Optimal control of linear retarded systems to small solutions

Fritz Colonius

Angaben zur Veröffentlichung / Publication details:

Colonius, Fritz. 1982. "Optimal control of linear retarded systems to small solutions." *International Journal of Control* 36 (4): 675–93.
https://doi.org/10.1080/00207178208932923.

Optimal control of linear retarded systems to small solutions

FRITZ COLONIUS†

This paper deals with necessary optimality conditions for control of general linear retarded systems to small solutions. Here the final states at time T are required to generate small solutions of the uncontrolled system vanishing after $T-h+\alpha$, where $\alpha \geqslant 0$ is fixed. This generalizes the fixed final state problem where $\alpha=0$, and for $\alpha=h$ includes problems with fixed reduced final state Fx_T, where F is the structural operator of the delay system. Necessary optimality conditions in the form of a maximum principle are valid if a certain space called the 'small attainability subspace' is closed in a Sobolev space. A sufficient criterion for this closedness property is indicated. Finally, an example is discussed where the maximum principle for the fixed final state problem is not valid, while the new problem with $\alpha=h$ can be solved by an application of the obtained results.

1. Notation

For a Banach space Z, $\langle z^*, z\rangle_Z = z^*(z)$ denotes the value of the bounded linear functional z^* on Z in $z \in Z$. For $1 < p < \infty$, we define q by $1/p + 1/q = 1$, and let $M^p := \mathbb{R}^n \times L_p([-h, 0], \mathbb{R}^n)$. For $k \in \mathbb{N}$, we denote by $W^{k,p} = W^{k,p}([-h, 0], \mathbb{R}^n)$ the Banach space of functions $x : [-h, 0] \to \mathbb{R}^n$ having an absolutely continuous $(k-1)$st derivative $x^{(k-1)}$ with derivative $x^{(k)} \in L^p$ and $\|x\| := |(|x(-h)|, |\dot{x}(-h)|, \ldots, |x^{(k-1)}(-h)|, \|x^{(k)}\|L_p)|$, where $|\cdot|$ denotes the euclidean norm in finite dimensional space. The dual spaces of $W^{k,p}$ and M^p are identified with $W^{k,q}$ and M^q, respectively. The natural embedding $\Lambda_{k,p}$ of $W^{k,p}$ into M^p is given by

$$[\Lambda_{k,p}\phi]^0 := \phi(0), \quad [\Lambda_{k,p}\phi]^1(s) := \phi(s), \quad s \in [-h, 0]$$

We omit the indices k and p. The inverse $\Lambda^{-1} : M^p \to W^{k,p}$ has domain $\Lambda W^{k,p}$. The abbreviation 'a.a.' means: for almost all.

Finally, $S(t)$ is the semigroup of operators on M^p corresponding to the free motions of (2.1), i.e. $S(t)$ maps $\psi \in M^p$ onto the corresponding solution segment $(x(t), x_t) \in M^p$ of (2.1) with $u=0$ and initial state (2.2).

2. Introduction

The purposes of this paper are: (a) to prove optimality conditions for control of linear retarded systems to small solutions, and (b) to clarify the relation between optimal control of these systems with general 'function space targets' and state space theory based on the structural operators D, F, and G introduced by Delfour and Manitius (1980). We consider linear autonomous systems of the form

$$\dot{x}(t) = L(x_t) + B_0 u(t) \quad \text{a.a.} \quad t \in [0, T] \tag{2.1}$$

with initial condition

$$x(0) = \psi^0, \quad x(t) = \psi^1(t) \quad \text{a.a.} \quad t \in [-h, 0] \tag{2.2}$$

Received 25 April 1982.

† Universität Bremen FB3, FS Dynamische Systeme, Bibliothekstrasse, D-28 Bremen 33, W. Germany.

where
$$x_t(s) := x(t+s) \in \mathbb{R}^n, \quad s \in [-h, 0], \quad u(t) \in \mathbb{R}^m$$
and $T > h > 0$, $L : C([-h, 0], \mathbb{R}^n) \to \mathbb{R}^n$ is linear and bounded, $B_0 \in \mathbb{R}^{n \times m}$, $\bar{\psi} = (\bar{\psi}^0, \bar{\psi}^1) \in M^p := \mathbb{R}^n \times L_p([-h, 0], \mathbb{R}^n)$, and $1 < p < \infty$.

Optimal control problems for (2.1) and (2.2) with integral cost functional and fixed final state
$$x_T = \bar{\phi} \in W^{1,p}([-h, 0], \mathbb{R}^n) \tag{2.3}$$
have found some interest in the last ten years (see, for example, Banks and Manitius (1974), Bien and Chyung (1980), Colonius (1982)). Motivated by recent work of Manitius and others on the structural operator F (Bernier and Manitius 1978, Delfour and Manitius 1980, Manitius 1979, 1980, Salamon 1981), we will treat problems where the end condition (2.3) (for simplicity consider $\bar{\phi} = 0$) is replaced by
$$x_{T+\alpha} = 0 \tag{2.4}$$
where $x(t)$, $t \in [T-h, T+\alpha]$, is the solution of (2.1) with initial state $(x(T), x_T)$ at time T and *zero control* $u(t) = 0$, $t > T$; furthermore $\alpha \geqslant 0$ is fixed. For $\alpha = 0$ this recovers (2.3). For $\alpha = h$ it means that *the reduced state* $F(x(T), x_T)$ is zero. For general α, condition (2.4) means that the state x_T generates a solution of the homogeneous part of (2.1) vanishing after $t = T + \alpha - h$; then the state of the system ' automatically ' becomes zero at $T + \alpha$.

Most interesting is the case where $x(t)$ vanishes for $t \geqslant T$ (i.e. $\alpha = h$).

By a classical result due to Henry (1970), each small solution (i.e. a solution decreasing faster than any exponential function) vanishes after $(n-1)h$. Hence it is sufficient to consider $\alpha \in [0, nh]$, and (2.4) with $\alpha = nh$ means that we want to reach an *arbitrary small solution* at time T.

Suppose that the homogeneous part of (2.1) has only the trivial small solution (i.e. Ker $F = \{0\}$). Then clearly the only way to bring the system to the null terminal state is to control it until the last moment, that is $\alpha = 0$.

For general target functions $\bar{\phi}$, we replace the end condition (2.3) by the following one: The difference $x_T - \bar{\phi}$ is to generate a small solution of the homogeneous part of (2.1) vanishing after $t = T + \alpha - h$. This might be called ' optimal control to final states which are fixed modulo small solutions '.

The optimal control problem with end condition (2.4) lies in between the fixed final state problems (with end condition of the form (2.3)) on the intervals $[0, T]$ and $[0, T+\alpha]$, since the control must be zero on $[T, T+\alpha]$, and say for a standard quadratic cost functional, these three problems will in general lead to different optimal controls. The advantage of this ' intermediate ' problem formulation lies in the following.

For fixed final state problems, the existence of a Lagrange multiplier in $W^{1,2}([-h, 0], \mathbb{R}^n)$ (i.e. the validity of a maximum principle) can be assured for each smooth cost functional, if and only if the attainability subspace is closed in $W^{1,2}([-h, 0], \mathbb{R}^n)$. In the case of simple delay systems of the form
$$\dot{x}(t) = A_0 x(t) + A_1 x(t-h) + B_0 u(t) \tag{2.5}$$
where $A_0, A_1 \in \mathbb{R}^{n \times n}$, $B_0 \in \mathbb{R}^{n \times m}$. Kurcyusz and Olbrot (1977) completely characterized this closedness property; in particular, this property is *independent* of the length $T > h$ of the time interval. Hence it is the same for T

and $T+\alpha$. However, we will show that for optimal control to final states which are *fixed modulo small solutions* a maximum principle is valid under a generalized closedness condition, which naturally depends on α. As an example illustrates, by this way one gets a maximum principle for a new class of linear retarded systems.

Section 3 contains known results and some complements on the strongly continuous semigroup of operators on M^p belonging to the free motions of (2.1), the adjoint and the transposed semigroup, and the related structural operators D, F, and G. This is used in § 4 for the proof of a maximum principle involving the transposed semigroup for problems with general function space end condition. Several concrete forms of the optimality conditions are derived.

Section 5 discusses the application to the problem of control to small solutions. Here the maximum principle is valid if the 'small attainability subspace of order α' (defined as the set of function segments $x_{T+\alpha}$ as in (2.4)) is closed in a certain Sobolev space. We give a sufficient condition for this closedness property. Finally, an example is discussed where the maximum principle for the fixed final state problem is not valid, while the new problem with end condition (2.4) and $\alpha=h$ can be solved by an application of the obtained results.

3. On the structural theory of linear retarded systems

In this section, we collect some basic facts on the semigroup of operators corresponding to the free motions of (2.1), the adjoint and the transposed semigroup and the structural operators D, F, and G (cp. Delfour and Manitius 1980, Manitius 1980, Salamon 1981) and give some complements.

Let $S(t)$, $t\geqslant 0$, denote the strongly continuous semigroup of bounded linear operators on M^p corresponding to

$$\dot{x}(t)=L(x_t)=\int_{-h}^{0} d\eta(s)x(t+s) \tag{3.1}$$

where $\eta:(-\infty,0]\to\mathbb{R}^{n\times n}$ is a $n\times n$ matrix function of bounded variation, right continuous on $(-h,0)$ with $\eta(0)=0$ and $\eta(s)=\eta(-h)$ for $s<-h$.

The transposed semigroup on M^q is denoted by $S^T(t)$ and corresponds to

$$\dot{x}(t)=L^T(x_t)=\int_{-h}^{0} d\eta^T(s)x(t+s) \tag{3.2}$$

Finally, $S^*(t)$, $t\geqslant 0$ denotes the adjoint semigroup of $S(t)$ on M^q.

These three semigroups are related to each other by the bounded linear structural operators F, $G: M^p\to M^p$. For $\psi=(\psi^0,\psi^1)\in M^p$, the value $F\psi$ is implicitly defined by Hale's bilinear form

$$\langle\bar{\psi}, F\psi\rangle_M=\bar{\psi}^{0T}\psi^0+\int_{-h}^{0}\int_{s}^{0}\bar{\psi}^{1T}(s-\sigma)\,d\eta(\sigma)\psi^1(s)\,ds \tag{3.3}$$

for all $\bar{\psi}=(\bar{\psi}^0,\bar{\psi}^1)\in M^q$.

Observe that F leaves the finite dimensional part ψ^0 of ψ invariant.

The operator G is defined by

$$\left.\begin{aligned}[G\psi]^1(s) &:= X(h+s)\psi^0 + \int_{-h}^{0} X(h+s+\sigma)\psi^1(\sigma)\,d\sigma, \quad s\in[-h,0],\\ [G\psi]^0 &:= [G\psi]^1(0)\end{aligned}\right\} \tag{3.4}$$

where $X(\cdot)$ is the fundamental solution of (3.1). The importance of the operators F and G becomes apparent in the following theorem (Manitius 1980).

Theorem 3.1

(i) Suppose that $x(t)$, $t \geqslant -h$, is a solution of (3.1) for $t \geqslant 0$ with initial condition $(x(0), x_0) = \psi\in M^p$. Then $x(t)=0$ for all $t\geqslant 0$ if and only if $F\psi=0$.

(ii) The following relations hold

$$S(h)=GF, \quad S^T(h)=G^*F^*, \quad S^*(t)F^*=F^*S^T(t), \quad t\geqslant 0$$

For the canonical embedding $\Lambda : W^{k,p}\rightarrow M^p$ defined above one has that im $G=\Lambda W^{1,p}$, and the map $G_W : M^p\rightarrow W^{1,p}$ defined by $(G_W\psi)(s) := [G\psi]^1(s)$, $s\in[-h, 0]$, is an isomorphism. This twin of G and its adjoint $G_W{}^* : W^{1,q}\rightarrow M^q$ will play an important role in the optimality conditions. The following relations are easily seen to be valid

$$G=\Lambda G_W, \quad G^*=G_W{}^*\,\Lambda^*, \quad \Lambda^{-1}S(h)=G_W F \tag{3.5}$$

Lemma 3.1

The adjoint $\Lambda^* : M^q\rightarrow W^{1,q}$ is given by

$$(\Lambda^*\psi)(-h)=\psi^0+\int_{-h}^{0}\psi^1(s)\,ds$$

$$\frac{d}{ds}(\Lambda^*\psi)(s)=\psi^0+\int_{s}^{0}\psi^1(\sigma)\,d\sigma, \quad s\in[-h,0]$$

Furthermore, Λ^*M^q is dense in $W^{1,q}$.

The proof of this lemma can be given using partial integration and the duality relations between M^p and M^q, respectively $W^{1,p}$ and $W^{1,q}$. Similarly, one gets the following proposition.

Proposition 3.1

The adjoint $G_W{}^* : W^{1,q}\rightarrow M^q$ is given by

$$[G_W{}^*\,\phi]^0=\phi(-h)+\int_{-h}^{0}\frac{\partial}{\partial s}X^T(h+s)\dot{\phi}(s)\,ds$$

$$[G_W{}^*\,\phi]^1(s)=\dot{\phi}(-h-s)+\int_{-h-s}^{0}\frac{\partial}{\partial\sigma}X^T(h+\sigma+s)\dot{\phi}(\sigma)\,d\sigma, \quad s\in[-h,0]$$

Remark 3.1

Observe that $[G_W{}^*\,\phi]^1(s)-\dot{\phi}(-h-s)$, $s\in[-h,0]$, is a continuous function having the value $[G_W{}^*\,\phi]^0-\phi(-h)$ at $s=0$.

Define inversion operators $\mathscr{J}^1$ and $\mathscr{J}$ in the following way

$$\mathscr{J}^1 : L_p([-h,0],\mathbb{R}^n)\to L_p([-h,0],\mathbb{R}^n),\quad (\mathscr{J}^1\xi)(s) := \xi(-h-s),\quad s\in[-h,0]$$

$$\mathscr{J} : M^p\to M^p,\quad [\mathscr{J}\psi]^0 := \psi^0,\quad [\mathscr{J}\psi]^1 := \mathscr{J}^1\psi^1$$

Consider the equation

$$\dot{y}(t) = -\int_{-h}^{0} d\eta^T(s)y(t-s),\quad \text{a.a.}\quad t\leqslant 0 \tag{3.6}$$

with initial data

$$(y(0), y_h) = \psi\in M^q \tag{3.7}$$

and define

$$\tilde{S}(t)\psi := (y(t,\psi), y_{t+h}(\cdot,\psi)),\quad t\leqslant 0$$

where $y(\cdot,\psi)$ is the unique solution of (3.6) on $(-\infty, 0]$ with initial condition (3.7). Then

$$\tilde{S}(-t) = S^T(t),\quad t\geqslant 0 \tag{3.8}$$

(cp. Bernier and Manitius 1978, proposition 5.1).

Proposition 3.2

Suppose that $y(\cdot)$ solves

$$\left.\begin{aligned} \dot{y}(t) &= -\int_{-h}^{0} d\eta^T(s)y(t-s)+f(t),\quad \text{a.a.}\quad t\leqslant 0 \\ (y(0), y_h) &= \psi\in M^q \end{aligned}\right\} \tag{3.9}$$

where $f : (-\infty, 0]\to\mathbb{R}^n$ is an L^q-function on compact subintervals. Then y solves the Volterra integral equation

$$y(t) = y(0) + \int_{-h}^{0} \eta^T(s)y(-s)\,ds - \int_{t}^{h} \eta^T(t-s)y(s) - \int_{t}^{0} f(s)\,ds,\quad t\leqslant 0\,. \tag{3.10}$$

Proof

Suppose that $\psi\in\Lambda W^{1,q}\subset M^q$. Then one finds for the corresponding solution y of (3.9) by partial integration and Fubini's Theorem for $t\leqslant 0$

$$\begin{aligned} &\int_{t}^{0}\int_{-h}^{0} d\eta^T(s)y(\tau-s)\,d\tau \\ &= -\int_{t}^{0} \eta^T(-h)y(\tau+h)\,d\tau + \int_{-h}^{0}\int_{t}^{0} \eta^T(s)\dot{y}(\tau-s)\,d\tau\,ds \\ &= -\int_{t+h}^{h} \eta^T(-h)y(\tau)\,d\tau + \int_{-h}^{0} \eta^T(s)y(-s)\,ds - \int_{-h}^{0} \eta^T(s)y(t-s)\,ds \\ &= -\int_{t}^{h} \eta^T(t-s)y(s)\,ds + \int_{-h}^{0} \eta^T(s)y(-s)\,ds \end{aligned}$$

Since the right-hand and the left-hand sides of this equality depend continuously on $\psi\in M^q$ (compare Delfour and Manitius 1980, Theorem 3.1), this equality holds for all $\psi\in M^q$. Now the assertion of proposition 3.2 follows easily. □

Proposition 3.3

For each $\phi \in W^{1,q}$, the function $\xi := [G_W^{\,*} \phi]^1 \in L_q$ satisfies the Volterra integral equation

$$\xi(t) = \dot{\phi}(-h-t) - \int_{-h}^{t} \eta^T(s-t)\xi(s)\,ds, \quad t \in [-h, 0]$$

Proof

Using the fundamental solution, Manitius (1980, pp. 7–8) has shown that for each $\psi \in M^p$ the L_p-function

$$\zeta(t) := [G\psi]^1(t-h), \quad t \in [0, h]$$

satisfies

$$\dot{\zeta}(t) = L(\zeta_t) + \psi^1(-t), \quad t \in [0, h]$$

with

$$\zeta(t) := 0, \quad t < 0 \quad \text{and} \quad \zeta(0) := \psi^0$$

Since G^* is related to the transposed equation as G is to the original, then $\xi := [G^*\psi]^1$ satisfies a certain transposed equation, which we write in integrated form

$$\xi(t) = \psi^0 + \int_{-h}^{t} \int_{-\tau-h}^{0} d\eta^T(s)\xi(\tau+s)\,d\tau + \int_{0}^{t+h} \psi^1(-\tau)\,d\tau, \quad t \in [-h, 0] \tag{3.11}$$

After partial integration and application of Fubini's theorem, this yields

$$\xi(t) = \psi^0 - \int_{-h}^{t} \eta^T(s-t)\xi(s)\,ds + \int_{-t-h}^{0} \psi^1(\tau)\,d\tau, \quad t \in [-h, 0]$$

Define

$$\Gamma : W^{1,q} \to L_q \quad \text{as } \Gamma\phi := \mathscr{J}^1\dot{\phi}$$

Then lemma 3.1 shows that

$$\xi(t) = (\Gamma\Lambda^*\psi)(t) - \int_{-h}^{t} \eta^T(s-t)\xi(s)\,ds, \quad t \in [-h, 0]$$

This is a Volterra integral equation in L_q with inhomogeneous term $\Gamma\Lambda^*\psi$, and the solution operator Σ is an isomorphism on L_q (see Dunford and Schwartz 1967, IV.9.53). Thus for $\psi \in M^q$

$$[G_W^{\,*}\,\Lambda^*\psi]^1 = [G^*\psi]^1 = \xi$$
$$= \Sigma\Gamma\Lambda^*\psi$$

Since Γ is continuous on $W^{1,q}$ and $\Lambda^* M^q$ is by lemma 3.1 dense in $W^{1,q}$, it follows that

$$[G_W^{\,*}]^1 = \Sigma\Gamma$$

This shows that for all $\phi \in W^{1,q}$ the integral equation in the proposition is satisfied. □

In the following section we need two operators related to the input. Define

$$B : \mathbb{R}^m \to M^p \quad \text{as } [Bu]^0 := B_0 u, \quad [Bu]^1 := 0$$

$$D : L_p \to M^p \quad \text{as } [D\xi]^0 := 0, \quad [D\xi]^1(s) := B_0\xi(s), \quad s \in [-h, 0]$$

We note the following lemma.

Lemma 3.2

The adjoint $D^* : M^q \to L_q$ is given by

$$(D^*\psi)(s) = B_0^T \psi^1(s), \quad s \in [-h, 0]$$

By the variation of constants formula, we obtain

$$GD\xi = \int_0^h S(h-s)B\xi(-s)\, ds \tag{3.12}$$

Hence by theorem 3.1 and (3.5)

$$\begin{aligned}
&\Lambda^{-1} \int_0^T S(T-s)Bu(s)\, ds \\
&= \Lambda^{-1} S(h) \int_0^{T-h} S(T-h-s)Bu(s)\, ds + \Lambda^{-1} \int_0^h S(h-s)Bu(T-h-s)\, ds \\
&= G_W \left[F \int_0^{T-h} S(T-h-s)Bu(s)\, ds + D\mathscr{J}^1 u_T \right]
\end{aligned} \tag{3.13}$$

where $u_T := u(T+s)$, $s \in [-h, 0]$.

Finally, define the attainable subspace $\mathscr{A}_t$ at time t by

$$\mathscr{A}_t := \left\{ \int_0^t S(t-s)Bu(s)\, ds, \quad u \in L_p \right\}$$

Then for $t \geqslant h$

$$\mathscr{A}_t = S(h)\mathscr{A}_{t-h} + \mathscr{A}_h \tag{3.14}$$

4. Optimality conditions

First, we formulate an optimal control problem for the linear retarded system (2.1) with general function space end condition. Then, under a certain closedness assumption necessary optimality conditions involving the transposed equation are proved. Various concrete forms of the optimality conditions are derived. We consider the following optimal control problem.

Problem 1

$$\underset{u \in L_p([0,\, T],\, \mathbb{R}^m)}{\text{Minimize}} \int_0^{T-h} g_0(x(t), u(t))\, dt + g_1(x_T, u_T)$$

subject to (2.1), (2.2), and

$$Cx_T = z \in Z \tag{4.1}$$

where $g_0 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $g_1 : W^{1,p} \times L_p \to \mathbb{R}$ are continuously Fréchet differentiable, and for all $k > 0$ there are m_1, $m_2 > 0$ such that for all $|x| < k$ and all $u \in \mathbb{R}^m$

$$\left| \frac{\partial}{\partial u} g_0(x, u) \right| \leqslant m_1 + m_2 |u|^{p-1}$$

the derivatives $(\partial/\partial x)g_1(x_T, u_T)$ and $(\partial/\partial u)g_1(x_T, u_T)$ will be identified with the respective elements of $W^{1,q}$ and L_q ; the map $C : W^{1,p} \to Z$ is linear and bounded, where Z is a Banach space.

Remark 4.1

The values Cx_T and $g_1(x_T, u_T)$ are well defined, since $x_t \in W^{1,p}$ for $t \geqslant h$.

Remark 4.2

The differentiability condition for g_1 is for example satisfied if g_1 has the integral form of the first term in the performance index and satisfies the same conditions. The dependence on $x(t)$ and $u(t)$, $t \in [T-h, T]$ is formulated via g_1, since this part of the performance index plays a special role which is more clearly expressed via g_1. Define $g : M^p \times \mathbb{R}^m \to \mathbb{R}$ as $g((\psi^0, \psi^1), u) := g_0(\psi^0, u)$, then by the variation of constants formula, problem 1 is equivalent to the following problem expressed in terms of the state space equation.

Problem 2

$$\text{Minimize } G(y, u) := \int_0^{T-h} g(y(t), u(t))\, dt + g_1(\Lambda^{-1} y(T), u_T)$$

subject to

$$y(t) = S(t)\bar{\psi} + \int_0^t S(t-s) B u(s)\, ds, \quad t \in [0, T]$$

$$C\Lambda^{-1} y(T) = z$$

The next theorem gives optimality conditions involving the transposed equation.

Theorem 4.1

Suppose that $C\Lambda^{-1}\mathscr{A}_T$ is closed in Z and $(\bar{x}, \bar{u})$ is an optimal solution of problem 1. Then there exists a solution $p(\cdot) = (p^0(\cdot), p^1(\cdot))$ of the transposed equation in M^q

$$\text{(i)} \quad p(t) = S^T(T-h-t) p(T-h) + \int_t^{T-h} S^T(s-t) \frac{\partial}{\partial x} g((\bar{x}(s), \bar{x}_s), \bar{u}(s))\, ds, \quad t \in [0, T-h]$$

such that the 'maximum conditions'

$$\text{(ii)} \quad B^* p(t) + \frac{\partial}{\partial u} g_0(\bar{x}(t), \bar{u}(t)) = 0 \quad \text{a.a.} \quad t \in [0, T-h]$$

$$D^* \mathscr{J} p(T-h) + \frac{\partial}{\partial u} g_1(\bar{x}_T, \bar{u}_T) = 0$$

and the following transversality condition hold

$$\text{(iii)} \quad p(T-h) - G_W{}^* \frac{\partial}{\partial x} g_1(\bar{x}_T, \bar{u}_T) \perp \operatorname{Ker}(CG_W)$$

Proof

Define $\bar{y}$ as the state space trajectory corresponding to $\bar{u}$. Then $(\bar{y}, \bar{u})$ is an optimal solution of problem 2. This problem may be considered as a minimization problem with respect to u subject to an equality constraint defined by the map associating with $u \in L_p([0, T], \mathbb{R}^m)$ the element

$$C\Lambda^{-1}\left[S(T)\bar{\psi} + \int_0^T S(T-s) B u(s)\, ds\right] - z \quad \text{of} \quad C\Lambda^{-1}\mathscr{A}_T$$

Since $C\Lambda^{-1}\mathscr{A}_T$ is by assumption closed in Z, it is a Banach space. Hence, by the Lagrange multiplier theorem (Luenberger 1968, p. 243), there exists $1\in(C\Lambda^{-1}\mathscr{A}_T)^*$ such that

$$\frac{d}{du}G(\bar{y},\bar{u})u+\left\langle 1,\, C\Lambda^{-1}\int_0^T S(T-s)Bu(s)\,ds\right\rangle_Z=0 \tag{4.2}$$

for all $u\in L_p$. Here the derivative exists by the chain rule, the assumptions on g_0, g_1, and Vainberg (1964, theorem 21.1).

Employing the chain rule, Fubini's theorem, theorem 3.1, and (3.5), one finds that the left-hand side of (4.2) equals

$$\begin{aligned}&\int_0^T\left\langle\int_s^{T-h}S^T(t-s)\frac{\partial}{\partial x}g((\bar{x}(t),\bar{x}_t),\bar{u}(t))\,dt\right.\\&\qquad\left.+S^T(T-h-s)G_W{}^*\frac{\partial}{\partial x}g_1(\bar{x}_T,\bar{u}_T),\,Bu(s)\right\rangle_M ds\\&\qquad+\left\langle G_W{}^*\frac{\partial}{\partial x}g_1(\bar{x}_T,\bar{u}_T),\,D\mathscr{J}^1u_T\right\rangle_L+\int_0^T\frac{\partial}{\partial u}g_0(\bar{x}(t),\bar{u}(t))\,dt\\&\qquad+\left\langle\frac{\partial}{\partial u}g_1(\bar{x}_T,\bar{u}_T),\,u_T\right\rangle_L+\int_0^{T-h}\langle S^T(T-h-s)G_W{}^*C^*1,\,Bu(s)\rangle_M\,ds\\&\qquad\qquad+\langle G_W{}^*C^*1,\,D\mathscr{J}^1u_T\rangle_M\end{aligned}$$

Define $p(\cdot)$ as the solution of (i) with

$$p(T-h):=\ G_W{}^*C^*1+G_W{}^*\frac{\partial}{\partial x}g_1(\bar{x}_T,\bar{u}_T) \tag{4.3}$$

Then (iii) holds, and (ii) follows from

$$\begin{aligned}&\int_0^{T-h}\langle p(s),\,Bu(s)\rangle_M\,ds+\int_0^{T-h}\frac{\partial}{\partial u}g_0(\bar{x}(t),\bar{u}(t))u(t)\,dt\\&\qquad+\left\langle\mathscr{J}^1D^*p(T-h)+\frac{\partial}{\partial u}g_1(\bar{x}_T,\bar{u}_T),\,u_T\right\rangle_L=0\end{aligned} \tag{4.4}$$

for all $u\in L_p$.

Equation (4.4) is equivalent to (4.2). This proves theorem 4.1. □

Remark 4.3

In a more general context, Kurcyusz (1976) has shown that the closedness assumption in theorem 4.1 is necessary in order to guarantee the existence of Lagrange multipliers for all differentiable performance indices. In particular, if C is the identity map on $W^{1,p}$, the attainability subspace must be closed in $W^{1,p}$.

Remark 4.4

If g_0 and g_1 are convex, the conditions (i)–(iii) above are also sufficient for optimality. Here the closedness assumption may be omitted (cp. Colonius and Hinrichsen 1978, remark 1.4).

Remark 4.5

There are two essential difficulties in deriving theorem 4.1.

(*a*) The transposed semigroup $S^T(\cdot)$ does not coincide with the (rather complicated) functional analytic adjoint $S^*(\cdot)$ of $S(\cdot)$. The remedy for this is the intertwining relation in Theorem 3.1 between $S^T(\cdot)$ and $S^*(\cdot)$ involving the operator F^*.

(*b*) The operator C in the end condition is only defined on $W^{1,p}$ (or $\Lambda W^{1,p}$), not on the whole state space M^p, since $x_T \in \mathscr{A}_T \subset W^{1,p}$. Thus C^*1 is in $W^{1,q}$, but not necessarily in $\Lambda^* M^q \nsubseteq W^{1,q}$.

Hence *prima facie* the adjoint equation should be an equation in $W^{1,q}$, not in M^q. However, due to the *smoothing property* of $S(h) = GF$—being incorporated in G, respectively $G_W : M^p \to W^{1,p}$—the adjoint $G_W^* : W^{1,q} \to M^q$ brings the adjoint equation down to M^q after time h. Thus the (backward) adjoint equation (i) in M^q holds only on the interval $[0, T-h]$.

In the following remarks, various concrete forms of the optimality conditions are discussed.

Remark 4.6

Define

$$\pi(t) := p^0(t), \quad t \in [0, T-h] \tag{4.5}$$

$$(\pi(T-h), \pi_T) := \mathscr{I} p(T-h) \tag{4.6}$$

Then π satisfies the equation

$$\dot{\pi}(t) = -\int_{-h}^{0} d\eta^T(s)\pi(t-s) - \frac{\partial}{\partial x} g_0(\bar{x}(t), \bar{u}(t)), \quad \text{a.a.} \quad t \in [0, T-h] \tag{4.7}$$

This follows from (i) and (3.8).

Remark 4.7

By the previous remark and proposition 3.2 one gets that π satisfies the following Volterra integral equation

$$\pi(t) = -\int_t^T \eta^T(t-s)\pi(s)\,ds + \pi(T-h) + \int_{-h}^{0} \eta^T(s)\pi(T-h-s)\,ds$$
$$+ \int_t^{T-h} \frac{\partial}{\partial x} g_0(\bar{x}(s), \bar{u}(s))\,ds, \quad t \in [0, T-h] \tag{4.8}$$

Remark 4.8

As discussed in remark 4.5, the transposed equation (i) for p is valid only on $[0, T-h]$. However, in a form similar to (4.8), it can be extended to the final interval $[T-h, T]$. By (4.3) and (4.6)

$$\mathscr{I}(\pi(T-h), \pi_T) = (\pi(T-h), \mathscr{I}^1 \pi_T) = p(T-h) = G_W^* \, \omega$$

where

$$\omega := C^*1 + \frac{\partial}{\partial x} g_1(\bar{x}_T, \bar{u}_T) \in W^{1,q}$$

Then, since $\mathscr{J}^1\pi_T = [G_W{}^* \omega]^1$, proposition 3.3 yields

$$(\mathscr{J}^1\pi_T)(t) = \dot{\omega}(-h-t) - \int_{-h}^{t} \eta^T(s-t)(\mathscr{J}^1\pi_T)(s)\, ds$$

Thus for $t\in[T-h, T]$

$$\pi(t) = \dot{\omega}(t-T) - \int_{t}^{T} \eta^T(t-s)\pi(s)\, ds \tag{4.9}$$

Formula (4.9) is an 'adjoint' equation for π on $[T-h, T]$ in the form of a Volterra integral equation involving $\omega\in W^{1,q}$. The function ω satisfies the transversality condition

$$\omega - \frac{\partial}{\partial x} g_1(\bar{x}_T, \bar{u}_T) \perp \mathrm{Ker}\, C \tag{4.10}$$

In the following remarks 4.9–4.11 we consider the case where g_1 is given by

$$g_1(x_T, u_T) = \int_{T-h}^{T} g_0(x(t), u(t))\, dt \tag{4.11}$$

with g_0 as above. This, certainly, is the most interesting case. First, we observe that (4.11) allows to write an 'adjoint' equation on T_1 which is different from (4.9). Then we show that the adjoint equations (4.8) and (4.9) for π can be simplified and unified if the new variable $\zeta := C^*1\in W^{1,q}$ is introduced. This leads also to simple forms of the maximum conditions and the transversality condition.

Remark 4.9

Suppose that g_1 is given by (4.11). Then for $\phi\in W^{1,p}$

$$\frac{\partial}{\partial x} g_1(\bar{x}_T, \bar{u}_T)\phi = \int_{-h}^{0} \frac{\partial}{\partial x} g_0(\bar{x}(T+s), \bar{u}(T+s))\phi(s)\, ds \tag{4.12}$$

In this special situation, there is an essentially different way to get an 'adjoint' equation on $[T-h, T]$. It is easily seen that $(\partial/\partial x)g_1(\bar{x}_T, \bar{u}_T)$ can uniquely be extended to a bounded linear functional on M^p. In this sense it can be identified with the following function $\gamma_M = (\gamma_M{}^0, \gamma_M{}^1)\in M^q$

$$\gamma_M{}^0 := 0, \quad \gamma_M{}^1(s) := \frac{\partial}{\partial x} g_0(\bar{x}(T+s), \bar{u}(T+s)), \quad s\in[-h, 0] \tag{4.13}$$

On the other hand, by lemma 3.1, $(\partial/\partial x)g_1(\bar{x}_T, \bar{u}_T)$—as a bounded linear functional on $W^{1,p}$—can be identified with the following function

$$\gamma_W = \Lambda^*\gamma_M \in W^{1,q} : \gamma_W(-h) := \int_{-h}^{0} \frac{\partial}{\partial x} g_0(\bar{x}(T+s), \bar{u}(T+s))\, ds$$

$$\frac{d}{ds}\gamma_W(s) := \int_{s}^{0} \frac{\partial}{\partial x} g_0(\bar{x}(T+\sigma), \bar{u}(T+\sigma))\, d\sigma, \quad s\in[-h, 0] \tag{4.14}$$

Then we have

$$G^*\gamma_M = G_W{}^* \Lambda^* \gamma_W$$
$$= G_W{}^* \frac{\partial}{\partial x} g_1(\bar{x}_T, \bar{u}_T) \tag{4.15}$$

By Manitius 1980, pp. (7–8), it is known that $G^*\gamma_M$ satisfies a transposed equation ; more specifically, define $\rho \in W^{1,q}([T-h, T], \mathbb{R}^n)$ by

$$\rho(t) := [G^*\gamma_M]^1(T-h-t), \quad t \in [T-h, T] \tag{4.16}$$

Then

$$\Lambda\rho = \mathscr{J} G_W{}^* \frac{\partial}{\partial x} g_1(\bar{x}_T, \bar{u}_T), \quad \rho(T-h) = 0$$

and

$$\dot{\rho}(t) = -L^T(\rho_t) - \frac{\partial}{\partial x} g_0(\bar{x}(t), \bar{u}(t)), \quad \text{a.a.} \quad t \in [T-h, T] \tag{4.17}$$

where

$$\rho(t) := 0 \quad \text{for } t \leqslant T-h$$

Transversality can here be expressed in terms of a function $\mu \in M^q$ defined by

$$\mu = (\mu^0, \mu^1) := G_W{}^* C^* 1 \tag{4.18}$$

Then

$$\mu \perp \operatorname{Ker}(CG_W) \tag{4.19}$$

The functions μ and ρ are related to π by

$$\pi(T-h) = \mu^0, \quad \pi(t) = \mu^1(T-h-t) + \rho(t), \quad t \in [T-h, T] \tag{4.20}$$

This follows since by (4.3), (4.15), and (4.16)

$$(\pi(T-h), \mathscr{J}^1 \pi_T) = G_W{}^* C^* 1 + G^* \gamma_M$$
$$= \mu + (0, \mathscr{J}^1 \rho_T)$$

Thus in the case of the special functional g_1 given by (4.11) (for which $(\partial/\partial x) g_1(\bar{x}_T, \bar{u}_T)$ can be extended to a bounded linear functional on Mp) the optimality conditions (4.17) and (4.19) for functions $\mu \in M^q$ and $\rho \in W^{1,q}$ related to π by (4.20) can replace the transversality condition (iii) in theorem 4.1 (the function π is related to p by (4.5) and (4.6)). The aesthetic advantage of this formulation is that here also on the final interval $[T-h, T]$ an adjoint equation in the form of a *transposed* equation is given.

Remark 4.10

For the special performance index given by (4.11), the integral equations for π described in remarks 4.7 and 4.8 can be simplified. Consider eqn. (4.9) and observe that

$$\omega = C^* 1 + \frac{\partial}{\partial x} g_1(\bar{x}_T, \bar{u}_T)$$
$$= \zeta + \gamma_W$$

where

$$\zeta := C^* 1 \in W^{1,q} \tag{4.21}$$

Then transversality is expressed by

$$\zeta \perp \operatorname{Ker} C \tag{4.22}$$

By (4.9) and (4.14) we find for a.a. $t \in [T-h, T]$

$$\begin{aligned}\pi(t) &= \dot{\zeta}(t-T) + \dot{\gamma}_W(t-T) - \int_t^T \eta^T(t-s)\pi(s)\, ds \\ &= \dot{\zeta}(t-T) + \int_t^T \frac{\partial}{\partial x} g_0(\bar{x}(s), \bar{u}(s))\, ds - \int_t^T \eta^T(t-s)\pi(s)\, ds\end{aligned} \tag{4.23}$$

Now consider eqn. (4.8). By remark 3.1 and proposition 3.3

$$\begin{aligned}\pi(T-h) &= [G_W{}^* \omega]^0 \\ &= -\int_{-h]}^0 \eta^T(\sigma)[G_W{}^* \omega]^1(\sigma)\, d\sigma + \omega(-h) \\ &= -\int_{T-h}^T \eta^T(T-h-\sigma)\pi(\sigma)\, d\sigma + \omega(-h)\end{aligned} \tag{4.24}$$

Furthermore by (4.14)

$$\begin{aligned}\omega(-h) &= \zeta(-h) + \gamma_W(-h) \\ &= \zeta(-h) + \int_{-h}^0 \frac{\partial}{\partial x} g_0(\bar{x}(T+s), \bar{u}(T+s))\, ds\end{aligned} \tag{4.25}$$

Insertion of (4.25) and (4.26) in (4.8) yields for a.a. $t \in [0, T-h]$

$$\pi(t) = -\int_t^T \eta^T(t-s)\pi(s)\, ds + \int_t^T \frac{\partial}{\partial x} g_0(\bar{x}(s), \bar{u}(s))\, ds + \zeta(-h) \tag{4.26}$$

In a different way, the adjoint equations (4.23) and (4.26) involving $\zeta \in W^{1,q}$ defined by (4.21) have been derived in Colonius and Hinrichsen (1978, theorem 3.1) (with a formal difference in the definition of η). This adjoint equation is a minor variant of (4.8), (4.9) based on the special form (4.11) of g_1. The advantage of this adjoint equation is that it has a short and concise form ; furthermore, transversality is directly expressible by ζ. However, the relation of the adjoint equation to the original retarded system is not immediately apparent.

Remark 4.11

If g_1 is given by (4.11), lemma 3.2 implies that the maximum conditions (ii) simplify to

$$B_0{}^T \pi(t) + \frac{\partial}{\partial u} g_0(\bar{x}(t), \bar{u}(t)) = 0 \quad \text{a.a.} \quad t \in [0, T] \tag{4.27}$$

Remark 4.12

Yet another form of the maximum principle has been given by Bien and Chyung (1980) (see also Colonius 1982).

Remark 4.13

Using non-linear semigroup theory, Barbu (1977) derived optimality conditions for non-differentiable convex control problems with functional differential systems, where the function η has finitely many jumps and an absolutely continuous part. The final version of the optimality conditions has a form similar to theorem 4.1.

5. Optimal control to final states which are fixed modulo small solutions

This problem is a special case of problem 1. Taking into account the smoothing property of retarded systems, a natural candidate for Z is

$$Z := W^{[\alpha/h]+1,p}([-h, 0], \mathbb{R}^n) \tag{5.1}$$

with

$$C := \Lambda^{-1}S(\alpha)\Lambda \tag{5.2}$$

Here $[\alpha/h]$ means the smallest integer equal to or less than α/h. $C : W^{1,p} \to W^{[\alpha/h]+1,p}$ is a bounded linear map, and (2.4) is a special case (with $\bar{\phi}=0$) of

$$C(x_T - \bar{\phi}) = \Lambda^{-1}S(\alpha)\Lambda(x_T - \bar{\phi}) = 0 \tag{5.3}$$

where

$$\bar{\phi} \in W^{1,p}([-h, 0], \mathbb{R}^n)$$

Definition 5.1

The subspace

$$\mathscr{A}_T^{\alpha} := \Lambda^{-1}S(\alpha)\Lambda\mathscr{A}_T \subset W^{[\alpha/h]+1,p}$$

is called the *small attainability subspace* of order α at time T.

Remark 5.1

In general, $\mathscr{A}_T^{\alpha} \subset \mathscr{A}_{T+\alpha}$. Hence if $T \geqslant nh$ one has (Salamon 1981, corollary 2.2) that $\mathscr{A}_T^{\alpha} \subset cl\mathscr{A}_{T+\alpha} = cl\mathscr{A}_T$, where the closure is taken in M^p. For the simple delay system (2.5), the attainability subspace is known to be constant for $T \geqslant nh$ (see Banks *et al.* 1975). Hence, $\mathscr{A}_T^{\alpha} \subset \mathscr{A}_{T+\alpha} = \mathscr{A}_T$, i.e. small attainability subspaces are contained in the attainability subspace.

The optimality conditions (i)–(iii) in theorem 4.1 hold for problem 1 with end condition (5.3) if the small attainability subspace $\mathscr{A}_T^{\alpha}$ is closed in $W^{[\alpha/h]+1,p}$. Before we analyse this closedness property, we remark that here the transversality condition (4.22) $\zeta \perp \operatorname{Ker} C$ has the following concrete form. Suppose that for $\phi \in W^{1,p}$ the corresponding solution x of

$$x(t) = \phi(t), \quad t \in [-h, 0], \quad \dot{x}(t) = L(x_t) \quad \text{a.a.} \quad t \in [0, \alpha]$$

satisfies $x_\alpha = 0$ (i.e. ϕ generates a small solution vanishing after time α).

Then it follows that

$$\langle \zeta, \phi \rangle_W = 0 \tag{5.4}$$

In order to give a sufficient condition for closedness of $\mathscr{A}_T^{\alpha}$, the following definition is convenient.

Definition 5.2

Let V_1, V_2 be subspaces of a vector space V_3. Then $V_1 \subset_f V_2$ if and only if there is a finite dimensional subspace V of V_3 with

$$V_1 \subset V_2 + V$$

Theorem 5.1

The small attainability subspace $\mathscr{A}_T^\alpha$ is closed in $W^{[\alpha/h]+1,p}$ if the following two conditions are satisfied

(i) $$\mathscr{A}_{T-h}^{\alpha+h} \subset_f \mathscr{A}_h^\alpha$$

(ii) $$\mathscr{A}_h^\alpha \text{ is closed in } W^{[\alpha/h]+1,p}$$

Proof

By (i) and (3.14)

$$S(\alpha)\mathscr{A}_T = S(\alpha+h)\mathscr{A}_{T-h} + S(\alpha)\mathscr{A}_h$$

Since $S(\alpha)\mathscr{A}_h \subset S(\alpha)\mathscr{A}_T$ this implies that there is a finite dimensional space with $S(\alpha)\mathscr{A}_T = S(\alpha)\mathscr{A}_h + V$. Then it follows by (ii) that $\mathscr{A}_T^\alpha = \Lambda^{-1}S(\alpha)\mathscr{A}_T$ is closed in $W^{[\alpha/h]+1,p}$. □

Remark 5.2

For $\alpha = 0$, condition (ii) is satisfied if and only if Im F is closed in M^p; this follows from $\Lambda^{-1}\mathscr{A}_h = G_W$ Im F since $G_W : M^p \to W^{1,p}$ is an isomorphism. For the simple delay system (2.5), Im F is always closed in M^p, since Im $F = \mathbb{R}^n \times L_p([-h, 0], \text{Im } A_1)$. General necessary and sufficient conditions for closedness of Im F seem to be an open problem (see Delfour and Manitius 1980, remark 2.7).

Remark 5.3

Condition (i) means that $\mathscr{A}_T^\alpha = \mathscr{A}_h^\alpha + V$, where V is a finite dimensional subspace of $W^{[\alpha/h]+1,p}$.

Remark 5.4

For system (2.5) and $\alpha = 0$, condition (i) is also necessary for closedness of $\mathscr{A}_T$ in $W^{1,p}$. This follows from Kurcyusz and Olbrot (1977, corollary 2).

In the following example, the attainability subspace is not closed in $W^{1,2}$ and there is no Lagrange multiplier 1 in $W^{1,2}$. However, the small attainability subspace of order $\alpha = 1$ is closed in $W^{2,2}$ and the maximum principle derived in § 4 can be used in order to compute the optimal solution for control to a small solution. The problem is taken from Kurcyusz (1973, example 2).

Example 5.1

Minimize $G(u) = \frac{1}{2} \int_0^3 [u(t) - v(t)]^2 \, dt$
subject to

$$\left.\begin{aligned} \dot{x}^1(t) &= u(t), && t \in [0, 3] \\ \dot{x}^2(t) &= x^1(t-1), && t \in [0, 3] \end{aligned}\right\} \quad (5.5)$$

$$x^1(t) = x^2(t) = 0, \quad t \in [-1, 0] \tag{5.6}$$

$$x_3 = \bar{\phi} \tag{5.7}$$

where

$$\bar{\phi}(t) := \begin{pmatrix} t+2 \\ 1/2(t+1)^2 \end{pmatrix}, \quad t \in [-1, 0]$$

and

$$v(t) := \begin{cases} 0, & 0 \leqslant t \leqslant \frac{3}{2} \\ 1, & \frac{3}{2} < t \leqslant 3 \end{cases}$$

Here

$$A_0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad B_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Since $A_1 B_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, it follows by Kurcyusz and Olbrot (1977, corollary 2) that $\mathscr{A}_3$ is not closed in $W^{1,2}$.

The unique optimal control $\bar{u}$ is

$$\bar{u}(t) = \begin{cases} 0, & 0 \leqslant t \leqslant 1 \\ 1, & 1 < t \leqslant 3 \end{cases}$$

with corresponding trajectory $\bar{x} = (\bar{x}^1, \bar{x}^2)$

$$\bar{x}^1(t) := \begin{cases} 0, & -1 \leqslant t \leqslant 1 \\ t-1, & 1 \leqslant t \leqslant 3 \end{cases}$$

$$\bar{x}^2(t) := \begin{cases} 0, & -1 \leqslant t \leqslant 2 \\ \frac{1}{2}(t-2)^2, & 2 \leqslant t \leqslant 3 \end{cases}$$

Then $G(\bar{u}) = \frac{1}{4}$.

The derivative of G at $\bar{u}$ can be identified with the following L_2-function

$$\frac{\partial}{\partial u} g_0(\bar{u}(t)) = \begin{cases} 0, & 0 \leqslant t \leqslant 1 \\ 1, & 1 < t \leqslant \frac{3}{2} \\ 0, & \frac{3}{2} < t \leqslant 3 \end{cases}$$

Then Kurcyusz showed that $\mathscr{A}_3$ is dense in $W^{1,2}$, and that there exist no Lagrange multiplier $l \in W^{1,p}$.

Let us consider the weakened version of the end condition with $\alpha = 1$

$$S(1)(x(3), x_3) = S(1)(\bar{\phi}(0), \bar{\phi})$$

i.e. $x(t)$, $t\in[3, 4]$ coincides with the solution $y=(y^1, y^2)$ of

$$\dot{y}^1(t)=0, \qquad t\in[3, 4]$$
$$\dot{y}^2(t)=y^1(t-1), \quad t\in[3, 4]$$
$$y_3=\bar{\phi}$$

One finds that for $t\in[3, 4]$

$$x^1(t)=y^1(t)=y^1(3)=\bar{\phi}^1(0)=2$$
$$x^2(t)=y^2(t)=\bar{\phi}^2(0)+\int_3^t y_1(s-1)\,ds=\tfrac{1}{2}(t-2)^2$$

This yields for the optimal solution $\tilde{u}$, $\tilde{x}=(\tilde{x}^1, \tilde{x}^2)$ on $[2, 3]$

$$\tilde{x}^1(t)=\dot{\tilde{x}}^2(t+1)=t-1$$
$$\tilde{u}(t)=\dot{\tilde{x}}^1(t)=1$$

In particular, we have

$$\tilde{x}^1(2)=1 \quad\text{and}\quad \tilde{x}^2(3)=\tfrac{1}{2}$$

By theorem 5.1, the small attainability subspace $\mathscr{A}_3{}^1$ is closed in $W^{2,2}$. Observe that

$$\mathscr{A}_1{}^1=\{(\phi^1, \phi^2)\in W^{2,2} : \phi^1=\text{constant}, \ \phi^2(-h)=\dot{\phi}^2(-h)=0\}$$

and

$$\mathscr{A}_3{}^1\subset\{(\phi^1, \phi^2)\in W^{2,2} : \phi^1=\text{constant}\}$$

Hence (i) and (ii) in theorem 5.1 hold. The transposed equation for $\pi=(\pi^1, \pi^2)$ (cp. (4.5), (4.6)) has the form

$$\dot{\pi}^1(t)=-\pi^2(t+1), \quad t\in[0, 2]$$
$$\dot{\pi}^2(t)=0, \qquad t\in[0, 2]$$

and the maximum condition yields (see Remark 4.11)

$$\pi^1(t)=v(t)-u(t), \quad t\in[0, 3]$$

The transversality condition for $\zeta\in W^{1,2}$ (see (5.4)) has the form

$$\langle\zeta, x_1\rangle_W=\zeta(-1)x(0)+\int_{-1}^{0}\dot{\zeta}(s)\dot{x}(1+s)\,ds=0$$

where $x=(x^1, x^2)$ is the solution of

$$\dot{x}^1(t)=0$$
$$\dot{x}^2(t)=x^1(t-1) \quad \text{for } t\in[0, 1]$$

and x_0 is any element in $W^{1,2}([-h, 0], \mathbb{R}^2)$. Hence $\zeta(0)=0$ and $\dot{\zeta}^2=0$. Then the adjoint equation (4.23) implies that

$$\pi^2(t)=\pi^2(t)-\dot{\zeta}^2(t-3)=0 \quad \text{a.a.} \quad t\in[2, 3]$$

The adjoint equation shows that

$$\pi^2(t)=a\in\mathbb{R} \quad \text{on } [0, 2]$$

and

$$\pi^1(t)=b\in\mathbb{R} \quad \text{on } [1, 2]$$
$$\pi^1(t)=-a(t-1)+b \quad \text{on } [0, 1]$$

The maximum condition yields

$$\tilde{u}(t) = \begin{cases} a(t-1)-b & t\in[0, 1] \\ -b & t\in(1, \frac{3}{2}] \\ 1-b & t\in(\frac{3}{2}, 2] \\ 1 & t\in(2, 3] \end{cases}$$

The constants a and b can be determined by the boundary conditions

$$\tilde{x}^1(2)=1 \quad \text{and} \quad \tilde{x}^2(3)=\tfrac{1}{2}$$

One obtains

$$a=\tfrac{3}{16} \quad \text{and} \quad b=-\tfrac{19}{64}$$

By standard arguments, $\tilde{u}$ can be seen to be optimal, and the minimal cost

$$J(u)\approx 0{\cdot}03$$

Remark 5.5

Kurcyusz (1973) proposed a different approach to the considered problem. He observed that for the fixed final state problem Lagrange multipliers in a stronger (Sobolev-space) topology exist. However, in general it is not clear, if an appropriate Sobolev-space topology exists and what the concrete form of the optimality condition is.

Remark 5.6

In the considered example we have rank $A_1<n$. For systems of the simple type (2.5) this—not very restrictive—condition is equivalent to the existence of non-trivial small solutions. Hence the introduced concept for the final condition makes sense in this case.

Remark 5.7

In Colonius (1982 a) we show that for the delay systems (2.5) the conditions (i) and (ii) of theorem 5.1 are not only sufficient, but also necessary for closedness of $\mathscr{A}_T{}^{\alpha}$. Furthermore it is shown that closedness of $\mathscr{A}_T{}^{\alpha}$ is in an algebraic sense ' typical ' for this type of system provided that the number of linearly independent delays (that is rank A_1) is not greater than the dimension m of the control space. More precisely the following statement is proven : in the irreducible variety V defined by

$$V := \{(A_0, A_1, B_0)\in R^{n\times n}\times R^{n\times n}\times R^{n\times m} : \text{rank } A_1 \leqslant m\}$$

the subset of triples (A_0, A_1, B_0) not satisfying conditions (i) and (ii) of theorem 5.1 is contained in a proper subvariety of V.

ACKNOWLEDGMENT

I thank Dietmar Salamon and Diederich Hinrichsen for stimulating discussions on state theory of retarded systems. Furthermore, I thank the referees for several remarks which helped to improve the present paper.

References

BANKS, H. T., JACOBS, M. Q., and LANGENHOP, C. E., 1974, *Function Space Controllability for Linear Functional Differential Equations, Differential Games and Control Theory*, edited by E. O. Roxin, P. Liu, and R. L. Sternberg (New York, Basel : Marcel Dekker), pp. 81–98.

BANKS, H. T., and MANITIUS, A., 1974, *I.E.E.E. Trans. autom. Control*, **19**, 524.

BARBU, V., 1977, *Ricerche Mat.*, **26**, 1.

BERNIER, C., and MANITIUS, A., 1978, *Can. J. Math.*, **30**, 897.

BIEN, Z., and CHYUNG, D. H., 1980, *Int. J. Control*, **32**, 539.

COLONIUS, F., 1981, *Optimal control of delay systems to small solutions, Functional-Differential Systems and Related Topics*, Vol. II, edited by M. Kisielewicz (Zielona Gora : The Higher College of Engineering in Zielona Gora), pp. 78–87 ; 1982, *SIAM Jl. Control*, **20**, No. 6, 1982 a, *SIAM Jl. Control* (submitted for publication).

COLONIUS, F., and HINRICHSEN, F., 1978, *SIAM Jl Control*, **16**, 861.

DELFOUR, M. C., and MANITIUS, A., 1980, *J. Math. Anal. Appl.*, **73**, 466.

DUNFORD, N., and SCHWARTZ, J. T., 1967, *Linear Operators, Part I : General Theory* (New York : Wiley–Interscience).

HENRY, D., 1970, *J. Diff. Eqns*, **8**, 494.

KURCYUSZ, S., 1973, *Control Cyber.*, **2**, 99 ; 1976, *J. Opt. Theory Appl.*, **20**, 81.

KURCYUSZ, S., and OLBROT, A. W., 1977, *J. Diff. Eqns*, **24**, 29.

LUENBERGER, D. G., 1968, *Optimization by Vector Space Methods* (New York : John Wiley).

MANITIUS, A., 1981, *SIAM Jl. Control.*, **19**, 516 ; 1980, *J. Diff. Eqns*, **35**, 1.

SALAMON, D., 1981, Report Nr. 38, FS ‘ Dynamische Systeme ’, Universität Bremen (submitted for publication).

VAINBERG, M. M., 1964, *Variational Methods for the Study of Nonlinear Operators* (San Francisco : Holden-Day).