

A CONDITIONAL GRADIENT METHOD FOR A CLASS OF TIME-LAG OPTIMAL CONTROL PROBLEMS

K. H. WONG and K. L. TEO¹

(Received 22 February 1982, revised 17 January 1983)

Abstract

In this paper, we consider a class of optimal control problems with discrete time delayed arguments and bounded control region. A computational algorithm for solving this class of time lag optimal control problems is developed by means of the conditional gradient technique. The convergence property of the algorithm is also investigated.

1. Introduction

Time-lag optimal control problems with bounded control region have been extensively studied in [1], [2], [4]–[9], [11], [13]–[16], [18]. Except for [13], [16] and [18], these papers deal only with questions relating to necessary conditions for optimality and existence of optimal controls. The main concern of [13], [16] and [18] is in computational algorithms. Some important convergence results of the algorithm are also established in [18]. More precisely, it is shown that the sequence of controls generated by the algorithm is a minimizing sequence which converges to the optimal control, both in the weak* topology of L_∞ and in the almost everywhere topology. However, these results are valid only for the case in which the dynamical system and the cost functional are, respectively, linear and convex.

Since the past is history and hence is beyond control, the control variables at negative times are taken as given functions in [4], [9], [17], and in the present paper.

In [12] a computational algorithm based on the strong variational technique is obtained for a class of nonlinear optimal control problems involving only

¹School of Mathematics, University of New South Wales, P.O. Box 1, Kensington, N.S.W. 2033.

© Copyright Australian Mathematical Society 1984, Serial-fee code 0334-2700/84

non-delayed systems. Furthermore, it is shown that any L_∞ -accumulation point of the sequence of controls constructed by the algorithm, if it exists, satisfies a necessary condition for optimality. This convergence result is certainly weaker than that of [18]. However, it is so far the best one can do for general nonlinear optimal control problems, even in the case of non-delayed systems. Note that this type of convergence concept was introduced to optimal control problems for the first time in [12]. Since then, it has become a common type of convergence result in the area of optimal control algorithms. Our main convergence result to be reported in Theorem 4.2 is also of the same type. Note also that in the strong variational technique, a control is required to be perturbed over a subset of the time interval $[0, T]$ in a specific way so that an improved control is obtained.

The computational algorithm reported in [13] is also based on the strong variational technique. However, the convergence results in the sense of [12] (and hence of Theorem 4.2) is not available, because of the time lag effect. This is an open problem.

In the conditional gradient method, an improved control is obtained by perturbing the given control over the whole of the time interval $[0, T]$. This method is applicable to both the delayed and non-delayed systems.

In any optimal control problem, the cost functional is required to be minimized over a particular function space. Furthermore, the control region is not necessarily a convex polyhedron. Normally, the projected gradient method requires projection of the gradient directions onto the boundary of the function space after some iterations. Practically, this is highly undesirable. Thus, the projected gradient method is not a suitable method.

In this paper, the class of time-lag optimal control problems considered is similar to that of [13]. From the above discussions, we see that the conditional gradient method is an appropriate one for such a problem. Thus, we shall use it to devise a computational algorithm. The convergence of the algorithm in the sense of [12] is established in Theorem 4.2.

2. Statement of the problem, basic definitions and assumptions

Consider the following delay-differential equation on the fixed time interval $(0, T]$.

$$\dot{x}(t) = \sum_{j=0}^s f^j(t - h_j, x(t - h_j), u(t - h_j)), \quad (1a)$$

where $x \equiv [x_1, \dots, x_n]^T \in R^n$, $u \equiv [u_1, \dots, u_r]^T \in R^r$ denote, respectively, the state and control vectors, $f^j \equiv [f^j_1, \dots, f^j_n]^T \in R^n$ ($j = 0, \dots, s$), and the superscript T denotes the transpose. The h_k are the time-delays, ordered so that

$$0 = h_0 < h_1 < \dots < h_s < T, \quad s < \infty.$$

The initial function for the differential equation (1a) is

$$x(t) = \phi(t), \quad t \in [-h_s, 0); \quad x(0) = x_0 \tag{1b}$$

where $\phi \equiv [\phi_1, \dots, \phi_n]^T$ is a given, piecewise continuous function on $[-h_s, 0)$ with values in R^n , and x_0 is a given vector in R^n .

Let $\gamma: [-h_s, 0) \rightarrow R^r$ be a given piecewise continuous function and let U be a compact and convex subset of R^r . A measurable function u from $[-h_s, T]$ into R^r is said to be an admissible control if $u(t) = \gamma(t)$ on $[-h_s, 0)$ and $u(t) \in U$ on $[0, T]$. Let \mathcal{Q} denote the class of all such admissible controls.

Let L^q_∞ denote the Banach space $L_\infty([-h_s, T], R^q)$ of all essentially bounded measurable functions from $[-h_s, T]$ into R^q . Its norm is defined as

$$\|u\|_\infty \equiv \text{ess sup}_{t \in [-h_s, T]} \left(\sum_{i=1}^q (u_i(t))^2 \right)^{1/2}.$$

For each $u \in L^q_\infty$, let $x(u)$ be the corresponding vector-valued function which is absolutely continuous on $(0, T]$ and satisfies the differential equation (1a) almost everywhere on $(0, T]$ and the initial condition (1b) everywhere on $[-h_s, 0]$. This function $x(u)$ is called the solution of the system (1) corresponding to $u \in L^q_\infty$.

We may now state our optimal control problem, denoted by (P) , as follows: subject to the system (1), find a control $u \in \mathcal{Q}$ that minimizes the cost functional J defined by

$$J(u) = g(x(u)(T)) + \int_0^T \sum_{j=0}^s f^j(t - h_j, x(u)(t - h_j), u(t - h_j)) dt \tag{2}$$

over \mathcal{Q} , where g and f^j ($j = 0, 1, \dots, s$), are real-valued functions.

Let $\langle \cdot, \cdot \rangle$ and $|\cdot|$ denote, respectively, the usual inner product and norm in any finite dimensional Euclidean space. For any function $G: R^{n_1} \rightarrow R^{n_2}$, let

$$\frac{\partial G}{\partial y} \equiv \begin{bmatrix} \frac{\partial G_1}{\partial y_1} & \dots & \frac{\partial G_{n_2}}{\partial y_1} \\ \vdots & & \vdots \\ \frac{\partial G_1}{\partial y_{n_1}} & \dots & \frac{\partial G_{n_2}}{\partial y_{n_1}} \end{bmatrix}$$

and

$$\left| \frac{\partial G}{\partial y} \right| = \left(\sum_{\substack{i=1, \dots, n_1 \\ j=1, \dots, n_2}} \left| \frac{\partial G_j}{\partial y_i} \right|^2 \right)^{1/2}$$

We assume throughout this paper that the following conditions are satisfied.

(Ai) $f^j: R^1 \times R^n \times R^r \rightarrow R^n$ ($j = 0, \dots, s$); and $f_0^j: R^1 \times R^n \times R^r \rightarrow R^1$ ($j = 0, \dots, s$).

(Aii) f_i^j ($i = 0, \dots, n; j = 0, \dots, s$), are piecewise continuous on $[-h, T]$ for each $(x, u) \in R^n \times R^r$.

(Aiii) There exist constants K_1 and K_2 such that, for all $j = 0, \dots, s$,

$$|F_\beta^j(t, x^1, u^1)| \leq K_1;$$

and

$$|F_\beta^j(t, x^2, u^2) - F_\beta^j(t, x^1, u^1)| \leq K_2(|x^2 - x^1| + |u^2 - u^1|)$$

for all $t \in [-h, T]$, $x^1, x^2 \in R^n$, and $u^1, u^2 \in R^r$, where F_β^j stands for any of the derivative matrices

$$F_x^j \equiv \left[\frac{\partial f_i^j}{\partial x_k} \right]_{i=0, \dots, n, k=1, \dots, n}$$

and

$$F_u^j = \left[\frac{\partial f_i^j}{\partial u_k} \right]_{i=0, \dots, n, k=1, \dots, r}$$

(Aiv) $g: R^n \rightarrow R^1$.

(Av) There exist constants K_3 and K_4 such that, for all $j = 1, \dots, n$,

$$\left| \frac{\partial g(x^1)}{\partial x_j} \right| \leq K_3$$

and

$$\left| \frac{\partial g(x^2)}{\partial x_j} - \frac{\partial g(x^1)}{\partial x_j} \right| \leq K_4|x^2 - x^1|$$

for all $x^1, x^2 \in R^n$.

(Avi) $\phi: [-h, 0) \rightarrow R^n$ and $\gamma: [-h, 0) \rightarrow R^r$ are piecewise continuous functions.

Using exactly the same argument as that given for Theorem 2.3 of [13], we can show that, corresponding to each $u \in L_\infty^r$, the system (1) has a unique solution $x(u)$.

REMARK 2.1. The cost functional J , defined by (2), is also well defined in L'_∞ .

For each $u \in L'_\infty$, let the function

$$\psi(u) \equiv [\psi_0(u), \psi_1(u), \dots, \psi_n(u)]^T: [-h_s, T + h_s] \rightarrow R^{n+1}$$

be the solution of the adjoint system:

$$\dot{\psi}_i(t) = - \sum_{k=0}^n \sum_{j=0}^s \frac{\partial f_k^i(t, x(u)(t), u(t))}{\partial x_i} \psi_k(t + h_j) \quad (i = 1, \dots, n), \quad t \in [0, T], \quad (3a)$$

$$\psi_0(t) = -1, \quad t \in [0, T], \quad (3b)$$

with the conditions

$$\psi_i(T) = - \frac{\partial g(x(u)(T))}{\partial x_i} \quad (i = 1, \dots, n), \quad (3c)$$

$$\psi_i(t) = 0, \quad t \in (T, T + h_s] \quad (i = 0, \dots, n), \quad (3d)$$

and

$$\psi(u)(t) \equiv 0, \quad t \in [-h_s, 0). \quad (3e)$$

Using the abbreviation $\Psi(t)$ for $\psi(t + h_0), \psi(t + h_1), \dots, \psi(t + h_s)$, we define the Hamiltonian function

$$H: [-h_s, T] \times R^n \times R^r \times R^{(n+1)(s+1)} \rightarrow R^1$$

to be

$$H(t, x, v, \Psi(t)) \equiv \sum_{k=0}^n \sum_{j=0}^s f_k^i(t, x, v) \psi_k(t + h_j) e(t + h_j), \quad (4)$$

where $e(\tau)$ is the step function defined by

$$e(\tau) = \begin{cases} 1 & \text{if } \tau > 0, \\ 0 & \text{if } \tau \leq 0. \end{cases}$$

3. The gradient of cost functionals

We begin with

DEFINITION 3.1. Let $u^0 \in L'_\infty$ and J a functional defined on L'_∞ . Then, J is said to be Fréchet differentiable at u^0 if there is a continuous linear functional $J_{u^0}: L'_\infty \rightarrow R^1$ such that

$$\lim_{\|u - u^0\|_\infty \rightarrow 0} \frac{|J(u) - J(u^0) - J_{u^0}(u - u^0)|}{\|u - u^0\|_\infty} = 0.$$

Basically, the goals of this section are to devise an explicit formula for the Fréchet derivative of the cost functional (2) (Theorem 3.1) and to show that this derivative is uniformly continuous in $u \in L'_\infty$ (Theorem 3.2). We have relegated all proofs of the lemmas in this section to the appendix.

For any $u \in L'_\infty$, $x(u)$ is the corresponding solution of the system (1). Let x be considered as a mapping from L'_∞ into L^n_∞ , where the norm in the space L^n_∞ is again denoted by $\|\cdot\|_\infty$.

LEMMA 3.1. *The mapping x from L'_∞ into L^n_∞ is uniformly continuous. That is to say, for any $u^1, u^2 \in L'_\infty$, $\|x(u^2) - x(u^1)\|_\infty \rightarrow 0$ whenever $\|u^2 - u^1\|_\infty \rightarrow 0$.*

THEOREM 3.1. *The cost functional $J: L'_\infty \rightarrow R^1$, defined by (2) is Fréchet differentiable everywhere on L'_∞ . Furthermore, the derivative of J at $u^0 \in L'_\infty$ is given by*

$$\begin{aligned} J_{u^0}(u) &= - \int_{-h}^T \left\langle \frac{\partial}{\partial u} H(t, x(u^0)(t), u^0(t), \Psi(u^0)(t)), u(t) \right\rangle dt \\ &= - \int_{-h}^T \sum_{i=1}^r \frac{\partial H(t, x(u^0)(t), u^0(t), \Psi(u^0)(t))}{\partial u_i} u_i(t) dt, \end{aligned} \tag{5}$$

for all $u \in L'_\infty$.

PROOF. Let $\psi(u)$ be the solution of the adjoint system (3) corresponding to the control $u \in L'_\infty$ and let

$$\hat{\psi}(u^0)(t) \equiv [\psi_1(u^0)(t), \dots, \psi_n(u^0)(t)]^T.$$

Then, from (2), (3b) and (1a), we get

$$\begin{aligned} J(u) - J(u^0) &= g(x(u)(T)) - g(x(u^0)(T)) \\ &\quad - \sum_{j=0}^s \int_{-h}^{T-h_j} \sum_{k=0}^n [f'_k(t, x(u)(t), u(t)) \psi_k(u^0)(t + h_j) \\ &\quad \quad \quad - f'_k(t, x(u^0)(t), u^0(t)) \psi_k(u^0)(t + h_j)] dt \\ &\quad + \int_0^T \langle \dot{x}(u)(t) - \dot{x}(u^0)(t), \hat{\psi}(u^0)(t) \rangle dt. \end{aligned} \tag{6}$$

From (4), (3d), (3e), and (Aii), (6) can be written as

$$\begin{aligned} J(u) - J(u^0) &= g(x(u)(T)) - g(x(u^0)(T)) \\ &\quad - \int_{-h}^T [H(t, x(u)(t), u(t), \Psi(u^0)(t)) \\ &\quad \quad \quad - H(t, x(u^0)(t), u^0(t), \Psi(u^0)(t))] dt \\ &\quad + \int_0^T \langle \dot{x}(u)(t) - \dot{x}(u^0)(t), \hat{\psi}(u^0)(t) \rangle dt. \end{aligned} \tag{7}$$

In view of (3c) and (1b), it follows from integration by parts that

$$\begin{aligned} & \int_0^T \langle \dot{x}(u)(t) - \dot{x}(u^0)(t), \dot{\psi}(u^0)(t) \rangle dt \\ &= - \left\langle \frac{\partial}{\partial x} g(x(u^0)(T)), x(u)(T) - x(u^0)(T) \right\rangle \\ & \quad - \int_0^T \langle x(u)(t) - x(u^0)(t), \dot{\psi}(u^0)(t) \rangle dt. \end{aligned} \tag{8}$$

By the Mean Value Theorem, we get

$$\begin{aligned} & g(x(u)(T)) - g(x(u^0)(T)) \\ &= \left\langle \frac{\partial}{\partial x} g(x_\varepsilon(T)), x(u)(T) - x(u^0)(T) \right\rangle, \end{aligned} \tag{9}$$

where x_ε is an intermediate value determined by the Mean Value Theorem. Furthermore, since $x(u)(\cdot)$ is absolutely continuous on $[0, T]$ for each $u \in \mathcal{U}$, it follows from Lemma 3.1 that

$$\lim_{\|u - u^0\|_\infty \rightarrow 0} x_\varepsilon(T) \rightarrow x(u^0)(T). \tag{10}$$

From (9), (Av), (A4) (of the Appendix), and (10), we deduce that

$$\begin{aligned} & \lim_{\|u - u^0\|_\infty \rightarrow 0} \left\{ \left| g(x(u)(T)) - g(x(u^0)(T)) \right. \right. \\ & \quad \left. \left. - \left\langle \frac{\partial}{\partial x} g(x(u^0)(T)), x(u)(T) - x(u^0)(T) \right\rangle \right\} / \|u - u^0\|_\infty \\ &= 0. \end{aligned} \tag{11}$$

Thus, it can be verified from (7), (8) and (11) that

$$\begin{aligned} & \lim_{\|u - u^0\|_\infty \rightarrow 0} \frac{J(u) - J(u^0)}{\|u - u^0\|_\infty} \\ &= \lim_{\|u - u^0\|_\infty \rightarrow 0} \left[- \int_{-h_s}^T \{ \dot{H}(t, x(u)(t), u(t), \Psi(u^0)(t)) \right. \\ & \quad \left. - H(t, x(u^0)(t), u^0(t), \Psi(u^0)(t)) \} / \|u - u^0\|_\infty dt \right. \\ & \quad \left. - \int_0^T \langle x(u)(t) - x(u^0)(t), \dot{\psi}(u^0)(t) \rangle / \|u - u^0\|_\infty dt \right]. \end{aligned} \tag{12}$$

By the Mean Value Theorem, we get

$$\begin{aligned} & H(t, x(u)(t), u(t), \Psi(u^0)(t)) - H(t, x(u^0)(t), u^0(t), \Psi(u^0)(t)) \\ &= \left\langle \frac{\partial}{\partial x} H(t, x_\varepsilon(t), u_\varepsilon(t), \Psi(u^0)(t)), x(u)(t) - x(u^0)(t) \right\rangle \\ & \quad + \left\langle \frac{\partial}{\partial u} H(t, x_\varepsilon(t), u_\varepsilon(t), \Psi(u^0)(t)), u(t) - u^0(t) \right\rangle, \end{aligned} \tag{13}$$

$$\begin{aligned}
 & + \int_{-h_s}^T \lim_{\|u-u^0\|_\infty \rightarrow 0} \left\{ \left| \left\langle \frac{\partial}{\partial u} H(t, x(u^0)(t), u^0(t), \Psi(u^0)(t)) \right. \right. \right. \\
 & \qquad \qquad \qquad \left. \left. \left. - \frac{\partial}{\partial u} H(t, x_\varepsilon(t), u_\varepsilon(t), \Psi(u^0)(t)), \right. \right. \right. \\
 & \qquad \qquad \qquad \left. \left. \left. u(t) - u^0(t) \right\rangle / \|u - u^0\|_\infty \right\} dt \\
 & \leq C_3 \int_0^T \lim_{\|u-u^0\|_\infty \rightarrow 0} \left| \frac{\partial}{\partial x} H(t, x(u^0)(t), u^0(t), \Psi(u^0)(t)) \right. \\
 & \qquad \qquad \qquad \left. - \frac{\partial}{\partial x} H(t, x_\varepsilon(t), u_\varepsilon(t), \Psi(u^0)(t)) \right| dt \\
 & + \int_{-h_s}^T \lim_{\|u-u^0\|_\infty \rightarrow 0} \left| \frac{\partial}{\partial u} H(t, x(u^0)(t), u^0(t), \Psi(u^0)(t)) \right. \\
 & \qquad \qquad \qquad \left. - \frac{\partial}{\partial u} H(t, x_\varepsilon(t), u_\varepsilon(t), \Psi(u^0)(t)) \right| dt. \tag{15}
 \end{aligned}$$

Using (4), (Aiii) and (14), we obtain

$$\begin{aligned}
 & \lim_{\|u-u^0\|_\infty \rightarrow 0} \left| \frac{\partial}{\partial x} H(t, x(u^0)(t), u^0(t), \Psi(u^0)(t)) \right. \\
 & \qquad \qquad \qquad \left. - \frac{\partial}{\partial x} H(t, x_\varepsilon(t), u_\varepsilon(t), \Psi(u^0)(t)) \right| = 0 \tag{16}
 \end{aligned}$$

for almost all $t \in [0, T]$; and

$$\begin{aligned}
 & \lim_{\|u-u^0\|_\infty \rightarrow 0} \left| \frac{\partial}{\partial u} H(t, x(u^0)(t), u^0(t), \Psi(u^0)(t)) \right. \\
 & \qquad \qquad \qquad \left. - \frac{\partial}{\partial u} H(t, x_\varepsilon(t), u_\varepsilon(t), \Psi(u^0)(t)) \right| = 0 \tag{17}
 \end{aligned}$$

for almost all $t \in [-h_s, T]$.

By virtue of (15), (16) and (17), we conclude that the cost functional is Fréchet differentiable everywhere on L'_∞ and its derivative is given by (5). This completes the proof.

For any $u \in L'_\infty$, $\psi(u)$ is the solution of the system (3). Let Ψ be considered as a mapping from L'_∞ into $L_\infty([-h_s, T + h_s], R^{n+1})$, where the norm in the space $L_\infty([-h_s, T + h_s], R^{n+1})$ is again denoted by $\|\cdot\|_\infty$.

LEMMA 3.2. *The mapping ψ from L'_∞ into $L_\infty([-h_s, T + h_s], R^{n+1})$ satisfies the following properties:*

- (i) *There exists a positive constant K_5 , independent of $u \in L'_\infty$, such that $\|\psi(u)\|_\infty \leq K_5$.*
- (ii) *ψ is uniformly continuous.*

With the help of Lemma 3.1 and Lemma 3.2, we can now derive a continuity theorem concerning the derivative of the cost functional J .

THEOREM 3.2. *The derivative J_u of the cost functional $J(u)$, which maps L'_∞ into $(L'_\infty)^*$ (the dual space of the Banach space L'_∞) is uniformly continuous on L'_∞ .*

PROOF. For any $u^0 \in L'_\infty$, it follows from Theorem 3.1 that

$$J_{u^0}(u) = - \int_{-h_s}^T \left\langle \frac{\partial}{\partial u} H(t, x(u^0)(t), u^0(t), \Psi(u^0)(t)), u(t) \right\rangle dt$$

for any $u \in L'_\infty$.

Thus, for any $u^1, u^2 \in L'_\infty$,

$$|J_{u^2}(u) - J_{u^1}(u)| \leq \|u\|_\infty \int_{-h_s}^T \left| \frac{\partial}{\partial u} H(t, x(u^1)(t), u^1(t), \Psi(u^1)(t)) - \frac{\partial}{\partial u} H(t, x(u^2)(t), u^2(t), \Psi(u^2)(t)) \right| dt.$$

In view of (4) and Cauchy's inequality, we obtain

$$\begin{aligned} \|J_{u^2} - J_{u^1}\| &\equiv \sup_{u \in L'_\infty} |J_{u^2}(u) - J_{u^1}(u)| / \|u\| \\ &\leq \sum_{j=0}^s \int_{-h_s}^T \left[\sum_{k=0}^n \left| \frac{\partial f'_k(t, x(u^1)(t), u^1(t))}{\partial u} - \frac{\partial f'_k(t, x(u^2)(t), u^2(t))}{\partial u} \right| \right. \\ &\quad \times |\psi_k(u^2)(t + h_j)| \\ &\quad \left. + \sum_{k=0}^n |\psi_k(u^1)(t + h_j) - \psi_k(u^2)(t + h_j)| \right. \\ &\quad \left. \times \left| \frac{\partial f'_k(t, x(u^1)(t), u^1(t))}{\partial u} \right| \right] dt. \end{aligned}$$

Thus, by using Lemma 3.2(i), (Aiii) and (A4), we get

$$\begin{aligned} J_{u^2} - J_{u^1} &\leq (s + 1)K_5K_2(C_3 + 1) \int_{-h_s}^T |u^2(t) - u^1(t)| dt \\ &\quad + \sum_{j=0}^s \int_{-h_s}^T K_1 |\psi(u^1)(t + h_j) - \psi(u^2)(t + h_j)| dt. \end{aligned}$$

Therefore, it follows from Lemma 3.2(ii) that

$$\|J_{u^2} - J_{u^1}\| \rightarrow 0 \quad \text{whenever } \|u^2 - u^1\|_\infty \rightarrow 0.$$

This completes the proof.

4. A conditional gradient algorithm

We are now in a position to devise a conditional gradient algorithm to solve the problem (P).

The algorithm is as follows:

STEP 1. Let $\alpha, \beta \in (0, 1)$ be given constants. Choose an initial control $u^0 \in \mathcal{Q}$. Set $i = 0$.

STEP 2. Solve for $x(u^i)$ in the range $[0, T]$.

STEP 3. Solve for $\psi(u^i)$ in the range $[0, T]$.

STEP 4. Find a control $\bar{u}^i \in \mathcal{Q}$ such that

$$\begin{aligned} & \left\langle \frac{\partial}{\partial u} H(t, x(u^i)(t), u^i(t), \Psi(u^i)(t)), u(t) \right\rangle \\ & \geq \left\langle \frac{\partial}{\partial u} H(t, x(u^i)(t), u^i(t), \Psi(u^i)(t)), \bar{u}^i(t) \right\rangle \end{aligned} \tag{18}$$

for all $u \in L_\infty([0, T], U)$ and for almost all $t \in [0, T]$.

STEP 5. Calculate J_{u^i} by using the formula

$$J_{u^i}(u) = - \int_0^T \left\langle \frac{\partial}{\partial u} H(t, x(u^i)(t), u^i(t), \Psi(u^i)(t)), u(t) \right\rangle dt \tag{19}$$

for all $u \in L_\infty([0, T], U)$.

If $J_{u^i}(u^i - u^i) = 0$, set $u^{i+j} = u^i$ for all positive integers j and stop; otherwise go to Step 6.

STEP 6. Choose σ^i to be first element in the sequence $1, \beta, \beta^2, \dots$, such that

$$J(u^i + \sigma^i(\bar{u}^i - u^i)) - J(u^i) \leq \sigma^i \alpha J_{u^i}(\bar{u}^i - u^i). \tag{20}$$

STEP 7. Set

$$u^{i+1} = u^i + \sigma^i(\bar{u}^i - u^i). \tag{21}$$

Go to Step 2.

This algorithm is well-defined. Steps 1, 2, 3 and 5 require no explanation, while the existence of a $u^i \in \mathcal{Q}$ as required in Step 4, is to be proved in Lemma 4.1.

Now, consider Step 6. It follows from Step 4 and Step 5 that $J_{u^i}(\bar{u}^i) \leq J_{u^i}(u^i)$, or equivalently, $J_{u^i}(u^i - \bar{u}^i) \leq 0$. Moreover, since \mathcal{Q} is convex, $u^i + \sigma(u^i - \bar{u}^i) \in \mathcal{Q}$ for all $\sigma \in [0, 1]$. By the definition of $J_{u^i}(u^i - \bar{u}^i)$, we have

$$\lim_{\sigma \rightarrow 0} \frac{J(u^i + \sigma(\bar{u}^i - u^i)) - J(u^i)}{\sigma} = J_{u^i}(\bar{u}^i - u^i),$$

and therefore, since $0 < \alpha < 1$, we have

$$J(u^i + \sigma(\bar{u}^i - u^i)) - J(u^i) \leq \sigma \alpha J_{u^i}(\bar{u}^i - u^i)$$

for σ sufficiently small. Thus, σ^i in Step 6 is well defined. Finally, as above, the convexity of \mathcal{Q} implies that $u^{i+1} \in \mathcal{Q}$ and so Step 7 is also well defined.

REMARK 4.1. *The rate of convergence of the algorithm depends very much on the sizes of α and β chosen in Step 1. The algorithm usually works better whenever the size of α is small and the size of β is large. From our experience, it appears that the algorithm usually converges reasonably well with $\alpha = 0.2$ and $\beta = 0.8$.*

LEMMA 4.1. *For each $u^i \in \mathcal{Q}$, there exists a control $\bar{u}^i \in \mathcal{Q}$ which minimizes the linear functional $J_{u^i}(u)$ over \mathcal{Q} .*

PROOF. Let

$$w(t) = \max_{v \in U} \left\{ \left\langle \frac{\partial}{\partial u} H(t, x(u^i)(t), u^i(t), \Psi(u^i)(t)), v \right\rangle \right\}.$$

Clearly, $w(t)$ is measurable. For each $t \in [0, T]$,

$$w(t) \in \left\{ \left\langle \frac{\partial}{\partial u} H(t, x(u^i)(t), u^i(t), \Psi(u^i)(t)), v \right\rangle : v \in U \right\}.$$

Thus, it follows from the Filippov implicit function lemma ([10], Theorem 3) that there exists a measurable function $m \in L_\infty([0, T], U)$ such that

$$w(t) = \left\langle \frac{\partial}{\partial u} H(t, x(u^i)(t), u^i(t), \Psi(u^i)(t)), m(t) \right\rangle.$$

Define

$$\bar{u}^i(t) = \begin{cases} m(t), & t \in (0, T], \\ \gamma(t), & t \in [-h_s, 0]. \end{cases}$$

This implies that $\bar{u}^i \in \mathcal{Q}$ and satisfies the condition (18). Thus, by (19), \bar{u}^i minimizes the functional $J_{u^i}(u)$ over \mathcal{Q} . This completes the proof.

REMARK 4.2. From the proof of Lemma 4.1, we observe that the control $u^i \in \mathcal{U}$ is the one which maximizes, for each $t \in [0, T]$, the linear function

$$\left\langle \frac{\partial}{\partial u} H(t, x(u^i)(t), u^i(t), \Psi(u^i)(t)), \cdot \right\rangle$$

over U . In theory, there may exist a nonmeasurable function which will do the same job. This is, however, a very rare situation in practice. Normally, there exists only a unique piecewise constant function on $[0, T]$ which maximizes such a linear function for each $t \in [0, T]$. Furthermore, the calculation of this piecewise constant function is, in general, trivial.

THEOREM 4.1. Consider the problem (P). If $u^* \in \mathcal{U}$ is an optimal control, then

$$J_{u^*}(u - u^*) \geq 0 \tag{22}$$

for all $u \in \mathcal{U}$.

PROOF. The proof follows easily from the convexity of \mathcal{U} and the definition of u^* .

THEOREM 4.2. Let $\{u^i\}$ be a sequence of admissible controls generated by the conditional gradient algorithm. If u^* is an accumulation point of $\{u^i\}$ in the L_∞ topology, then u^* satisfies the necessary condition for optimality (22).

PROOF. Since u^* is an accumulation point of $\{u^i\}$, there is a subsequence, again denoted by $\{u^i\}$, converging to u^* .

There are two cases to consider.

CASE 1. There exists an \bar{u}^i such that $J_{\bar{u}^i}(\bar{u}^i - u^i) = 0$. Then, by Step 5 of the algorithm, $u^j = u^i$ for all $j \geq i$ and so $u^* = u^i$. Using (18) and (19), we have

$$J_{u^*}(u - u^*) = J_{\bar{u}^i}(u - u^i) = J_{\bar{u}^i}(u - \bar{u}^i) + J_{\bar{u}^i}(\bar{u}^i - u^i) \geq 0$$

for all $u \in \mathcal{U}$, and so u^* satisfies the necessary condition for optimality (22).

CASE 2. For all $i = 0, 1, 2, \dots$, we have

$$J_{u^i}(\bar{u}^i - u^i) < 0.$$

It then follows from (20) and (21) that

$$\sigma^i \alpha |J_{u^i}(\bar{u}^i - u^i)| \leq |J(u^{i+1}) - J(u^i)|$$

for all i , and so

$$\sigma^i \alpha J_{u^i}(\bar{u}^i - u^i) \rightarrow 0 \quad \text{as } i \rightarrow \infty. \tag{23}$$

We want to show that

$$J_{u^i}(\bar{u}^i - u^i) \rightarrow 0 \quad \text{as } i \rightarrow \infty, \tag{24}$$

for then by taking limits in the inequality

$$J_{u'}(\bar{u}^i - u^i) \leq J_{u'}(u - u^i), \quad u \in \mathcal{U},$$

we obtain

$$0 \leq J_{u^*}(u - u^*), \quad u \in \mathcal{U},$$

the necessary condition for optimality.

Thus, it remains for us to show (24) holds for the sequence $\{u^i\}$ (or at least for a subsequence). If σ^i equals 1 for all but a finite number of i , then (24) follows trivially from (23). Therefore, assume $\{\hat{u}^i\}$ is a subsequence for which $\sigma^i < 1$, for all i .

Then, it follows from the definition of σ^i that

$$J\left(u^i + \frac{\sigma^i}{\beta}(\bar{u}^i - u^i)\right) - J(u^i) > \frac{\sigma^i}{\beta} \alpha J_{u'}(\bar{u}^i - u^i), \tag{25}$$

for all i . Using Taylor's Theorem in Remainder form, the inequality (25) can be written as

$$J_{\hat{u}^i}(\bar{u}^i - u^i) - J_{u^i}(\bar{u}^i - u^i) > (\alpha - 1)J_{u'}(\bar{u}^i - u^i)$$

where $\hat{u}^i = u^i + (\sigma^i \theta / \beta)(\bar{u}^i - u^i) \in \mathcal{U}$, and $0 \leq \theta \leq 1$. Thus, we obtain

$$|J_{u'}(\bar{u}^i - u^i)| \leq \frac{2c}{1 - \alpha} \|J_{u'} - J_{\hat{u}^i}\| \tag{26}$$

where $c = \sup_{u \in \mathcal{U}} \|u\|_\infty$.

Suppose (24) is not true. Then, there exists an $\epsilon > 0$ and a subsequence, again denoted by $\{u^i\}$, such that

$$J_{u'}(\bar{u}^i - u^i) < -\epsilon < 0, \quad \text{for all } i \tag{27}$$

and by (23), it also follows that

$$\sigma^i \rightarrow 0, \quad i \rightarrow \infty.$$

However, then it is clear that $\|\hat{u}^i - u^i\|_\infty \rightarrow 0$ as $i \rightarrow \infty$ and so, by Theorem 3.2,

$$\|J_{\hat{u}^i} - J_{u^i}\| \rightarrow 0 \quad \text{as } i \rightarrow \infty.$$

Finally, using (26) we obtain a contradiction to (27). Thus, (24) holds and the theorem is proved.

5. An illustrative example

Consider the problem of minimizing

$$J(u) = \frac{3}{2} [x(u)(3)]^2 + \frac{1}{2} \int_0^3 [u(t)]^2 dt$$

subject to the differential equation

$$\dot{x}(t) = x(t-1) + u(t), \quad 0 < t \leq 3,$$

with the initial condition $x(t) = 1$ for $t \in [-1, 0]$, where u is a control function with values in $[-3, 3]$.

This problem is adapted from example 4.1 of [3], page 189. The optimal control of this problem is given by

$$\bar{u}(t) = \begin{cases} \delta\{-(t-2)^2/2 - 3/2\}, & 0 \leq t \leq 1, \\ \delta(t-3), & 1 \leq t \leq 2, \\ -\delta, & 2 \leq t \leq 3 \end{cases}$$

where δ is given (approximately) by $\delta = .56231$. The true optimal cost is approximately 1.7338.

This problem has been solved by the conditional gradient method using different sets of parameters $\alpha, \beta \in (0, 1)$. The initial control is chosen to be zero throughout the interval $[0, 3]$ for each set of parameters used. From the numerical results listed in Table 1, it appears that the conditional gradient algorithm works reasonably well for this problem with the 5 different sets of parameters used. However, the rate of convergence for this problem is fastest with the set of parameters $\alpha = 0.2$ and $\beta = 0.8$, and is slowest with the set of parameters $\alpha = 0.8$ and $\beta = 0.8$. The value of the optimal cost obtained by using $\alpha = 0.2$ and $\beta = 0.8$ after 120 iterations or by using $\alpha = 0.1$ and $\beta = 0.9$ after 123 iterations is 1.7342. It is extremely close to the true optimal cost 1.7338 of this problem. Graphs of the true optimal controls, u^{120} with $\alpha = 0.2$ and $\beta = 0.8$, and u^{123} with $\alpha = 0.1$ and $\beta = 0.9$ are plotted in Figure 1. The combination of the fourth order Runge-Kutta integration scheme together with the method of steps is used to integrate the system forward in time and the adjoint system backward in time, both over the interval $[0, 3]$ which is divided into 300 netpoints.

TABLE 1
(Initial Control = 0.0)

<i>i</i>	<i>J(u')</i>				
	$\alpha = 0.8$ $\beta = 0.8$	$\alpha = 0.5$ $\beta = 0.5$	$\alpha = 0.5$ $\beta = 0.8$	$\alpha = 0.1$ $\beta = 0.9$	$\alpha = 0.2$ $\beta = 0.8$
20	1.8050	1.7774	1.7618	1.8515	1.7414
40	1.7681	1.7544	1.7513	1.7802	1.7363
60	1.7606	1.7491	1.7471	1.7435	1.7350
80	1.7542	1.7462	1.7432	1.7388	1.7346
100	1.7506	1.7446	1.7409	1.7362	1.7344
120	1.7492	1.7444	1.7392	1.7343	1.7342

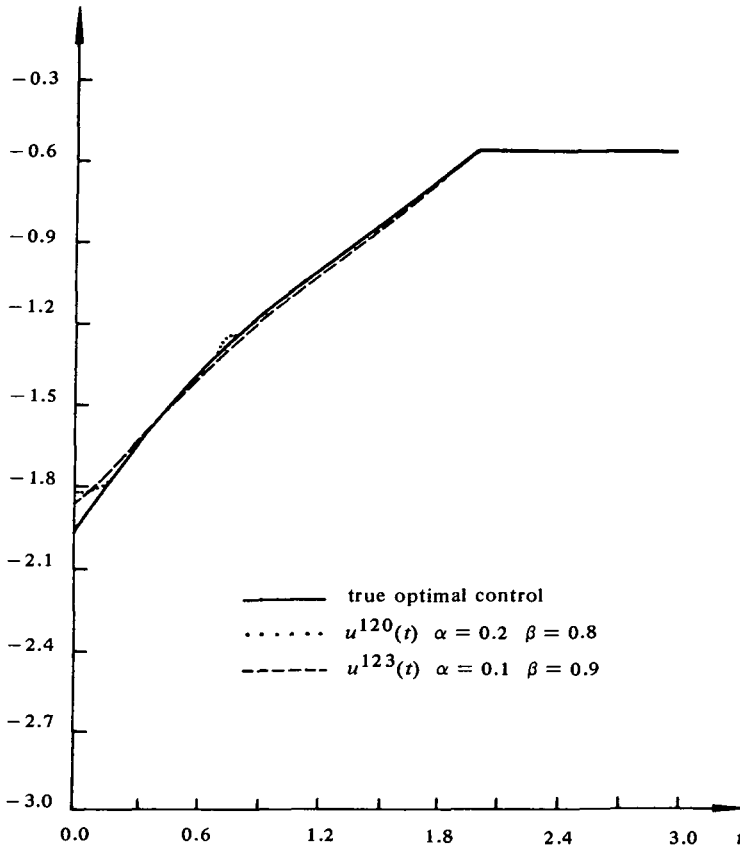


Figure 1. The comparison of controls.

Appendix

Proof of Lemma 3.1. From (1), we get

$$x(u)(t) = x_0 + \int_0^t \sum_{j=0}^s f^j(\tau - h_j, x(u)(\tau - h_j), u(\tau - h_j)) d\tau$$

for all $\tau \in [0, T]$. Thus, it follows that

$$\begin{aligned} &|x(u^2)(t) - x(u^1)(t)| \\ &\leq \int_{-h_s}^t \sum_{j=0}^s [|f^j(\tau, x(u^2)(\tau), u^2(\tau)) - f^j(\tau, x(u^1)(\tau), u^2(\tau))| \\ &\quad + |f^j(\tau, x(u^1)(\tau), u^2(\tau)) - f^j(\tau, x(u^1)(\tau), u^1(\tau))|] d\tau. \end{aligned}$$

By (Aii) and (Aiii), we obtain

$$\begin{aligned}
 &|x(u^2)(t) - x(u^1)(t)| \\
 &\leq K_1(s + 1) \int_{-h_s}^t [|x(u^2)(\tau) - x(u^1)(\tau)| + |u^2(\tau) - u^1(\tau)|] d\tau.
 \end{aligned}
 \tag{A1}$$

In view of (1b), the inequality (A1) becomes

$$\begin{aligned}
 \|x(u^2) - x(u^1)\|_\infty &\leq K_1(s + 1) \int_0^t |x(u^2)(\tau) - x(u^1)(\tau)| d\tau \\
 &\quad + K_1(s + 1)(T + h_s) \|u^2 - u^1\|_\infty \\
 &\leq C_1 [N_1(\|x(u^2) - x(u^1)\|_\infty)] + C_2 \|u^2 - u^1\|_\infty
 \end{aligned}
 \tag{A2}$$

where $C_1 \equiv K_1(s + 1)$, $C_2 \equiv K_1(s + 1)(T + h_s)$ and $N_1(w)(\tau) \equiv \int_0^\tau w(\tau) d\tau$. On this basis, it can be shown, by induction that r , that

$$\begin{aligned}
 \|x(u^2) - x(u^1)\|_\infty &\leq \left\{ \sum_{i=0}^{r-1} C_i' [N_i'(C_2 \|u^2 - u^1\|_\infty)] \right\} \\
 &\quad + C_r' [N_r'(\|x(u^2) - x(u^1)\|_\infty)].
 \end{aligned}
 \tag{A3}$$

Finally, noting that for any constant $C > 0$, $N_r'(C) = C r! / r! \leq C T^r / r!$ and taking the limit as $r \rightarrow \infty$ in (A3), we obtain

$$\|x(u^2) - x(u^1)\|_\infty \leq C_2 \|u^2 - u^1\|_\infty \sum_{i=0}^\infty \frac{C_i' T^i}{i!} = C_3 \|u^2 - u^1\|_\infty,
 \tag{A4}$$

where $C_3 \equiv C_2 \sum_{i=0}^\infty (K_5^i T^i / i!)$. This completes the proof of the lemma.

Proof of Lemma 3.2. From the equation (3a), we get

$$\begin{aligned}
 \psi_i(u)(t) &= \psi_i(u)(T) + \int_t^T \sum_{k=0}^n \sum_{j=0}^s \frac{\partial f_k^j(\tau, x(u)(\tau), u(\tau))}{\partial x_i} \psi_k(u)(\tau + h_j) d\tau \\
 &\hspace{20em} (i = 1, \dots, n),
 \end{aligned}$$

for all $t \in [0, T]$. Thus, by using (Aiii), we deduce from Cauchy's inequality that

$$|\psi_i(u)(t)| \leq |\psi_i(u)(T)| + (s + 1) K_1 (N_2(\|\psi(u)\|_\infty))
 \tag{A5}$$

for all $t \in [0, T]$, where

$$N_2(w)(t) = \int_t^T w(\tau) d\tau.$$

Therefore, from (A5), (3b), (3c) and (Av), it can be verified that

$$\|\psi(u)\|_\infty \leq C_4 + C_5 N_2(\|\psi(u)\|_\infty)$$

where $C_4 \equiv (1 + nk_3^2)^{1/2}$ and $C_5 \equiv (s + 1)(n + 1)^{1/2}K_1$. Now, by using the same argument as that used to obtain the inequality (A4) from the inequality (A2), we obtain

$$\|\psi(u)\|_\infty \leq \sum_{i=0}^\infty \frac{C_5' C_4 T^i}{i!} \equiv K_5. \tag{A6}$$

This completes the proof of part (i) of the lemma.

Using (3a), and (Aiii), we deduce from Cauchy's inequality that, for any $u^1, u^2 \in L'_\infty$,

$$\begin{aligned} & |\psi_i(u^2)(t) - \psi_i(u^1)(t)| \\ & \leq |\psi_i(u^2)(T) - \psi_i(u^1)(T)| \\ & \quad + \left| \int_t^T \sum_{k=0}^n \sum_{j=0}^s \left\{ \left[\frac{\partial f_k^j(\tau, x(u^2)(\tau), u^2(\tau))}{\partial x_i} \right. \right. \right. \\ & \quad \quad \quad \left. \left. \left. - \frac{f \partial f_k^j(\tau, x(u^1)(\tau), u^1(\tau))}{\partial x_i} \right] \psi_k(u^1)(\tau + h_j) \right. \right. \\ & \quad \quad \quad \left. \left. + [\psi_k(u^2)(\tau + h_j) - \psi_k(u^1)(\tau + h_j)] \right. \right. \\ & \quad \quad \quad \left. \left. \times \frac{\partial f_k^j(\tau, x(u^2)(\tau), u^2(\tau))}{\partial x_i} \right\} d\tau \right. \\ & \leq |\psi_i(u^2)(T) - \psi_i(u^1)(T)| \\ & \quad + K_5 \int_t^T \sum_{j=0}^s \left| \frac{\partial f^j(\tau, x(u^2)(\tau), u^2(\tau))}{\partial x_i} - \frac{\partial f^j(\tau, x(u^1)(\tau), u^1(\tau))}{\partial x_i} \right| d\tau \\ & \quad + (s + 1)K_1 [N_2(\|\psi(u^2) - \psi(u^1)\|_\infty)] \quad (i = 1, \dots, n), \end{aligned} \tag{A7}$$

for all $t \in [0, T]$.

From (Av) and Lemma 3.1, we note that, for a given $\epsilon > 0$, there exists a $\delta_1 > 0$ such that, for any $u^1, u^2 \in L'_\infty$,

$$\left| \frac{\partial}{\partial x_i} g(x(u^2)(T)) - \frac{\partial}{\partial x_i} g(x(u^1)(T)) \right| < \epsilon \tag{A8}$$

for all $i = 1, \dots, n$, whenever $\|u^2 - u^1\|_\infty < \delta_1$.

In view of (Aiii) and Lemma 3.1, we observe that, for a given $\varepsilon > 0$, there exists a $\delta_2 > 0$ such that, for any $u^1, u^2 \in L'_\infty$,

$$\left\| \frac{\partial f_j}{\partial x_i}(\cdot, x(u^2)(\cdot), u^2(\cdot)) - \frac{\partial f_j}{\partial x_i}(\cdot, x(u^1)(\cdot), u^1(\cdot)) \right\|_\infty < \varepsilon \tag{A9}$$

for all $i = 1, \dots, n$, and for all $j = 0, \dots, s$, whenever $\|u^2 - u^1\|_\infty < \delta_2$.

Using (A7), it follows from (3c), (A8) and (A9) that

$$\begin{aligned} |\psi_i(u^2)(t) - \psi_i(u^1)(t)| &< \varepsilon + K_5(s + 1)T\varepsilon \\ &\quad + (s + 1)K_1(N_2(\|\psi(u^2) - \psi(u^1)\|_\infty)) \end{aligned} \tag{A10}$$

for all $i = 1, 2, \dots, n$, and for all $t \in [0, T]$, whenever $\|u^2 - u^1\|_\infty < \delta \equiv \min(\delta_1, \delta_2)$. It then follows from (A10) and (3b) that

$$\|\psi(u^2) - \psi(u^1)\|_\infty < C_6\varepsilon + C_7(N_2\|\psi(u^2) - \psi(u^1)\|_\infty)$$

where $C_6 \equiv (n + 1)^{1/2}(1 + K_5(s + 1)T)$ and $C_7 \equiv (n + 1)^{1/2}(s + 1)K_1$.

Now, by using the same argument as that used to obtain the inequality (A4) from the inequality (A2), we get

$$\|\psi(u^2) - \psi(u^1)\|_\infty \leq \sum_{i=0}^{\infty} C_6 T^i C_7^i \varepsilon / i! \equiv C_8 \varepsilon.$$

Therefore,

$$\|\psi(u^2) - \psi(u^1)\|_\infty \rightarrow 0 \quad \text{whenever} \quad \|u^2 - u^1\|_\infty \rightarrow 0.$$

This completes the proof.

Acknowledgement

The authors wish to thank the referees for their comments, and Dr. D. J. Clements for his helpful assistance in shortening the paper.

References

- [1] N. U. Ahmed, "Existence of optimal controls for a class of hereditary systems with lagging control", *J. Inform. Control* 26 (1974), 178–185.
- [2] H. T. Banks, "Necessary conditions for control problems with variable time lags", *SIAM J. Control* 6 (1968), 9–47.
- [3] H. T. Banks and J. A. Burns, "Hereditary control problems: numerical methods based on averaging approximations", *SIAM J. Control Optim.* 6 (1978), 169–208.
- [4] J. M. Blatt, "Control systems with time delays", *J. Austral. Math. Soc. Ser. B* 19 (1976), 478–492.

- [5] A. Friedman, "Optimal control for hereditary processes", *Arch. Rational Mech. Anal.* 15 (1964), 396–416.
- [6] N. D. Georganas, "Optimal control for a class of hereditary systems", Ph. D. Thesis, University of Ottawa, Ottawa, Canada, 1970.
- [7] N. D. Georganas, "Optimization in a class of hereditary systems with direct penalization of the initial data", *Internat. J. Systems Sci.* 5 (1974), 295–299.
- [8] N. D. Georganas and N. U. Ahmed, "Optimal initial data and control for a class of linear hereditary systems", *J. Inform. Control* 22 (1973), 394–402.
- [9] G. Halanay, "Optimal controls for systems with time-lag", *SIAM J. Control* 6 (1968), 215–235.
- [10] C. J. Himmelberg, M. Q. Jacobs and F. S. Van Vleck, "Measurable multi-functions, selectors, and Filippov's implicit functions lemma", *J. Math. Anal. Appl.* 25 (1969), 276–284.
- [11] E. B. Lee, "Variational problems for systems having delay in the control action", *IEEE Trans. Automat. Control* 13 (1969), 697–699.
- [12] D. Q. Mayne and E. Polak, "First-order strong variation algorithms for optimal control", *J. Optim. Theory Appl.* 16 (1975), 303–325.
- [13] J. M. Murray and K. L. Teo, "On a computational algorithm for a class of optimal control problems involving discrete time delayed arguments", *J. Austral. Math. Soc. Ser. B* 20 (1978), 315–343.
- [14] M. N. Oguztoreli, *Time-lag control systems* (Academic Press, New York, 1966).
- [15] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mischenko, *The mathematical theory of optimal processes* (John Wiley and Sons (Interscience Publishers), New York, 1962), 213–226.
- [16] K. L. Teo and B. D. Craven, "On a computational algorithm for time-lag optimal control problems with restricted phase coordinates", *J. Austral. Math. Soc. Ser. B* 21 (1980), 385–397.
- [17] K. L. Teo and E. J. Moore, "Necessary conditions for optimality for control problems with time delays appearing in both state and control variables", *J. Optim. Theory Appl.* 23 (1977), 413–428.
- [18] K. L. Teo, Z. S. Wu and D. J. Clements, "A computational method for convex optimal control problems involving linear hereditary system", *Internat. J. System Sci.* 12 (1981), 1045–1060.