \int_0^t G(t-s) f(s)\, ds. \tag{2.23}$$

Since $f \in L_1(R^+; X)$ and $G(t)$ satisfies (2.20), we have by applying the convolution theorem on Laplace transforms to (2.23) that

$$\hat{y}(\lambda) = \hat{G}(\lambda)\, g^0 + \hat{G}(\lambda) \left(\sum_{r=1}^{m} A_r e^{-\lambda h_r} \int_{-h_r}^{0} e^{-\lambda s} g^1(s)\, ds \right)$$
$$+ \hat{G}(\lambda) \left(\int_{-h}^{0} e^{\lambda\xi} D(\xi) \int_{\xi}^{0} e^{-\lambda s} g^1(s)\, ds\, d\xi \right) + \hat{G}(\lambda) \hat{f}(\lambda)$$

(Fubini's theorem)

$$\equiv \hat{G}(\lambda)(g^0 + F(\lambda; g^1) + \hat{f}(\lambda)), \tag{2.24}$$

where $\hat{G}(\lambda)$ and $\hat{f}(\lambda)$ denote the Laplace transforms of $G(t)$ and $f(t)$, respectively. On the other hand, since $x(t)$ satisfies (2.4), the Laplace transform $\hat{x}(\lambda)$ of $x(t)$ is given by

$$\hat{x}(\lambda) = R(\lambda; A_0)(g^0 + \hat{f}(\lambda) + \widehat{Ex}(\lambda)), \tag{2.25}$$

where $R(\lambda; A_0)$ is the resolvent of A_0 and $\widehat{Ex}(\lambda)$ is given by

$$\widehat{Ex}(\lambda) = \int_0^{\infty} e^{-\lambda t} \int_{-h}^{0} d\eta(\xi)\, x(t+\xi)\, dt.$$

Noticing that $x(t)=g^1(t)$ a.e. $t\in[-h,0)$, we use Fubini's theorem again to obtain

$$\begin{aligned}\widehat{Ex}(\lambda)&=\sum_{r=1}^{m} A_r e^{-\lambda h_r}\int_{-h_r}^{0} e^{-\lambda s}g^1(s)\,ds+\int_{-h}^{0} e^{\lambda\xi}D(\xi)\int_{\xi}^{0} e^{-\lambda s}g^1(s)\,ds\,d\xi\\&\quad+\left(\sum_{r=1}^{m} A_r e^{-\lambda h_r}+\int_{-h}^{0} e^{\lambda s}D(s)\,ds\right)\hat{x}(\lambda)\\&\equiv F(\lambda;g^1)+\left(\int_{-h}^{0} e^{\lambda s}\,d\eta(s)\right)\hat{x}(\lambda).\end{aligned}\tag{2.26}$$

Then by (2.25) and (2.26),

$$\left[I-R(\lambda;A_0)\int_{-h}^{0} e^{\lambda s}\,d\eta(s)\right]\hat{x}(\lambda)=R(\lambda;A_0)(g^0+F(\lambda;g^1)+\hat{f}(\lambda)).\tag{2.27}$$

Now we see for Re $\lambda>\gamma_0$,

$$\left\|R(\lambda;A_0)\int_{-h}^{0} e^{\lambda s}\,d\eta(s)\right\|\leqslant\frac{M}{\operatorname{Re}\lambda-\omega}\cdot\operatorname{Var}\eta<1,$$

and hence by (2.27),

$$\hat{x}(\lambda)=\left[I-R(\lambda;A_0)\int_{-h}^{0} e^{\lambda s}d\eta(s)\right]^{-1}R(\lambda;A_0)(g^0+F(\lambda;g^1)+\hat{f}(\lambda)).\tag{2.28}$$

The Laplace transform of (2.19) yields

$$\hat{G}(\lambda)=R(\lambda;A_0)+R(\lambda;A_0)\int_{-h}^{0} e^{\lambda s}\,d\eta(s)\cdot\hat{G}(\lambda),$$

so that

$$\hat{G}(\lambda)=\left[I-R(\lambda;A_0)\int_{-h}^{0} e^{\lambda s}\,d\eta(s)\right]^{-1}R(\lambda;A_0).\tag{2.29}$$

Therefore, from (2.24), (2.28), and (2.29) it follows that

$$\hat{x}(\lambda)=\hat{G}(\lambda)(g^0+F(\lambda;A_0)+\hat{f}(\lambda))=\hat{y}(\lambda)\qquad\text{for Re }\lambda>\gamma_0.$$

By the uniqueness of Laplace transforms [16, p. 626] and the strong continuity of $x(t)$ and $y(t)$ on R^+, we obtain

$$x(t)=y(t)\qquad\text{for all }t\in R^+,$$

which proves (2.21). Lastly we shall prove (2.21) without assuming $f \in L_q(R^+; X) \cap L_1(R^+; X)$. For this it sufficies to prove (2.21) for $t \in [0, L]$ with any fixed $L > 0$. For a given $f \in L_q^{\mathrm{loc}}(R^+; X)$ and $L > 0$ we define the truncated function $f_L(t)$ by $f_L(t) = \chi_{[0, L]}(t) f(t)$. Then $f_L \in L_q(R^+; X) \cap L_1(R^+; X)$ and the corresponding solution $x_L(t)$ of (2.4) satisfies (2.21) for all $t \geqslant 0$. Since $x_L(t) = x(t)$ for $t \in [0, L]$, (2.21) is true for all $t \in [0, L]$. This finishes the proof.

Remark 2.2. Assume $D(\cdot) \in C(I_h; \mathscr{L}(X))$. Then the operator $U_t(s)$ in (2.22) is piecewise strongly continuous on I_h. The discontinuity of $U_t(s)$ yields from the first term of the right-hand side of (2.22) and the discontinuous points are $s = -h_r, r = 1,..., m-1$ and $s = t - h_r \in [-h, 0)$, $r = 1,..., m$. If $t > h$, the discontinuity of $U_t(s)$ occurs only at $s = -h_r$, $r = 1,..., m-1$. The second integral term of the right-hand side of (2.22) is strongly continuous on I_h. The fact can be proved by using Lebesgue's dominated convergence theorem. We denote the jump $U_t(s+0) - U_t(s-0)$ at $s \in I_h$ by $\delta U_t(s)$. Then

$$\delta U_t(-h_r) = G(t) A_r \qquad (r = 1,..., m-1),$$

$$\delta U_t(t - h_r) = -A_r \chi_{[-h, 0)}(t - h_r) \qquad (r = 1,..., m).$$

These jumps are closely related to the degeneracy phenomena of retarded systems in infinite-dimensional space [30].

When X is reflexive, the mild solution $x(t)$ of (E) is a weak solution in the sense given below. A function $x(t)$, $t \in [-h, \infty)$, is said to be a weak solution of (E) if

(i) $x \in C(R^+; X)$;

(ii) for each $x^* \in D(A_0^*)$, the function $\langle x(t), x^* \rangle$ is absolutely continuous and satisfies

$$\frac{d}{dt} \langle x(t), x^* \rangle = \langle x(t), A_0^* x^* \rangle + \left\langle \int_{-h}^{0} d\eta(s)\, x(s+t), x^* \right\rangle + \langle f(t), x^* \rangle \qquad \text{a.e.} \quad t \geqslant 0;$$

(iii) $x(0) = g^0$, $x(s) = g^1(s)$ a.e. $s \in [-h, 0)$.

COROLLARY 2.1. *Let the assumption in Theorem* 2.1 *be satisfied and let X be reflexive. Then the mild dsolution $x(t)$ given in* (2.21) *is a weak solution of* (E).

Proof. See Appendix 2.

Finally, we give a condition which implies that the mild solution of (E) becomes a strong solution of (E), which is a function $x\colon [-h, \infty) \to X$ such that

(i) $x \in C(R^+; X) \cap W_p^{(1)}([0, T]; X)$ for all $T > 0$;

(ii) $x(t) \in D(A_0)$ for a.e. $t \geqslant 0$, $x(t)$ is strongly differentiable and satisfies (2.1) a.e. $t \geqslant 0$;

(iii) $x(0) = g^0$, $x(s) = g^1(s)$ a.e. $s \in [-h, 0)$.

COROLLARY 2.2. *Let the assumption in Corollary 2.1 be satisfied. If* $g = (g^0, g^1)$ *and* f *satisfy*

$$g^1 \in W_p^{(1)}(I_h; X), \qquad g^1(0) = g^0 \in D(A_0),$$
$$f \in W_p^{(1)}([0, T]; X) \qquad \text{for each } T > 0,$$

then the function $x(t)$ *given in* (2.21) *is a strong solution of* (E).

Proof. See Appendix 3.

3. Optimal Control Problems and Adjoint System

Let $T > 0$ be fixed and let $I = [0, T]$. We consider the following hereditary controlled system on X:

$$(\mathrm{CS}) \begin{cases} \dfrac{dx(t)}{dt} = A_0 x(t) + \displaystyle\int_{-h}^{0} d\eta(s)\, x(s+t) + f(t) + B(t)\, u(t) \quad \text{a.e. } t \in I, & (3.1) \\ x(0) = g^0, \qquad x(s) = g^1(s) \qquad \text{a.e. } s \in [-h, 0), & (3.2) \\ u \in U_{\mathrm{ad}}, & (3.3) \end{cases}$$

where $A_0, \eta, g = (g^0, g^1)$ are given in Section 2 and $f \in L_p(I; X)$, $U_{\mathrm{ad}} \subset L_p(I; Y)$, $p \in [1, \infty]$, and $B \in L_\infty(I; \mathscr{L}(Y, X))$.

The quantities $x(t)$, $u(t)$, $B(t)$, and U_{ad} in (CS) denote a system state (or a trajectory), a control, a controller, and a class of admissible controls, respectively.

Let $G(t)$ be the fundamental solution of (E) and the assumption H_0^p be satisfied. Then the function

$$x(t) = x(t; f, g) + \int_0^t G(t-s)\, B(s)\, u(s)\, ds \qquad (3.4)$$

is the mild solution of (3.1), (3.2), and a member of $C(I; X)$, where $x(t; f, g)$

is given in (2.21). Since we use the class of mild solutions (3.4) to investigate the control problems for (CS), the assumption H_0^p is always assumed.

In what follows the admissible set U_{ad} is assumed to be closed and convex in $L_p(I; Y)$. We sometimes denote $x(t)$ in (3.4) by $x_u(t)$ to express the dependence on $u \in U_{\mathrm{ad}}$. The function x_u is called the trajectory corresponding to a control u.

We shall shortly explain the results obtained in this paper.

Let $J = J(u, x)$ be the integral convex cost given by

$$J = \phi_0(x(T)) + \int_I (f_0(x(t), t) + k_0(u(t), t))\, dt, \tag{3.5}$$

where $\phi_0 \colon X \to R$, $f_0 \colon X \times I \to R$, $k_0 \colon Y \times I \to R$. We study the following control problems P_1 and P_2 on the finite interval $I = [0, T]$, $T > 0$.

P_1. Find a control $u \in U_{\mathrm{ad}}$ which minimizes the cost J subject to the constraint (CS).

P_2. Find optimality conditions for $(\bar{u}, x_{\bar{u}}) \in U_{\mathrm{ad}} \times C(I; X)$ such that

$$\inf_{u \in U_{\mathrm{ad}}} J(u, x) = J(\bar{u}, x_{\bar{u}}). \tag{3.6}$$

In P_1 such a $u \in U_{\mathrm{ad}}$ is called an optimal control for the cost J. In P_2 the pair $(\bar{u}, x_{\bar{u}})$ is called the optimal solution for J. We will solve P_1 partly by showing the existence of optimal controls in Section 4 and solve P_2 by deriving necessary optimality conditions of integral type in Section 5. Further properties such as maximum principle and bang-bang principle are studied in Sections 6 and 7. At the same time the problem P_1 is solved completely in some specific problems.

Let W be a weakly compact set in X. Define

$$U_0 = \{u \in U_{\mathrm{ad}} \colon x_u(t) \in W \text{ for some } t \in I\} \tag{3.7}$$

and suppose that $U_0 \neq \phi$. For each $u \in U_0$ we can define the transition time that is the first time $\tilde{t}(u)$ such that $x_u(\tilde{t}) \in W$. The set W is called a target set. The time optimal control problem P_3 with a target set W is formulated as

P_3. Find a control $\bar{u} \in U_0$ such that

$$\tilde{t}(\bar{u}) \leqslant \tilde{t}(u) \qquad \text{for all } u \in U_0 \tag{3.8}$$

subject to the constraint (CS).

In P_3 such a $\bar{u} \in U_{ad}$ is called a time optimal control and $\tilde{t}(\bar{u})$ is called an optimal time. In Section 8 we study the problem P_3. First we give an existence theorem of time optimal controls. Next we establish the maximum principle and the bang-bang principle in the case where W has non-empty interior. A convergence theorem of time optimal controls for target sets converging to a point target set is also given in Section 8.

To give a concrete form of those optimality conditions some knowledge on the adjoint system is required. In the sequel we introduce and investigate the retarded system mainly in the case where X is reflexive. First consider the case X is reflexive. Let $q_0^* \in X^*$ and $q_1^* \in L_1(I; X^*)$. The retarded adjoint system (AS) on X^* is defined by

$$
(\mathrm{AS}) \quad \begin{cases} \dfrac{dp(t)}{dt} + A_0^* p(t) + \displaystyle\int_{-h}^{0} d\eta^*(s)\, p(t-s) - q_1^*(t) = 0 \quad \text{a.e.} \quad t \in I & (3.9) \\ p(T) = -q_0^*, \qquad p(s) = 0 \qquad \text{a.e. } s \in (T, T+h], & (3.10) \end{cases}
$$

where $\eta^*(s)$ denotes the adjoint of $\eta(s)$. Since X is reflexive, it is shown in [32] that the adjoint operator A_0^* of A_0 generates a C_0-semigroup $T^*(t)$ on X^* which is the adjoint of $T(t)$, $t \geqslant 0$. Hence we can construct the fundamental solution $G_*(t)$ as in Section 2 (remark that $\int_{I_h} \|D(s)\|\, ds = \int_{I_h} \|D^*(s)\|\, ds < \infty$). That is, $G_*(t)$ is characterized as the (unique) solution of

$$
G_*(t) = \begin{cases} T^*(t) + \displaystyle\int_0^t T^*(t-s) \int_{-h}^{0} d\eta^*(\xi)\, G_*(\xi + s)\, ds, & t \geqslant 0 \\ 0, & t < 0. \end{cases} \tag{3.11}
$$

We denote by $G^*(t)$ the adjoint of $G(t)$, $t \in R$. Then it is verified that $G^*(t) = G_*(t)$ (see Appendix 1). Hence $G^*(t)$ is strongly continuous on R^+. By changing time direction in (AS), we have the following system on X^*:

$$
(\mathrm{CS})^* \quad \begin{cases} \dfrac{dw(t)}{dt} = A_0^* w(t) + \displaystyle\int_{-h}^{0} d\eta^*(s)\, w(t+s) + q_1^*(T-t) \qquad \text{a.e.} \quad t \in I \\ w(0) = -q_0^*, \qquad w(s) = 0 \qquad s \in [-h, 0). \end{cases}
$$

The mild solution $w(t)$ of $(\mathrm{CS})^*$ is represented by

$$
w(t) = G^*(t)(-q_0^*) + \int_0^t G^*(t-s)\, q_1^*(T-s)\, ds. \tag{3.12}
$$

It is easily seen that the system (CS)* is transformed to the system (AS) by a change of variable $t \to T-t$. Hence by (3.12) the function

$$p(t) = w(T-t) = G^*(T-t)(-q_0^*) + \int_t^T G^*(s-t)(-q_1^*(s))\,ds, \quad t \in I \quad (3.13)$$

may be called the mild solution of (AS). The function $p(t)$ in (3.13) is called the adjoint state. In the sense of Corollary 2.1 we often say that $p(t)$ solves (AS) in the weak sense.

When X is not reflexive, the adjoint system can be defined in the following manner (cf. [25, 32]). Define $X_S^* \subset X^*$ by

$$X_S^* = \{x^* \in X^* : \lim_{t \to 0+} |T^*(t)\,x^* - x^*|_* = 0\}.$$

Then the linear subspace X_S^* is invariant under $T^*(t)$, i.e., $T^*(t)\,X_S^* \subset X_S^*$ holds for all $t \geqslant 0$. Note that X_S^* is closed in X^* with respect to the norm topology of X^*. We define the semigroup $T_S^*(t)$ on X_S^* by the restriction $T^*(t)|_{X_S^*}$. Then $T_S^*(t)$ is a C_0-semigroup on the Banach space X_S^*, so that the infinitesimal generator $A_{0,S}^*$ of $T_S^*(t)$ can be determined uniquely. Concerning other operators in $\eta^*(s)$, we suppose that $A_r^*(X_S^*) \subset X_S^*$, $r = 1,..., m$, and $D^*(s)(X_S^*) \subset X_S^*$ a.e. $s \in I_h$. We denote the restrictions $A_r^*|_{X_S^*}$ and $D^*(s)|_{X_S^*}$ by $A_{r,S}^*$ and $D_S^*(s)$. Then it can be verified that $A_{r,S}^* \in \mathscr{L}(X_S^*)$, $r = 1,..., m$, and $D_S^*(\cdot) \in L_1(I_h; \mathscr{L}(X_S^*))$. Let $q_0^* \in X_S^*$, $q_1^* \in L_1(I; X_S^*)$, and η_S^* be the Stieltjes measure corresponding to $A_{r,S}^*$ and $D_S^*(s)$. Now we define the adjoint system $(\mathrm{AS})_S$ on X_S^* by

$$(\mathrm{AS})_S \begin{cases} \dfrac{dp(t)}{dt} + A_{0,S}^*\,p(t) + \displaystyle\int_{-h}^0 d\eta_S^*(s)\,p(t-s) - q_1^*(t) = 0 & \text{a.e.} \quad t \in I, \\ p(T) = -q_0^*, \qquad p(s) = 0, \qquad s \in (T, T+h], \end{cases}$$

when X is not reflexive. Since the structure of X_S^* is not clear for general non-reflexive Banach spaces, we do not use the adjoint system $(\mathrm{AS})_S$ in this paper. However, quite analogous results in terms of the above adjoint system hold true for non-reflexive Banach spaces.

4. Existence of Optimal Control

This section is concerned with the existence of optimal controls for the cost problem P_1. It is assumed in this section that Y is reflexive and $1 < p < \infty$. We consider two cases to study P_1, one is the case where U_{ad} is bounded and the other is where U_{ad} is unbounded in $L_p(I; Y)$. The following assumption H_1 on ϕ_0, f_0, and k_0 is for a bounded U_{ad}.

H_1. (1) *$\phi_0\colon X \to R$ is continuous and convex;* (4.1)

(2) *$f_0\colon X \times I \to R$ is measurable in $t \in I$ for each $x \in X$ and continuous and convex in $x \in X$ for a.e. $t \in I$ and further for each bounded set $K \subset X$ there exists a measurable function $m_K \in L_1(I; R)$ such that*

$$\sup_{x \in K} |f_0(x, t)| \leqslant m_K(t) \qquad a.e.\ t \in I; \tag{4.2}$$

(3) *$k_0\colon Y \times I \to R$ satisfies that for any $u \in U_{\text{ad}}$, $k_0(u(t), t)$ is integrable on I and the functional $\Gamma_0\colon U_{\text{ad}} \to R$ given by*

$$\Gamma_0(u) = \int_I k_0(u(t), t)\, dt \tag{4.3}$$

is continuous and convex.

THEOREM 4.1. *Let U_{ad} be bounded and H_1 be satisfied. Then there exists a control $u_0 \in U_{\text{ad}}$ that minimizes the cost J in (3.5).*

Proof. Let $\{u_n\}$ be a minimizing sequence of J such that

$$\inf_{u \in U_{\text{ad}}} J = \lim_{n \to \infty} J(u_n, x_n) = m_0,$$

where x_n is the trajectory corresponding to u_n. Since U_{ad} is bounded and weakly closed, there exist a subsequence (which we denote again by $\{u_n\}$) of $\{u_n\}$ and a $u_0 \in U_{\text{ad}}$ such that

$$u_n \to u_0 \qquad \text{weakly in } L_p(I; Y). \tag{4.4}$$

We denote by x_0 the trajectory corresponding to u_0. Let $x^* \in X^*$ and $t \in I$ be fixed. Since $G(t) = 0$ if $t < 0$, then

$$\langle x_n(t), x^* \rangle = \langle x(t; f, g), x^* \rangle + \int_I \langle u_n(s), B^*(s)\, G^*(t-s)\, x^* \rangle_{Y, Y^*}\, ds. \tag{4.5}$$

Since $B \in L_\infty(I; \mathscr{L}(Y, X))$ and $G(t)$ is strongly continuous on I, it is easy to see that the function $B^*(\cdot)\, G^*(t - \cdot)\, x^*$ belongs to $L_{p'}(I; Y^*)$, $1/p + 1/p' = 1$. Hence by (4.4), (4.5),

$$\begin{aligned}
\langle x_n(t), x^* \rangle &\to \langle x(t; f, g), x^* \rangle + \int_I \langle u_0(s), B^*(s)\, G^*(t-s)\, x^* \rangle_{Y, Y^*}\, ds \\
&= \langle x(t; f, g), x^* \rangle + \left\langle \int_0^t G(t-s)\, B(s)\, u_0(s)\, ds, x^* \right\rangle \\
&= \langle x_0(t), x^* \rangle \qquad \text{as} \quad n \to \infty,
\end{aligned}$$

i.e.,

$$x_n(t) \to x_0(t) \qquad \text{weakly in } X. \tag{4.6}$$

It is well known that continuity plus convexity imply weak lower semi-continuity. Then (4.1) and (4.6) with $t = T$ imply

$$\varliminf_{n\to\infty} \phi_0(x_n(T)) \geqslant \phi_0(x_0(T)). \tag{4.7}$$

By the same reason we have

$$\varliminf_{n\to\infty} f_0(x_n(t), t) \geqslant f_0(x_0(t), t) \qquad \text{a.e. } t \in I. \tag{4.8}$$

Since U_{ad} is bounded, then by Hölder inequality the set $K = \cup\{x_n(t): t \in I, n = 1, 2,...\}$ is shown to be bounded in X. So from (4.2), there exists an $m_K \in L_1(I; R)$ such that

$$|f_0(x_n(t), t)| \leqslant m_K(t) \qquad \text{a.e. } t \in I, n = 1, 2,.... \tag{4.9}$$

Hence from (4.8) and (4.9) it follows via the Lebesgue–Fatou lemma that

$$\begin{aligned} \varliminf_{n\to\infty} \int_I f_0(x_n(t), t)\, dt &\geqslant \int_I (\varliminf_{n\to\infty} f_0(x_n(t), t))\, dt \\ &\geqslant \int_I f_0(x_0(t), t)\, dt. \end{aligned} \tag{4.10}$$

Concerning the term $\int_I k_0(u_n(t), t)\, dt$, it is lear from $H_1(3)$ that

$$\varliminf_{n\to\infty} \Gamma_0(u_n) \geqslant \Gamma_0(u_0) = \int_I k_0(u_0(t), t)\, dt. \tag{4.11}$$

Therefore by (4.7), (4.10), and (4.11) we have

$$\begin{aligned} m_0 = \inf_{u\in U_{\mathrm{ad}}} J &\geqslant \varliminf_{n\to\infty} \phi_0(x_n(T)) + \varliminf_{n\to\infty} \int_I f_0(x_n(t), t)\, dt + \varliminf_{n\to\infty} \Gamma_0(u_n) \\ &\geqslant \phi_0(x_0(T)) + \int_I (f_0(x_0(t), t) + k_0(u_0(t), t))\, dt \\ &= J(u_0, x_0) > -\infty, \end{aligned}$$

so that $m_0 = J(u_0, x_0)$. This proves that (u_0, x_0) is the optimal solution for J.

The condition $H_1(3)$ seems artificial, but the condition is automatically satisfied in many applications given in later sections.

We next consider the case where U_{ad} is unbounded. In this case we suppose H_1 and the next additional assumption H_2.

H_2. (1) *There exists a constant* $c_0 > 0$ *such that* $\phi_0(x) \geqslant -c_0$ *on* X;

(2) *there exists a constant* $c_1 > 0$ *such that* $f_0(x, t) \geqslant -c_1$ *on* $X \times I$;

(3) *there exists a monotone increasing function* $\theta_0 \in C(R^+; R)$ *such that* $\lim_{r \to \infty} \theta_0(r) = \infty$ *and*

$$\Gamma_0(u) = \int_I k_0(u(t), t)\, dt \geqslant \theta_0(\|u\|_{p,I}) \qquad \text{for } u \in U_{\text{ad}}.$$

THEOREM 4.2. *Let* H_1 *and* H_2 *be satisfied. Then there exists a control* $u_0 \in U_{\text{ad}}$ *which minimizes the cost* J *in* (3.5).

Proof. By virtue of H_2,

$$J \geqslant \theta_0(\|u\|_{p,I}) - c_0 - c_1 T \qquad \text{for } u \in U_{\text{ad}}.$$

Hence a standard argument with $\lim_{r \to \infty} \theta_0(r) = \infty$ implies that the minimizing sequence $\{u_n\}$ is bounded in $L_p(I; Y)$. Then as in the proof of Theorem 4.1, the conclusion of this theorem follows.

Remark 4.1. In Theorem 4.2, the condition (4.2) in $H_1(2)$ can be removed if we use the Fatou lemma instead of the Lebesgue–Fatou lemma.

Remark 4.2. The above existence theorems can be extended to include more general cost functions, for instance,

$$J = \phi_0(x(t_1),..., x(t_k)) + \int_I (f_0(x(t), x(t+s_1),..., x(t+s_k), t) + k_0(u(t), t))\, dt,$$

where $t_i \in I$ and $s_i \in I_h$ $(i = 1,..., k)$. We will use such an extension in later sections.

5. OPTIMALITY CONDITION

In this section we are going to solve the problem P_2. That is, we seek necessary optimality conditions of the optimal solution (u, x) for J in (3.5). The existence of optimal solutions is assumed but the closedness of U_{ad} is not assumed in this section and Sections 6 and 7. In order to give two types of optimality conditions, we require the following assumptions H_3 and H_3^w.

H_3. (1) $\phi_0: X \to R$ *is continuous and Gateau differentiable, and the Gateau derivative* $d\phi_0(x) \in X^*$ *for each* $x \in X$;

(2) $f_0: X \times I \to R$ *is measurable in* $t \in I$ *for each* $x \in X$ *and continuous and convex on* X *for a.e.* $t \in I$ *and further there exist functions* $\partial_1 f_0$: $X \times I \to X^*$, $\theta_1 \in L_1(I; R)$, $\theta_2 \in C(R^+; R)$ *such that*

(a) $\partial_1 f_0$ *is measurable in* $t \in I$ *for each* $x \in X$ *and continuous in* $x \in X$ *for a.e.* $t \in I$ *and the value* $\partial_1 f_0(x, t)$ *is the Gateau derivative of* $f_0(x, t)$ *in the first argument for* $(x, t) \in X \times I$, *and*

(b) $|\partial_1 f_0(x, t)|_* \leqslant \theta_1(t) + \theta_2(|x|)$ *for* $(x, t) \in X \times I$;

(3) $k_0: Y \times I \to R$ *is measurable in* $t \in I$ *for each* $u \in Y$ *and continuous and convex on* Y *for a.e.* $t \in I$ *and further there exist functions* $\partial_1 k_0: Y \times I \to Y^*$, $\theta_3 \in L_{p'}(I; R)$, *and* $M_4 > 0$ *such that*

(a) $\partial_1 k_0$ *is measurable in* $t \in I$ *for each* $u \in Y$ *and continuous in* $u \in Y$ *for a.e.* $t \in I$ *and the value* $\partial_1 k_0(u, t)$ *is the Gateau derivative of* $k_0(u, t)$ *in the first argument for* $(u, t) \in Y \times I$, *and*

(b) $|\partial_1 k_0(u, t)|_{Y^*} \leqslant \theta_3(t) + M_4 |u|_Y^{p/p'}$ *for* $(u, t) \in Y \times I$.

Next we give the condition $(3)^w$ which is different from $H_3(3)$.

H_3. $(3)^w$ $k_0: Y \times I \to R$ *is measurable in* $t \in I$ *for each* $u \in Y$ *and continuous and convex on* Y *for a.e.* $t \in I$ *and further there exist* $\theta_5 \in L_1(I; R)$ *and* $M_6 > 0$ *such that*

$$|k_0(u, t)| \leqslant \theta_5(t) + M_6 |u|_Y^p \qquad \textit{for} \quad (u, t) \in Y \times I.$$

The assumption H_3^w is the set of conditions $H_3(1)$, (2), and $(3)^w$. The assumption H_3 is for differentiable costs and H_3^w is for non-differentiable costs. The following is the main theorem in this section.

THEOREM 5.1. *Let* H_3(*resp.* H_3^w) *be satisfied and let* $(u, x) \in U_{\text{ad}} \times C(I; X)$ *be an optimal solution for* J *in* (3.5). *Then the integral inequality*

$$\int_I \langle v(t) - u(t), \partial_1 k_0(u(t), t) - B^*(t) p(t) \rangle_{Y, Y^*} dt \geqslant 0 \qquad \textit{for all } v \in U_{\text{ad}} \tag{5.1}$$

$$(\textit{resp.} \int_I \langle v(t) - u(t), -B^*(t) p(t) \rangle_{Y, Y^*} dt$$

$$+ \int_I (k_0(v(t), t) - k_0(u(t), t))\, dt \geqslant 0 \qquad \textit{for all } v \in U_{\text{ad}}) \tag{5.2}$$

holds, where

$$p(t) = -G^*(T - t)\, d\phi_0(x(T)) - \int_t^T G^*(s - t)\, \partial_1 f_0(x(s), s)\, ds. \tag{5.3}$$

If X is reflexive, $p \in C(I; X^*)$ satisfies

$$
\text{(AS)} \quad \begin{cases} \dfrac{dp(t)}{dt} + A_0^* p(t) + \displaystyle\int_{-h}^{0} d\eta^*(s)\, p(t-s) - \partial_1 f_0(x(t), t) = 0 & a.e. \quad t \in I, \\ p(T) = -d\phi_0(x(T)), \qquad p(s) = 0 \qquad s \in (T, T+h] \end{cases}
$$

in the weak sense.

Proof. Let H_3 be satisfied. Then the cost J in (3.5) is Gateau differentiable. We know [26, p. 10] that the necessary optimality condition is given by the variational inequality

$$J'(u)(v-u) \geqslant 0 \qquad \text{for all } v \in U_{\mathrm{ad}} \tag{5.5}$$

when J is differentiable. By virtue of H_3, we have by Lebesgue's dominated convergence theorem that

$$
\begin{aligned}
J'(u)(v-u) = {} & \left\langle \int_I G(T-s)\, B(s)(v(s)-u(s))\, ds,\, d\phi_0(x(T)) \right\rangle \\
& + \int_I \left\langle \int_0^s G(s-\tau)\, B(\tau)(v(\tau)-u(\tau))\, d\tau,\, \partial_1 f_0(x(s), s) \right\rangle ds \\
& + \int_I \langle v(s)-u(s),\, \partial_1 k_0(u(s), s) \rangle_{Y, Y^*}\, ds.
\end{aligned} \tag{5.6}
$$

We remark that all integrals in (5.6) are well defined by making use of H_3. The first term of (5.6) can be rewritten as

$$\int_I \langle v(s)-u(s),\, B^*(s)\, G^*(T-s)\, d\phi_0(x(T)) \rangle_{Y, Y^*}\, ds. \tag{5.7}$$

Using Fubini's theorem the second term of (5.6) is transformed as

$$
\begin{aligned}
& \int_I \int_0^s \langle G(s-\tau)\, B(\tau)(v(\tau)-u(\tau)),\, \partial_1 f_0(x(s), s) \rangle\, d\tau\, ds \\
& \quad = \int_I \left\langle v(s)-u(s),\, B^*(s) \int_s^T G^*(\tau-s)\, \partial_1 f_0(x(\tau), \tau)\, d\tau \right\rangle_{Y, Y^*} ds.
\end{aligned} \tag{5.8}
$$

If we define $p(t)$ by (5.3), then from (5.5)–(5.8) the inequality (5.1) follows. Next let H_3^w be satisfied. Then we can use the variational inequality

$$(J - \Gamma_0)'(u)(v-u) + (\Gamma_0(v) - \Gamma_0(u)) \geqslant 0 \qquad \text{for all } v \in U_{\mathrm{ad}}$$

in [26, p. 12] to obtain the condition (5.2), where Γ_0 is given in (4.3). The last statement is clear from the argument in Section 3.

Remark 5.1. If $U_{ad} = L_p(I; Y)$ in Theorem 5.1, then the condition (5.1) is reduced so that

$$\partial_1 k_0(u(t), t) - B^*(t)\, p(t) = 0 \qquad \text{a.e.} \quad t \in I. \tag{5.9}$$

Remark 5.2. From (5.2) the following "integral" maximum principle holds:

$$\max_{v \in U_{ad}} \int_I (\langle v(s), B^*(s)\, p(s) \rangle_{Y, Y^*} - k_0(v(s), s))\, ds$$
$$= \int_I (\langle u(s), B^*(s)\, p(s) \rangle_{Y, Y^*} - k_0(u(s), s))\, ds. \tag{5.10}$$

Remark 5.3. Consider the special case where Y is a Hilbert space, $p = 2$, $U_{ad} = \{u \in L_2(I; Y) \colon \|u\|_{2,I} \leqslant \alpha\}$, and H_3 is satisfied. In this case the optimal control u is characterized by the relation

$$u = -\alpha \frac{\Lambda^{-1} K(u)}{\|\Lambda^{-1} K(u)\|_{2,I}},$$

where Λ is the canonical isomorphism of $L_2(I; Y)$ into $L_2(I; Y^*)$ and $K(u)(t) = \partial_1 k_0(u(t), t) - B^*(t)\, p(t)$ a.e. $t \in I$.

We now give applications of Theorem 5.1 to the regulator problem and the uniqueness of averaging observation control.

EXAMPLE 5.1 (Regulator problem). Let X and Y be Hilbert spaces with inner products $(\ ,\ )$ and $\langle\ ,\ \rangle_Y$, respectively. We suppose $U_{ad} = L_2(I; Y)$. The spaces X and X^* are identified. The cost J_1 is given by

$$J_1 = (x(T), Nx(T)) + \int_I (x(t), W(t)\, x(t))\, dt + \Gamma_Q(u), \tag{5.11}$$

where

$$\Gamma_Q(u) = \frac{1}{2} \int_I \langle u(t), Q(t)\, u(t) \rangle_Y\, dt. \tag{5.12}$$

In (5.11) and (5.12) we assume that $N \in \mathscr{L}(X)$, $W(\cdot) \in L_\infty(I; \mathscr{L}(X))$, $Q(\cdot) \in L_\infty(I; \mathscr{L}(Y))$; N, $W(t)$, $Q(t)$ are positive and symmetric for each $t \in I$; there exists a constant $c > 0$ such that

$$\langle u, Q(t)\, u \rangle_Y \geqslant c\, |u|_Y^2 \qquad \text{for a.e.} \quad t \in I.$$

Under the above conditions it is verified that $\Gamma_Q(u)$ is strongly continuous and strictly convex in $L_2(I; Y)$ [26, Chap. III]. Hence the assumptions H_1 and H_2 are satisfied for the cost J_1. In addition J_1 is strictly convex. Then by Theorem 4.2 there exists a unique optimal control for J_1. Thus, from Theorem 5.1 and Remark 5.1 we obtain

COROLLARY 5.1. *Let the cost* J_1 *be given by* (5.11), (5.12). *Then there exists a unique optimal solution* $(u, x) \in L_2(I; Y) \times C(I; X)$ *for* J_1. *The optimal control* $u(t)$ *is given by*

$$u(t) = Q^{-1}(t)\, B^*(t)\, p(t) \qquad \text{a.e.} \quad t \in I,$$

where the pair $(x, p) \in C(I; X) \times C(I; X)$ *satisfies the system of equations*

$$\begin{cases} \dfrac{dx(t)}{dt} = A_0 x(t) + \displaystyle\int_{-h}^{0} d\eta(s)\, x(t+s) + B(t)\, Q^{-1}(t)\, B^*(t)\, p(t) + f(t) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{a.e.} \quad t \in I, \\ x(0) = g^0, \qquad x(s) = g^1(s) \qquad \text{a.e.} \quad s \in [-h, 0), \\ \dfrac{dp(t)}{dt} + A_0^* p(t) + \displaystyle\int_{-h}^{0} d\eta^*(s)\, p(t-s) - W(t)\, x(t) = 0 \qquad \text{a.e.} \quad t \in I, \\ p(T) = -Nx(T), \qquad p(s) = 0 \qquad s \in (T, T+h] \end{cases}$$

in the weak sense.

The above cost problem is called the regulator problem and is very important in system design and synthesis. There are many researchers who discuss the problem in both finite- and infinite-dimensional systems. We refer to the books [1, 14, 26] for infinite-dimensional systems without delay and [15, 16, 19] for finite-dimensional retarded systems. But the literature dealing with infinite-dimensional retarded systems is small [2, 35].

The optimality condition (5.1) or (5.2) is often used to derive the uniqueness of optimal control. To give such an application we need the following lemma, which is well known for C_0-semigroups [33, Chap. 7].

LEMMA 5.1. *Let* $f \in L_p(I; X)$, $1 \leqslant p \leqslant \infty$. *If*

$$\int_0^t G(t-s) f(s)\, ds = 0 \qquad \text{for all } t \in I,$$

then $f(t) = 0$ *a.e.* $t \in I$.

Proof. Put $y(t)=\int_0^t G(t-s)f(s)\,ds$. Then by Theorem 2.2, $y(t)$ satisfies

$$y(t)=\begin{cases}\int_0^t T(t-s)\int_{-h}^0 d\eta(\xi)\,y(s+\xi)\,ds+\int_0^t T(t-s)f(s)\,ds, & t\in I\\ 0, \quad t\in[-h,0). & \end{cases} \tag{5.13}$$

Since $y(t)=0$, $t\in I$, it follows from (5.13) that

$$\int_0^t T(t-s)f(s)\,ds=0 \qquad \text{for all } t\in I.$$

Hence by the property of a C_0-semigroup $T(t)$, we have

$$f(t)=0 \qquad \text{a.e.} \quad t\in I.$$

EXAMPLE 5.2 (Uniqueness of averaging observation control). Let Z be a Hilbert space with inner product $\langle , \rangle_Z$ and norm $|\cdot|_Z$. Let $C(\cdot)\in L_2(I;\mathscr{L}(X,Z))$ and $p=2$. The cost J_2 is given by

$$J_2=\frac{1}{2}\int_I |C(t)\,x(t)-z_d(t)|_Z^2\,dt, \qquad z_d\in L_2(I;Z). \tag{5.14}$$

Note that the cost J_2 is not strictly convex.

COROLLARY 5.2. *Let the cost J_2 be given by* (5.14) *and U_{ad} be bounded and closed in $L_2(I;Y)$. Then there exists an optimal control u for J_2. If both $B(t)$ and $C(t)$ are one to one for a.e. $t\in I$, then the optimal control for J_2 is unique.*

Proof. Since U_{ad} is bounded and closed, the existence of an optimal control follows from Theorem 4.1. It is sufficient to show the uniqueness of optimal control. Let u_1, u_2 be optimal controls for J_2 and x_1, x_2 be the corresponding trajectories to u_1, u_2, respectively. Then as in the proof of Theorem 5.1, we have for $i=1,2$

$$J_2'(u_i)(v-u_i)=\int_I \left\langle \int_0^t G(t-\tau)\,B(\tau)(v(\tau)-u_i(\tau))\,d\tau,\right.$$

$$\left. C^*(t)(C(t)\,x_i(t)-z_d(t))\right\rangle dt\geqslant 0 \qquad \text{for all} \quad v\in U_{\mathrm{ad}}. \tag{5.15}$$

By substituting $v=u_2$ if $i=1$ and $v=u_1$ if $i=2$ in (5.15) and adding these inequalities, we obtain

$$\int_I \left\langle \int_0^t G(t-\tau)\, B(\tau)(u_1(\tau)-u_2(\tau))\, d\tau,\, C^*(t)(C(t)(x_1(t)-x_2(t))) \right\rangle dt \leqslant 0,$$

so that from the representation (3.4),

$$\int_I |C(t)(x_1(t)-x_2(t))|_Z^2\, dt \leqslant 0. \tag{5.16}$$

Since $C(t)$ is one to one, then by (5.16)

$$x_1(t)-x_2(t)=\int_0^t G(t-s)\, B(s)(u_1(s)-u_2(s))\, ds=0, \qquad t\in I. \tag{5.17}$$

Applying Lemma 5.1 to (5.17), we have

$$B(t)(u_1(t)-u_2(t))=0 \qquad \text{a.e.} \quad t\in I.$$

Since $B(t)$ is also one to one, $u_1(t)=u_2(t)$ a.e. $t\in I$. That is, the optimal control for J_2 is unique.

6. Maximum Principle

The purpose of this section is to establish the "pointwise" maximum principle for time varying control domain with the convex cost J in (3.5). The assumption H_3^w is assumed in this section. Let the admissible set U_{ad} be

$$U_{\mathrm{ad}}=\{u\in L_p(I;\, Y)\colon u(t)\in U(t) \quad \text{a.e.} \quad t\in I\}, \tag{6.1}$$

where the (time varying) control domain $U(t)\subset Y$, $t\in I$, satisfies

H_4. (1) *$U(t)$ is closed and convex in Y for each $t\in I$;*

(2) *for any $t\in I$, $v\in \operatorname{Int} U(t)$, there exists an $\varepsilon_0>0$ such that*

$$v\in\left(\bigcap_{s\in(t,t+\varepsilon_0)} U(s)\right)\cup\left(\bigcap_{s\in(t-\varepsilon_0,t)} U(s)\right). \tag{6.2}$$

It is evident from $H_4(1)$ that U_{ad} is convex.

Remark 6.1. If $U(t)$ varies continuously with respect to the Hausdorff metric or $U(t)$ is monotone increasing or decreasing, then the condition $H_4(2)$ is satisfied.

The following "pointwise" maximum principle is deduced from the optimality condition (5.2). Compare with (5.10).

THEOREM 6.1. *Let U_{ad} be given by* (6.1) *and H_4 be satisfied. Let $(u, x) \in U_{ad} \times C(I; X)$ be an optimal solution for J in* (3.5). *Then*

$$\max_{v \in U(t)} \{\langle B(t)\, v, p(t)\rangle - k_0(v, t)\} = \langle B(t)\, u(t), p(t)\rangle - k_0(u(t), t) \qquad \text{a.e.} \quad t \in I, \quad (6.3)$$

where $p(t)$ is given by (5.3). *If X is reflexive, then $p(\cdot) \in C(I; X^*)$ and is the mild solution of* (AS) *given in* (5.4).

Proof. Let $t \in I$ and $v \in \operatorname{Int} U(t)$. Since v satisfies (6.2), we suppose, e.g., $v \in \bigcap_{s \in (t, t+\varepsilon_0)} U(s)$. Then it is easily seen that the function

$$v_\varepsilon(s) = \begin{cases} u(s), & s \in I - (t, t+\varepsilon) \\ v, & s \in [t, t+\varepsilon] \end{cases}$$

belongs to U_{ad} for each $\varepsilon \in (0, \varepsilon_0)$. Substituting v_ε for v in (5.2) and deviding the resulting inequality by ε, we obtain

$$\frac{1}{\varepsilon}\int_t^{t+\varepsilon} \{\langle v - u(s), -B^*(s)\, p(s)\rangle_{Y, Y^*} + (k_0(v, s) - k_0(u(s), s))\}\, ds \geqslant 0. \quad (6.4)$$

Since all the integrands in (6.4) are integrable on I by virtue of H_3^w, the Lebesgue density theorem [33, Chap. 1] can apply. Then by letting $\varepsilon \to 0$ in (6.4), we have

$$\langle v, B^*(t)\, p(t)\rangle_{Y, Y^*} - k_0(v, t) \leqslant \langle u(t), B^*(t)\, p(t)\rangle_{Y, Y^*} - k_0(u(t), t) \qquad \text{a.e.} \quad t \in I. \quad (6.5)$$

Let $t \in I$ be fixed for which $u(t) \in U(t)$ and (6.5) holds. Since the duality pairing $\langle v, B^*(t)\, p(t)\rangle_{Y, Y^*}$ is continuous in v, we see from (6.5) that the maximum principle (6.3) is true for such $t \in I$.

Remark 6.2. Let H_3 be satisfied and U_{ad} be as in Theorem 6.1. Then the "pointwise" optimality condition that for a.e. $t \in I$

$$\langle v - u(t), \partial_1 k_0(u(t), t) - B^*(t)\, p(t)\rangle_{Y, Y^*} \geqslant 0 \qquad \text{for all} \quad v \in U(t)$$

holds. This fact is proved by applying the Lebesgue density theorem to the condition (5.1).

Remark 6.3. Consider the extended cost J in Remark 4.2. Under some suitable conditions on the total Fréchet differentiability of ϕ_0 and f_0, the (modified) maximum principle (6.3) also holds in which $p(t)$ is replaced by

$$p(t) = -\sum_{i=1}^{k} G^*(t_i - t)\, \partial_i \phi_0(x(t_1),..., x(t_k))$$
$$-\sum_{i=1}^{k} \int_t^T G^*(s-t)\, \partial_i f_0(x(s), x(s+s_1),..., x(s+s_k), s)\, ds,$$

where ∂_i denotes the Fréchet derivative in the ith argument.

Before giving applications of Theorem 6.1 we shall show the next lemma.

LEMMA 6.1. *Let Y be reflexive and $1 < p < \infty$. Let U_{ad} be given by* (6.1) *and $H_4(1)$ be satisfied. If*

$$\bigcup_{t \in I} U(t) \qquad \text{is bounded in } Y, \tag{6.6}$$

then U_{ad} is weakly (*and hence strongly*) *closed and weakly compact in $L_p(I; Y)$.*

Proof. We shall prove that U_{ad} is weakly closed. Let $\{u_n\}$ be a sequence in U_{ad} such that $u_n \to u_0$ weakly in $L_p(I; Y)$. By (6.6), $\{u_n\}$ is uniformly bounded in Y, i.e., there exists a constant $M > 0$ such that

$$|u_n(t)|_Y \leqslant M \qquad \text{for all } t \in I \text{ and } n = 1, 2,...,$$

Hence by the lemma in Kato [23, Lemma 8], we have

$$u_0(t) \in \tilde{U}(t) \qquad \text{for a.e.} \quad t \in I,$$

where $\tilde{U}(t)$ denotes the closed convex hull of the weak closure of $U(t)$. Since $H_4(1)$ is satisfied, then $\tilde{U}(t) = U(t)$, $t \in I$, and hence $u_0 \in U_{\mathrm{ad}}$. This shows U_{ad} is weakly closed. Because $L_p(I; Y)$ is reflexive and U_{ad} is bounded (by (6.6)), convex and weakly closed, we see from the Eberlein–Šmulian theorem [17, p. 430] that U_{ad} is weakly compact.

In what follows we consider the special cost functionals $J_2 - J_6$ in Examples 6.1–6.5. Such costs are important in practical applications and are studied in [1, 3, 4, 14, 18, 26] for systems without delay. In all examples given below the assumptions on U_{ad} in Lemma 6.1 are supposed. Then the existence of an optimal solution for each J_i, $i = 2, 3, 4, 5, 6$, is assured by Theorem 4.1, Remark 4.2, and Lemma 6.1.

EXAMPLE 6.1. (Averaging observation control problem). Consider the same convex cost problem for J_2 in Example 5.2. In this case the maximum principle is represented by the following

COROLLARY 6.1. *Let $(u, x) \in U_{\text{ad}} \times C(I; X)$ be an optimal solution for J_2 in* (5.14). *Then*

$$\max_{v \in U(t)} \langle v, B^*(t)\, p(t) \rangle_{Y, Y^*} = \langle u(t), B^*(t)\, p(t) \rangle_{Y, Y^*} \qquad a.e. \quad t \in I,$$

where

$$p(t) = \int_t^T G^*(s-t)\, C^*(s)(z_d(s) - C(s)\, x(s))\, ds, \qquad t \in I. \tag{6.7}$$

If X is reflexive, $p(t)$ in (6.7) *is strongly continuous on X^* and satisfies*

$$\begin{cases} \dfrac{dp(t)}{dt} + A_0^* p(t) + \displaystyle\int_{-h}^{0} d\eta^*(s)\, p(t-s) + C^*(t)(z_d(t) - C(t)\, x(t)) = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad a.e. \quad t \in I \\ p(s) = 0 \qquad s \in [T, T+h] \end{cases}$$

in the weak sense.

EXAMPLE 6.2. (Special linearized Bolza problem). The cost J_3 is given by

$$J_3 = \langle x(T), \psi_0^* \rangle + \int_I \langle x(t), \psi_1^*(t) \rangle\, dt, \tag{6.8}$$

where $\psi_0^* \in X^*$ and $\psi_1^* \in L_1(I; X^*)$. Then we have

COROLLARY 6.2. *Let $(u, x) \in U_{\text{ad}} \times C(I; X)$ be an optimal solution for J_2 in* (6.8). *Then*

$$\max_{v \in U(t)} \langle B(t)\, v, p(t) \rangle = \langle B(t)\, u(t), p(t) \rangle \qquad a.e. \quad t \in I,$$

where

$$p(t) = -G^*(T-t)\, \psi_0^* - \int_t^T G^*(s-t)\, \psi_1^*(s)\, ds, \qquad t \in I. \tag{6.9}$$

If X is reflexive, $p(t)$ in (6.9) *belongs to $C(I; X^*)$ and satisfies*

$$\begin{cases} \dfrac{dp(t)}{dt} + A_0^* p(t) + \displaystyle\int_{-h}^{0} d\eta^*(s)\, p(t-s) - \psi_1^*(t) = 0 \qquad a.e. \quad t \in I \\ p(T) = -\psi_0^*, \qquad p(s) = 0 \qquad s \in (T, T+h] \end{cases}$$

in the weak sense.

EXAMPLE 6.3 (Terminal value control problem). Let X be a Hilbert space. As usual we identify X and X^*. The cost J_4 is given by

$$J_4 = \tfrac{1}{2}\, |x(T) - x_d|^2, \qquad x_d \in X. \tag{6.10}$$

COROLLARY 6.3. *Let $(u, x) \in U_{\mathrm{ad}} \times C(I; X)$ be an optimal solution for J_4 in (6.10). Then*

$$\max_{v \in U(t)} (B(t)\, v, p(t)) = (B(t)\, u(t), p(t)) \qquad a.e. \quad t \in I,$$

where $p(t)$ is given by

$$p(t) = G^*(T - t)(x_d - x(T)), \qquad t \in I. \tag{6.11}$$

The adjoint state $p \in C(I; X)$ in (6.11) satisfies

$$\begin{cases} \dfrac{dp(t)}{dt} + A_0^* p(t) + \displaystyle\int_{-h}^{0} d\eta^*(s)\, p(t-s) = 0 \qquad a.e. \quad t \in I \\ p(T) = x_d - x(T), \qquad p(s) = 0 \qquad s \in (T, T+h] \end{cases}$$

in the weak sense ($p(t)$ may be identically zero).

EXAMPLE 6.4 (Minimum energy problem). Let X and Y be Hilbert spaces. The cost J_5 is given by

$$J_5 = \int_I (\lambda^2\, |x(t)|^2 + |u(t)|_Y^2)\, dt, \tag{6.12}$$

where $\lambda > 0$. Then by Theorem 6.1 and Corollary 2.2 we have

COROLLARY 6.4. *Let $(u, x) \in U_{\mathrm{ad}} \times C(I; X)$ be an optimal solution for J_5 in (6.11). Then*

$$\max_{v \in U(t)} (B(t)\, v, p(t)) - |v|_Y^2 = (B(t)\, u(t), p(t)) - |u(t)|_Y^2 \qquad a.e. \quad t \in I,$$

where

$$p(t) = -\int_t^T G^*(s - t)(2\lambda^2 x(s))\, ds \in X^* = X, \qquad t \in I \tag{6.13}$$

satisfies

$$\begin{cases} \dfrac{dp(t)}{dt} + A_0^* p(t) + \displaystyle\int_{-h}^{0} d\eta^*(s)\, p(t-s) - 2\lambda^2 x(t) = 0 \qquad a.e. \quad t \in I \\ p(s) = 0 \qquad s \in [T, T+h] \end{cases} \tag{6.14}$$

in the weak sense. If $f(t)$, $(g^0, g^1(s))$, *and* $B(t)u(t)$ *in* (CS) *satisfy, respectively,* (2.30), (2.31), *and* $B(\cdot)u(\cdot) \in W_p^{(1)}(I; X)$, *then* $p(t)$ *in* (6.13) *is absolutely continuous on* I *and satisfies* (6.14).

EXAMPLE 6.5 (Intermediate values control problem). Let Z be a Hilbert space and let $\{(t_i, C_i, z_i): i = 1,..., k\} \subset I \times \mathscr{L}(X, Z) \times Z$. The cost J_6 is given by

$$J_6 = \frac{1}{2} \sum_{i=1}^{k} |C_i x(t_i) - z_i|_Z^2. \tag{6.15}$$

From Remark 6.3, we have

COROLLARY 6.5. *Let* $(u, x) \in U_{\text{ad}} \times C(I; X)$ *be an optimal solution for* J_6 *in* (6.15). *Then*

$$\max_{v \in U(t)} \langle B(t)v, p(t) \rangle = \langle B(t)u(t), p(t) \rangle \qquad a.e. \quad t \in I,$$

where

$$p(t) = \sum_{i=1}^{k} G^*(t_i - t)\, C_i^*(z_i - C_i x(t_i)), \qquad t \in I.$$

7. BANG–BANG PRINCIPLE

Let the admissible set U_{ad} be given in Section 6. In this section we consider the terminal value cost J given by

$$J = \phi_0(x(T)), \tag{7.1}$$

where ϕ_0 satisfies $H_1(1)$ and $H_3(1)$. We investigate the possibility of the so-called bang-bang control for J in (7.1) for the time varying control domain $U(t)$. In general it is known that the bang-bang control does not hold for the retarded system even in finite-dimensional space [20, p. 60]. However, by restricting the cost J to the terminal value cost, we can prove that the bang-bang control is possible under some regularity condition on the adjoint system. Let X be reflexive in this section. Consider the adjoint system (AS) given in (3.9), (3.10). We denote by $p(t; q_0^*, q_1^*)$ the mild solution of (AS). We now give the condition

C_w: $q_0^* = 0$ in X^* follows from the existence of a set $E \subset I$ such that

$$\text{meas}\, E > 0 \qquad \text{and} \qquad p(t; q_0^*, 0) = 0 \qquad \text{for all} \quad t \in E. \tag{7.2}$$

We say that the adjoint system (AS) is weakly regular if the condition C_w is satisfied. Examples for which the system (AS) is weakly regular are given in [18, p. 41], but those systems do not involve time delay.

EXAMPLE 7.1. Consider the control system (CS) enjoying the following conditions:

(i) A_0 generates an analytic semigroup;

(ii) the Stieltjes measure η is given by $\eta(s) = -\chi_{(-\infty, -h]}(s) A_1$;

(iii) the system (CS) is pointwise complete for all $t > 0$.

The condition (iii) means that for any $f \in L_p^{\mathrm{loc}}(R^+; X)$,

$$\mathrm{Cl}\{x(t; f, g): g \in M_p\} = X \qquad \text{for each } t > 0,$$

where Cl denotes the closure in X. If (i), (ii), (iii) are satisfied, the adjoint system of (CS) is weakly regular (see Appendix 4).

Let us define the reachable set $\Omega(t)$, $t \in I$, by

$$\Omega(t) = \{y \in X: y = x_u(t), u \in U_{\mathrm{ad}}\}. \tag{7.3}$$

It is verified that $\Omega(t)$ is convex, closed, and weakly compact in X provided that U_{ad} is weakly compact in $L_p(I; Y)$ (cf. Lemma 6.1). The following assumption is needed in proving the bang-bang principle.

H_5. *$d\phi_0(y) \neq 0$ in X^* for all $y \in \Omega(T)$.*

THEOREM 7.1. *Let the cost J be given in (7.1). Assume that the adjoint system (AS) is weakly regular and $B^*(t)$ is one to one for a.e. $t \in I$. If H_5 is satisfied, then the optimal control $u(t)$ is a bang-bang control, i.e., $u(t)$ satisfies*

$$u(t) \in \partial U(t) \qquad a.e. \quad t \in I, \tag{7.4}$$

where $\partial U(t)$ denotes the boundary of $U(t)$.

Proof. For the terminal value cost J in (7.1) the maximum principle is written by

$$\max_{v \in U(t)} \langle v, B^*(t)\, p(t) \rangle_{Y, Y^*} = \langle u(t), B^*(t)\, p(t) \rangle_{Y, Y^*} \qquad \text{a.e.} \quad t \in I, \tag{7.5}$$

where $p(t) = p(t; d\phi_0(x(T)), 0)$ and $x(t)$ is the trajectory corresponding to the optimal control $u(t)$. Then it is sufficient to show (7.4) that

$$B^*(t)\, p(t) \neq 0 \text{ in } Y^* \qquad \text{a.e.} \quad t \in I. \tag{7.6}$$

Suppose to the contrary that there exists a set E such that meas $E > 0$ and $B^*(t)\,p(t) = 0$ for $t \in E$. Since $B^*(t)$ is one to one and (AS) is weakly regular, we have by C_w that $d\phi_0(x(T)) = 0$. Because $x(T) \in \Omega(T)$, the condition $d\phi_0(x(T)) = 0$ is impossible by H_5. Hence (7.6) is shown.

EXAMPLE 7.2. Let the assumption in Theorem 7.1 be satisfied and let X be a Hilbert space. We consider two costs $J_4 = \frac{1}{2}|x(T) - x_d|^2$ and $J_7 = (x(T), \psi_0)$, $\psi_0 \in X$. If there exists no trajectory x_u, $u \in U_{\mathrm{ad}}$, such that $x_u(T) = x_d$, then the optimal control $u(t)$ for J_4 is a bang-bang control. For the cost J_7 the bang-bang principle (7.4) holds for any $\psi_0 \neq 0$.

Let U be a convex set in Y. The set U is said to be strictly convex if u, v, $(u+v)/2 \in U$ imply $u = v$. We know that the non-void closed ball in a Hilbert space is strictly convex. The next corollaries are immediate from Theorem 7.1.

COROLLARY 7.1. *Let the assumption in Theorem* 7.1 *be satisfied and let $U(t)$ be strictly convex for all $t \in I$. Then the optimal control $u(t)$ for J in* (7.1) *is unique.*

COROLLARY 7.2. *Let the assumption in Theorem* 7.1 *be satisfied. Let Y be a Hilbert space and $U(t)$ be given by*

$$U(t) = \{u \in Y\colon |u - y(t)|_Y \leqslant r(t)\}, \qquad t \in I, \tag{7.7}$$

where $y(\cdot) \in C(I; Y)$ and $r(\cdot) \in C(I; R^+\text{-}\{0\})$. Then the optimal control $u(t)$ for J in (7.1) *is unique and is given by*

$$u(t) = y(t) + r(t) \cdot \frac{\Lambda_Y^{-1} B^*(t)\,p(t)}{|\Lambda_Y^{-1} B^*(t)\,p(t)|_Y} \qquad a.e. \quad t \in I,$$

where Λ_Y is the canonical isomorphism of Y onto Y^* and $p(t) = G^*(T-t)\,d\phi_0(x(T))$, $t \in I$.

8. TIME OPTIMAL CONTROL

In this section we study the time optimal control problem P_3. Throughout this section it is assumed that X is reflexive, W is weakly compact in X, and U_{ad} is weakly compact in $L_p(I; Y)$. Let U_0 be given in (3.7). Since $x_u \in C(I; X)$, the transition time $\tilde{t}(u)$ is well defined for each $u \in U_{\mathrm{ad}}$.

THEOREM 8.1. *Assume that $U_0 \neq \varnothing$. Then there exists a time optimal control for* P_3.

Proof. Put $t_0 = \inf\{\tilde{t}(u) : u \in U_0\}$. Let $\{u_n, x_n\}$ be a minimizing sequence such that

$$x_n(t_n) = x(t_n; f, g) + \int_0^{t_n} G(t_n - s)\, B(s)\, u_n(s)\, ds \in W, \qquad u_n \in U_0, \quad (8.1)$$

where $t_n = \tilde{t}(u_n) \downarrow t_0$ as $n \to \infty$. We denote $\{x_n(t_n)\}$ by $\{w_n\}$. Since W and U_{ad} are weakly compact, there exist an $u_0 \in U_{\mathrm{ad}}$, $w_0 \in W$ and subsequences, which are denoted again by $\{u_n\}$, $\{w_n\}$, and $\{t_n\}$, such that

$$\begin{aligned} &u_n \to u_0 && \text{weakly in } L_p(I; Y), \\ &w_n \to w_0 && \text{weakly in } X, \\ &t_n \downarrow t_0 && \text{in } I. \end{aligned} \qquad (8.2)$$

Let $x^* \in X^*$. Then

$$\begin{aligned} \langle w_n, x^* \rangle = \langle x(t_n; f, g), x^* \rangle &+ \int_0^{t_0} \langle G(t_n - s)\, B(s)\, u_n(s), x^* \rangle\, ds \\ &+ \int_{t_0}^{t_n} \langle G(t_n - s)\, B(s)\, u_n(s), x^* \rangle\, ds. \end{aligned} \qquad (8.3)$$

Since $G(t)$ satisfies (2.19), then

$$\begin{aligned} G(t+\varepsilon) - T(\varepsilon)\, G(t) &= T(t+\varepsilon) + \int_0^{t+\varepsilon} T(t+\varepsilon-s) \int_{-h}^0 d\eta(\xi)\, G(s+\xi)\, ds \\ &\quad - T(\varepsilon)\left(T(t) + \int_0^t T(t-s) \int_{-h}^0 d\eta(\xi)\, G(s+\xi)\, ds \right) \\ &= \int_t^{t+\varepsilon} T(t+\varepsilon-s) \int_{-h}^0 d\eta(\xi)\, G(s+\xi)\, ds \\ &= \int_0^{\varepsilon} T(\varepsilon-s) \int_{-h}^0 d\eta(\xi)\, G(s+t+\xi)\, ds, \qquad \varepsilon \geqslant 0. \end{aligned}$$

Hence, the second term in the right-hand side of (8.3) is written by

$$\begin{aligned} &\int_0^{t_0} \langle G(t_0 - s)\, B(s)\, u_n(s), T^*(t_n - t_0)\, x^* \rangle\, ds \\ &\quad + \int_0^{t_0} \langle K(t_n, t_0; s)\, B(s)\, u_n(s), x^* \rangle\, ds, \end{aligned} \qquad (8.4)$$

where

$$K(t_n, t_0; s) = \int_0^{t_n - t_0} T(t_n - t_0 - v) \int_{-h}^{0} d\eta(\xi)\, G(v + t_0 - s + \xi)\, dv. \quad (8.5)$$

By the expression (8.5), we have

$$\|K(t_n, t_0; s)\| \leqslant (\sup_{t \in I} \|T(t)\|) \cdot \operatorname{Var} \eta \cdot (\sup_{t \in I} \|G(t)\|)(t_n - t_0)$$
$$\equiv M_7(t_n - t_0), \qquad s \in [0, t_0].$$

So that by Hölder inequality,

$$\left| \int_0^{t_0} \langle K(t_n, t_0; s)\, B(s)\, u_n(s), x^* \rangle\, ds \right|$$
$$\leqslant M_7\, \|B(\cdot)\|_{\infty, I} \cdot \|u_n(\cdot)\|_{p,I} \cdot t_0^{1/p'} \cdot |x^*|_* \cdot (t_n - t_0), \quad (8.6)$$

where $1/p + 1/p' = 1$. Similarly the last term in (8.3) is estimated as

$$\left| \int_{t_0}^{t_n} \langle G(t_n - s)\, B(s)\, u_n(s), x^* \rangle\, ds \right|$$
$$\leqslant (\sup_{t \in I} \|G(t)\|) \cdot \|B(\cdot)\|_{\infty, I} \cdot \|u_n(\cdot)\|_{p,I} \cdot |x^*|_* \cdot (t_n - t_0)^{1/p'}. \quad (8.7)$$

Since $x(t; f, g)$ is strongly continuous in t (Theorem 2.1),

$$x(t_n; f, g) \to x(t_0; f, g) \qquad \text{strongly in } X. \quad (8.8)$$

Moreover, by (8.2)

$$T^*(t_n - t_0)\, x^* \to x^* \qquad \text{strongly in } X^*, \quad (8.9)$$

$$G(t_0 - \cdot\,)\, B(\cdot)\, u_n(\cdot) \to G(t_0 - \cdot\,)\, B(\cdot)\, u_0(\cdot) \qquad \text{weakly in } L_p([0, t]; X). \quad (8.10)$$

Therefore, by tending $n \to \infty$ in (8.3) it follows from (8.4)–(8.10) that

$$\langle w_0, x^* \rangle = \langle x(t_0; f, g), x^* \rangle + \int_0^{t_0} \langle G(t_0 - s)\, B(s)\, u_0(s), x^* \rangle\, ds.$$

Since x^* is arbitrarily chosen,

$$w_0 = x(t; f, g) + \int_0^{t_0} G(t_0 - s)\, B(s)\, u_0(s)\, ds \in W,$$

and hence $u_0 \in U_0$. It is obvious by definition that $t_0 = \tilde{t}(u_0) \leqslant \tilde{t}(u)$ for all $u \in U_0$. This shows u_0 is a time optimal control for P_3.

Next we consider the possibility of maximum principle and bang-bang principle for time optimal controls. Probably, the most simple case in which the maximum principle holds is given by the following

THEOREM 8.2. *Assume that W is convex, closed, and has non-empty interior. Let u be a time optimal control for* P_3 *and let t_0 be its optimal time. Then there exists a non-zero $q^* \in X^*$ such that*

$$\max_{v \in U_{\mathrm{ad}}} \int_0^{t_0} \langle v(s), B^*(s)\, G^*(t_0 - s)\, q^* \rangle_{Y,Y^*}\, ds$$

$$= \int_0^{t_0} \langle u(s), B^*(s)\, G^*(t_0 - s)\, q^* \rangle_{Y,Y^*}\, ds. \tag{8.11}$$

Furthermore if U_{ad} is given by (6.1) *and the control domain $U(t)$ satisfies H_4, then*

$$\max_{v \in U(t)} \langle v, B^*(t)\, G^*(t_0 - t)\, q^* \rangle_{Y,Y^*} = \langle u(t), B^*(t)\, G^*(t_0 - t)\, q^* \rangle_{Y,Y^*}$$

$$a.e. \quad t \in [0, t_0]. \tag{8.12}$$

Proof. Let $\Omega(t_0)$ be the reachable set at time t_0 in (7.3). We shall show $(\mathrm{Int}\, W) \cap \Omega(t_0) = \varnothing$. Suppose to the contrary that there exists a $y \in (\mathrm{Int}\, W) \cap \Omega(T_0)$. Then there exists a control $v \in U_{\mathrm{ad}}$ such that $y = x_v(t_0) \in \mathrm{Int}\, W$. Since $x_v(t)$ is continuous in t, there is $t_1 < t_0$ such that $x_v(t_1) \in W$, which contradicts that t_0 is an optimal time. Then $(\mathrm{Int}\, W) \cap \Omega(t_0) = \varnothing$. It is clear that both $\Omega(t_0)$ and $\mathrm{Int}\, W\, (\neq \varnothing)$ are convex in X. Hence, by the separating hyperplane theorem [17, p. 417], there exists a non-zero $q^* \in X^*$ such that

$$\sup_{y \in \Omega(t_0)} \langle y, q^* \rangle \leqslant \inf_{y \in \mathrm{Int}\, W} \langle y, q^* \rangle. \tag{8.13}$$

Since W is convex and closed, $W = \mathrm{Cl}(\mathrm{Int}\, W)$. So that by continuity and (8.13),

$$\sup_{y \in \Omega(t_0)} \langle y, q^* \rangle \leqslant \inf_{y \in W} \langle y, q^* \rangle \leqslant \langle x_u(t_0), q^* \rangle. \tag{8.14}$$

By the definition of $\Omega(t_0)$, the condition (8.14) is reduced to

$$\sup_{v \in U_{\mathrm{ad}}} \int_0^{t_0} \langle v(s), B^*(s)\, G^*(t_0 - s)\, q^* \rangle_{Y,Y^*}\, ds$$

$$\leqslant \int_0^{t_0} \langle u(s), B^*(s)\, G^*(t_0 - s)\, q^* \rangle_{Y,Y^*}\, ds. \tag{8.15}$$

Therefore (8.11) follows. In the case where U_{ad} is given by (6.1), we can obtain (8.12) from (8.15) by applying the Lebesgue density theorem.

COROLLARY 8.1. *Let W be a closed and convex set in X with non-empty interior. Let the assumption in Theorem 7.1 in which T is replaced by t_0 be satisfied, where t_0 is the optimal time for P_3. Then the time optimal control $u(t)$ is a bang-bang control on $I_0 = [0, t_0]$, i.e., $u(t)$ satisfies*

$$u(t) \in \partial U(t) \qquad a.e. \quad t \in I_0.$$

Proof. The proof is similar to that given in Theorem 7.1. Note that $q^* \neq 0$ in X^*.

COROLLARY 8.2. *Let the assumption in Corollary 8.1 be satisfied. Let $U(t)$ be strictly convex for all $t \in I_0 = [0, t_0]$. Then there exists a unique time optimal solution $(u, x) \in U_{ad} \times C(I_0; X)$. In addition, if Y is a Hilbert space, $p = 2$, and $U(t)$ is given by (7.7) in which I is replaced by I_0, then the time optimal control $u(t)$ is given by*

$$u(t) = y(t) + r(t) \cdot \frac{\Lambda_Y^{-1} B^*(t) p(t)}{|\Lambda_Y^{-1} B^*(t) p(t)|_Y} \qquad a.e. \quad t \in I_0,$$

where $p(t) = G^*(t_0 - t) q^*$, $t \in I_0$, and q^* is as given in Theorem 8.1.

Lastly we consider the case $W = \{w_0\}$, a single point. In this case the time optimal control problem can be considered as a limit of those problems for target sets with non-empty interior. Let $\{W_n\}$ be a sequence of convex and weakly compact sets in X such that

$$w_0 \in \bigcap_{n=1}^{\infty} W_n, \qquad \text{Int } W_n \neq \emptyset, \quad n = 1, 2, ..., \ W_1 \supset W_2 \supset \cdots \supset W_n \supset \cdots,$$

$$\text{dist}(w_0, W_n) = \sup_{x \in W_n} |x - w_0| \to 0 \qquad \text{as} \quad n \to \infty. \tag{8.16}$$

Put $U_0^n = \{u \in U_{ad} : x_u(t) \in W_n \text{ for some } t \in I\}$.

THEOREM 8.3. *Let $\{W_n\}$ be a sequence of closed convex sets in X satisfying the condition (8.16). Assume $U_0^n \neq \emptyset$ for all $n = 1, 2, ...$ and let $\{u_n\}$ be a sequence such that u_n is the time optimal control with the optimal time t_n to the target set W_n, $n = 1, 2, ...$ Then there exists a time optimal control $u_0(t)$ with the optimal time $t_0 = \sup_{n \geq 1}\{t_n\}$ to a point target set $\{w_0\}$ which is given by the weak limit of some subsequence of $\{u_n\}$ in $L_p([0, t_0]; Y)$.*

Proof. Let $t_0 = \sup_{n \geq 1}\{t_n\}$ and let $x_n(t)$, $t \in I$, be the trajectory corresponding to u_n. Since (8.16) is satisfied and U_{ad} is weakly compact, there exist $u_0 \in U_{\text{ad}}$ and subsequences (which are denoted again by $\{u_n\}$, $\{t_n\}$, and $\{w_n\}$) such that

$$\begin{aligned} &u_n \to u_0 && \text{weakly in } L_p(I;\, Y), \\ &t_n = t(u_n) \uparrow t_0 && \text{in } I, \\ &w_n = x_n(t_n) \in W_n \to w_0 && \text{strongly in } X. \end{aligned} \tag{8.17}$$

First we shall show

$$u_0 \in U_0 \equiv \{u \in U_{\text{ad}} : x_u(t) = w_0 \text{ for some } t \in I\}. \tag{8.18}$$

Let $x^* \in X^*$. Then by noticing $G(t) = 0$ if $t < 0$,

$$\langle w_n, x^* \rangle = \langle x(t_n; f, g), x^* \rangle + \int_0^{t_0} \langle u_n(s), B^*(s)\, G^*(t_n - s)\, x^* \rangle_{Y, Y^*}\, ds. \tag{8.19}$$

Since X is reflexive, $G^*(t)$ is strongly continuous on R^+, so that

$$\lim_{n \to \infty} G^*(t_n - s)\, x^* = G^*(t_0 - s)\, x^* \qquad \text{a.e.} \quad s \in I_0 = [0, t_0].$$

Then by the Lebesgue dominated convergence theorem,

$$\begin{aligned} &\int_{I_0} |B^*(s)(G^*(t_0 - s) - G^*(t_n - s))\, x^*|_Y^{p'}\, ds \\ &\quad \leqslant \|B(\cdot)\|_{\infty, I_0} \cdot \int_{I_0} |G^*(t_0 - s)\, x^* - G^*(t_n - s)\, x^*|_*^{p'}\, ds \to 0 \qquad \text{as} \quad n \to \infty, \end{aligned}$$

where $1/p + 1/p' = 1$. This proves

$$B^*(\cdot)\, G^*(t_n - \cdot\,)\, x^* \to B^*(\cdot)\, G^*(t_0 - \cdot\,)\, x^* \quad \text{strongly in } L_{p'}(I_0;\, Y^*). \tag{8.20}$$

We then apply (8.17) and (8.20) to (8.19) and obtain

$$\langle w_0, x^* \rangle = \langle x(t_0; f, g), x^* \rangle + \int_0^{t_0} \langle u_0(s), B^*(s)\, G^*(t_0 - s)\, x^* \rangle_{Y, Y^*}\, ds,$$

and hence

$$w_0 = x(t_0; f, g) + \int_0^{t_0} G(t_0 - s)\, B(s)\, u_0(s)\, ds.$$

Then (8.18) is shown. Next we shall prove that u_0 is the time optimal control and t_0 is the optimal time to the target $\{w_0\}$. Suppose to the contrary that there exists a $v \in U_{\mathrm{ad}}$ such that $x_v(t_1) = w_0$ and $t_1 < t_0$. We choose a large integer n_0 such that $t_1 < t_{n_0} \leqslant t_0$, then by the third term in (8.16), $v \in U_0^{n_0}$. Since u_{n_0} is the time optimal control with the optimal time t_{n_0}, $t_{n_0} \leqslant \tilde{t}(u)$ for all $u \in U_0^{n_0}$, so that $t_{n_0} \leqslant \tilde{t}(v) \leqslant t_1$, a contradiction yields.

APPENDIX 1

Let X be reflexive. Let $G_*(t)$ be the adjoint operator of $G(t)$ and $G_*(t)$ be the solution of (3.11). Then

$$G^*(t) = G_*(t), \qquad t \in R. \tag{A.1}$$

Proof. We shall prove (A.1) by using Laplace transforms. Since $G(t)$ satisfies (2.19), then

$$G^*(t) = T^*(t) + \int_0^t \left(\int_{-h}^0 G^*(s+\xi)\, d\eta^*(\xi) \right) T^*(t-s)\, ds, \qquad t \geqslant 0. \tag{A.2}$$

Clearly, $T^*(t)$ is strongly continuous on R^+. Then by (A.2) and using the Lebesgue dominated convergence theorem, $G^*(t)$ is also strongly continuous on R^+. It is easy to see that $G^*(t)\, x^*$ and $G_*(t)\, x^*$ are of exponential order for each $x^* \in X^*$. Hence both $G^*(t)$ and $G_*(t)$ are Laplace transformable. Taking Laplace transform of (3.11), we have

$$\hat{G}_*(\lambda) = R(\lambda; A_0^*) + R(\lambda; A_0^*) \int_{-h}^0 e^{\lambda s}\, d\eta^*(s) \cdot \hat{G}_*(\lambda) \qquad \text{for } \operatorname{Re} \lambda \text{ large,} \tag{A.3}$$

where $R(\lambda; A_0^*)$ denotes the resolvent of A_0^*. Thus

$$\hat{G}_*(\lambda) = \left[I - R(\lambda; A_0^*) \int_{-h}^0 e^{\lambda s}\, d\eta^*(s) \right]^{-1} R(\lambda; A_0^*). \tag{A.4}$$

We now recall the following relation proved in [30];

$$\hat{G}(\lambda) \left(\lambda I - A_0 - \int_{-h}^0 e^{\lambda s}\, d\eta(s) \right) = I \qquad \text{for } \operatorname{Re} \lambda \text{ large.} \tag{A.5}$$

Substituting $\lambda = \bar{\lambda}$ (complex conjugate) in (A.5) and taking their adjoints, we obtain

$$
\begin{aligned}
I &= \left(\bar{\lambda} I - A_0 - \int_{-h}^{0} e^{\bar{\lambda} s}\, d\eta(s)\right)^* (\hat{G}(\bar{\lambda}))^* \\
&= \left(\lambda I - A_0^* - \int_{-h}^{0} e^{\lambda s}\, d\eta^*(s)\right) \hat{G}^*(\lambda) \\
&= (\lambda I - A_0^*) \left[I - R(\lambda; A_0^*) \int_{-h}^{0} e^{\lambda s}\, d\eta^*(s) \right] \hat{G}^*(\lambda),
\end{aligned}
$$

so that

$$
\hat{G}^*(\lambda) = \left[I - R(\lambda^*; A_0^*) \int_{-h}^{0} e^{\lambda s} d\eta^*(s) \right]^{-1} R(\lambda; A_0^*). \tag{A.6}
$$

Hence by (A.4) and (A.6),

$$
\hat{G}^*(\lambda) = \hat{G}_*(\lambda) \qquad \text{for Re } \lambda \text{ large,}
$$

and then by the uniqueness of Laplace transforms,

$$
G^*(t) = G_*(t) \qquad \text{a.e.} \quad t \in R^+. \tag{A.7}
$$

Since both $G^*(t)$ and $G_*(t)$ are strongly continuous on R^+ and $G^*(t) = G_*(t) = 0$ if $t < 0$, we have (A.1) from (A.7).

APPENDIX 2: Proof of Corollary 2.1

It is shown in [30] that for each $x^* \in D(A_0^*)$, $G^*(t)\, x^*$ is absolutely continuous and satisfies

$$
\frac{dG^*(t)\, x^*}{dt} = G^*(t)\, A_0^* x^* + \int_{-h}^{0} G^*(t+s)\, d\eta^*(s)\, x^* \qquad \text{a.e.} \quad t \geqslant 0, \tag{A.8}
$$

provided that X is reflexive. Let $x(t)$ be the mild solution (2.21). We put $x(t; f, (g^0, 0)) = x_0(t)$ and $x(t; 0, (0, g^1)) = x_1(t)$, $t \geqslant 0$. Then the scalar function $\langle x(t), x^* \rangle$, $x^* \in D(A_0^*)$, is represented by

$$
\begin{aligned}
\langle x(t), x^* \rangle &= \langle x_0(t), x^* \rangle + \langle x_1(t), x^* \rangle \\
&= \left(\langle g^0, G^*(t)\, x^* \rangle + \int_0^t \langle f(s), G^*(t-s)\, x^* \rangle\, ds \right) \\
&\quad + \int_{-h}^{0} \langle U_t(s)\, g^1(s), x^* \rangle\, ds.
\end{aligned}
$$

Since $G^*(t)$ is strongly continuous and $G^*(0)=I$ (Appendix 1), then by (A.8) and Fubini's theorem we have

$$\begin{aligned}
\frac{d}{dt}\langle x_0(t), x^*\rangle &= \langle g^0, G^*(t) A_0^* x^*\rangle + \left\langle g^0, \int_{-h}^{0} G^*(t+s)\, d\eta^*(s)\, x^*\right\rangle \\
&\quad + \langle f(t), x^*\rangle + \int_0^t \langle f(s), G^*(t-s) A_0^* x^*\rangle\, ds \\
&\quad + \int_0^t \left\langle f(s), \int_{-h}^{0} G^*(t-s+\xi)\, d\eta^*(\xi)\, x^*\right\rangle ds \\
&= \langle G(t) g^0, A_0^* x^*\rangle + \left\langle \int_{-h}^{0} d\eta(s)\, G(t+s)\, g^0, x^*\right\rangle \\
&\quad + \langle f(t), x^*\rangle + \left\langle \int_0^t G(t-s) f(s)\, ds, A_0^* x^*\right\rangle \\
&\quad + \left\langle \int_{-h}^{0} d\eta(s) \int_0^{t+s} G(t+s-\xi) f(\xi)\, d\xi, x^*\right\rangle \\
&= \langle x_0(t), A_0^* x^*\rangle \\
&\quad + \left\langle \int_{-h}^{0} d\eta(s)\, x_0(t+s), x^*\right\rangle + \langle f(t), x^*\rangle \qquad \text{a.e.} \quad t \geqslant 0.
\end{aligned} \tag{A.9}$$

We next use the relation

$$\int_{-h}^{0} U_t(s)\, g^1(s)\, ds = \int_0^t G(t-s)\, E(s)\, ds, \qquad t \geqslant 0, \tag{A.10}$$

where

$$E(s) = \begin{cases} \displaystyle\int_{-h}^{-s-0} d\eta(\xi)\, g^1(s+\xi), & 0 \leqslant s \leqslant h \\ 0, & s > h. \end{cases}$$

For $t<0$ we put $x_1(t)=0$. Then as calculated in (A.9) we obtain by (A.10) and (2.12) that if $t \in [0, h]$,

$$\begin{aligned}
&\frac{d}{dt}\langle x_1(t), x^*\rangle \\
&\quad = \langle E(t), x^*\rangle + \langle x_1(t), A_0^* x^*\rangle + \left\langle \int_{-h}^{0} d\eta(\xi)\, x_1(t+\xi), x^*\right\rangle
\end{aligned}$$

$$= \langle x_1(t), A_0^* x^* \rangle + \left\langle \int_{-h}^{-t-0} d\eta(\xi)\, g^1(t+\xi) + \int_{-t}^{0} d\eta(\xi)\, x_1(t+\xi), x^* \right\rangle$$

$$= \langle x_1(t), A_0^* x^* \rangle + \left\langle \int_{-h}^{0} d\eta(s)\, x_1(t+s), x^* \right\rangle \qquad \text{a.e.} \quad t \in [0, h]. \quad \text{(A.11)}$$

If $t \geqslant h$, then $E(t) = 0$ and $t + s \geqslant 0$ for $s \in I_h$, so that

$$\frac{d}{dt} \langle x_1(t), x^* \rangle = \langle x_1(t), A_0^* x^* \rangle + \left\langle \int_{-h}^{0} d\eta(s)\, x_1(t+s), x^* \right\rangle \qquad \text{a.e.} \quad t \geqslant h. \quad \text{(A.12)}$$

Hence from (A.9), (A.11), and (A.12) it follows that

$$\frac{d}{dt} \langle x(t), x^* \rangle = \langle x(t), A_0^* x^* \rangle + \left\langle \int_{-h}^{0} d\eta(s)\, x(t+s), x^* \right\rangle + \langle f(t), x^* \rangle \qquad \text{a.e.} \quad t \geqslant 0.$$

This proves that $x(t)$ is a weak solution of (E).

APPENDIX 3: Proof of Corollary 2.2

From (2.4), (2.30), and (2.31) it can be verified that

$$x(\cdot) \in W_p^{(1)}([-h, T]; X) \cap C([-h, T]; X) \qquad \text{for any } T > 0.$$

Then

$$E(\cdot) = \int_{-h}^{0} d\eta(s)\, x(\cdot + s) \in W_p^{(1)}([0, T]; X) \qquad \text{for any } T > 0. \quad \text{(A.13)}$$

Since X is reflexive, we have by (2.31) and (A.13) that

$$y_i(t) = \int_0^t T(t-s)\, F_i(s)\, ds \in D(A_0) \qquad \text{a.e.} \quad t \geqslant 0,$$

$y_i(t)$ is strongly differentiable for a.e. $t \geqslant 0$ and satisfies

$$\frac{d}{dt} y_i(t) = A_0 y_i(t) + F_i(t) \qquad \text{a.e.} \quad t \geqslant 0, i = 1, 2,$$

where $F_1(t) = f(t)$ and $F_2(t) = E(t)$ [9, p. 32]. This implies by (2.4) and (2.30) that $x(t) \in D(A_0)$ a.e. $t \geqslant 0$ and

$$\frac{d}{dt} x(t) = A_0 T(t) g^0 + \int_{-h}^{0} d\eta(s)\, x(s+t) + f(t)$$

$$+ A_0 \int_0^t T(t-s) f(s)\, ds + A_0 \int_0^t T(t-s) \int_{-h}^{0} d\eta(\xi)\, x(\xi+s)\, ds$$

$$= A_0 x(t) + \int_{-h}^{0} d\eta(s)\, x(s+t) + f(t) \qquad \text{a.e.} \quad t \geqslant 0.$$

Hence $x(t)$ is a strong solution of (E).

APPENDIX 4

The retarded adjoint system given in Example 7.1 is weakly regular.

Proof. It is proved in [30] that the system (CS) is pointwise complete for all $t > 0$ if and only if

$$\bigcap_{s \geqslant t} \operatorname{Ker} G^*(s) = \{0\} \qquad \text{for each } t > 0. \tag{A.14}$$

Since the conditions (i), (ii) in Example 7.1 are satisfied, it can be checked that $G(t)$ is piecewise analytic, i.e., $G(t)\,x$ is analytic on each $((k-1)\,h, kh]$ $(k = 1, 2, ...)$ for any $x \in X$. Hence the mild solution $p(t; q_0^*, 0) = G^*(T-t)(-q_0^*)$ is also piecewise analytic. If the condition (7.2) is satisfied, then by analytic continuation and strong continuity of $G^*(t)$ there exists an integer j such that

$$G^*(t)\, q_0^* = 0 \qquad \text{for all} \quad t \in [jh, (j+1)\,h). \tag{A.15}$$

Since the adjoint system is autonomous, we have by (A.15) that

$$G^*(t)\, q_0^* = 0 \qquad \text{for all } t \in [jh, \infty),$$

or

$$q_0^* \in \bigcap_{t \geqslant jh} \operatorname{Ker} G^*(t). \tag{A.16}$$

So, by (A.14) and (A.16), $q_0^* = 0$ in X^*. Thus C_w is satisfied.

ACKNOWLEDGMENT

The author wishes to thank Professor Hiroki Tanabe for his helpful comments.

References

1. N. U. Ahmed and K. L. Teo, "Optimal Control of Distributed Parameter Systems," North-Holland, New York, 1981.
2. M. Artola, Equations paraboliques à retardement, *C. R. Acad. Sci. Paris* **264** (1967), 668–671.
3. A. V. Balakrishnan, Optimal control problems in Banach spaces, *SIAM J. Control* **3** (1965), 152–180.
4. A. V. Balakrishnan, "Applied Functional Analysis," 2nd ed., Springer-Verlag, Berlin/Heidelberg/New York, 1981.
5. H. T. Banks, Necessary conditions for control problems with variable time-lags, *SIAM J. Control* **6** (1969), 9–47.
6. H. T. Banks and M. Q. Jacobs, An attainable sets approach to optimal control of functional differential equations with function space boundary conditions, *J. Differential Equations* **13** (1973), 127–149.
7. H. T. Banks and G. T. Kent, Control of functional differential equations of retarded and neutral type with target sets in function space, *SIAM J. Control* **10** (1972), 567–593.
8. H. T. Banks and A. Manitius, Application of abstract variational theory to hereditary systems—A survey, *IEEE Trans. Automat. Control* **AC-19** (1974), 524–533.
9. V. Barbu, "Nonlinear Semi-groups and Differential Equations in Banach Spaces," Noordhoff, Leyden, the Netherlands, 1976.
10. Z. Bien and D. H. Chyung, Optimal control of delay systems with a final function condition, *Internat. J. Control* **32** (1980), 539–560.
11. D. H. Chyung and E. B. Lee, Linear optimal systems with time delays, *SIAM J. Control* **4** (1966), 548–575.
12. F. Colonius, The maximum principle for relaxed hereditary differential systems with function space end condition, *SIAM J. Control Optim.* **20** (1982), 695–712.
13. F. Colonius and D. Hinrichsen, Optimal control of functional differential systems, *SIAM J. Control Optim.* **16** (1978), 861–879.
14. R. F. Curtain and A. J. Pritchard, "Infinite Dimensional Linear Systems Theory," Lecture Notes in Control and Information Science, Vol. 8, Springer-Verlag, Berlin/Heidelberg/New York, 1978.
15. M. C. Delfour, The linear quadratic optimal control problem for hereditary differential systems: Theory and numerical solution, *Appl. Math. Optim.* **3** (1977), 101–162.
16. M. C. Delfour and S. K. Mitter, Controllability, observability and optimal feedback control of affine hereditary differential systems, *SIAM J. Control*, **10** (1972), 298–328.
17. N. Dunford and J. T. Schwartz, "Linear Operators, Part I," Interscience, New York, 1966.
18. A. Friedman, Optimal control in Banach spaces, *J. Math. Anal. Appl.* **19** (1967), 35–55.
19. J. S. Gibson, Linear-quadratic optimal control of hereditary differential systems: Infinite dimensional Ricati equations and numerical approximations, *SIAM J. Control Optim.* **21** (1983), 95–139.
20. J. K. Hale, "Theory of Functional Differential Equations," Springer-Verlag, New York/Heidelberg/Berlin, 1977.
21. E. Hille and R. S. Phillips, "Functional Analysis and Semi-group," Colloquium Publications, Vol. 31, *Amer. Math. Soc.*, Providence, R.I., 1957.
22. M. Q. Jacobs and T. J. Kao, An optimum setting problem for time lag systems, *J. Math. Anal. Appl.* **40** (1972), 687–707.
23. T. Kato, Accretive operators and non-linear evolution equations in Banach spaces, *in* "Proceedings, Symposium on Nonlinear Funtional Analysis, Chicago," pp. 138–161, Amer. Math. Soc., Providence, R.I., 1968.

24. T. Kato, "Perturbation Theory for Linear Operators," 2nd ed., Springer-Verlag, Berlin/Heidelberg/New York, 1976.
25. H. Komatsu, Semi-groups of operators in locally convex spaces, *J. Math. Soc. Japan* **16** (1964), 230–262.
26. J. L. Lions, "Optimal Control of Systems Governed by Partial Differential Equations," Springer-Verlag, Berlin/Heidelberg/New York, 1971.
27. S. Nababan and K. L. Teo, On the existence of optimal controls of the first boundary value problems for parabolic delay-differential equations in divergence form, *J. Math. Soc. Japan* **32** (1980), 343–362.
28. S. Nababan and K. L. Teo, Necessary conditions for optimality of Cauchy problems for parabolic partial delay-differential equations, *J. Optim. Theory Appl.* **34** (1981), 117–155.
29. S. Nakagiri, On the fundamental solution of delay-differential equations in Banach spaces, *J. Differential Equations* **41** (1981), 349–368.
30. S. Nakagiri, Pointwise completeness and degeneracy of functional differential equations in Banach spaces, I, II, to appear.
31. M. N. Oguztöreli, "Time-Lag Control Systems," Academic Press, New York, 1966.
32. R. S. Phillips, The adjoint semi-group, *Pacific J. Math.* **5** (1955), 269–283.
33. H. Tanabe, "Equations of Evolution," Pitman, New York, 1979.
34. K. L. Teo, Optimal control of systems governed by time delayed second order linear parabolic partial differential equations with a first boundary condition, *J. Optim. Theory Appl.* **29** (1979), 437–481.
35. P. K. C. Wang, Optimal control of parabolic systems with boundary conditions involving time delays, *SIAM J. Control* **13** (1975), 274–293.
36. J. Warga, "Optimal Control of Differential and Functional Equations," Academic Press, New York, 1972.